

Entkopplung von Wachstum und Überproduktion von Chemikalien in *Escherichia coli*

Decoupling of Growth and Overproduction of Chemicals in *Escherichia coli*

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

Dem Fachbereich Biologie der
Philipps-Universität Marburg vorgelegt von

Thorben Schramm

geboren am 16. Dezember 1991
in Oldenburg in Holstein

Marburg, August 2022

Das Originaldokument ist gespeichert auf dem Publikationsserver der
Philipps-Universität Marburg (<http://archiv.ub.uni-marburg.de>).



Dieses Werk bzw. Inhalt steht unter einer
Creative Commons
Namensnennung
Nicht-kommerziell
Weitergabe unter gleichen Bedingungen
4.0 International
(CC BY-NC-SA 4.0)
Lizenz.

Ausgenommen davon sind die bereits publizierten Teile dieses Werks, für die die
jeweiligen genannten, ursprünglichen Lizenzen gelten.

Die vollständige Lizenz finden Sie unter:
<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.en>

Die Arbeit zur vorliegenden Dissertation wurde vom Juni 2017 bis September 2020 in der Arbeitsgruppe *Dynamic Control of Metabolic Networks* am Max-Planck-Institut für terrestrische Mikrobiologie, Marburg, und vom Oktober 2020 bis August 2022 in der Arbeitsgruppe *Bacterial Metabolomics* an der Eberhard Karls Universität Tübingen unter der Leitung von Prof. Dr.-Ing. Hannes P. Link angefertigt.

Vom Fachbereich Biologie der Philipps-Universität Marburg (Hochschulkennziffer 1180)
als Dissertation angenommen am: 12.08.2022

Erstgutacher:	Prof. Dr.-Ing. Hannes P. Link
Zweitgutachter:	Prof. Dr. Lennart Randau
Weitere Mitglieder der Prüfungskommission:	Prof. Dr. Martin R. Thanbichler Prof. Dr. Victor Sourjik Prof. Dr. Lars-Oliver Essen
Tag der Disputation (mündlichen Prüfung):	<u>15.11.2022</u>

Teile dieser Arbeit wurden in folgenden Artikeln veröffentlicht:

Schramm, T., Link, H., Von der Stöchiometrie zur Kontrolle metabolischer Netzwerke, **2021**, *BIOspektrum*, 27, pp. 34-36, DOI: <https://doi.org/10.1007/s12268-021-1538-0>
(*non-peer-reviewed review*)

Schramm, T., Lempp, M., Beuter, D., Sierra, S.G., Glatter, T., Link, H., High-throughput enrichment of temperature-sensitive argininosuccinate synthetase for two-stage citrulline production in *E. coli*, **2020**, *Metabolic Engineering*, 60, pp. 14-24, DOI: <https://doi.org/10.1016/j.ymben.2020.03.004>

Boecker, S., Zahoor, A., Schramm, T., Link, H., Klamt, S., Broadening the Scope of Enforced ATP Wasting as a Tool for Metabolic Engineering in *Escherichia coli*, **2019**, *Biotechnology Journal*, 14, 1800438, pp. 1-9, DOI: <https://doi.org/10.1002/biot.201800438>

Boecker, S.* , Slaviero, G.* , Schramm, T., Szymanski, W., Steuer, R., Link, H., Klamt, S., Deciphering the physiological response of *Escherichia coli* under high ATP demand, **2021**, *Molecular Systems Biology*, 17:e10504, DOI: <https://doi.org/10.15252/msb.202110504> (*contributed equally)

Folgender Teil dieser Arbeit befindet sich als Manuskript bei einem einschlägigen Journal unter Begutachtung (*Stand: 12.08.2022*):

Schramm, T.*, **Farke, N.***, **Verhülsdonk, A.**, **Link, H.**, A network approach identifies in-source modifications of primary metabolites during flow-injection mass spectrometry, **2022**, *under revision at ACS Analytical Chemistry* (*contributed equally)

Folgender Teil dieser Arbeit wird für die Begutachtung und Publikation in einem einschlägigen Journal vorbereitet (*Stand: 12.08.2022*):

Schramm, T., **Pahl, V.**, **Link, H.**, A pooled CRISPR screen and metabolomics map temperature-sensitive *Escherichia coli* alleles enabling dynamic control of metabolism, **2022**

Folgender Artikel ist nicht Teil dieser Dissertation und wurde während der Promotionszeit in Koautorenschaft veröffentlicht:

Schubert, C., **Winter, M.**, **Ebert-Jung, A.**, **Kierszniowska, S.**, **Schramm, T.**, **Link, H.**, **Winter, S.**, **Unden, G.**, C4-dicarboxylates and L-aspartate utilization by *Escherichia coli* K-12 in the mouse intestine: L-aspartate as a major substrate for fumarate respiration and as a nitrogen source, **2021**, *Environmental Microbiology*, 23(5), pp. 2564-2577, DOI: <https://doi.org/10.1111/1462-2920.15478>

Diese Arbeit widme ich meinen Großeltern.

WOLFGANG & CHRISTINE SCHRAMM

UVE & EVA ASSMANN

Contents

Summary		1
Zusammenfassung		5
Chapter 1	Introduction	9
Chapter 2	Von der Stöchiometrie zur Kontrolle metabolischer Netzwerke	65
Chapter 3	A network approach identifies in-source modifications of primary metabolites during flow-injection mass spectrometry	73
Chapter 4	High-throughput enrichment of temperature-sensitive argininosuccinate synthetase for two-stage citrulline production in <i>E. coli</i>	99
Chapter 5	A pooled CRISPR screen and metabolomics map temperature-sensitive <i>Escherichia coli</i> alleles enabling dynamic control of metabolism	141
Chapter 6	Broadening the scope of enforced ATP wasting as a tool for metabolic engineering in <i>Escherichia coli</i>	181
Chapter 7	Deciphering the physiological response of <i>Escherichia coli</i> under high ATP demand	205
Conclusion and outlook		249

Summary

Metabolic engineering enables the construction of microbial strains that can effectively overproduce chemicals. However, overproduction strains use substrates not only for the product formation but also for growth. Thus, there is a trade-off between product formation and cell growth that can result in sub-optimal production performance. A solution is to decouple growth and overproduction in two-stage bioprocesses. After a first stage, in which cells grow without product formation, growth is stopped in the second stage while production is induced. There are two key challenges in the creation of two-stage bioprocesses: (1) the microbial metabolism and cell growth needs to be dynamically controlled to achieve a transition between the two stages, and (2) metabolic activity and production rates are typically low in non-growing cells, which needs to be adjusted by genetic engineering. For both challenges, precise knowledge about metabolite concentrations in the engineered strains is critical to guide the design efforts. Therefore, quantitative and high-throughput mass spectrometry-based metabolomics methods are developed that enable metabolite measurements in large numbers of engineered strains. One of the fastest methods is flow-injection mass spectrometry.

In the here presented work, we study aspects of the dynamic control of metabolism and growth, metabolic activity under growth arrest, and flow-injection mass spectrometry. Studying all of these aspects is important to improve our understanding of how to decouple growth and overproduction of chemicals. **Chapter 1** provides a general introduction to metabolic engineering, mass spectrometry-based metabolomics, and two-stage bioprocesses. **Chapter 2** is a short review of metabolic networks.

Flow-injection mass spectrometry (FI-MS) is a metabolomics method that can detect hundreds of metabolites with measurement times in the second scale. However, FI-MS does not rely on chromatographic separation of metabolites prior to analysis. Since all metabolites arrive simultaneously at the mass spectrometer, this could lead to negative effects like in-source modifications and false-positive annotations. With **Chapter 3**, we provide a systematic study of in-source modifications during FI-MS. Key in our analysis was the use of 160 authentic metabolite standards added to a metabolite extract sam-

ple. A network approach and information about metabolite fragmentation identified abundant in-source modifications and showed that even sequential modification events occur. Our analysis approach could explain a large fraction of these modifications. The here presented data are a valuable resource and can be helpful to avoid false-positive metabolite annotations.

Temperature-sensitive proteins carry amino acid substitutions rendering them active at low temperatures. Yet, at higher temperatures, at which the wild type proteins are still active, temperature-sensitive proteins are inactive. Here, we study temperature-sensitive proteins as a tool for metabolic engineering to dynamically control cell growth and metabolism. A goal was to use temperature-sensitive mutants to decouple growth and overproduction of chemicals and create two-stage bioprocesses. Since the identification of temperature-sensitive mutants can be challenging, a focus of our work was also on the development of high-throughput approaches to find temperature-sensitive mutants.

With **Chapter 4**, we present a high-throughput method to enrich temperature-sensitive mutants of a single essential gene in *Escherichia coli*. The method coupled a TIMER protein-based single cell growth rate reporter with fluorescence activated cell sorting. This allowed us to screen millions of cells and enrich temperature-sensitive mutants of argininosuccinate synthetase ArgG. We showed that temperature-sensitive ArgG functions as a metabolic valve that allows for gradual control of growth by temperature. At the same time, it also allows for the overproduction of citrulline, which is the substrate of the ArgG-catalysed reaction. Using temperature-sensitive ArgG, we achieved a two-stage bioprocess that, within 45 h, produced 3 g/L citrulline on a 1 L-bioreactor scale.

We follow up on the study of temperature-sensitivity as tool for metabolic engineering and describe an approach to generate and identify temperature-sensitive mutants in many different genes (**Chapter 5**). We adapted a barcoded CRISPR/Cas9 genome editing method and used a custom design approach to create a pooled *E. coli* strain library with 15,120 members. Each strain carried a mutation causing a single amino acid substitutions in one of 346 essential proteins. In competitive fitness assays at two temperatures, we tracked the abundance of single strains in the pooled strain library by deep sequencing of plasmid-borne barcodes. This allowed us to identify 1,045 temperature-sensitive candidate strains. After isolating a subset of 92 strains, we validated the function of 42 temperature-sensitive enzymes as metabolic valves by FI-MS. As final step, we applied seven temperature-sensitive strains in the two-stage overproduction of chemicals.

A promising approach to achieve high metabolic activity under growth arrest is *en-*

forced ATP wasting. With **Chapter 6**, we provide a study on *enforced ATP wasting* in *E. coli*. Overexpression of ATPase resulted in strongly increased glucose uptake rates in anaerobic conditions under nitrogen starvation. Fermentation products accumulated rapidly until glucose was depleted from the medium.

Following up on our study on *enforced ATP wasting*, we analysed how different levels of ATPase overexpression affected energy metabolism in *E. coli* (**Chapter 7**). Increasing ATPase levels also increased glucose uptake rates up to a critical point. Increasing the expression levels beyond this critical point resulted in a sharp decrease in the glucose uptake rate below the rate of a wild type strain. We showed that this effect is caused by an enzyme in upper glycolysis: phosphofructokinase, which has ATP as substrate and is allosterically activated by ADP. These findings contribute to a better understanding of *E. coli* energy metabolism. They also show how effective *enforced ATP wasting* is at increasing metabolic activity in growth arrested cells making it a very powerful tool in metabolic engineering.

Zusammenfassung

Das *Metabolic Engineering* ermöglicht es, mikrobielle Stämme zu konstruieren, die Chemikalien effektiv überproduzieren. Allerdings nutzen die Produktionsstämme Substrate nicht nur für die Überproduktion von Chemikalien sondern auch um zu wachsen. Also besteht ein *trade-off* zwischen der Produktion und dem Wachstum, der zu sub-optimaler Produktionsleistung führen kann. Eine Lösung dafür ist die Entkopplung des Zellwachstums von der Überproduktion mittels Zwei-Phasen-Bioprozessen. Nach einer ersten Phase, in der Zellen wachsen ohne zu produzieren, wird das Wachstum in der zweiten Phase gestoppt und die Produktion gestartet. Bei der Kreierung eines Zwei-Phasen-Bioprozesses gibt es zwei zentrale Herausforderungen: (1) Der mikrobielle Stoffwechsel und das Zellwachstum müssen dynamisch kontrolliert werden, um einen Übergang zwischen den beiden Prozess-Phasen zu erzielen. (2) Die Stoffwechselaktivität und Produktionsraten sind in nicht-wachsenden Zellen typischerweise niedrig, was durch genetische Modifikationen korrigiert werden muss. Für beide Herausforderungen ist die genaue Kenntnis über die Metabolitkonzentrationen in den modifizierten Stämmen kritisch, um entscheiden zu können, welche genetischen Modifikationen zum Erfolg führen. Daher werden quantitative und Hochdurchsatz-Massenspektrometrie-Methoden entwickelt, die es ermöglichen Metabolitkonzentrationen in einer großen Anzahl an modifizierten Stämmen in kurzer Zeit zu untersuchen. Eine der schnellsten Methoden ist die *flow-injection* Massenspektrometrie.

In dieser Arbeit untersuchen wir Aspekte der dynamischen Kontrolle des Stoffwechsels und Wachstums, der metabolischen Aktivität in nicht-wachsenden Zellen und der *flow-injection* Massenspektrometrie. Es ist wichtig all diese Aspekte zu untersuchen, um besser zu verstehen, wie das Zellwachstum und die Überproduktion von Chemikalien entkoppelt werden kann. Das **Kapitel 1** dient als Einleitung in das Metabolic Engineering, die Massenspektrometrie-basierte Metabolomik, und Zwei-Phasen-Bioprozessen. Das **Kapitel 2** ist ein kurze Abhandlung über metabolische Netzwerke.

Die *flow-injection* Massenspektrometrie (FI-MS) ist eine Metabolomik-Methode, die hunderte Metabolite detektieren kann und Messzeiten im Sekundenbereich hat. Aller-

ings werden Metabolite vor einer Messung mit FI-MS nicht durch chromatographische Methoden aufgetrennt. Dadurch erreichen alle Metabolite das Massenspektrometer zur gleichen Zeit. Das kann zu negativen Effekten führen, wie zum Beispiel Metabolit-Modifikationen in der Ionenquelle und falschen Annotation. Mit dem **Kapitel 3**, stellen wir eine systematische Untersuchung von Metabolit-Modifikationen in der Ionenquelle während der Messungen mittels FI-MS vor. Dabei war die Nutzung von 160 authentischen Metabolitstandards, die wir einer Metabolitextrakt-Probe hinzugegeben haben, zentral für unsere Analyse. Mittels einer Netzwerk-Analyse und Informationen über Metabolit-Fragmentierung konnten wir zeigen, dass Metabolit-Modifikationen in der Ionenquelle abundant sind und sogar sequentielle Mehrfachmodifikationen vorkommen. Mit unserem Analyse-Ansatz konnten wir einen großen Anteil dieser Modifikationen erklären. Die hier vorgestellten Daten sind eine wertvolle Ressource und sind hilfreich, um falsche Annotationen zu vermeiden.

Temperatur-sensitive Proteine haben Aminosäure-Substitutionen, die eine Aktivität bei niedrigen Temperaturen erlauben. Bei höheren Temperaturen, bei denen Wildtyp-Proteine noch aktiv sind, sind Temperatur-sensitive Proteine hingegen inaktiv. Als Teil dieser Arbeit haben wir untersucht, ob Temperatur-sensitive Proteine als Hilfsmittel für das *Metabolic Engineering* nützlich sind und eine dynamische Kontrolle des Zellwachstums und Stoffwechsels ermöglichen. Ein Ziel war es Temperatur-sensitive Proteine zu nutzen, um Zwei-Phasen-Bioprozesse zu kreieren. Dadurch, dass es schwierig sein kann Temperatur-sensitive Mutanten zu finden, lag ein Fokus unserer Arbeit auf der Entwicklung von Hochdurchsatzverfahren zur Erzeugung von Temperatur-sensitiven Mutanten.

Mit dem **Kapitel 4**, stellen wir eine Hochdurchsatz-Methode zur Anreicherung von Temperatur-sensitiven Mutanten eines einzelnen essentiellen Genes in *Escherichia coli* vor. Die Methode verbindet einen Einzelzell-Wachstumsraten-Sensor, der auf einem TIMER-Protein basiert, mit Fluoreszenz-aktivierter Zellsortierung. Dies hat es uns ermöglicht, Millionen von Zellen zu screenen und Temperatur-sensitive Mutanten der Argininosuccinat Synthetase ArgG anzureichern. Wir konnten zeigen, dass ArgG ähnlich wie ein Ventil für Stoffwechselwege funktionieren kann, und, dass es uns ermöglicht das Wachstum graduell mittels der Temperatur zu steuern. Gleichzeitig erlaubt es auch die Überproduktion von Citrullin, das das Substrat der ArgG-katalysierten Reaktion ist. Mittels Temperatur-sensitivem ArgG, haben wir einen Zwei-Phasen-Bioprozess kreiert, mit dem innerhalb von 45 Stunden 3 g/L Citrullin in 1 L-Bioreaktoren hergestellt werden konnte.

Anschließend zu der Studie über Temperatur-Sensitivität als Hilfsmittel für das *Metabo-*

lic Engineering, beschreiben wir in **Kapitel 5** eine Methode, um Temperatur-sensitive Mutanten in vielen unterschiedlichen Genen zu erzeugen und zu identifizieren. Wir haben eine CRISPR/Cas9-basierte Methode zur Genomeditierung adaptiert und einen angepassten Designansatz genutzt, um eine gepoolte Sammlung von 15.120 *E. coli* Stämmen zu erzeugen. Jeder dieser Stämme hat eine Mutation, die eine einzelne Aminosäure-Substitution in einem von 346 essentiellen Proteinen hervorruft. In kompetitiven Fitness-Untersuchungen bei zwei unterschiedlichen Temperaturen haben wir die Abundanz einzelner Stämme mit Hilfe von *deep sequencing* der Plasmid-basierten barcodes verfolgt. Das hat es uns ermöglicht, 1.045 Temperatur-sensitive Kandidaten zu identifizieren. Nachdem wir 92 Stämme isoliert hatten, konnten wir die Funktion von 42 Temperatur-sensitiven Enzymen als *metabolische Ventile* mittels FI-MS bestätigen. Als letzten Schritt haben wir sieben Temperatur-sensitive Stämme für die Überproduktion von Chemikalien in Zwei-Phasen-Prozessen genutzt.

Das *enforced ATP wasting* (erzwungene ATP-Verschwendung) ist ein vielversprechender Ansatz, um hohe metabolische Aktivitäten während des Wachstum-Stopps zu erzielen. Mit **Kapitel 6** stellen wir eine Studie zu dem *enforced ATP wasting* in *E. coli* vor. Die Überexpression von ATPase führte unter anaeroben Bedingungen und Stickstofflimitierung zu stark erhöhten Glukose-Aufnahmeraten. Fermentationsprodukte akkumulierten bis die Glukose im Medium erschöpft war.

Wir haben an der ersten Studie zum *enforced ATP wasting* angeknüpft und untersucht, wie der Energiestoffwechsel durch unterschiedliche Expressionsstärken der ATPase beeinflusst wird (**Kapitel 7**). Mit steigender ATPase-Expressionsstärke stieg auch die Glukose-Aufnahmerate bis zu einem kritischen Punkt. Eine weitere Erhöhung der Expressionsstärke über diesen kritischen Punkt hinweg führte zu einem rapiden Abfall in der Glukose-Aufnahmerate, sogar unter das Niveau eines Wildtyp-Stamms. Wir konnten zeigen, dass dieser Effekt auf ein Enzym in der oberen Glykolyse zurückzuführen ist: die Phosphofruktokinase, welche ATP als Substrat hat und die durch ADP allosterisch aktiviert wird. Diese Ergebnisse tragen zu einem besseren Verständnis des Energiestoffwechsels in *E. coli* bei. Darüber hinaus zeigen unsere Ergebnisse auch, dass *enforced ATP wasting* sehr effektiv ist, um die Glukose-Aufnahmeraten in nicht-wachsenden Zellen deutlich zu steigern. Das macht *enforced ATP wasting* zu einem sehr nützlichen Mittel für das *Metabolic Engineering*.

Chapter 1

Introduction

Thorben Schramm

Background

Modern society relies on the constant production of chemicals that are used to produce plastics, pharmaceuticals, fuels, and many more. Vast quantities of these so-called commodity chemicals are required at a low price. Still today, oil- or natural gas-based chemical processes account for the major share of the global production of commodity chemicals (Yadav et al., 2020). However, the use of fossil resources like crude oils pose the serious threats of global climate change (Pörtner et al., 2022) and pollution (Kotcher et al., 2019). Therefore, one of the great challenges of this century is the transition to a sustainable bioeconomy with less or no environmental impact. Cornerstones for this transition are bioprocesses that use renewable instead of fossil resources.

Looking at history, human relied on microbes for the production and preservation of food since thousands of years. Nearly every human culture knows fermentation products as part of their diet (Tamang et al., 2020). In the ancient Egypt, bread and beer was an integral part of the culture (Samuel, 1996), and, in Scandinavia, the fermentation of fish can be traced back to 9000 years ago (Boethius, 2016). In the beginning of the 20th century, the earliest industrial-scale bioprocesses were established that focussed also on the production of non-food commodity chemicals. Especially, the first world war was a major driver for the commercialization of the aceton-butanol-ethanol fermentation process, whose products were in high demand in the explosive production (Fernbach and Strange, 1912; Weizmann, 1919; Sauer, 2016). However, despite those early successes in industrial bioprocessing, the fuel and chemical industry would strongly rely on crude oils in the the 20th century and beyond. The reasons for this development are manifold (S. J. Bennett and Pearson, 2009). Advances in chemical synthesis were made, and new materials like plastic or pharmaceuticals could be synthesized from oil-derived chemicals (S. J. Bennett and Pearson, 2009). During the second world war, military aviation drove the production of high-octane fuels (Yadav et al., 2020), and after the second world war, the car industry, which then mainly used gasoline or diesel engines, thrived and expanded largely in the private sector.

Since crude oils were very abundant throughout the 20th century, prices for oil-based products remained low, and interest in the development of bioprocesses for fuel or chemical production was low. Also, high development costs in the range between 10 and 50 Million \$ for feasible bioprocesses (M. D. Lynch et al., 2016) and development times between 6 and 8 years (Nielsen and Keasling, 2016) hamper a transition from fossil fuels to renewables. Only in the 70s, when oil-prices tripled or even quadrupled during

an international political crisis, bioprocesses for the production of fuels were revisited. Since then, a large ethanol economy emerged that focussed on the production of industrial ethanol by fermentation mainly from grains such as wheat and maize (Sheehan and Himmel, 1999). This so-called bioethanol is an additive to conventional gasoline lowering the environmental impact of the fuel. Despite these developments, crude oils still remain the main raw material for fuel and chemical industry today, and it is unclear how long global reserves will last (Bardi, 2019). However, fossil resources are fundamentally limited by definition. Therefore, together with environmental concerns, this will inevitably foster the development of new and economically competitive bioprocesses.

At the heart of every bioprocess are microbes that convert substrates like sugars into biomass and other chemicals. With recent technological breakthroughs in genetic engineering, DNA sequencing, mass spectrometry-based metabolomics, and many more, we are currently facing great opportunities in the biological sciences. We can now study and engineer microbial life at the molecular level better than ever before, and we are using our insights to improve microbial strains to produce chemicals faster, to higher concentrations, and with better yields. Key for the improvement of bioprocesses is the **metabolic engineering**.

1 Metabolic engineering

Metabolic engineering emerged as a distinctly recognized scientific field in the years around 1990 (Bailey, 1991; Stephanopoulos and Vallino, 1991; Keasling, 2010; Woolston et al., 2013). A goal of the field is to develop and improve genetically modified microbial organism for biotechnological applications (Koffas et al., 1999; Becker and Wittmann, 2015; Liao et al., 2016). Main aspects in metabolic engineering are to understand the function of microbial metabolic networks, to design new metabolic networks, and to genetically encode the networks into microbes (Nielsen, 2001). Together, these aspects can form a **Metabolic Engineering Cycle** (Bailey, 1991) or synonymously a **Design-Build-Test-Learn** cycle (Turdo et al., 2022; Carbonell et al., 2018; Opgenorth et al., 2019) (Fig. 1.1). These cycles describe an iterative, practical approach, in which genetic modifications are rationally designed first, then implemented, the performance of the modifications like production rates measured, and finally the design adapted and improved. A motivation for such extensive optimization cycles in metabolic engineering is that the final microbial strains should allow for economically viable bioprocesses (Woolston et al., 2013).

Although the Metabolic Engineering Cycle is meant literally to optimize a specific system, each single step in the cycle can be the subject of dedicated studies that, for example, focus on improving genetic engineering methods (Larson et al., 2013; Qi et al., 2013; Garst et al., 2017; Na et al., 2013), investigating new bioprocessing concepts (Boecker et al., 2019; Klamt et al., 2018), advance bioinformatic tools (Monk et al., 2017; Norsigian et al., 2018; Xu et al., 2015; Nothias et al., 2020; Chambers et al., 2012), or focus on the precise measurements of products (Fuhrer et al., 2011; Guder et al., 2017; B. D. Bennett et al., 2009; L. Chen et al., 2021). However, as a starting point for most design efforts in metabolic engineering, detailed knowledge about metabolic networks is crucial.

1.1 Metabolic networks

Metabolism is a complex network of biochemical reactions, in which most of the reactions are enzymatically catalysed (Monk et al., 2017; A. C. Guo et al., 2012; Orth et al., 2011). For a single enzymatically-catalysed reaction, the reaction rates, at which the products and substrates are converted, depend on the abundance of the enzyme, the activity of the enzyme, and the concentrations of the substrates. The Michaelis-Menten-equation describes this dependency for a very simple case without inhibition or other

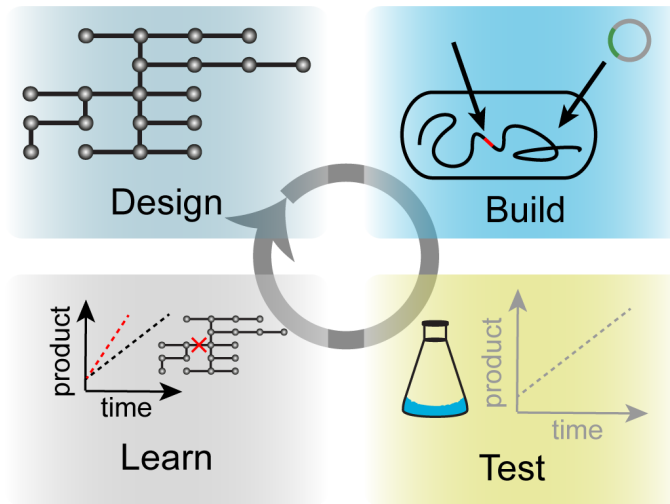


Figure 1.1: The Metabolic Engineering Cycle. Metabolic engineering often relies on an iterative optimization approach that is called the Metabolic Engineering Cycle or also Design-Build-Test-Learn cycle (Bailey, 1991; Gurdo et al., 2022; Carbonell et al., 2018; Opgenorth et al., 2019). Initial modifications to metabolic networks (native or synthetic) are rationally designed. After implementing these modification by genetic methods, the resulting production strains are tested. Based on these results, initial designs are changed and further improved.

effects (eq. 1.1, Michaelis et al., 2011; Rogers and Gibon, 2009). At low substrate concentrations $[S]$, the reaction rate v is very sensitive to changes in the substrate concentration (Fig. 1.2.a). However, with increasing substrate concentration, the reaction rate becomes more and more insensitive to changes in the substrate concentration and converges to a maximum reaction rate v_{max} . The maximum reaction rate also contains information about the enzyme abundance and is the product of the turnover number k_{cat} and the enzyme concentration $[E]$ (eq. 1.2). The Michaelis-Menten-constant K_m is the substrate concentration, at which the reaction rate is equal to half of v_{max} . Both, k_{cat} and K_m , values are unique properties of an enzyme, and they can vary largely among different enzymes. For example, fumarase can have a k_{cat} value of 1,150 1/s, whereas rubisco has a k_{cat} of 2.5 1/s (Rogers and Gibon, 2009).

$$v = v_{max} \cdot \frac{[S]}{K_m + [S]} \quad (1.1)$$

$$v_{max} = k_{cat} \cdot [E] \quad (1.2)$$

In a metabolic network, most biochemical reactions are only one element in chain of

sequential reactions, in which the product of one reaction is the substrate of another reaction. These sequential reactions are the **metabolic pathways** of a network, which make metabolism highly structured. Within a metabolic pathway, the reaction rates are usually in steady state (Fig. 1.2.b) and equal. Thus, the reaction rates in a pathway are also called the **metabolic flux**. The slowest reaction in a pathway, which is the rate limiting step, determines the flux. If the rate of a reaction drops below the otherwise common metabolic flux in the pathway, the pathway has a metabolic bottleneck and is not in steady state any more. Such a bottleneck is usually accompanied by the accumulation of the reaction's substrate (Donati et al., 2021, 2018, Chapter 4 and Chapter 5). To avoid bottlenecks and ensure robust synthesis of the building blocks, metabolism is often precisely adjusted and controlled by regulation mechanisms (Sander, Farke, et al., 2019; Matsuoka and Shimizu, 2011; Alon, 2007; Metallo and Vander Heiden, 2013). In *E. coli*, metabolite concentrations are kept relatively constant under different conditions and growth rates (Radoš et al., 2022), and the metabolic regulation is so effective that it requires major perturbations like starvation to cause metabolite levels to change substantially (Radoš et al., 2022). This is in stark contrast to protein concentrations that correlate with growth rates and are adjusted strongly dependent on the condition (Schmidt et al., 2016).

Regulation of metabolism occurs at the transcriptional, translational, or enzyme level (Fig. 1.3) (Donati et al., 2018). Sigma factors are one example for regulatory proteins. They are DNA-binding subunits of RNA polymerase and usually influence the transcription of a large number of genes. Thus, they are part of so-called global regulation. Sigma factors like σ^{70} (RpoD) of *E. coli* are even essential for cell function (Goodall et al., 2018; Baba et al., 2006). However, non-essential sigma factors like σ^{38} (RpoS) are also involved in large regulatory programs and often respond to major stresses to the cell (Schellhorn, 2020).

Alternatively, transcription factors are also DNA-binding proteins that can modulate the transcription of a gene. Many transcription factors are sensitive to environmental cues or stresses including temperature (Weber, 2003), pH (Barda et al., 2020), salts (Diray-Arce et al., 2019), light (E. M. Zhao et al., 2018), or nutrients (Kim et al., 2018) and can thus alter metabolism for the adaptation to a new environment. In some cases, transcription factors influence the expression of other, subordinate transcription factors and large parts of metabolism. These transcription factors are also called global regulators. Examples for global regulators are the transcription factors Crp and Cra that are responsible for coordinating carbon metabolism in *E. coli* (Matsuoka and Shimizu, 2011; Kim

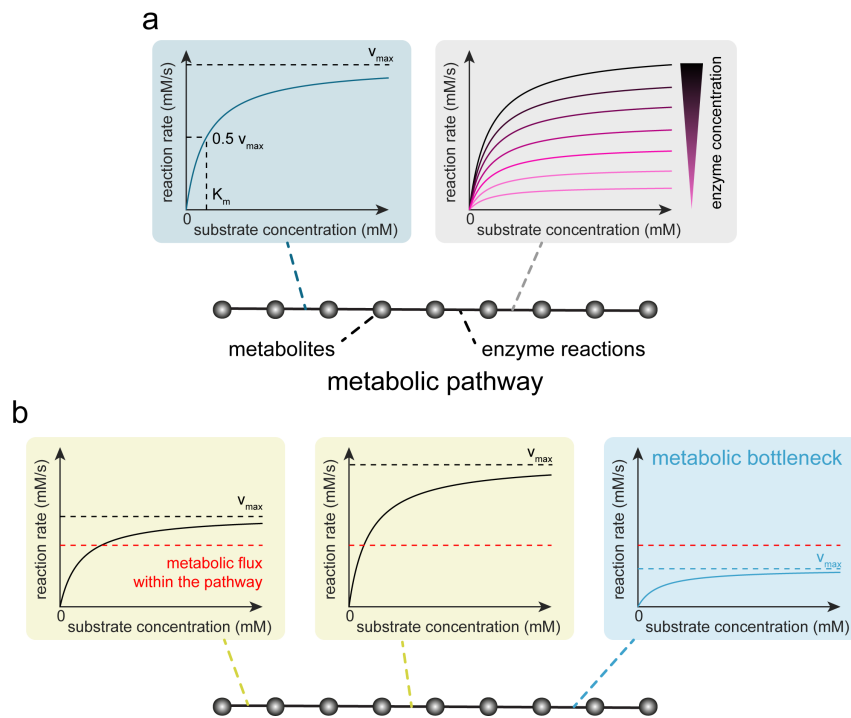


Figure 1.2: Michaelis-Menten-Theory in the context of metabolic pathways. (a) The chart in the blue box shows the relationship between the rate of an enzymatically catalysed reaction and the substrate concentration according to Michaelis et al., 2011. v_{max} is the maximum reaction rate. K_m is the Michaelis-Menten-constant and is the substrate concentration, at which the reaction rate is $0.5 \cdot v_{max}$. The chart in the grey box illustrates how the reaction rate is affected by the enzyme concentration. In metabolic pathways, the products of an enzyme reaction are usually the substrates for subsequent reactions. Products and substrates are thus metabolites. **(b)** The charts show Michaelis-Menten-kinetics similar to (a). The reaction rates in a metabolic pathway are usually in steady state, equal, and called metabolic flux (red lines). An enzyme reaction with a reaction rate below the usually common metabolic flux in a pathway is a metabolic bottleneck (blue box). (adapted from Schramm and Link, 2021)

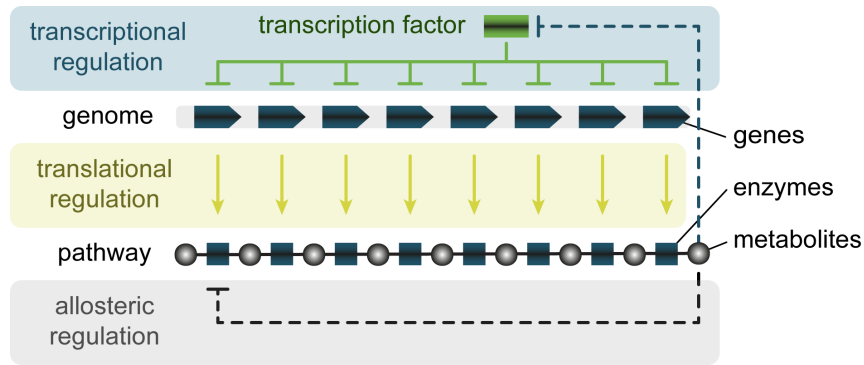


Figure 1.3: Feedback regulation of metabolic pathways. Metabolic feedback regulation is commonly observed in *E. coli* amino acid metabolism (Sander, Farke, et al., 2019). The regulation occurs at the transcriptional level by transcription factors, translational level by attenuation, or on enzyme level by direct allosteric interaction of metabolites.

et al., 2018; Perrenoud and Sauer, 2005). Other transcription factors operate on a much smaller scale and affect only genes from a particular metabolic pathway. In general, the regulation by transcription factors can be mapped by network approaches resulting in transcription networks (Alon, 2007; Ament et al., 2018). By combining global and local regulation, the allocation of resources in the metabolic network can be precisely controlled.

Transcription factors often interact allosterically with metabolites (Lempp et al., 2019) and thus are modulated in their activity. As described by the Michaelis-Menten-kinetic, enzyme concentrations influence reactions rates and consequently metabolite levels. By allosteric interactions with transcription factors, metabolites can in return influence gene expression and enzyme concentrations. As consequence, there is a mutual influence between metabolism and gene expression. This effect is also called crosstalk (Donati et al., 2018; Sander, Farke, et al., 2019; Grüning et al., 2010). Metabolites can not only interact with transcription factors to control gene expression but also with certain mRNAs to influence translation rates (Gollnick and Babitzke, 2002). This type of regulation is called attenuation.

In summary, metabolites can influence gene expression, and thus the concentration of an enzyme, by transcription factor-mediated regulation or by attenuation. However, allosteric interactions can occur also directly between metabolites and enzymes, which means that the gene expression and thus enzyme levels are not affected. Instead, the activity of an enzyme is modulated directly by the interaction with a metabolite.

Amino acid synthesis pathways in *E. coli* are well studied (Sander, Farke, et al., 2019) and examples of metabolic feedback regulation that uses both, regulation of enzyme lev-

els and regulation of enzyme activity (Fig. 1.3). In most amino acid synthesis pathways, the first enzyme is allosterically regulated by the concentration of the pathway's end product, the amino acid. The enzyme levels in the pathway are also controlled by the end product by transcriptional regulation. The interplay between both feedback mechanism ensures that the enzyme levels in the pathway are actually higher than required at the bare minimum to ensure the desired metabolic flux. Thus, enzymes in the pathway are not operating at full capacity and v_{max} . This effect is called enzyme overabundance and enables high robustness of the pathway against perturbations (Sander, Farke, et al., 2019).

The metabolic pathways in a cell are highly connected to form a large network. In literature, the network is often separated by two major functions, although a strict separation is not possible in reality (Russell and Cook, 1995). First, in catabolism, substrates like sugars are broken down to generate energy equivalents and precursor metabolites. These precursor metabolites are then used in anabolism for synthesis of a cell's building blocks, which include amino acids, nucleotides, lipids, peptidoglycan precursors, co-factors, and many more.

It was found that the core functions of metabolism including amino acid, energy, nucleotide, and lipid metabolism are highly conserved among species (Peregrín-Alvarez et al., 2009). This is astonishing, considering that demands on metabolism can vary greatly and that life can thrive within a broad range of environmental conditions (Merino et al., 2019) including pH values from at least -3.6 (Nordstrom et al., 2000) to 13 (Czop et al., 2011) and temperatures from -20 °C (Rivkina et al., 2000) to 122 °C (Takai et al., 2008). Since metabolism is highly conserved and genes that encode for enzymes are often homologous (Peregrín-Alvarez et al., 2009), genome sequencing and comparative genomics can thus reveal the topologies of metabolic networks (Durot et al., 2009; Haggart et al., 2011). In the short review *Von der Stöchiometrie zur Kontrolle metabolischer Netzwerke* (Chapter 2), we also provided a brief overview of the current state and opportunities in the study of bacterial metabolic networks and started with the topology.

In many cases, the topology of the metabolic network is so well understood that closed mass balances can be formulated and so-called genome-scale metabolic models constructed (Durot et al., 2009; Haggart et al., 2011), which contain all metabolic reactions of an organism. These models are available for a variety of organisms such as *Escherichia coli* (Monk et al., 2017; Orth et al., 2011), *Saccharomyces cerevisiae* (Lopes and Rocha, 2017), human cells (Brunk et al., 2018; Masid et al., 2020), and many more (Norsigian et al., 2018; H. Wang et al., 2021; Seif et al., 2018). As an example, the most recent genome-scale

metabolic model of *E. coli* contains 1,515 genes, 1,515 proteins, 1,192 metabolites, and 2,719 biochemical reactions (Monk et al., 2017). In recent years, genome-scale metabolic models were also created that cover multiple strains of a species and are mainly used as highly curated reference model or to study inter-strain variations (Seif et al., 2018; Norsigian et al., 2020; Monk et al., 2013).

In the context of metabolic engineering, genome-scale metabolic models are very important because they can be used to guide the design process in the Metabolic Engineering Cycle (Curran and Alper, 2012). One method, which relies on genome-scale metabolic models and is often used, is **flux balance analysis** (Schilling et al., 1999; Edwards and Palsson, 2000; Edwards et al., 2002). This method enables the estimation of fluxes in the network. A more experimental approach to determine fluxes in metabolism is **metabolic flux analysis** (Antoniewicz, 2015; Zamboni et al., 2009; Wiechert, 2001). It combines genome-scale metabolic models with isotope tracer experiments. Most often the stable ^{13}C -isotope is used to label metabolites, and the labelling patterns and dynamics can then be measured by mass spectrometry-based metabolomics and reveals the carbon flow within the network.

1.2 Design strategies in metabolic engineering

By access to genome-scale metabolic models and estimation of flux distributions, the design process in metabolic engineering can be guided. However, it is often useful to start with a strain that can already overproduce the product metabolite and that is used as a basis for further optimization. To create such an initial overproduction strain, there are some general design strategies in metabolic engineering (Fig. 1.4). Since microbial metabolism is extensively regulated, and metabolite concentrations remain stable over a broad range of conditions (Radoš et al., 2022), genetic modifications are required to break metabolite homeostasis and to overproduce metabolites.

Pathway dysregulation In case of the feedback regulated arginine synthesis pathway of *E. coli*, it was possible to remove both, allosteric and transcriptional regulation, and overproduce arginine to high concentrations (1.4.a) (Sander, Wang, et al., 2019). In this work, we also present two studies, in which pathway dysregulation in the arginine pathway was pivotal for high level productions of citrulline (Chapter 4) and arginine (Chapter 5). Removing the feedback regulation of pathways worked also in many other pathways (Lee and Wendisch, 2017). Notably, it was also shown that it can be useful to not entirely remove feedback regulation by

the transcriptional repressor ArgR but to reduce its level and only partially dysregulate the arginine pathway (Sander, Wang, et al., 2019). The reason for this was that full dysregulation of the pathways drained carbamoyl phosphate and caused a limitation in the pyrimidine pathway that also requires carbamoyl phosphate. Due to this limitation, the production performance was sub-optimal.

Synthetic pathways Another design strategy is to create entirely new pathways (Fig. 1.4.b): by expressing heterologous genes, the native metabolic network of a strain can be extended. This enables the overproduction of foreign metabolites such as opioids or cannabinoids in yeast (Galanie et al., 2015; Luo et al., 2019), glycerol in *E. coli* (C.-Y. Wang et al., 2021), or the use of CO₂ as carbon source in *E. coli* (Gleizer et al., 2019). However, since the native metabolic network does often not support the synthetic pathways, additional genetic modifications are required to compensate detrimental effects of the synthetic pathway expression.

Removing competing pathways By adding a synthetic pathway, we often create a new branching point in the metabolic pathway such that a substrate can be converted further by two different routes. Many of these branching points exist also natively in metabolic networks. Thus, the production pathways can perform below full capacity because other pathways compete for the same substrates. By removing such competing pathways, flux can be directed into the production pathway (Fig. 1.4.c) (Koffas et al., 1999; Ma et al., 2017). It is not always trivial to identify competing pathways and to optimize fluxes in the metabolic network, which is a reason why flux balance analysis and metabolic flux analysis are very valuable tools.

Blocking product degradation End products of metabolic pathways are usually used as building blocks or substrates in other metabolic processes. In case of amino acids, there are even additional pathways that are expressed upon high concentrations of the metabolite and that focus solely on their degradation (Zampieri et al., 2019). Removing degradation pathways can thus improve the overproduction (Fig. 1.4.d) (Ma et al., 2017).

Product export If the intracellular product concentration reaches very high levels, the product can become toxic or inhibit the production pathway (Keasling, 2010). As a solution, it is possible to engineer and overexpress transport systems that export the product outside the cell (Fig. 1.4.e) (Ma et al., 2017). In a here presented study,

we also applied this strategy (Chapter 5) and overexpressed the arginine exporter ArgO in an arginine overproduction strain.

Cofactor engineering Many production pathways require an extensive amount of reducing equivalents or other cofactors. Improving cofactor availability and engineering the cofactor metabolism can thus enhance the production (Koffas et al., 1999; Ma et al., 2017; Chemler et al., 2010). For example, by creating and using a glycerol 3-phosphate dehydrogenase that can use both NADH and NADPH, lysine production in *Corynebacterium glutamicum* could be improved by 60 % (Bommareddy et al., 2014).

Alternative substrates Many microbial overproduction strains are developed that use sugars like D-glucose, which can also be used as food or feed. On large scale, this could lead to high substrate costs making the bioprocess economically infeasible. Therefore, many metabolic engineering studies focus on the development of base strains that can use cheap alternative substrates that do not compete with food or feed industry (Ferreira et al., 2020). These substrates can be waste biomasses or alternative sugars like arabinose and xylose that can be generated from cellulose and hemicellulose (Ma et al., 2017).

1.3 Lambda Red recombineering and barcoded, multiplexed CRISPR/Cas9 genome editing in *E. coli*

Having a set of design strategies and genome-scale metabolic models for further optimization attempts, the next challenge is to physically encode any proposed modifications into the genome of a microbial production strain. Here, we focus on two key approaches in the genome editing of *E. coli*.

In many cases, it is useful to remove native genes encoding for transcription factors or unwanted enzymes. To achieve targeted gene deletions (knockouts) in *E. coli*, homologous recombination is the most common approach and usually coupled to a selective pressure to remove the gene. Since wild type *E. coli* is not proficient of homologous recombination, recombinases need to be heterologously expressed to achieve high cloning efficiencies. The Red recombinase of the *Escherichia virus Lambda* proved to be very useful (Y. Zhang et al., 2000; Datsenko and Wanner, 2000). Genome editing in *E. coli* that relies on Lambda Red is also referred to as **Lambda Red recombineering**.

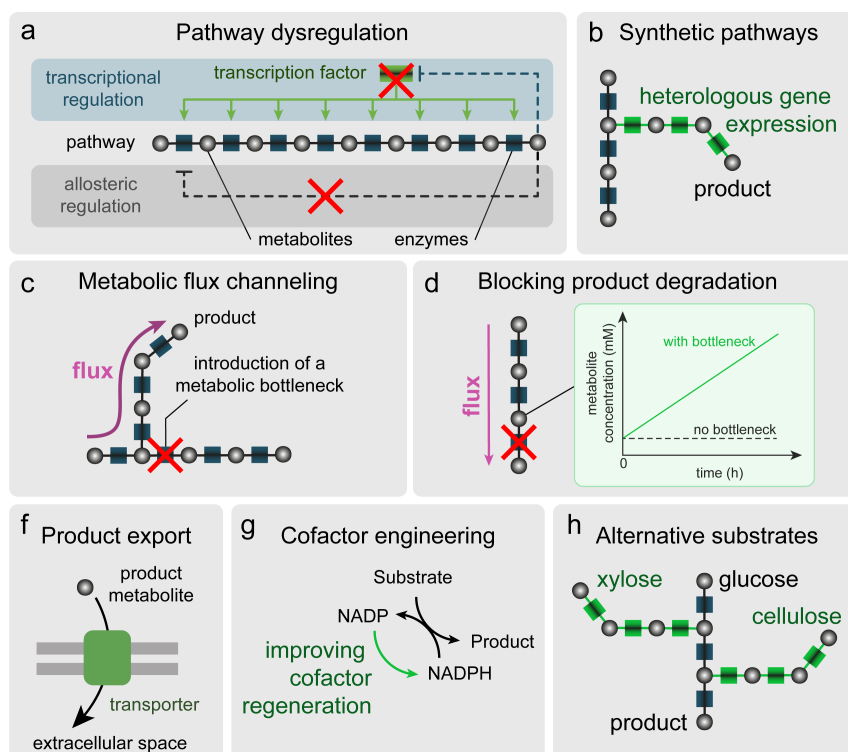


Figure 1.4: Common design strategies in metabolic engineering. (a) Removing native regulation can lead to overproduction of the pathway's end product (Sander, Farke, et al., 2019). (b) Heterologous gene expression enables the creation of synthetic metabolic networks with new products (Galanie et al., 2015; Luo et al., 2019; C.-Y. Wang et al., 2021; Gleizer et al., 2019). (c) To maximize flux in the production pathway, competing pathways can be removed at branching points (Koffas et al., 1999; Ma et al., 2017). (d) Often, the product metabolites are degraded or further processed in the native metabolic network. Removing subsequent reactions can improve the production (Ma et al., 2017). (e) Intracellular product concentrations can reach toxic levels inhibiting the production. Overexpression or engineering product export systems can help (Ma et al., 2017). (f) Some pathways require extensive amounts of cofactors such that enhanced cofactor synthesis can improve the production (Koffas et al., 1999; Chemler et al., 2010; Bom-mareddy et al., 2014). (g) The economic feasibility can be enhanced by using alternative and cheaper substrates for the bioprocess (Ma et al., 2017).

The Lambda Red recombineering method by Datsenko and Wanner was pivotal in the creation of a large *E. coli* strain library bearing a gene knockout of every non-essential gene (KEIO collection, Baba et al., 2006; Datsenko and Wanner, 2000). In this method, a linear template DNA was created directly by PCR without further cloning that had an antibiotic resistance gene and at each flanking ends two special sites: the inner regions were flipase recognition sites, the outer were regions homologous to regions on the *E. coli* genome (Fig. 1.5.a). Cells were then transformed with the template DNA and subjected to the antibiotic according to the resistance gene on the DNA template. By Lambda Red-mediated homologous recombination, a target gene was then replaced by the template DNA. Optionally, the antibiotic gene could be removed using the FRT sites and a flipase, which left a single FRT site as a scar on the genome. If the antibiotic gene is not removed from the genome, it is very easy to propagate a gene knockout in one *E. coli* strain to another one using P1 phage transduction (Thomason et al., 2007). The KEIO collection has been a very important resource and contributed substantially to metabolic engineering and systems biology. A major limitation of the KEIO collection is however that essential genes are not covered because a deletion would be lethal to the organism.

In recent years, another powerful method was developed based on **CRISPR/Cas9** by many scientist around Virginijus Siksnys, Emmanuelle Charpentier, Jennifer A. Doudna, Feng Zhang, Blake Wiedenheft, and George Church (Lander, 2016; Ledford, 2016). For their discoveries, Charpentier and Doudna received the Nobel Prize in Chemistry 2020. Originally, CRISPR/Cas9 is part of a bacterial defence system protecting cells from foreign DNA (Haurwitz et al., 2010). The single guide RNA (sgRNA) of the CRISPR/Cas9 system has a so-called protospacer, which is homologous to a target DNA, and a secondary structure interacting and binding to the endonuclease Cas9. By the protospacer, the sgRNA-Cas9 complex is then guided to the target DNA site, where the Cas9 cuts the DNA if a small DNA motif (protospacer adjacent motif: PAM) is present. However, the CRISPR/Cas9 system can be repurposed for genetic engineering, for which it provides selective pressure by double strand brakes to modify the target DNA. In most cases, homologous recombination is used for introducing edits to the genome yet other mechanism exist, such as base-editing and prime-editing, that do not require Cas9-mediated cutting of the DNA (Kantor et al., 2020).

In 2017, Garst *et al.* reported a **barcoded, multiplexed CRISPR/Cas9 genome editing** approach that aims at introducing thousands of different single codon mutations to the genome of *E. coli* simultaneously (Fig. 1.5) (Garst et al., 2017). This approach can make use of oligonucleotide pools and thus enables a high degree of multiplexing. In

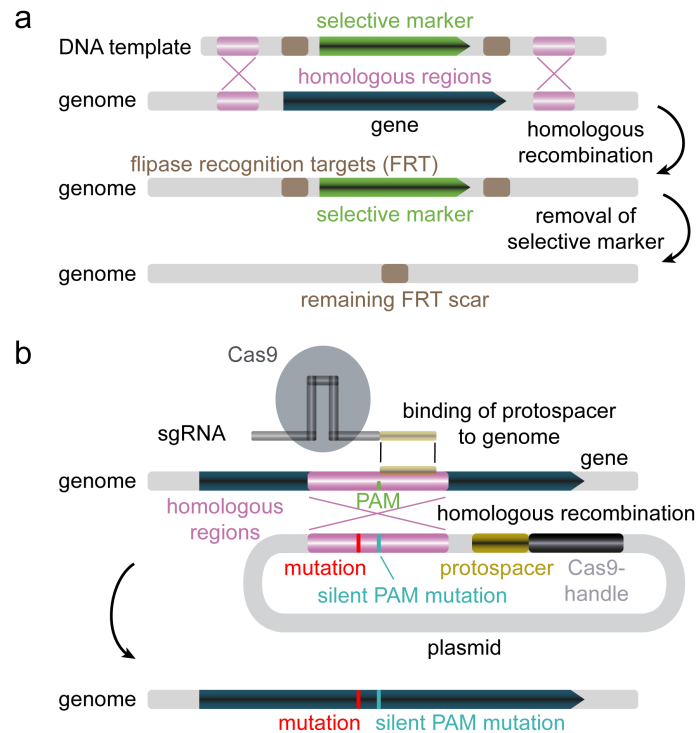


Figure 1.5: Two genome editing methods in *E. coli*. (a) The approach by Datsenko and Wanner to delete a genomic gene in *E. coli* is based on homologous recombination by Lambda Red (Datsenko and Wanner, 2000). A template DNA is provided that carries a gene encoding a selective marker and two flanking regions at each of the 5'- and 3'-ends that are flipase recognition sites (FRT) and regions that are homologous to a genomic target site. Upon applying selective pressure and by homologous recombination, the genomic target gene is replaced by the template DNA. Using the FRT sites and flipase, the selective marker can be removed in a second step leaving a single FRT site as scar on the genome. (b) By barcoded CRISPR/Cas9 genome editing, it is possible to mutate thousands of single codons on the *E. coli* genome in one batch (Garst et al., 2017). Homologous DNA with the desired mutation and a silent PAM mutation is provided on a plasmid. The plasmid also carries the sgRNA for the CRISPR/Cas9 system in close proximity to the homologous DNA such that oligonucleotide pools can be used for cloning. By expressing the CRISPR/Cas9 strain and Lambda Red, the mutations are introduced to the genome by homologous recombination. Since the sgRNA-carrying plasmid is maintained throughout the experiments, it can function as NGS-barcodes of the mutant strains.

the method, the template DNA is provided on a plasmid for Lambda Red recombineering. Since the plasmid is maintained throughout all further experiments, it can function as a barcode for amplicon deep sequencing. The power of this method is that, due to the deep sequencing of the barcodes, the abundance of single mutants in pooled strain libraries can be measured. With this method, saturation mutagenesis of a gene can be achieved and all possible single amino acid substitutions of a protein tested (Garst et al., 2017). Several studies created large strain libraries by the method, and it was investigated, which out of over 16,000 mutations in the *E. coli* lysine pathway can contribute to resistance against the cell toxic L-lysine analog S-(2-aminoethyl)-L-cysteine (Bassalo et al., 2018). In another study, 110,120 mutations in 82 regulators were created in *E. coli* to study regulation on a global scale (R. Liu et al., 2020). Five strain libraries, with each 5,260 to 7,340 mutations in different global regulators, were used to investigate antibiotic resistance in *E. coli* (C. Chen et al., 2020). Additional two libraries with around 40,000 and 162,000 members were used in metabolic engineering attempts to improve 3-hydroxypropionic acid production (R. Liu, Liang, Choudhury, et al., 2018; R. Liu, Liang, Garst, et al., 2018). In the latter two studies, several iterative rounds of recombineering were applied. In one of the here presented studies, we adapted and improved the barcoded, multiplexed CRISPR/Cas9 approach (Chapter 5).

2 Mass spectrometry-based metabolomics

The goal of metabolomics is to study and quantify small molecules up to circa 1500 Da that are part of biological processes (Markley et al., 2017). The concentrations of metabolites provide insights into the function of metabolic networks and can be biomarkers for disease (Koulman et al., 2009). They can also contain information about many other things like whether an athlete is doping (Keen et al., 2022), the composition of food (Adebo et al., 2017), or whether environments are polluted (Bundy et al., 2009).

The number of unique metabolites on earth is unknown. According to the latest genome-scale metabolic model, there are 1,192 metabolites in the bacterium *E. coli* (Monk et al., 2017). This number is however dwarfed by the expected number of unique metabolites in human, which are over 100,000 (Markley et al., 2017; Alseekh and Fernie, 2018). It is estimated that between 100,000 and 1,000,000 metabolites are in the plant kingdom (Alseekh and Fernie, 2018). Large databases like the Human Metabolome Database (Wishart et al., 2022), the Global Natural Products Social Molecular Networking platform (M. Wang et al., 2016), or METLIN (Guijas et al., 2018) collect information about

measured metabolites and are valuable reference resources.

Metabolomics includes a wide range of analytical methods, and many different techniques have been used over the past. Already in ancient Greek, India, and China, the glucose levels in the urine of a patient were measured by the sweetness of the urine and used to diagnose diabetes (Karamanou et al., 2016). Today's methods are more selective and quantitative. Mass spectrometry (MS) is currently the dominant analytical approach in metabolomics followed by nuclear magnetic resonance (NMR) spectroscopy (Markley et al., 2017; Raftery, 2014; A.-H. Emwas et al., 2019; A.-H. M. Emwas, 2015).

In brief, advantages of NMR over MS are that sample preparation is even simpler, samples are not destroyed during the analysis, measurements are more quantitative and reproducible, and isotopomers can be easily distinguished (Markley et al., 2017; A.-H. Emwas et al., 2019). A crucial disadvantage is however that NMR is much less sensitive than MS and often can only detect metabolites with concentrations greater than 1 μM (A.-H. Emwas et al., 2019). Thus, sample volumes in NMR can be much larger than in MS, and the number of quantified metabolites much smaller. At the same time, MS offers high selectivity, especially if it is combined with chromatography.

The beginnings of MS can be traced back to the early 20th century when physicists were studying electrons and the masses of atoms (Griffiths, 2008). The fundamental principle of MS is to separate ions under vacuum by their mass to charge ratio (m/z) using electromagnetic fields. Although MS is mainly used in analytics nowadays, it is possible to use the principles of MS to enrich ions with a specific mass. This technique is called preparative mass spectrometry and was used during the Manhattan project to enrich the uranium isotope 235 (Griffiths, 2008; Yergey and Yergey, 1997). In modern MS approaches, there are four major components that frequently differ between instruments (Fig. 1.6): (1) separation by chromatography prior to analysis, (2) ion sources, (3) collision induced dissociation, and (4) mass analyser.

(1) Chromatography coupled to MS In most cases, metabolites are separated by chromatography prior to analysis by MS (Fig. 1.6). This increases the selectivity but can also decrease negative effects in the ion source like ion suppression and space charge effects (Alseekh et al., 2021). High performance liquid chromatography (HPLC) and gas chromatography (GC) are the dominant techniques, followed by capillary electrophoresis (CE) (Alseekh et al., 2021), whereas supercritical fluid chromatography (SFC) remained a niche application so far (van de Velde et al., 2020). Capillary electrophoresis separates metabolites in a liquid phase and uses an electric field to separate polar or charged compounds within a capillary (W.

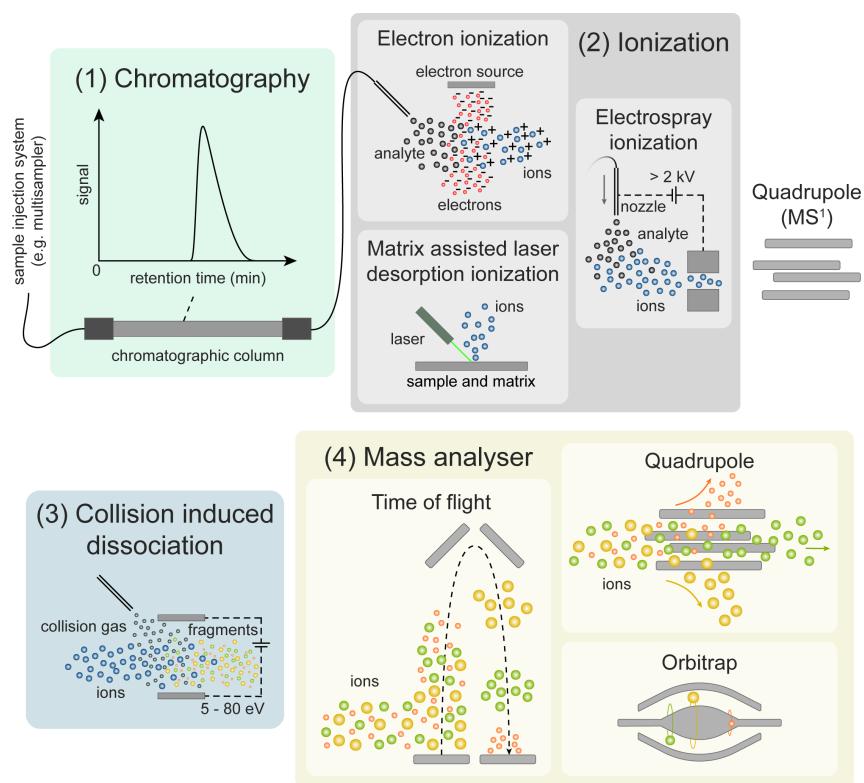


Figure 1.6: Technical configurations of mass spectrometers. This figure gives an overview of different optional configurations of current mass spectrometers. **(1)** MS is often coupled with chromatography for metabolite separation prior to analysis by MS (Bjerrum, 2015). **(2)** The analyte molecules need to be ionized for MS analysis. Common ionization techniques of MS approaches for metabolomics are electron ionization (Famiglini et al., 2021), matrix-assisted desorption ionization (Jurinke et al., 2004), and electrospray ionization (Wilm, 2011). **(3)** Optionally, mass spectrometers have a collision cell, in which ions are accelerated, collide with an inert gas, and fragment. This procedure is called collision induced dissociation (CID) (Martin Somer et al., 2020). Prior to CID parent ions are usually selected by a quadrupole. **(4)** Three different mass analyser types are commonly used in MS methods for metabolomics: time-of-flight mass analyser (Boesl, 2017), quadrupoles (Glish and Vachet, 2003), and orbitraps (Zubarev and Makarov, 2013).

Zhang and Ramautar, 2021). Gas chromatography is used for separation of volatile and low molecular weight molecules in a gas phase (Alseekh et al., 2021). GC often requires derivatisation of non-volatile compounds (Bjerrum, 2015; Gowda and Djukovic, 2014). HPLC does not require derivatisation and can use a variety of different mobile phases like water, isopropanol, methanol, or acetonitril. It has access to a broad range of different chromatographic columns. This makes it possible to measure unpolar compounds, lipids, and small polar metabolites by HPLC. Reverse phase columns are often C8- or C18-based and used for separation of unpolar metabolites (Bjerrum, 2015), whereas polar metabolites can be separated by normal phase columns, which can rely on hydrophilic interactions (Perez de Souza et al., 2021).

(2) Ion sources Metabolites need to be ionized before they can enter the mass analyser.

Common ionization techniques include electron ionization (EI), matrix-assisted laser desorption ionization (MALDI), and electrospray ionization (ESI) (Fig. 1.6) (Famiglini et al., 2021; Jurinke et al., 2004; Wilm, 2011). EI was one of the earlier developed ionization techniques and mainly used for small volatile compounds in gas or liquid phases (Famiglini et al., 2021). EI uses electrons with ~ 70 eV energy to ionize the analyte. During EI, only positively charged ions of the analyte can be generated, and metabolites are often fragmented due to the high energy of the electrons. Thus, EI is considered a hard ionization technique. In contrast, MALDI and ESI are considered soft ionization techniques since they mainly leave the analytes intact (Jurinke et al., 2004; Steckel and Schlosser, 2019). The samples for MALDI are usually embedded into a solid matrix of organic compounds (Jurinke et al., 2004). Irradiating the matrix with laser beams leads to a desorption effect and generates the sample ions. MALDI finds application in DNA analysis (Jurinke et al., 2004), but it also enables spatial analyses of tissues, cells, bones, and more (C. Zhao and Cai, 2022; Good et al., 2022). The two or three dimensional analysis by mass spectrometry is also called imaging mass spectrometry (Good et al., 2022; Schnackenberg et al., 2022). During ESI, a continuous flow of solvent that carries the sample is dispersed into droplets by applying a high voltage of typically more than 2000 V to the nozzle of the injection capillary and vaporized by heat (Wilm, 2011; Glish and Vachet, 2003; Fenn et al., 1989, 1990; Luedtke et al., 2008). ESI can produce positively and negatively charged ions and is an ambient pressure ionization technique.

(3) Collision induced dissociation Some mass spectrometers can select specific ions (parent ions) prior to detection and fragment them by collision induced dissociation (CID) (Fig. 1.6) (Martin Somer et al., 2020): the parent ions are accelerated to energies in the range of 5 to 80 eV. By collision with an inert gas like nitrogen or noble gases, the ions fragment into product ions, which are then analysed (Glish and Vachet, 2003). The fragmentation pattern are highly specific and thus provide a high degree of confidence that a certain metabolite is measured and not a false-positive. Quadrupoles are typically used for the selection of parent ions. Since both, the parent ion selection and the analysis of the product ions, are performed by MS, this approach is also called tandem mass spectrometry, MS^2 , or MS/MS. Using linear ion traps, it is even possible to fragment product ions further in sequential fragmentation steps (multistage MS, MS^n) (Sun et al., 2018). Using CID, it is possible to perform multiple reaction monitoring (MRM), in which, repeatedly, a list of parent masses are sequentially selected, fragmented, and the product ions measured. This enables the simultaneous measured of multiple metabolites with high selectivity.

(4) Mass analyser Although all mass spectrometers have in common that they distinguish ions by their mass over charge, there are very different mass analyser designs. Common analysers with different functional principles are quadrupoles, time-of-flight analysers, and orbitraps (Fig. 1.6) (Glish and Vachet, 2003). Quadrupole mass analyser use four rods that have high electric currents alternating at specific frequencies to deflect all ions with undesired m/z values (Glish and Vachet, 2003). Only ions within a selected m/z range have trajectories that can pass the quadrupole. Quadrupoles are often used to select specific ions for CID and can be coupled with other mass analysers. Quadrupole time-of-flight mass spectrometers (QTOF) and quadrupole-orbitrap mass spectrometers are common. Since collision cells for CID are typically also quadrupoles, tandem mass spectrometers with two quadrupole mass analyser and a quadrupole collision cell are called triple quadrupole mass spectrometers (QQQ). Orbitraps are Fourier transform type mass spectrometers, in which ions rotate within a radial electric field around an inner spindle-like electrode (Zubarev and Makarov, 2013). Depending on the m/z , ions show different oscillations axial to the inner electrode (Zubarev and Makarov, 2013). Time-of-flight (TOF) mass analyser accelerate ions in batches within an electric field (Boesl, 2017). Depending on their m/z , ions arrive at different time points at the detector. Therefore, knowing the flight time of an ion reveals

its m/z . TOF mass spectrometers can easily be recognized by their characteristic flight tube. The most recent major development in MS is the trapped ion mobility spectrometry (TIMS) that is usually coupled to a TOF mass analyser (TIMS-TOF) (Jeanne Dit Fouque and Fernandez-Lima, 2019; Ridgeway et al., 2019). In TIMS, ions are kept stationary in a linear ion trap, whereas a gas flows through the trap (Jeanne Dit Fouque and Fernandez-Lima, 2019; Ridgeway et al., 2019). Depending on the molecule size and shape, ions are retained or washed out of the trap towards the mass analyser. Thus, TIMS data provides a retention time, similar to chromatography. The strength of this method is that many isomers can be separated, which are difficult to resolve by chromatography. A current limitation of TIMS in metabolomics is however its lower sensitivity compared to mass spectrometers without TIMS. For all the different mass analysers, it is possible to perform targeted or untargeted analyses. In targeted MS, prior knowledge about the m/z values of a metabolite is required, whereas untargeted cover entire ranges of m/z values. Targeted approaches can focus on quantitative measurements of known metabolites (Guder et al., 2017; Ribbenstedt et al., 2018). Untargeted approaches are useful in the natural product discovery (Leao et al., 2021; Gauglitz et al., 2022).

Metabolism can change within seconds (Link et al., 2013). Thus, sampling for metabolomics has to be performed rapidly and usually involves quenching of metabolic reactions. Common methods involve the quenching with $-20\text{ }^{\circ}\text{C}$ cold 40:40:20 acetonitrile:methanol:water (Harwood and Wipat, 2012) or $80\text{ }^{\circ}\text{C}$ hot ethanol (van Gulik, 2010). It is also possible to measure intracellular metabolites by (fast-)filtration methods (Harwood and Wipat, 2012), extracellular metabolites by removing cells using centrifugation, or metabolites in the whole culture broth (van Gulik, 2010).

Metabolomics plays an important role in metabolic engineering since it provides measures for the performance of a bioprocess. Production rates, substrate import rates, by-product formation but also more intricate metabolite data and metabolic flux analysis guide design efforts in metabolic engineering. With high-throughput cloning techniques like pooled CRISPR/Cas9 genome editing or even error-prone-PCR, it is easily possible to generate hundreds of thousands of strains (R. Liu, Liang, Choudhury, et al., 2018; R. Liu, Liang, Garst, et al., 2018; Schramm et al., 2020). However, the analysis of only ten thousand strains is already very difficult with current methods, and faster metabolomics methods are required.

In 2017, we published a fast, quantitative metabolomics method that uses isotope ra-

tios (Guder et al., 2017). In brief, ^{12}C -metabolite samples are mixed with a fully labelled ^{13}C internal standard sample before analysis by LC-MS/MS. Calibrating the ^{13}C internal standard with authentic metabolite standards enabled the calculation of absolute metabolite concentrations. Using the ratio between the ^{12}C -signal and the ^{13}C -signal of a metabolite for quantification improves the linearity (Wu et al., 2005) and provided additional confidence in the metabolite annotation. This allowed us to achieve run times of 2 min per injection.

Another approach enables even faster measurements. Injecting samples without chromatographic separation into a continuous flow of a mobile phase and, subsequently, into super-high-resolution TOF mass spectrometers achieves measurement times in the second range (Fuhrer et al., 2011). This analysis approach is also called flow-injection mass spectrometry (FI-MS). Although, ESI is considered a soft ionization method, in-source fragmentation of metabolites is still a commonly observed phenomenon (Guder et al., 2017; J. Guo et al., 2021; Senan et al., 2019). A sample arrives at the ESI in a single plug during FI-MS. We thus were interested about in-source fragmentation during FI-MS and whether it poses a risk for metabolite missannotations. In this work, we present a study about the systematic identification of in-source modifications by a network approach (Chapter 3).

3 Decoupling of growth and overproduction

The performance of bioprocesses is mainly characterized by product **Titers**, production **Rates**, and product **Yields**. Together, they are the **TRY criterion** (Nielsen and Keasling, 2016). In bioprocessing, microbial cells convert a substrate into the desired product. However, the substrate is also consumed to generate microbial cells (biomass). We thus face a trade-off between biomass and product formation (Klamt et al., 2018). This results either in low product yields but high production rates or in high product yields but low production rates (Fig. 1.7.a). One solution to this challenge is to dynamically control microbial metabolism and to decouple the cell growth and the overproduction of chemicals (Lalwani et al., 2018). By switching between different metabolic states, two-stage bioprocesses can be achieved (Fig. 1.7.b) (Hartline et al., 2021): in a first stage, biomass accumulates by cell growth without product formation. In a second stage, growth is arrested, and the production of the desired chemical induced. Using two-stage bioprocesses, titers, rates, and yields can be improved beyond what one-stage bioprocesses can achieve (Klamt et al., 2018).

There are two major challenges in creating two-stage bioprocesses: one challenge is to dynamically control cell growth and metabolism, and the second is to engineer metabolism to be highly active under growth arrest.

3.1 Dynamic control of cell growth and metabolism

To create a transition between two bioprocess stages, cell growth and metabolism needs to be dynamically controlled (Hartline et al., 2021; Brockman and Prather, 2015a; Venayak et al., 2015; Cress et al., 2015; Burg et al., 2016). Given the great innovations in the genetic engineering during the last decade, we now have a plethora of options to control cellular metabolism dynamically. Most genetic systems have in common that they require an input signal for induction, and, in the context of metabolic engineering, many different input signals have been studied: chemical inducers were used to control expression of a single gene (Soma et al., 2014) or to control a CRISPR interference system that subsequently controls the transcription of a target gene (S. Li et al., 2016). A quorum sensing system allows for cell density dependent autoinduction of a two-stage process (Gupta et al., 2017), and also light can be used to control a bioprocess (E. M. Zhao et al., 2018). Some of the systems allow the process operator to actively control the bioprocess, and others like the quorum-sensing based system are auto-regulated (Gupta et al., 2017; Lo et al., 2016). Although these approaches enable the dynamic control of metabolism,

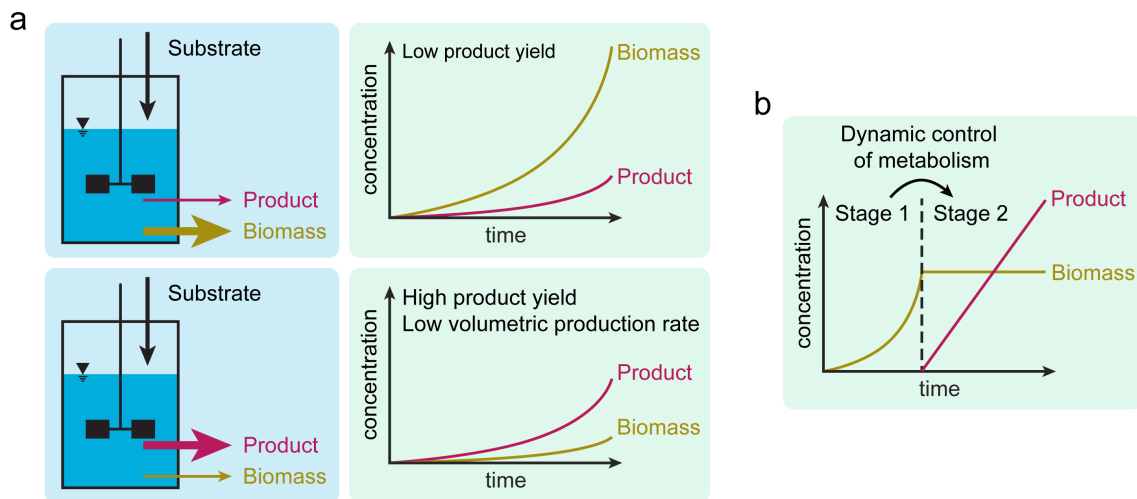


Figure 1.7: Concept figure of one- and two-stage bioprocesses. (a) The upper scheme shows a theoretical one-stage bioprocess, in which most of the substrate is converted into biomass, and few substrate into the product. The biomass concentration reaches high levels over the time, whereas the product concentrations remain low. The substrate-specific product yield is low. In the lower scheme, a one-stage bioprocess is depicted, in which most of the substrate is converted into the product, and not into the biomass. The biomass, which functions as a catalyst and converts the substrate into the product, remains low throughout the cultivation causing low volumetric production rates. The substrate-specific product yield is high. (b) The chart shows the biomass and product concentrations over time for a theoretical two-stage bioprocess. In the first stage, only biomass accumulates. Then, by dynamically controlling the metabolism of the overproduction strain, growth is arrested, and the production is induced in the second stage of the bioprocess. Consequently, the product accumulates.

a limitation of many of the genetic systems is that they cannot arrest cell growth fully, which is desired in two-stage bioprocesses.

A stable growth arrest can be effectively achieved by nutrient limitations and starvation. For example, growth is arrested by omitting essential nutrients such as phosphorus, sulphur, magnesium, and nitrogen sources from the medium (Chubukov and Sauer, 2014). Especially, nitrogen limitation was often used to stop growth in metabolic engineering projects to study cells under growth arrest (Boecker et al., 2019; Chubukov et al., 2017; Michalowski et al., 2017; Sonderegger et al., 2005). A practical limitation of this approach is however that nutrients cannot be easily removed from the cultivation medium and need to be diluted or depleted during the cultivation.

Another option to robustly arrest growth and create two-stage bioprocesses is to use gene deletion strains that are auxotrophic for specific metabolites. Then, growth can be controlled by supplementing or omitting these specific metabolites. In one of the here presented studies (Chapter 4), we used this approach to investigate the citrulline overproduction under growth arrest in *E. coli* strains with different feedback regulations removed. Dependent on the price of the fed metabolites, this approach can however face economic drawbacks. A solution to avoid supplementations with expensive metabolites is to dynamically introduce auxotrophies during the fermentation (Gupta et al., 2017).

An important concept in metabolic engineering is the **metabolic valve** (Gupta et al., 2017; Brockman and Prather, 2015b; Solomon et al., 2012; Venayak et al., 2018, Chapter 4 and Chapter 5). A metabolic valve is an enzyme reaction that can be dynamically tuned during the bioprocess and that introduces a metabolic bottleneck on demand. This can be achieved by modulating the enzyme abundance or the enzyme activity (also see Section 1.1). Blocking the metabolic flux at a specific location in the metabolic pathway can have multiple benefits like shutting-off degradation or competing pathways (also see Section 1.2). Yet, it can also lead to the accumulation of the substrate of the perturbed enzyme reaction (Donati et al., 2021, 2018; Gupta et al., 2017, Chapter 4 and Chapter 5) and to a growth arrest if a downstream metabolite is essential (Harder et al., 2018; Cho et al., 2012, Chapter 4 and Chapter 5). Thus, converting an essential enzyme into a metabolic valve allows us to dynamically introduce auxotrophies and to control growth and metabolism.

However, there are also alternative approaches to arrest growth that do not rely on nutrient limitations or metabolic bottlenecks. These approaches are especially interesting for modifying conventional overproduction strain that usually operate in one-stage bioprocesses and for making them applicable in two-stage bioprocesses. Cell growth can be arrested by overexpressing the small RNA Rcd in an *E. coli* strain with a mutation in

the global transcription factor H-NS (Rowe and Summers, 1999; Hommais et al., 2001). This enables the overproduction of proteins under growth arrest. These metabolically active but non-growing cells were called "quiescent" (Rowe and Summers, 1999). Further, inducible RNA polymerase expression (*rpoB* and *rpoC*) can be used to control growth of *E. coli* cells (Izard et al., 2015). This enabled the construction of a two-stage bioprocess producing glycerol in *E. coli* that had improved glycerol yields (Izard et al., 2015). Targeting genes related DNA replication (*dnaA* and *oriC*) by CRISPR interference allowed also for controlling growth of *E. coli* (S. Li et al., 2016).

3.2 Metabolic activity under growth arrest

Besides achieving tight control over metabolism and cell growth, there is another challenge during the development of two-stage bioprocesses: microbial cells in growth arrest often show lower metabolic rates than growing cells (Chubukov and Sauer, 2014; Chubukov et al., 2017; Sonderegger et al., 2005; Chesbro et al., 1990; Nyström, 1998; Venturi, 2003). Low metabolic rates can cause low production rates and are thus undesired in bioprocessing. In contrast, cells that are metabolically highly active under growth arrest are sought-after and called quiescent (Sonderegger et al., 2005). The substrate uptake rates are often used as a proxy for metabolic activity under growth arrest, and they can vary greatly between different types of nutrient limitations (Chubukov et al., 2014). In *E. coli* for example, nitrogen starvation resulted in the lowest glucose uptake rates compared to phosphate, sulfur, magnesium, tryptophan, and leucine starvation (Chubukov and Sauer, 2014). It was shown that *E. coli* can produce proteins under growth arrest at a constant rate for multiple days (Gefen et al., 2014). This indicated that *E. coli* has principally the capacity for prolonged productions under growth arrest. However, regulatory responses to the growth arrest, which, for example, reduce substrate import rates, can explain low metabolic activities.

One of the responses to starvation in *E. coli* is the stringent response (Traxler et al., 2008; Brown et al., 2014), in which RelA and SpoT generate the signal metabolite ppGpp and trigger a global regulatory mechanism with broad implications for cell physiology. In a metabolic engineering study, the stringent response was engineered in *E. coli* to keep glucose uptake rates high under nitrogen limitation (Michalowski et al., 2017). To achieve this, the *relA* gene was deleted from the genome, two point mutations leading to Arg290Glu and Lys292Asp were introduced to *spoT*, and a point mutation leading to Gly267Cys was introduced to *aceE*. The resulting engineered strain showed a specific glucose uptake rate of $3.3 \text{ mmol g}_{DW}^{-1} \text{ h}^{-1}$ during nitrogen starvation, which is over 2-

times higher than for the comparable wild-type strain ($1.4 \text{ mmol g}_{DW}^{-1} \text{ h}^{-1}$). In other studies, mutants were identified that had also increased glucose uptake rates under nitrogen starvation (Chubukov et al., 2017; Sonderegger et al., 2005). However, the mechanisms by which the mutations achieved the increased uptake rate are not entirely clear.

In contrast to engineering substrate uptake rates directly, *enforced ATP-wasting* can function as an indirect driver for metabolism (Hädicke et al., 2015; Jensen et al., 2003; Semkiv et al., 2016; J. Liu et al., 2016; Holm et al., 2010; Koebmann et al., 2002; Chao and Liao, 1994, Chapter 6 and Chapter 7). The idea is to include a futile cycle that converts the main energy equivalent of microbial cells ATP into ADP, which stores less energy than ATP. Often the ATP synthesis can also be coupled directly to the product formation. Wasting ATP in *E. coli* increases glucose import and production rates even under growth arrest (Chapter 6 and Chapter 7). In the here presented work, we studied the *enforced ATP-wasting* concept further.

4 Temperature-sensitive proteins

Environmental factors including the temperature, salt concentrations, pH, and pressure affect the stability of a protein (Honig and Yang, 1995; Dill, 1990). Under certain conditions, a protein can be destabilized, undergoes a conformational change, and unfolds. This denaturation effect has practical applications and, for example, is needed in the preparation of paneer, an Indian cottage cheese (Khan and Pal, 2011): whey proteins in milk are denatured by heat and lowering the pH by citric acid or vinegar. Subsequently, denatured proteins coagulate and form the curd that is separated from the whey by filtration and, eventually, becomes the cheese.

The stability and the point, at which a protein unfolds, varies greatly among different proteins but also different species (Jarzab et al., 2020). For example, in thermophilic organisms, enzymes have a higher overall enzyme rigidity (Radestock and Gohlke, 2011) and increased numbers of β -sheets, higher lysine and arginine content, as well as lower threonine, glutamine, and aspartate content compared to enzymes in mesophiles (Leuenberger et al., 2017; Mallik and Kundu, 2013). Despite high levels of thermal adaption, thermozyms and proteins in mesophiles are very similar with sequence similarities in the range of 40 to 85 % (Vieille and Zeikus, 2001).

One of the main contributing factors to the stability of a protein is hydrophobic interactions between non-polar parts of the protein, buried within a protein's three-dimensional structure (Shriver, 2009). Other factors of protein stability are hydrogen bonding, van der Waals interactions, and electrostatic effects (Shriver, 2009; Goldenzweig and Fleishman, 2018; Shirley, 1995). To describe and quantify the stability of a protein with a single folded state and an unfolded state, we can use following definition of the unfolding free energy $\Delta G_{un\text{folded}}$ (Shriver, 2009):

$$\Delta G_{un\text{folded}} = -RT \cdot \ln\left(\frac{f_{un\text{folded}}}{f_{\text{folded}}}\right) \quad (1.3)$$

, whereas R is the ideal gas constant, T the temperature and $f_{\text{folded/unfolded}}$ the fractions of folded or unfolded protein. Usually, the unfolding free energy is between 5 and 20 kcal/mol (Dill, 1990; Goldenzweig and Fleishman, 2018; Varadarajan et al., 1996). This is comparably low considering that single hydrogen bonds in a protein are estimated to have ΔG values between 1 to 12 kcal/mol (Honig and Yang, 1995; Goldenzweig and Fleishman, 2018). At the so-called melting temperature T_m , half of the protein is unfolded, which makes the T_m a good parameter to compare the temperature-sensitivity of different proteins. The lower the T_m the higher the temperature-sensitivity.

The T_m of a protein can be determined by thermal shift assays that often rely on fluorescent dyes that can bind to denatured proteins (Huynh and Partch, 2015). With thermal proteome profiling (TPP) (Franken et al., 2015), which is a mass spectrometry-based proteomics method, it is possible to determine melting curves of proteins on a large scale. TPP also measures precipitation and not only thermal unfolding. Thus, the T_m values determined by TPP are not equal to the T_m values as defined only by thermal unfolding of proteins. However, investigations of 13 species including psychrophilic, mesophilic, and thermophilic organisms as well as bacteria, one archaeon, and eukaryotes by TPP revealed basic principles of thermal stability of proteins (Jarzab et al., 2020). The optimal growth temperature was distinctly lower than the median T_m values for eukaryotes, but, for bacteria, the T_m values and optimal growth temperatures could be much closer (Jarzab et al., 2020). In case of *E. coli*, the median T_m was 50 °C and, thus, 13 °C higher than the optimal growth temperature (Jarzab et al., 2020). 90 % of the T_m values determined for *E. coli* were within a 20 °C range (Jarzab et al., 2020). Five proteins in *E. coli* had T_m values below 40 °C (DnaC, AroH, YraR, HisG, CsrD) (Jarzab et al., 2020).

Mutations can have a great impact on the stability of a protein. Changes in the unfolding free energy ($\Delta\Delta G$) between a wild type and a mutant protein are calculated to quantify the effect of a mutation (Sanavia et al., 2020; Pancotti et al., 2022):

$$\Delta\Delta G_{unfolded} = \Delta G_{unfolded}^{mutant} - \Delta G_{unfolded}^{wildtype} \quad (1.4)$$

On average, a single amino acid substitution in 22 enzymes caused a destabilization with a $\Delta\Delta G$ value of 1.3 kcal/mol (Tokuriki and Tawfik, 2009). It is also estimated that around 80 % of the deleterious mutations that cause monogenic diseases in human are due to effects on protein stability (Tokuriki and Tawfik, 2009; Yue et al., 2005), and, on average, mutations are in 36% of the cases deleterious (Tokuriki and Tawfik, 2009; Smith and Raines, 2006; Bershtein et al., 2006; Camps et al., 2007). Since the design space of proteins is incomprehensibly large and many mutations are deleterious, the fitness landscape of a protein is thus characterized by island of high fitness (Smith and Raines, 2006; Romero and Arnold, 2009).

Knowing and engineering protein stability is crucial in the development of therapeutic agents (Sanavia et al., 2020; Leader et al., 2008) and biotechnological applications like biocatalysis (Walls and Loughran, 2011). Therefore, measured $\Delta\Delta G$ values are collected in a large database (Xavier et al., 2021) and many mathematical models have been developed that predict $\Delta G_{unfolded}$ and $\Delta\Delta G$ values (Sanavia et al., 2020; Pancotti et al., 2022; Guerois et al., 2002; Schymkowitz et al., 2005; Yang et al., 2019; Cao et al., 2019).

The $\Delta\Delta G$ values can lead design efforts to create proteins with, typically, an increased stability. Notably, increasing an enzyme's thermal stability does not always imply that the activity at low temperatures is reduced (Vieille and Zeikus, 2001). This effect can be explained by local high flexibility in the catalytic center but an overall increased rigidity (Radestock and Gohlke, 2011; Vieille and Zeikus, 2001). However, stability has an effect on enzyme activity, especially if the enzyme unfolds at high temperatures.

There are different mathematical models that approximate the maximum reaction rate v_{max} of an enzyme as a function of the temperature T (Daniel and Danson, 2013). One of the simpler models considers two main factors (eq. 1.5) (Fig. 1.8.a) (Daniel and Danson, 2013). First, the turnover number k_{cat} increases with increasing temperature as described by the Arrhenius equation (eq. 1.6). Second, the fraction of active enzyme decreases over time t and with increasing temperature (eq. 1.7), whereas thermal unfolding and other temperature-dependent inactivation effects are not distinguished and considered as a single effect.

$$v_{max} = k_{cat} \cdot [E_0] \cdot e^{-k_{inact} \cdot t} \quad (1.5)$$

$$k_{cat} = \frac{k_B \cdot T}{h} \cdot e^{-\frac{\Delta G_{cat}^\ddagger}{RT}} \quad (1.6)$$

$$k_{inact} = \frac{k_B \cdot T}{h} \cdot e^{-\frac{\Delta G_{inact}^\ddagger}{RT}} \quad (1.7)$$

In the model, $[E_0]$ is the initial concentration of the active enzyme, k_B the Boltzmann constant, h the Planck constant, R the ideal gas constant, ΔG_{cat}^\ddagger the activation energy of the reaction, and $\Delta G_{inact}^\ddagger$ the activation energy of the thermal inactivation. To calculate the reaction rates from the maximum reaction rate, other mathematical models like the Michaelis-Menten kinetic can be used (Fig. 1.2).

Mutations change the stability of a protein and thus can change the relationship between the maximum reaction rate of an enzyme and the temperature. Theoretically, it could be possible that mutations mainly influence the temperature, at which the v_{max} is maximal (Fig. 1.8.b), or mutations could change only the maximum of the v_{max} . However, realistically, we would expect a much complexer change in the temperature-dependency that affects both, the maximum of v_{max} and T_{max} , but also the slopes at which reaction rates increase or decrease. Among the different possible changes of a protein's thermal property due to mutations, there is however a special type of mutation with a distinct characteristic: In 1934, Timoféeff and Ressovsky were one of the first to report mu-

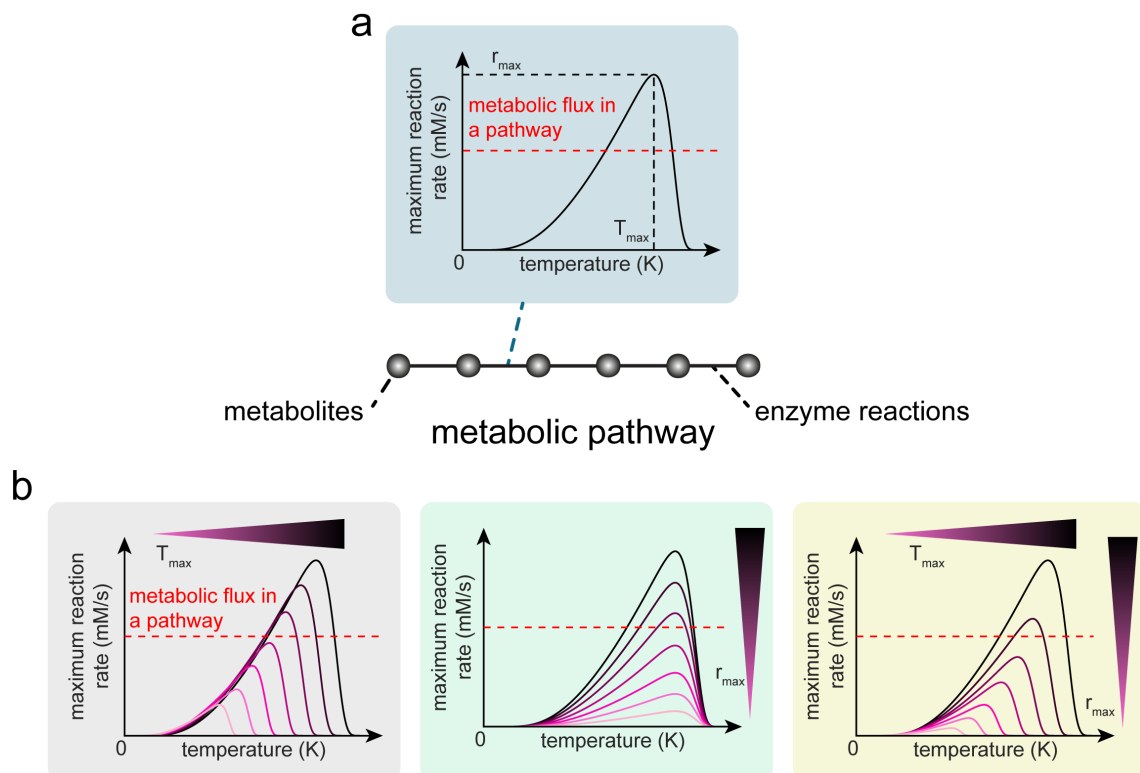


Figure 1.8: Maximum reaction rates as a function of temperature. (a) The chart shows the maximum reaction rate v_{max} of an enzyme reaction as a function of temperature at a single time point. The mathematical model that is used for the function considers increasing reaction rates according to Arrhenius and decreasing reaction rates due to thermal inactivation of the enzyme (Daniel and Danson, 2013). Here, we define T_{max} as the temperature, at which v_{max} is maximal, and r_{max} as the maximal v_{max} . The red line indicates the metabolic flux in a hypothetical metabolic pathway (b) Upon substituting amino acid residues in an enzyme, changes in the temperature dependency of the maximum reaction rate can be caused. The here displayed charts are theoretical, should only illustrate some of the possible outcomes, and thus are not exhaustive. A mutation could cause a change in the T_{max} (grey box) or in the r_{max} (green box). Most likely, a mutation causes a much more complex change in the temperature dependency of the maximum reaction rate and changes both, the T_{max} and the r_{max} (yellow box). Depending on the mutations, the maximum reaction rate can be lower than the metabolic flux in a pathway (red line). Then, a metabolic bottleneck is caused. Gain of function mutations are not considered here.

tants of *Drosophila* that had different viabilities as a function of temperature (Timoféeff-Ressovsky, 1934). Dobzhansky and Spassky reported similar effects in 1943 (Dobzhansky and Spassky, 1944). These observations describe what we now know as temperature-sensitive mutants. Temperature-sensitive proteins maintain function at a low temperature, which is called the permissive temperature. However, at higher (non-permissive) temperatures, the function of the protein is impaired causing a phenotype. In the context of a metabolic pathway, temperature-sensitive enzymes maintain the metabolic flux in a pathway at the permissive temperature but with increasing temperature they become a metabolic bottleneck (1.8.b). The great advantage of temperature-sensitive mutants is that they enable us to study the function of essential genes, which is usually not possible by gene knockouts.

Already in 1952, Maas and Davis reported an *E. coli* mutant with a temperature-sensitive enzyme in the pantothenate synthesis (Maas and Davis, 1952). To obtain the temperature-sensitive mutant, they used a *E. coli* mutant that was auxotrophic for pantothenate at any temperature. They plated a large number of cells of this mutant strain onto minimal agar medium and isolated secondary mutants at 25 °C. One of these mutant was then found to be temperature-sensitive.

In the following decades after discovery, temperature-sensitive mutants contributed to major findings in biology: temperature-sensitive *E. coli* mutants helped in understanding the cell cycle (Hansen and Atlung, 2018) and T4 phages with temperature-sensitive ligase and DNA polymerase contributed to understanding DNA replication and Okazaki fragments (Okazaki et al., 1968; Pauling and Hamm, 1969; Waard, 1965).

Finding temperature-sensitive mutants in DNA replication was possible by using 5-bromouracil that, if incorporated into the DNA, makes cells much more sensitive to UV light (Carl, 1970). Initially, *E. coli* cells were mutagenized by N-methyl-N-nitrosoguanidine. The pooled strain library was then cultivated in the presence of 5-bromouracil. Strains that were not temperature-sensitive in DNA replication incorporated the 5-bromouracil. Treating the cells by UV light led to an enrichment of temperature-sensitive mutants, which had not incorporated 5-bromouracil. After the enrichment, mutants still had to be isolated and tested for temperature-sensitivity.

Over the years, many temperature-sensitive alleles in *E. coli* were found. Currently, the *E. coli* Genetic Stock Centre at Yale University (CGSC, <https://cgsc.biology.yale.edu/>, accessed on 28.07.2022) collected 184 mutations in *E. coli* that are labelled as temperature-sensitive. The mutations include 117 unique genes. However, most strains with a temperature-sensitive mutation contain mutations in other genes, are not sequenced, or

have multiple mutations in the gene encoding for the temperature-sensitive protein. Many temperature-sensitive mutants without sequencing data were discovered before DNA sequencing was broadly available during a time, in which P1 transduction mapping (Bitner and Kuempel, 1982) or restriction endonuclease mapping (Yasuda and Hirota, 1977) was used to create maps of genomes.

In the years around 2010, large scale screenings identified hundreds of temperature-sensitive mutants in yeast (*Saccharomyces cerevisiae*) (Ben-Aroya et al., 2008; Z. Li et al., 2011; Kofoed et al., 2015). To create these strains, the *diploid shuffling* method was used (Kofoed et al., 2015; Ben-Aroya et al., 2010) that is not accessible in *E. coli*. In brief, a gene of interest is amplified and mutagenized by error-prone-PCR. The resulting fragments are cloned in between two selective markers into a plasmid. After digestion, linear DNA fragments of the mutant gene of interest with the flanking markers are then used to transform diploid yeast strains. These strains carry a wild type version of the gene of interest on one genome copy and a knockout of the gene on the other genome copy. Homologous regions to the linear DNA fragment are at the gene knockout site. Thus, the mutant gene of interest is inserted into the genome at the knockout site. The diploid strain is brought to sporulation. Using the second selective marker, only haploid strains with the mutant gene survive. By replica plating and cultivations at different temperatures, temperature-sensitive strains can be identified. Since the last steps of the protocol are performed on agar plates, this method can make full use of multiplexing and enables semi-automated high-throughput screenings. These large libraries of temperature-sensitive mutants were combined with gene knockout libraries and enabled the construction of gene-gene interaction networks (Costanzo et al., 2010; Costanzo et al., 2016).

Although many bioinformatic tools were developed to predict the impact of mutations on protein stability, only few were developed to predict mutations that induce temperature-sensitivity (Varadarajan et al., 1996; Tan et al., 2014; Poultney et al., 2011). These tools make use of buried sites of a protein that are amino acid sites within a protein with little access to solvent-accessible surface area (< than 10 % in Poultney et al., 2011). After identifying the buried sites, putative mutations at the sites are then further analysed and proposed as candidates depending on the algorithms.

In the context of metabolic engineering, using the temperature as signal to control cell growth and metabolism is very attractive. It allows the operator to directly control the bioprocess. No addition of chemicals is required, and the risk of contamination is reduced. Also, every bioreactor is already equipped for tight control of the cultivation

temperature. Thus, no additional control technology is required. Therefore, different systems were studied to control metabolism by temperature. The temperature-sensitive transcription factor CI857 of *Escherichia* virus Lambda was used to dynamically control the expression of the essential isocitrate dehydrogenase Icd in *E. coli* (Harder et al., 2018). This enabled the two-stage overproduction of itaconic acid with a final product concentration of 47 g/L (Harder et al., 2018). Other examples of temperature-sensitive transcription factors are RheA of *Streptomyces albus* (Weber, 2003; Servant et al., 2000) and temperature-sensitive mutants of TetR (Pearce et al., 2017).

Alternatively, temperature-sensitive enzymes can be used directly as metabolic valves (also see Section 3.1) to control metabolism and cell growth (M. D. Lynch et al., 2016; Cho et al., 2012; M. Lynch et al., 2019, Chapter 4 and Chapter 5). In the here presented work, we further studied temperature-sensitive proteins for the use in metabolic engineering (Chapter 4 and Chapter 5).

5 Objectives of the thesis

With current genetic tools, it is possible to generate vast quantities of mutant strains. In order to test large numbers of strain, high-throughput metabolomics methods like flow-injection mass spectrometry (FI-MS) are required. Since FI-MS, does not use chromatographic separation of metabolites prior to analysis, samples arrive simultaneously at the ion source. This could lead to in-source modifications and false-positive metabolite annotations. One goal of the presented work is to study in-source modifications during FI-MS. With **Chapter 3**, we present a network approach that allows us to identify in-source modifications and shows limitations but also opportunities of FI-MS.

During the overproduction of chemicals by engineered microbial strains, we face a trade-off between biomass and product formation that results in suboptimal product titers, rates, or yields. A solution to the trade-off is to decouple growth and overproduction and to establish two-stage bioprocesses. The metabolism of a microbial strain needs to be dynamically controlled to enable a transition between the two stages. To achieve the dynamic control of metabolism, overproduction of chemicals, and prolonged production under growth arrest, genetic engineering of microbial strains is required.

Goals of this work are to study temperature-sensitivity as a tool to control metabolism and to decouple growth and overproduction in two-stage bioprocesses. This includes the development of new approaches to generate and identify temperature-sensitive mutants in *E. coli*. With **Chapter 4** and **Chapter 5**, we present two different approaches. The goal of one of the approaches is to find many temperature-sensitive mutants of a single gene (*argG*) to establish a two-stage bioprocess overproducing citrulline. The other approach aims at finding temperature-sensitive mutants in many different essential genes in *E. coli*. Subsequently, the identified mutants should be tested for the dynamic control of metabolism. This includes overproducing chemicals in two-stage bioprocesses by selected strains.

Microbial strains often show reduced metabolic activity and substrate uptake rates during growth arrest. *Enforced ATP-wasting* is a promising concept to keep metabolic activity at a high level. With **Chapter 6** and **Chapter 7**, we provide studies on the concept that rely on overexpression of ATPase. Open questions were how elevated levels of ATPase influence the production rates under growth arrest but also how *E. coli* buffers perturbations in ATP metabolism.

6 Abbreviations

Table 1.1: Abbreviations.

ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
CE	Capillary electrophoresis
CID	Collision induced dissociation
CRISPR	Clustered regularly interspaced short palindromic repeats
$\Delta\Delta G$	Change of unfolding free energy (kcal/mol)
$\Delta G_{un\ folded}^{mutant/wildtype}$	Unfolding free energy in mutant or wildtype strains (kcal/mol)
$\Delta G_{cat}^{\ddagger}$	Activation energy of an enzyme reaction (kcal/mol)
$\Delta G_{inact}^{\ddagger}$	Activation energy of the thermal inactivation of an enzyme reaction (kcal/mol)
DNA	Deoxyribonucleic acid
DW	Dry weight (mg)
$[E]$	Concentration of active enzyme (mM)
$[E_0]$	Concentration of the active enzyme at time point zero (mM)
EI	Electron ionization
ESI	Electrospray ionization
FI-MS	Flow-injection mass spectrometry
FRT	Flipase recognition site
$f_{folded/unfolded}$	Fractions of folded or unfolded protein (-)
h	Planck's constant ($6.6260701 \cdot 10^{-34} \text{ J Hz}^{-1}$)
HPLC	High-performance liquid chromatography
k_B	Boltzmann constant ($1.380649 \cdot 10^{-23} \text{ J/K}$)
k_{cat}	Turnover number (1/s)
k_{inact}	Inactivation parameter
K_m	Michaelis-Menten-constant (mM)
LC	Liquid chromatography
LC-MS/MS	Liquid chromatography tandem mass spectrometry
MALDI	Matrix-assisted laser desorption ionization
MRM	Multiple reaction monitoring
mRNA	Messenger RNA
MS	Mass spectrometry
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NMR	Nuclear magnetic resonance
PAM	Protospacer adjacent motif
PCR	Polymerase chain reaction
QQQ	Triple quadruple mass spectrometer
QTOF	Tandem mass spectrometer using a TOF mass analyser
r_{max}	Maximal maximum reaction rate (mM/s)
R	Ideal gas constant ($8.31446261815324 \text{ J K}^{-1} \text{ mol}^{-1}$)
RNA	Ribonucleic acid
$[S]$	Substrate concentration (mM)
SFC	Supercritical fluid chromatography
sgRNA	single guide RNA
T	Temperature (K or °C)
TIMS	Trapped ion mobility spectrometry
TIMS-TOF	Trapped ion mobility spectrometry coupled to a TOF mass analyser
T_m	Melting temperature (K or °C)
T_{max}	Temperature, at which r_{max} is reached (K or °C)
TOF	Time-of-flight mass analyser
TRY criterion	A criterion to evaluate bioprocesses based on product titers, production rates, and product yields
UV	Ultraviolet
v	Reaction rate (mM/s)
v_{max}	Maximum reaction rate (mM/s)

References

- Yadav, V. G., Yadav, G. D., & Patankar, S. C. (2020). The production of fuels and chemicals in the new world: Critical analysis of the choice between crude oil and biomass vis-à-vis sustainability and the environment. *Clean Technologies and Environmental Policy*, 22(9), 1757–1774. <https://doi.org/10.1007/s10098-020-01945-5> (cited on page 10).
- Pörtner, H.-Ö., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegria, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., & Rama, B. (2022). IPCC, 2022: Climate change 2022: Impacts, adaptation, and vulnerability. contribution of working group II to the sixth assessment report of the intergovernmental panel on climate change. (Cited on page 10).
- Kotcher, J., Maibach, E., & Choi, W.-T. (2019). Fossil fuels are harming our brains: Identifying key messages about the health effects of air pollution from fossil fuels. *BMC Public Health*, 19(1), 1079. <https://doi.org/10.1186/s12889-019-7373-1> (cited on page 10).
- Tamang, J. P., Cotter, P. D., Endo, A., Han, N. S., Kort, R., Liu, S. Q., Mayo, B., Westerik, N., & Hutkins, R. (2020). Fermented foods in a global age: East meets west. *Comprehensive Reviews in Food Science and Food Safety*, 19(1), 184–217. <https://doi.org/https://doi.org/10.1111/1541-4337.12520> (cited on page 10).
- Samuel, D. (1996). Investigation of ancient egyptian baking and brewing methods by correlative microscopy. *Science*, 273(5274), 488–490. <https://doi.org/10.1126/science.273.5274.488> (cited on page 10).
- Boethius, A. (2016). Something rotten in scandinavia: The world’s earliest evidence of fermentation. *Journal of Archaeological Science*, 66, 169–180. <https://doi.org/10.1016/j.jas.2016.01.008> (cited on page 10).
- Fernbach, A., & Strange, E. H. (1912, November 12). *Fermentation process for the production of acetone and higher alcohols from starch, sugars, and other carbohydrate material* (U.S. pat. No. 1044368A). Retrieved November 17, 2020, from <https://patents.google.com/patent/US1044368A/en> (cited on page 10).
- Weizmann, C. (1919, September 9). *Production of acetone and alcohol by bacteriological processes* (U.S. pat. No. 1315585A). Retrieved November 17, 2020, from <https://patents.google.com/patent/US1315585A/en> (cited on page 10).
- Sauer, M. (2016). Industrial production of acetone and butanol by fermentation—100 years later (R. Boden, Ed.). *FEMS Microbiology Letters*, 363(13), fnw134. <https://doi.org/10.1093/femsle/fnw134> (cited on page 10).
- Bennett, S. J., & Pearson, P. J. (2009). From petrochemical complexes to biorefineries? the past and prospective co-evolution of liquid fuels and chemicals production in the UK. *Chemical Engineering Research and Design*, 87(9), 1120–1139. <https://doi.org/10.1016/j.cherd.2009.02.008> (cited on page 10).
- Lynch, M. D., Gill, R. T., & Lipscomb, T. E. W. (2016, August 30). *Method for producing 3-hydroxypropionic acid and other products* (U.S. pat. No. 9428778B2). Retrieved August 10, 2022, from <https://patents.google.com/patent/US9428778B2/en> (cited on pages 10, 42).
- Nielsen, J., & Keasling, J. D. (2016). Engineering cellular metabolism. *Cell*, 164(6), 1185–1197. <https://doi.org/10.1016/j.cell.2016.02.004> (cited on pages 10, 31).

- Sheehan, J., & Himmel, M. (1999). Enzymes, energy, and the environment: A strategic perspective on the u.s. department of energy's research and development activities for bioethanol. *Biotechnology Progress*, 15(5), 817–827. <https://doi.org/10.1021/bp990110d> (cited on page 11).
- Bardi, U. (2019). Peak oil, 20 years later: Failed prediction or useful insight? *Energy Research & Social Science*, 48, 257–261. <https://doi.org/10.1016/j.erss.2018.09.022> (cited on page 11).
- Bailey, J. E. (1991). Toward a science of metabolic engineering. *Science*, 252(5013), 1668–1675. <https://doi.org/10.1126/science.2047876> (cited on pages 12, 13).
- Stephanopoulos, G., & Vallino, J. J. (1991). Network rigidity and metabolic engineering in metabolite overproduction. *Science*, 252(5013), 1675–1681. <https://doi.org/10.1126/science.1904627> (cited on page 12).
- Keasling, J. D. (2010). Manufacturing molecules through metabolic engineering. *Science*, 330(6009), 1355–1358. <https://doi.org/10.1126/science.1193990> (cited on pages 12, 19).
- Woolston, B. M., Edgar, S., & Stephanopoulos, G. (2013). Metabolic engineering: Past and future. *Annual Review of Chemical and Biomolecular Engineering*, 4(1), 259–288. <https://doi.org/10.1146/annurev-chembioeng-061312-103312> (cited on page 12).
- Koffas, M., Roberge, C., Lee, K., & Stephanopoulos, G. (1999). Metabolic engineering. *Annual Review of Biomedical Engineering*, 1, 535–557. <https://doi.org/10.1146/annurev.bioeng.1.1.535> (cited on pages 12, 19–21).
- Becker, J., & Wittmann, C. (2015). Advanced biotechnology: Metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products. *Angewandte Chemie International Edition*, 54(11), 3328–3350. <https://doi.org/10.1002/anie.201409033> (cited on page 12).
- Liao, J. C., Mi, L., Pontrelli, S., & Luo, S. (2016). Fuelling the future: Microbial engineering for the production of sustainable biofuels. *Nature Reviews Microbiology*, 14(5), 288–304. <https://doi.org/10.1038/nrmicro.2016.32> (cited on page 12).
- Nielsen, J. (2001). Metabolic engineering. *Applied Microbiology and Biotechnology*, 55(3), 263–283. <https://doi.org/10.1007/s002530000511> (cited on page 12).
- Gurdo, N., Volke, D. C., & Nikel, P. I. (2022). Merging automation and fundamental discovery into the design–build–test–learn cycle of nontraditional microbes. *Trends in Biotechnology*, S0167779922000725. <https://doi.org/10.1016/j.tibtech.2022.03.004> (cited on pages 12, 13).
- Carbonell, P., Jervis, A. J., Robinson, C. J., Yan, C., Dunstan, M., Swainston, N., Vinaixa, M., Hollywood, K. A., Currin, A., Rattray, N. J. W., Taylor, S., Spiess, R., Sung, R., Williams, A. R., Fellows, D., Stanford, N. J., Mulherin, P., Le Feuvre, R., Barran, P., ... Scrutton, N. S. (2018). An automated design-build-test-learn pipeline for enhanced microbial production of fine chemicals. *Communications Biology*, 1(1), 66. <https://doi.org/10.1038/s42003-018-0076-9> (cited on pages 12, 13).
- Opgenorth, P., Costello, Z., Okada, T., Goyal, G., Chen, Y., Gin, J., Benites, V., de Raad, M., Northen, T. R., Deng, K., Deutsch, S., Baidoo, E. E. K., Petzold, C. J., Hillson, N. J., Garcia Martin, H., & Beller, H. R. (2019). Lessons from two design–build–test–learn cycles of dodecanol production in *Escherichia coli* aided by machine learning. *ACS Synthetic Biology*, 8(6), 1337–1351. <https://doi.org/10.1021/acssynbio.9b00020> (cited on pages 12, 13).
- Larson, M. H., Gilbert, L. A., Wang, X., Lim, W. A., Weissman, J. S., & Qi, L. S. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols*, 8(11), 2180–2196. <https://doi.org/10.1038/nprot.2013.132> (cited on page 12).

- Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, *152*(5), 1173–1183. <https://doi.org/10.1016/j.cell.2013.02.022> (cited on page 12).
- Garst, A. D., Bassalo, M. C., Pines, G., Lynch, S. A., Halweg-Edwards, A. L., Liu, R., Liang, L., Wang, Z., Zeitoun, R., Alexander, W. G., & Gill, R. T. (2017). Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nature Biotechnology*, *35*(1), 48–55. <https://doi.org/10.1038/nbt.3718> (cited on pages 12, 22–24).
- Na, D., Yoo, S. M., Chung, H., Park, H., Park, J. H., & Lee, S. Y. (2013). Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nature Biotechnology*, *31*(2), 170–174. <https://doi.org/10.1038/nbt.2461> (cited on page 12).
- Boecker, S., Zahoor, A., Schramm, T., Link, H., & Klamt, S. (2019). Broadening the scope of enforced ATP wasting as a tool for metabolic engineering in *Escherichia coli*. *Biotechnology Journal*, *14*(9), 1800438. <https://doi.org/10.1002/biot.201800438> (cited on pages 12, 33).
- Klamt, S., Mahadevan, R., & Hädicke, O. (2018). When do two-stage processes outperform one-stage processes? *Biotechnology Journal*, *13*(2), 1700539. <https://doi.org/10.1002/biot.201700539> (cited on pages 12, 31).
- Monk, J. M., Lloyd, C. J., Brunk, E., Mih, N., Sastry, A., King, Z., Takeuchi, R., Nomura, W., Zhang, Z., Mori, H., Feist, A. M., & Palsson, B. O. (2017). iML1515, a knowledgebase that computes *Escherichia coli* traits. *Nature Biotechnology*, *35*(10), 904–908. <https://doi.org/10.1038/nbt.3956> (cited on pages 12, 17, 18, 24).
- Norsigian, C. J., Kavvas, E., Seif, Y., Palsson, B. O., & Monk, J. M. (2018). iCN718, an updated and improved genome-scale metabolic network reconstruction of *Acinetobacter baumannii* AYE. *Frontiers in Genetics*, *9*, 121. <https://doi.org/10.3389/fgene.2018.00121> (cited on pages 12, 17).
- Xu, Y.-F., Lu, W., & Rabinowitz, J. D. (2015). Avoiding misannotation of in-source fragmentation products as cellular metabolites in liquid chromatography–mass spectrometry-based metabolomics. *Analytical Chemistry*, *87*(4), 2273–2281. <https://doi.org/10.1021/ac504118y> (cited on page 12).
- Nothias, L.-F., Petras, D., Schmid, R., Dührkop, K., Rainer, J., Sarvepalli, A., Protosyuk, I., Ernst, M., Tsugawa, H., Fleischauer, M., Aicheler, F., Aksenov, A. A., Alka, O., Allard, P.-M., Barsch, A., Cachet, X., Caraballo-Rodriguez, A. M., Da Silva, R. R., Dang, T., ... Dorrestein, P. C. (2020). Feature-based molecular networking in the GNPS analysis environment. *Nature Methods*, *17*(9), 905–908. <https://doi.org/10.1038/s41592-020-0933-6> (cited on page 12).
- Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M.-Y., Paulse, C., Creasy, D., ... Mallick, P. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, *30*(10), 918–920. <https://doi.org/10.1038/nbt.2377> (cited on page 12).
- Fuhrer, T., Heer, D., Begemann, B., & Zamboni, N. (2011). High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection–time-of-flight mass spectrometry. *Analytical Chemistry*, *83*(18), 7074–7080. <https://doi.org/10.1021/ac201267k> (cited on pages 12, 30).
- Guder, J. C., Schramm, T., Sander, T., & Link, H. (2017). Time-optimized isotope ratio LC–MS/MS for high-throughput quantification of primary metabolites. *Analytical Chemistry*, *89*(3), 1624–1631. <https://doi.org/10.1021/acs.analchem.6b03731> (cited on pages 12, 29, 30).

- Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., & Rabinowitz, J. D. (2009). Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nature Chemical Biology*, 5(8), 593–599. <https://doi.org/10.1038/nchembio.186> (cited on page 12).
- Chen, L., Lu, W., Wang, L., Xing, X., Chen, Z., Teng, X., Zeng, X., Muscarella, A. D., Shen, Y., Cowan, A., McReynolds, M. R., Kennedy, B. J., Lato, A. M., Campagna, S. R., Singh, M., & Rabinowitz, J. D. (2021). Metabolite discovery through global annotation of untargeted metabolomics data. *Nature Methods*, 18(11), 1377–1385. <https://doi.org/10.1038/s41592-021-01303-3> (cited on page 12).
- Guo, A. C., Jewison, T., Wilson, M., Liu, Y., Knox, C., Djoumbou, Y., Lo, P., Mandal, R., Krishnamurthy, R., & Wishart, D. S. (2012). ECMDDB: The *E. coli* metabolome database. *Nucleic Acids Research*, 41, D625–D630. <https://doi.org/10.1093/nar/gks992> (cited on page 12).
- Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., Nam, H., Feist, A. M., & Palsson, B. Ø. (2011). A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011. *Molecular Systems Biology*, 7, 535. <https://doi.org/10.1038/msb.2011.65> (cited on pages 12, 17).
- Michaelis, L., Menten, M. L., Johnson, K. A., & Goody, R. S. (2011). The original Michaelis constant: Translation of the 1913 Michaelis-Menten paper. *Biochemistry*, 50(39), 8264–8269. <https://doi.org/10.1021/bi201284u> (cited on pages 13, 15).
- Rogers, A., & Gibon, Y. (2009). Enzyme kinetics: Theory and practice. In J. Schwender (Ed.), *Plant metabolic networks* (pp. 71–103). Springer New York. https://doi.org/10.1007/978-0-387-78745-9_4 (cited on page 13).
- Donati, S., Kuntz, M., Pahl, V., Farke, N., Beuter, D., Glatter, T., Gomes-Filho, J. V., Randau, L., Wang, C.-Y., & Link, H. (2021). Multi-omics analysis of CRISPRi-knockdowns identifies mechanisms that buffer decreases of enzymes in *E. coli* metabolism. *Cell Systems*, 12(1), 56–67.e6. <https://doi.org/10.1016/j.cels.2020.10.011> (cited on pages 14, 33).
- Donati, S., Sander, T., & Link, H. (2018). Crosstalk between transcription and metabolism: How much enzyme is enough for a cell? *WIREs Systems Biology and Medicine*, 10(1). <https://doi.org/10.1002/wsbm.1396> (cited on pages 14, 16, 33).
- Sander, T., Farke, N., Diehl, C., Kuntz, M., Glatter, T., & Link, H. (2019). Allosteric feedback inhibition enables robust amino acid biosynthesis in *E. coli* by enforcing enzyme overabundance. *Cell Systems*, 8(1), 66–75.e8. <https://doi.org/10.1016/j.cels.2018.12.005> (cited on pages 14, 16, 17, 21).
- Matsuoka, Y., & Shimizu, K. (2011). Metabolic regulation in *Escherichia coli* in response to culture environments via global regulators. *Biotechnology Journal*, 6(11), 1330–1341. <https://doi.org/10.1002/biot.201000447> (cited on page 14).
- Alon, U. (2007). Network motifs: Theory and experimental approaches. *Nature Reviews Genetics*, 8(6), 450–461. <https://doi.org/10.1038/nrg2102> (cited on pages 14, 16).
- Metallo, C. M., & Vander Heiden, M. G. (2013). Understanding metabolic regulation and its influence on cell physiology. *Molecular Cell*, 49(3), 388–398. <https://doi.org/10.1016/j.molcel.2013.01.018> (cited on page 14).
- Radoš, D., Donati, S., Lempp, M., Rapp, J., & Link, H. (2022). Homeostasis of the biosynthetic *E. coli* metabolome. *iScience*, 25(7), 104503. <https://doi.org/10.1016/j.isci.2022.104503> (cited on pages 14, 18).
- Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrne, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebbersold, R., & Heinemann, M. (2016). The quantitative and condition-dependent-*Escherichia-*

- coli proteome. *Nature Biotechnology*, 34(1), 104–110. <https://doi.org/10.1038/nbt.3418> (cited on page 14).
- Schramm, T., & Link, H. (2021). Von der Stöchiometrie zur Kontrolle metabolischer Netzwerke. *BIOspektrum*, 27(1), 34–36. <https://doi.org/10.1007/s12268-021-1538-0> (cited on page 15).
- Goodall, E. C. A., Robinson, A., Johnston, I. G., Jabbari, S., Turner, K. A., Cunningham, A. F., Lund, P. A., Cole, J. A., & Henderson, I. R. (2018). The essential genome of escherichia coli k-12. *mBio*, 9(1), e02096–17. <https://doi.org/10.1128/mBio.02096-17> (cited on page 14).
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of escherichia coli k-12 in-frame, single-gene knockout mutants: The keio collection. *Molecular Systems Biology*, 2, 2006.0008. <https://doi.org/10.1038/msb4100050> (cited on pages 14, 22).
- Schellhorn, H. E. (2020). Function, evolution, and composition of the RpoS regulon in escherichia coli. *Frontiers in Microbiology*, 11, 560099. <https://doi.org/10.3389/fmicb.2020.560099> (cited on page 14).
- Weber, W. (2003). Conditional human VEGF-mediated vascularization in chicken embryos using a novel temperature-inducible gene regulation (TIGR) system. *Nucleic Acids Research*, 31(12), 69e–69. <https://doi.org/10.1093/nar/gng069> (cited on pages 14, 42).
- Barda, O., Maor, U., Sadhasivam, S., Bi, Y., Zakin, V., Prusky, D., & Sionov, E. (2020). The pH-responsive transcription factor PacC governs pathogenicity and ochratoxin a biosynthesis in aspergillus carbonarius. *Frontiers in Microbiology*, 11. Retrieved July 27, 2022, from <https://www.frontiersin.org/articles/10.3389/fmicb.2020.00210> (cited on page 14).
- Diray-Arce, J., Knowles, A., Suvorov, A., O'Brien, J., Hansen, C., Bybee, S. M., Gul, B., Khan, M. A., & Nielsen, B. L. (2019). Identification and evolutionary characterization of salt-responsive transcription factors in the succulent halophyte suaeda fruticosa. *PLoS ONE*, 14(9), e0222940. <https://doi.org/10.1371/journal.pone.0222940> (cited on page 14).
- Zhao, E. M., Zhang, Y., Mehl, J., Park, H., Lalwani, M. A., Toettcher, J. E., & Avalos, J. L. (2018). Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature*, 555(7698), 683–687. <https://doi.org/10.1038/nature26141> (cited on pages 14, 31).
- Kim, D., Seo, S. W., Gao, Y., Nam, H., Guzman, G. I., Cho, B.-K., & Palsson, B. O. (2018). Systems assessment of transcriptional regulation on central carbon metabolism by cra and CRP. *Nucleic Acids Research*, 46(6), 2901–2917. <https://doi.org/10.1093/nar/gky069> (cited on page 14).
- Perrenoud, A., & Sauer, U. (2005). Impact of global transcriptional regulation by ArcA, ArcB, cra, crp, cya, fnr, and mlc on glucose catabolism in escherichia coli. *Journal of Bacteriology*, 187(9), 3171–3179. <https://doi.org/10.1128/JB.187.9.3171-3179.2005> (cited on page 16).
- Ament, S. A., Pearl, J. R., Cantle, J. P., Bragg, R. M., Skene, P. J., Coffey, S. R., Bergey, D. E., Wheeler, V. C., MacDonald, M. E., Baliga, N. S., Rosinski, J., Hood, L. E., Carroll, J. B., & Price, N. D. (2018). Transcriptional regulatory networks underlying gene expression changes in huntington's disease. *Molecular Systems Biology*, 14(3). <https://doi.org/10.15252/msb.20167435> (cited on page 16).
- Lempp, M., Farke, N., Kuntz, M., Freibert, S. A., Lill, R., & Link, H. (2019). Systematic identification of metabolites controlling gene expression in e. coli. *Nature Communications*, 10(1), 4463. <https://doi.org/10.1038/s41467-019-12474-1> (cited on page 16).

- Grüning, N.-M., Lehrach, H., & Ralser, M. (2010). Regulatory crosstalk of the metabolic network. *Trends in Biochemical Sciences*, 35(4), 220–227. <https://doi.org/10.1016/j.tibs.2009.12.001> (cited on page 16).
- Gollnick, P., & Babitzke, P. (2002). Transcription attenuation. *Biochimica Et Biophysica Acta*, 1577(2), 240–250. [https://doi.org/10.1016/s0167-4781\(02\)00455-4](https://doi.org/10.1016/s0167-4781(02)00455-4) (cited on page 16).
- Russell, J. B., & Cook, G. M. (1995). Energetics of bacterial growth: Balance of anabolic and catabolic reactions. *Microbiological Reviews*, 59(1), 48–62. <https://doi.org/10.1128/mr.59.1.48-62.1995> (cited on page 17).
- Peregrín-Alvarez, J. M., Sanford, C., & Parkinson, J. (2009). The conservation and evolutionary modularity of metabolism. *Genome Biology*, 10(6), R63. <https://doi.org/10.1186/gb-2009-10-6-r63> (cited on page 17).
- Merino, N., Aronson, H. S., Bojanova, D. P., Feyhl-Buska, J., Wong, M. L., Zhang, S., & Giovannelli, D. (2019). Living at the extremes: Extremophiles and the limits of life in a planetary context. *Frontiers in Microbiology*, 10, 780. <https://doi.org/10.3389/fmicb.2019.00780> (cited on page 17).
- Nordstrom, D. K., Alpers, C. N., Ptacek, C. J., & Blowes, D. W. (2000). Negative pH and extremely acidic mine waters from iron mountain, california. *Environmental Science & Technology*, 34(2), 254–258. <https://doi.org/10.1021/es990646v> (cited on page 17).
- Czop, M., Motyka, J., Sracek, O., & Szuwarzyński, M. (2011). Geochemistry of the hyperalkaline gorka pit lake (pH > 13) in the chrzanow region, southern poland. *Water, Air, & Soil Pollution*, 214(1), 423–434. <https://doi.org/10.1007/s11270-010-0433-x> (cited on page 17).
- Rivkina, E. M., Friedmann, E. I., McKay, C. P., & Gilichinsky, D. A. (2000). Metabolic activity of permafrost bacteria below the freezing point. *Applied and Environmental Microbiology*, 66(8), 3230–3233. <https://doi.org/10.1128/AEM.66.8.3230-3233.2000> (cited on page 17).
- Takai, K., Nakamura, K., Toki, T., Tsunogai, U., Miyazaki, M., Miyazaki, J., Hirayama, H., Nakagawa, S., Nunoura, T., & Horikoshi, K. (2008). Cell proliferation at 122°C and isotopically heavy CH₄ production by a hyperthermophilic methanogen under high-pressure cultivation. *Proceedings of the National Academy of Sciences*, 105(31), 10949–10954. <https://doi.org/10.1073/pnas.0712334105> (cited on page 17).
- Durot, M., Bourguignon, P.-Y., & Schachter, V. (2009). Genome-scale models of bacterial metabolism: Reconstruction and applications. *FEMS Microbiology Reviews*, 33(1), 164–190. <https://doi.org/10.1111/j.1574-6976.2008.00146.x> (cited on page 17).
- Haggart, C. R., Bartell, J. A., Saucerman, J. J., & Papin, J. A. (2011). Whole-genome metabolic network reconstruction and constraint-based modeling. In *Methods in enzymology* (pp. 411–433, Vol. 500). Elsevier. <https://doi.org/10.1016/B978-0-12-385118-5.00021-9> (cited on page 17).
- Lopes, H., & Rocha, I. (2017). Genome-scale modeling of yeast: Chronology, applications and critical perspectives. *FEMS Yeast Research*, 17(5). <https://doi.org/10.1093/femsyr/fox050> (cited on page 17).
- Brunk, E., Sahoo, S., Zielinski, D. C., Altunkaya, A., Dräger, A., Mih, N., Gatto, F., Nilsson, A., Preciat Gonzalez, G. A., Aurich, M. K., Prlić, A., Sastry, A., Danielsdottir, A. D., Heinken, A., Noronha, A., Rose, P. W., Burley, S. K., Fleming, R. M. T., Nielsen, J., ... Pálsson, B. O. (2018). Recon3d enables a three-dimensional view of gene variation in human metabolism. *Nature Biotechnology*, 36(3), 272–281. <https://doi.org/10.1038/nbt.4072> (cited on page 17).

- Masid, M., Ataman, M., & Hatzimanikatis, V. (2020). Analysis of human metabolism by reducing the complexity of the genome-scale models using redHUMAN. *Nature Communications*, *11*(1), 2821. <https://doi.org/10.1038/s41467-020-16549-2> (cited on page 17).
- Wang, H., Robinson, J. L., Kocabas, P., Gustafsson, J., Anton, M., Cholley, P.-E., Huang, S., Gobom, J., Svensson, T., Uhlen, M., Zetterberg, H., & Nielsen, J. (2021). Genome-scale metabolic network reconstruction of model animals as a platform for translational research. *Proceedings of the National Academy of Sciences*, *118*(30), e2102344118. <https://doi.org/10.1073/pnas.2102344118> (cited on page 17).
- Seif, Y., Kavvas, E., Lachance, J.-C., Yurkovich, J. T., Nuccio, S.-P., Fang, X., Catoiu, E., Raffatellu, M., Palsson, B. O., & Monk, J. M. (2018). Genome-scale metabolic reconstructions of multiple salmonella strains reveal serovar-specific metabolic traits. *Nature Communications*, *9*(1), 3771. <https://doi.org/10.1038/s41467-018-06112-5> (cited on pages 17, 18).
- Norsigian, C. J., Fang, X., Seif, Y., Monk, J. M., & Palsson, B. O. (2020). A workflow for generating multi-strain genome-scale metabolic models of prokaryotes. *Nature Protocols*, *15*(1), 1–14. <https://doi.org/10.1038/s41596-019-0254-3> (cited on page 18).
- Monk, J. M., Charusanti, P., Aziz, R. K., Lerman, J. A., Premyodhin, N., Orth, J. D., Feist, A. M., & Palsson, B. Ø. (2013). Genome-scale metabolic reconstructions of multiple escherichia coli strains highlight strain-specific adaptations to nutritional environments. *Proceedings of the National Academy of Sciences*, *110*(50), 20338–20343. <https://doi.org/10.1073/pnas.1307797110> (cited on page 18).
- Curran, K. A., & Alper, H. S. (2012). Expanding the chemical palate of cells by combining systems biology and metabolic engineering. *Metabolic Engineering*, *14*(4), 289–297. <https://doi.org/10.1016/j.ymben.2012.04.006> (cited on page 18).
- Schilling, C. H., Edwards, J. S., & Palsson, B. O. (1999). Toward metabolic phenomics: Analysis of genomic data using flux balances. *Biotechnology Progress*, *15*(3), 288–295. <https://doi.org/10.1021/bp9900357> (cited on page 18).
- Edwards, J. S., & Palsson, B. O. (2000). Metabolic flux balance analysis and the in silico analysis of escherichia coli k-12 gene deletions. *BMC bioinformatics*, *1*, 1. <https://doi.org/10.1186/1471-2105-1-1> (cited on page 18).
- Edwards, J. S., Covert, M., & Palsson, B. (2002). Metabolic modelling of microbes: The flux-balance approach. *Environmental Microbiology*, *4*(3), 133–140. <https://doi.org/10.1046/j.1462-2920.2002.00282.x> (cited on page 18).
- Antoniewicz, M. R. (2015). Methods and advances in metabolic flux analysis: A mini-review. *Journal of Industrial Microbiology and Biotechnology*, *42*(3), 317–325. <https://doi.org/10.1007/s10295-015-1585-x> (cited on page 18).
- Zamboni, N., Fendt, S.-M., Rühl, M., & Sauer, U. (2009). 13c-based metabolic flux analysis. *Nature Protocols*, *4*(6), 878–892. <https://doi.org/10.1038/nprot.2009.58> (cited on page 18).
- Wiechert, W. (2001). 13c metabolic flux analysis. *Metabolic Engineering*, *3*(3), 195–206. <https://doi.org/10.1006/mben.2001.0187> (cited on page 18).
- Sander, T., Wang, C. Y., Glatter, T., & Link, H. (2019). CRISPRi-based downregulation of transcriptional feedback improves growth and metabolism of arginine overproducing e. coli. *ACS Synthetic Biology*, *8*(9), 1983–1990. <https://doi.org/10.1021/acssynbio.9b00183> (cited on pages 18, 19).

- Lee, J.-H., & Wendisch, V. F. (2017). Production of amino acids – genetic and metabolic engineering approaches. *Bioresource Technology*, *245*, 1575–1587. <https://doi.org/10.1016/j.biortech.2017.05.065> (cited on page 18).
- Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M., & Smolke, C. D. (2015). Complete biosynthesis of opioids in yeast. *Science*, *349*(6252), 1095–1100. <https://doi.org/10.1126/science.aac9373> (cited on pages 19, 21).
- Luo, X., Reiter, M. A., d'Espaux, L., Wong, J., Denby, C. M., Lechner, A., Zhang, Y., Grzybowski, A. T., Harth, S., Lin, W., Lee, H., Yu, C., Shin, J., Deng, K., Benites, V. T., Wang, G., Baidoo, E. E. K., Chen, Y., Dev, I., ... Keasling, J. D. (2019). Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature*, *567*(7746), 123–126. <https://doi.org/10.1038/s41586-019-0978-9> (cited on pages 19, 21).
- Wang, C.-Y., Lempp, M., Farke, N., Donati, S., Glatter, T., & Link, H. (2021). Metabolome and proteome analyses reveal transcriptional misregulation in glycolysis of engineered e. coli. *Nature Communications*, *12*(1), 4929. <https://doi.org/10.1038/s41467-021-25142-0> (cited on pages 19, 21).
- Gleizer, S., Ben-Nissan, R., Bar-On, Y. M., Antonovsky, N., Noor, E., Zohar, Y., Jona, G., Krieger, E., Shamshoum, M., Bar-Even, A., & Milo, R. (2019). Conversion of escherichia coli to generate all biomass carbon from CO₂. *Cell*, *179*(6), 1255–1263.e12. <https://doi.org/10.1016/j.cell.2019.11.009> (cited on pages 19, 21).
- Ma, Q., Zhang, Q., Xu, Q., Zhang, C., Li, Y., Fan, X., Xie, X., & Chen, N. (2017). Systems metabolic engineering strategies for the production of amino acids. *Synthetic and Systems Biotechnology*, *2*(2), 87–96. <https://doi.org/10.1016/j.synbio.2017.07.003> (cited on pages 19–21).
- Zampieri, M., Hörl, M., Hotz, F., Müller, N. F., & Sauer, U. (2019). Regulatory mechanisms underlying coordination of amino acid and glucose catabolism in escherichia coli. *Nature Communications*, *10*(1), 3354. <https://doi.org/10.1038/s41467-019-11331-5> (cited on page 19).
- Chemler, J. A., Fowler, Z. L., McHugh, K. P., & Koffas, M. A. (2010). Improving NADPH availability for natural product biosynthesis in escherichia coli by metabolic engineering. *Metabolic Engineering*, *12*(2), 96–104. <https://doi.org/10.1016/j.ymben.2009.07.003> (cited on pages 20, 21).
- Bommareddy, R. R., Chen, Z., Rappert, S., & Zeng, A.-P. (2014). A de novo NADPH generation pathway for improving lysine production of corynebacterium glutamicum by rational design of the coenzyme specificity of glyceraldehyde 3-phosphate dehydrogenase. *Metabolic Engineering*, *25*, 30–37. <https://doi.org/10.1016/j.ymben.2014.06.005> (cited on pages 20, 21).
- Ferreira, S., Pereira, R., Wahl, S. A., & Rocha, I. (2020). Metabolic engineering strategies for butanol production in *Escherichia coli*. *Biotechnology and Bioengineering*, *117*(8), 2571–2587. <https://doi.org/10.1002/bit.27377> (cited on page 20).
- Zhang, Y., Muyrers, J. P., Testa, G., & Stewart, A. F. (2000). DNA cloning by homologous recombination in escherichia coli. *Nature Biotechnology*, *18*(12), 1314–1317. <https://doi.org/10.1038/82449> (cited on page 20).
- Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in escherichia coli k-12 using PCR products. *Proceedings of the National Academy of Sciences*, *97*(12), 6640–6645. <https://doi.org/10.1073/pnas.120163297> (cited on pages 20, 22, 23).

- Thomason, L. C., Costantino, N., & Court, D. L. (2007). E. coli genome manipulation by p1 transduction. *Current Protocols in Molecular Biology*, 79(1), 1.17.1–1.17.8. <https://doi.org/10.1002/0471142727.mb0117s79> (cited on page 22).
- Lander, E. S. (2016). The heroes of CRISPR. *Cell*, 164(1), 18–28. <https://doi.org/10.1016/j.cell.2015.12.041> (cited on page 22).
- Ledford, H. (2016). The unsung heroes of CRISPR. *Nature*, 535(7612), 342–344. <https://doi.org/10.1038/535342a> (cited on page 22).
- Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K., & Doudna, J. A. (2010). Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science*, 329(5997), 1355–1358. <https://doi.org/10.1126/science.1192272> (cited on page 22).
- Kantor, A., McClements, M., & MacLaren, R. (2020). CRISPR-cas9 DNA base-editing and prime-editing. *International Journal of Molecular Sciences*, 21(17), 6240. <https://doi.org/10.3390/ijms21176240> (cited on page 22).
- Bassalo, M. C., Garst, A. D., Choudhury, A., Grau, W. C., Oh, E. J., Spindler, E., Lipscomb, T., & Gill, R. T. (2018). Deep scanning lysine metabolism in *Escherichia coli*. *Molecular Systems Biology*, 14(11). <https://doi.org/10.15252/msb.20188371> (cited on page 24).
- Liu, R., Liang, L., Freed, E. F., Choudhury, A., Eckert, C. A., & Gill, R. T. (2020). Engineering regulatory networks for complex phenotypes in e. coli. *Nature Communications*, 11(1), 4050. <https://doi.org/10.1038/s41467-020-17721-4> (cited on page 24).
- Chen, C., Choudhury, A., Zhang, S., Garst, A. D., Song, X., Liu, X., Chen, T., Gill, R. T., & Wang, Z. (2020). Integrating CRISPR-enabled trackable genome engineering and transcriptomic analysis of global regulators for antibiotic resistance selection and identification in escherichia coli (Z. Bulman, Ed.). *mSystems*, 5(2), e00232–20. <https://doi.org/10.1128/mSystems.00232-20> (cited on page 24).
- Liu, R., Liang, L., Choudhury, A., Bassalo, M. C., Garst, A. D., Tarasava, K., & Gill, R. T. (2018). Iterative genome editing of escherichia coli for 3-hydroxypropionic acid production. *Metabolic Engineering*, 47, 303–313. <https://doi.org/10.1016/j.ymben.2018.04.007> (cited on pages 24, 29).
- Liu, R., Liang, L., Garst, A. D., Choudhury, A., Nogué, V. S. i., Beckham, G. T., & Gill, R. T. (2018). Directed combinatorial mutagenesis of escherichia coli for complex phenotype engineering. *Metabolic Engineering*, 47, 10–20. <https://doi.org/10.1016/j.ymben.2018.02.007> (cited on pages 24, 29).
- Markley, J. L., Brüschweiler, R., Edison, A. S., Eghbalnia, H. R., Powers, R., Raftery, D., & Wishart, D. S. (2017). The future of NMR-based metabolomics. *Current Opinion in Biotechnology*, 43, 34–40. <https://doi.org/10.1016/j.copbio.2016.08.001> (cited on pages 24, 25).
- Koulman, A., Lane, G. A., Harrison, S. J., & Volmer, D. A. (2009). From differentiating metabolites to biomarkers. *Analytical and Bioanalytical Chemistry*, 394(3), 663–670. <https://doi.org/10.1007/s00216-009-2690-3> (cited on page 24).
- Keen, B., Cawley, A., Reedy, B., & Fu, S. (2022). Metabolomics in clinical and forensic toxicology, sports anti-doping and veterinary residues. *Drug Testing and Analysis*, 14(5), 794–807. <https://doi.org/10.1002/dta.3245> (cited on page 24).
- Adebo, O. A., Njobeh, P. B., Adebisi, J. A., Gbashi, S., & Kayitesi, E. (2017, August 2). Food metabolomics: A new frontier in food analysis and its application to understanding fermented foods. In M. C. Hueda (Ed.), *Functional food - improve health through adequate food*. InTech. <https://doi.org/10.5772/intechopen.69171> (cited on page 24).

- Bundy, J. G., Davey, M. P., & Viant, M. R. (2009). Environmental metabolomics: A critical review and future perspectives. *Metabolomics*, 5(1), 3–21. <https://doi.org/10.1007/s11306-008-0152-0> (cited on page 24).
- Alseekh, S., & Fernie, A. R. (2018). Metabolomics 20 years on: What have we learned and what hurdles remain? *The Plant Journal*, 94(6), 933–942. <https://doi.org/10.1111/tpj.13950> (cited on page 24).
- Wishart, D. S., Guo, A., Oler, E., Wang, F., Anjum, A., Peters, H., Dizon, R., Sayeeda, Z., Tian, S., Lee, B. L., Berjanskii, M., Mah, R., Yamamoto, M., Jovel, J., Torres-Calzada, C., Hiebert-Giesbrecht, M., Lui, V. W., Varshavi, D., Varshavi, D., ... Gautam, V. (2022). HMDB 5.0: The human metabolome database for 2022. *Nucleic Acids Research*, 50, D622–D631. <https://doi.org/10.1093/nar/gkab1062> (cited on page 24).
- Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., Nguyen, D. D., Watrous, J., Kapon, C. A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V., Meehan, M. J., Liu, W.-T., Crüsemann, M., Boudreau, P. D., Esquenazi, E., Sandoval-Calderón, M., ... Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nature Biotechnology*, 34(8), 828–837. <https://doi.org/10.1038/nbt.3597> (cited on page 24).
- Guijas, C., Montenegro-Burke, J. R., Domingo-Almenara, X., Palermo, A., Warth, B., Hermann, G., Koelensperger, G., Huan, T., Uritboonthai, W., Aisporna, A. E., Wolan, D. W., Spilker, M. E., Benton, H. P., & Siuzdak, G. (2018). METLIN: A technology platform for identifying knowns and unknowns. *Analytical Chemistry*, 90(5), 3156–3164. <https://doi.org/10.1021/acs.analchem.7b04424> (cited on page 24).
- Karamanou, M., Protogerou, A., Tsoucalas, G., Androutsos, G., & Poulakou-Rebelakou, E. (2016). Milestones in the history of diabetes mellitus: The main contributors. *World Journal of Diabetes*, 7(1), 1–7. <https://doi.org/10.4239/wjd.v7.i1.1> (cited on page 25).
- Raftery, D. (Ed.). (2014). *Mass spectrometry in metabolomics* (Vol. 1198). Springer New York. <https://doi.org/10.1007/978-1-4939-1258-2> (cited on page 25).
- Emwas, A.-H., Roy, R., McKay, R. T., Tenori, L., Saccenti, E., Gowda, G. A. N., Raftery, D., Alahmari, F., Jaremko, L., Jaremko, M., & Wishart, D. S. (2019). NMR spectroscopy for metabolomics research. *Metabolites*, 9(7), 123. <https://doi.org/10.3390/metabo9070123> (cited on page 25).
- Emwas, A.-H. M. (2015). The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. In J. T. Bjerrum (Ed.), *Metabonomics* (pp. 161–193, Vol. 1277). Springer New York. https://doi.org/10.1007/978-1-4939-2377-9_13 (cited on page 25).
- Griffiths, J. (2008). A brief history of mass spectrometry. *Analytical Chemistry*, 80(15), 5678–5683. <https://doi.org/10.1021/ac8013065> (cited on page 25).
- Yergey, A. L., & Yergey, A. K. (1997). Preparative scale mass spectrometry: A brief history of the calutron. *Journal of the American Society for Mass Spectrometry*, 8(9), 943–953. [https://doi.org/10.1016/S1044-0305\(97\)00123-2](https://doi.org/10.1016/S1044-0305(97)00123-2) (cited on page 25).
- Bjerrum, J. T. (Ed.). (2015). *Metabonomics* (Vol. 1277). Springer New York. <https://doi.org/10.1007/978-1-4939-2377-9> (cited on pages 26, 27).

- Famiglioni, G., Palma, P., Termopoli, V., & Cappiello, A. (2021). The history of electron ionization in LC-MS, from the early days to modern technologies: A review. *Analytica Chimica Acta*, 1167, 338350. <https://doi.org/10.1016/j.aca.2021.338350> (cited on pages 26, 27).
- Jurinke, C., Oeth, P., & van den Boom, D. (2004). MALDI-TOF mass spectrometry : A versatile tool for high-performance DNA analysis. *Molecular Biotechnology*, 26(2), 147–164. <https://doi.org/10.1385/MB:26:2:147> (cited on pages 26, 27).
- Wilm, M. (2011). Principles of electrospray ionization. *Molecular & Cellular Proteomics*, 10(7). <https://doi.org/10.1074/mcp.M111.009407> (cited on pages 26, 27).
- Martin Somer, A., Macaluso, V., Barnes, G. L., Yang, L., Pratihar, S., Song, K., Hase, W. L., & Spezia, R. (2020). Role of chemical dynamics simulations in mass spectrometry studies of collision-induced dissociation and collisions of biological ions with organic surfaces. *Journal of the American Society for Mass Spectrometry*, 31(1), 2–24. <https://doi.org/10.1021/jasms.9b00062> (cited on pages 26, 28).
- Boesl, U. (2017). Time-of-flight mass spectrometry: Introduction to the basics. *Mass Spectrometry Reviews*, 36(1), 86–109. <https://doi.org/10.1002/mas.21520> (cited on pages 26, 28).
- Glish, G. L., & Vachet, R. W. (2003). The basics of mass spectrometry in the twenty-first century. *Nature Reviews Drug Discovery*, 2(2), 140–150. <https://doi.org/10.1038/nrd1011> (cited on pages 26–28).
- Zubarev, R. A., & Makarov, A. (2013). Orbitrap mass spectrometry. *Analytical Chemistry*, 85(11), 5288–5296. <https://doi.org/10.1021/ac4001223> (cited on pages 26, 28).
- Alseekh, S., Aharoni, A., Brotman, Y., Contrepois, K., D’Auria, J., Ewald, J., C. Ewald, J., Fraser, P. D., Giavalisco, P., Hall, R. D., Heinemann, M., Link, H., Luo, J., Neumann, S., Nielsen, J., Perez de Souza, L., Saito, K., Sauer, U., Schroeder, F. C., ... Fernie, A. R. (2021). Mass spectrometry-based metabolomics: A guide for annotation, quantification and best reporting practices. *Nature Methods*, 18(7), 747–756. <https://doi.org/10.1038/s41592-021-01197-1> (cited on pages 25, 27).
- van de Velde, B., Guillarme, D., & Kohler, I. (2020). Supercritical fluid chromatography – mass spectrometry in metabolomics: Past, present, and future perspectives. *Journal of Chromatography B*, 1161, 122444. <https://doi.org/10.1016/j.jchromb.2020.122444> (cited on page 25).
- Zhang, W., & Ramautar, R. (2021). CE-MS for metabolomics: Developments and applications in the period 2018–2020. *ELECTROPHORESIS*, 42(4), 381–401. <https://doi.org/10.1002/elps.202000203> (cited on page 25).
- Gowda, G. A. N., & Djukovic, D. (2014). Overview of mass spectrometry-based metabolomics: Opportunities and challenges. In D. Raftery (Ed.), *Mass spectrometry in metabolomics* (pp. 3–12, Vol. 1198). Springer New York. https://doi.org/10.1007/978-1-4939-1258-2_1 (cited on page 27).
- Perez de Souza, L., Alseekh, S., Scossa, F., & Fernie, A. R. (2021). Ultra-high-performance liquid chromatography high-resolution mass spectrometry variants for metabolomics research. *Nature Methods*, 18(7), 733–746. <https://doi.org/10.1038/s41592-021-01116-4> (cited on page 27).
- Steckel, A., & Schlosser, G. (2019). An organic chemist’s guide to electrospray mass spectrometric structure elucidation. *Molecules*, 24(3), 611. <https://doi.org/10.3390/molecules24030611> (cited on page 27).
- Zhao, C., & Cai, Z. (2022). Three-dimensional quantitative mass spectrometry imaging in complex system: From subcellular to whole organism. *Mass Spectrometry Reviews*, 41(3), 469–487. <https://doi.org/10.1002/mas.21674> (cited on page 27).

- Good, C. J., Neumann, E. K., Butrico, C. E., Cassat, J. E., Caprioli, R. M., & Spraggins, J. M. (2022). High spatial resolution MALDI imaging mass spectrometry of fresh-frozen bone. *Analytical Chemistry*, *94*(7), 3165–3172. <https://doi.org/10.1021/acs.analchem.1c04604> (cited on page 27).
- Schnackenberg, L. K., Thorn, D. A., Barnette, D., & Jones, E. E. (2022). MALDI imaging mass spectrometry: An emerging tool in neurology. *Metabolic Brain Disease*, *37*(1), 105–121. <https://doi.org/10.1007/s11011-021-00797-2> (cited on page 27).
- Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., & Whitehouse, C. M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science*, *246*(4926), 64–71. <https://doi.org/10.1126/science.2675315> (cited on page 27).
- Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., & Whitehouse, C. M. (1990). Electrospray ionization—principles and practice. *Mass Spectrometry Reviews*, *9*(1), 37–70. <https://doi.org/10.1002/mas.1280090103> (cited on page 27).
- Luedtke, W. D., Landman, U., Chiu, Y.-H., Levandier, D. J., Dressler, R. A., Sok, S., & Gordon, M. S. (2008). Nanojets, electrospray, and ion field evaporation: Molecular dynamics simulations and laboratory experiments. *The Journal of Physical Chemistry A*, *112*(40), 9628–9649. <https://doi.org/10.1021/jp804585y> (cited on page 27).
- Sun, S., Huang, C., Wang, Y., Liu, Y., Zhang, J., Zhou, J., Gao, F., Yang, F., Chen, R., Mulloy, B., Chai, W., Li, Y., & Bu, D. (2018). Toward automated identification of glycan branching patterns using multistage mass spectrometry with intelligent precursor selection. *Analytical Chemistry*, *90*(24), 14412–14422. <https://doi.org/10.1021/acs.analchem.8b03967> (cited on page 28).
- Jeanne Dit Fouque, K., & Fernandez-Lima, F. (2019). Recent advances in biological separations using trapped ion mobility spectrometry – mass spectrometry. *TrAC Trends in Analytical Chemistry*, *116*, 308–315. <https://doi.org/10.1016/j.trac.2019.04.010> (cited on page 29).
- Ridgeway, M. E., Bleiholder, C., Mann, M., & Park, M. A. (2019). Trends in trapped ion mobility – mass spectrometry instrumentation. *TrAC Trends in Analytical Chemistry*, *116*, 324–331. <https://doi.org/10.1016/j.trac.2019.03.030> (cited on page 29).
- Ribbenstedt, A., Ziarrusta, H., & Benskin, J. P. (2018). Development, characterization and comparisons of targeted and non-targeted metabolomics methods. *PLOS ONE*, *13*(11), e0207082. <https://doi.org/10.1371/journal.pone.0207082> (cited on page 29).
- Leao, T. F., Clark, C. M., Bauermeister, A., Elijah, E. O., Gentry, E. C., Husband, M., Oliveira, M. F., Bandeira, N., Wang, M., & Dorrestein, P. C. (2021). Quick-start infrastructure for untargeted metabolomics analysis in GNPS. *Nature Metabolism*, *3*(7), 880–882. <https://doi.org/10.1038/s42255-021-00429-0> (cited on page 29).
- Gauglitz, J. M., West, K. A., Bittremieux, W., Williams, C. L., Weldon, K. C., Panitchpakdi, M., Di Ottavio, F., Aceves, C. M., Brown, E., Sikora, N. C., Jarmusch, A. K., Martino, C., Tripathi, A., Meehan, M. J., Dorrestein, K., Shaffer, J. P., Coras, R., Vargas, F., Goldasich, L. D., ... Dorrestein, P. C. (2022). Enhancing untargeted metabolomics using metadata-based source annotation. *Nature Biotechnology*, 1–6. <https://doi.org/10.1038/s41587-022-01368-1> (cited on page 29).
- Link, H., Kochanowski, K., & Sauer, U. (2013). Systematic identification of allosteric protein-metabolite interactions that control enzyme activity in vivo. *Nature Biotechnology*, *31*(4), 357–361. <https://doi.org/10.1038/nbt.2489> (cited on page 29).

- Harwood, C. R., & Wipat, A. (2012). *Systems biology of bacteria* (1st ed). Elsevier Academic Press. (Cited on page 29).
- van Gulik, W. M. (2010). Fast sampling for quantitative microbial metabolomics. *Current Opinion in Biotechnology*, 21(1), 27–34. <https://doi.org/10.1016/j.copbio.2010.01.008> (cited on page 29).
- Schramm, T., Lempp, M., Beuter, D., Sierra, S. G., Glatter, T., & Link, H. (2020). High-throughput enrichment of temperature-sensitive argininosuccinate synthetase for two-stage citrulline production in *e. coli*. *Metabolic Engineering*. <https://doi.org/10.1016/j.ymben.2020.03.004> (cited on page 29).
- Wu, L., Mashego, M. R., van Dam, J. C., Proell, A. M., Vinke, J. L., Ras, C., van Winden, W. A., van Gulik, W. M., & Heijnen, J. J. (2005). Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly ¹³C-labeled cell extracts as internal standards. *Analytical Biochemistry*, 336(2), 164–171. <https://doi.org/10.1016/j.ab.2004.09.001> (cited on page 30).
- Guo, J., Shen, S., Xing, S., Yu, H., & Huan, T. (2021). ISFrag: De novo recognition of in-source fragments for liquid chromatography–mass spectrometry data. *Analytical Chemistry*, 93(29), 10243–10250. <https://doi.org/10.1021/acs.analchem.1c01644> (cited on page 30).
- Senan, O., Aguilar-Mogas, A., Navarro, M., Capellades, J., Noon, L., Burks, D., Yanes, O., Guimerà, R., & Sales-Pardo, M. (2019). CliqueMS: A computational tool for annotating in-source metabolite ions from LC-MS untargeted metabolomics data based on a coelution similarity network. *Bioinformatics*, 35(20), 4089–4097. <https://doi.org/10.1093/bioinformatics/btz207> (cited on page 30).
- Lalwani, M. A., Zhao, E. M., & Avalos, J. L. (2018). Current and future modalities of dynamic control in metabolic engineering. *Current Opinion in Biotechnology*, 52, 56–65. <https://doi.org/10.1016/j.copbio.2018.02.007> (cited on page 31).
- Hartline, C. J., Schmitz, A. C., Han, Y., & Zhang, F. (2021). Dynamic control in metabolic engineering: Theories, tools, and applications. *Metabolic Engineering*, 63, 126–140. <https://doi.org/10.1016/j.ymben.2020.08.015> (cited on page 31).
- Brockman, I. M., & Prather, K. L. J. (2015a). Dynamic metabolic engineering: New strategies for developing responsive cell factories. *Biotechnology Journal*, 10(9), 1360–1369. <https://doi.org/10.1002/biot.201400422> (cited on page 31).
- Venayak, N., Anesiadis, N., Cluett, W. R., & Mahadevan, R. (2015). Engineering metabolism through dynamic control. *Current Opinion in Biotechnology*, 34, 142–152. <https://doi.org/10.1016/j.copbio.2014.12.022> (cited on page 31).
- Cress, B. F., Trantas, E. A., Ververidis, F., Linhardt, R. J., & Koffas, M. A. (2015). Sensitive cells: Enabling tools for static and dynamic control of microbial metabolic pathways. *Current Opinion in Biotechnology*, 36, 205–214. <https://doi.org/10.1016/j.copbio.2015.09.007> (cited on page 31).
- Burg, J. M., Cooper, C. B., Ye, Z., Reed, B. R., Moreb, E. A., & Lynch, M. D. (2016). Large-scale bioprocess competitiveness: The potential of dynamic metabolic control in two-stage fermentations. *Current Opinion in Chemical Engineering*, 14, 121–136. <https://doi.org/10.1016/j.coche.2016.09.008> (cited on page 31).
- Soma, Y., Tsuruno, K., Wada, M., Yokota, A., & Hanai, T. (2014). Metabolic flux redirection from a central metabolic pathway toward a synthetic pathway using a metabolic toggle switch. *Metabolic Engineering*, 23, 175–184. <https://doi.org/10.1016/j.ymben.2014.02.008> (cited on page 31).

- Li, S., Jendresen, C. B., Grünberger, A., Ronda, C., Jensen, S. I., Noack, S., & Nielsen, A. T. (2016). Enhanced protein and biochemical production using CRISPRi-based growth switches. *Metabolic Engineering*, 38, 274–284. <https://doi.org/10.1016/j.ymben.2016.09.003> (cited on pages 31, 34).
- Gupta, A., Reizman, I. M. B., Reisch, C. R., & Prather, K. L. J. (2017). Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nature Biotechnology*, 35(3), 273–279. <https://doi.org/10.1038/nbt.3796> (cited on pages 31, 33).
- Lo, T.-M., Chng, S. H., Teo, W. S., Cho, H.-S., & Chang, M. W. (2016). A two-layer gene circuit for decoupling cell growth from metabolite production. *Cell Systems*, 3(2), 133–143. <https://doi.org/10.1016/j.cels.2016.07.012> (cited on page 31).
- Chubukov, V., & Sauer, U. (2014). Environmental dependence of stationary-phase metabolism in bacillus subtilis and escherichia coli. *Applied and Environmental Microbiology*, 80(9), 2901–2909. <https://doi.org/10.1128/AEM.00061-14> (cited on pages 33, 34).
- Chubukov, V., Desmarais, J. J., Wang, G., Chan, L. J. G., Baidoo, E. E., Petzold, C. J., Keasling, J. D., & Mukhopadhyay, A. (2017). Engineering glucose metabolism of escherichia coli under nitrogen starvation. *npj Systems Biology and Applications*, 3(1), 16035. <https://doi.org/10.1038/npjbsa.2016.35> (cited on pages 33–35).
- Michalowski, A., Siemann-Herzberg, M., & Takors, R. (2017). Escherichia coli HGT: Engineered for high glucose throughput even under slowly growing or resting conditions. *Metabolic Engineering*, 40, 93–103. <https://doi.org/10.1016/j.ymben.2017.01.005> (cited on pages 33, 34).
- Sonderegger, M., Schümperli, M., & Sauer, U. (2005). Selection of quiescent escherichia coli with high metabolic activity. *Metabolic Engineering*, 7(1), 4–9. <https://doi.org/10.1016/j.ymben.2004.05.005> (cited on pages 33–35).
- Brockman, I. M., & Prather, K. L. J. (2015b). Dynamic knockdown of e. coli central metabolism for redirecting fluxes of primary metabolites. *Metabolic Engineering*, 28, 104–113. <https://doi.org/10.1016/j.ymben.2014.12.005> (cited on page 33).
- Solomon, K. V., Sanders, T. M., & Prather, K. L. J. (2012). A dynamic metabolite valve for the control of central carbon metabolism. *Metabolic Engineering*, 14(6), 661–671. <https://doi.org/10.1016/j.ymben.2012.08.006> (cited on page 33).
- Venayak, N., von Kamp, A., Klamt, S., & Mahadevan, R. (2018). MoVE identifies metabolic valves to switch between phenotypic states. *Nature Communications*, 9(1), 1–9. <https://doi.org/10.1038/s41467-018-07719-4> (cited on page 33).
- Harder, B.-J., Bettenbrock, K., & Klamt, S. (2018). Temperature-dependent dynamic control of the TCA cycle increases volumetric productivity of itaconic acid production by *Escherichia coli*. *Biotechnology and Bioengineering*, 115(1), 156–164. <https://doi.org/10.1002/bit.26446> (cited on pages 33, 42).
- Cho, H.-S., Seo, S. W., Kim, Y. M., Jung, G. Y., & Park, J. M. (2012). Engineering glyceraldehyde-3-phosphate dehydrogenase for switching control of glycolysis in escherichia coli. *Biotechnology and Bioengineering*, 109(10), 2612–2619. <https://doi.org/10.1002/bit.24532> (cited on pages 33, 42).
- Rowe, D. C. D., & Summers, D. K. (1999). The quiescent-cell expression system for protein synthesis in *Escherichia coli*. *Applied and Environmental Microbiology*, 65(6), 2710–2715. <https://doi.org/10.1128/AEM.65.6.2710-2715.1999> (cited on page 34).

- Hommais, F., Krin, E., Laurent-Winter, C., Soutourina, O., Malpertuy, A., Le Caer, J.-P., Danchin, A., & Bertin, P. (2001). Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, h-NS. *Molecular Microbiology*, *40*(1), 20–36. <https://doi.org/10.1046/j.1365-2958.2001.02358.x> (cited on page 34).
- Izard, J., Gomez Balderas, C. D., Ropers, D., Lacour, S., Song, X., Yang, Y., Lindner, A. B., Geiselmann, J., & Jong, H. (2015). A synthetic growth switch based on controlled expression of RNA polymerase. *Molecular Systems Biology*, *11*(11), 840. <https://doi.org/10.15252/msb.20156382> (cited on page 34).
- Chesbro, W., Arbige, M., & Eifert, R. (1990). When nutrient limitation places bacteria in the domains of slow growth: Metabolic, morphologic and cell cycle behavior. *FEMS Microbiology Ecology*, *7*(2), 103–119. <https://doi.org/10.1111/j.1574-6941.1990.tb01677.x> (cited on page 34).
- Nyström, T. (1998). To be or not to be: The ultimate decision of the growth-arrested bacterial cell. *FEMS Microbiology Reviews*, *21*(4), 283–290. [https://doi.org/10.1016/S0168-6445\(97\)00060-0](https://doi.org/10.1016/S0168-6445(97)00060-0) (cited on page 34).
- Venturi, V. (2003). Control of rpoS transcription in escherichia coli and pseudomonas: Why so different?: Regulation of rpoS expression. *Molecular Microbiology*, *49*(1), 1–9. <https://doi.org/10.1046/j.1365-2958.2003.03547.x> (cited on page 34).
- Chubukov, V., Gerosa, L., Kochanowski, K., & Sauer, U. (2014). Coordination of microbial metabolism. *Nature Reviews Microbiology*, *12*(5), 327–340. <https://doi.org/10.1038/nrmicro3238> (cited on page 34).
- Gefen, O., Fridman, O., Ronin, I., & Balaban, N. Q. (2014). Direct observation of single stationary-phase bacteria reveals a surprisingly long period of constant protein production activity. *Proceedings of the National Academy of Sciences*, *111*(1), 556–561. <https://doi.org/10.1073/pnas.1314114111> (cited on page 34).
- Traxler, M. F., Summers, S. M., Nguyen, H.-T., Zacharia, V. M., Hightower, G. A., Smith, J. T., & Conway, T. (2008). The global, ppGpp-mediated stringent response to amino acid starvation in escherichia coli. *Molecular Microbiology*, *68*(5), 1128–1148. <https://doi.org/10.1111/j.1365-2958.2008.06229.x> (cited on page 34).
- Brown, D. R., Barton, G., Pan, Z., Buck, M., & Wigneshweraraj, S. (2014). Nitrogen stress response and stringent response are coupled in escherichia coli. *Nature Communications*, *5*(1), 4115. <https://doi.org/10.1038/ncomms5115> (cited on page 34).
- Hädicke, O., Bettenbrock, K., & Klamt, S. (2015). Enforced ATP futile cycling increases specific productivity and yield of anaerobic lactate production in escherichia coli. *Biotechnology and Bioengineering*, *112*(10), 2195–2199. <https://doi.org/10.1002/bit.25623> (cited on page 35).
- Jensen, P. R., Snoep, J. L., & Westerhoff, H. V. (2003, January 28). *Method of improving the production of biomass or a desired product from a cell* (pat. No. 6511836B1). Retrieved November 2, 2020, from <https://patents.google.com/patent/US6511836B1/en> (cited on page 35).
- Semkiv, M. V., Dmytruk, K. V., Abbas, C. A., & Sibirny, A. A. (2016). Activation of futile cycles as an approach to increase ethanol yield during glucose fermentation in saccharomyces cerevisiae. *Bioengineered*, *7*(2), 106–111. <https://doi.org/10.1080/21655979.2016.1148223> (cited on page 35).
- Liu, J., Kandasamy, V., Würtz, A., Jensen, P. R., & Solem, C. (2016). Stimulation of acetoin production in metabolically engineered lactococcus lactis by increasing ATP demand. *Applied Microbiology and Biotechnology*, *100*(22), 9509–9517. <https://doi.org/10.1007/s00253-016-7687-1> (cited on page 35).

- Holm, A. K., Blank, L. M., Oldiges, M., Schmid, A., Solem, C., Jensen, P. R., & Vemuri, G. N. (2010). Metabolic and transcriptional response to cofactor perturbations in *escherichia coli*. *The Journal of Biological Chemistry*, 285(23), 17498–17506. <https://doi.org/10.1074/jbc.M109.095570> (cited on page 35).
- Koebmann, B. J., Westerhoff, H. V., Snoep, J. L., Solem, C., Pedersen, M. B., Nilsson, D., Michelsen, O., & Jensen, P. R. (2002). The extent to which ATP demand controls the glycolytic flux depends strongly on the organism and conditions for growth. *Molecular Biology Reports*, 29(1), 41–45. <https://doi.org/10.1023/A:1020398117281> (cited on page 35).
- Chao, Y. P., & Liao, J. C. (1994). Metabolic responses to substrate futile cycling in *escherichia coli*. *The Journal of Biological Chemistry*, 269(7), 5122–5126 (cited on page 35).
- Honig, B., & Yang, A.-S. (1995, January 1). Free energy balance in protein folding. In C. B. Anfinsen, F. M. Richards, J. T. Edsall, & D. S. Eisenberg (Eds.), *Advances in protein chemistry* (pp. 27–58, Vol. 46). Academic Press. [https://doi.org/10.1016/S0065-3233\(08\)60331-9](https://doi.org/10.1016/S0065-3233(08)60331-9) (cited on page 36).
- Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry*, 29(31), 7133–7155. <https://doi.org/10.1021/bi00483a001> (cited on page 36).
- Khan, S. U., & Pal, M. A. (2011). Paneer production: A review. *Journal of Food Science and Technology*, 48(6), 645–660. <https://doi.org/10.1007/s13197-011-0247-x> (cited on page 36).
- Jarzab, A., Kurzawa, N., Hopf, T., Moerch, M., Zecha, J., Leijten, N., Bian, Y., Musiol, E., Maschberger, M., Stoehr, G., Becher, I., Daly, C., Samaras, P., Mergner, J., Spanier, B., Angelov, A., Werner, T., Bantscheff, M., Wilhelm, M., ... Kuster, B. (2020). Meltome atlas—thermal proteome stability across the tree of life. *Nature Methods*, 17(5), 495–503. <https://doi.org/10.1038/s41592-020-0801-4> (cited on pages 36, 37).
- Radestock, S., & Gohlke, H. (2011). Protein rigidity and thermophilic adaptation: Protein rigidity and thermophilic adaptation. *Proteins: Structure, Function, and Bioinformatics*, 79(4), 1089–1108. <https://doi.org/10.1002/prot.22946> (cited on pages 36, 38).
- Leuenberger, P., Ganschä, S., Kahraman, A., Cappelletti, V., Boersema, P. J., von Mering, C., Claassen, M., & Picotti, P. (2017). Cell-wide analysis of protein thermal unfolding reveals determinants of thermostability. *Science*, 355(6327), eaai7825. <https://doi.org/10.1126/science.aai7825> (cited on page 36).
- Mallik, S., & Kundu, S. (2013). A comparison of structural and evolutionary attributes of *escherichia coli* and *thermus thermophilus* small ribosomal subunits: Signatures of thermal adaptation (G. E. Fox, Ed.). *PLoS ONE*, 8(8), e69898. <https://doi.org/10.1371/journal.pone.0069898> (cited on page 36).
- Vieille, C., & Zeikus, G. J. (2001). Hyperthermophilic enzymes: Sources, uses, and molecular mechanisms for thermostability. *Microbiology and Molecular Biology Reviews*, 65(1), 1–43. <https://doi.org/10.1128/MMBR.65.1.1-43.2001> (cited on pages 36, 38).
- Shriver, J. W. (Ed.). (2009). *Protein structure, stability, and interactions* (Vol. 490). Humana Press. <https://doi.org/10.1007/978-1-59745-367-7> (cited on page 36).
- Goldenzweig, A., & Fleishman, S. J. (2018). Principles of protein stability and their application in computational design. *Annual Review of Biochemistry*, 87(1), 105–129. <https://doi.org/10.1146/annurev-biochem-062917-012102> (cited on page 36).
- Shirley, B. A. (Ed.). (1995). *Protein stability and folding: Theory and practice*. Humana Press. (Cited on page 36).

- Varadarajan, R., Nagarajaram, H. A., & Ramakrishnan, C. (1996). A procedure for the prediction of temperature-sensitive mutants of a globular protein based solely on the amino acid sequence. *Proceedings of the National Academy of Sciences*, 93(24), 13908–13913. <https://doi.org/10.1073/pnas.93.24.13908> (cited on pages 36, 41).
- Huynh, K., & Partch, C. L. (2015). Analysis of protein stability and ligand interactions by thermal shift assay. *Current Protocols in Protein Science*, 79(1), 28.9.1–28.9.14. <https://doi.org/10.1002/0471140864.ps2809s79> (cited on page 37).
- Franken, H., Mathieson, T., Childs, D., Sweetman, G. M. A., Werner, T., Tögel, I., Doce, C., Gade, S., Bantscheff, M., Drewes, G., Reinhard, F. B. M., Huber, W., & Savitski, M. M. (2015). Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nature Protocols*, 10(10), 1567–1593. <https://doi.org/10.1038/nprot.2015.101> (cited on page 37).
- Sanavia, T., Birolo, G., Montanucci, L., Turina, P., Capriotti, E., & Fariselli, P. (2020). Limitations and challenges in protein stability prediction upon genome variations: Towards future applications in precision medicine. *Computational and Structural Biotechnology Journal*, 18, 1968–1979. <https://doi.org/10.1016/j.csbj.2020.07.011> (cited on page 37).
- Pancotti, C., Benevenuta, S., Birolo, G., Alberini, V., Repetto, V., Sanavia, T., Capriotti, E., & Fariselli, P. (2022). Predicting protein stability changes upon single-point mutation: A thorough comparison of the available tools on a new dataset. *Briefings in Bioinformatics*, 23(2), bbab555. <https://doi.org/10.1093/bib/bbab555> (cited on page 37).
- Tokuriki, N., & Tawfik, D. S. (2009). Stability effects of mutations and protein evolvability. *Current Opinion in Structural Biology*, 19(5), 596–604. <https://doi.org/10.1016/j.sbi.2009.08.003> (cited on page 37).
- Yue, P., Li, Z., & Moulton, J. (2005). Loss of protein structure stability as a major causative factor in monogenic disease. *Journal of Molecular Biology*, 353(2), 459–473. <https://doi.org/10.1016/j.jmb.2005.08.020> (cited on page 37).
- Smith, B. D., & Raines, R. T. (2006). Genetic selection for critical residues in ribonucleases. *Journal of Molecular Biology*, 362(3), 459–478. <https://doi.org/10.1016/j.jmb.2006.07.020> (cited on page 37).
- Bershtein, S., Segal, M., Bekerman, R., Tokuriki, N., & Tawfik, D. S. (2006). Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature*, 444(7121), 929–932. <https://doi.org/10.1038/nature05385> (cited on page 37).
- Camps, M., Herman, A., Loh, E., & Loeb, L. A. (2007). Genetic constraints on protein evolution. *Critical Reviews in Biochemistry and Molecular Biology*, 42(5), 313–326. <https://doi.org/10.1080/10409230701597642> (cited on page 37).
- Romero, P. A., & Arnold, F. H. (2009). Exploring protein fitness landscapes by directed evolution. *Nature Reviews Molecular Cell Biology*, 10(12), 866–876. <https://doi.org/10.1038/nrm2805> (cited on page 37).
- Leader, B., Baca, Q. J., & Golan, D. E. (2008). Protein therapeutics: A summary and pharmacological classification. *Nature Reviews Drug Discovery*, 7(1), 21–39. <https://doi.org/10.1038/nrd2399> (cited on page 37).
- Walls, D., & Loughran, S. T. (Eds.). (2011). *Protein chromatography: Methods and protocols* (Vol. 681). Humana Press. <https://doi.org/10.1007/978-1-60761-913-0> (cited on page 37).

Chapter 1

- Xavier, J. S., Nguyen, T.-B., Karmarkar, M., Portelli, S., Rezende, P. M., Velloso, J. P. L., Ascher, D. B., & Pires, D. E. V. (2021). ThermoMutDB: A thermodynamic database for missense mutations. *Nucleic Acids Research*, 49, D475–D479. <https://doi.org/10.1093/nar/gkaa925> (cited on page 37).
- Guerois, R., Nielsen, J. E., & Serrano, L. (2002). Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *Journal of Molecular Biology*, 320(2), 369–387. [https://doi.org/10.1016/S0022-2836\(02\)00442-4](https://doi.org/10.1016/S0022-2836(02)00442-4) (cited on page 37).
- Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., & Serrano, L. (2005). The FoldX web server: An online force field. *Nucleic Acids Research*, 33, W382–W388. <https://doi.org/10.1093/nar/gki387> (cited on page 37).
- Yang, Y., Ding, X., Zhu, G., Niroula, A., Lv, Q., & Vihinen, M. (2019). ProTstab – predictor for cellular protein stability. *BMC Genomics*, 20(1), 804. <https://doi.org/10.1186/s12864-019-6138-7> (cited on page 37).
- Cao, H., Wang, J., He, L., Qi, Y., & Zhang, J. Z. (2019). DeepDDG: Predicting the stability change of protein point mutations using neural networks. *Journal of Chemical Information and Modeling*, 59(4), 1508–1514. <https://doi.org/10.1021/acs.jcim.8b00697> (cited on page 37).
- Daniel, R. M., & Danson, M. J. (2013). Temperature and the catalytic activity of enzymes: A fresh understanding. *FEBS Letters*, 587(17), 2738–2743. <https://doi.org/10.1016/j.febslet.2013.06.027> (cited on pages 38, 39).
- Timoféeff-Ressovsky, N. W. (1934). Über die Vitalität einiger Genmutationen und ihrer Kombinationen bei *Drosophila funebris* und ihre Abhängigkeit vom „genotypischen“ und vom äußeren Milieu. *Zeitschrift für Induktive Abstammungs- und Vererbungslehre*, 66(1), 319–344. <https://doi.org/10.1007/BF01739846> (cited on page 40).
- Dobzhansky, T., & Spassky, B. (1944). Genetics of natural populations. xi. manifestation of genetic variants in *Drosophila pseudoobscura* in different environments. *Genetics*, 29(3), 270–290. <https://doi.org/10.1093/genetics/29.3.270> (cited on page 40).
- Maas, W. K., & Davis, B. D. (1952). Production of an altered pantothenate-synthesizing enzyme by a temperature-sensitive mutant of *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 38(9), 785–797. <https://doi.org/10.1073/pnas.38.9.785> (cited on page 40).
- Hansen, F. G., & Atlung, T. (2018). The DnaA tale. *Frontiers in Microbiology*, 9, 319. <https://doi.org/10.3389/fmicb.2018.00319> (cited on page 40).
- Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K., & Sugino, A. (1968). Mechanism of DNA chain growth. i. possible discontinuity and unusual secondary structure of newly synthesized chains. *Proceedings of the National Academy of Sciences*, 59(2), 598–605. <https://doi.org/10.1073/pnas.59.2.598> (cited on page 40).
- Pauling, C., & Hamm, L. (1969). PROPERTIES OF a TEMPERATURE-SENSITIVE, RADIATION-SENSITIVE MUTANT OF *Escherichia coli*, II. DNA REPLICATION*. *Proceedings of the National Academy of Sciences*, 64(4), 1195–1202. <https://doi.org/10.1073/pnas.64.4.1195> (cited on page 40).
- Waard, A. D. (1965). The structural gene for deoxyribonucleic acid polymerase in bacteriophages t4 and t5. 54, 8 (cited on page 40).
- Carl, P. L. (1970). *Escherichia coli* mutants with temperature-sensitive synthesis of DNA. *Molecular and General Genetics MGG*, 109(2), 107–122. <https://doi.org/10.1007/BF00269647> (cited on page 40).

- Bitner, R. M., & Kuempel, P. L. (1982). P1 transduction mapping of the *trg* locus in *rac*⁺ and *rac* strains of *Escherichia coli* K-12. *Journal of Bacteriology*, *149*(2), 529–533. <https://doi.org/10.1128/jb.149.2.529-533.1982> (cited on page 41).
- Yasuda, S., & Hirota, Y. (1977). Cloning and mapping of the replication origin of *Escherichia coli*. *Proceedings of the National Academy of Sciences*, *74*(12), 5458–5462. <https://doi.org/10.1073/pnas.74.12.5458> (cited on page 41).
- Ben-Aroya, S., Coombes, C., Kwok, T., O'Donnell, K. A., Boeke, J. D., & Hieter, P. (2008). Toward a comprehensive temperature-sensitive mutant repository of the essential genes of *Saccharomyces cerevisiae*. *Molecular Cell*, *30*(2), 248–258. <https://doi.org/10.1016/j.molcel.2008.02.021> (cited on page 41).
- Li, Z., Vizeacoumar, F. J., Bahr, S., Li, J., Warringer, J., Vizeacoumar, F. S., Min, R., VanderSluis, B., Bellay, J., DeVit, M., Fleming, J. A., Stephens, A., Haase, J., Lin, Z.-Y., Baryshnikova, A., Lu, H., Yan, Z., Jin, K., Barker, S., ... Boone, C. (2011). Systematic exploration of essential yeast gene function with temperature-sensitive mutants. *Nature Biotechnology*, *29*(4), 361–367. <https://doi.org/10.1038/nbt.1832> (cited on page 41).
- Kofoed, M., Milbury, K. L., Chiang, J. H., Sinha, S., Ben-Aroya, S., Giaever, G., Nislow, C., Hieter, P., & Stirling, P. C. (2015). An updated collection of sequence barcoded temperature-sensitive alleles of yeast essential genes. *Genetics*, *195*(9), 1879–1887. <https://doi.org/10.1534/g3.115.019174> (cited on page 41).
- Ben-Aroya, S., Pan, X., Boeke, J. D., & Hieter, P. (2010, January 1). Chapter 8 - making temperature-sensitive mutants. In *Methods in enzymology* (pp. 181–204, Vol. 470). Academic Press. [https://doi.org/10.1016/S0076-6879\(10\)70008-2](https://doi.org/10.1016/S0076-6879(10)70008-2) (cited on page 41).
- Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E. D., Sevier, C. S., Ding, H., Koh, J. L., Toufighi, K., Mostafavi, S., Prinz, J., St. Onge, R. P., VanderSluis, B., Makhnevych, T., Vizeacoumar, F. J., Alizadeh, S., Bahr, S., Brost, R. L., Chen, Y., ... Boone, C. (2010). The genetic landscape of a cell. *Science*, *327*(5964), 425–431. <https://doi.org/10.1126/science.1180823> (cited on page 41).
- Costanzo, M., VanderSluis, B., Koch, E. N., Baryshnikova, A., Pons, C., Tan, G., Wang, W., Usaj, M., Hanchard, J., Lee, S. D., Pelechano, V., Styles, E. B., Billmann, M., van Leeuwen, J., van Dyk, N., Lin, Z.-Y., Kuzmin, E., Nelson, J., Piotrowski, J. S., ... Boone, C. (2016). A global genetic interaction network maps a wiring diagram of cellular function. *Science*, *353*(6306), aaf1420–aaf1420. <https://doi.org/10.1126/science.aaf1420> (cited on page 41).
- Tan, K. P., Khare, S., Varadarajan, R., & Madhusudhan, M. S. (2014). TSpred: A web server for the rational design of temperature-sensitive mutants. *Nucleic Acids Research*, *42*, W277–W284. <https://doi.org/10.1093/nar/gku319> (cited on page 41).
- Poultney, C. S., Butterfoss, G. L., Gutwein, M. R., Drew, K., Gresham, D., Gunsalus, K. C., Shasha, D. E., & Bonneau, R. (2011). Rational design of temperature-sensitive alleles using computational structure prediction (V. N. Uversky, Ed.). *PLoS ONE*, *6*(9), e23947. <https://doi.org/10.1371/journal.pone.0023947> (cited on page 41).
- Servant, P., Grandvalet, C., & Mazodier, P. (2000). The RheA repressor is the thermosensor of the HSP18 heat shock response in *Streptomyces albus*. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(7), 3538–3543. Retrieved October 29, 2020, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC16275/> (cited on page 42).

Chapter 1

- Pearce, S. C., McWhinnie, R. L., & Nano, F. E. (2017). Synthetic temperature-inducible lethal gene circuits in *Escherichia coli*. *Microbiology*, *163*(4), 462–471. <https://doi.org/10.1099/mic.0.000446> (cited on page 42).
- Lynch, M., Louie, M., Copley, S., Spindler, E., Prather, B., Lipscomb, M., Lipscomb, T., Liao, H., Hogsett, D., & Evans, R. (2019, July 2). *Microorganisms and methods for the production of fatty acids and fatty acid derived products* (U.S. pat. No. 10337038B2). <https://patents.google.com/patent/US10337038B2/en> (cited on page 42).

Chapter 2

Von der Stöchiometrie zur Kontrolle metabolischer Netzwerke

Thorben Schramm, Hannes Link

This chapter was published in: Schramm and Link, 2021, *BIOspektrum*, volume 27, 2021, pages 34–36, DOI: <https://doi.org/10.1007/s12268-021-1538-0>, (*non-peer-reviewed review*), (license: Creative Commons "Namensnennung 4.0 International" (CC BY 4.0), <http://creativecommons.org/licenses/by/4.0/deed.de>, no modifications). My contribution to this review was in share with Prof. Dr.-Ing. Hannes P. Link the conceptualization, the writing, and the preparation of figures.

Abstract

Cellular metabolism is very complex and extensively regulated. For many organisms we know almost the complete set of biochemical reactions in their metabolic network. However, it is not well understood how these reactions are regulated and how they interact in order to enable cellular functions. In this review, we describe recent methodological advances to study metabolic networks with a focus on bacterial metabolism.

Metabolische Netzwerke gehören zu den komplexesten und zugleich zu den am besten untersuchten biologischen Netzwerken, die wir kennen. Zum Beispiel sind am metabolischen Netzwerk von *Escherichia coli* 1.192 Metabolite, 1.515 Gene, 1.515 Proteine und 2.719 biochemische Reaktionen beteiligt (Monk et al., 2017). Für viele Organismen ist der Aufbau des metabolischen Netzwerks ähnlich gut beschrieben. Allerdings ist unklar, wie metabolische Netzwerke so genau reguliert werden, dass alle Zell- und Energiebausteine kontinuierlich und in der erforderlichen Menge produziert werden. Um diese komplexe Aufgabe zu meistern, haben Zellen Regulationsmechanismen entwickelt, die alle Reaktionen im Netzwerk präzise aufeinander abstimmen. Diese Regulationsmechanismen zu finden und ihre Funktion im Netzwerk zu verstehen, ist ein aktuelles Ziel der Stoffwechselforschung. Ein umfassendes Verständnis metabolischer Netzwerke und deren Regulation ist besonders für neue biotechnologische Anwendungen wichtig. Die synthetische Biologie entwickelt z. B. bereits sehr effiziente und neuartige Stoffwechselwege und ermöglicht damit komplexe Synthesen von Naturstoffen in Hefen (Galanie et al., 2015) oder CO₂-Fixierung in *E. coli* (Gleizer et al., 2019). Allerdings sind diese synthetischen Stoffwechselwege oft sehr langsam, weil Regulationsmechanismen fehlen, sodass der restliche Stoffwechsel überlastet wird. Doch bevor wir synthetische Regulation entwickeln können, müssen wir das Zusammenspiel aller Reaktionen in natürlichen metabolischen Netzwerken verstehen.

Modelle metabolischer Netzwerke aus Genomen

Die Stöchiometrie metabolischer Netzwerke ist der Ausgangspunkt für fast alle Analysen und Modelle des Stoffwechsels. Dank der biochemischen Erkenntnisse des 20. Jahrhunderts sind die Reaktionen metabolischer Netzwerke und deren Stöchiometrie

sehr gut verstanden. Zusammen mit der Sequenzierung ganzer Genome können so großskalige Modelle metabolischer Netzwerke erstellt werden. Solche großen metabolischen Modelle (oder genome-scale models) existieren inzwischen für viele Organismen, von Bakterien bis hin zu menschlichen Zellen, und sie verbinden lückenlos alle metabolischen Reaktionen miteinander. Dadurch können wir geschlossene Massenbilanzen aufstellen und theoretische Stoffflüsse berechnen, beispielsweise für maximales Wachstum oder maximale Produktion eines gewünschten Stoffs (Monk et al., 2017). Allerdings sind diese Vorhersagen statisch, und durch die Berechnungen kennen wir zwar die optimalen Stoffflüsse im Netzwerk, leider jedoch nicht, wie diese erreicht und eingestellt werden. Die dazu erforderlichen dynamischen Modelle, die die Regulation und Kinetik aller Enzyme berücksichtigen, können aktuell nur für sehr kleine Netzwerke erstellt werden.

Hunderte Enzyme aufeinander abstimmen

Die Regulation metabolischer Netzwerke stellt sicher, dass die Raten aller Reaktionen genau aufeinander abgestimmt sind. Denn nur, wenn sich alle Reaktionen in einem Fließgleichgewicht befinden, entsteht ein kontinuierlicher Stofffluss durch das Netzwerk, der genau an den Bedarf der Zelle angepasst ist (Abb. 2.1.A). Weicht eine Reaktion vom Fließgleichgewicht ab, könnte das zu einem Engpass führen, der schnell das ganze Netzwerk stört. Ein Engpass im Biosyntheseweg einer einzelnen Aminosäure kann z. B. innerhalb kürzester Zeit die gesamte Proteinbiosynthese und damit das Zellwachstum stören. Da Bakterien aber sehr kontinuierlich und konstant wachsen (Wang et al., 2010), können wir davon ausgehen, dass metabolische Netzwerke sehr gut eingestellt sind und dass die Reaktionsraten aller Enzyme immer wieder in Echtzeit präzise aufeinander abgestimmt werden. Aber wie gelingt die Abstimmung bei Hunderten von Enzymen? Im einfachsten Fall folgt ein Enzym einer Michaelis-Menten-Kinetik (Abb. 2.1.B) und arbeitet mit maximaler Geschwindigkeit (v_{max}). Dann bestimmt nur die Konzentration des Enzyms die Reaktionsrate. Experimentell lässt sich aber zeigen, dass Enzyme nicht nahe am v_{max} arbeiten, weil Enzymkonzentrationen relativ stark abfallen können, bevor ein Wachstumsdefekt eintritt (Donati et al., 2021). Das deutet darauf hin, dass die meisten Enzyme Überkapazitäten haben, die z. B. durch Änderung der Substratkonzentration sofort abgerufen werden können. Allerdings ist die Expression von Enzymen auch mit Kosten verbunden, weshalb Enzyme nicht zu weit weg vom v_{max} arbeiten sollten. Die Konsequenz daraus ist, dass die Expression von Enzymen oft stark reguliert ist. Allosterische Interaktionen zwischen Transkriptionsfaktoren und Metaboliten spie-

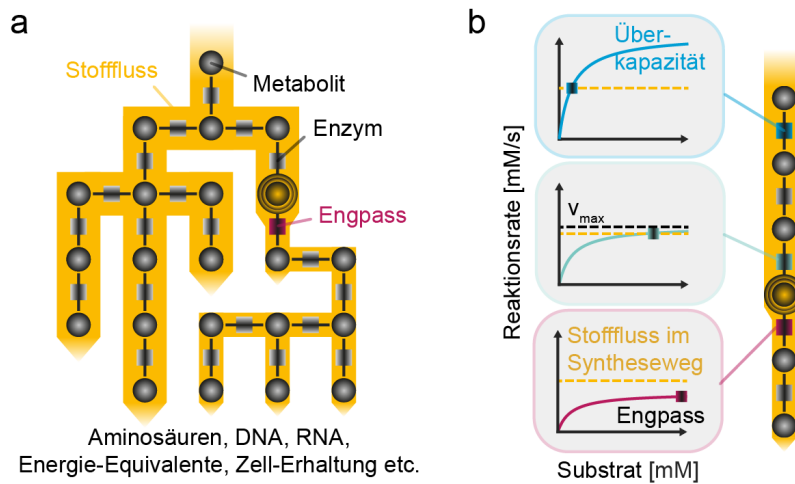


Figure 2.1: Fließgleichgewicht in metabolischen Netzwerken. A, Metabolische Netzwerke bestehen aus Enzymen (Quadrate) und Metaboliten (Kreise), die durch biochemische Reaktionen miteinander verbunden sind. Das aktuelle Modell des metabolischen Netzwerks von *Escherichia coli* umfasst 1.192 Metabolite, 1.515 Gene, 1.515 Proteine und 2.719 biochemische Reaktionen (Monk et al., 2017). B, Im Fließgleichgewicht verlaufen alle Reaktionen mit gleicher Rate, die dem Stofffluss durch das Netzwerk (gelb gestrichelt) entspricht. Die Rate einer Reaktion folgt im einfachsten Fall einer Michaelis-Menten-Kinetik und wird durch die Konzentrationen des Substrats und des Enzyms beeinflusst. Die Konzentration des Enzyms entscheidet, ob der Stofffluss in dem Syntheseweg aufrechterhalten werden kann, ob an der Reaktion ein Engpass entsteht oder die Reaktion Überkapazitäten hat. v_{max} gibt die maximale Reaktionsgeschwindigkeit bei einer bestimmten Enzymkonzentration an.

len dabei eine wesentliche Rolle und erlauben komplexe Wechselwirkungen und Rückkopplungen zwischen metabolischen und genetischen Netzwerken. So konnten wir am Beispiel der Aminosäuresynthese von *E. coli* zeigen, dass ein Zusammenspiel mehrerer negativer Rückkopplungen zwischen Aminosäuren und Proteinen die Enzymkonzentrationen so präzise einstellt, dass dies gleichzeitig die Effizienz und die Robustheit der Stoffwechselwege maximiert (Sander et al., 2019). Die Konzentrationen von Metaboliten sind also Schlüsselsignale für die Regulation metabolischer Netzwerke.

Metabolite als Informationsträger

Auf das gesamte metabolische Netzwerk bezogen wissen wir nicht, welche Metabolite Schlüsselsignale sind und welche Einfluss auf die Genexpression nehmen. Zurzeit gibt es auch kaum Methoden, um Interaktionen zwischen Metaboliten und der Genexpression systematisch zu finden. Da unser Wissen hauptsächlich auf in vitro-Messungen

basiert, kennen wir selbst in *E. coli* nur insgesamt 137 Interaktionen zwischen Metaboliten und Transkriptionsfaktoren (Lempp et al., 2019), obwohl tausende solcher Interaktionen möglich wären. Um Metabolite mit regulatorischer Funktion zu finden, ist es wichtig, intrazelluläre Metabolite mit größtmöglicher Abdeckung und Präzision zu messen. Metabolomik-Methoden, die auf Massenspektrometrie basieren, ermöglichen es heute, hunderte Metabolite in Zellextrakten zu messen, entweder durch ungezielte Analysen (Fuhrer et al., 2011; Lu et al., 2020) oder gezielt (Guder et al., 2017). Aber selbst mit modernen Massenspektrometrie-Methoden ist die Anzahl an messbaren Metaboliten bisher limitiert, da die Metabolite einer Zelle chemisch zu divers sind, um sie alle mit einer einzigen chromatographischen Methode aufzutrennen. Eine Lösung dafür könnte die direkte Injektion der Proben in besonders hoch auflösende Massenspektrometer sein, ohne eine chromatographische Auftrennung. Ein großer Vorteil dieser Massenspektrometrie-Methoden ist die Geschwindigkeit, mit der Proben prozessiert und gemessen werden können, was es sogar ermöglicht, Metabolite in lebenden Zellen in Echtzeit zu messen (Link et al., 2015). Da wir nun metabolische Netzwerke vermessen können, stellt sich die Frage, welche Experimente geeignet sind, um Regulation nachzuweisen und zu verstehen. Ein klassisches Vorgehen ist, das Netzwerk zu stören und von den gemessenen Veränderungen auf Regulationsmechanismen rückzuschließen. In *stimulus-response*-Experimenten wird ein äußerer Faktor wie die Kohlenstoffquelle verändert, um dann zeitlich aufgelöst Metabolitkonzentrationen zu messen. Mit *stimulus-response*-Experimenten wurden die ersten *in vivo*-Parameter von Enzymen abgeschätzt (Wu et al., 2006), und vor kurzem konnten wir damit neue Metabolit-Transkriptionsfaktor-Interaktionen identifizieren (Lempp et al., 2019). Man kann das metabolische Netzwerk aber auch gezielter stören. Beispielsweise wurde in 4.913 Hefestämmen jeweils ein Gen aus dem Genom entfernt und die intrazellulären Aminosäurekonzentrationen gemessen. Die Aminosäuren zeigten Signaturen, die so spezifisch waren, dass sie Rückschlüsse auf die Funktion der einzelnen Gene zuließen (Mülleider et al., 2016). Zum Beispiel hatten Gene mit ähnlicher Funktion auch ähnliche Aminosäuresignaturen. Solche Experimente lassen erahnen, wieviel Information die Konzentration der Metabolite des gesamten Netzwerks enthalten. Um diese Informationen zu entschlüsseln, sind allerdings sehr große Datensätze notwendig, die möglichst alle Metabolite und tausende Störungen des metabolischen Netzwerks umfassen. Solche Störungen können mit CRISPR-basierten Methoden präzise und mit massivem Durchsatz erzeugt werden, da Millionen von mit CRISPR modifizierten Stämmen gebündelt in einem Reaktionsgefäß kloniert werden können. Mit CRISPR-Interferenz (CRISPRi) konnten wir die Expression jedes Gens im metabolis-

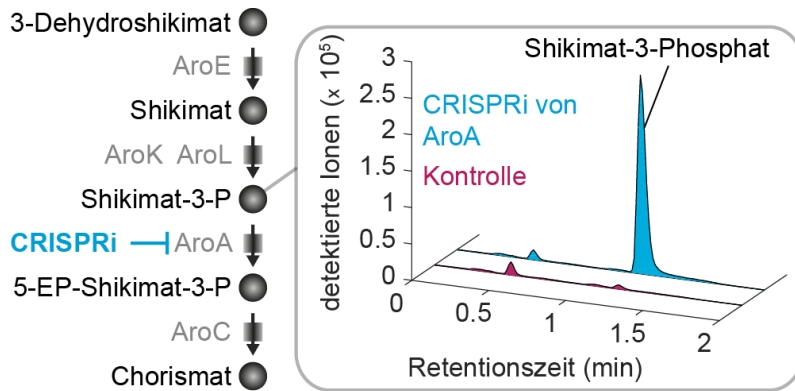


Figure 2.2: Mithilfe von CRISPR-Interferenz (CRISPRi) kann ganz gezielt die Konzentration eines Enzyms und somit auch der Stofffluss gestört werden. Häufig akkumuliert das Substrat einer gestörten Reaktion. Im Biosyntheseweg von aromatischen Aminosäuren konnte z. B. bei der Störung der von AroA-katalysierten Reaktion eine Akkumulation von Shikimat-3-Phosphat gemessen werden (Donati et al., 2021). Zum Vergleich wurde in Kontrollproben das CRISPRi-System nicht induziert. Die Messung der Metabolite erfolgte mittels Hochdurchsatz-Flüssigkeitschromatographie gekoppelt an Massenspektrometrie (Guder et al., 2017).

chen Netzwerk von *E. coli* stören, was sehr spezifische und lokale Veränderungen von Metaboliten erzeugte (Donati et al., 2021). Die stärksten Änderungen zeigten Metabolite, die einen direkten Einfluss auf die Aktivität des gestörten Enzyms haben, z. B. Substrate (Abb. 2.2), aber auch allosterische Effektoren. Mit CRISPR-Methoden können auch Punktmutationen präzise erzeugt werden, um allosterische Rückkopplungen zu entfernen und deren regulatorische Rolle zu untersuchen (Sander et al., 2019).

Design neuer Stoffwechsel-Netzwerke

Nachdem in der Vergangenheit ein Fokus auf der Kartierung von Genen, biochemischen Reaktionen und Transkriptionsfaktoren lag, geht es heute darum, neue metabolische Netzwerke zu designen. Die Erkenntnisse der letzten Jahre haben gezeigt, dass metabolische Netzwerke durch Regulation sowohl effizienter als auch robuster gegenüber Störungen werden. Metabolitkonzentrationen nehmen dabei als Signale eine zentrale Rolle ein. Um in Zukunft neue und bessere metabolische Netzwerke für medizinische und biotechnologische Anwendungen zu entwerfen, gilt es daher zu verstehen, welche Metabolite regulatorische Aufgaben haben und wie wir diese Regulation nutzen und gezielt verändern können.

References

- Schramm, T., & Link, H. (2021). Von der Stöchiometrie zur Kontrolle metabolischer Netzwerke. *BIOSpektrum*, 27(1), 34–36. <https://doi.org/10.1007/s12268-021-1538-0> (cited on page 65).
- Monk, J. M., Lloyd, C. J., Brunk, E., Mih, N., Sastry, A., King, Z., Takeuchi, R., Nomura, W., Zhang, Z., Mori, H., Feist, A. M., & Palsson, B. O. (2017). iML1515, a knowledgebase that computes escherichia coli traits. *Nature Biotechnology*, 35(10), 904–908. <https://doi.org/10.1038/nbt.3956> (cited on pages 66–68).
- Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M., & Smolke, C. D. (2015). Complete biosynthesis of opioids in yeast. *Science*, 349(6252), 1095–1100. <https://doi.org/10.1126/science.aac9373> (cited on page 66).
- Gleizer, S., Ben-Nissan, R., Bar-On, Y. M., Antonovsky, N., Noor, E., Zohar, Y., Jona, G., Krieger, E., Shamshoum, M., Bar-Even, A., & Milo, R. (2019). Conversion of escherichia coli to generate all biomass carbon from CO₂. *Cell*, 179(6), 1255–1263.e12. <https://doi.org/10.1016/j.cell.2019.11.009> (cited on page 66).
- Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A., & Jun, S. (2010). Robust growth of escherichia coli. *Current biology: CB*, 20(12), 1099–1103. <https://doi.org/10.1016/j.cub.2010.04.045> (cited on page 67).
- Donati, S., Kuntz, M., Pahl, V., Farke, N., Beuter, D., Glatter, T., Gomes-Filho, J. V., Randau, L., Wang, C.-Y., & Link, H. (2021). Multi-omics analysis of CRISPRi-knockdowns identifies mechanisms that buffer decreases of enzymes in e. coli metabolism. *Cell Systems*, 12(1), 56–67.e6. <https://doi.org/10.1016/j.cels.2020.10.011> (cited on pages 67, 70).
- Sander, T., Farke, N., Diehl, C., Kuntz, M., Glatter, T., & Link, H. (2019). Allosteric feedback inhibition enables robust amino acid biosynthesis in e. coli by enforcing enzyme overabundance. *Cell Systems*, 8(1), 66–75.e8. <https://doi.org/10.1016/j.cels.2018.12.005> (cited on pages 68, 70).
- Lempp, M., Farke, N., Kuntz, M., Freibert, S. A., Lill, R., & Link, H. (2019). Systematic identification of metabolites controlling gene expression in e. coli. *Nature Communications*, 10(1), 4463. <https://doi.org/10.1038/s41467-019-12474-1> (cited on page 69).
- Fuhrer, T., Heer, D., Begemann, B., & Zamboni, N. (2011). High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection–time-of-flight mass spectrometry. *Analytical Chemistry*, 83(18), 7074–7080. <https://doi.org/10.1021/ac201267k> (cited on page 69).
- Lu, W., Xing, X., Wang, L., Chen, L., Zhang, S., McReynolds, M. R., & Rabinowitz, J. D. (2020). Improved annotation of untargeted metabolomics data through buffer modifications that shift adduct mass and intensity. *Analytical Chemistry*, 92(17), 11573–11581. <https://doi.org/10.1021/acs.analchem.0c00985> (cited on page 69).
- Guder, J. C., Schramm, T., Sander, T., & Link, H. (2017). Time-optimized isotope ratio LC–MS/MS for high-throughput quantification of primary metabolites. *Analytical Chemistry*, 89(3), 1624–1631. <https://doi.org/10.1021/acs.analchem.6b03731> (cited on pages 69, 70).
- Link, H., Fuhrer, T., Gerosa, L., Zamboni, N., & Sauer, U. (2015). Real-time metabolome profiling of the metabolic switch between starvation and growth. *Nature Methods*, 12(11), 1091–1097. <https://doi.org/10.1038/nmeth.3584> (cited on page 69).
- Wu, L., Mashego, M. R., Proell, A. M., Vinke, J. L., Ras, C., van Dam, J., van Winden, W. A., van Gulik, W. M., & Heijnen, J. J. (2006). In vivo kinetics of primary metabolism in saccharomyces cerevisiae

Chapter 2

studied through prolonged chemostat cultivation. *Metabolic Engineering*, 8(2), 160–171. <https://doi.org/10.1016/j.ymben.2005.09.005> (cited on page 69).

Mülleder, M., Calvani, E., Alam, M. T., Wang, R. K., Eckerstorfer, F., Zelezniak, A., & Ralser, M. (2016). Functional metabolomics describes the yeast biosynthetic regulome. *Cell*, 167(2), 553–565.e12. <https://doi.org/10.1016/j.cell.2016.09.007> (cited on page 69).

Chapter 3

A network approach identifies in-source modifications of primary metabolites during flow-injection mass spectrometry

Thorben Schramm¹, Niklas Farke¹, Andreas Verhülsdonk, Hannes Link

Under revision at *ACS Analytical Chemistry*, 2022 (Stand: 12.08.2022). Following the CRediT authorship contribution statements, my contributions included the conceptualization, investigation, software, visualization, writing - original draft. Andreas Verhülsdonk and I handled the metabolite standards. I prepared metabolite extract samples and performed measurements by FI-MS. I was involved in the development of the processing software of FI-MS raw data. I was involved in the preparation of figures. I co-wrote the paper.

¹These authors contributed equally to this work.

Abstract

Flow-injection mass spectrometry (FI-MS) enables metabolomics studies with a very high sample-throughput. In most FI-MS methods, samples are directly injected into the electrospray ionization (ESI) source of a high-resolution mass spectrometer, and metabolites are annotated to ion peaks in the MS1 spectrum based on their exact mass over charge. Although ESI is considered a soft ionization technique, it can cause in-source modifications of analytes that are then misannotated to metabolites. In-source modifications include adduct formation, fragmentation, and other chemical reactions of metabolites. These effects are especially prominent in FI-MS because all analytes and the sample matrix enter the ESI at the same time. Here, we spiked authentic standards of 160 primary metabolites individually into an *Escherichia coli* metabolite extract and measured the thus derived 160 spike-in samples by FI-MS. Out of the 160 metabolites, 154 were annotated in their protonated or deprotonated form to ion peaks in the MS1 spectrum, and 134 of these ion peaks increased in the respective spike-in the standard. These results demonstrated that FI-MS can capture a wide-range of chemically diverse analytes within 30 seconds measurement time. However, the data also revealed extensive in-source modifications: across all 160 spike-in samples, we identified significant increases of 11,013 ion peaks in positive and negative mode combined. To explain these unknown m/z features, we connected them to the m/z feature of the (de-)protonated metabolite using information about mass differences and MS2 spectra. This resulted in networks that explained on average 49% of all significant m/z features. The networks showed that a single metabolite undergoes compound specific and often sequential in-source modifications like adductions, chemical reactions, and fragmentations. Taken together, our results show that FI-MS generates complex MS1 spectra, which can lead to a 68-fold overestimation of significant features. Yet, known mass differences and MS2 level information can explain these features and can therefore avoid misannotation of metabolites in FIMS analyses.

1 Introduction

Flow-injection mass spectrometry (FI-MS) does not rely on chromatographic separation of analytes (Fuhrer et al., 2011; Beckmann et al., 2008). Instead, samples are injected into the mobile phase that directly enters a mass spectrometer. Metabolites are then distinguished solely by their mass to charge ratio (m/z) in the MS1 spectrum. This makes FI-MS faster than methods with chromatographic separation (Reiter et al., 2021; Sarvin et al., 2020) and enables run times on the second time scale or even real-time metabolomics with living cells (Link et al., 2015).

FI-MS has been applied to measure the metabolome in various organisms including *Escherichia coli*, yeast, ruminants, and human cancer cell lines (Fuhrer et al., 2017; Anglada-Girotto et al., 2022; Holbrook-Smith et al., 2022; Dubuis et al., 2018; Rathahao-Paris et al., 2019). In these studies, hundreds or even thousands of strains or conditions could be analyzed due to the fast measurement time of FI-MS. Although FI-MS detects usually a very large number of m/z features (ion peaks in the MS1 spectrum), only a small fraction of m/z features can be annotated to metabolites. Thus, there is a large number of unexplained m/z features in FI-MS analyses, which could mean that either many metabolites are not known or that single metabolites produce multiple m/z features. Annotation of unknown m/z features is a general challenge in all untargeted metabolomics methods (Hartl et al., 2020; L. Wang et al., 2019; Kachman et al., 2020; Sindelar and Patti, 2020). For example, an untargeted LC-MS analysis suggested that out of 25,000 measured m/z features less than 1,000 originated from unique metabolites (Mahieu and Patti, 2017). The high number of m/z features in untargeted metabolomic methods is often attributed to contaminants, isotopes, modification of metabolites in the ion-source, and other mass spectrometry artifacts.

In-source fragmentation is one example of such mass spectrometry artifacts that increase the number of m/z features per metabolite. The conditions in the ESI can lead to fragmentation because metabolites are subjected to high temperatures (150 °C to 400 °C) and electric potentials between 2000 V and 4000 V. While ESI sources are usually designed to minimize in-source fragmentation, it is also possible to promote in-source fragmentation such that MS1 spectra resemble MS2 spectra that were obtained by collision induced dissociation (Xue et al., 2020). Apart from in-source fragmentation, other modifications of metabolites in the ion-source includes the formation of adducts (e.g. with Na, K, ammonia, sulfate), gains or losses of functional groups by chemical reactions (methylation, phosphorylation), or formation of homo- and heterodimers. Even

self-cyclization has been observed for glutamate and glutamine (Purwaha et al., 2014).

A common approach to identify in-source modifications and improve m/z feature annotation is based on chromatographic peak shape correlation analysis (Guo et al., 2021; Schmid et al., 2021; Senan et al., 2019). This approach considers that m/z features from the same metabolite must have the same elution profile (Xu et al., 2015). Chromatographic peak shape correlation analysis is especially effective if it is combined with MS2 spectra (Schmid et al., 2021; Dührkop et al., 2021; Chen et al., 2021) or isotope labelled substrates (Hartl et al., 2020; Xu et al., 2015). Some recent molecular networking approaches (Schmid et al., 2021; Chen et al., 2021; Nothias et al., 2020) combine similarities of elution profiles and MS2 spectra to identify in-source modifications and to increase annotation confidence. In isotope labelling approaches, metabolites are labelled by feeding cells with ^{13}C -carbon or ^{15}N -nitrogen sources (Hartl et al., 2020; Xu et al., 2015), which changes the mass of all metabolites (N- or C- containing) but not their retention times. Analyzing the mass differences of m/z features with the same retention time can then improve annotation confidence and identification of in-source modifications or contaminants.

Because FI-MS lacks a chromatographic separation, it is not possible to detect in-source effects by chromatographic peak shape correlation analysis. Therefore, approaches to consider in-source effects in FI-MS are limited and currently based on extending the list of reference masses (Fuhrer et al., 2011; Stricker et al., 2021). This means that, instead of annotating m/z features only to (de-)protonated metabolites, they are also annotated to the most prevalent adducts and neutral losses or gains. However, this approach cannot identify complex sequential in-source modifications due to combinatorial explosion of the reference list. Moreover, it is difficult to unequivocally annotate m/z features to a single entry in a reference list, especially if they include a large number of metabolites and derivatives with the same mass.

Here, we used an experimental approach to identify in-source modifications of metabolites in FI-MS. Therefore, we spiked 160 metabolite standards individually into an *E. coli* extract and measured MS1 spectra by FI-MS. We then searched for m/z features that increased in a spike-in sample relative to all other spike-in samples. On average 68 m/z features increased per spike-in standard suggesting extensive in-source modifications. While some spike-in standards showed hundreds of significant m/z features that should all originate from a single metabolite standard, others showed only increases of the m/z feature that matched the (de-)protonated metabolite standard. We could explain 49 % of the significant m/z features by connecting them in networks that represent known

in-source reactions, adducts, isotope patterns, and in-source fragments.

2 Materials and Methods

Authentic metabolite standards were purchased from Merck KGaA (former Sigma-Aldrich, Germany). The standards were dissolved in water to a concentration of 1 mM if not stated otherwise (Supporting Information: Table A). Standards were then further diluted with acetonitrile and methanol to a final concentration of 10 μM in 40:40:20 acetonitrile:methanol:water. The 10 μM metabolite standards were then added to an *E. coli* metabolite extract to yield a final concentration of 1 μM . *E. coli* cultures were in a M9 minimal medium, which contained: 22 mM KH_2PO_4 , 42.2 mM Na_2HPO_4 , 11.3 mM $(\text{NH}_4)_2\text{SO}_4$, 8.56 mM NaCl, 100 μM $\text{CaCl}_2 \times 2 \text{H}_2\text{O}$, 1 mM $\text{MgSO}_4 \times 7 \text{H}_2\text{O}$, 60 μM FeCl_3 , 6.3 μM $\text{ZnSO}_4 \times 7 \text{H}_2\text{O}$, 7.6 μM $\text{CoCl}_2 \times 6 \text{H}_2\text{O}$, 7.1 μM , 7 μM $\text{CuCl}_2 \times 2 \text{H}_2\text{O}$, and $\text{MnSO}_4 \times 2 \text{H}_2\text{O}$.

2.1 Chemicals and materials

5 mL LB medium was inoculated with *E. coli* MG1655 from a cryo stock. After 6 - 7 h of cultivation at 37 °C, 10 μL of the culture was transferred to 5 mL M9 minimal medium with 5 g/L glucose. For ^{13}C -labelled extracts, uniformly labelled ^{13}C -glucose was used (# CLM-1396, Cambridge Isotope Laboratories Inc., USA). The M9 precultures were grown overnight at 37 °C and at 220 rpm shaking. 20 mL of M9 with 5 g/L ^{12}C - or ^{13}C -glucose was inoculated with the overnight culture to an optical density at 600 nm (OD) of 0.05. At an OD of 1, aliquots of 4 mL of the culture were vacuum-filtrated using 0.45 μm pore size filters (HVLP02500, Merck Millipore). The filters were transferred to -20 °C cold 40:40:20 acetonitrile:methanol:water for metabolite extraction. After at least 30 min at -20 °C, the metabolite extracts were centrifuged for 30 min at -9 °C and 4,000 rpm. The supernatant was stored at -80 °C.

2.2 Mass spectrometry

Samples were analyzed by FI-MS on an Agilent 6546 Series quadrupole time-of-flight mass spectrometer (Agilent Technologies, USA). The electrospray source was operated in negative and positive ionization mode. The mobile phase was 60:40 isopropanol:water buffered with 10 mM ammonium carbonate $(\text{NH}_4)_2\text{CO}_3$ and 0.04 % (v/v) ammonium hydroxide for both ionization modes, and the flow rate was 0.15 mL/min. For online mass

axis correction, 2-propanol (in the mobile phase) and HP-921 were used for negative mode and purine and HP-921 were used for positive mode. Mass spectra were recorded in profile mode from 50 to 1100 m/z with a frequency of 1.4 spectra/s for 0.5 min using 10 Ghz resolving power. Source temperature was set to 225 °C, with 1 L/min drying gas and a nebulizer pressure of 20 psi. Fragmentor, skimmer, and octupole voltages were set to 120 V, 65 V, and 650 V, respectively.

2.3 Data preprocessing

Raw files were converted into "mzXML" format by MSConvert (Chambers et al., 2012). Further data processing was performed using MATLAB version R2021a (The Mathworks, Inc., USA). For each sample, an average spectrum was calculated from the ten scans with the highest ion counts. The spectra were resampled to 10^6 data points to align m/z values of all samples. Ion peaks were picked with the "findpeaks" function of MATLAB, using a peak height and prominence cutoff of 1,000 units. Hierarchical clustering with a tolerance of 7.5 mDa was used to bin peaks. For each peak bin, we calculated a centroid m/z value from the individual peak m/z values. Peaks were annotated to metabolites using the centroid m/z value with a tolerance of 3 mDa. Z-scores were calculated from logarithmic mean signal intensities (triplicates). Z-scores above three were considered significant.

2.4 Calculation of mass differences

Mass differences between all significant features were calculated and combined for positive and negative ionization mode. Using the MATLAB function "histcounts", all mass differences were assigned to one of 5×10^5 bins, which accounted for a mass resolution of ca. 2 mDa. The total number of mass differences in each bin is the frequency of a mass difference. In the resulting neutral loss spectrum (x-axis is the mass difference and y-axis the frequency), peaks were picked with 3 mDa tolerance using the "findpeaks" function with prominence and height cutoffs of 10 units. 51 peaks that matched mass differences in the literature (Chen et al., 2021) were then used for further analysis (Supporting Information: Table B).

2.5 Construction of feature networks

Networks of significant m/z features were constructed for each spike-in sample in positive and negative ionization mode. Nodes are significant m/z features, and edges are

putative modifications like adducts or chemical reactions, or isotopes. Edges are drawn between each pair of nodes if the mass difference between them matches the mass differences in the list of 51 explained mass differences (Supporting Information: Table B). The fraction of explained features are calculated by counting all features that are either directly or indirectly connected to the (de-)protonated metabolite or fragment ion versus the total number of significant features for each spike-in standard. The feature networks were built with python v. 3.8.5 using the "networkx" toolbox. MS2 spectra were obtained from the Human Metabolome Database when experimental spectra were reported for a spike-in standard²⁷ (Supporting Information: Table C). The experimental MS2 spectra from HMDB contained data that was acquired by high- and low-resolution mass spectrometers. Therefore, we matched our significant m/z features to the MS2 spectra with a tolerance of 100 mDa.

3 Results

3.1 FI-MS with 160 authentic metabolite standards

We prepared 160 authentic standards of primary metabolites and spiked them individually into a metabolite extract from glucose-fed *E. coli* cells (Supporting Information: Table A). The 160 standards fall into six functional categories: amino acid metabolism, nucleotide metabolism, central metabolism, cofactor metabolism, antioxidants, and others. Each metabolite standard was added to the *E. coli* metabolite extract at a final concentration of 1 μM and was measured by FI-MS in both positive and negative ionization mode (three analytical replicates) (Figure 3.1a).

Out of 160 metabolite standards, 154 were annotated in their protonated or deprotonated form to an ion peak in the MS1 spectrum (Figure 3.1b). Six metabolites were not annotated, either due to low abundant ion peaks ($< 1,000$ counts: menadione, 3,4-dihydroxy-L-phenylalanine, tetra-hydrofolic acid, carbamoyl-P, and L-cysteine) or because the ion peak prominence was too low ($< 1,000$ counts: argininosuccinic acid).

Next, we inspected if the addition of a metabolite standard led to increases of the respective ion peak. For example, spike-in samples with ATP, GTP, CTP, and UTP showed increases of ion peaks that matched the protonated and deprotonated form of these metabolites (Figure 3.1c). Notably, increases of all nucleotides were consistent between three analytical replicates, showing that FI-MS is reproducible. Ion peaks of ATP, GTP, CTP, and UTP were also present in the other spike-in samples (black spectra in Figure

3.1c), but the corresponding ion intensities were often low and close to the baseline signal. These "near-baseline" ion peaks of ATP, GTP, CTP, and UTP were not present in a ^{13}C -labelled *E. coli* extract, thus confirming that these peaks originate from endogenous *E. coli* nucleotides (Supporting Information: Figure 3.6).

In 134 spike-in samples, the ion peaks that matched the (de-)protonated metabolites were significantly increased in either ionization mode (z -score > 3 , Figure 3.1d, Supporting Information: Table D). In negative ionization mode, 120 spike-in samples showed increased signals as deprotonated metabolites ([Metabolite-H]⁻). In positive ionization mode, 105 peaks increased as protonated metabolites ([Metabolite+H]⁺). In the following, we will refer to significantly changed ion peaks with a z -score > 3 as "significant features". 26 spike-in standards did not show a significant feature at the (de-)protonated ion peak. One explanation for this is that the metabolites have already a high concentration in the *E. coli* metabolite extract and that an addition of 1 μM does not lead to a strong increase with a z -score > 3 . For example, reduced glutathione is one of the most abundant metabolites in *E. coli* (Guder et al., 2017), and the addition of glutathione standard hardly increased its concentration in the spike-in sample (Supporting Information: Figure 3.7).

FI-MS was reproducible because the median relative standard deviation (RSD) between the three analytical replicates was below 5 % for signals from metabolite standards in negative and positive ionization mode (Figure 3.1e). Signals from endogenous metabolites, had a median RSD of 12.4 % in positive and 19.4 % negative mode (Supporting Information: Fig. 3.8).

These results suggested that FI-MS can detect concentration changes of chemically diverse metabolites, which are in the physiological range of intracellular metabolites (1 μM in the final sample corresponds to ca. 1 mM intracellularly). However, we noticed that many significant features did not match the (de-)protonated form of the metabolite in the spike-in sample (e.g. significant features that are off the diagonal in Figure 3.1d). Thus, we next investigated all significant features in all spike-in samples.

3.2 Single metabolites can produce extensive in-source derivatives

Most spike-in samples showed significant features (ion peaks with a z -score > 3) that matched the protonated or deprotonated metabolite (Figure 3.1d). However, most spike-in samples had more significant features than only the (de-)protonated metabolite standard (Figure 3.2a). On average, we found 68 significant features per spike-in sample, and 11 spike-in samples showed more than 100 significant features. The glycerol 3-

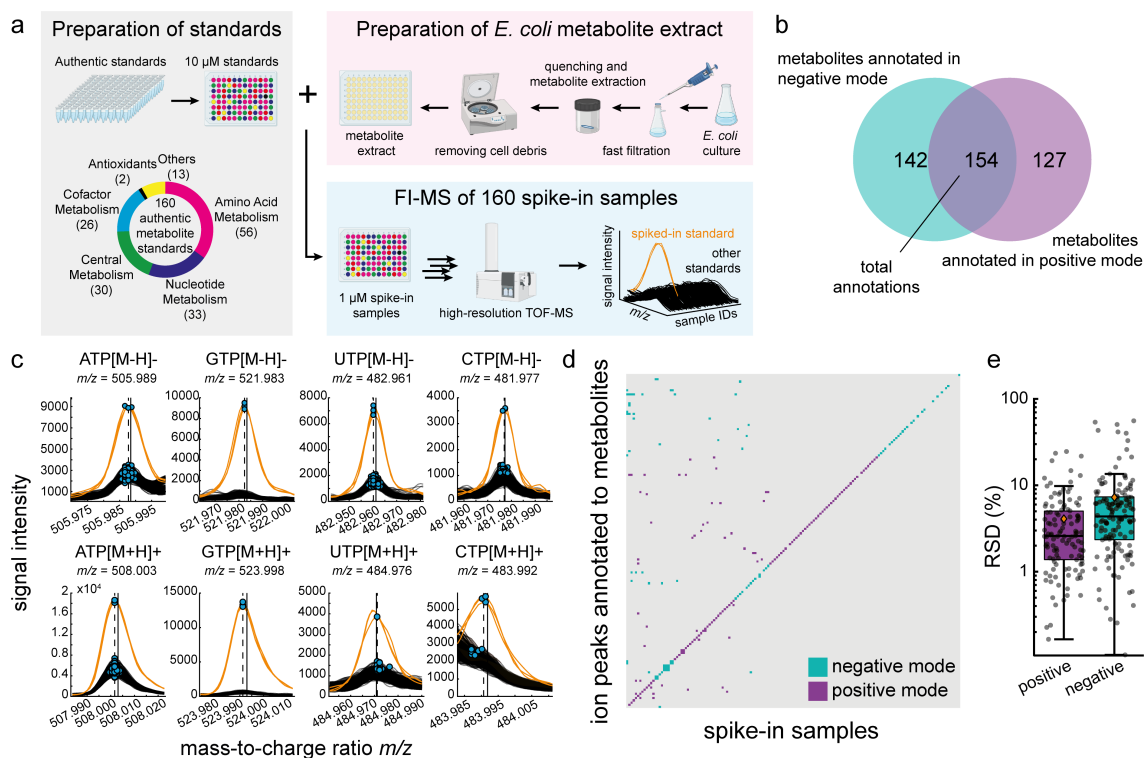


Figure 3.1: FI-MS with 160 authentic metabolite standards. a) 160 authentic metabolite standards were spiked into an *E. coli* metabolite extract at a final concentration of 1 μM . Spike-in samples were measured by FI-MS in analytical triplicates in positive and negative ionization mode. b) Number of ion peaks (*m/z* features) that were annotated to the protonated (positive mode) and deprotonated (negative mode) form of the 160 metabolite standards. c) Ion peaks that are annotated to four nucleotides (ATP, CTP, GTP, UTP) in positive and negative ionization mode. The spike-in sample that contains the respective nucleotide is indicated in orange, the other 159 spike-in samples are black. Blue dots indicate *m/z* features in single samples, and vertical dotted lines are merged and centroided *m/z* features. Vertical solid lines indicate the monoisotopic masses of the nucleotides plus/minus the mass of a proton. d) The binary heatmap shows increases of *m/z* features that are annotated to metabolite standards in the spike-in samples. Significant increases of *m/z* features (*z*-score > 3) are shown in blue (negative mode) and purple (positive mode). Columns are the spike-in samples and rows the respective *m/z* features. e) Boxplots show the relative standard deviation (RSD) of metabolite standards. Black dots are the RSD for each spike-in metabolite ($n = 3$). Orange diamonds are the means.

phosphate (G3P) spike-in sample had the highest number of significant features (5,464, Figure 3.2a). Across all 160 spike-in samples, FI-MS in positive mode showed 10,206 significant features and 807 in negative mode (Supporting Information: Table E). The significant features were distributed over the entire mass spectrum and occurred even in the higher mass range of 800 – 1,000 m/z (Figure 3.2b).

To understand the origin of these significant features, we first calculated the mass differences ($\Delta m/z$) between all pairs of significant features in a single spike-in sample (Figure 3.2c). Several mass differences ($\Delta m/z$) occurred frequently across the 160 spike-in samples, thus indicating common in-source effects like neutral losses, adduct formation, and chemical reactions that are prevalent for many different compounds (Figure 3.2d and 3.2e). 51 mass differences that appeared more than ten times matched known in-source effects and isotope pattern reported in the literature (Chen et al., 2021, Supporting Information: Table B). Out of these 51 known mass differences, 23 were chemical reactions, 26 were adducts, and 2 were natural isotopes (^{13}C and ^{18}CO). The 23 chemical reactions account for 61.5 % of the frequent mass differences (> 10 times in all samples), the 22 adducts for 31.5 %, and the isotopomers containing ^{13}C or ^{18}CO for 7 % (Figure 3.2f). For example, the 21.982 Da mass difference of a Na-H neutral loss occurred in total 699 times and was among the most frequent ones (Figure 3.2e). The ten most frequent mass differences occurred more than 1,000 times across all 160 spike-in samples, and eight of them could be explained with the mass differences in the literature (Figure 3.2d).

We then wondered whether certain mass differences occurred more frequently for metabolites of a specific functional category than for metabolites of other categories. For example, only 21 % of the 160 standards were metabolites from nucleotide metabolism. Yet, they accounted for 45 % of all explainable mass differences (Figure 3.2g). This indicated that metabolites from nucleotide metabolism were more susceptible to modifications than metabolites in other categories. In contrast, 35 % of the 160 metabolite standards were part of amino acid metabolism but they covered only 23 % of the explainable mass differences indicating that metabolites from amino acid biosynthesis were less prone to modifications in our reference list of mass differences than the other metabolites. Metabolites from central metabolism as well as co-factor biosynthesis accounted for 19 % and 16 % of the standard library, respectively, and they explained a similar fraction of mass differences (15 % and 13 %).

Since some metabolite categories were more often modified than others, we looked into individual mass differences and examined whether certain mass differences occurred preferably for specific categories (Figure 3.2h). Indeed, the data indicated that

individual mass differences were more frequent for some categories than for others. For example, the $C_6H_8O_6$ neutral loss occurred exclusively in metabolites from nucleotide metabolism. Similarly, many other modifications, including modifications with O, HNO_3 , $C_2H_4O_2$, or SO_3 , occurred more frequently with metabolites from nucleotide metabolism (Supporting Information: Table F).

3.3 A network approach explains significant m/z features in FI-MS spectra

We expected that the significant features can be linked to the (de-)protonated spike-in metabolite by single and multiple modification steps. To test this, we created a network for each spike-in sample, in which nodes represent all significant features. Then, we drew an edge between two nodes if the mass difference $\Delta m/z$ between them matched one of the 51 frequent mass differences (the mass differences identified above, see Figure 3.2h). Thus, edges represent in-source effects and nodes significant features (see schematic in Figure 3.3a).

The thus derived networks connected on average 43 % of the significant features in a spike-in sample to the m/z feature of the respective (de-)protonated metabolite (Figure 3.3a). Thus, 43 % of the significant features can be linked to a single metabolite and therefore, are explained by the 51 frequent mass differences. Only 20 % of the significant features were directly linked to the m/z feature of the (de-)protonated metabolite (Figure 3.3a, Supporting Information: Table G). This shows that single in-source modifications account only for half of the significant features and that sequential modifications are frequent.

For example, the glucosamine 6-phosphate (Ga6p) spike-in standard showed five significant features, which are all directly or indirectly connected to the m/z feature that matches protonated Ga6p (Figure 3.3b). Two significant features are directly connected to the protonated Ga6p mass, and they are likely a water loss (H_2O) and a sodium adduct (Na-H). Two other significant features ($m/z = 224.031$ and $m/z = 242.042$) were two steps away from protonated Ga6P, and they were explained by a double loss of water and a NaH_2PO_4 adduction to the sodium adduct. Thus, drawing edges in an unbiased way between all pairs of nodes resulted in a network that explained all significant features of the Ga6p spike-in sample.

In many networks, the nodes (significant m/z features) were connected to the (de-)protonated spike-in metabolite by different series of sequential modifications. One spe-

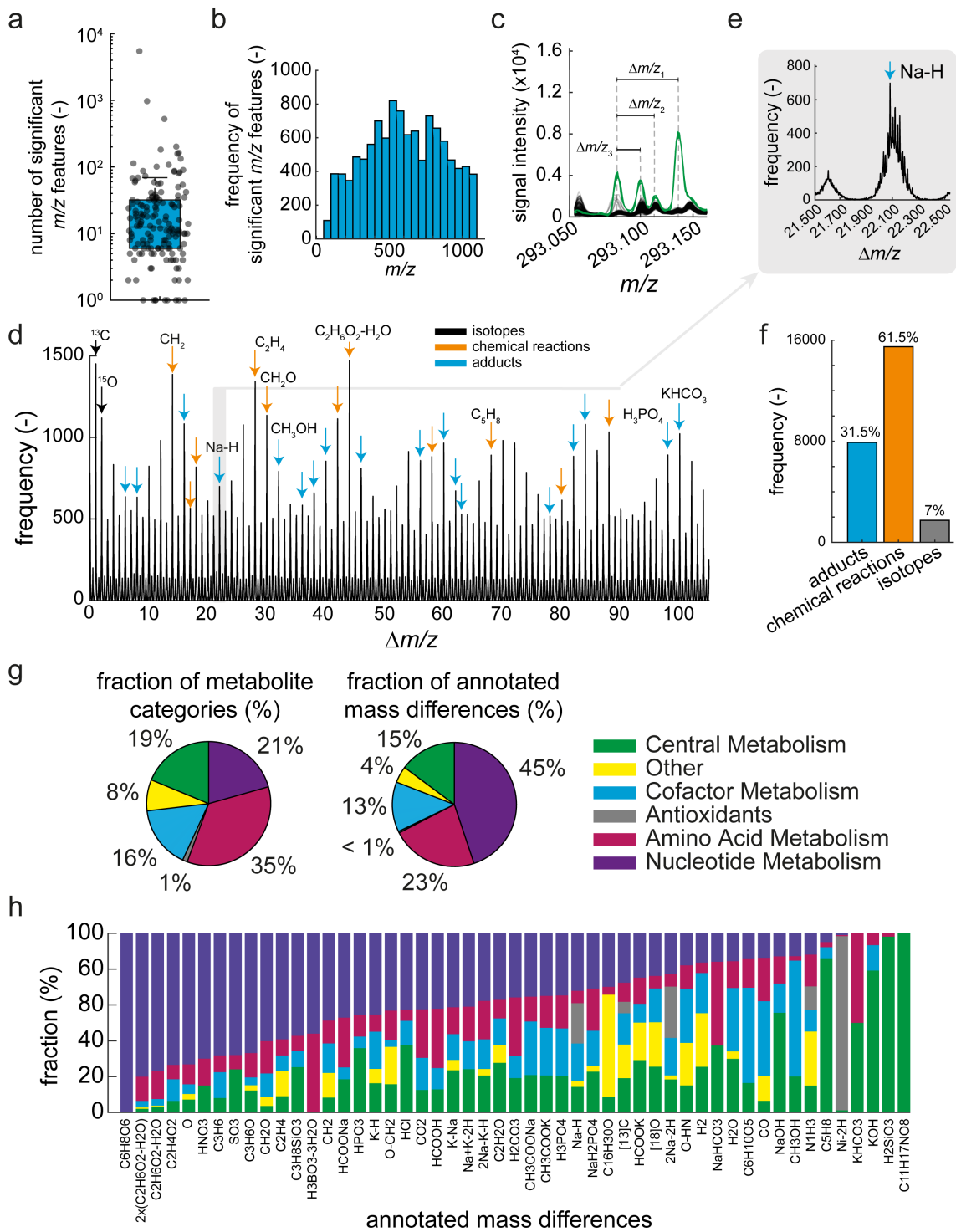


Figure 3.2: The figure caption is on the next page.

Figure 3.2: Systematic analysis of all m/z features that increase in spike-in samples. a) Number of m/z features with a z-score > 3 (significant m/z features) in each of the 160 spike-in samples (grey dots). The upper and lower edges of the box in the boxplot indicate the 25 % and 75 % percentiles, and the line is the median. b) Histogram showing the distribution of significant m/z features over the MS1 spectrum. c) Example of significant m/z features in the MS1 spectrum (100 mDa window) of the spike-in sample with glycerol 3-phosphate. Green lines are the glycerol 3-phosphate spike-in samples ($n = 3$), and black lines are the other spike-in samples. Rulers indicate the mass differences between two m/z features in the spectra. d) $\Delta m/z$ spectrum based on the pairwise mass differences between all significant m/z features in all 160 spike-in samples (shown is the $\Delta m/z$ range between 0 Da and 110 Da). The peak height corresponds to the frequency of a $\Delta m/z$ value. Arrows indicate $\Delta m/z$ peaks that match mass differences of known isotopes, chemical reactions or adducts (Supporting Information: Table B). e) Example of the $\Delta m/z$ peak that matches the sodium adduct [Na-H]. f) Fraction of $\Delta m/z$ peaks that match known isotopes, chemical reactions or adducts. g) The left pie chart shows the fraction of metabolite categories across the 160 standards. The right pie chart shows the fraction of annotated mass differences for each metabolite category. h) Stacked bar plot showing the relationship between the functional categories of the spiked-in metabolites and the annotated mass differences. The fraction indicates the ratio between the number of spike-in metabolites of a specific category, in which the mass difference occurred, and the total number of samples, in which the mass difference occurred. The spike-in samples of glycerol 3-phosphate and fumarate were left out.

cific series of sequential modifications could be an initial modification by Na-K that is followed by a second modification like Na-H. In some cases, different series of sequential modifications have very similar net mass changes and can explain the same significant feature. One example that illustrates this phenomenon is the feature network of 2,6-diaminopimelic acid (DAP) in positive ionization mode. The DAP spike-in sample showed six significant features, which were all connected with the m/z feature of protonated DAP (Figure 3.3c).

The network approach connected on average 43 % of all significant features of a spike-in sample to the (de-)protonated metabolite standard. Yet, some significant features had no connection to others or they formed sub-networks with no connection to the (de-)protonated metabolite (see schematic in Figure 3.3a). Therefore, we next examined whether sub-networks and isolated features are caused by in-source fragmentation, which can lead to similar effects as collision induced dissociation in tandem mass spectrometry (Xue et al., 2020).

3.4 MS2 information identifies significant features that are in-source fragments

To identify significant features that originate from in-source fragmentation of the metabolite standard, we used information about MS2 spectra in the human metabolome database (HMDB, Wishart et al., 2022). HMDB listed experimental MS2 spectra for 152 out of 160 metabolite standards (Figure 3.4a, Supporting Information: Table C). 103 standards had at least one significant feature that matched an ion peak in the MS2 spectrum (Supporting Information: Table H). On average, each spike-in sample had 3.4 in-source fragments indicating a substantial number of in-source fragmentation events during FI-MS (Figure 3.4a). In total, 551 MS2 features matched the significant m/z features. As expected, in-source fragments had masses in the lower m/z range, between 50 and 500 m/z (Figure 3.4b). One example of an in-source fragment is hypoxanthine, which is formed by fragmenting inosine and inosine monophosphate (IMP). Consequently, the ion peak of hypoxanthine increased in the IMP and inosine spike-in samples (Figure 3.4c).

Then, we tested whether fragments were present in sub-networks without a link to the metabolite standard. For example, the L-citrulline spike-in sample showed 8 significant features, 6 of which were connected with the m/z feature of the (de-)protonated metabolite (Figure 3.4d). Two significant features were isolated nodes with no connection, but they were in the MS2 spectrum of citrulline ($m/z = 177.111$ and $m/z = 176.112$).

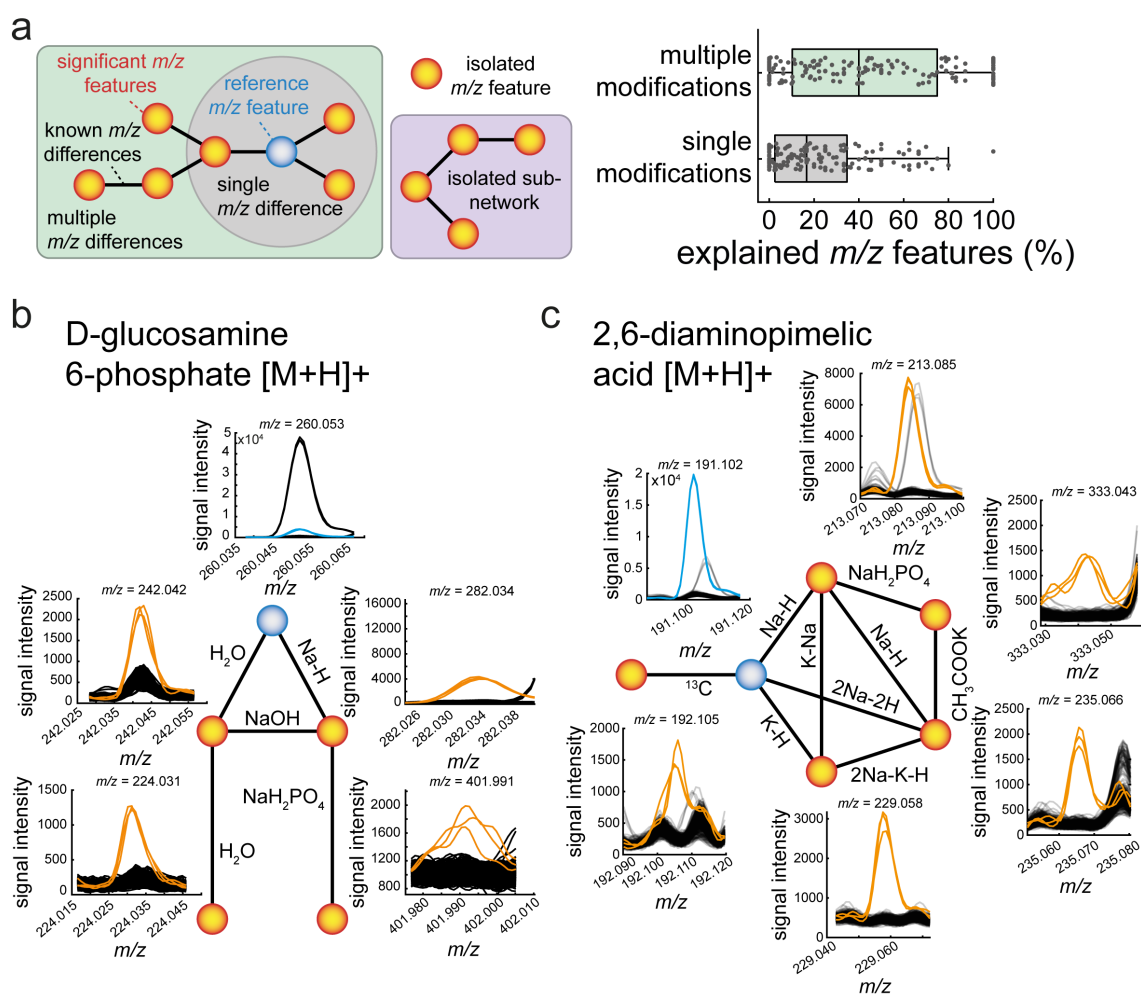


Figure 3.3: Networks of significant m/z features. a) Concept figure showing the structure of a m/z feature network: orange nodes are significant m/z features. The blue node is a significant m/z feature that is annotated to the (de-)protonated metabolite standard (reference m/z feature). Nodes are connected by edges that correspond to one of 51 known m/z differences. The grey circle indicates nodes that are directly connected to the reference m/z feature by a single m/z difference. The green box indicates a network, in which all nodes are connected with the reference node including multiple sequential combinations of m/z differences. Isolated sub-networks are not connected to the reference node (purple box). Isolated m/z features are not connected to any other node. The boxplots show the fraction of significant m/z features that are connected to the reference m/z feature by a single m/z difference or by multiple sequential combinations of m/z differences. b) Example of the feature network of the spike-in sample with D-glucosamine 6-phosphate (Ga6p). c) Same as b) for the diaminopimelic acid (DAP) spike-in sample in positive ionization mode.

The MS2 spectrum included another two m/z features that were already linked to the metabolite standard ($m/z = 113.070$ and $m/z = 159.076$), thus indicating that collision induced dissociation (CID) produces some of the 51 in-source modifications. In case of L-citrulline, these were a neutral loss of NH_3 and a neutral loss of a HCOOH group. To confirm that these losses occur by CID, we predicted fragment structures of L-citrulline by CFM-ID (F. Wang et al., 2021). Indeed, CFM-ID predicted the fragment structures that matched the masses of the two m/z features (113.070 and 159.076) and confirmed the neutral losses of NH_3 and HCOOH (Figure 3.4d).

In total, adding in source fragments to our networks explained another 6 % of significant m/z features. Thus, on average, 49 % of all significantly changed m/z features were connected with the correct metabolite standard either by known in-source modifications or by in-source fragmentation (Figure 3.4e, Supporting Information: Table G).

3.5 Misannotation of in-source derivatives to metabolites

A single metabolite can produce multiple significant features, and we wondered how many of these significant features were misannotated to a metabolite that was not spiked into the sample. To determine how many significant features were misannotated, we used a reference list of 961 *E. coli* metabolites from the genome-scale metabolic model iML1515 (Monk et al., 2017). Since FI-MS cannot resolve isomers, they were considered as a single metabolite.

In 54 % of our standards (87/160), at least one significant feature was falsely annotated to a metabolite (Figure 3.5). 18 standards had more than two misannotations. Overall, 64 % of the misannotations were in positive ionization mode (131 in total) and 37 % were in negative ionization mode (75 in total) (Figure 3.5, Supporting Information: Table I). This means that biological screens with FI-MS are prone to misannotations if only (de-)protonated masses are considered. Based on our results, an estimate is that one (true) increase of a single metabolite will cause one (false) increase of an ion peak that is misannotated to another metabolite.

4 Discussion

FI-MS methods have been used for metabolome analyses in various studies (Fuhrer et al., 2011; Beckmann et al., 2008; Sarvin et al., 2020; Link et al., 2015; Fuhrer et al., 2017; Anglada-Girotto et al., 2022; Holbrook-Smith et al., 2022; Dubuis et al., 2018; Rathahao-

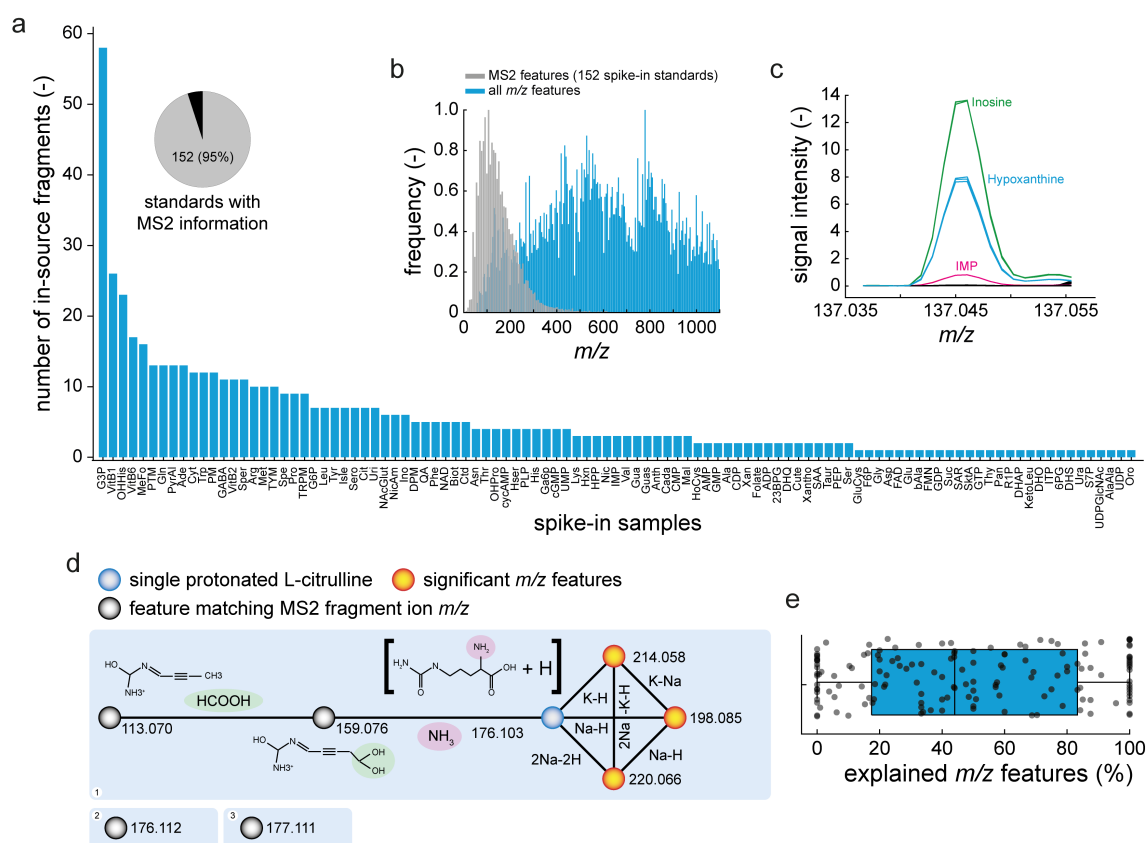


Figure 3.4: Identification of in-source fragments by MS2 spectra of metabolite standards. a) The pie chart shows the fraction of metabolite standards with MS2 information in the human metabolome database (HMDB). The bar plot shows the number of MS2 fragment masses that match the features of individual spike-in samples. b) Histogram showing the distribution of the m/z features. Blue are all significant m/z features from the spike-in samples. Grey are MS2 fragment masses. c) Example MS1 spectrum at the mass of hypoxanthine. The purple line is the IMP spike-in sample, the blue line is the hypoxanthine spike-in sample, and the green line is the inosine spike-in sample. d) Example network of L-citrulline. Nodes are features and edges are explained mass differences. Black nodes are features that matched MS2 fragment masses from HMDB. The blue node is the protonated mass of citrulline. The orange nodes are other features. For the features at $m/z = 113.070$ and $m/z = 159.076$, structures were predicted by CFM-ID. e) Boxplot showing the fraction of explained features for all spike-in samples. Each black point corresponds to a spike-in sample and shows the explained fraction of features. Upper and lower box edges indicate the 25 % and 75 % percentiles. The whiskers indicate the furthest point, at which samples were not considered as outliers. The black line indicates the median.

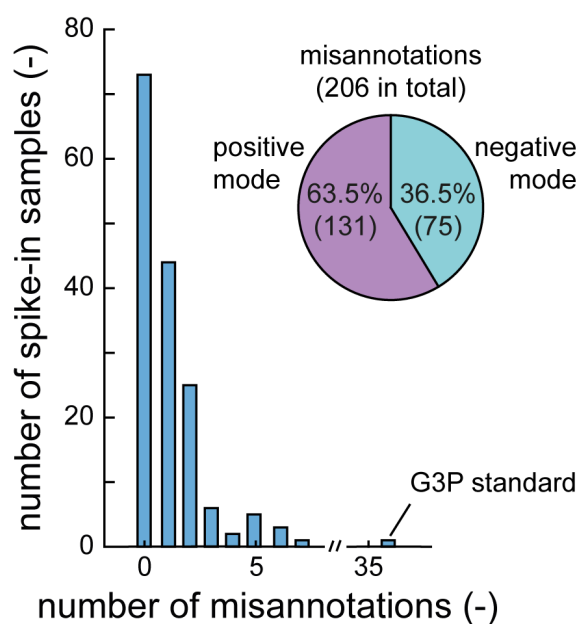


Figure 3.5: Misannotation of significant features to metabolites. The histogram shows the number of significant features that are misannotated to a metabolite. The pie chart shows the fraction of misannotations for positive and negative ionization mode.

Paris et al., 2019). Their advantages are fast analysis times (10 to 30 seconds per sample) and a high coverage of metabolites (often more than 1,000 putatively annotated metabolites (Fuhrer et al., 2017; Anglada-Girotto et al., 2022)). Disadvantages, however, are low confidence levels of metabolite annotation and a high susceptibility to matrix effects due to the lack of chromatographic separation. Here, we confirmed the broad metabolite coverage of FI-MS, which detected increases of 134 out of 160 metabolite standards based on (de-)protonated ion peaks in the MS1 spectrum. However, we also observed pervasive in-source modifications of metabolites. These in-source modifications lead to multiple ion peaks per metabolite in the MS1 spectrum and, in the worst case, to false positive hits in FI-MS analyses of biological samples. By systematically analyzing FI-MS data from 160 spike-in standards, we found that a single metabolite produces on average 68 significant features and that, in extreme cases, more than 1,000 significant features originate from only one metabolite. This observation matches previous LC-MS based studies where the majority of m/z features were attributed to in-source modifications and only few (3 – 5 %) m/z features were unique metabolites (Hartl et al., 2020; L. Wang et al., 2019; Mahieu and Patti, 2017; Mahieu et al., 2016).

Chromatographic peak shape correlation of m/z features can identify such confounding effects in LC-based methods (Schmid et al., 2021; Xu et al., 2015), but they are dif-

difficult to detect with FI-MS methods. Here, we used an experimental approach and examined significant m/z features in metabolite standards, which are most likely in-source derivatives. Connecting these features via 51 mass differences of neutral losses, adducts, and isotopes described in the literature (Chen et al., 2021) resulted in networks that explained the origin of 43 % of the significant features. MS2 spectra of the metabolite standards provided additional information about in-source fragmentation and explained another 6 % of the significant features.

Taken together, we found that FI-MS of single metabolites produces complex MS1 spectra, but they are explainable by known in-source modifications. The jury is out if in-source modifications are a bug or a feature for FI-MS data analysis: they may complicate or improve metabolite annotation. Therefore, the future challenge is to use FI-MS spectra of single metabolites to deconvolute FI-MS spectra from biological samples and, thereby, increase confidence of metabolite annotation. A first step is the construction of FI-MS databases with MS1 spectra of single metabolite standards to map in-source modifications across thousands of compounds. This is especially important since we found that the type of in-source modification depends on the metabolite classes (Figure 3.4d). Here, we provided a starting point with in-source modifications of 160 metabolites (Supporting Information: Table E) and mining MS1 spectra in existing databases like GNPS (M. Wang et al., 2016) and Metlin (Guijas et al., 2018) may provide additional reference data.

5 Supporting Information

Supporting Table (.xlsx) with additional information and data. Supporting Figures (.docx):

Figure S1: Ion peaks of nucleotides in negative ionization mode from fully labelled ^{12}C - and ^{13}C -labelled *E. coli* extracts.

Figure S2: Ion peak of reduced glutathione in negative ionization mode.

Figure S3: Boxplot showing the relative standard deviation (RSD) of the endogenous metabolites measured by FI-MS in positive and negative ionization mode.

6 Author Information

6.1 Corresponding Author

E-mail: hannes.link@uni-tuebingen.de

6.2 ORCID

Niklas Farke: 0000-0003-0106-5271

Thorben Schramm: 0000-0001-5013-6705

Andreas Verhülsdonk: 0000-0001-9187-5117

Hannes Link: 0000-0002-6677-555X

6.3 Author Contributions

^ψN. F. and T. S. contributed equally. N. F., T. S., and A. V. performed experiments. N. F. and T. S. analyzed data. N. F., T. S., and H. L. wrote the manuscript and designed the study. H. L. directed the project. All authors have given approval to the final version of the manuscript.

6.4 Notes

The authors declare no competing financial interest.

7 Acknowledgement

This work was supported by the ERC starting grant 715650.

8 Supporting Figures

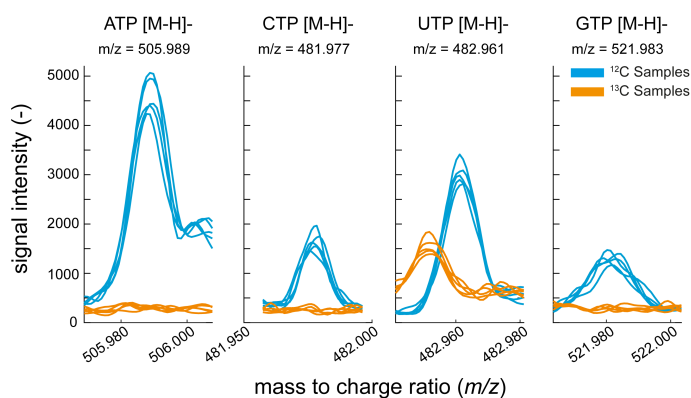


Figure 3.6: (also Figure S1.) Ion peaks that are annotated to four nucleotides (ATP, CTP, UTP, GTP) in negative ionization mode. ^{12}C -labelled *E. coli* samples are blue lines measured in five technical replicates. ^{13}C -labelled *E. coli* samples are orange lines measured in five technical replicates. ATP, adenosine triphosphate; CTP, cytidine triphosphate; UTP, uridine triphosphate; GTP, guanosine triphosphate.

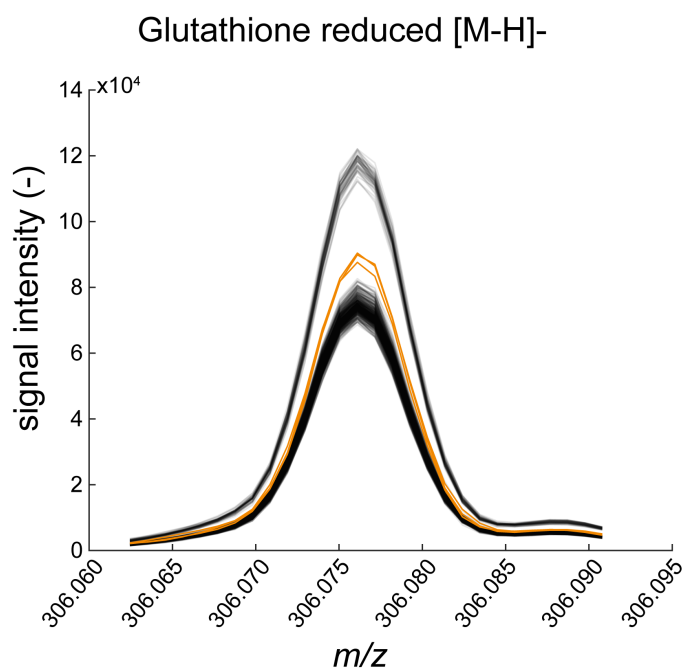


Figure 3.7: (also Figure S2.) Ion peak of reduced glutathione (deprotonated) in negative ionization mode. The spike-in sample that contains reduced glutathione is indicated in orange, the other 159 spike-in samples are black.

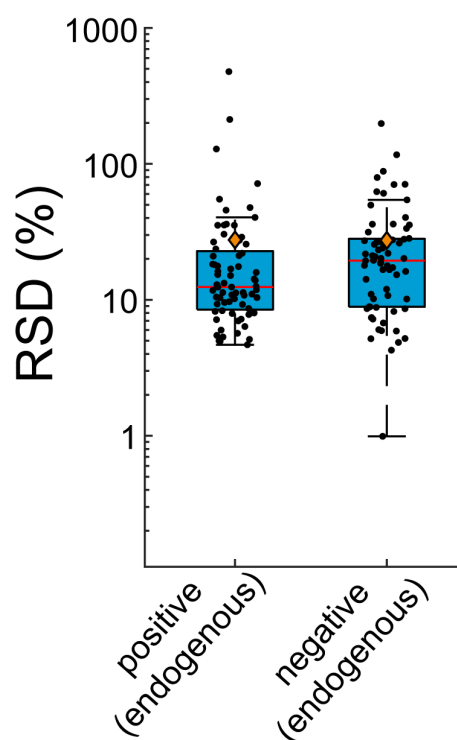


Figure 3.8: (also Figure S3.) Boxplot showing the relative standard deviation (RSD) of the endogenous metabolites measured by FI-MS in positive and negative ionization mode. Black dots are the RSD for each endogenous metabolite. Upper and lower box edges indicate the 25 % and 75 % percentiles. The whiskers indicate the furthest point, at which samples were not considered as outliers. The red line indicates the median. Orange diamonds are the means.

References

- Fuhrer, T., Heer, D., Begemann, B., & Zamboni, N. (2011). High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection–time-of-flight mass spectrometry. *Analytical Chemistry*, 83(18), 7074–7080. <https://doi.org/10.1021/ac201267k> (cited on pages 75, 76, 88).
- Beckmann, M., Parker, D., Enot, D. P., Duval, E., & Draper, J. (2008). High-throughput, nontargeted metabolite fingerprinting using nominal mass flow injection electrospray mass spectrometry. *Nature Protocols*, 3(3), 486–504. <https://doi.org/10.1038/nprot.2007.500> (cited on pages 75, 88).
- Reiter, A., Herbst, L., Wiechert, W., & Oldiges, M. (2021). Need for speed: Evaluation of dilute and shoot-mass spectrometry for accelerated metabolic phenotyping in bioprocess development. *Analytical and Bioanalytical Chemistry*, 413(12), 3253–3268. <https://doi.org/10.1007/s00216-021-03261-3> (cited on page 75).
- Sarvin, B., Lagziel, S., Sarvin, N., Mukha, D., Kumar, P., Aizenshtein, E., & Shlomi, T. (2020). Fast and sensitive flow-injection mass spectrometry metabolomics by analyzing sample-specific ion distributions. *Nature Communications*, 11(1), 3186. <https://doi.org/10.1038/s41467-020-17026-6> (cited on pages 75, 88).
- Link, H., Fuhrer, T., Gerosa, L., Zamboni, N., & Sauer, U. (2015). Real-time metabolome profiling of the metabolic switch between starvation and growth. *Nature Methods*, 12(11), 1091–1097. <https://doi.org/10.1038/nmeth.3584> (cited on pages 75, 88).
- Fuhrer, T., Zampieri, M., Sévin, D. C., Sauer, U., & Zamboni, N. (2017). Genomewide landscape of gene–metabolome associations in *Escherichia coli*. *Molecular Systems Biology*, 13(1), 907. <https://doi.org/10.15252/msb.20167150> (cited on pages 75, 88, 90).
- Anglada-Girotto, M., Handschin, G., Ortmayr, K., Campos, A. I., Gillet, L., Manfredi, P., Mulholland, C. V., Berney, M., Jenal, U., Picotti, P., & Zampieri, M. (2022). Combining CRISPRi and metabolomics for functional annotation of compound libraries. *Nature Chemical Biology*, 18(5), 482–491. <https://doi.org/10.1038/s41589-022-00970-3> (cited on pages 75, 88, 90).
- Holbrook-Smith, D., Durot, S., & Sauer, U. (2022). High-throughput metabolomics predicts drug–target relationships for eukaryotic proteins. *Molecular Systems Biology*, 18(2). <https://doi.org/10.15252/msb.202110767> (cited on pages 75, 88).
- Dubuis, S., Ortmayr, K., & Zampieri, M. (2018). A framework for large-scale metabolome drug profiling links coenzyme a metabolism to the toxicity of anti-cancer drug dichloroacetate. *Communications Biology*, 1(1), 101. <https://doi.org/10.1038/s42003-018-0111-x> (cited on pages 75, 88).
- Rathahao-Paris, E., Alves, S., Boussaid, N., Picard-Hagen, N., Gayrard, V., Toutain, P.-L., Tabet, J.-C., Rutledge, D. N., & Paris, A. (2019). Evaluation and validation of an analytical approach for high-throughput metabolomic fingerprinting using direct introduction–high-resolution mass spectrometry: Applicability to classification of urine of scrapie-infected ewes. *European Journal of Mass Spectrometry*, 25(2), 251–258. <https://doi.org/10.1177/1469066718806450> (cited on pages 75, 88).
- Hartl, J., Kiefer, P., Kaczmarczyk, A., Mittelviehhaus, M., Meyer, F., Vonderach, T., Hattendorf, B., Jenal, U., & Vorholt, J. A. (2020). Untargeted metabolomics links glutathione to bacterial cell cycle progression. *Nature Metabolism*, 2(2), 153–166. <https://doi.org/10.1038/s42255-019-0166-0> (cited on pages 75, 76, 90).

- Wang, L., Xing, X., Chen, L., Yang, L., Su, X., Rabitz, H., Lu, W., & Rabinowitz, J. D. (2019). Peak annotation and verification engine for untargeted LC-MS metabolomics. *Analytical Chemistry*, *91*(3), 1838–1846. <https://doi.org/10.1021/acs.analchem.8b03132> (cited on pages 75, 90).
- Kachman, M., Habra, H., Duren, W., Wigginton, J., Sajjakulnukit, P., Michailidis, G., Burant, C., & Karnovsky, A. (2020). Deep annotation of untargeted LC-MS metabolomics data with *Binner* (J. Kelso, Ed.). *Bioinformatics*, *36*(6), 1801–1806. <https://doi.org/10.1093/bioinformatics/btz798> (cited on page 75).
- Sindelar, M., & Patti, G. J. (2020). Chemical discovery in the era of metabolomics. *Journal of the American Chemical Society*, *142*(20), 9097–9105. <https://doi.org/10.1021/jacs.9b13198> (cited on page 75).
- Mahieu, N. G., & Patti, G. J. (2017). Systems-level annotation of a metabolomics data set reduces 25 000 features to fewer than 1000 unique metabolites. *Analytical Chemistry*, *89*(19), 10397–10406. <https://doi.org/10.1021/acs.analchem.7b02380> (cited on pages 75, 90).
- Xue, J., Domingo-Almenara, X., Guijas, C., Palermo, A., Rinschen, M. M., Isbell, J., Benton, H. P., & Siuzdak, G. (2020). Enhanced in-source fragmentation annotation enables novel data independent acquisition and autonomous METLIN molecular identification. *Analytical Chemistry*, *92*(8), 6051–6059. <https://doi.org/10.1021/acs.analchem.0c00409> (cited on pages 75, 86).
- Purwaha, P., Silva, L. P., Hawke, D. H., Weinstein, J. N., & Lorenzi, P. L. (2014). An artifact in LC-MS/MS measurement of glutamine and glutamic acid: In-source cyclization to pyroglutamic acid. *Analytical Chemistry*, *86*(12), 5633–5637. <https://doi.org/10.1021/ac501451v> (cited on page 76).
- Guo, J., Shen, S., Xing, S., Yu, H., & Huan, T. (2021). ISFrag: De novo recognition of in-source fragments for liquid chromatography–mass spectrometry data. *Analytical Chemistry*, *93*(29), 10243–10250. <https://doi.org/10.1021/acs.analchem.1c01644> (cited on page 76).
- Schmid, R., Petras, D., Nothias, L.-F., Wang, M., Aron, A. T., Jagels, A., Tsugawa, H., Rainer, J., Garcia-Aloy, M., Dührkop, K., Korf, A., Pluskal, T., Kameník, Z., Jarmusch, A. K., Caraballo-Rodríguez, A. M., Weldon, K. C., Nothias-Esposito, M., Aksenov, A. A., Bauermeister, A., ... Dorrestein, P. C. (2021). Ion identity molecular networking for mass spectrometry-based metabolomics in the GNPS environment. *Nature Communications*, *12*(1), 3832. <https://doi.org/10.1038/s41467-021-23953-9> (cited on pages 76, 90).
- Senan, O., Aguilar-Mogas, A., Navarro, M., Capellades, J., Noon, L., Burks, D., Yanes, O., Guimerà, R., & Sales-Pardo, M. (2019). CliqueMS: A computational tool for annotating in-source metabolite ions from LC-MS untargeted metabolomics data based on a coelution similarity network. *Bioinformatics*, *35*(20), 4089–4097. <https://doi.org/10.1093/bioinformatics/btz207> (cited on page 76).
- Xu, Y.-F., Lu, W., & Rabinowitz, J. D. (2015). Avoiding misannotation of in-source fragmentation products as cellular metabolites in liquid chromatography–mass spectrometry-based metabolomics. *Analytical Chemistry*, *87*(4), 2273–2281. <https://doi.org/10.1021/ac504118y> (cited on pages 76, 90).
- Dührkop, K., Nothias, L.-F., Fleischauer, M., Reher, R., Ludwig, M., Hoffmann, M. A., Petras, D., Gerwick, W. H., Rousu, J., Dorrestein, P. C., & Böcker, S. (2021). Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. *Nature Biotechnology*, *39*(4), 462–471. <https://doi.org/10.1038/s41587-020-0740-8> (cited on page 76).
- Chen, L., Lu, W., Wang, L., Xing, X., Chen, Z., Teng, X., Zeng, X., Muscarella, A. D., Shen, Y., Cowan, A., McReynolds, M. R., Kennedy, B. J., Lato, A. M., Campagna, S. R., Singh, M., & Rabinowitz, J. D. (2021). Metabolite discovery through global annotation of untargeted metabolomics data. *Nature*

- Methods*, 18(11), 1377–1385. <https://doi.org/10.1038/s41592-021-01303-3> (cited on pages 76, 78, 82, 91).
- Nothias, L.-F., Petras, D., Schmid, R., Dührkop, K., Rainer, J., Sarvepalli, A., Protsyuk, I., Ernst, M., Tsugawa, H., Fleischauer, M., Aicheler, F., Aksenov, A. A., Alka, O., Allard, P.-M., Barsch, A., Cachat, X., Caraballo-Rodriguez, A. M., Da Silva, R. R., Dang, T., ... Dorrestein, P. C. (2020). Feature-based molecular networking in the GNPS analysis environment. *Nature Methods*, 17(9), 905–908. <https://doi.org/10.1038/s41592-020-0933-6> (cited on page 76).
- Stricker, T., Bonner, R., Lisacek, F., & Hopfgartner, G. (2021). Adduct annotation in liquid chromatography/high-resolution mass spectrometry to enhance compound identification. *Analytical and Bioanalytical Chemistry*, 413(2), 503–517. <https://doi.org/10.1007/s00216-020-03019-3> (cited on page 76).
- Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M.-Y., Paulse, C., Creasy, D., ... Mallick, P. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, 30(10), 918–920. <https://doi.org/10.1038/nbt.2377> (cited on page 78).
- Guder, J. C., Schramm, T., Sander, T., & Link, H. (2017). Time-optimized isotope ratio LC–MS/MS for high-throughput quantification of primary metabolites. *Analytical Chemistry*, 89(3), 1624–1631. <https://doi.org/10.1021/acs.analchem.6b03731> (cited on page 80).
- Wishart, D. S., Guo, A., Oler, E., Wang, F., Anjum, A., Peters, H., Dizon, R., Sayeeda, Z., Tian, S., Lee, B. L., Berjanskii, M., Mah, R., Yamamoto, M., Jovel, J., Torres-Calzada, C., Hiebert-Giesbrecht, M., Lui, V. W., Varshavi, D., Varshavi, D., ... Gautam, V. (2022). HMDB 5.0: The human metabolome database for 2022. *Nucleic Acids Research*, 50, D622–D631. <https://doi.org/10.1093/nar/gkab1062> (cited on page 86).
- Wang, F., Liigand, J., Tian, S., Arndt, D., Greiner, R., & Wishart, D. S. (2021). CFM-ID 4.0: More accurate ESI-MS/MS spectral prediction and compound identification. *Analytical Chemistry*, 93(34), 11692–11700. <https://doi.org/10.1021/acs.analchem.1c01465> (cited on page 88).
- Monk, J. M., Lloyd, C. J., Brunk, E., Mih, N., Sastry, A., King, Z., Takeuchi, R., Nomura, W., Zhang, Z., Mori, H., Feist, A. M., & Palsson, B. O. (2017). iML1515, a knowledgebase that computes escherichia coli traits. *Nature Biotechnology*, 35(10), 904–908. <https://doi.org/10.1038/nbt.3956> (cited on page 88).
- Mahieu, N. G., Spalding, J. L., Gelman, S. J., & Patti, G. J. (2016). Defining and detecting complex peak relationships in mass spectral data: The mz.unity algorithm. *Analytical Chemistry*, 88(18), 9037–9046. <https://doi.org/10.1021/acs.analchem.6b01702> (cited on page 90).
- Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., Nguyen, D. D., Watrous, J., Kaponov, C. A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V., Meehan, M. J., Liu, W.-T., Crüsemann, M., Boudreau, P. D., Esquenazi, E., Sandoval-Calderón, M., ... Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nature Biotechnology*, 34(8), 828–837. <https://doi.org/10.1038/nbt.3597> (cited on page 91).
- Guijas, C., Montenegro-Burke, J. R., Domingo-Almenara, X., Palermo, A., Warth, B., Hermann, G., Koelensperger, G., Huan, T., Uritboonthai, W., Aisporna, A. E., Wolan, D. W., Spilker, M. E., Benton, H. P., & Siuzdak, G. (2018). METLIN: A technology platform for identifying knowns and un-

Chapter 3

knowns. *Analytical Chemistry*, 90(5), 3156–3164. <https://doi.org/10.1021/acs.analchem.7b04424>
(cited on page 91).

Chapter 4

High-throughput enrichment of temperature-sensitive argininosuccinate synthetase for two-stage citrulline production in *E. coli*

Thorben Schramm, Martin Lempp, Dominik Beuter, Silvia Gonzáles Sierra, Timo Glat-ter, Hannes Link

This chapter was published in: Schramm et al., 2020, *Metabolic Engineering*, **2020**, 60, pp. 14-24, DOI: <https://doi.org/10.1016/j.ymben.2020.03.004>, (license: Creative Commons "Attribution 4.0 International" (CC BY 4.0), <https://creativecommons.org/licenses/by/4.0/>, no modifications). Following the CRediT authorship contribution statements, my contributions included the conceptualization, investigation, visualization, project administration, writing - original draft, writing - review & editing, and formal analysis. I constructed plasmids and the mutant *argG* library, engineered *E. coli* strains, and performed growth and production experiments. Martin Lempp and I performed the bioreactor experiments together. Martin Lempp performed the off-gas analysis of the bioreactor experiments. Silvia Gonzales Sierra performed cell sorting by flow cytometry (FACS). I performed all metabolite measurements by LC-MS/MS and by a glucose assay kit. I purified enzymes and performed activity assays. I prepared samples for proteomics. Timo Glat-ter performed proteomics measurements and processed the related data. I analysed the data, prepared figures, and co-wrote the paper.

Abstract

Controlling metabolism of engineered microbes is important to modulate cell growth and production during a bioprocess. For example, external parameters such as light, chemical inducers, or temperature can act on metabolism of production strains by changing the abundance or activity of enzymes. Here, we created temperature-sensitive variants of an essential enzyme in arginine biosynthesis of *Escherichia coli* (argininosuccinate synthetase, ArgG) and used them to dynamically control citrulline overproduction and growth of *E. coli*. We show a method for high-throughput enrichment of temperature-sensitive ArgG variants with a fluorescent TIMER protein and flow cytometry. With 90 of the thus derived ArgG variants, we complemented an ArgG deletion strain showing that 90% of the strains exhibit temperature-sensitive growth and 69% of the strains are auxotrophic for arginine at 42°C and prototrophic at 30°C. The best temperature-sensitive ArgG variant enabled precise and tunable control of cell growth by temperature changes. Expressing this variant in a feedback-dysregulated *E. coli* strain allowed us to realize a two-stage bioprocess: a 33°C growth-phase for biomass accumulation and a 39°C stationary-phase for citrulline production. With this two-stage strategy, we produced 3 g/L citrulline during 45 h cultivation in a 1-L bioreactor. These results show that temperature-sensitive enzymes can be created en masse and that they may function as metabolic valves in engineered bacteria.

1 Introduction

The ability to switch overproduction strains between different physiological states enables bioprocesses with separate phases of growth and production (Brockman and Prather, 2015; Burg et al., 2016; Klamt et al., 2018; Lalwani et al., 2018; S. Li et al., 2016). A common parameter to control growth and production is the concentration of nutrients in bioreactors, which affects physiology and metabolism of production strains (Mears et al., 2017; Michalowski et al., 2017; Tokuyama et al., 2019). Alternatively, it is also possible to control metabolism directly by engineering enzymes that act as metabolic valves. To this end, either the abundance or activity of the respective enzyme responds to external parameters, such as light (Zhao et al., 2018), chemicals/auto-inducers (Brockman and Prather, 2015; Gupta et al., 2017; Soma et al., 2014), or temperature (B.-J. Harder et al., 2018).

Temperature-responsive control of metabolism is economically advantageous since most bioreactors are standardly equipped with temperature regulation. Another advan-

tage is that temperature dependent control is abundant in nature and occurs at DNA, RNA, and protein level (Klinkert and Narberhaus, 2009). As a consequence, many mutants of different organisms are known that show temperature-sensitivity (Z. Li et al., 2011; Lovato et al., 2009; Saluja and Godson, 1995). In yeast, collecting known and creating new temperature-sensitive mutants led to large strain libraries, which can be applied to systematically investigate the function of essential genes (Z. Li et al., 2011; Ben-Aroya et al., 2008; Kofoed et al., 2015).

Many temperature-responsive systems use transcription factors that show thermal sensitivity. For example, the temperature-sensitive transcription factor RheA of *Streptomyces albus* changes its activity reversibly upon temperature changes (Servant et al., 2000) and was used to control gene expression in chicken embryo cells (Weber, 2003). Temperature-sensitive variants of the transcription factor TetR were created to control expression of a restriction endonuclease in *E. coli*, which is a toxic product (Pearce et al., 2017). Another temperature-sensitive transcription factor is cI857 from *Escherichia* virus Lambda, which has a mutation (A66T) that is thought to introduce thermal instability (Nauta et al., 1997). This transcription factor was used in a biotechnological application to control the TCA cycle in *E. coli* (B.-J. Harder et al., 2018). In this study, a temperature downshift from 37°C to 28°C repressed expression of the TCA cycle enzyme isocitrate dehydrogenase and redirected metabolic flux from biomass formation to production of itaconic acid. As an alternative to temperature-sensitive control of transcription, RNA thermometers control enzyme abundance at the level of translation in *E. coli* (Neupert et al., 2008; Sen et al., 2017) and *Thermus thermophilus* (Verdú et al., 2019).

Apart from temperature-dependent control of transcription and translation, temperature-sensitive enzymes allow a more direct control of metabolism. At the non-permissive temperature, the enzyme is inactive, which can arrest growth if the enzyme is essential or it redirects metabolic flux. A temperature-sensitive variant of the fatty acid biosynthesis enzyme FabI was used to produce malonyl-CoA derived 3-hydroxypropionic acid (M. D. Lynch et al., 2016), as well as temperature-sensitive variants of the enzymes FabI, FabB and FabD for fatty acid production (M. Lynch et al., 2019). These temperature-sensitive enzyme variants were obtained by nitrosoguanidine- or ethyl methane sulfonate- based random mutagenesis techniques (Broekman and Steenbakkens, 1973; M. E. Harder et al., 1974; Russell and Pittard, 1971). A temperature-sensitive variant of glyceraldehyde-3-phosphate dehydrogenase (GapA) was used to control glycolysis (Cho et al., 2012). GapA mutants were generated by error-prone PCR and used to complement a *gapA* deletion strain. This made it possible to identify temperature-sensitive GapA variants by inspect-

ing colony growth at 30°C and 37°C.

Here, we show a method for high-throughput enrichment of essential temperature-sensitive enzyme variants with the single-cell growth rate reporter TIMER (Beuter et al., 2018; Claudi et al., 2014). As a case study, we used argininosuccinate synthetase (ArgG) in the arginine pathway and created ArgG variants by error-prone PCR. The fluorescent TIMER protein allowed us to enrich temperature-sensitive variants from the ArgG library, and we characterized 90 ArgG variants in more detail.

We then used one ArgG variant to dynamically control growth and production of citrulline in *E. coli*. Citrulline is the substrate of ArgG and an essential intermediate in the arginine biosynthesis pathway. Citrulline plays an important role in human metabolism (Curis et al., 2005) and is used as a dietary supplement. Approaches to produce citrulline include biocatalysis with purified enzymes or whole cell catalysis (Kakimoto et al., 1971; Song et al., 2015; Yamamoto et al., 1974), as well as fermentation processes with *Corynebacterium glutamicum* or *Bacillus subtilis* (Eberhardt et al., 2014; Hao et al., 2015; Ikeda et al., 2009; Shinji et al., 1966). An engineered strain of *C. glutamicum* produced the highest titers reported in the literature: 27 g/L citrulline at a final OD of ca. 70 (Ikeda et al., 2009). To our knowledge, citrulline has not been overproduced with engineered *E. coli*. Here, we produced 3 g/L citrulline in a 1-L bioreactor using a two-stage process strategy, which had a growth phase at 33°C and a production phase at 39°C.

2 Material and Methods

Chemicals were ordered from Merck KGaA or Carl Roth GmbH & Co. KG. MATLAB R2017b (MathWorks, Inc.) and BD FACSDiva 8.0.1 (BD Biosciences) were used for data analysis.

2.1 Construction of plasmids

Strains, plasmids, and oligonucleotides are listed in Suppl. Table 1. Q5 High-Fidelity DNA polymerase (New England BioLabs Inc. (NEB)) was used in PCRs. Plasmids were constructed using Gibson Assembly Master Mix (NEB). DNA fragments for Gibson assembly were purified after agarose gel-electrophoresis (NucleoSpin Gel and PCR Clean-up Kit, Macherey-Nagel GmbH & Co. KG). The DNA Clean & Concentrator Kit (Zymo Research Europe GmbH) was used for DNA clean-up after PCRs or Gibson assemblies. Plasmids were isolated from liquid cultures with the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific Inc.). Oligonucleotides were obtained from Eurofins Genomics Germany GmbH. The wild-type *argG* gene from *E. coli* was expressed from a lowcopy plasmid pTS036-*argG* with the pSC101 origin of replication (Suppl. Fig. 4.7) derived from the plasmid pUA66-*rrnBp* (Zaslaver et al., 2006). The pTS036-*argG* plasmid carries a chloramphenicol resistance *cmR*. The *argG* gene is under control of a tetR inducible promoter (pLetO-1) (Lutz, 1997) and a strong RBS (Elowitz and Leibler, 2000). For overexpression of *argG* and *argG*-G9, pTS049 and pTS050 were constructed with the pCA24N backbone from the ASKA library (Kitagawa et al., 2006). ArgG and ArgG-G9 had an N-terminal His-tag.

2.2 Construction of strains

The $\Delta argG$ and $\Delta argR$ strains originated from the KEIO collection (Baba et al., 2006). The kanamycin resistance was removed with FLP-recombinase on the pCP20 plasmid (Datsenko and Wanner, 2000). Gene deletions were propagated from KEIO strains to other strains by P1 phage transduction (Thomason et al., 2007). The *argA*-H15Y mutation was introduced into the genome of the $\Delta argG \Delta argR$ strain with a CRISPR-cas9 based method (Reisch and Prather, 2015; Sander et al., 2019). All modifications of genomic DNA were verified by sequencing.

2.3 Cultivations

M9 minimal medium was used that contained 5 g/L glucose, 42.2 mM Na₂HPO₄, 22 mM KH₂ PO₄, 11.3 mM (NH₄)₂SO₄, 8.56 mM NaCl, 1 mM MgSO₄ x 7 H₂O, 100 μM CaCl₂ x 2 H₂O, 60 μM FeCl₃, 7.6 μM CoCl₂ x 6 H₂O, 7.1 μM MnSO₄ x 2 H₂O, 7 μM CuCl₂ x 2 H₂O, and 6.3 μM ZnSO₄ x 7 H₂O. When required, 50 μg/mL kanamycin, 100 μg/mL carbenecillin, 30 μg/L chloramphenicol, 50 μg/mL spectinomycin, or 0.2 μM anhydrotetracycline (aTc) were added to the medium. LB and M9 minimal medium plates contained 1.5% agar. The conversion factor optical density (OD) to cell dry weight was 0.37 g_{DW} OD⁻¹ L⁻¹. The OD was measured at 600 nm.

Growth characterization of 90 strain carrying ArgG variants

0.5 mL LB precultures were inoculated with 90 randomly picked strains carrying ArgG variants from agar plate and incubated in a 2 mL deep-well plate, sealed with gas permeable foil (Diversified Biotech, Inc.), at 37°C for 6 h under shaking at 220 rpm. Glycerol stocks were prepared from the LB precultures. 150 μL M9 cultures were inoculated 1:75 with the LB precultures and incubated in a transparent 96-well plate (Greiner Bio-One GmbH) for 20 h at 30°C in a Biotek Epoch plate reader (BioTek Instruments, Inc.). The M9 cultures were re-diluted 1:30 and incubated for another 20 h at 42°C.

Probing conditional arginine auxotrophy of nine strains carrying ArgG variants

0.5 mL LB precultures were started from glycerol stocks and incubated in a 2 mL-deep well plate, sealed with gas permeable foil, for 6 h at 37°C under shaking at 220 rpm. 0.5 mL M9 precultures were inoculated 1:75 with LB precultures and incubated in a 2 mL-deep well plate at 30°C for 16 h under shaking at 220 rpm. 150 μL M9 main cultures, supplemented with and without 1 mM arginine, were inoculated 1:75 with the M9 precultures and incubated in a transparent 96-well plate at 42°C in a plate reader.

Citrulline production with *argG* deletion strains

5 mL LB precultures were inoculated from glycerol stocks and incubated for 6 h at 37°C in a rotary shaker. 5 mL M9 precultures, supplemented with 1 mM arginine, were inoculated 1:500 with the LB precultures and incubated overnight at 37°C in a rotary shaker. M9 precultures were washed twice by centrifugation (4000 rpm, 37°C, 5 min) with 5

mL M9 medium and resuspended in 5 mL M9 medium. 40 mL M9 main cultures supplemented with 100 μ M arginine were inoculated with the washed M9 precultures to a start OD of 0.05 and incubated in 500 mL shake flasks at 37°C under shaking at 220 rpm. After 3 h, the M9 main cultures were washed twice with M9 medium to remove the arginine from the medium. Subsequently, cells were resuspended in 30 mL of fresh M9 medium to an OD of 0.05, and incubated at 220 rpm and 37°C.

Growth and citrulline production screening of nine engineered strains carrying ArgG variants

5 mL LB precultures were inoculated from glycerol stocks and incubated for 6 h at 37°C under shaking at 220 rpm. 5 mL M9 precultures were inoculated 1:100 with the LB precultures and incubated for 16 h at 30°C under shaking of 220 rpm. 10 mL M9 main cultures were inoculated with the M9 precultures to a start OD of 0.05. The 10 mL M9 main cultures were split up to four parts, each 2 mL, and transferred to 5 mL-culture tubes that were incubated at different temperatures (30°C, 34°C, 37°C, and 42°C) for 7 h under shaking at 220 rpm.

Growth characterization of strains carrying the ArgG variant G9 and the wild-type ArgG

5 mL LB precultures were inoculated from glycerol stocks and incubated for 6 h at 37°C in a rotary shaker. 40 mL M9 precultures were inoculated 1:200 and incubated in 500 mL shake flasks for 16 h at 30°C and 220 rpm of shaking. The growing cultures were washed twice with M9 medium by centrifugation, resuspended, and diluted to yield 45 mL M9 main cultures with a start OD of 0.05. The 45 mL M9 main cultures were split up to three parts, each 14 mL, and transferred to 100 mL shake flasks that were incubated at three different temperatures (30°C/ 35°C/39°C or 33°C/37°C/42°C) under shaking at 220 rpm.

Two-stage bioreactor cultivation

5 mL LB precultures were inoculated from glycerol stocks and incubated for 6 h at 37°C in a rotary shaker. 40 mL M9 precultures were inoculated 1:500 with the LB precultures and incubated in 500 mL shake flasks overnight at 33°C under shaking at 220 rpm. The M9 precultures were diluted in 100 mL M9 to an OD of 0.15 and further cultivated in 1 L shake flasks at 33°C for 8 h under shaking at 220 rpm. Subsequently, 45 mL of the M9 precultures were washed twice with M9 medium by centrifugation, resuspended, and di-

luted to yield 500 mL M9 main cultures with a start OD of 0.05, which were supplemented with 12.5 g/L glucose and 4 g/L NH₄Cl and incubated in bioreactors (BioFlo/CelliGen 115 and BioFlo 120, Eppendorf AG). After 13 h of incubation at 33°C, the temperature was switched stepwise to 39°C (30 min 33°C – 36°C, 10 min 36°C – 38°C, 38°C – 39°C). After 20 h, 25 mL of a sterile 300 g/L glucose 30 g/L NH₄Cl solution was fed to the M9 main cultures.

2.4 Cultivation for proteome analysis at different temperatures

5 mL LB precultures were inoculated from glycerol stocks and incubated for 6 h at 37°C in a rotary shaker. 50 mL M9 precultures were inoculated 1:200 and incubated in 500 mL shake flasks overnight at 30°C and 220 rpm. The cultures were washed twice with M9 medium by centrifugation, resuspended, and diluted to yield 70 mL M9 main cultures with a start OD of 0.2. The 70 mL M9 main cultures were divided into three volumes, each 20 mL, and transferred to 100 mL shake flasks that were incubated at three different temperatures (33°C/39°C/ 42°C) under shaking at 220 rpm.

2.5 Enrichment of the temperature-sensitive ArgG variants

The *argG* gene was mutagenized with error-prone PCR according to the manufactures manual (Jena Bioscience GmbH, JBS Error-Prone Kit # PP-102). The template for the error-prone PCR was linear *argG* DNA (final concentration in the PCR: 0.1 ng/μL), which was amplified from the *E. coli* BW25113 genome and purified by agarose gel-electrophoresis. The mutagenized *argG* DNA was cleaned-up and inserted into the plasmid pTS036 with Gibson assembly (NEBuilder HiFi DNA Assembly Master Mix (NEB)). After DNA clean-up, electrocompetent cells (MegaX DH10β cells) were transformed with pTS036-*argG*(mutant), and plated on two 150 mm petri-dishes with LB agar. After incubation overnight at 37°C, colonies from two plates were collected in 20 mL LB medium and the plasmids were isolated, resulting in the final *argG* variant library pTS036-*argG*(mutant). *E. coli* Δ*argG* carrying the TIMER plasmid (pBR322_ TIMER) was transformed with the pTS036-*argG*(mutant) library by electroporation. After plating the cells, colonies were collected from the plates with LB medium to prepare glycerol stocks. For the selective enrichment of temperature-sensitive variants, 25 mL M9 cultures were inoculated from glycerol stock, diluted at different ratios, and incubated for 36 h at 30°C under shaking at 220 rpm. The culture with a final OD of 0.1 was re-diluted at different ratios and incubated for 6 h at 42°C under shaking at 220 rpm. Then, slow growing cells were sorted with

fluorescence-activated cell sorting (FACS) on a BD FACS Aria Fusion (BD Biosciences, NJ, USA). 561-nm lasers, 600 long pass and 632/22 bandpass filters were used to detect the red fraction of TIMER. 488-nm lasers, 500 long pass and a 520/30 band pass filters were used for green fluorescence. To identify cells in the forward/side scatter plot, 488-nm lasers were used. The sorted cells were plated on M9 agar medium (150 mm petri-dishes) at different dilutions and incubated at 30°C. The growth of 90 randomly picked strains was investigated in microtiter plates (Methods 2.3.1).

2.6 Protein purification and enzymatic characterization of *E. coli* argininosuccinate synthetase ArgG

5 mL TB precultures were inoculated from glycerol stocks and incubated overnight at 37°C in a rotary shaker. 200 mL TB main cultures in 1 L shake flasks were inoculated to a start OD of 0.04 with TB precultures and incubated at 37°C under shaking at 220 rpm. At OD 0.6, 2 mL 0.1 M IPTG was added to the cultures. The cultures were incubated overnight at 16°C under shaking at 220 rpm. Cells from 16°C overnight culture were harvested by centrifugation (4000 rpm, 4°C, 30 min). Following work was conducted at 4°C. Cells were resuspended in 2.8 mL LEW (50 mM NaH₂PO₄ 300 mM NaCl, pH 8.0), and 100 μ L protease inhibitor and DNase 1 was added to aliquots of 0.7 mL. Cells were lysed by sonication (3 \times 1 min, with 30 s of cooling breaks). After centrifugation (30 min, 4°C, 17.000 g), the supernatant was used for protein purification with Protino Ni-TED 1000 Packed Columns (Macherey-Nagel). Proteins were eluted in the first elution step. The protein concentration in the eluate was determined with the Pierce Microplate BCA Protein Assay Kit (Thermo Fisher, # 23252). Stock solutions were prepared with protein concentrations of 0.2 mg mL⁻¹. ArgG enzyme assays were conducted in 1.5 mL reaction tubes. 10 μ L of enzyme stock solution was added to 80 μ L reaction buffer that contained at final concentrations 10 mM ATP, 10 mM MgCl₂, 10 mM L-aspartate, and 20 mM HEPES (pH 7). To start the reaction, 10 μ L of L-citrulline was added (final concentration 10 mM). The formation of the reaction product, argininosuccinate, was measured with LC-MS/MS.

2.7 Citrulline, argininosuccinate, and glucose measurements

Citrulline was measured in the whole cultivation broth. Therefore, 5 μ L cultivation broth was transferred to 495 μ L 40:40:20 acetonitrile:methanol:H₂O at -20°C in 1.5 mL reaction tubes. Samples were centrifuged at -9°C and 17.000 g for 20 min, and the supernatant

was stored at -80°C until analysis with LC-MS/MS. Citrulline concentrations were determined with an Agilent 1290 Infinity II UHPLC system coupled to an Agilent 6495 triple quadrupole mass spectrometer (Agilent Technologies, Inc.) (Guder et al., 2017). The LC column was an Acquity BEH Amide, 30×2.1 mm, $1.7 \mu\text{m}$ particle size, column (Waters Corporation). The mobile phases were (A) water with 10 mM ammonium formate and 0.1 vol.-% formic acid and (B) acetonitrile with 0.1 vol.-% formic acid. The flow rate was 0.4 mL min^{-1} and the gradients were: 0 min 90% B; 1.3 min 40% B; 1.5 min 40% B; 1.7 min 90% B; 2 min 90% B. Citrulline was quantified with a ^{15}N -internal standard. The ^{14}N parent mass for citrulline was 176 and the product mass was 70, at a collision energy of 25 eV in positive mode (179 and 71 for ^{15}N citrulline).

Samples of ArgG enzyme assays were prepared by transferring $10 \mu\text{L}$ of the enzyme assay reaction solution to $90 \mu\text{L}$ of 45:45:10 acetonitrile: methanol:water at -20°C . Samples were stored at -80°C until centrifugation (17,000 g, -9°C). Argininosuccinate concentrations in the supernatants were determined by LC-MS/MS (Guder et al., 2017). The LC column was a HILICON iHILIC-Fusion(p), 50×2.1 mm, $5 \mu\text{m}$ particle size column. The mobile phases were (A) aqueous 10 mM ammonium carbonate and 0.2% ammonium hydroxide and (B) acetonitrile. The flow rate and gradients were the same as for citrulline quantification. Argininosuccinate was quantified with a ^{13}C -internal standard. The ^{12}C parent mass for argininosuccinate was 291 and the product mass 70, at a collision energy of 45 eV in positive mode (301 and 74 for ^{13}C argininosuccinate).

Supernatant samples for glucose measurements were prepared by centrifugation of 1 ml of the culture broth for 1 min at 17,000 g at room temperature. The supernatant was stored at -20°C . The glucose concentration in the supernatant was determined with an assay kit (DGlucose Assay Kit, GOPOD Format, Megazyme Inc).

2.8 Proteomics

Samples were subjected to tryptic digest as described in detail in the Supplementary Methods 4. $1 \mu\text{g}$ peptide was analyzed using liquid chromatography-mass spectrometry (LC-MS/MS). The LC-MS/MS analysis including label-free quantification was carried out as previously described in (Sander et al., 2019) with minor modifications. In short, LC-MS/MS analysis of protein digests was performed on Q-Exactive Plus mass spectrometer connected to an electrospray ion source (Thermo Fisher Scientific). Peptide separation was carried out using Ultimate 3000 nanoLC-system (Thermo Fisher Scientific), equipped with packed in-house C18 resin column (Magic C18 AQ $2.4 \mu\text{m}$, Dr. Maisch). The peptides were first loaded onto a C18 precolumn (preconcentration setup)

and then eluted in backflush mode with a gradient from 98% solvent A (0.15% formic acid) and 2% solvent B (99.85% acetonitrile, 0.15% formic acid) to 30% solvent B over 115 min. Label-free quantification was done using Progenesis QI software (Nonlinear Dynamics, v2.0), MS/MS search was performed in MASCOT (v2.5, Matrix Science) against the Uniprot *E. coli* protein database. The following search parameters were used: full tryptic search with two missed cleavage sites, 10 ppm MS1 and 0.02 Da fragment ion tolerance. Carbamidomethylation (C) as fixed, oxidation (M) and deamidation (N,Q) as variable modification. Progenesis outputs were further processed with SafeQuant (Glatzer et al., 2012).

3 Results

3.1 Enrichment of temperature-sensitive ArgG variants with a TIMER protein

Argininosuccinate synthetase (ArgG) catalyzes the seventh reaction in the arginine biosynthesis pathway of *E. coli*, and converts citrulline into argininosuccinate (Fig. 4.1a). To construct temperature-sensitive variants of ArgG, we used error-prone PCR for in vitro mutagenesis of the *argG* gene (Fig. 4.1b). The resulting library of ArgG variants was then used to complement an *argG* deletion strain ($\Delta argG$) (Baba et al., 2006). The $\Delta argG$ strain is auxotrophic for arginine and does not grow on a minimal glucose medium without arginine (Suppl. Fig. 4.8). Therefore, only cells with catalytically active ArgG variants would grow on glucose minimal medium, and we enriched these variants by culturing the library for 36 h at 30°C in shaking flasks. Then, we shifted the culture to 42°C and expected a growth-arrest of cells expressing temperature-sensitive ArgG variants (variants that are inactivated at 42°C). The challenge was to isolate the fraction of growth-arrested cells from the library, and for this purpose, we used the single-cell growth rate reporter TIMER. The TIMER protein is a dsRed variant, which indicates slow- and non-growing cells by a low ratio of green/red fluorescence (Beuter et al., 2018; Claudi et al., 2014). After switching the ArgG library from 30°C to 42°C, we waited 6 h to allow maturation of the TIMER protein and then isolated 81.000 cells with the lowest green/red ratio with FACS (Fig. 4.1c). The entire ArgG library had a wide-range of green/red ratios with a ratio median of 0.44, while the sorted fraction had a narrow green/red ratio with a median of 0.14. This indicates that the majority of cells was growing at 42°C and did not express temperature-sensitive ArgG variants.

To test if we enriched temperature-sensitive ArgG variants, we randomly selected 90 isolates from the sorted cells with low green/red ratios. The 90 isolates were then cultured in microtiter plates at both 30°C and 42°C. At 30°C, all isolates grew similar to a control strain that expressed the wild-type ArgG, thus indicating that all ArgG variants are catalytically active at 30°C (Fig. 4.1d and Suppl. Fig. 4.9). At 42°C in contrast, the majority of strains (90%) grew worse than the control strain, and 62 out of the 90 strains did not reach a final OD of 0.5. Thus, the single cell growth rate reporter successfully enriched temperature-sensitive ArgG variants. Moreover, the large variation of growth characteristics at 42°C suggests that the enzymes respond differently to a temperature increase.

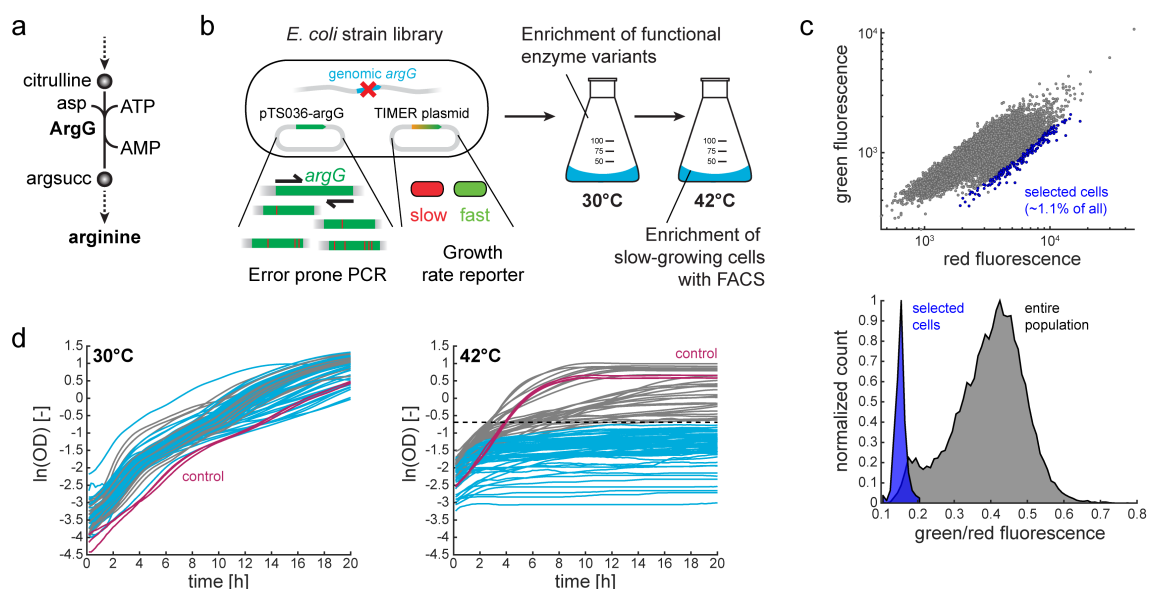


Figure 4.1: High-throughput enrichment of temperature-sensitive ArgG variants. (a) ArgG catalyzes the seventh reaction in the arginine biosynthesis pathway and converts citrulline into argininosuccinate (argsucc). asp, aspartate. (b) A plasmid library with mutagenized *argG* was generated with error-prone-PCR and used to complement an *argG* knockout strain of *E. coli* BW25113. The strain also carried a plasmid with the single-cell growth rate reporter TIMER. The pooled strain library was first incubated at 30°C to enrich ArgG variants that are catalytically active and support growth. Subsequently, the culture was shifted to 42°C to select non-growing cells with FACS based on the green/red signal of the TIMER protein. (c) Red and green fluorescence (top) of cells after 6 h culturing at 42°C. FACS was used to isolate the fraction with a low green/red ratio (shown in blue). Histogram (bottom) showing the distribution of cells according to the green to red ratio. (d) Growth at 30°C and 42°C of 90 single strains isolated from the fraction with a low green/red ratio (blue fraction in Fig. 1c). Strains that did not reach an OD of 0.5 at 42°C are shown in blue. A control strain expressing wild-type ArgG is shown in red.

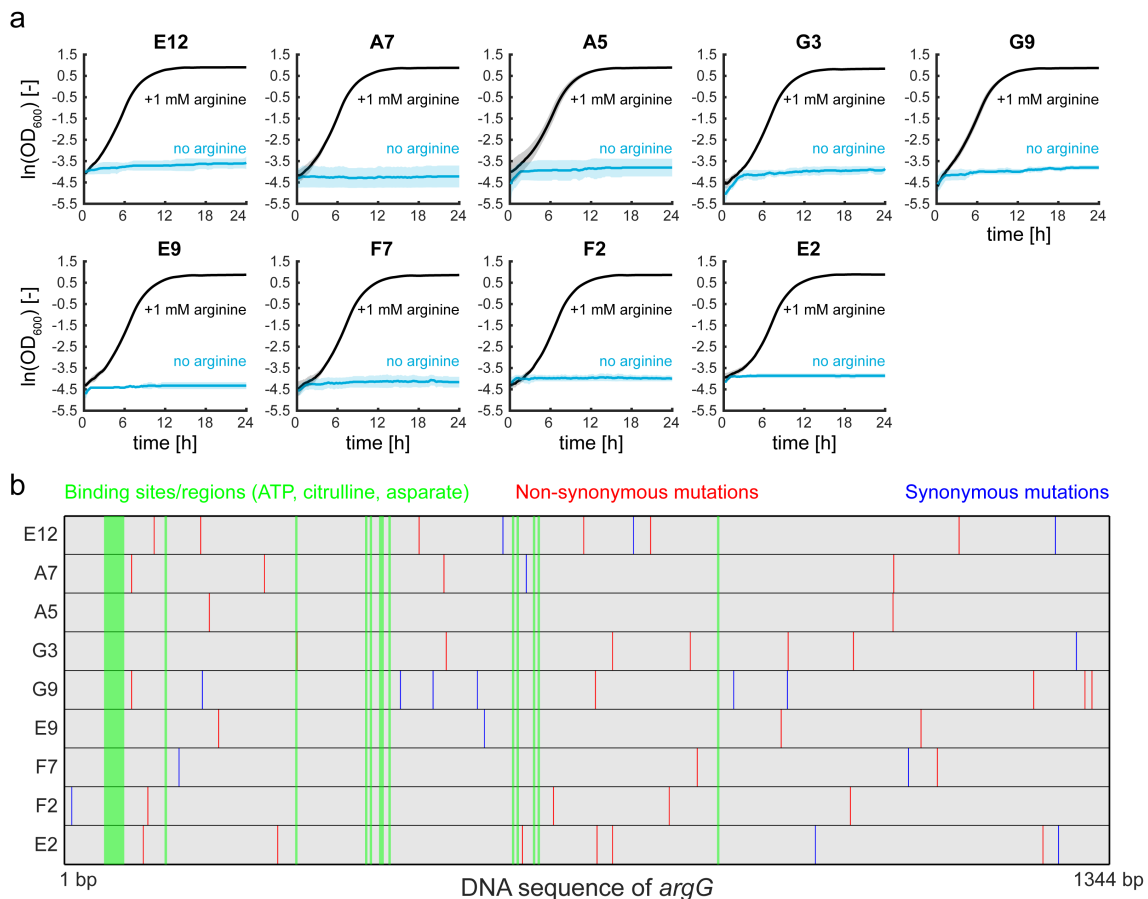


Figure 4.2: Temperature-dependent arginine auxotrophy and mutations of nine ArgG variants. (a) Growth of nine strains with ArgG variants at 42°C with supplementation of arginine (black lines) and without (blue lines). Lines show means and shades the standard deviation of $n = 3$ plate reader cultures. (b) Mutations of the nine ArgG variants. Binding sites of ATP, citrulline, and aspartate are shown in green. Non-synonymous mutations are red, synonymous are blue.

3.2 Characterization of nine temperature-sensitive ArgG variants

We selected nine ArgG variants that did not support growth at 42°C but achieved the highest growth rates at 30°C (Suppl. Fig. 4.10). First, we confirmed that the nine strains are indeed auxotrophic for arginine at 42°C by culturing them with and without supplementation of arginine (Figure 4.2a). Addition of arginine could fully restore growth at 42°C, with maximum growth rates of $0.70 \pm 0.02 \text{ h}^{-1}$. Consequently, the temperature-sensitivity of all nine strains results from an auxotrophy for arginine.

Sequencing of the nine ArgG variants revealed that each variant had between two and eleven mutations (Fig. 4.2b and Suppl. Table 2). 69% of all mutations were non-

synonymous and changed the amino acid sequence. All ArgG variants were unique, and only a few ArgG variants had amino acid exchanges at the same position (G9 and A7 at position 29, E2 and G3 at position 235, and A5 and A7 at position 356). The E235D mutation occurred twice in G3 and E2, and mutant A7 had a point mutation in the ribosomal binding site of *argG*. The number of synonymous mutations per gene varied between zero (variant A5) and six (variant G9). Because codon usage has an impact on the translational efficiency (Plotkin and Kudla, 2011), we expect that these mutations could play a role in tuning expression levels of ArgG. However, it seems unlikely that ArgG expression levels play a role in temperature-sensitivity, because this would also affect growth at 30°C. Mapping the mutations to the DNA sequence of *argG* revealed that only one out of 55 mutations (F100Y of G3) occurred at a known binding site for citrulline, ATP, or aspartate (Lemke and Howell, 2001, 2002) indicating conservation of those critical residues.

To test if these mutations have the potential to reduce protein stability we obtained $\Delta\Delta G$ values with FoldX (Guerois et al., 2002; Schymkowitz et al., 2005). $\Delta\Delta G$ are changes in Gibbs energy (ΔG) relative to the wild-type ArgG sequence. The average $\Delta\Delta G$ for the nine mutants was $6.5 \text{ kcal mol}^{-1}$ (Suppl. Table 2), which indicated strong destabilizing effects. As a reference, previous studies showed that the average $\Delta\Delta G$ for single amino acid substitution in 22 proteins was $1.3 \text{ kcal mol}^{-1}$, and only 15% of these mutations had a $\Delta\Delta G > 3 \text{ kcal mol}^{-1}$ (Tokuriki et al., 2008). Next, we used the nine ArgG variants to dynamically control the arginine pathway and to overproduce the substrate of ArgG, citrulline.

3.3 Citrulline overproduction in feedback-dysregulated *E. coli*

We expected that inactivation of ArgG blocks the arginine pathway and that this would lead to an accumulation of the substrate of ArgG, citrulline (Fig. 4.3a). To test the consequences of blocking arginine biosynthesis at the reaction catalyzed by ArgG (argininosuccinate synthetase), we used an *argG* deletion strain ($\Delta argG$). First, the $\Delta argG$ strain was grown in medium with externally supplemented arginine, which suppresses the physiological consequences of the *argG* deletion. Then, we washed and diluted the cells in medium without arginine so that cells would have to switch to de novo arginine biosynthesis. After dilution, we measured the OD and citrulline concentration in the whole cultivation broth for 54 h (Fig. 4.3b). As expected, cells did not grow after removing arginine from the medium, and the OD remained constant at 0.05 (Suppl. Fig. 4.8). Citrulline, in contrast, increased to a final concentration of 20 mg/L, showing that blocking ArgG results in accumulation of citrulline.

To improve citrulline production, we dysregulated the arginine pathway, which is feedback regulated at the transcriptional and allosteric level by arginine (Fig. 4.3a). Therefore, we deleted the transcriptional repressor of the arginine pathway (ArgR) and removed allosteric feedback inhibition by a point mutation in the first enzyme in the pathway (H15Y mutation of ArgA) (Rajagopal et al., 1998). Deletion of ArgR resulted in a 2-fold higher citrulline concentration at the end of the experiment (40 mg/L) (Fig. 4.3b). A doubly dysregulated strain ($\Delta argG \Delta argR argA(H15Y)$) achieved 3-fold higher citrulline concentrations (60 mg/L). The specific citrulline production rates of the three strains decreased over time. During the first 10 h, the doubly dysregulated strain had a biomass-specific citrulline production rate of $0.73 \text{ mmol g}_{DW}^{-1} \text{ h}^{-1}$, which decreased to $0.25 \text{ mmol g}_{DW}^{-1} \text{ h}^{-1}$ between 20 and 32 h. Citrulline production stopped between 44 and 54 h.

These results show that blocking arginine biosynthesis at the ArgG reaction leads to overproduction of citrulline, and that complete dysregulation of the arginine pathway enhances production. Next, we combined the doubly dysregulated strain ($\Delta argG \Delta argR argA(H15Y)$) with temperature-sensitive ArgG to control citrulline production dynamically.

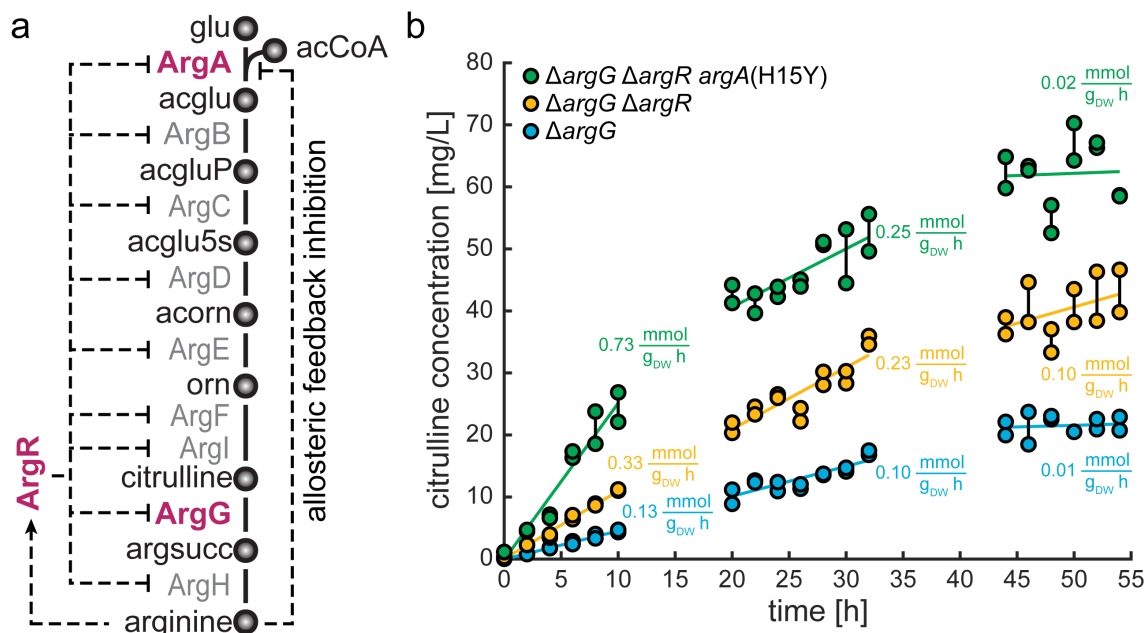


Figure 4.3: Overproduction of citrulline in an *argG* deletion strain. (a) Arginine biosynthesis in *E. coli* is feedback regulated by arginine at the level of transcription (*ArgR*) and allosteric control of the first enzyme (*ArgA*). The transcription factor *ArgR* regulates the expression of all genes in the pathway. The activity of the protein *ArgA*, which catalyzes the first reaction in the pathway, is directly regulated by allosteric interaction with arginine. Engineering targets are shown in red: deletion of *argG*, deletion of *argR*, H15Y mutation removes allosteric inhibition of *ArgA*. glu, L-glutamate; acglu, N-acetyl-L-glutamate; acgluP, N-acetylglutamyl-phosphate; acglu5s, N-acetyl-L-glutamate 5-semialdehyde; acorn, N-acetyl-L-ornithine; orn, L-ornithine; argsucc, L-arginino-succinate; acCoA, acetyl-coenzyme-A; (b) Citrulline concentration in the whole cultivation broth of three *argG* deletion strains after removing arginine from the cultivation medium. Blue: a strain with only *argG* deletion, Orange: a strain with *argG* deletion and additional deletion of the transcriptional repressor *argR*. Green: a strain with deletion of *argG*, *argR* and a point mutation (H15Y) in *argA* that removes inhibition of *ArgA* by arginine. Specific citrulline production rates were calculated by regression analysis in the three time intervals.

3.4 Growth and citrulline production screening of the temperature-sensitive ArgG variants

To control growth and overproduction of citrulline dynamically, we expressed the nine temperature-sensitive ArgG variants in the doubly dysregulated citrulline producer $\Delta argG \Delta argR argA(H15Y)$. As a control, we also expressed the wild-type ArgG enzyme in the same strain. We expected that temperature-sensitive ArgG functions as a metabolic valve that controls arginine biosynthesis flux in a temperature-dependent manner (Fig. 4.4a). At low temperatures, ArgG is active and sustains biosynthetic flux into arginine and growth. At high temperatures, ArgG is inactive and blocks the arginine pathway, resulting in overproduction of citrulline and a growth-arrest.

To test if the nine ArgG variants achieve this metabolic control in the citrulline producer, we cultured the nine strains and the control (Fig. 4.4b). M9 precultures (30°C) were used to inoculate 4 main cultures at OD 0.05, which were incubated at different temperatures (30°C/34°C/37°C/42°C). OD and citrulline were measured after 7 h. At 30°C, all strains with a temperature-sensitive ArgG variant grew to a similar OD, which was even higher than the OD of the control strain (Fig. 4c). Citrulline production at 30°C was similar in all strains, including the control (Fig. 4.4c). The basal production of citrulline indicates that ArgG is limiting the flux in the arginine pathway already at 30°C. This bottleneck at ArgG is probably caused by high abundances of the other enzymes in the arginine pathway due to the ArgR deletion (Sander et al., 2019b).

When we cultured the strains at 42°C, the OD of all temperature-sensitive variants increased just slightly (from 0.05 to an average of 0.08). Citrulline production of the variants increased on average 2.2- fold at 42°C when compared to 30°C. In contrast, the control strain reached almost the same OD and citrulline levels at 42°C and 30°C. This shows that all ArgG variants enable temperature-dependent control of both growth and citrulline production.

Notably, at intermediate temperatures (34°C and 37°C), the OD and citrulline levels differed across the nine ArgG variants (Suppl. Fig. 4.11), suggesting that the variants have different temperature dependencies. The ArgG variant G9 achieved the highest citrulline levels, and we tested temperature-sensitivity of this variant with in vitro enzyme assays. The results show that ArgG-G9 is indeed temperature-sensitive: the specific enzyme activity was $11.5 \mu\text{mol min}^{-1} \text{mg}^{-1}$ at 30°C and $4.1 \mu\text{mol min}^{-1} \text{mg}^{-1}$ at 42°C (Fig. 4.4d). In contrast, the activities of wildtype ArgG were $47.5 \mu\text{mol min}^{-1} \text{mg}^{-1}$ at 30°C and $114.5 \mu\text{mol min}^{-1} \text{mg}^{-1}$ at 42°C. The incubation time at 42°C did not affect activity of the

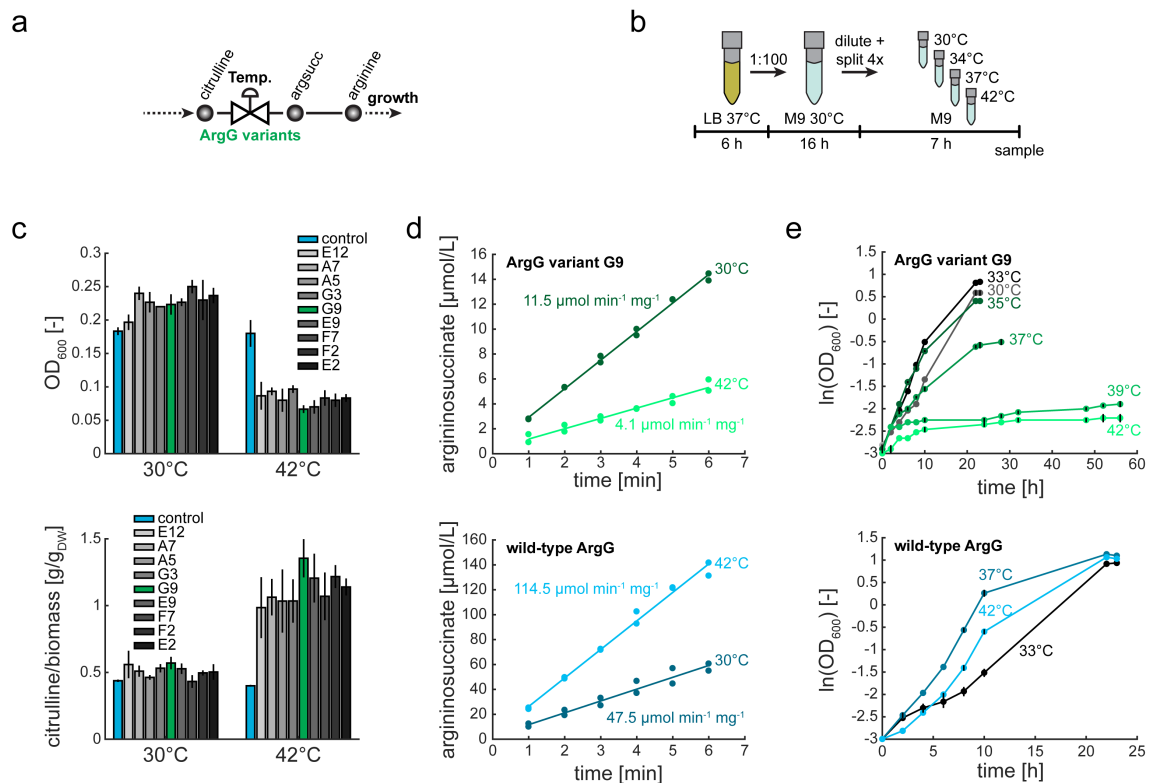


Figure 4.4: Growth and citrulline production of the doubly dysregulated citrulline producer in combination with the nine ArgG variants. (a) Schematic of a temperature-sensitive metabolic valve at ArgG. (b) Schematic of the experimental setup to screen citrulline production and growth of nine ArgG variants at different temperatures. (c) The nine ArgG variants were expressed in the doubly dysregulated citrulline producer ($\Delta argG \Delta argR argGA(H15Y)$). Shown is the OD (top) after 7 h cultivation in minimal medium at 30°C and 42°C. All cultures started at an OD of 0.05. Error bars show the standard deviation of $n = 3$ cultures. Biomass specific citrulline concentration (bottom) after 7 h cultivation in minimal medium at 30°C and 42°C. Error bars show the standard deviation of $n = 3$ cultures. (d) In vitro enzymatic assays with purified ArgG variant G9 (top) and wild-type ArgG (bottom) at different temperatures ($n = 2$ enzyme assays, proteins purified 2 times). Shown is the formation of the reaction product (argininosuccinate) after starting the reaction at $t = 0$ min. Specific enzyme activities were calculated with linear regression. (e) Growth of the doubly dysregulated citrulline producer ($\Delta argG \Delta argR argGA(H15Y)$) expressing the ArgG variant G9 (top), and wild-type ArgG (bottom). Colors indicate different temperatures. Dots are means, and error bars show the difference between $n = 2$ cultures.

enzymes, because activities did not change when incubating the enzymes for 10 min or 60 min at 42°C (wild type: 112.0 $\mu\text{mol min}^{-1} \text{mg}^{-1}$, variant G9: 4.0 $\mu\text{mol min}^{-1} \text{mg}^{-1}$) (Suppl. Fig. 4.12). This data indicated that temperature-sensitivity of the ArgG-G9 variant occurred at faster time-scales.

Cultivating the doubly dysregulated strain with ArgG variant G9 at different temperatures (30, 33, 35, 37, 39, 42°C) revealed that the best growth occurred at 33°C and that 39°C was sufficient to stop growth (Fig. 4.4e). Between 33°C and 39°C, growth decreased gradually, showing tunable growth control of this ArgG variant by temperature. The control strain, which expressed the wild-type ArgG, grew at all temperatures and best at 37°C (Fig. 4.4e).

Strains with ArgG-G9 grew faster than strains with wild-type ArgG at low temperatures (Fig. 4.4e and Suppl. Fig. 4.10) although the specific enzyme activity of ArgG-G9 was lower compared to the wild-type ArgG (Fig. 4.4d). A potential explanation is that the mutations (also synonymous) tuned and increased expression levels by codon usage or mRNA stability resulting in faster growth.

Heat stress and the accompanying metabolic burden could potentially pose a major challenge in temperature-controlled bioprocesses. Therefore, we measured the proteome of our strains at different temperatures and investigated the abundance of heat-shock related proteins (Fig. 4.5a). Heat-shock related proteins were stronger expressed at 42°C than at 39°C, which suggested that switching to 39°C was less burdensome than switching to 42°C.

We then inspected the abundance of the proteins in the arginine pathway. All arginine enzymes increased in the doubly dysregulated strain (average 6.9-fold), which was expected due to deletion of ArgR (Fig. 4.5b). The only exception was ArgG-G9, because this enzyme was expressed from a plasmid. ArgG-G9 levels were 5.9-fold below wildtype levels during exponential growth at 33°C. At 39°C and 42°C, ArgG-G9 levels decreased further, 6.3- and 7.4-fold, respectively. The lower abundance of ArgG-G9 at high temperatures indicated degradation of either the enzyme or the mRNA. Thus, it seems that high temperature affects ArgG-G9 two-fold: it reduces stability and abundance of the protein.

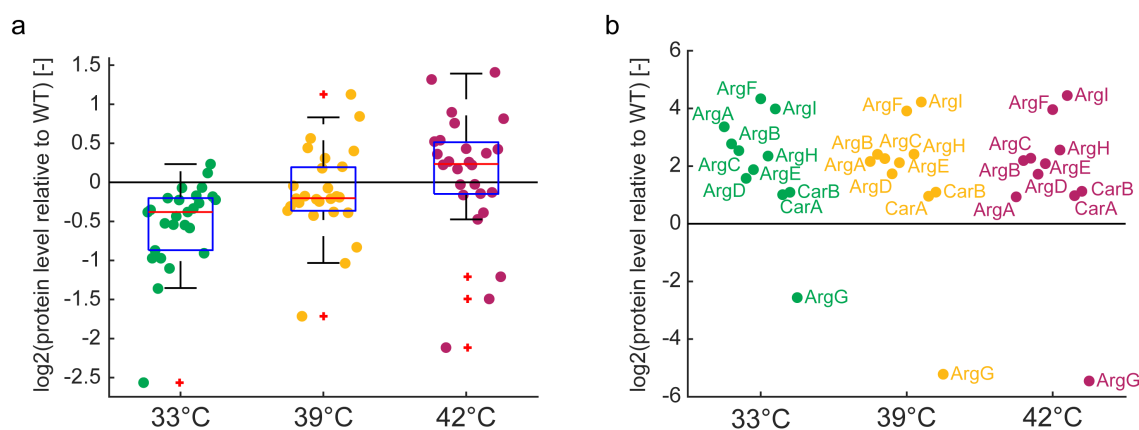


Figure 4.5: Proteome data of the doubly dysregulated citrulline producer ($\Delta argG \Delta argR argA(H15Y)$) expressing the ArgG variant G9 at different temperatures (33°C, 39°C, 42°C). Data is normalized to the proteome of exponentially growing wild-type cells at 37°C. Dots are the mean of independent replicates ($n = 3$). (a) Relative abundance of heat-shock proteins: IbpB, IbpA, DnaJ, GroS, DnaK, FxsA, GroL, ClpB, HtpX, HtpG, GrpE, Lon, YcjF, PrlC, HslV, MutM, HslU, YbbN, YbeZ, RpoD, YbeD, YcjX, LdhA, ClpP, ClpX, HslJ. Red lines indicate the medians. Boxes indicate the 25th and 75th percentiles. (b) Relative abundance of enzymes in the arginine biosynthesis pathway as well as CarA and CarB.

3.5 A temperature-controlled two-stage process for overproduction of citrulline

We used the doubly dysregulated strain ($\Delta argG \Delta argR argA(H15Y)$) with the temperature-sensitive ArgG variant G9 to produce citrulline in a 1 L-bioreactor. This strain grew best at 33°C and growth stopped already at 39°C. Therefore, we used these temperatures to separate a growth phase (33°C) and a production phase (39°C). We cultivated the strain in two independent bioreactors for 14 h at 33°C in minimal medium until the biomass reached an OD of approximately 1. During this time, the cells grew exponentially with growth rates of 0.31 h⁻¹ and 0.35 h⁻¹. Then, we slowly increased the temperature to 39°C over a time of 1 h. The slow temperature increase was necessary to avoid a temperature overshoot that may cause a heat shock. Growth stopped when both bioreactors reached 39°C, and cells entered a stationary phase (Fig. 4.6a).

Throughout the experiment, we measured glucose and citrulline concentrations to quantify specific uptake and production rates. The glucose concentrations in the supernatant decreased continuously (Fig. 4.6b), and during the first 5 h into the stationary phase the specific glucose uptake rates were 4.2 mmol g_{DW}⁻¹ h⁻¹ and 5.3 mmol g_{DW}⁻¹ h⁻¹. These rates correspond to 45% of the glucose uptake rate of glucose-fed wild-type *E. coli* during exponential growth at 37°C (Long et al., 2018). Thus, despite the growth arrest, cells remained a relatively high metabolic activity, which is about 45% of exponentially growing cells.

Apart from arresting growth, the temperature shift from 33°C to 39°C induced production of citrulline. During the 30 h stationary phase, specific citrulline production was constant at 0.92 mmol g_{DW}⁻¹ h⁻¹ (bioreactor 1) and 1.10 mmol g_{DW}⁻¹ h⁻¹ (bioreactor 2). These rates matched productions rates estimated from the initial screening (Fig. 4.4c) and were 38% higher than for the doubly dysregulated strain with deletion of ArgG (Fig. 4.3b). The final concentrations of citrulline in the two bioreactors were 3.3 g/L and 2.9 g/L (Fig. 6c), and the final biomass-specific citrulline yields were 4.88 g/g_{DW} and 6.07 g/g_{DW}. Titters, yields, and production rates of all experiments are listed in Suppl. Table 4.3.

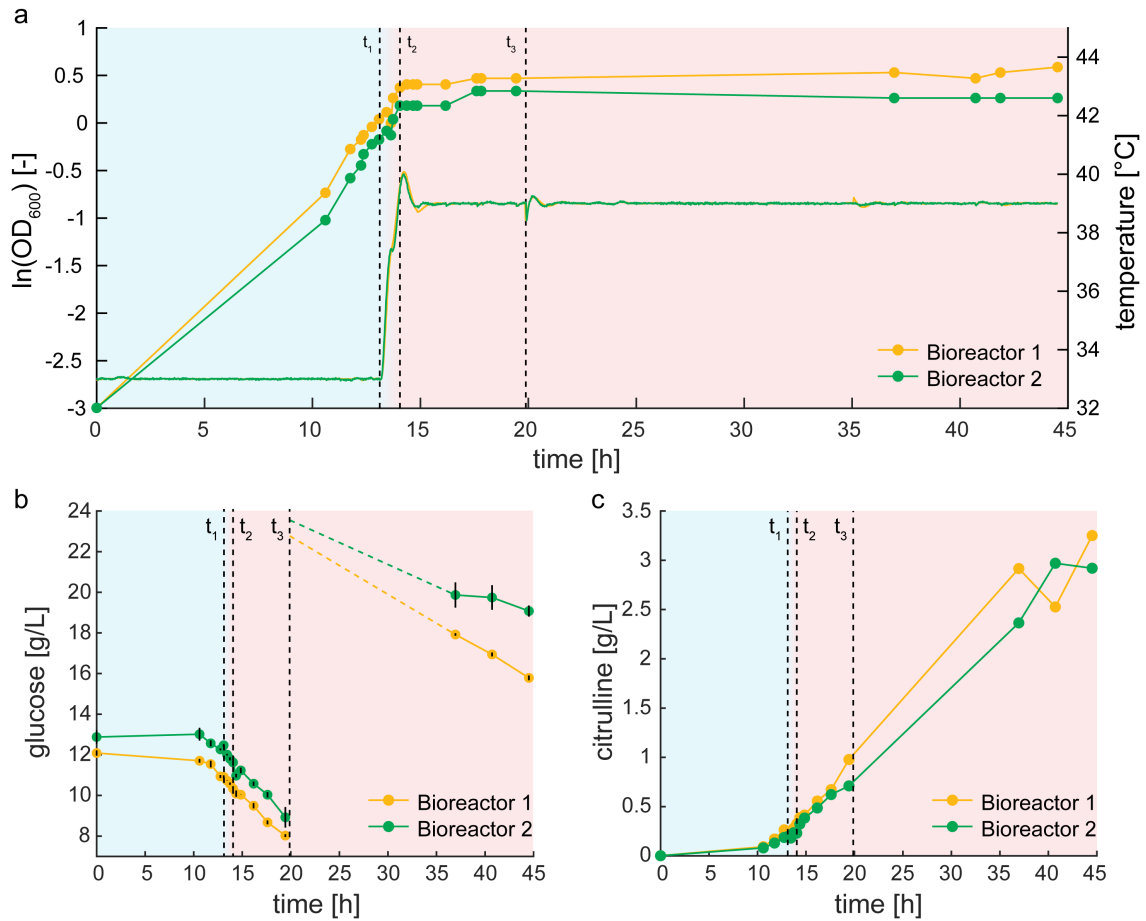


Figure 4.6: Two-stage production of citrulline with the ArgG variant G9. (a) The doubly dysregulated citrulline producer ($\Delta argG \Delta argR argA(H15Y)$) expressing the ArgG variant G9 was cultivated in two independent 1-L bioreactors. OD and the temperature are shown for bioreactor 1 (orange) and bioreactor 2 (green). t_1 and t_2 indicate the time window when temperature was increased from 33°C to 39°C. t_3 indicates the time when glucose and ammonium was fed. blue area: growth phase, red area: production phase. (b) Glucose concentration in the supernatant of the two bioreactors. Dots are the mean, and error bars are the standard deviation of $n = 4$ analytical replicates per bioreactor. (c) Citrulline concentration in the whole cultivation broth of the two bioreactors.

4 Discussion

We presented a method to enrich a large number of temperature-sensitive variants of an essential enzyme. The single-cell growth rate reporter TIMER was key to enrich thousands of potentially temperature-sensitive variants because it allowed us to isolate a small fraction of slow- or non-growing cells (1.1%) from a population of mainly growing cells. The ability to create temperature-sensitive enzymes en masse opens up opportunities to comprehensively map mutations that confer thermal sensitivity, for example by deep sequencing of all enzyme variants (Bassalo et al., 2018; Garst et al., 2017). Such approaches would advance our understanding about principles that underlie thermal sensitivity of enzymes (Leuenberger et al., 2017) and enable the prediction of temperature effects based on protein sequences (Chakshusmathi et al., 2004).

Apart from mapping mutations that confer thermal sensitivity, enriching a large number of temperature sensitive enzymes facilitates the identification of variants that are optimal for a particular application. For example, it could be possible to find enzymes that are partially inactive at a given temperature or enzymes that switch quickly forth and back upon temperature changes. Such gradual and reversible control of an enzyme by temperature would be an alternative to knockdown methods like CRISPR interference (S. Li et al., 2016; Larson et al., 2013; Qi et al., 2013), which requires expression of additional protein and RNA components in the host.

Our case study was the arginine biosynthesis enzyme argininosuccinate synthetase (ArgG), and we applied it for the overproduction of citrulline. Our data suggested that the ArgG variants have indeed different temperature-dependencies and that the temperature affects the ArgG catalyzed reaction, and consequently arginine biosynthesis, in a gradual and tunable way. This precise control of metabolic reactions and pathways renders temperature-sensitive enzymes an effective tool to implement metabolic valves in overproduction strains.

Recent computational studies identified targets that are particularly suited as metabolic valves in two-stage bioprocesses (Venayak et al., 2018), and the methods presented in this study can help implementing them in production strains. So far, the method is limited to valves that are essential proteins because selection of temperature-sensitive variants is dependent on growth. A solution to this problem is modifying strains such that non-essential targets become essential, for example by deletion of isoenzymes.

Creating many protein variants with different temperature-characteristics is especially important because computational analysis suggested that multiple metabolic valves

are necessary to achieve optimal dynamic metabolic control for certain products (Burg et al., 2016; Venayak et al., 2018). A practical problem in implementing multiple metabolic valves is probably to coordinate their switching behavior with a single input signal like temperature. To this end, a large panel of temperature-sensitive enzymes, each with different temperature dependencies, will help addressing this problem and finding the optimal combination of valves.

We showed that temperature-sensitive ArgG functions as a metabolic valve and that it switches the arginine pathway between synthesis of the amino acid end-product (arginine) and synthesis of the intermediate (citrulline). The ArgG deletion strain revealed the consequences of completely closing the metabolic valve: a growth arrest and production of citrulline. Removing allosteric feedback inhibition and transcriptional regulation in the arginine pathway increased citrulline production about 3-fold indicating that even in the absence of arginine the pathway is feedback inhibited. This basal repression results either from the remaining arginine levels (e.g. from protein degradation) or inhibitory effects of other metabolites. For example, lysine has been shown to be an additional activator of the arginine repressor ArgR (Lempp et al., 2019). Temperature-sensitive ArgG allowed us to establish a two-stage bioprocess, in which the metabolic valve was open at 33°C and closed at 39°C. In the future, it will be important to clarify if partially closing the metabolic valve results in better overproduction and avoids premature decreases of production rates (here after 44 h in the ArgG deletion strain). Another option to maintain a high productivity might be switching continuously between different temperatures. Such an approach requires temperature-sensitive enzymes that switch reversibly between an active and inactive state.

Temperature gradients and fluctuations in large-scale fermentations might be problematic if they cause an unintended growth arrest during the growth phase (e.g. if zones with 42°C exist during the 33°C growth phase). Therefore, future studies could examine the function of temperature-sensitive enzymes under fluctuating temperatures. Economically feasible production of citrulline at an industrial-scale would require higher titers of citrulline, which can be realized by prolonging both growth and production phases. Further improvements can be achieved by genomic integration of temperature-sensitive variants, expression of exporters, and elimination of competing pathways (e.g. the putrescine pathway).

In conclusion, temperature-sensitive enzymes are a promising tool for metabolic engineering that enable dynamic control of metabolism. In case of temperature-sensitive ArgG, the biggest advantage is the ability to gradually switch the arginine biosynthetic

pathway on and off. Thus, together with thermo-sensitive transcription or translation, temperature-sensitive enzymes open up novel applications and process strategies in industrial biotechnology.

Funding

This work was supported by the ERC starting grant 715650.

CRedit authorship contribution statement

Thorben Schramm: Conceptualization, Investigation, Visualization, Project administration, Writing - original draft, Writing - review & editing, Formal analysis. Martin Lempp: Investigation. Dominik Beuter: Investigation. Silvia González Sierra: Investigation, Resources. Timo Glatter: Investigation, Resources, Formal analysis. Hannes Link: Conceptualization, Supervision, Project administration, Visualization, Writing - original draft, Writing - review & editing, Formal analysis, Funding acquisition.

Declaration of competing interest

None.

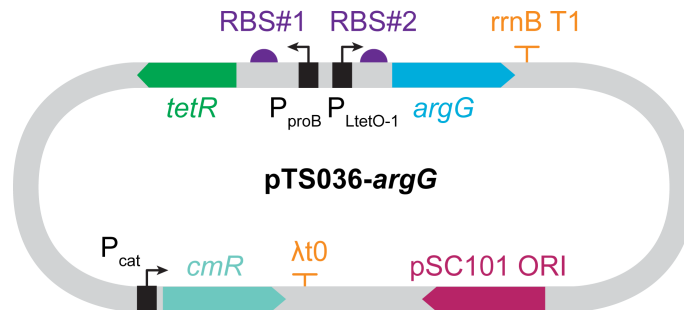
Acknowledgements

We thank L. Søgaard-Andersen and I. Bischofs-Pfeifer for discussions.

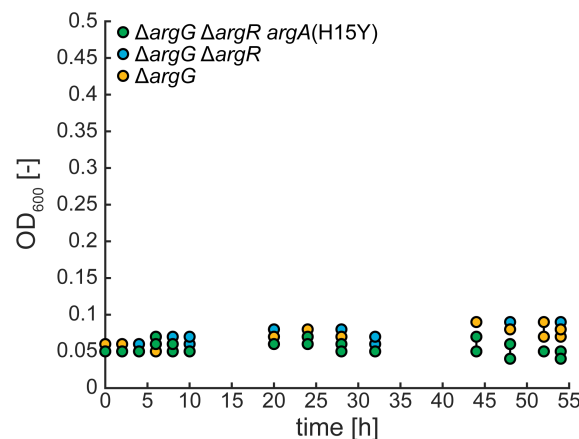
Appendix A. Supplementary data

Supplementary data to this article can be found online at
<https://doi.org/10.1016/j.ymben.2020.03.004>

Appendix B. Supplementary figures



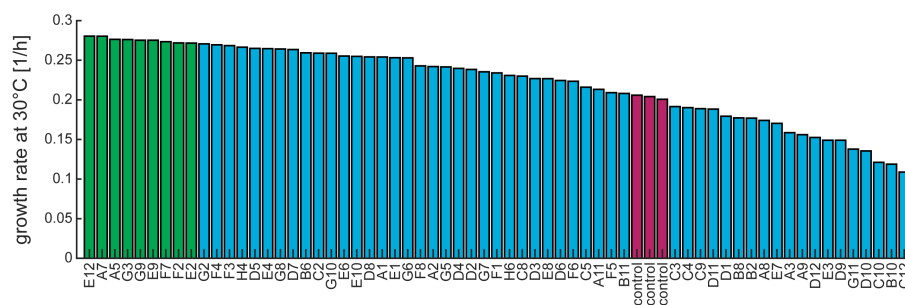
Suppl. Fig. 4.7: Plasmid map of pTS036-*argG*. RBS# 1 and P_{proB} =ribosomal binding site and promoter of *proB* (Davis et al., 2011). RBS# 2=strong ribosomal binding site (Elowitz and Leibler, 2000). $P_{LtetO-1}$ = TetR inducible promoter (Lutz and Bujard, 1997). Backbone derives from pUA66-*rrnBp* (Zaslaver et al., 2006). *tetR* and *cmR* derived from pdCas9 (Addgene # 44249) (Qi et al., 2013).



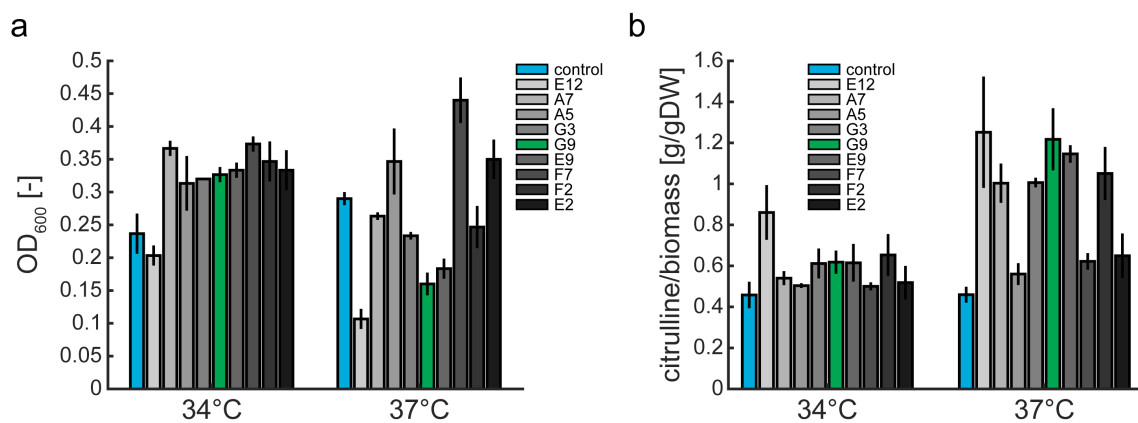
Suppl. Fig. 4.8: Growth of *argG* deletion strains during arginine starvation. Cells were first grown in minimal medium supplemented with arginine. Then, arginine was removed by washing, and cells were resuspended in minimal medium without arginine. The graph depicts the time course of the OD after the washing.



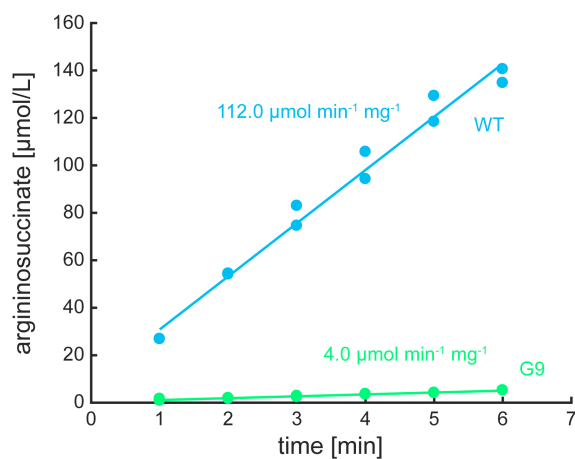
Suppl. Fig. 4.9: Growth of 90 randomly selected strains from the enriched *argG* library. Strains were cultivated in minimal medium for 20 h at 30°C, re-diluted by a factor of 30, and further incubated for 20 h at 42°C. 62 strains (blue) did not reach an OD of 0.5 after cultivation at 42°C. A control strain with wild-type ArgG is indicated in red.



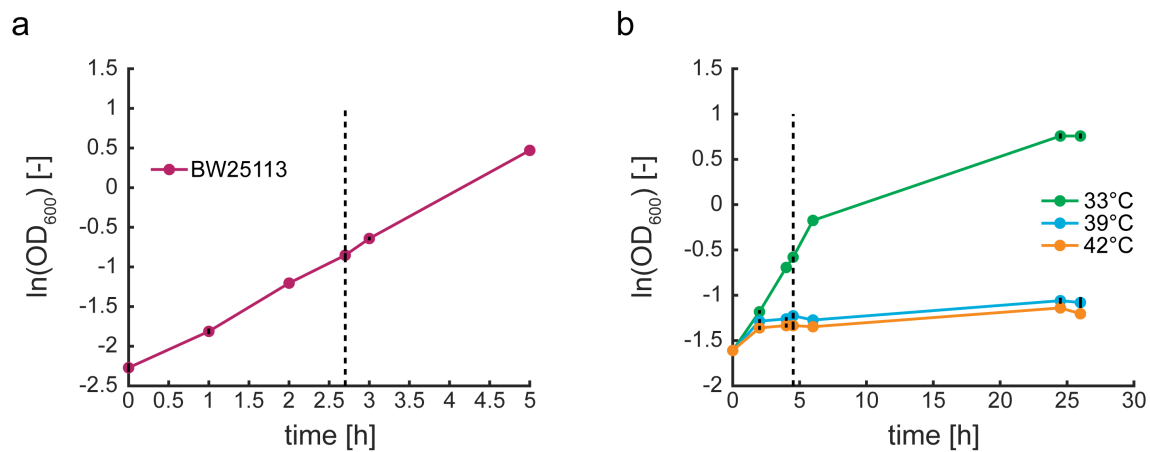
Suppl. Fig. 4.10: Growth rates at 30°C of 90 randomly selected strains from the enriched *argG* library. The nine fastest growing strains (green bars) were sequenced. A control strain with wild-type ArgG is indicated with red bars.



Suppl. Fig. 4.11: Growth and citrulline production of the doubly dysregulated citrulline producer in combination with the nine ArgG variants at intermediate temperatures. (a) Shown is the OD after 7 h of cultivation in minimal medium at 34°C and 37°C. Error bars show the standard deviation of n=3 cultures. (b) Biomass specific citrulline concentrations after 7 h of cultivation in minimal medium at 34°C and 37°C. Same strains as in (a). Error bars show the standard deviation of n=3 cultures.



Suppl. Fig. 4.12: Specific enzyme activity of *E. coli* wildtype ArgG (WT) and ArgG variant G9 at 42°C. Enzymes were incubated for 1 h at 42°C before starting the reaction.



Suppl. Fig. 4.13: Growth of (a) wildtype *E. coli* and (b) the doubly dysregulated citrulline overproducer strain with the ArgG variant G9. Dotted lines indicate the time, at which samples for proteomics were taken. Dots are means, and error bars show the standard deviation of n=3 cultures.

Supplementary Methods

S1. Sample preparation for proteomics

Samples for proteomics were prepared by transferring 109 cells to a 15 mL reaction tube and centrifugation (4°C, 4000 rpm, 10 min). The supernatant was removed, and the cell pellet resuspended in 10 mL 4°C-cold phosphate buffered saline (PBS, 8 g/L NaCl, 0.2 g/L KCL, 0.2 g/L KH₂PO₄, 1.15 g/L Na₂HPO₄, pH 7.4). After centrifugation (4°C, 4000 rpm, 10 min), the cell pellet was resuspended in 1 mL 4°C-cold PBS and transferred to 2 mL reaction tubes. Cells were pelleted by centrifugation (4°C, 17.000 g, 5 min), the supernatant removed, and the pellet stored at -80°C. 300 µL lysis buffer (2% Na-lauroyl sarcosinate in 100 mM NH₄HCO₃) were added to the cell pellet, and incubated at 90°C for 16 min under shaking at 1200 rpm. Remaining cell pellet was resuspended by 25 s of sonication. Samples were centrifuged (15.000 rpm, 20°C, 5 min). The protein concentration in the supernatant was determined with a BCA Protein Assay Kit (Thermo Fisher, # 23252). 7.5 µL of 5 mM tris(2-carboxyethyl)phosphine (TCEP) were added to the samples and incubated at 90°C for 10 min under shaking at 1200 rpm. After cooling of the samples, 10 µL of 10 mM iodacetamide were added and incubated at 25°C for 30 min under shaking of 500 rpm. 50 µg of protein was transferred to new reaction tubes. 10 µL of 0.1 g/L trypsin were added and incubated overnight at 30°C under shaking of 750 rpm. 90 µL of 5% trifluoroacetic acid and 50 µL of 10% HCl were added and incubated for 10 min at room temperature. Samples were centrifuged for (10 min, 15.000 rpm, 4°C). The supernatant was transferred to CHROMABOND Spincolumns (Macherey-Nagel) that were conditioned with 500 µL of acetonitrile and equilibrated with 500 µL and 150 µL 0.1% TFA. After loading, the peptides were washed with 500 µL 0.1% TFA in 5:95 acetonitrile:water and eluted with 400 µL 0.1% TFA in 50:50 acetonitrile:water. Peptides were concentrated and dried under vacuum at 50°C. Peptides were dissolved in 100 µL 0.1% TFA by 25 s of sonication and incubation at 22°C under shaking at 1200 rpm for 5 min.

Supplementary Tables

Suppl. Table 1.a: Strains and plasmids used in this study.

Strains	Note	Reference/source
<i>Escherichia coli</i> MegaX DH10B T1R		Invitrogen, Thermo Fischer Scientific
<i>E. coli</i> DH5 α	Cat# C640003, used during cloning	Invitrogen, Thermo Fischer Scientific
<i>E. coli</i> BL21(DE3)	Cat# 18265017, used during cloning	Invitrogen, Thermo Fischer Scientific
<i>E. coli</i> BW25113	Cat# C600003, used for protein expression	(Datsenko and Wanner, 2000)
<i>E. coli</i> JW3140	parent strain of KEIO collection	(Baba et al., 2006)
<i>E. coli</i> JW3140 (-Kan)	Δ argG (also BW25113 Δ argG)	This study
<i>E. coli</i> JW3206	Δ argR	(Baba et al., 2006)
<i>E. coli</i> JW3140 Δ argR	Δ argG Δ argR	This study
<i>E. coli</i> JW3140 Δ argR argA(H15Y)	Δ argG Δ argR argA(H15Y)	This study
<i>E. coli</i> JW3140/pBR322_T1MER	Used for transformation with the argG plasmid library	This study
<i>E. coli</i> JW3140/pBR322_T1MER/pT5036-argG	non-screened argG library generated by error-prone PCR control in screening experiment	This study
<i>E. coli</i> JW3140 Δ argR argA(H15Y)/pBR322_T1MER/pT5036-argG	variants as mentioned in the text	This study
<i>E. coli</i> JW3140 Δ argR argA(H15Y)/pBR322_T1MER/pT5036-argG	dysregulated crutline overproduction strain with wild-type argG	This study
<i>E. coli</i> JW3140 Δ argR argA(H15Y)/pT5036-argG	overexpression of wildtype ArgG	This study
<i>E. coli</i> BL21(DE3)/pT5049	overexpression of ArgG variant G9	This study
<i>E. coli</i> BL21(DE3)/pT5050		
Plasmids	Note	Reference/source
pBR322_T1MER	single cell growth reporter	(Claudi et al., 2014)
pCP20	curation of kanamycin resistance from KEIO strain	(Datsenko and Wanner, 2000)
pKDsRNA-argA(H15Y)	sgRNA targeting argA near-H15	(Sander et al., 2019)
pKDsRNA-p15	curation of the Cas9 plasmid	Addgene plasmid # 62656 (Reisch and Prather, 2015)
pCas9-CR4	Cas9 plasmid for noSCAR	Addgene plasmid # 44249 (Qi et al., 2013)
pDCas9	template for cloning only	(Zaslaver et al., 2006)
pUA66-rmbp	template for cloning only	(Kitagawa et al., 2006)
pCA24N-argR	pSC101, cmR, tetR under control of <i>prof</i> -promoter, wild-type <i>E. coli</i> argG under control of pLtetO	This study
pT5036-argG	plasmid library with mutagenized argG variants as mentioned in the text	This study
pT5036-argG(mutant)		This study
pT5036-# (variant)		This study
pT5036-G9	pCA24N derivative with <i>E. coli</i> wild-type ArgG	This study
pT5049	pCA24N derivative with ArgG variant G9	This study
pT5050		

Suppl. Table 1.b: Oligonucleotides used in this study.

Oligonucleotides	Identifier	Use
ATGACGACGATTTCTGAGCATCTC	<i>argG-amp-F</i>	pTS036- <i>argG</i> , pTS036-G9
TTACTGGCCCTTGTTTTCAGATTTC	<i>argG-amp-R</i>	pTS036- <i>argG</i> , pTS036-G9
AATCTGAAAACAAAGGCCAGTAAATGTGGATCCCGACCTGCAGGCATG	<i>argG-psC101-F</i>	pTS036- <i>argG</i> , pTS049
GAGATGCTTGAACATGTCGTCATGATGATCTCTTCTCTTAAATCTTTTGAATTTCTTGATC	<i>argG-psC101-R</i>	pTS036- <i>argG</i>
TTACTGGCCCTTGTTTTCAGACTTC	<i>G9-ampR</i>	pTS036-G9, pTS050
GCCTTCTCCTGCTTCCCTTAAGCCG	<i>argG-F</i>	sequencing
GAGACTACCGGTTGGATGCAAAACCATG	<i>argG-R</i>	sequencing
GATAAAGTAAAGATGTTGCTCCGCTGCCG	<i>argR-F</i>	sequencing
GTATTCAITTTGTTGATACATGTCGCAAG	<i>argR-R</i>	sequencing
GTGGTAAAGGAACTAAAAACCAGATTTG	<i>argA-F</i>	sequencing
TTAGCCCTAATCCCGCATCAACAC	<i>argA-R</i>	sequencing
GTGTTAAAGGAACGTAAAAACCAGATTGTTGTCGAGGATTCGCCGTATTCCTATATCATACCAACCGGGGAA	<i>argA_H151_rec2</i>	nosCAR <i>argA</i> (H151) (Sander et al., 2019)
ACGACGATTTCTAAGCATCTCCG	<i>argG-moATG-F</i>	pTS049, pTS050
TCTGGAAAACAAAGGCCAGTAAAGGCCATATGCCGCCGCTAAAGGTCGACCTG	<i>ASKA-His-ArgG-F</i>	pTS049, pTS050
GAGATGCTTGAAGATGCTGTCGCCCTCAAGGCCCGGATCCGTAATGGTG	<i>ASKA-His-ArgG-R</i>	pTS049, pTS050

Suppl. Table 4.2: Mutations and predicted $\Delta\Delta G$ values of 9 ArgG variants.

ArgG variant	Non-synonymous mutations	No. of synonymous mutations	$\Delta\Delta G^*$ [kcal mol ⁻¹]
E12	Y39H, R59S, Q152H, V223A, K252E, P384L	3	6.1
A7	L29Q, Q86L, I163F, I356N	1	4.4
A5	E62D, I356F	0	8.8
G3	F100Y, D164V, E235D, G269S, I311F, S339P	1	9.0
G9	L29P, E228V, R416H, Q438R, N441S	6	6.0
E9	E66D, L308M, N368D	1	1.1
F7	G272S, T375S	2	7.5
F2	A36V, L210H, M260V, F337L	1	9.9
E2	K34R, N92D, M197V, S229C, E235D, F420Y	2	5.3

* $\Delta\Delta G = \Delta G_{mutant}^{fold} - \Delta G_{wildtype}^{fold}$, estimated with FoldX 5.0 (Guerois et al., 2002; Schymkowitz et al., 2005) using a re-refined ArgG crystal structure (1k92 from PDB_REDO) (Lemke and Howell, 2001; Joosten et al., 2011).

Suppl. Table 4.3: Citrulline concentrations, biomass-specific citrulline yields, and citrulline production rates in different experiments.

	Doubly dysregulated ArgG deletion strain (Fig. 4.3)(n=2)	G9-based strain at 42°C (Fig. 4.4)(n=3)	G9-based strain in bioreactor (Fig. 4.5)(n=2)
citrulline concentrations [g L ⁻¹]	0.024 ± 0.003 (after 10 h)	0.033 ± 0.002 (after 7 h)	3.09 ± 0.23 (after 45 h)
yields [g g _{DW} ⁻¹]	1.32 ± 0.18 (after 10 h)	1.35 ± 0.14 (after 7 h)	5.48 ± 0.84 (after 45 h)
production rates [mmol g _{DW} ⁻¹ h ⁻¹]	0.73 ± 0.07 (first 10 h)	1.10 ± 0.12	1.01 ± 0.13 (only production phase)

References

- Schramm, T., Lempp, M., Beuter, D., Sierra, S. G., Glatter, T., & Link, H. (2020). High-throughput enrichment of temperature-sensitive argininosuccinate synthetase for two-stage citrulline production in *e. coli*. *Metabolic Engineering*. <https://doi.org/10.1016/j.ymben.2020.03.004> (cited on page 99).
- Brockman, I. M., & Prather, K. L. J. (2015). Dynamic metabolic engineering: New strategies for developing responsive cell factories. *Biotechnology Journal*, *10*(9), 1360–1369. <https://doi.org/10.1002/biot.201400422> (cited on page 100).
- Burg, J. M., Cooper, C. B., Ye, Z., Reed, B. R., Moreb, E. A., & Lynch, M. D. (2016). Large-scale bioprocess competitiveness: The potential of dynamic metabolic control in two-stage fermentations. *Current Opinion in Chemical Engineering*, *14*, 121–136. <https://doi.org/10.1016/j.coche.2016.09.008> (cited on pages 100, 123).
- Klamt, S., Mahadevan, R., & Hädicke, O. (2018). When do two-stage processes outperform one-stage processes? *Biotechnology Journal*, *13*(2), 1700539. <https://doi.org/10.1002/biot.201700539> (cited on page 100).
- Lalwani, M. A., Zhao, E. M., & Avalos, J. L. (2018). Current and future modalities of dynamic control in metabolic engineering. *Current Opinion in Biotechnology*, *52*, 56–65. <https://doi.org/10.1016/j.copbio.2018.02.007> (cited on page 100).
- Li, S., Jendresen, C. B., Grünberger, A., Ronda, C., Jensen, S. I., Noack, S., & Nielsen, A. T. (2016). Enhanced protein and biochemical production using CRISPRi-based growth switches. *Metabolic Engineering*, *38*, 274–284. <https://doi.org/10.1016/j.ymben.2016.09.003> (cited on pages 100, 122).
- Mears, L., Stocks, S. M., Sin, G., & Gernaey, K. V. (2017). A review of control strategies for manipulating the feed rate in fed-batch fermentation processes. *Journal of Biotechnology*, *245*, 34–46. <https://doi.org/10.1016/j.jbiotec.2017.01.008> (cited on page 100).
- Michalowski, A., Siemann-Herzberg, M., & Takors, R. (2017). *Escherichia coli* HGT: Engineered for high glucose throughput even under slowly growing or resting conditions. *Metabolic Engineering*, *40*, 93–103. <https://doi.org/10.1016/j.ymben.2017.01.005> (cited on page 100).
- Tokuyama, K., Toya, Y., Matsuda, F., Cress, B. F., Koffas, M. A. G., & Shimizu, H. (2019). Magnesium starvation improves production of malonyl-CoA-derived metabolites in *Escherichia coli*. *Metabolic Engineering*, *52*, 215–223. <https://doi.org/10.1016/j.ymben.2018.12.002> (cited on page 100).
- Zhao, E. M., Zhang, Y., Mehl, J., Park, H., Lalwani, M. A., Toettcher, J. E., & Avalos, J. L. (2018). Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature*, *555*(7698), 683–687. <https://doi.org/10.1038/nature26141> (cited on page 100).
- Gupta, A., Reizman, I. M. B., Reisch, C. R., & Prather, K. L. J. (2017). Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nature Biotechnology*, *35*(3), 273–279. <https://doi.org/10.1038/nbt.3796> (cited on page 100).
- Soma, Y., Tsuruno, K., Wada, M., Yokota, A., & Hanai, T. (2014). Metabolic flux redirection from a central metabolic pathway toward a synthetic pathway using a metabolic toggle switch. *Metabolic Engineering*, *23*, 175–184. <https://doi.org/10.1016/j.ymben.2014.02.008> (cited on page 100).
- Harder, B.-J., Bettenbrock, K., & Klamt, S. (2018). Temperature-dependent dynamic control of the TCA cycle increases volumetric productivity of itaconic acid production by *Escherichia coli*. *Biotechnology and Bioengineering*, *115*(1), 156–164. <https://doi.org/10.1002/bit.26446> (cited on pages 100, 101).

- Klinkert, B., & Narberhaus, F. (2009). Microbial thermosensors. *Cellular and Molecular Life Sciences*, 66(16), 2661–2676. <https://doi.org/10.1007/s00018-009-0041-3> (cited on page 101).
- Li, Z., Vizeacoumar, F. J., Bahr, S., Li, J., Warringer, J., Vizeacoumar, F. S., Min, R., VanderSluis, B., Bellay, J., DeVit, M., Fleming, J. A., Stephens, A., Haase, J., Lin, Z.-Y., Baryshnikova, A., Lu, H., Yan, Z., Jin, K., Barker, S., ... Boone, C. (2011). Systematic exploration of essential yeast gene function with temperature-sensitive mutants. *Nature Biotechnology*, 29(4), 361–367. <https://doi.org/10.1038/nbt.1832> (cited on page 101).
- Lovato, T. L., Adams, M. M., Baker, P. W., & Cripps, R. M. (2009). A molecular mechanism of temperature sensitivity for mutations affecting the drosophila muscle regulator myocyte enhancer factor-2. *Genetics*, 183(1), 107–117. <https://doi.org/10.1534/genetics.109.105056> (cited on page 101).
- Saluja, D., & Godson, G. N. (1995). Biochemical characterization of escherichia coli temperature-sensitive dnaB mutants dnaB8, dnaB252, dnaB70, dnaB43, and dnaB454. *Journal of bacteriology*, 177(4), 1104–1111. <https://doi.org/10.1128/JB.177.4.1104-1111.1995> (cited on page 101).
- Ben-Aroya, S., Coombes, C., Kwok, T., O'Donnell, K. A., Boeke, J. D., & Hieter, P. (2008). Toward a comprehensive temperature-sensitive mutant repository of the essential genes of saccharomyces cerevisiae. *Molecular Cell*, 30(2), 248–258. <https://doi.org/10.1016/j.molcel.2008.02.021> (cited on page 101).
- Kofoed, M., Milbury, K. L., Chiang, J. H., Sinha, S., Ben-Aroya, S., Giaever, G., Nislow, C., Hieter, P., & Stirling, P. C. (2015). An updated collection of sequence barcoded temperature-sensitive alleles of yeast essential genes. 5(9), 1879–1887. <https://doi.org/10.1534/g3.115.019174> (cited on page 101).
- Servant, P., Grandvalet, C., & Mazodier, P. (2000). The RheA repressor is the thermosensor of the HSP18 heat shock response in streptomyces albus. *Proceedings of the National Academy of Sciences of the United States of America*, 97(7), 3538–3543. Retrieved October 29, 2020, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC16275/> (cited on page 101).
- Weber, W. (2003). Conditional human VEGF-mediated vascularization in chicken embryos using a novel temperature-inducible gene regulation (TIGR) system. *Nucleic Acids Research*, 31(12), 69e–69. <https://doi.org/10.1093/nar/gng069> (cited on page 101).
- Pearce, S. C., McWhinnie, R. L., & Nano, F. E. (2017). Synthetic temperature-inducible lethal gene circuits in escherichia coli. *Microbiology*, 163(4), 462–471. <https://doi.org/10.1099/mic.0.000446> (cited on page 101).
- Nauta, A., Burg, B. v. d., Karsens, H., Venema, G., & Kok, J. (1997). Design of thermolabile bacteriophage repressor mutants by comparative molecular modeling. *Nature Biotechnology*, 15(10), 980–983. <https://doi.org/10.1038/nbt1097-980> (cited on page 101).
- Neupert, J., Karcher, D., & Bock, R. (2008). Design of simple synthetic RNA thermometers for temperature-controlled gene expression in escherichia coli. *Nucleic Acids Research*, 36(19), e124–e124. <https://doi.org/10.1093/nar/gkn545> (cited on page 101).
- Sen, S., Apurva, D., Satija, R., Siegal, D., & Murray, R. M. (2017). Design of a toolbox of RNA thermometers. *ACS Synthetic Biology*, 6(8), 1461–1470. <https://doi.org/10.1021/acssynbio.6b00301> (cited on page 101).
- Verdú, C., Sanchez, E., Ortega, C., Hidalgo, A., Berenguer, J., & Mencía, M. (2019). A modular vector toolkit with a tailored set of thermosensors to regulate gene expression in thermus thermophilus. *ACS Omega*, 4(11), 14626–14632. <https://doi.org/10.1021/acsomega.9b02107> (cited on page 101).

- Lynch, M. D., Gill, R. T., & Lipscomb, T. E. W. (2016, August 30). *Method for producing 3-hydroxypropionic acid and other products* (U.S. pat. No. 9428778B2). Retrieved August 10, 2022, from <https://patents.google.com/patent/US9428778B2/en> (cited on page 101).
- Lynch, M., Louie, M., Copley, S., Spindler, E., Prather, B., Lipscomb, M., Lipscomb, T., Liao, H., Hogsett, D., & Evans, R. (2019, July 2). *Microorganisms and methods for the production of fatty acids and fatty acid derived products* (U.S. pat. No. 10337038B2). <https://patents.google.com/patent/US10337038B2/en> (cited on page 101).
- Broekman, J. H. F. F., & Steenbakkers, J. F. (1973). Growth in high osmotic medium of an unsaturated fatty acid auxotroph of escherichia coli k-12. *Journal of Bacteriology*, *116*(1), 285–289. Retrieved October 29, 2020, from <https://jb.asm.org/content/116/1/285> (cited on page 101).
- Harder, M. E., Ladenson, R. C., Schimmel, S. D., & Silbert, D. F. (1974). Mutants of escherichia coli with temperature-sensitive malonyl coenzyme a-acyl carrier protein transacylase. *The Journal of Biological Chemistry*, *249*(23), 7468–7475 (cited on page 101).
- Russell, R. R. B., & Pittard, A. J. (1971). Mutants of escherichia coli unable to make protein at 42 c. *Journal of Bacteriology*, *108*(2), 790–798. Retrieved October 29, 2020, from <https://jb.asm.org/content/108/2/790> (cited on page 101).
- Cho, H.-S., Seo, S. W., Kim, Y. M., Jung, G. Y., & Park, J. M. (2012). Engineering glyceraldehyde-3-phosphate dehydrogenase for switching control of glycolysis in escherichia coli. *Biotechnology and Bioengineering*, *109*(10), 2612–2619. <https://doi.org/10.1002/bit.24532> (cited on page 101).
- Beuter, D., Gomes-Filho, J. V., 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*, *7*(12), 2775–2782. <https://doi.org/10.1021/acssynbio.8b00379> (cited on pages 102, 110).
- Claudi, B., Spröte, P., Chirkova, A., Personnic, N., Zankl, J., Schürmann, N., Schmidt, A., & Bumann, D. (2014). Phenotypic variation of salmonella in host tissues delays eradication by antimicrobial chemotherapy. *Cell*, *158*(4), 722–733. <https://doi.org/10.1016/j.cell.2014.06.045> (cited on pages 102, 110, 131).
- Curis, E., Nicolis, I., Moinard, C., Osowska, S., Zerrouk, N., Bénazeth, S., & Cynober, L. (2005). Almost all about citrulline in mammals. *Amino Acids*, *29*(3), 177–205. <https://doi.org/10.1007/s00726-005-0235-4> (cited on page 102).
- Kakimoto, T., Shibatani, T., Nishimura, N., & Chibata, I. (1971). Enzymatic production of l-citrulline by pseudomonas putida. *Applied Microbiology*, *22*(6), 992–999. <https://doi.org/10.1128/AEM.22.6.992-999.1971> (cited on page 102).
- Song, W., Sun, X., Chen, X., Liu, D., & Liu, L. (2015). Enzymatic production of l-citrulline by hydrolysis of the guanidinium group of l-arginine with recombinant arginine deiminase. *Journal of Biotechnology*, *208*, 37–43. <https://doi.org/10.1016/j.jbiotec.2015.05.012> (cited on page 102).
- Yamamoto, K., Sato, T., Tosa, T., & Chibata, I. (1974). Continuous production of l-citrulline by immobilized Pseudomonas putida cells. *Biotechnology and Bioengineering*, *16*(12), 1589–1599. <https://doi.org/10.1002/bit.260161203> (cited on page 102).
- Eberhardt, D., Jensen, J. V. K., & Wendisch, V. F. (2014). L-citrulline production by metabolically engineered corynebacterium glutamicum from glucose and alternative carbon sources. *AMB Express*, *4*(1), 85. <https://doi.org/10.1186/s13568-014-0085-0> (cited on page 102).

- Hao, N., Mu, J., Hu, N., Xu, S., Yan, M., Li, Y., Guo, K., & Xu, L. (2015). Improvement of l-citrulline production in *Corynebacterium glutamicum* by ornithine acetyltransferase. *Journal of Industrial Microbiology & Biotechnology*, 42(2), 307–313. <https://doi.org/10.1007/s10295-014-1561-x> (cited on page 102).
- Ikedo, M., Mitsuhashi, S., Tanaka, K., & Hayashi, M. (2009). Reengineering of a *Corynebacterium glutamicum* l-arginine and l-citrulline producer. *Applied and Environmental Microbiology*, 75(6), 1635–1641. <https://doi.org/10.1128/AEM.02027-08> (cited on page 102).
- Shinji, O., Masao, S., Shimpachi, K., Teruo, S., & Noboru, K. (1966, November 1). *Method of producing citrulline by bacterial fermentation* (pat. No. 3282794A). Retrieved October 29, 2020, from <https://patents.google.com/patent/US3282794A/en> (cited on page 102).
- Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., Liebermeister, W., Surette, M. G., & Alon, U. (2006). A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nature Methods*, 3(8), 623–628. <https://doi.org/10.1038/nmeth895> (cited on pages 103, 131).
- Lutz, R. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/i1-i2 regulatory elements. *Nucleic Acids Research*, 25(6), 1203–1210. <https://doi.org/10.1093/nar/25.6.1203> (cited on page 103).
- Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767), 335–338. <https://doi.org/10.1038/35002125> (cited on page 103).
- Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., & Mori, H. (2006). Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): Unique resources for biological research. *DNA Research*, 12(5), 291–299. <https://doi.org/10.1093/dnares/dsi012> (cited on pages 103, 131).
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Molecular Systems Biology*, 2, 2006.0008. <https://doi.org/10.1038/msb4100050> (cited on pages 103, 110, 131).
- Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, 97(12), 6640–6645. <https://doi.org/10.1073/pnas.120163297> (cited on pages 103, 131).
- Thomason, L. C., Costantino, N., & Court, D. L. (2007). *E. coli* genome manipulation by p1 transduction. *Current Protocols in Molecular Biology*, 79(1), 1.17.1–1.17.8. <https://doi.org/10.1002/0471142727.mb0117s79> (cited on page 103).
- Reisch, C. R., & Prather, K. L. J. (2015). The no-SCAR (scarless Cas9 assisted recombineering) system for genome editing in *Escherichia coli*. *Scientific Reports*, 5(1), 15096. <https://doi.org/10.1038/srep15096> (cited on pages 103, 131).
- Sander, T., Farke, N., Diehl, C., Kuntz, M., Glatter, T., & Link, H. (2019). Allosteric feedback inhibition enables robust amino acid biosynthesis in *E. coli* by enforcing enzyme overabundance. *Cell Systems*, 8(1), 66–75.e8. <https://doi.org/10.1016/j.cels.2018.12.005> (cited on pages 103, 108, 131, 132).
- Guder, J. C., Schramm, T., Sander, T., & Link, H. (2017). Time-optimized isotope ratio LC-MS/MS for high-throughput quantification of primary metabolites. *Analytical Chemistry*, 89(3), 1624–1631. <https://doi.org/10.1021/acs.analchem.6b03731> (cited on page 108).
- Glatter, T., Ludwig, C., Ahrné, E., Aebersold, R., Heck, A. J. R., & Schmidt, A. (2012). Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage ef-

- iciency of tandem lys-c/trypsin proteolysis over trypsin digestion. *Journal of Proteome Research*, 11(11), 5145–5156. <https://doi.org/10.1021/pr300273g> (cited on page 109).
- Plotkin, J. B., & Kudla, G. (2011). Synonymous but not the same: The causes and consequences of codon bias. *Nature Reviews Genetics*, 12(1), 32–42. <https://doi.org/10.1038/nrg2899> (cited on page 113).
- Lemke, C. T., & Howell, P. (2001). The 1.6 Å crystal structure of e. coli argininosuccinate synthetase suggests a conformational change during catalysis. *Structure*, 9(12), 1153–1164. [https://doi.org/10.1016/S0969-2126\(01\)00683-9](https://doi.org/10.1016/S0969-2126(01)00683-9) (cited on pages 113, 133).
- Lemke, C. T., & Howell, P. L. (2002). Substrate induced conformational changes in argininosuccinate synthetase. *Journal of Biological Chemistry*, 277(15), 13074–13081. <https://doi.org/10.1074/jbc.M112436200> (cited on page 113).
- Guerois, R., Nielsen, J. E., & Serrano, L. (2002). Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *Journal of Molecular Biology*, 320(2), 369–387. [https://doi.org/10.1016/S0022-2836\(02\)00442-4](https://doi.org/10.1016/S0022-2836(02)00442-4) (cited on pages 113, 133).
- Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., & Serrano, L. (2005). The FoldX web server: An online force field. *Nucleic Acids Research*, 33, W382–W388. <https://doi.org/10.1093/nar/gki387> (cited on pages 113, 133).
- Tokuriki, N., Stricher, F., Serrano, L., & Tawfik, D. S. (2008). How protein stability and new functions trade off. *PLOS Computational Biology*, 4(2), e1000002. <https://doi.org/10.1371/journal.pcbi.1000002> (cited on page 113).
- Rajagopal, B. S., DePonte, J., Tuchman, M., & Malamy, M. H. (1998). Use of inducible feedback-ResistantN-acetylglutamate synthetase (argA) genes for enhanced arginine biosynthesis by genetically EngineeredEscherichia coli k-12 strains. *Applied and Environmental Microbiology*, 64(5), 1805–1811. <https://doi.org/10.1128/AEM.64.5.1805-1811.1998> (cited on page 114).
- Long, C. P., Gonzalez, J. E., Feist, A. M., Palsson, B. O., & Antoniewicz, M. R. (2018). Dissecting the genetic and metabolic mechanisms of adaptation to the knockout of a major metabolic enzyme in escherichia coli. *Proceedings of the National Academy of Sciences*, 115(1), 222–227. <https://doi.org/10.1073/pnas.1716056115> (cited on page 120).
- Bassalo, M. C., Garst, A. D., Choudhury, A., Grau, W. C., Oh, E. J., Spindler, E., Lipscomb, T., & Gill, R. T. (2018). Deep scanning lysine metabolism in *Escherichia coli*. *Molecular Systems Biology*, 14(11). <https://doi.org/10.15252/msb.20188371> (cited on page 122).
- Garst, A. D., Bassalo, M. C., Pines, G., Lynch, S. A., Halweg-Edwards, A. L., Liu, R., Liang, L., Wang, Z., Zeitoun, R., Alexander, W. G., & Gill, R. T. (2017). Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nature Biotechnology*, 35(1), 48–55. <https://doi.org/10.1038/nbt.3718> (cited on page 122).
- Leuenberger, P., Ganscha, S., Kahraman, A., Cappelletti, V., Boersema, P. J., von Mering, C., Claassen, M., & Picotti, P. (2017). Cell-wide analysis of protein thermal unfolding reveals determinants of thermostability. *Science*, 355(6327), eaai7825. <https://doi.org/10.1126/science.aai7825> (cited on page 122).
- Chakshusmathi, G., Mondal, K., Lakshmi, G. S., Singh, G., Roy, A., Ch., R. B., Madhusudhanan, S., & Varadarajan, R. (2004). Design of temperature-sensitive mutants solely from amino acid sequence. *Proceedings of the National Academy of Sciences*, 101(21), 7925–7930. <https://doi.org/10.1073/pnas.0402222101> (cited on page 122).

- Larson, M. H., Gilbert, L. A., Wang, X., Lim, W. A., Weissman, J. S., & Qi, L. S. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols*, *8*(11), 2180–2196. <https://doi.org/10.1038/nprot.2013.132> (cited on page 122).
- Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, *152*(5), 1173–1183. <https://doi.org/10.1016/j.cell.2013.02.022> (cited on pages 122, 126, 131).
- Venayak, N., von Kamp, A., Klamt, S., & Mahadevan, R. (2018). MoVE identifies metabolic valves to switch between phenotypic states. *Nature Communications*, *9*(1), 1–9. <https://doi.org/10.1038/s41467-018-07719-4> (cited on pages 122, 123).
- Lempp, M., Farke, N., Kuntz, M., Freibert, S. A., Lill, R., & Link, H. (2019). Systematic identification of metabolites controlling gene expression in *e. coli*. *Nature Communications*, *10*(1), 4463. <https://doi.org/10.1038/s41467-019-12474-1> (cited on page 123).
- Joosten, R. P., te Beek, T. A. H., Krieger, E., Hekkelman, M. L., Hooft, R. W. W., Schneider, R., Sander, C., & Vriend, G. (2011). A series of PDB related databases for everyday needs. *Nucleic Acids Research*, *39*, D411–D419. <https://doi.org/10.1093/nar/gkq1105> (cited on page 133).

Chapter 5

A pooled CRISPR screen and metabolomics map temperature-sensitive *Escherichia coli* alleles enabling dynamic control of metabolism

Thorben Schramm, Vanessa Pahl, Hannes Link

Manuscript in preparation (*Stand: 12.08.2022*). Following the CRediT authorship contribution statements, my contributions included the conceptualization, investigation, visualization, project administration, software, writing - original draft, formal analysis. I adapted the bar-coded multiplex CRISPR/Cas9 system. I designed the temperature-sensitive *E. coli* strain library using my own software. I cloned and constructed the temperature-sensitive *E. coli* strain library. I performed cultivation and overproduction experiments. Vanessa Pahl constructed the *dnaX*(Leu289Gln)-based arginine overproduction strain, performed the arginine overproduction two-stage experiment, and measured the arginine levels by LC-MS/MS. I prepared samples for next generation sequencing (NGS). I measured metabolite concentrations by FI-MS and LC-MS/MS. I analysed the data, created the figures, and co-wrote the manuscript.

Abstract

Temperature-sensitive proteins are a unique tool to control engineered microbes in biomedical and biotechnological applications. Here, we screened for temperature-sensitive alleles across all 352 essential genes of *Escherichia coli* and used metabolomics for functional profiling. We created a pooled strain library by CRISPR/Cas9 genome editing, which included 15,120 *E. coli* mutants each with a different amino acid substitution in an essential protein. We measured temperature-sensitivity of the mutants in a competitive fitness assay and identified 1,045 strains with putatively temperature-sensitive alleles in 250 genes. Growth analysis of 92 isolated temperature-sensitive mutants revealed that the growth rate-temperature relationships varied greatly between mutants and that growth rates could be gradually controlled by temperature. By measuring the metabolome, we investigated the metabolic implications of the mutations. For 42 out of 80 mutants with temperature-sensitive enzymes, we confirmed the function as metabolic valves that enabled the introduction of metabolic bottlenecks. In some cases, metabolic profiles were more complex and indicated that single mutations caused secondary metabolic bottlenecks in distal pathways. Examples for this effect are folate metabolism and purine synthesis, arginine and purine synthesis (and vice-versa), as well as cysteine and methionine synthesis. In a final step, we employed a selection of temperature-sensitive metabolic valves for the overproduction of chemicals. A (non-enzymatic) temperature-sensitive mutant of *dnaX* was feasible to control growth of an arginine overproduction strain enabling a two-stage bioprocess. With this study, we consolidate temperature-sensitivity as a powerful and versatile tool for metabolic engineering.

1 Keywords

Temperature-sensitive mutants, CRISPR/Cas9, metabolic engineering, *Escherichia coli*, competitive fitness assay, metabolomics, next generation sequencing

2 Introduction

Microbial strains can be engineered to produce a wide variety of industrially relevant chemicals. Though, many bioprocesses suffer from low production yields since substrates are converted into both, biomass and the product chemical. One solution to the trade-off between biomass and product formation is to dynamically switch the metabolism of a production strain between growth and production stages (Burg et al., 2016; Cress et al., 2015; Hartline et al., 2021; Klamt et al., 2018).

To control microbial growth and product formation, different strategies have been used. Nutrient limitations are very effective to switch off growth (Chubukov et al., 2017; Sonderegger et al., 2005; Tokuyama et al., 2019) but, at the same time, they can trigger cellular stress responses and impair production rates. Chemical inducers (Soma et al., 2014), optogenetics (Zhao et al., 2018), quorum sensing systems (Gupta et al., 2017), degradation tags (Brockman and Prather, 2015), or CRISPR interference (Lv et al., 2015) were used to control metabolic pathways. However, a limitation of these methods is that they often fail to stop growth fully because they modulate enzyme abundance and residual functional enzyme can enable growth.

As alternative, temperature-sensitive proteins can be used to control microbial growth and metabolism by temperature (M. D. Lynch et al., 2016; M. Lynch et al., 2019; Schramm et al., 2020). For example, a temperature-sensitive transcription factor has been used to control of gene expression in chicken embryo cells (Weber, 2003). Large yeast libraries have been constructed and collected with hundreds of temperature-sensitive alleles (Ben-Aroya et al., 2008; Kofoed et al., 2015) that were pivotal in understanding gene-gene interactions (Li et al., 2011). So far, many temperature-sensitive alleles have been found for *E. coli*. Very often, sequencing data is however missing or mutants contain multiple mutations. Finding temperature-sensitive mutations can be tedious, and many methods rely on random mutagenesis approaches (Schramm et al., 2020; Ben-Aroya et al., 2010).

Here, we used a barcoded CRISPR-Cas9 genome editing method to generate around 46 mutants in each of 346 essential genes. In a competitive fitness assay, we screened for temperature-sensitive mutants. After isolating 92 temperature-sensitive strains, we analyzed their function by metabolomics. We show gradual and dynamic control of growth and endogenous as well as heterogeneous overproduction of metabolites for a set of example strains.

3 Results

3.1 A CRISPR library with temperature-sensitive variants of all essential genes in *E. coli*.

We used a CRISPR-Cas9 method (Garst et al., 2017) to create a library of temperature-sensitive *E. coli* strains. As a starting point, we selected 352 genes that are essential for growth on minimal glucose medium (Goodall et al., 2018; Patrick et al., 2007). For each of the 352 essential genes, we predicted amino acid substitutions that had a high probability of increasing temperature-sensitivity of the respective protein (Tan et al., 2014; Varadarajan et al., 1996). The CRISPR-Cas9 genome editing method (Garst et al., 2017) uses 85 bp-long homologous DNA sequences (integration cassette) to insert the desired mutations into the *E. coli* genome. Because the distance between the PAM site and the mutation site affects recombination efficiency, we ranked all the desired amino acid substitutions based on this distance but also the number of possible amino acid substitutions and off-targets. We then selected the 10 best predicted amino acid substitution sites with each 5 substitutions (Fig. 5.1.a). This would yield 50 amino acid substitutions per essential gene.

For 154 genes, we found less than 50 substitutions, either because the number of predictions for amino acid substitutions were limited or only cassettes with putatively poor recombination efficiencies were available. For example, the algorithms predicted no temperature sensitive alleles in *rpmA* and *rpmH*, which encode small (85 aa and 45 aa) ribosomal proteins. In another set of genes (*leuL*, *rplU*, *rpmC*, and *rpsI*) we found only cassettes with putatively poor recombination efficiencies available and were discarded.

In total, our CRISPR library contained 16,038 member and covered 346 genes with a mean of 46.4 amino acid substitutions per gene, 9.28 amino acid substitution sites per gene, and a mean of 4.91 substitutions per amino acid substitution site. Out of the 346 essential genes, the largest fractions of genes were part of amino acid metabolism (69 genes, 19.9 %), and cofactor and prosthetic group biosynthesis (58 genes, 16.8 %) (Fig. 1.b). With 3.2 % to 3.8 % library shares, the smallest fractions covered genes in central metabolism, cell division, or were translation associated.

Using a 200 bp array-synthesized oligonucleotide pool, we first cloned the 16,038 cassettes on plasmids. Deep sequencing of this pooled plasmid library showed that 97 % of the designed cassettes were present. During cloning of the plasmids, the library composition was conserved with exception of library members targeting ribosomal subunits. These members were diluted from 10.4 % share in the library before cloning to 4.7 % after

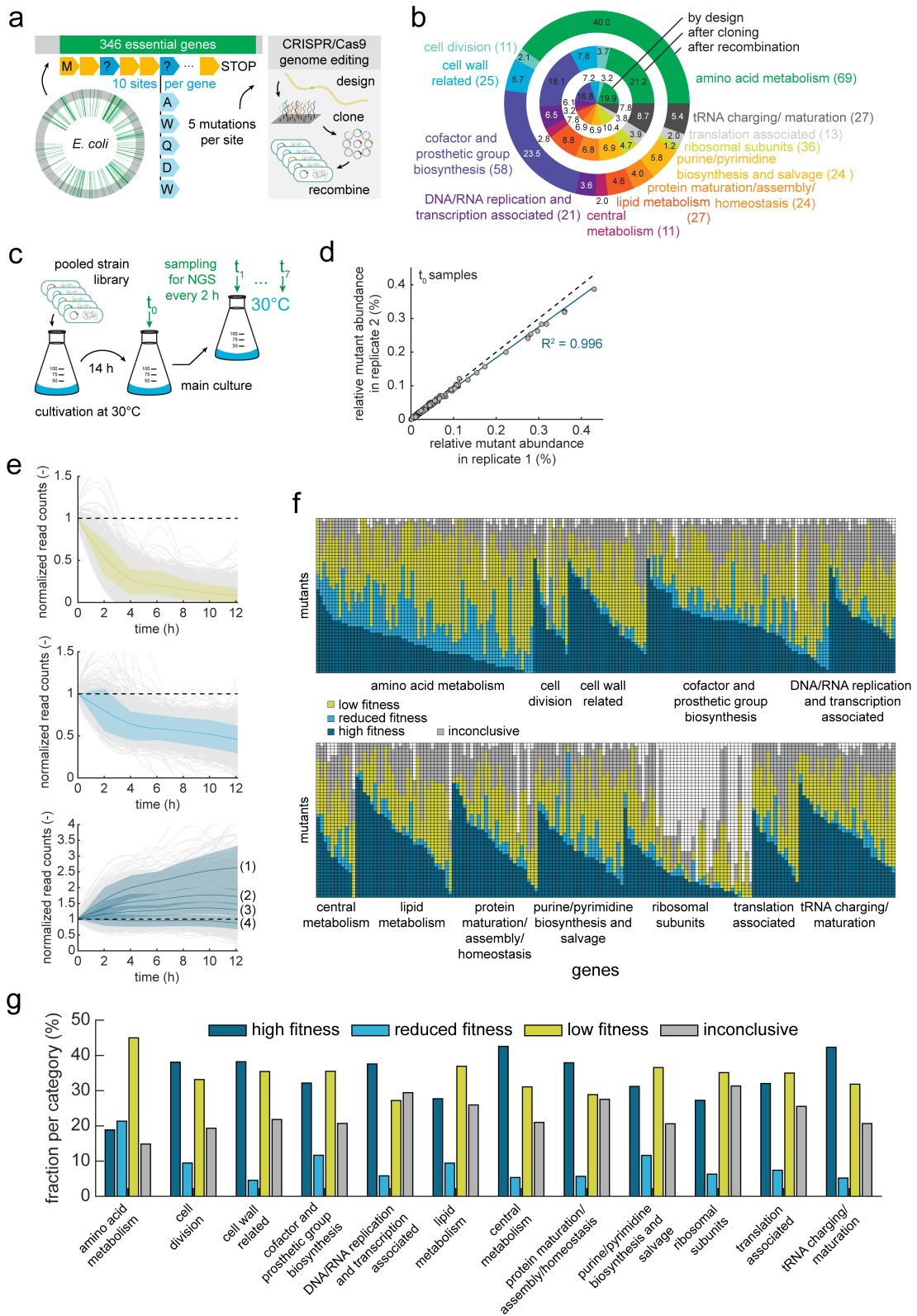


Figure 5.1: The figure caption is on the next page.

Figure 5.1: A pooled CRISPR library of *E. coli* strains bearing single amino acid substitutions. (a) An overview is provided showing key steps in the creation of the CRISPR library. (b) The 346 genes covered by the library were assigned to functional categories. The share of each category in the library is shown for different steps in the creation of the library. (c) The scheme shows our experimental approach of a competitive fitness assay using NGS to read out plasmid-borne barcodes. (d) The relative abundance of each strain detected by NGS in the sample at $t = 0$ in the competitive fitness assay. R^2 is Pearson's correlation coefficient between replicates ($n = 2$). (e) The normalized read counts are shown for the competitive fitness assay. Lines were calculated by smoothing through data of 7 time points and indicate the abundance changes for single strains in the library. Each subfigure shows data for strains that are assigned to different clusters with different fitness (yellow: low fitness, light blue: reduced fitness, dark blue: high fitness). Colored lines show the mean of entire clusters, the colored shades show the standard deviation. The four blue clusters containing strains with neutral or increased fitness were plotted into a single subfigure. (f) Heatmap showing which strain was assigned to which of the clusters. (g) Bar plot showing the fraction of strains within a functional gene category that had high, reduced, or low fitness. The fraction of strains with inconclusive data is grey.

cloning (Fig. 5.1.b). An explanation for this effect could be that the DNA sequences on the plasmid or the protospacer in the sgRNA, which are homologous to ribosomal genes, are toxic to the cells. Next, we transformed an *E. coli* strain, which carried a second plasmid with Cas9 and the lambda red system, with the plasmid library (Suppl. Fig. 5.6). In these strains, we induced Cas9 expression and Lambda Red mediated recombination to obtain the final CRISPR library. This library contained 15,120 of the 16,038 designed cassettes (94 %) and targeted all 346 genes that we included in the initial library design. In contrast to the cloning step, the library composition changed during recombination (Fig. 5.1.b). Mutants in amino acid metabolism were strongly enriched from 21.2 % library share after cloning to 40.0 % after recombination. Mutants in the cofactor and prosthetic group biosynthesis were slightly enriched (18.1 % to 23.5 %). Mutants in all other functional groups were diluted. A reason for the enrichment of mutants in the amino acid metabolism is that many of the respective genes are conditional essential. Mutations in the amino acid metabolism that are lethal in M9 glucose medium can often be complemented by supplementation of amino acids in the medium. During recombineering, complex media that also contain amino acids are commonly used to reduce cellular stress and maximize efficiency of the genome editing. Therefore, we expected that a large fraction of the mutants in the amino acid metabolism were in fact unviable in minimal glucose medium.

Having created a pooled strain library with mutations in 346 essential genes, we then sought to measure growth associated fitness of the 15,120 mutants at 30 °C. For this purpose, we used a growth competition assay. After cultivating the pooled strain library in duplicate for 14 h in minimal glucose medium, we started main cultures at t_0 and took samples for next generation sequencing (NGS) every 2 hours (t_1 to t_7) for a total of 12 h (Fig. 5.1.c). Oxygen and nutrient limitations in the medium were avoided by diluting the cultures back every 3 hours.

The NGS was reproducible as indicated by high correlation scores ($R^2 \geq 0.96$) between duplicates for all time points (t_0 data as example: Fig. 5.1.d). We then used the normalized read counts as proxy for the fitness and growth of the mutants. To calculate the normalized read counts, we first normalized the read counts to the total number of reads per sample, which gave the relative abundances of the mutants in a sample. Second, we normalized these values to the t_0 sample. This yielded curves that indicated how the abundance of the mutants changed in the pooled library throughout the cultivation.

Clustering of these curves helped us to classify the mutants into three categories (Fig. 5.1.e): (1) strains that had a low fitness at 30 °C and were rapidly diluted in the population, (2) strains that had reduced fitness and were slowly diluted, and (3) strains that had a high fitness and were maintained or even enriched throughout the experiment. The strains with high fitness were distributed over 4 clusters with different degrees of enrichment. 7,154 strains had less than 15 reads at t_0 , which was our minimum cut off. However, 4,178 strains were present during the cloning and recombination steps, with read counts higher than our cut off, and were then diluted to zero reads throughout the experiment. These strains were also likely to have strong fitness defects and were included. In total, 5,748 strains had low fitness (35.8 %), 1,712 had reduced fitness (10.7 %), and 5,118 strain had a high fitness (31.9 %) (Fig. 1.f). For 3,460 strains (21.6 %), the data was inconclusive, which means that the mutants were not present even during the construction of the library or their read count was too low to draw conclusions about the fitness.

We wondered whether some functional categories (Fig. 5.1.b) had more mutants with fitness defects than other categories. Among the categories, genes from the amino acid metabolism stood out. They had the lowest fractions of mutants with a high fitness or inconclusive data and the highest fractions with reduced or low fitness (Fig. 5.1.g). Mutants in the amino acid metabolism were enriched during recombineering in complex medium (Fig. 5.1.b), and we expected already that a large fraction of the respective mutants is unviable in minimal glucose medium. This idea is supported by the results from the fitness measurements.

In summary, we created a CRISPR library with 15,120 *E. coli* mutants, each with a single amino acid substitution in one of 346 essential proteins. We measured fitness of the mutants at 30 °C by tracking relative strain abundance during a competitive growth assay. 31.9 % of the mutants had a high fitness at 30 °C and grew. Another 35.8 % had a low fitness and were likely to have a defective gene product causing a growth defect. Around 10.7 % had a reduced fitness indicating partial function of the gene product and slow growth.

3.2 A time-resolved competitive fitness assay reveals temperature-sensitive mutants

Having measured the fitness of 15,120 mutants in the CRISPR library at 30 °C, we sought to identify temperature-sensitive mutants that grow at 30 °C but have a growth defect at 42 °C. To measure the fitness of the mutants at 42 °C, we extended our competitive growth experiment (Fig. 5.1.c), split up the main cultures at t_0 , and also tracked relative strain abundances during cultivation at 42 °C (Fig. 5.2.a).

With an average growth rate of 0.61 h^{-1} , the total population grew faster at 42 °C than at 30 °C (ca. 0.37 h^{-1}) (Fig. 5.2.b). These growth rates were similar to wild type *E. coli* (Fig. 5.2.c) and indicated that a considerable fraction of the CRISPR library was not temperature-sensitive. Most mutants that had a low or reduced fitness already at 30 °C (Fig. 5.1.e) also had a reduced fitness at 42 °C (Fig. 5.2.d). Mutants with a high fitness at 30 °C showed much more diverse phenotypes at 42 °C. Apart from strains with different degrees of high fitness at both temperatures, we also detected mutants in the library whose data indicated temperature-sensitivity. Many of these temperature-sensitive strains showed very different dynamics in the relative read counts at 42 °C (Fig. 5.2.e). Some mutants like *dnaN*(Val70Trp) were immediately depleted from the population after shifting the cultures to 42 °C while others like *panB*(Leu42Gln), *carA*(Phe266Pro), *secA*(Val71Pro), and *accC*(Val335Gln) were depleted later. These strains showed also different dynamics at 30 °C (Fig. 5.2.e), and *panB*(Leu42Gln) had already a mild growth defect at 30 °C while *carA*(Phe266Pro) was enriched in the population under this condition.

Apart from differences in the dynamics, we also observed that the normalized read counts of some temperature-sensitive strains like *aroB*(Val273Gln) converged not to zero but to another finite value (Fig. 5.2.f). We believe that this effect is caused by a subset of strains that carry a plasmid with a correct NGS barcode but that do not have a correct

genomic edit. Because of this convergence effect due to false-positive strains but also complex dynamics in the reads counts, non-supervised approaches (clustering) provided insufficient results to detect temperature-sensitive strains. Instead, we used a set of six cut-offs to classify whether a strain was temperature-sensitive (Fig. 5.2.f). With these cut-offs, we identified 1,045 putatively temperature-sensitive strains covering 250 genes (Fig. 5.2.g). 743 of these mutants had a high fitness at 30 °C (Fig. 2.h), 297 a reduced fitness, and 5 a low fitness. Thus, the fraction of temperature-sensitive strains was highest among mutants with reduced fitness (17.4 %). In comparison, 14.5 % of the mutants with high fitness at 30 °C were temperature-sensitive.

Besides temperature-sensitivity, the competitive fitness assay also provided insight into how different mutations at a single amino acid site could affect protein function. For example, our NGS data suggested that different mutations in DnaE at Leu634 cause different phenotypes (Fig. 5.2.i). Leu634Trp likely resulted in the disruption of the protein function, Leu634Pro and Leu634Asp caused temperature-sensitivity, and Leu634Ala and Leu634Gln had no apparent effects. In the future, our pooled approach of measuring fitness of mutants with single codon mutations could be combined with protein structure analysis and machine learning (Baek and Baker, 2022; Jumper et al., 2021) to automatize and learn de novo design of proteins (Lovelock et al., 2022).

We constructed a CRISPR-Cas9 strain library that, after cloning, had 15,120 single codon mutations in 346 essential genes in *E. coli*. The library was used for a competitive fitness assay that resulted in the identification of 1,045 putatively temperature-sensitive mutants. The results indicated that mutants with a reduced fitness at 30 °C were more likely to be temperature-sensitive than strains with high or low fitness at 30 °C. The NGS data of this screen also allowed us to study how different amino acid substitutions at single sites affected the fitness of the strain.

3.3 Temperature-dependent growth of 92 mutants

Out of the previously identified 1,045 temperature-sensitive candidate strains, we chose one strain per gene, yielding a total of 250 strains, and cloned them in two independent batches using a new oligonucleotide pool. Sequencing of the library after recombineering revealed that the construction was reproducible between batches, which was indicated by R^2 values of ≥ 0.95 (Suppl. Fig. 5.7). In the strain library, each of the 250 strains was present. However, the mean relative abundance of individual strains varied between 0.0006 % and 2.53 %, with a median fraction of 0.1 % for all strains. Based on this broad distribution, isolation of every mutant from the pooled library was improba-

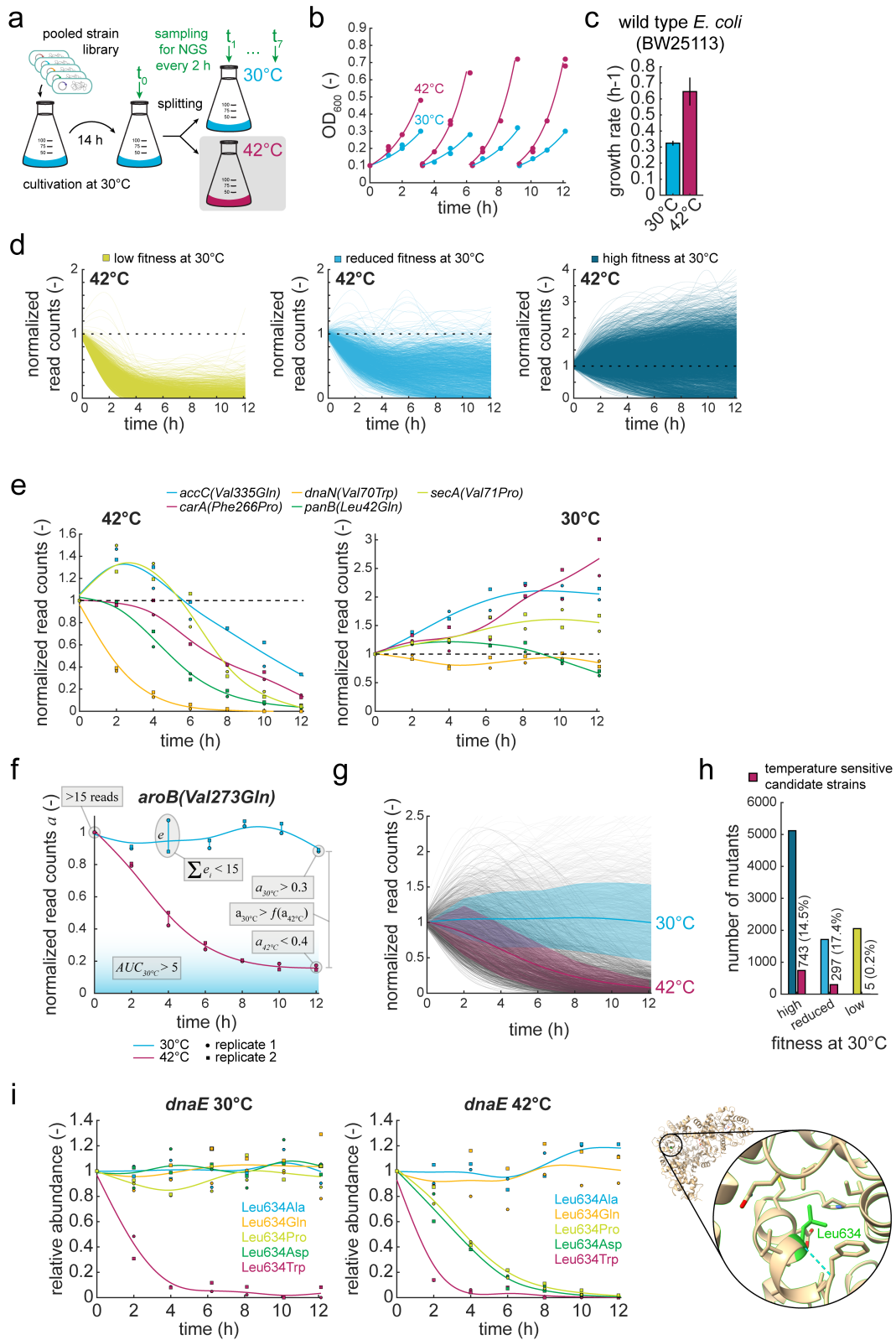


Figure 5.2: The figure caption is on the next page.

Figure 5.2: A competitive fitness assay identifies temperature-sensitive strains. (a) Scheme showing the experimental approach of the competitive fitness assay. (b) A chart showing the optical density (at 600 nm) of the pooled strain library in the competitive fitness assay at two different temperatures ($n = 2$). (c) Bar plot showing the maximum growth rate of wild type *E. coli* BW25113 at 30 °C and 42 °C. (d) Charts showing the normalized read counts of strains from three different clusters (also see Fig. 1.e) during the competitive fitness assay at 42 °C. Lines were calculated by a smoothing algorithm. (e) Charts showing the normalized read counts over the time course of the competitive fitness assay for five selected example strains at 30 °C and 42 °C. Squares indicate replicate 1, circles replicate 2. Lines are calculated by a smoothing algorithm. (f) Same as in (e). A series of cut-offs is displayed that was used to identify temperature-sensitive strains. e is an error measure and the relative difference between two replicates. $AUC_{30^{\circ}C}$ is the area under the curve for the data of a single strain at 30 °C. $a_{30^{\circ}C}$ and $a_{42^{\circ}C}$ are the normalized read counts of a strain at the indicated temperatures and at the end of the fitness assay. $f(a)$ is a custom function with $a_{42^{\circ}C}$ as input (also see section M.7). (g) Same as (d). Lines in light grey show strains at 30 °C, lines in dark grey strains at 42 °C. The blue line is the mean over all strains at 30 °C, the red line the mean over all strains at 42 °C. Shades indicate the respective standard deviations. (h) Bar plot showing the number of strains from clusters with different fitness and their respective number of temperature-sensitive strains. (i) Same as (e). The protein structure of DnaE is shown, and the location of the mutant residue highlighted.

ble. We chose to screen 2016 isolates in 96-well liquid cultures for growth phenotypes. We then Sanger-sequenced 456 strains that showed temperature-sensitivity and, finally, yielded 92 unique temperature-sensitive mutants each covering a different gene.

Next, we analyzed the growth of the mutants in minimal M9 medium at different temperatures. All strains grew at 30 °C and had a growth rate that was lower at 42 °C than for a control strain without a mutation (Fig. 5.3.a and Suppl. Fig. 5.8). At least 50 % of the mutants grew faster at 34 °C than at 30 °C but slower at 38 °C (Fig. 5.3.a). 8 strains had already strong growth defects at 34 °C, with growth rates below 0.1 1/h: *argG*(Leu114Pro), *dfp*(Leu347Asp), *gapA*(Val17Trp), *glnA*(Ile88Gln), *leuA*(Ile7Trp), *ligA*(Val121Asp), *panB*(Leu42Gln), and *panC*(Ile263Trp).

For many strains, the growth rate would first increase with increasing temperature before sharply decreasing. Similar behavior is commonly observed for enzyme activities (Daniel and Danson, 2013). Thus, multiple growth rate models have been developed before that assume that a single enzyme becomes the rate limiting step for growth while increasing temperature (Sharpe and DeMichele, 1977; Zwietering et al., 1991). In our temperature-sensitive mutants, it is very likely that the mutant proteins become indeed the single point of failure or the rate limiting step in the cells. Thus, we fitted our growth

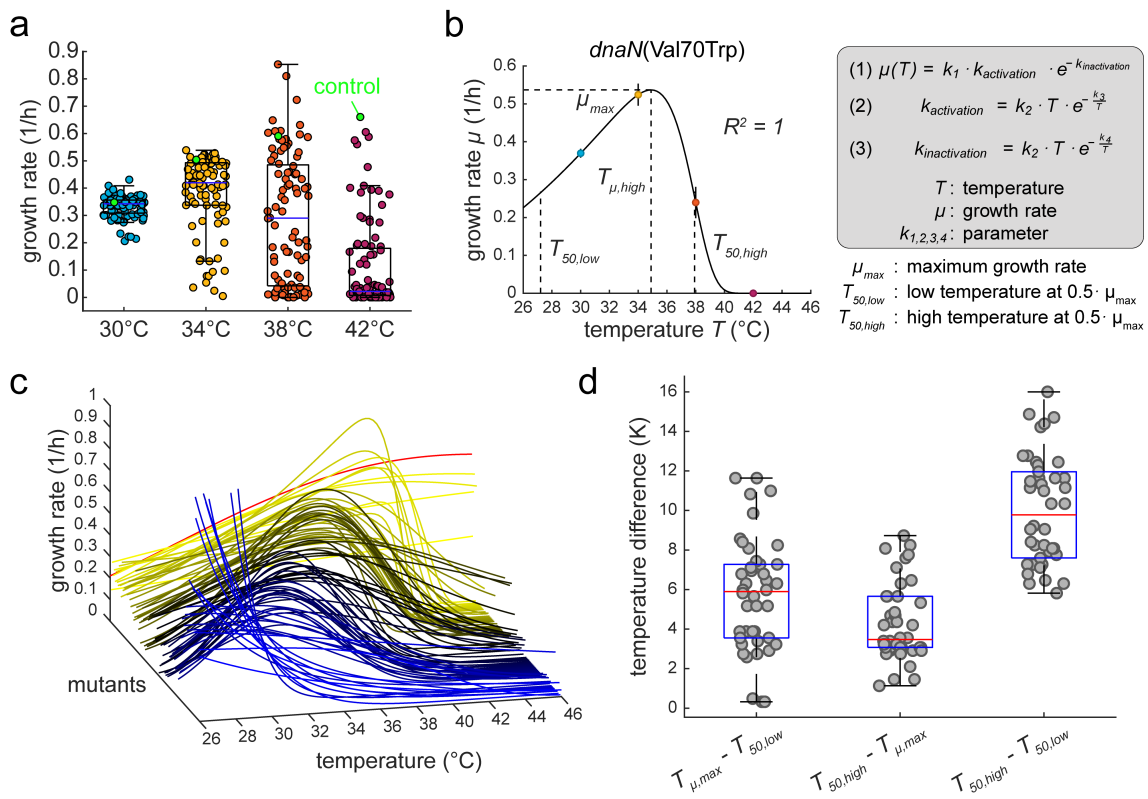


Figure 5.3: The growth rates of 92 temperature-sensitive strains as functions of temperature. (a) A dot plot showing the maximum growth rates of 92 temperature sensitive strains and a control strain (green circles) without genomic edit at different temperatures. (b) A chart showing the growth rate μ as a function of temperature T for the strain *dnaN*(Val70Trp). The function was determined using a mathematical model using three terms, which are given in the grey box. Dots in the chart indicate the mean of measured growth rates ($n = 3$). Vertical black lines are the standard deviation. R^2 is the coefficient of determination. Dotted black lines indicate the maximum growth rate and the relative temperature ($T_{\mu,max}$) as well as temperatures, at which 50 % of the maximum growth rate was reached ($T_{50,low}$ and $T_{50,high}$). (c) The same as (b) for all 93 strains except that only the fitted functions are shown. Blue indicates a low temperature, at which the maximum growth rate was reached, yellow a high temperature. The red line shows the control strain. (d) Dot plot showing temperature differences between combinations of $T_{50,low}$, $T_{50,high}$, and $T_{\mu,max}$ for 42 strains, for which both, $T_{50,low}$ and $T_{50,high}$, could be calculated.

rate data to a model that assumes Arrhenius-like increase of the growth rate at low temperatures and growth rate decreases at high temperatures as observed for enzyme inactivation (Daniel and Danson, 2013) (Fig. 5.3.b). We observed a very good fit with our data indicated by a median R^2 of > 0.987 .

Analyzing the sensitivity of the individual strains to changes in the temperature, we noticed that the growth rate of some mutants decreased sharply within a few Kelvin while other mutants had a much more gradual decrease (Fig. 5.3.c). To quantify this effect, we determined the temperatures, at which the growth rate is maximal or 50 % of the maximum growth rate ($T_{\mu,max}$, $T_{50,low}$, and $T_{50,high}$). These values indicated how sensitive the strains were to changes in temperature. In 51 cases, at least one of these values could not be determined since 50 % of the growth rate was not reached at any temperature. Yet, for the remaining 42 mutants, we observed that the $T_{50,low}$ was on median 5.9 K lower than the temperature at maximum growth rate (Fig. 5.3.d). The $T_{50,high}$ was on median only 3.5 K above the $T_{\mu,max}$. By this analysis, the three most sensitive strains were *purM*(Ile305Trp), *lpxD*(Phe141Asp), and *ppc*(Val809Asp), and they each had $T_{50,high}$ values that were < 1.5 K above the $T_{\mu,max}$. The three most insensitive strains were *accC*(VAL335GLN), *frr*(Val117Gln), *metC*(Met106Asp) with > 8 Kelvin difference between $T_{\mu,max}$ and $T_{50,high}$.

To conclude, having isolated 92 temperature-sensitive strains, we characterized the growth of the strains at different temperatures. This analysis revealed that the strains were temperature-sensitive at different degrees. In the extreme cases, strains would grow at maximum growth rate and then, almost switch-like, stop growing after increasing the temperature only by a few Kelvin. In contrast, other strains allowed a precise, gradual control of growth by increasing the temperature in single Kelvin steps. Both cases, sensitive or less sensitive switching, are equally useful for controlling metabolism dynamically, depending of the application. Thus, we continued with all 92 temperature-sensitive strains and investigated metabolic implications of the mutations.

3.4 Functional screening of 92 temperature-sensitive mutants by metabolomics

To study the function of the mutations and the effects on metabolism, we performed high-throughput metabolomics based on flow-injection mass spectrometry (Fuhrer et al., 2011). We measured 325 metabolites in the 92 temperature-sensitive strains and a control strain after prolonged incubation (16 h) in 96-well microtiter plates at 42 °C ($n = 3$).

Out of the 92 mutant genes, 80 encode for enzymes that catalyze metabolic reactions or directly use a metabolite to, for example, charge a tRNA. For reactions that are catalyzed by temperature-sensitive enzymes, we expected that substrates of the reaction accumulate and products decrease at 42 °C. For example, LysA catalyzes the reaction from meso-diaminopimelate to L-lysine. In the *lysA*(Leu280Asp) strain, we measured an increase of the substrate and a decrease of the product (Fig. 5.4.a). Since the flow-injection mass spectrometry method cannot resolve isomers, all figures are denoted without information about isomers. In the *lysA* case, the metabolic reaction is controlled by temperature and can be turned off by increasing the temperature. Thus, the mutant protein LysA(Leu280Asp) functions as a metabolic valve that can be tuned by temperature, similar to temperature-sensitive mutants of ArgG (Schramm et al., 2020). In 26 temperature-sensitive strains, either the product or the substrate changed in a similar manner to the *lysA*(Leu280Asp) strain (Fig. 5.4.b). This provided evidence that the temperature-sensitivity of these 26 strains was indeed due to the mutations in the respective genes.

We then wondered about the remaining 65 temperature-sensitive strains. In some cases, we were not able to measure the direct substrates or products of the relevant enzyme reaction. However, we could often detect metabolites that are in close proximity to the putatively perturbed reaction and that are substrates in up- or downstream reactions. In the *aroC*(Val339Pro) strain, for example, we measured an increase in shikimate phosphate, which is the substrate in the reaction one step upstream from *aroC* in the shikimate and aromatic amino acid synthesis pathway (Fig. 5.4.c). We also observed that shikimate phosphate was increased in strains with mutations in the leucine synthesis pathway. A possible explanation for this could be a regulatory crosstalk between the leucine and aromatic amino acid synthesis. Shikimate phosphate was decreased in strains with mutations in cysteine metabolism.

Another case, in which metabolites in close proximity to a perturbed reaction showed the function of the temperature-sensitive strain, was *gapA*(Val17Trp). Dihydroxyacetone phosphate (DHAP) accumulated in glycolysis instead of glyceraldehyde phosphate (Fig. 5.4.d), which is the substrate of GapA, and phosphoenolpyruvate (PEP), three reactions downstream of GapA, decreased.

Another interesting cases, in which two synthesis pathways were perturbed simultaneously, were the temperature-sensitive mutants *carA*(Phe266Pro) and *carB*(Val322As). First, we observed a decrease of orotate, downstream of the CarA and CarB catalyzed reactions in the pyrimidine synthesis pathway. Second, we measured an increase of acetyl-ornithine in the arginine synthesis pathway. The reaction product of CarA and

Figure 5.4: Functional analysis of 92 temperature-sensitive strains by metabolomics. (a) Charts displaying flow-injection mass spectrometry raw data in the mass over charge range relevant for diaminopimelate and lysine. Both metabolites are reactants of the LysA catalyzed metabolic reaction. (b) Dot plot showing the metabolomics data of 92 temperature sensitive strains and a control strain without genomic edit. The data was normalized by multiple steps including a modified z-scoring (based on median values, also see Methods part 5.10). Dots are the mean of three replicates. Blue dots indicate strains with mutations in genes that encode for enzymes that have the indicated metabolite as substrate or product. (c, d, e, and f) Same as (b). The schematics show metabolic pathways of *E. coli*. Yellow circles are metabolites, for which metabolomics data is given in one of the charts. Metabolic valves in blue were confirmed by data in (b). Metabolic valves were confirmed by more complex metabolic profiles, and according genes are shown as yellow circles in the charts.

CarB, carbamoyl phosphate, is not only used in the pyrimidine synthesis but also used in a reaction in the arginine synthesis pathway (ArgF/ArgI catalyzed). Acetyl-ornithine is upstream of this carbamoyl phosphate-using reaction. Thus, its increase indicated that the carbamoyl phosphate using reaction in the arginine synthesis pathway became a metabolic bottleneck, which lead to an accumulation of acetyl-ornithine. The connection between pyrimidine and arginine synthesis by carbamoyl phosphate already attracted attention in a metabolic engineering study, in which arginine overproduction caused a carbamoyl phosphate shortage and, thus, a limitation in purine synthesis (Sander, Wang, et al., 2019). Interestingly, in the *purK*(Ile213Asp), *purC*(Val188Asp), and *purE*(Ile29Trp) strains from the purine synthesis pathway, we measured elevated levels of acetyl-glutamate 5-semialdehyde (Fig. 5.4.e), which is also an intermediate in the arginine synthesis. This data suggested also crosstalk between the purine and arginine pathway.

In the temperature-sensitive strain *panC*(Ile263Trp), acetyl-CoA levels were decreased. PanC is, likewise to Dfp and CoaD, involved in the coenzyme A synthesis (Fig. 5.4.f). Direct substrate or product level changes showed already the function of the temperature-sensitive mutants *dfp*(Leu347Asp) and *coaD*(Leu113Trp). However, in both, *dfp*(Leu347Asp) and *coaD*(Leu113Trp), acetyl-CoA levels were also decreased like in *panC*(Ile263Trp). This suggested that a limitation in coenzyme A synthesis causes decreased acetyl-CoA levels and explained the *panC*(Ile263Trp) strain. Similar to the *carA* and *carB* mutants, *glyA*(Met201Asp) and *folA*(Met92Pro) mutants caused a metabolic bottleneck in another pathway. In these mutants, we detected three increased metabolites in the purine synthesis pathway (AICAR, AIR, and SAICAR)(Fig. 5.4.e). The three metabolites are all up-

stream of the PurH catalyzed reaction that requires a tetrahydrofolate as cofactor. Thus, a perturbed tetrahydrofolate metabolism in the *glyA* and *folA* mutants could explain why PurH became a metabolic bottleneck.

Notably, we also observed that specifically AIR accumulated strongly in the *argG*(Leu-114Pro) strain (Fig. 5.4.f). Since the substrate of the ArgG reaction (citrulline) also accumulated strongly in this strain, there was likely a metabolic bottleneck in the arginine biosynthesis pathway. Such bottleneck usually triggers metabolic feedback regulation in the arginine synthesis (Schramm et al., 2020; Sander, Wang, et al., 2019; Sander, Farke, et al., 2019), and we expected that the *argG* mutant had increased expression of CarA and CarB draining intracellular hydrogen carbonate levels. The reaction in the purine synthesis pathway that uses AIR as substrate, which accumulated in the *argG* mutant, requires hydrogen carbonate as cofactor. Thus, a depletion of hydrogen carbonate could trigger a metabolic bottleneck in purine synthesis and explain the increased AIR level.

The Asd catalyzed reaction fuels the lysine, methionine, and isoleucine synthesis pathways. We measured low lysine levels in the *asd*(Ile110Asp) strain indicating that the strain indeed had a metabolic bottleneck at the Asd reaction. During this analysis, we also noticed that lysine levels were high in the *argA*(Met173Pro) strain. This was expected for a bottleneck in the arginine pathway since lysine and arginine are structurally very similar and both interact each with the transcription factors of both pathways, ArgR and ArgP (Lempp et al., 2019; Nguyen Le Minh et al., 2018).

Lastly, metabolomics revealed that mutants in the cysteine metabolism (*cysB*(Val171Gln), *cysE*(Val226Ala), and *cysH*(Val49Gln)) caused perturbations in the methionine and threonine synthesis. We were not able to measure cysteine. However, these three strains had decreased levels of reduced glutathione (Fig. 5.4.f), which is formed in two steps from cysteine, indicating that cysteine synthesis was limited. Since cysteine is also required in the methionine synthesis pathway, a cysteine limitation explains low levels of methionine that we measured in the *cysB*(Val171Gln), *cysE*(Val226Ala), and *cysH*(Val49Gln) strains. Since homoserine and threonine are isomers, we were not able to distinguish between them with our metabolomics approach. Still, the combined value of both metabolites provided insights: increased levels of threonine/homoserine and phospho-homoserine indicated that a metabolite accumulation upstream of the reaction in the methionine synthesis pathway that uses cysteine would redirect flux into the threonine pathway. In the *cysS*(Val45Trp) strain, we expected high levels of cysteine levels that could explain the elevated levels of sulfate that we measured in this strain (Fig. 5.4.f). Although the *cysD*(Ile269Gln) strain had also low levels of reduced glutathione, we did not observe

strong changes (mod. z-score > 3) in the methionine or threonine pathways.

Apart from the temperature-sensitive strains covering enzymes, 12 strains had mutations in genes not encoding for enzymes and included: *dnaE*(Leu634Asp), *dnaN*(Val70Trp), *dnaX*(Leu289Gln), *frr*(Val117Gln), *ftsQ*(Ile74Gln), *ftsY*(Ile457Gln), *grpE*(Val190Gln), *gyrB*(Phe514Trp), *mreC*(Val238Pro), *rpoH*(Phe60Pro), *secA*(Val71Pro), and *ssb*(Val104Pro). Most of these strains had either no significant metabolic changes compared to a control strains (modified z-score > 3 or < -3) or cryptic changes. Yet, one exception was the *ssb*(Val104Pro) strain. In this strain, thymine, guanine, inosine, and guanosine levels were increased (Suppl. Fig. 5.9). Since Ssb binds ssDNA and protects it from degradation (Shereda et al., 2008), malfunction of the temperature-sensitive Ssb at 42 °C and subsequent degradation of ssDNA explains the measured increase in the thymine, guanine, inosine, and guanosine levels.

To conclude, out of 80 temperature-sensitive strains covering enzymes, we could confirm the function as a metabolic valve for 42 strains by metabolomics. The analysis revealed that in many cases (26) the substrate or product levels of perturbed reactions were directly affected. However, since not all metabolites were measurable, we relied also on inferring the function as metabolic valve from very specific metabolic pattern. By this analysis, we could confirm another 16 metabolic valves, and observed that many metabolic valves can cause secondary metabolic bottlenecks in distal pathways. The data indicated that (mis-)regulation of metabolism could be a major factor for this effect. In case of the 12 temperature-sensitive strains that had mutations in genes not encoding for enzymes, only *ssb*(Val104Pro) had a specific metabolic profile that allowed to draw a connection to the function of Ssb.

3.5 Temperature-sensitive mutants applied in metabolic engineering

Out of 92 temperature-sensitive strains, 80 had mutations in genes encoding for enzymes. We already confirmed the function as metabolic valves for 42 of the 80 mutant enzymes. A main characteristic of a metabolic valves is that, if it is closed, it introduces a bottleneck into a metabolic pathway. Usually, this is accompanied by the accumulation of the substrate of the perturbed enzyme reaction (Donati et al., 2021, 2018). If the metabolic valve is essential, not only metabolite accumulation but also growth can be controlled (Schramm et al., 2020; Harder et al., 2018; Cho et al., 2012) that enables two-stage bioprocesses with optimized product yields. Thus, we selected a subset of

our temperature-sensitive strains to test them in two-stage bioprocesses to overproduce metabolites.

Based on our previous metabolomics results (Fig. 5.4), we selected *lysA*(Leu280Asp) for the diaminopimelate production, *argG*(Leu114Pro) for the citrulline production, *aroC*-(Val339Pro) for the shikimate phosphate production, and *metA*(Phe285Trp) and *thrB*(Phe267Asp) for the homoserine/threonine production. Since homoserine is the substrate for the enzyme reactions of MetA and ThrB, we also included a strain that carried both temperature-sensitive mutations *metA*(Phe285Trp) + *thrB*(Phe267Asp). Homoserine and threonine were inseparable by LC-MS/MS, and, thus, we always reference both. After cultivating the strains at 30 °C, we started main cultures for 24 h of incubation at 42 °C. All temperature-sensitive strains did not grow at 42 °C, except the *aroC*(Val339Pro) strain that had a growth defect compared to a control strain without a mutation (Fig. 5.5.a). Throughout the experiment, we took metabolite samples from the culture broth and determined absolute metabolite concentrations by targeted LC-MS (Guder et al., 2017). Isomers cannot be distinguished by this method such that metabolite labels are given without indication of isomerism.

The *aroC*(Val339Pro) strain produced 103 $\mu\text{mol/L}$ shikimate phosphate, whereas the production rate was 28 $\mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$ in the initial 6 h. The *argG*(Leu114Pro) strain produced 186 $\mu\text{mol/L}$ citrulline, and the production rate remained stable for 24 h at 301 $\mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$ (Fig. 5.5.c). The *lysA*(Leu280Asp) strain accumulated diaminopimelate (DAP) with a production rate of 31 $\mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$ during the first 6 hours (Fig. 5.5.d). Then, the production rate declined. The final concentration of DAP was 20 $\mu\text{mol/L}$ after 24 h. The *metA*(Phe285Trp) strain produced 30 $\mu\text{mol/L}$ homoserine/threonine in 24 h, and the *thrB*(Phe267Asp) strain 189 $\mu\text{mol/L}$. A strain with both mutation produced 254 $\mu\text{mol/L}$ (Fig. 5.5.e). In all three homoserine/threonine production strains, the production rates were stable over 24 h and were 48 $\mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$, 270 $\mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$, and 477 $\mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$ for *metA*(Phe285Trp), *thrB*(Phe267Asp), and *metA*(Phe285Trp)+*thrB*(Phe267Asp), respectively.

These results showed that a wild type *E. coli* strain can be converted into an overproduction strain, whose growth can be controlled by temperature, by introduction of a single amino acid substitution. However, the measured production rates varied greatly between 28 and 477 $\mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$. To have a point of reference for these values, we used the iJO1366 genome scale model of *E. coli* (Orth et al., 2010) and calculated the theoretical metabolite flux in the amino acid synthesis pathways at exponential growth (0.6 h^{-1}).

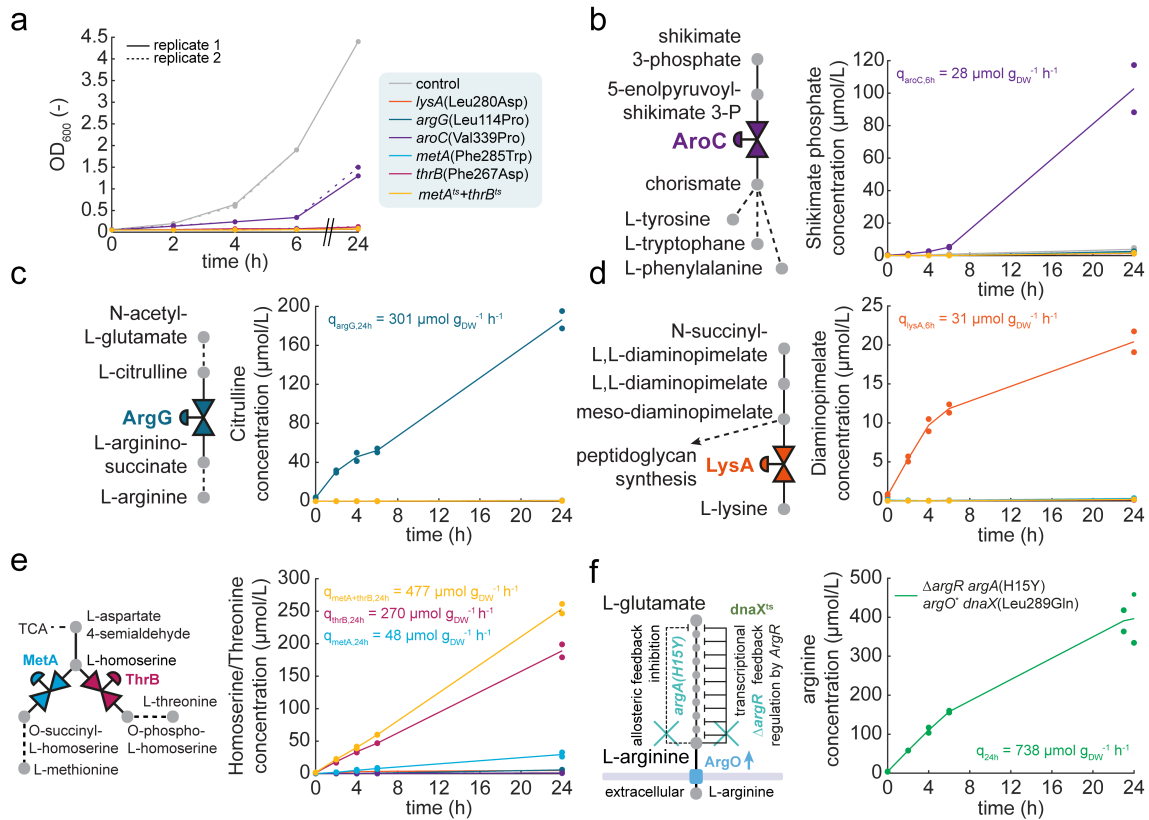


Figure 5.5: Benchmarking temperature-sensitive strains in two-stage metabolite overproduction experiments. (a) Chart showing the OD₆₀₀ of six temperature-sensitive strains (as described in the grey box) and a control strain without genomic edit during shake flask cultivations at 42 °C over 24 h. Lines and dashed lines indicate replicates. (b, c, d, e) Schemes showing parts of synthesis pathway in *E. coli*. The charts show the metabolite concentrations in the whole culture broth of strains from (a) over the time course of the cultivation. $q_{gene,time}$ is the biomass-specific production rate of a specific strain in a time range from the start of the experiment up to the indicated hour. (f) Same as (b-e). Growth data is given in Suppl. Fig. 5.10. The arginine strain is transcriptionally and allosterically dysregulated in the arginine synthesis pathway and carries a plasmid for overexpression of ArgO, an arginine exporter.

The flux in the arginine synthesis pathway is estimated to be $172 \mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$ during exponential growth. In the *argG*(Leu114Pro) strain we measured rates almost double as high during growth arrest ($301 \mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$). These measured rates were similar to production rates in an *E. coli* strain with deletion of *argG* and *argR* (Schramm et al., 2020). ArgR is the transcription factor controlling expression of the arginine pathway genes by arginine-mediated feedback regulation. Thus, we expected that expression of arginine genes were increased in the *argG*(Leu114Pro) at 42 °C, which can explain the measured production rates.

The shikimate phosphate production rate in the *aroC*(Val339Pro) strain was lower than the theoretical flux in the aromatic amino acid pathway ($222 \mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$). This was expected because the strain was still growing at reduced rate at 42 °C implying that AroC(Val339Pro) was partially active.

In case of the DAP production, we observed only $31 \mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$, which was around 7 times lower than the estimated flux in the lysine pathway at exponential growth ($200 \mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$). In contrast to the *argG* strain, the *lysA*(Leu280Asp) strain stayed well below its usual capacity at exponential growth, similarly to the *aroC*(Val339Pro) strain. However, the *lysA*(Leu280Asp) did not grow unlike the *aroC* mutant. Thus, we did not expect low production rate in the *lysA*(Leu280Asp) strain. However, there are two major differences in comparison to the *argG* strain. First, DAP is also a substrate for peptidoglycan synthesis. Second, the lysine synthesis pathway starts with a substrate (L-aspartate 4-semialdehyde) that is also used in methionine, threonine, and subsequently also isoleucine synthesis (Fig. 5.4.f). Thus, these synthesis pathways competing for the same substrates could drain flux from the lysine synthesis pathway causing the lower-than-expected production rates.

Just one enzyme reaction downstream of L-aspartate 4-semialdehyde, L-homoserine is another branching-point, from which the L-threonine and isoleucine synthesis as well as methionine synthesis pathways are fed (Fig. 5.4.f). MetA catalyzes the first step in L-methionine synthesis, and ThrB the first step in L-threonine and isoleucine synthesis. The theoretical flux in the L-methionine synthesis pathway is $90 \mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$, and $317 \mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$ in the L-threonine and L-isoleucine pathways combined. Having a bottleneck in either, ThrB or MetA, resulted in the overproduction of homoserine/threonine. Yet, the production rates were lower than the theoretical ones. In a strain with both mutations (*metA*(Phe285Trp) and *thrB*(Phe267Asp)), the production rate was however higher than the theoretical rate ($407 \mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$). This, again, indicated that flux could be redirected at a metabolic branching point resulting in reduced production

rates.

Apart from metabolic valves, we also tested a temperature-sensitive mutant of *dnaX* in the overproduction of arginine and used it to control growth by temperature. The production strain had an allosterically and transcriptionally dysregulated arginine synthesis pathway (Sander, Wang, et al., 2019) and carried a plasmid expressing the arginine exporter ArgO. Additionally, the strain carried the *dnaX*(Leu289Gln) mutation that we identified in this study. At 42 °C, the strain did not grow (Suppl. Fig. 5.10) and accumulated 478 $\mu\text{mol/L}$ arginine during 24 h. The arginine production rate was 738 $\mu\text{mol g}_{DW}^{-1} \text{h}^{-1}$, which was around three times less than in a growing arginine overproduction strain (Sander, Wang, et al., 2019). However, the biomass-specific arginine yield was, to our knowledge, with a value of 14.2 mmol/g_{DW} the highest reported so far. This showed the strength of two-stage bioprocesses to improve production yields but also showed that temperature-sensitive *dnaX* can be used to control growth of an overproduction strain.

4 Discussion

In this study, we explored temperature-sensitivity as a tool for metabolic engineering. Using a barcoded CRISPR-Cas9 genome editing method, we designed and created a pooled strain library with 15,120 *E. coli* strains covering 346 essential genes. Subsequently, we investigated the fitness of the strains by a competitive fitness assay and found that around 35.8 % of the strains had already a very low fitness at 30 °C, and 10.7 % a reduced fitness. Only 31.9 % of strains were unaffected by the mutations. In comparison, around 30 % to 40 % of spontaneous or induced mutations were lethal in yeast (Eyre-Walker and Keightley, 2007; Wloch et al., 2001). This showed that the selection of putatively temperature-sensitivity inducing mutations by the TSpred algorithm (Tan et al., 2014; Varadarajan et al., 1996) resulted in similar fractions of strains with fitness defects. An estimation of how well the prediction tool performed is difficult since reference data is not present. However, future studies investigating temperature-sensitivity could focus on site saturation mutation libraries to understand the underlying mechanisms better and improve temperature-sensitivity prediction tools.

As next step, we extended the competitive fitness assay to 42 °C and identified 1,045 putatively temperature-sensitive strains covering 250 unique genes. Our data indicated that strains that had already a mild fitness defect at 30 °C were most likely to be temperature-sensitive (17.4 %). Strains with high or neutral fitness at 30 °C were in 14.5 %

of the cases temperature-sensitive. Like in our study, in most other studies that screen for temperature-sensitive strains, not all theoretically possible mutations of a gene were tested. The main reason for this is that the number of possible mutations is usually too large for testing. For example, the average bacterial protein has 320 amino acids (Tiessen et al., 2012), which means that around 6,000 mutations are needed to be tested if only single amino acid substitutions were considered. As a consequence, it is not known how many of the single amino acid substitutions cause temperature-sensitivity for an average protein.

For each of the 250 genes with temperature-sensitive strains, we chose one strain for a new strain library, from which we isolated 92 strains. Characterization of the growth rate-temperature relationship revealed that the individual temperature-sensitive strains differed greatly between each other. The temperature, at which the strains grew with maximum growth rate, spanned over at least 10 K between the strains. The strains also showed different slopes, at which the growth rates decreased with temperature, that indicated different degrees of temperature-sensitivity. For example, for some strains the temperature with maximum growth rate and a temperature without growth were separated by only less than 4 K, whereas other strains had greater temperature differences (> 8 K in extreme cases). On median, 3.5 K separated the temperature with the maximum growth rate and a higher temperature with 50 % of the maximum growth rate. A gradual control of growth by temperature-sensitivity has been reported before (Schramm et al., 2020; Li et al., 2011; Chakshusmathi et al., 2004). Our observations on the temperature dependency of the growth rate was very similar to the enzyme activity-temperature relationship (Daniel and Danson, 2013). In metabolic engineering approaches, different switching characteristics could be used to fine tune fluxes in the metabolic network and, for examples, turn off competing pathways entirely but leaving other required pathways at reduced activity. Relying on different switching behaviors of the temperature-sensitive mutants, it would also be possible to combine multiple mutations to establish triple-stage bioprocesses with three distinct cultivation temperatures: an initial growth stage is followed by a conditioning phase with reduced growth, in which the enzymes for overproduction of a metabolite are overexpressed. In a third stage, growth is arrested fully, and only the desired product metabolite formed.

By high-throughput metabolomics, we confirmed the function of 42 temperature-sensitive enzymes as metabolic valves. 26 of the strains had either direct substrates or products of the perturbed enzyme reaction accumulating or decreasing. In further 16 strains, the function as metabolic valve was inferred from complex metabolic profiles.

The data showed that metabolic bottlenecks can cause secondary bottlenecks in distal pathways. For example, perturbations in the folate metabolism caused a bottleneck in purine synthesis. Another metabolic bottleneck in purine metabolism was observed in the *argG*(Leu114Pro) strain that can be explained by a depletion of hydrogen carbonate due to increased expression of CarAB. Perturbations in the cysteine synthesis suggested a metabolic bottleneck in methionine synthesis, which is accompanied by flux redistribution from the methionine pathway towards threonine synthesis.

Out of the 42 temperature-sensitive strains that function as metabolic valves, we selected 5 different strains and constructed one strain with two temperature-sensitive mutations for the overproduction of metabolites. Our results show the potential of the metabolic valves as a tool for metabolic engineering to precisely control metabolic reactions by temperature and to establish two-stage bioprocesses with improved product yields. However, not only the metabolic valves are useful for metabolic engineering: the temperature-sensitive mutant *dnaX*(Leu289Gln) allowed us to construct an arginine overproduction strain, whose growth can be controlled by temperature. In a two-stage bioprocess, we produced arginine with unprecedentedly high biomass-specific yields.

Taken together, our study consolidates that temperature-sensitive mutants are a versatile and powerful tool for metabolic engineering. Finding these mutants can however be tedious. With the here presented screening approach based on barcoded pooled CRISPR-Cas9 genome engineering coupled to competitive fitness assays, we were able to find 92 temperature-sensitive strains of *E. coli* that each carried a single amino acid substitution in a unique, essential gene. These mutants are a valuable resource for future systems biology and metabolic engineering attempts.

5 Material and Methods

5.1 Construction of plasmids

We adapted a CRISPR-Cas9 genome editing method that is called CREATE (Garst et al., 2017) and cloned two plasmids (pTS040 and pTS041) by Gibson assembly that resemble the function of CREATE. pTS040 had the p15A origin of replication and carried a chloramphenicol resistance gene, a cassette with the homology arm for recombination, and the guide RNA of the CRISPR system under control of a constitutive promoter (P_{J23119}). pTS041 had the pSC101 origin of replication and carried a kanamycin resistance gene, a gene for the anhydrotetracycline(aTc)-sensitive repressor *tetR*, *cas9* under control of

the aTc controlled P_{LtetO1} promoter, the arabinose-sensitive repressor *araC*, and the *Escherichia virus Lambda* genes red under control of the arabinose-controlled promoter P_{araBAD}. pTS056 had a p15A origin of replication, an ampicillin resistance gene, a gene for the anhydrotetracycline(aTc)-sensitive repressor *tetR*, and *argO* encoding for an arginine exporter under the aTc controlled P_{LtetO1} promoter. pTS056 was based on a plasmid from (Sander, Wang, et al., 2019). For cloning, Q5 High-fidelity DNA polymerase (New England BioLabs Inc., NEB) and Gibson Assembly Master Mix (NEB) was used.

5.2 Design of the temperature sensitive *E. coli* library

As targets for editing, we chose all genes (352) in *E. coli* that are known to be essential on M9 minimal medium supplemented with glucose (Goodall et al., 2018; Patrick et al., 2007). Using the TSpred online tool (Tan et al., 2014; Varadarajan et al., 1996), we gathered predictions for every target gene. It is predicted, which amino acid of a protein, upon substitution by one of the five amino acids alanine, tryptophan, glutamine, aspartate, or proline, is likely to introduce temperature-sensitivity. If possible, we used crystal structures of the target proteins as input for the algorithm. Elsewise, we used amino acid sequences. We preferred partial crystal structures over amino acid sequences. After collecting all predictions, we then checked possible designs for the homology arms for recombination.

In general, the quality of a protospacer for CRISPR-Cas9 genome editing is strongly affected by the distance of the PAM to the target site (Garst et al., 2017) and its off-targets. We considered every PAM within 30 bp distance for every predicted site and checked if a silent PAM-mutation was available. In some cases, the PAM and the target site overlapped such that not every amino acid substitution was available to disrupt the PAM. Further, we used the Cas-OFFinder (Bae et al., 2014) to check for off-targets up to 4 mismatches. We also checked, whether the protospacers had a 11 PAM-proximal perfect match to multiple PAMs, which is good practice for CRISPRi protospacer design (Rousset et al., 2018). Based on these results, we then ranked each available design for every site with a custom scoring system and chose 10 predicted sites for each gene that had the highest-ranking designs. We excluded designs that did not reach a certain minimum score such that some target genes yielded no or less than 10 available designs. The final library contained 16,038 members covering 346 genes.

Similarly as reported by (Garst et al., 2017), the final oligonucleotides in the library were in total 200 bp long and contained in the following order: a spacer sequence ('TCCTCTGGCGGAAAGCC'), a homology sequence with the desired mutation and a silent

PAM-mutation, another spacer ('GATC'), the J23119 promoter ('TTGACAGCTAGCTCAG-TCCTAGGTATAATACTAGT'), a protospacer, and a part of the sgRNA-Cas9 handle ('GTTT-TAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAG').

5.3 Cloning and recombination of the libraries

The oligonucleotide pools were manufactured by Twist Bioscience (South San Francisco, United States). The oligonucleotides were used as template for PCR amplification (oligonucleotide concentration: 0.1 μ M; 15 cycles). The PCR products of correct size were purified by agarose gel electrophoresis. The purified linear DNA was used for cloning of pTS040 by Gibson assembly (NEBuilder HiFi DNA Assembly Reaction, NEB) and electroporation of *E. coli* MegaX DH10B T1R cells (Invitrogen, Thermo Fisher Scientific Inc.). *E. coli* BW25113 carrying pTS041 was cultured in LB medium at 37 °C under shaking of 220 rpm until exponential growth. Expression of the Lambda red genes was induced with L-arabinose (7.5 g/L). After 30 min, the culture was harvested for electroporation with the pooled pTS041 plasmid library. Cells were recovered in SOC medium with kanamycin and 1 μ M aTc for Cas9 induction at 30 °C for 2 h and streaked out on to LB agar plates with kanamycin, chloramphenicol, and 1 μ M aTc. After incubation overnight at 30 °C, colonies were pooled by flushing the agar plates with LB medium, glycerol added (final concentration: 22 vol.-%), the OD was determined, and the strain library stored as cryo stocks.

5.4 Cultivations

If not stated otherwise, minimal medium (M9) was used for the experiments and contained 42.2 mM Na₂HPO₄, 11.3 mM (NH₄)₂SO₄, 22 mM KH₂PO₄, 8.56 mM NaCl, 1 mM MgSO₄ x 7 H₂O, 100 μ M CaCl₂ x 2 H₂O, 60 μ M FeCl₃, 6.3 μ M ZnSO₄ x 7 H₂O, 7 μ M CuCl₂ x 2 H₂O, 7.1 μ M, MnSO₄ x 2 H₂O, 2.8 μ M thiamine-HCL, and 7.6 μ M CoCl₂ x 6 H₂O. Standardly, 5 g/L glucose was used as substrate. Minimal medium (M9) and LB agar plates contained 1.5 % agar. 30 μ g/L chloramphenicol, 50 μ g/mL kanamycin, and 50 μ g/mL spectinomycin were added to the media when required.

Competition experiment and sampling for amplicon sequencing

The ts plasmid library (before electroporation of *E. coli* BW25113//pTS041) was used as a sample for amplicon sequencing ("sample before recombination"). Plasmids were extracted from the cryo stock of the ts strain library (*E. coli* BW25113//pTS041//pTS040(ts-

library), "sample after recombination"). 75 mL M9 medium was inoculated with 200 μ L of the ts strain library from cryo stock and incubated in a 500 mL shake flask for 15 h at 30 °C under shaking of 220 rpm. 10 mL of the exponentially growing culture were used for plasmid extraction ("sample time point zero"). 300 mL of M9 medium was inoculated with the previous culture to a start OD of 0.1. The 300 mL culture was split up to each 150 mL for cultivation in 1 L-shake flasks at 30 °C and 42 °C under shaking of 220 rpm. Every 3 h, the 150 mL cultures were back diluted to an OD of 0.1. Every 2 h, a sample for plasmid extraction was taken (sample volume \times OD \geq 5).

96-well cultivation and sampling for metabolomics by flow-injection mass spectrometry

500 μ L of LB in 2-mL deep well plates (96-well) were inoculated from cryo stock, covered with Breathe-Easy (Diversified Biotech BEM-1) adhesive membrane, and incubated for 6 h at 30 °C under shaking at 220 rpm. 495 μ L M9 (5 g/L glucose) were inoculated with 5 μ L of the LB precultures and incubated for 24 h at 30 °C in 2 mL deep-well plates, under shaking of 220 rpm. 100 μ L M9 precultures were transferred to 900 μ L of fresh M9 + 5 g/L glucose in 2 mL deep-well plates and incubated for 16 h at 42 °C, under shaking of 220 rpm. 850 μ L of the liquid culture was centrifuged in 2 mL deep-well plates for 15 min at 4,000 rpm at 4 °C. The supernatant was removed and the cell pellets stored at -80 °C. 100 μ L of -20 °C-cold 40:40:20 acetonitrile:methanol:water was added to the frozen cell pellets and incubated for 4 h at -20 °C. The plate was vortexed and 80 μ L of the cell extract transferred to v-bottomed 96-well storage plates. The cell extracts were stored at -80 °C until further analysis by flow-injection mass spectrometry.

5.5 Plate reader cultivations

500 μ L of LB in 2-mL deep well plates (96-well) were inoculated from cryo stock, covered with Breathe-Easy (Diversified Biotech BEM-1) adhesive membrane, and incubated for 6 h at 30 °C under shaking at 220 rpm. 500 μ L M9 (5 g/L glucose) were inoculated with 1 μ L of the LB precultures and incubated overnight at 30 °C in 2 mL deep-well plates under shaking of 220 rpm. 297 μ L of M9 (5 g/L glucose) were inoculated with 3 μ L of the overnight cultures in 96-well Greiner plates (flat-bottomed). 150 μ L were transferred to a second 96-well Greiner plate. Each plate was incubated at each two different temperatures. Biotek Epoch 2 or TECAN Synergy plate readers were used for incubation and measurements of OD at 600 nm every 10 min. Maximum growth rates were calculated

in exponential growth phases if applicable.

5.6 Two-stage production of metabolites and sampling for LC-MS/MS

5 mL LB cultures were started from cryo stock. After ca. 6 h at 30 °C, 5 mL M9 + 5 g/L glucose overnight cultures were started using 25 μ L of the LB culture for inoculation. Overnight M9 cultures were washed: cultures were pelletized by centrifugated at 4,000 rpm and 40 °C. After removing the supernatant, fresh 5 mL of M9 glucose medium was added for resuspending cells. This step was repeated further 2 times. Final 15 mL cultures were started in 100 mL shaking flasks at an OD of 0.5 and incubated for 24 h under shaking of 220 rpm and 42 °C. At the start of the cultivation and after 2 h, 4 h, 6 h, and 24 h incubation, the OD₆₀₀ was measured and samples for LC-MS/MS taken: 100 μ L of the culture broth transferred to 400 μ L -20 °C cold 50:50 acetonitrile:methanol in 1.5 mL reaction tubes. The samples were stored at -80 °C until further processing. The samples were centrifugated for 15 min at 17,000 g and -9 °C. Metabolite concentrations in the supernatant were analyzed by LC-MS/MS.

5.7 Sample processing for NGS

Using 3 ng total plasmid DNA, a plasmid part covering the homology sequences and protospacers was amplified (15 cycles) using two primers suited for further indexing PCRs (forward: 'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTATCACGAGGCAGATCCTCTG', reverse: 'GTCTCGTGGCTCGGAGATGTGTATAAGAGACAGACTCGGTGCCACTTTTCAAGTT'). Amplicons were purified by AM-Pure XP PCR beads (Beckman Coulter, # A63881). Using standard Illumina indexing primers, amplicons were indexed in a second PCR and again purified by bead-clean up. Amplicons were pooled and sequenced on an Illumina NextSeq500 (paired-end, NextSeqTM 500 Mid Output Kit v2.5, # 20024908, 300 cycles). Two cartridges were required to yield the desired sequencing depth of around 4 million reads per sample.

5.8 NGS data analysis

Demultiplexed paired-end reads were aligned, merged (based on overlapping sequences), and trimmed to the region of interest using a custom Matlab script. The resulting processed reads were mapped against the designed sequences of the library. For each library member, the number of matching reads was counted. Only reads that shared a

100 % identity with a designed sequence were counted since mutations could indicate a malfunction of the CRISPR-Cas9 genome editing system with no genomic edit. For each sample, the library member counts were normalized to the total number of reads of the sample. The samples of a replicate were further normalized to the first sample of the replicate. Using the normalized data for the ts library, the area under the curve (AUC) was determined for the 30 °C and 42 °C time series of each library member. For each library member of the CRISPR library i , an error e was estimated using the normalized read counts \bar{n} for each replicate (A and B) and time point t normalized to the mean normalized read counts:

$$e_i = \sum_t \frac{|\bar{n}_{i,A,t,30^\circ C} - \bar{n}_{i,B,t,30^\circ C}|}{\frac{\bar{n}_{i,A,t,30^\circ C} + \bar{n}_{i,B,t,30^\circ C}}{2}} + \frac{|\bar{n}_{i,A,t,42^\circ C} - \bar{n}_{i,B,t,42^\circ C}|}{\frac{\bar{n}_{i,A,t,42^\circ C} + \bar{n}_{i,B,t,42^\circ C}}{2}} \quad (5.1)$$

5.9 Selecting temperature-sensitive candidate strains from the CRISPR library

Only strains were considered that met following criteria:

- They had at least 15 reads at time point zero ($\bar{r}_{t=0h}$).
- The mean normalized read count of the last sample at 30 °C ($\bar{n}_{i,t=12h,30^\circ C}$) was greater than 0.3.
- The mean normalized read count of the last sample at 42 °C ($\bar{n}_{i,t=12h,42^\circ C}$) was lower than 0.4.
- The error e_i was lower than 15.
- The area under the curve for the 30 °C time series ($AUC_{i,30^\circ C}$) was greater than 5.
- The mean normalized read count of the last samples for the different temperatures fulfilled following criterion:

$$\bar{n}_{i,t=12h,30^\circ C} > \frac{1}{1 + \left(\frac{0.25}{\bar{n}_{i,t,12h,42^\circ C}}\right)} + 0.3 \quad (5.2)$$

For each gene, the candidates were sorted by the number of reads at time point zero $\bar{r}_{t=0h}$, the relative area under the curve difference, the difference between the normalized read counts at 30 °C and 42 °C at the last time point $diff_{AUC} = \frac{AUC_{30^\circ C} - AUC_{42^\circ C}}{AUC_{42^\circ C}}$ ($diff_{t=12h}$), and the error e . Based on placement in the sortings ($rank$), a score was calculated for each candidate i :

$$score_i = rank_{i, \bar{n}_{t=0h}} \cdot 1.25 + rank_{i, diff_{AUC}} \cdot 2 + rank_{i, diff_{t=12h}} \cdot 1.5 + rank_{i, c} \cdot 0.75 \quad (5.3)$$

The candidates for each gene with the lowest score were selected for a new (sub-)library.

5.10 Metabolomics

Flow-injection mass spectrometry (FI-MS)

An Agilent 6546 QTOF mass spectrometer (Agilent Technologies, Santa Clara, USA) was used to analyze metabolite levels in the cell extracts. The source parameters were: source gas 225 °C, flow rate of the drying gas 11 L/min, nebulizer pressure 20 psi, sheath gas temperature 350 °C, sheath gas flow 10 L/min, nozzle voltage. Spectra in a 50-1100 m/z range were acquired in 10 Ghz mode with an acquisition rate of 1.4 spectrum/s. The mobile phase was 10 mM (NH₄)CO₃, 0.04 % NH₄OH, 60:40 Isopropanol:H₂O. The reference masses for online mass calibration in negative mode were 59.050 Da (C₃H₈O, Isopropanol) and 1033.988 Da (C₁₈H₁₈F₂₄N₃O₆P₃, HP-921); in positive mode, 121.050 Da (C₅H₄N₄, Purine) and 922.009 Da (C₁₈H₁₈F₂₄N₃O₆P₃, HP-921).

FI-MS data analysis

Raw data files were converted into ".mzXML" files by MSConvert (Chambers et al., 2012). Following data analysis was performed by custom MATLAB scripts that utilized MATLAB functions (The MathWorks, Inc., Massachusetts, USA). The 10 spectra with the highest signal in the TIC were summed, the resulting spectra were resampled ("msresample"). Peaks with a minimum peak height of 1000 units and a peak peak prominence of 500 units were selected ("findpeaks"). Peaks were annotated with a 3 mDa tolerance by matching monoisotopic masses of metabolites with a single proton loss for negative mode and single proton gain in positive mode. Double annotations (positive and negative mode) were manually cured based on peak shape and height. For each metabolite, the maximum height of the annotated peak was taken for further analysis. The data was then normalized to the control strain and converted into log-2 space. Subsequently, modified z-scores were calculated as following:

$$mod.z - core = \frac{0.6745 \cdot (x_i - median(x))}{median(abs(x - median(x)))} \quad (5.4)$$

Targeted metabolomics by LC-MS/MS

Whole culture broth samples were analyzed by an isotope-ratio based LC-MS/MS method (Guder et al., 2017). Changes to the LC parameters were made to reduce exposure of the ESI to inorganic salts from the samples: in the initial 0.3 min, the analyte was discarded into the waste. Between 0.3 and 2.0 min the analyte was injected to the ESI. 2.0 to 2.3 min the analyte was discarded into the waste again. An internal, fully ^{13}C -labelled standard was calibrated with authentic ^{12}C -metabolite standards. Based on the calibrated ^{13}C -standard and isotope ratios, absolute metabolite concentrations in the samples were calculated. Homoserine and threonine could not be distinguished. We used an authentic homoserine standard to calculate absolute concentrations.

Acknowledgements

Amplicon next generation sequencing was performed and supported by Janina Geißert and the NGS Competence Center Tübingen (NCCT) and its technology platforms.

CRedit authorship contribution statement

Thorben Schramm: Conceptualization, Investigation, Visualization, Project administration, writing, Formal analysis.

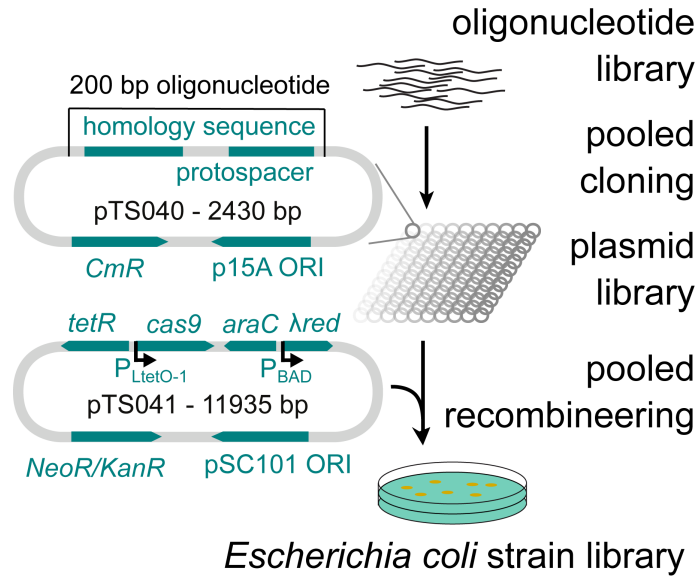
Vanessa Pahl: Investigation (arginine overproduction).

Hannes Link: Conceptualization, Supervision, Project administration, Visualization, Writing, Funding acquisition.

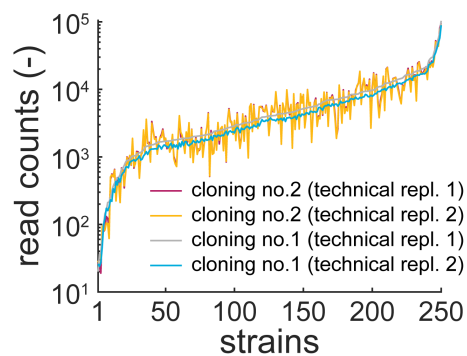
Declaration of interests

The authors declare no competing interests.

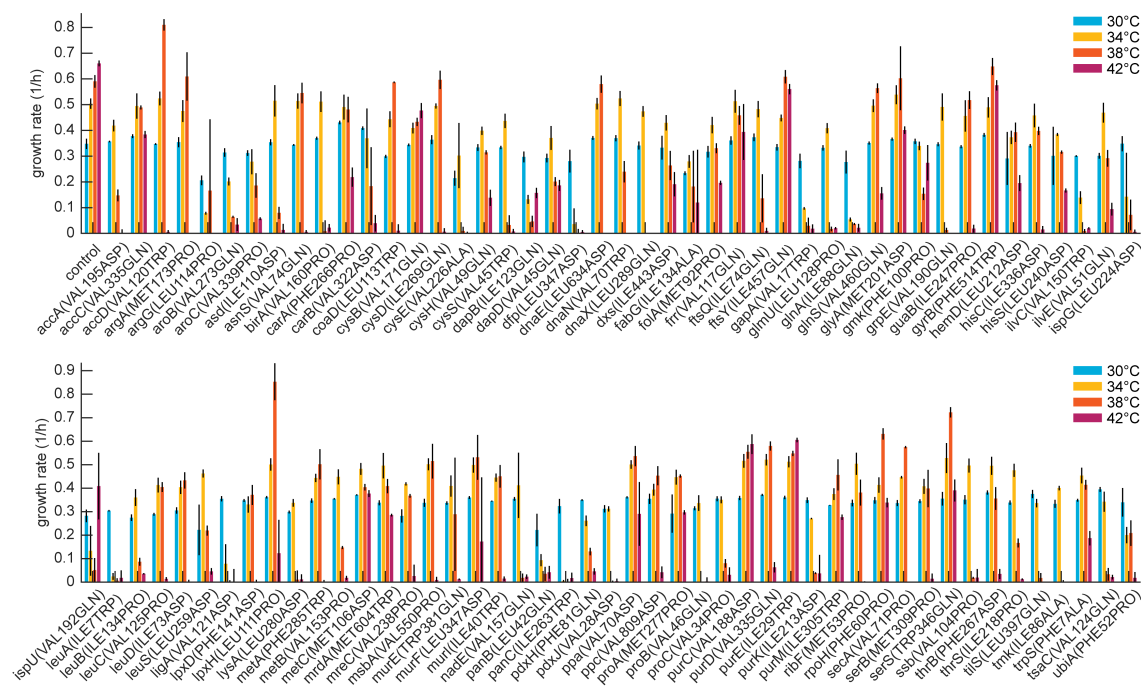
6 Supplementary figures



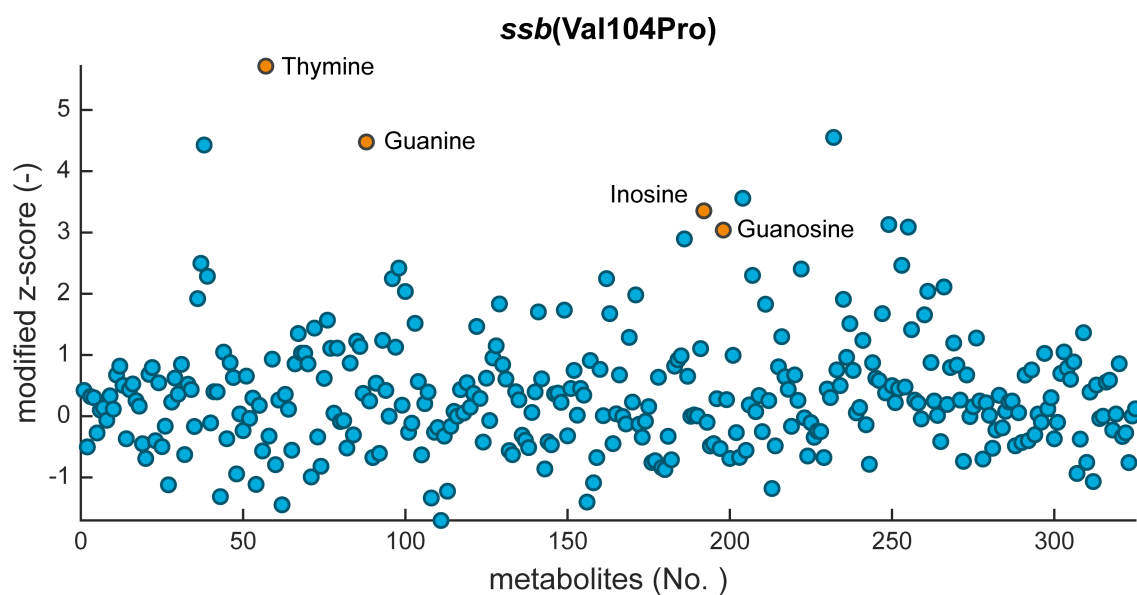
Suppl. Fig. 5.6: Cloning and recombineering by an adapted CRISPR-Cas9 genome editing method. A oligonucleotide pool was used to clone a pooled plasmid library. The plasmid library was used in a pooled recombination step to introduce specific mutations into an *E. coli* genome.



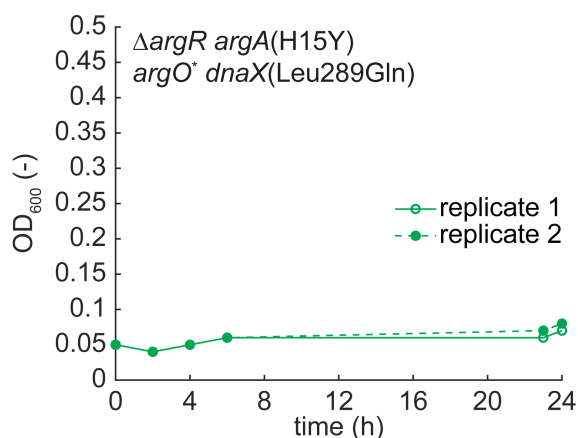
Suppl. Fig. 5.7: Reproducible construction of a pooled strain library. The chart shows the read counts of each strain in the library in two different cloning batches with each two technical replicates.



Suppl. Fig. 5.8: The maximum growth rates of 92 temperature-sensitive strains at different temperatures. The bar plot shows the mean maximum growth rates of indicated strains at 30 °C (blue), 34 °C (yellow), 38 °C (orange), 42 °C (red) during cultivation for 24 h in 96-well microtiter plates. Black vertical lines indicate the standard deviations ($n = 3$).



Suppl. Fig. 5.9: Metabolomics data of the temperature-sensitive strain *ssb(Val104Pro)*. The modified z-score was calculated as given in M.8.2 and is shown on the y-axis of the chart. The x-axis shows the different measured metabolites. Yellow indicate some of metabolites that had a z-score > 3. Dots are the mean of three replicates.



Suppl. Fig. 5.10: OD₆₀₀-time course of a temperature-sensitive arginine overproduction strain. Lines and dashed lines indicate replicates.

References

- Burg, J. M., Cooper, C. B., Ye, Z., Reed, B. R., Moreb, E. A., & Lynch, M. D. (2016). Large-scale bioprocess competitiveness: The potential of dynamic metabolic control in two-stage fermentations. *Current Opinion in Chemical Engineering*, *14*, 121–136. <https://doi.org/10.1016/j.coche.2016.09.008> (cited on page 143).
- Cress, B. F., Trantas, E. A., Ververidis, F., Linhardt, R. J., & Koffas, M. A. (2015). Sensitive cells: Enabling tools for static and dynamic control of microbial metabolic pathways. *Current Opinion in Biotechnology*, *36*, 205–214. <https://doi.org/10.1016/j.copbio.2015.09.007> (cited on page 143).
- Hartline, C. J., Schmitz, A. C., Han, Y., & Zhang, F. (2021). Dynamic control in metabolic engineering: Theories, tools, and applications. *Metabolic Engineering*, *63*, 126–140. <https://doi.org/10.1016/j.ymben.2020.08.015> (cited on page 143).
- Klamt, S., Mahadevan, R., & Hädicke, O. (2018). When do two-stage processes outperform one-stage processes? *Biotechnology Journal*, *13*(2), 1700539. <https://doi.org/10.1002/biot.201700539> (cited on page 143).
- Chubukov, V., Desmarais, J. J., Wang, G., Chan, L. J. G., Baidoo, E. E., Petzold, C. J., Keasling, J. D., & Mukhopadhyay, A. (2017). Engineering glucose metabolism of escherichia coli under nitrogen starvation. *npj Systems Biology and Applications*, *3*(1), 16035. <https://doi.org/10.1038/npsjba.2016.35> (cited on page 143).
- Sonderregger, M., Schümperli, M., & Sauer, U. (2005). Selection of quiescent escherichia coli with high metabolic activity. *Metabolic Engineering*, *7*(1), 4–9. <https://doi.org/10.1016/j.ymben.2004.05.005> (cited on page 143).
- Tokuyama, K., Toya, Y., Matsuda, F., Cress, B. F., Koffas, M. A. G., & Shimizu, H. (2019). Magnesium starvation improves production of malonyl-CoA-derived metabolites in escherichia coli. *Metabolic Engineering*, *52*, 215–223. <https://doi.org/10.1016/j.ymben.2018.12.002> (cited on page 143).
- Soma, Y., Tsuruno, K., Wada, M., Yokota, A., & Hanai, T. (2014). Metabolic flux redirection from a central metabolic pathway toward a synthetic pathway using a metabolic toggle switch. *Metabolic Engineering*, *23*, 175–184. <https://doi.org/10.1016/j.ymben.2014.02.008> (cited on page 143).
- Zhao, E. M., Zhang, Y., Mehl, J., Park, H., Lalwani, M. A., Toettcher, J. E., & Avalos, J. L. (2018). Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature*, *555*(7698), 683–687. <https://doi.org/10.1038/nature26141> (cited on page 143).
- Gupta, A., Reizman, I. M. B., Reisch, C. R., & Prather, K. L. J. (2017). Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nature Biotechnology*, *35*(3), 273–279. <https://doi.org/10.1038/nbt.3796> (cited on page 143).
- Brockman, I. M., & Prather, K. L. J. (2015). Dynamic metabolic engineering: New strategies for developing responsive cell factories. *Biotechnology Journal*, *10*(9), 1360–1369. <https://doi.org/10.1002/biot.201400422> (cited on page 143).
- Lv, L., Ren, Y.-L., Chen, J.-C., Wu, Q., & Chen, G.-Q. (2015). Application of CRISPRi for prokaryotic metabolic engineering involving multiple genes, a case study: Controllable p(3hb-co-4hb) biosynthesis. *Metabolic Engineering*, *29*, 160–168. <https://doi.org/10.1016/j.ymben.2015.03.013> (cited on page 143).

- Lynch, M. D., Gill, R. T., & Lipscomb, T. E. W. (2016, August 30). *Method for producing 3-hydroxypropionic acid and other products* (U.S. pat. No. 9428778B2). Retrieved August 10, 2022, from <https://patents.google.com/patent/US9428778B2/en> (cited on page 143).
- Lynch, M., Louie, M., Copley, S., Spindler, E., Prather, B., Lipscomb, M., Lipscomb, T., Liao, H., Hogsett, D., & Evans, R. (2019, July 2). *Microorganisms and methods for the production of fatty acids and fatty acid derived products* (U.S. pat. No. 10337038B2). <https://patents.google.com/patent/US10337038B2/en> (cited on page 143).
- Schramm, T., Lempp, M., Beuter, D., Sierra, S. G., Glatter, T., & Link, H. (2020). High-throughput enrichment of temperature-sensitive argininosuccinate synthetase for two-stage citrulline production in *e. coli*. *Metabolic Engineering*. <https://doi.org/10.1016/j.ymben.2020.03.004> (cited on pages 143, 154, 157, 158, 161, 163).
- Weber, W. (2003). Conditional human VEGF-mediated vascularization in chicken embryos using a novel temperature-inducible gene regulation (TIGR) system. *Nucleic Acids Research*, *31*(12), 69e–69. <https://doi.org/10.1093/nar/gng069> (cited on page 143).
- Ben-Aroya, S., Coombes, C., Kwok, T., O'Donnell, K. A., Boeke, J. D., & Hieter, P. (2008). Toward a comprehensive temperature-sensitive mutant repository of the essential genes of *saccharomyces cerevisiae*. *Molecular Cell*, *30*(2), 248–258. <https://doi.org/10.1016/j.molcel.2008.02.021> (cited on page 143).
- Kofoed, M., Milbury, K. L., Chiang, J. H., Sinha, S., Ben-Aroya, S., Giaever, G., Nislow, C., Hieter, P., & Stirling, P. C. (2015). An updated collection of sequence barcoded temperature-sensitive alleles of yeast essential genes. *5*(9), 1879–1887. <https://doi.org/10.1534/g3.115.019174> (cited on page 143).
- Li, Z., Vizeacoumar, F. J., Bahr, S., Li, J., Warringer, J., Vizeacoumar, F. S., Min, R., VanderSluis, B., Bellay, J., DeVit, M., Fleming, J. A., Stephens, A., Haase, J., Lin, Z.-Y., Baryshnikova, A., Lu, H., Yan, Z., Jin, K., Barker, S., ... Boone, C. (2011). Systematic exploration of essential yeast gene function with temperature-sensitive mutants. *Nature Biotechnology*, *29*(4), 361–367. <https://doi.org/10.1038/nbt.1832> (cited on pages 143, 163).
- Ben-Aroya, S., Pan, X., Boeke, J. D., & Hieter, P. (2010, January 1). Chapter 8 - making temperature-sensitive mutants. In *Methods in enzymology* (pp. 181–204, Vol. 470). Academic Press. [https://doi.org/10.1016/S0076-6879\(10\)70008-2](https://doi.org/10.1016/S0076-6879(10)70008-2) (cited on page 143).
- Garst, A. D., Bassalo, M. C., Pines, G., Lynch, S. A., Halweg-Edwards, A. L., Liu, R., Liang, L., Wang, Z., Zeitoun, R., Alexander, W. G., & Gill, R. T. (2017). Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nature Biotechnology*, *35*(1), 48–55. <https://doi.org/10.1038/nbt.3718> (cited on pages 144, 165).
- Goodall, E. C. A., Robinson, A., Johnston, I. G., Jabbari, S., Turner, K. A., Cunningham, A. F., Lund, P. A., Cole, J. A., & Henderson, I. R. (2018). The essential genome of *escherichia coli* k-12. *mBio*, *9*(1), e02096–17. <https://doi.org/10.1128/mBio.02096-17> (cited on pages 144, 165).
- Patrick, W. M., Quandt, E. M., Swartzlander, D. B., & Matsumura, I. (2007). Multicopy suppression underpins metabolic evolvability. *Molecular Biology and Evolution*, *24*(12), 2716–2722. <https://doi.org/10.1093/molbev/msm204> (cited on pages 144, 165).
- Tan, K. P., Khare, S., Varadarajan, R., & Madhusudhan, M. S. (2014). TSpred: A web server for the rational design of temperature-sensitive mutants. *Nucleic Acids Research*, *42*, W277–W284. <https://doi.org/10.1093/nar/gku319> (cited on pages 144, 162, 165).

- Varadarajan, R., Nagarajaram, H. A., & Ramakrishnan, C. (1996). A procedure for the prediction of temperature-sensitive mutants of a globular protein based solely on the amino acid sequence. *Proceedings of the National Academy of Sciences*, 93(24), 13908–13913. <https://doi.org/10.1073/pnas.93.24.13908> (cited on pages 144, 162, 165).
- Baek, M., & Baker, D. (2022). Deep learning and protein structure modeling. *Nature Methods*, 19(1), 13–14. <https://doi.org/10.1038/s41592-021-01360-8> (cited on page 149).
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2> (cited on page 149).
- Lovelock, S. L., Crawshaw, R., Basler, S., Levy, C., Baker, D., Hilvert, D., & Green, A. P. (2022). The road to fully programmable protein catalysis. *Nature*, 606(7912), 49–58. <https://doi.org/10.1038/s41586-022-04456-z> (cited on page 149).
- Daniel, R. M., & Danson, M. J. (2013). Temperature and the catalytic activity of enzymes: A fresh understanding. *FEBS Letters*, 587(17), 2738–2743. <https://doi.org/10.1016/j.febslet.2013.06.027> (cited on pages 151, 153, 163).
- Sharpe, P. J. H., & DeMichele, D. W. (1977). Reaction kinetics of poikilotherm development. *Journal of Theoretical Biology*, 64(4), 649–670. [https://doi.org/10.1016/0022-5193\(77\)90265-X](https://doi.org/10.1016/0022-5193(77)90265-X) (cited on page 151).
- Zwietering, M. H., de Koos, J. T., Hasenack, B. E., de Witt, J. C., & van't Riet, K. (1991). Modeling of bacterial growth as a function of temperature. *Applied and Environmental Microbiology*, 57(4), 1094–1101. <https://doi.org/10.1128/aem.57.4.1094-1101.1991> (cited on page 151).
- Fuhrer, T., Heer, D., Begemann, B., & Zamboni, N. (2011). High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection–time-of-flight mass spectrometry. *Analytical Chemistry*, 83(18), 7074–7080. <https://doi.org/10.1021/ac201267k> (cited on page 153).
- Sander, T., Wang, C. Y., Glatter, T., & Link, H. (2019). CRISPRi-based downregulation of transcriptional feedback improves growth and metabolism of arginine overproducing e. coli. *ACS Synthetic Biology*, 8(9), 1983–1990. <https://doi.org/10.1021/acssynbio.9b00183> (cited on pages 156, 157, 162, 165).
- Sander, T., Farke, N., Diehl, C., Kuntz, M., Glatter, T., & Link, H. (2019). Allosteric feedback inhibition enables robust amino acid biosynthesis in e. coli by enforcing enzyme overabundance. *Cell Systems*, 8(1), 66–75.e8. <https://doi.org/10.1016/j.cels.2018.12.005> (cited on page 157).
- Lempp, M., Farke, N., Kuntz, M., Freibert, S. A., Lill, R., & Link, H. (2019). Systematic identification of metabolites controlling gene expression in e. coli. *Nature Communications*, 10(1), 4463. <https://doi.org/10.1038/s41467-019-12474-1> (cited on page 157).
- Nguyen Le Minh, P., Velázquez Ruiz, C., Vandermeeren, S., Abwoyo, P., Bervoets, I., & Charlier, D. (2018). Differential protein-DNA contacts for activation and repression by ArgP, a LysR-type (LTTR) transcriptional regulator in escherichia coli. *Microbiological Research*, 206, 141–158. <https://doi.org/10.1016/j.micres.2017.10.009> (cited on page 157).

- Shereda, R. D., Kozlov, A. G., Lohman, T. M., Cox, M. M., & Keck, J. L. (2008). SSB as an organizer/mobilizer of genome maintenance complexes. *Critical Reviews in Biochemistry and Molecular Biology*, 43(5), 289–318. <https://doi.org/10.1080/10409230802341296> (cited on page 158).
- Donati, S., Kuntz, M., Pahl, V., Farke, N., Beuter, D., Glatter, T., Gomes-Filho, J. V., Randau, L., Wang, C.-Y., & Link, H. (2021). Multi-omics analysis of CRISPRi-knockdowns identifies mechanisms that buffer decreases of enzymes in e. coli metabolism. *Cell Systems*, 12(1), 56–67.e6. <https://doi.org/10.1016/j.cels.2020.10.011> (cited on page 158).
- Donati, S., Sander, T., & Link, H. (2018). Crosstalk between transcription and metabolism: How much enzyme is enough for a cell? *WIREs Systems Biology and Medicine*, 10(1). <https://doi.org/10.1002/wsbm.1396> (cited on page 158).
- Harder, B.-J., Bettenbrock, K., & Klamt, S. (2018). Temperature-dependent dynamic control of the TCA cycle increases volumetric productivity of itaconic acid production by *Escherichia coli*. *Biotechnology and Bioengineering*, 115(1), 156–164. <https://doi.org/10.1002/bit.26446> (cited on page 158).
- Cho, H.-S., Seo, S. W., Kim, Y. M., Jung, G. Y., & Park, J. M. (2012). Engineering glyceraldehyde-3-phosphate dehydrogenase for switching control of glycolysis in *Escherichia coli*. *Biotechnology and Bioengineering*, 109(10), 2612–2619. <https://doi.org/10.1002/bit.24532> (cited on page 158).
- Guder, J. C., Schramm, T., Sander, T., & Link, H. (2017). Time-optimized isotope ratio LC–MS/MS for high-throughput quantification of primary metabolites. *Analytical Chemistry*, 89(3), 1624–1631. <https://doi.org/10.1021/acs.analchem.6b03731> (cited on pages 159, 171).
- Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis? *Nature Biotechnology*, 28(3), 245–248. <https://doi.org/10.1038/nbt.1614> (cited on page 159).
- Eyre-Walker, A., & Keightley, P. D. (2007). The distribution of fitness effects of new mutations. *Nature Reviews. Genetics*, 8(8), 610–618. <https://doi.org/10.1038/nrg2146> (cited on page 162).
- Wloch, D. M., Szafraniec, K., Borts, R. H., & Korona, R. (2001). Direct estimate of the mutation rate and the distribution of fitness effects in the yeast *Saccharomyces cerevisiae*. *Genetics*, 159(2), 441–452. <https://doi.org/10.1093/genetics/159.2.441> (cited on page 162).
- Tiessen, A., Pérez-Rodríguez, P., & Delaye-Arredondo, L. J. (2012). Mathematical modeling and comparison of protein size distribution in different plant, animal, fungal and microbial species reveals a negative correlation between protein size and protein number, thus providing insight into the evolution of proteomes. *BMC Research Notes*, 5(1), 85. <https://doi.org/10.1186/1756-0500-5-85> (cited on page 163).
- Chakshusmathi, G., Mondal, K., Lakshmi, G. S., Singh, G., Roy, A., Ch., R. B., Madhusudhanan, S., & Varadarajan, R. (2004). Design of temperature-sensitive mutants solely from amino acid sequence. *Proceedings of the National Academy of Sciences*, 101(21), 7925–7930. <https://doi.org/10.1073/pnas.0402222101> (cited on page 163).
- Bae, S., Park, J., & Kim, J.-S. (2014). Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*, 30(10), 1473–1475. <https://doi.org/10.1093/bioinformatics/btu048> (cited on page 165).
- Rousset, F., Cui, L., Siouve, E., Becavin, C., Depardieu, F., & Bikard, D. (2018). Genome-wide CRISPR-dCas9 screens in *E. coli* identify essential genes and phage host factors. *PLoS Genetics*, 14(11), e1007749. <https://doi.org/10.1371/journal.pgen.1007749> (cited on page 165).

Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M.-Y., Paule, C., Creasy, D., ... Mallick, P. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, *30*(10), 918–920. <https://doi.org/10.1038/nbt.2377> (cited on page 170).

Chapter 6

Broadening the scope of enforced ATP wasting as a tool for metabolic engineering in *Escherichia coli*

Simon Boecker, Ahmed Zahoor, Thorben Schramm, Hannes Link, Steffen Klamt

This chapter was published in: Boecker et al., 2019, *Biotechnology Journal*, 2019, 14, 1800438, pp. 1-9, DOI: <https://doi.org/10.1002/biot.201800438>, (license: Creative Commons "Attribution-NonCommercial-NoDerivatives 4.0 International" (CC BY-NC-ND 4.0), <http://creativecommons.org/licenses/by-nc-nd/4.0/>, no modifications). My contributions were metabolite measurements by LC-MS/MS and processing of related data. Related data is shown in Fig. 6.3.A.

Abstract

The targeted increase of cellular adenosine triphosphate (ATP) turnover (enforced ATP wasting) has recently been recognized as a promising tool for metabolic engineering when product synthesis is coupled with net ATP formation. The goal of the present study is to further examine and to further develop the concept of enforced ATP wasting and to broaden its scope for potential applications. In particular, considering the fermentation products synthesized by *Escherichia coli* under anaerobic conditions as a proxy for target chemical(s), i) a new genetic module for dynamic and gradual induction of the F₁-part of the ATPase is developed and it is found that ii) induction of the ATPase leads to higher metabolic activity and increased product formation in *E. coli* under anaerobic conditions, and that iii) ATP wasting significantly increases substrate uptake and productivity of growth-arrested cells, which is vital for its use in two-stage processes. To the best of the authors' knowledge, the glucose uptake rate of $6.49 \text{ mmol g}_{CDW}^{-1} \text{ h}^{-1}$ achieved with enforced ATP wasting is the highest value reported for nongrowing *E. coli* cells. In summary, this study shows that enforced ATP wasting can be used to improve yield and titer (in growth-coupled processes) as well as volumetric productivity (in two-stage processes) depending on which of the performance measures is more crucial for the process and product of interest.

1 Introduction

The development of bio-based production processes for fuels, commodity chemicals, and high-value products plays a pivotal role in making the chemical industry more sustainable and environmentally friendly (Liao et al., 2016; Becker and Wittmann, 2015; Lee et al., 2019). To be economically competitive with conventional fossil-based manufacturing and to justify industrial applications, these processes need to be optimized in terms of the three key performance measures: yield, titer, and (volumetric) productivity (Liao et al., 2016; Cho et al., 2015). Metabolic engineering of the microbial production organisms is one of the main approaches to improve these parameters (Keasling, 2010).

Manipulating the adenosine triphosphate (ATP) pool has been a major target for enhancing the performance of production organisms (Hädicke and Klamt, 2015; Singh et al., 2011; Hädicke et al., 2015; Bj et al., 2002; Liu et al., 2016; de Kok et al., 2012; Koebmann et al., 2002). For pathways with ATP limitations, increasing the pool of available ATP may improve the production of desired compounds, e.g., of succinic acid (Singh et al., 2011; Liu et al., 2016) or of recombinant proteins (Kim et al., 2012). A contrary strategy for metabolic engineering has been proposed more recently based on enforced ATP consumption (or enforced ATP wasting) (Hädicke et al., 2015; Liu et al., 2016; Semkiv et al., 2016; Jensen et al., 2003). The main idea behind this approach is as follows: if product synthesis is coupled to net ATP synthesis, an increased ATP drain should—not only for thermodynamic reasons, but also due to evolutionary pressure—lead to an increased flux along the product pathway. Enforced ATP hydrolysis has been implemented either by the introduction of short futile cycles (Hädicke et al., 2015; Semkiv et al., 2016; Patnaik et al., 1992) or by a more "direct" approach via expressing ATP-hydrolyzing enzymes such as the (uncoupled) cytosolic F_1 -subunit of the ATPase from *Escherichia coli* (Liu et al., 2016; de Kok et al., 2012; Koebmann et al., 2002). Early studies (Koebmann et al., 2002; Holm et al., 2010; Chao and Liao, 1994) focused mainly on the physiological response of the cells upon exposure to elevated ATP drain and did not aim to increase the synthesis of certain products (Table 6.1). In these works, the authors consistently found that the substrate uptake rates in *E. coli* increased with elevated ATP consumption, indicating the potential of ATP wasting to improve the properties of microbial cell factories. Koebmann et al. (Koebmann et al., 2002) already envisioned the application of ATP wasting for metabolic engineering purposes; however, only very recently, first concrete application examples have been published targeting lactate synthesis in *E. coli*, (Hädicke et al., 2015) acetoin synthesis in *Lactococcus lactis* (Liu et al., 2016), and ethanol production in

Saccharomyces cerevisiae (Semkiv et al., 2016) (Table 6.1).

In general, using enforced ATP wasting as a tool for metabolic engineering requires that synthesis of the desired product is coupled with net ATP synthesis, ideally in an obligatory manner where the pathway from substrate to product is the only pathway that leads to net ATP production. In the case of anaerobic ethanol production with yeast, (Semkiv et al., 2016) this is naturally the case. For lactate production with *E. coli* and acetoin production with *L. lactis*, alternative (fermentation) pathways for ATP synthesis have to be blocked by knocking out certain metabolic genes. Note that, to obtain coupling, ATP synthesis need not necessarily be achieved along the pathway from a precursor to the product. For example, the pathway from pyruvate to lactate does not lead to the production of ATP; however, ATP synthesis from glucose via glycolysis under anaerobic conditions essentially requires the lactate pathway for balancing redox and replenishing NAD if all alternative pathways have been blocked.

In the above-cited works, (Hädicke et al., 2015; Liu et al., 2016; Semkiv et al., 2016) it could be shown that in strains where ATP synthesis is coupled with product synthesis, a higher ATP demand imposed by a futile cycle or an uncoupled ATPase forces the cells to generate more ATP, leading to higher (specific) productivities, substrate uptake rates, and product yields accompanied with decreased growth rate and biomass yield. Based on the results of these proofs of principle, we here aim to further develop and broaden the scope of ATP wasting as a tool for metabolic engineering. One major aspect studied herein concerns the fact that due to the reduced formation of biomass (the catalyst of the bioconversion), the volumetric productivity (amount of product per time and volume) will likely decrease compared to the strain without ATP wasting (Hädicke and Klamt, 2015).

One approach to overcome such inherent trade-offs between high product yield and high volumetric productivity is to use two-stage fermentation (TSF) (Burg et al., 2016; Venayak et al., 2015; Klamt et al., 2018) with decoupled growth and production phase in contrast to one-stage fermentation (OSF) with growth-coupled product synthesis as used in the aforementioned studies. In a recent theoretical study where the productivities of OSF and TSF processes were systematically compared (Klamt et al., 2018), we found that enforced ATP wasting in the second (production) phase can significantly increase the productivity and thus the competitiveness of TSFs as it keeps a high driving force for substrate uptake also in the production phase where cell growth (usually consuming a large fraction of the substrate taken up) is missing (Harder et al., 2018; Chubukov and Sauer, 2014). Accordingly, we here want to give experimental evidence that ATP

Table 6.1: Aspects investigated in this study and other published papers on enforced ATP wasting.

	Chao and Liao, 1994	Koebmann et al., 2002	Holm et al., 2010	Hädicke et al., 2015	Liu et al., 2016	Semkiv et al., 2016	This study
Direct ATP wasting via F ₁ -ATPase	No (futile cycle)	Yes	Yes	No (futile cycle)	Yes	No (futile cycle)	Yes
Inducible ATPase	No	No	No	No	No	No	Yes
Organism	<i>E. coli</i>	<i>E. coli</i>	<i>E. coli</i>	<i>E. coli</i>	<i>L. lactis</i>	<i>S. cerevisiae</i>	<i>E. coli</i>
Anaerobic conditions	No	No	No	Yes	No	Yes	Yes
Use of ATP wasting for growth-coupled product synthesis	No	No	No	Yes (lactate)	Yes (acetone)	Yes (ethanol)	Yes (fermentation products)
Use of ATP wasting for growth-decoupled product synthesis	No	No	No	No	No	No	Yes (fermentation products)

wasting can serve as a tool to elevate substrate uptake and product synthesis rates in growth-arrested cells.

A second important goal of this study is to test the use of an inducible ATPase as an ATP wasting mechanism under anaerobic conditions, the preferred operation mode for industrial applications. So far there is only one study examining ATP wasting under anaerobic conditions in *E. coli* (Hädicke et al., 2015) (see also Table 6.1). This study used the pyruvate kinase/phosphoenolpyruvate synthase futile cycle, which is an indirect way to waste ATP and could have undesired side effects. In this study, we therefore intended to use the ATP-hydrolyzing F₁-subunit of the *E. coli* ATPase as a direct mechanism for enforced ATP wasting. This method was described earlier for use in *E. coli* and *L. lactis*, (Liu et al., 2016; Koebmann et al., 2002) however, only under aerobic cultivation conditions. In addition, in these studies, the ATPase was set under the control of a constitutive promoter. If a TSF or multistage fermentation process is meant to be applied, one needs dynamic control over the relevant reactions for the synthesis of the product (Venayak et al., 2015, 2018). Therefore, we put the ATPase encoding genes *atpAGD* under the control of a promoter inducible by *m*-toluate (Balzer et al., 2013).

As an application example to analyze the effect of ATP wasting with an inducible ATPase under anaerobic conditions in both growth-coupled (OSF) as well as growth-decoupled (for TSF) mode, we chose the *E. coli* wild-type strain MG1655 and considered the standard fermentation byproducts (ethanol, formate, acetate, lactate, and succinate) as proxies for product synthesis since formation of these metabolites is naturally coupled to ATP synthesis under anaerobic conditions (Clark, 1989). ATP wasting was induced in growth-coupled fermentation conditions (OSF) and in conditions where growth was arrested by the lack of a nitrogen source in the medium. We found that in growth-

coupled conditions, yield, titer, as well as specific productivity can be improved by ATP wasting, while volumetric productivity decreased due to a lower growth rate. In the growth-decoupled production phase, however, even volumetric productivity could be more than doubled by ATP wasting as high glucose uptake rates could be maintained during stationary phase. Our results demonstrate that ATP wasting is a promising general approach for metabolic engineering, as it can be used to adjust both factors of the trade-off of optimizing yield/titer or productivity, depending on which of the parameters is more crucial for the process and product of interest.

2 Experimental Section

2.1 Strain and Plasmid Construction

The strains, plasmids, and primers used in this study are listed in Table 6.2. The standard molecular cloning techniques followed the protocols described earlier (Sambrook and Russell, 2001). Polymerase chain reactions (PCRs) were performed using the Q5 Hot Start High-Fidelity DNA Polymerase (New England Biolabs) according to the manufacturer's protocol. For construction of pSB38.2, the kanamycin resistance cassette was cut out from pSB-M1g-1-17 with restriction enzyme PstI (New England Biolabs) and substituted with the PstI-digested ampicillin resistance cassette created by PCR-amplification from pKD3 with the primer pair Amp_PstI_fw/Amp_PstI_rv (Table 6.2). For the construction of the ATPase expression plasmid pSB44.1, genes encoding the ATPase F₁-subunit were amplified by PCR from pCP41::atpAGD with primer pair atpAGD_mono_fw/atpAGD_mono_rv (Table 6.2). gfpmut3 was cut out from plasmid pSB38.2 with restriction enzymes NdeI (New England Biolabs) and BamHI (New England Biolabs) and substituted with the NdeI/BamHI-digested atpAGD PCR product. For construction of the control vector pSB43.1, the 5'-overhangs of the NdeI/BamHI-digested plasmid pSB38.2 were filled-in using the Klenow Fragment (Thermo Scientific) and the blunt-ended DNA fragment was self-ligated. pSB43.1 and pSB44.1 were transformed into the *E. coli* wild-type strain MG1655, generating the control strain and the ATPase strain, respectively (Table 6.2).

2.2 Media and Cultivation Conditions

When needed, 100 $\mu\text{g mL}^{-1}$ of ampicillin was added to cultures. For growth assays, 3 mL of LB0 medium (5 g L⁻¹ yeast extract, 10 g⁻¹ tryptone, 5 g⁻¹ NaCl) was inoculated with the corresponding strain at 37°C and 150 rpm for 5 h. 100 μL of the LB0 culture was used to inoculate 50 mL of minimal medium (MM) adapted from Tanaka et al. (Tanaka et al., 1967) with the pH adjusted to 7.0 and 0.4% of glucose added as the sole carbon source. The expression of ATPase was induced with 0.1 mM (for growth-coupled production) or 0.5 mM (for growth-decoupled production) of m-toluate, and the medium was incubated without shaking at 37°C overnight. For growth-coupled production, cells from the overnight culture were washed and used to inoculate fresh MM (with 0.1 mM of m-toluate) to an optical density at 420 nm (OD₄₂₀) of 0.2. The medium was filled into 5 mL screw-cap glass vials (completely filled to the top), and the vials were incubated at

Table 6.2: Strains, plasmids, and primers used in this study.

Strain or plasmid	Relevant characteristics	Source
<i>E. coli</i> NEB 5-alpha	Competent cells for heat shock transformation	New England Biolabs
<i>E. coli</i> MG1655	<i>E. coli</i> wild type	Blattner et al., 1997
Control strain	<i>E. coli</i> MG1655 transformed with pSB43.1	This study
ATPase strain	<i>E. coli</i> MG1655 transformed with pSB44.1	This study
pCP42::atpAGD	<i>atpAGD</i> under control of CP41- <i>lacLM</i> promoter, <i>Erm^r</i>	Koebmann et al., 2002
pSB-M1g-1-17	<i>m</i> -Toluate inducible <i>xylS/Pm</i> promoter (variant ML1-17), <i>gfpmut3</i> , <i>Kan^r</i>	Balzer et al., 2013
pKD3	Donor for <i>Amp^r</i> -cassette, <i>Amp^r</i>	Datsenko and Wanner, 2000
pSB38.2	pSB-M1g-1-17 derivative, kanamycin resistance cassette switched to ampicillin resistance cassette, <i>Amp^r</i>	This study
pSB43.1	pSB38.2 without <i>gfpmut3</i> reporter gene (empty control plasmid), <i>Amp^r</i>	This study
pSB44.1	pSB43.1 with <i>atpAGD</i> gene (ATPase plasmid), <i>Amp^r</i>	This study
Primer	Sequence (5' → 3')	
Amp_PstI_fw	CGTACTGCAGAATGTGCGCGGAACCCCTATTTG	
Amp_PstI_rv	CGTACTGCAGCGTACTATCAACAGGTTGAAC	
atpAGD_mono_fw	CATGAACATATGCAACTGAATTCACCGAAATC	
atpAGD_mono_rv	CTAGAGGATCCTTAAAGTTTTTTGGCTTTTTC	

37°C without shaking. For every time point, new vials were opened to guarantee anaerobic conditions. For growth-decoupled production, cultivation conditions were the same as described above, but MM without $(\text{NH}_4)_2\text{SO}_4$ as a nitrogen source was used. The medium was inoculated with an OD420 of 2, and 0.5 mM of *m*-toluate was added for ATPase expression. Cell growth was monitored measuring the OD420 and using a factor of 0.22 to convert one OD420 unit to gram cell dry weight (g_{CDW}) L^{-1} . All cultivations were performed in biological triplicates.

2.3 Analytical Methods

Extracellular glucose and ethanol concentrations in the medium were measured using the corresponding kits from Megazyme. Extracellular lactate, acetate, formate, and succinate were quantified by the high-performance liquid chromatography (HPLC) method described by Harder et al. (Harder et al., 2016) and using a mix of the organic acids as an external standard. Pyruvate, orotate, and fumarate were measured by the same method but were not secreted in significant amounts by the strains. The ATPase activity of the cell lysate was measured using the ATPase Activity Assay Kit (Colorimetric) from BioVision (# K417). Cells from 15 mL of medium (OD420 1.7–2.3) were harvested by centrifugation, and the cell pellet was resuspended in the supplied buffer (175 μL buffer/OD420). The cells were disrupted by sonication, and the ATPase activity in the lysate was measured according to the manufacturer's protocol. ATPase activity was normalized to the overall protein content of the lysate, which was measured by the method described by Bradford (Bradford, 1976).

For determination of intracellular adenosine monophosphate (AMP), adenosine diphosphate (ADP), and ATP concentrations, cells (≈ 0.5 mg of biomass) were applied to filter

disks (polyvinylidene difluoride, 0.45 μm , 25 mm; Merck-Millipore), while a N_2 flow was used to keep the environment oxygen-free. The medium was removed by suction filtration, and the filter disks were immediately transferred to 1 mL of a -20°C cold acetonitrile/methanol/water (40:40:20) quenching solution. After incubation at -20°C for 30 min, the samples were shaken vigorously, and 500 μL of the mixture was centrifuged at 13 000 rpm and -9°C for 15 min. Next, 400 μL of the supernatant was kept at -80°C until metabolite quantification by liquid chromatography-tandem mass spectrometry, which was performed as previously described (Guder et al., 2017) using an Agilent 6495 triple quadrupole mass spectrometer (Agilent Technologies). Absolute ATP, ADP, and AMP concentrations were determined with ^{13}C internal standard and authentic standards (Guder et al., 2017). The intracellular adenosine energy charge was calculated with the formula $([\text{ATP}] + 0.5[\text{ADP}])/([\text{ATP}] + [\text{ADP}] + [\text{AMP}])$.

For growth-coupled production, the growth, productivity, and glucose uptake rates were calculated for the exponential growth phase, while the yield and titer were determined at the end of the cultivation. For growth-decoupled production, biomass concentration was assumed to be constant and the average of the measured biomass concentrations during the cultivation was used for calculating productivity and glucose uptake rate.

3 Results

3.1 Construction of an Inducible ATP-Wasting Mechanism for *E. coli*

In order to have a dynamic and gradual control over the expression of the ATP wasting machinery, the genes encoding the F1-part (consisting of the α -, γ -, and β -subunits) of the ATPase from *E. coli* (atpAGD) were cloned into the XylS/Pm expression system, inducible by the addition of m-toluate (Balzer et al., 2013). Herein, we used higher inducer concentrations for the growthdecoupled than for the growth-coupled batch process to obtain the highest ATPase expression under starvation.

3.2 Effect of ATP Wasting on Growth-Coupled Product Formation

The effect of ATPase overexpression was first tested during growth-coupled production of fermentation products (OSF) under anaerobic conditions. Plasmids pSB43.1 (empty vector) and pSB44.1 (ATPase expression vector) were transformed into *E. coli* MG1655; in the following, these two strains are referred to as control strain and ATPase strain, respectively. Both strains were cultivated anaerobically, and as the formation of the main fermentation products - ethanol, formate, acetate, lactate, and succinate - is naturally coupled to ATP synthesis, they were used as proxies to evaluate the influence of ATP wasting on product synthesis.

Overexpression of the ATPase increased the specific glucose uptake rate by 18.5% and the specific productivity by 17.2% compared to the control strain (Table 6.3 and Figure 6.1C). Note that for a better comparability, the sum of all carbon atoms incorporated in the five main fermentation products was considered to be the overall product. At the same time, the yield and titer also increased by 6.8% and 8.7%, respectively (Figure 6.1E,F). With a yield of 88% product C-atoms/glucose C-atoms, the control strain is already close to the theoretical maximum yield. With the help of ATP wasting, more than half of the remaining 12% could be captured in the form of fermentative products.

In contrast, the expression of the ATPase decreased the growth rate, which dropped from 0.40 h^{-1} (control strain) to 0.28 h^{-1} (ATPase strain) (Figure 6.1A and Table 6.3). Similarly, although the specific productivity increased significantly, the volumetric productivity of the ATPase strain decreased by 25.3% compared to the control strain during exponential growth (Figure 6.1D). This is due to the lower growth rate and the associ-

Table 6.3: Summary of performance parameters with and without ATP wasting under growth-coupled and growth-decoupled production conditions.

	Growth-coupled production		Growth-decoupled production	
	Control strain	ATPase strain	Control strain	ATPase strain
μ [h^{-1}]	0.40 ± 0.01	0.28 ± 0.01	≈ 0	≈ 0
r_{Glucose} [$\text{mmol gCDW}^{-1} \text{h}^{-1}$]	11.68 ± 0.35	13.84 ± 0.51	2.82 ± 0.02	6.49 ± 0.19
q_{Glucose} [$\text{mmol L}^{-1} \text{h}^{-1}$]	1.97 ± 0.11	1.46 ± 0.04	1.29 ± 0.01	2.75 ± 0.08
$r_{\Sigma \text{ C-atoms in products}}$ [$\text{mmol gCDW}^{-1} \text{h}^{-1}$]	66.31 ± 2.09	77.75 ± 2.52	15.37 ± 0.71	37.30 ± 2.49
$q_{\Sigma \text{ C-atoms in products}}$ [$\text{mmol L}^{-1} \text{h}^{-1}$]	11.19 ± 0.60	8.36 ± 0.33	7.02 ± 0.39	15.79 ± 0.99
$Y_{\Sigma \text{ C-atoms in products/ glucose C-atoms}}$ [mol mol^{-1}]	0.88 ± 0.03	0.94 ± 0.02	0.92 ± 0.03	1.01 ± 0.08

ated decrease in biomass, which acts as the biocatalyst. At the beginning of the cultivation (the first ≈ 3.5 h), the higher specific productivity can compensate for the reduced biomass formation since product concentrations are slightly higher in the ATPase strain than in the control strain even though less biomass is present (Figures 6.1B and 6.2B). In addition to the volumetric productivity, the volumetric glucose uptake rate of the ATPase strain is also higher at the beginning of the cultivation (Figure 6.2A). However, when the difference in biomass surpasses a certain value, the higher specific productivity cannot make up for the reduced amount of biomass and only at the end of the cultivation, the product concentration of the ATPase strain surpasses the control strain, leading to a higher yield and titer.

3.3 Determination of ATPase Activity and Intracellular ATP Concentration

To examine whether the observed effects on growth, productivity, and yield are due to an increased ATPase activity in the cytosol and not a result of the burden to overexpress three proteins (sum of molecular weights of the three ATPase subunits: 137.12 kDa), the ATPase activity in the lysate of both strains was measured. With 21.5 mU mg^{-1} protein, the ATPase activity was almost twice as high in the ATPase strain as in the control, showing that the overexpression leads indeed to an increased ATPase activity in the cytosol (Figure 6.3B).

Furthermore, the influence of the expressed ATPase on the intracellular concentrations of AMP, ADP, and ATP was examined. Samples for intracellular metabolite quantification were taken at the end of the exponential growth phase (after 7.3 h). The mea-

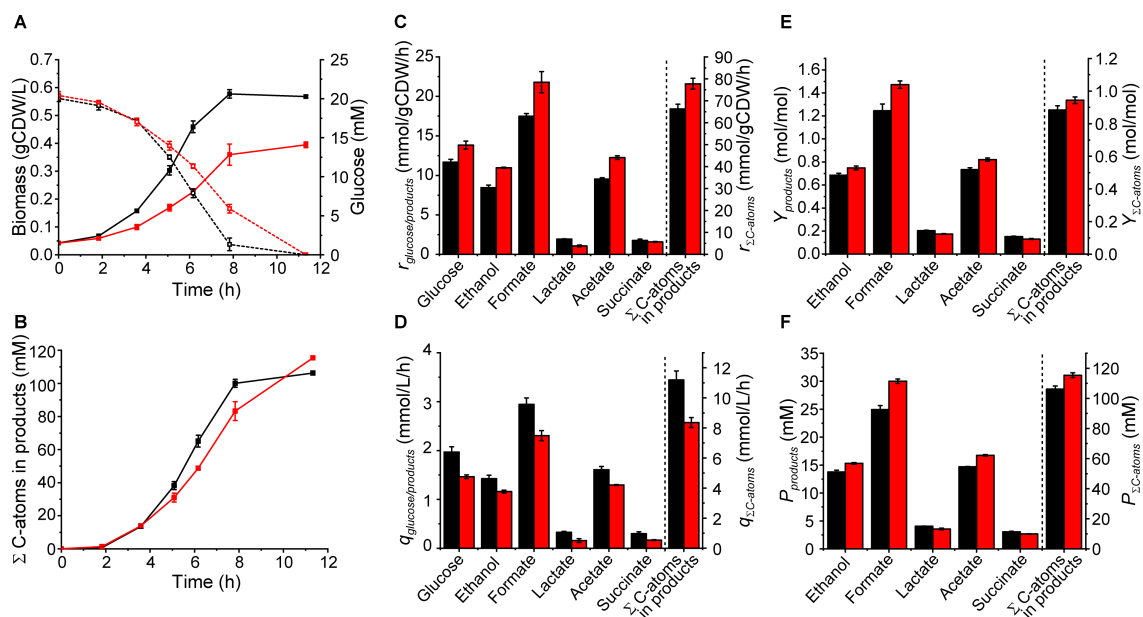


Figure 6.1: Concentration profiles, productivities, yield, and titer of the control strain (black) and the ATPase strain (red) under growth-coupled production conditions. A) Biomass concentration (solid line) and glucose concentration (dashed line) during cultivation, B) product (sum of carbon atoms incorporated in main fermentation products) concentration during cultivation, C) specific glucose uptake rate and specific productivity, D) volumetric glucose uptake rate and volumetric productivity, E) yield, and F) titer.

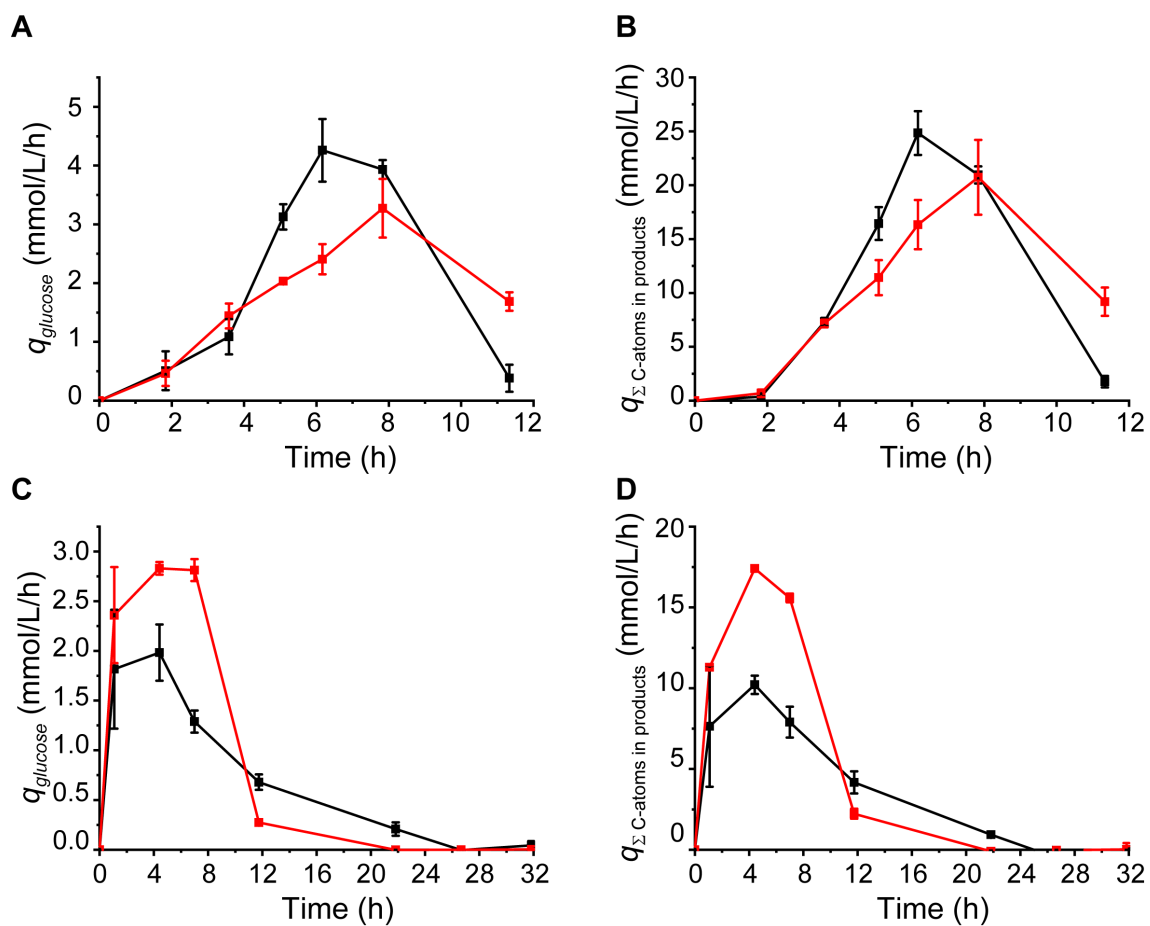


Figure 6.2: A,C) Volumetric glucose uptake rate and B,D) volumetric productivity (A,B) during growth-coupled and (C,D) growth-decoupled production of the control strain (black) and the ATPase strain (red).

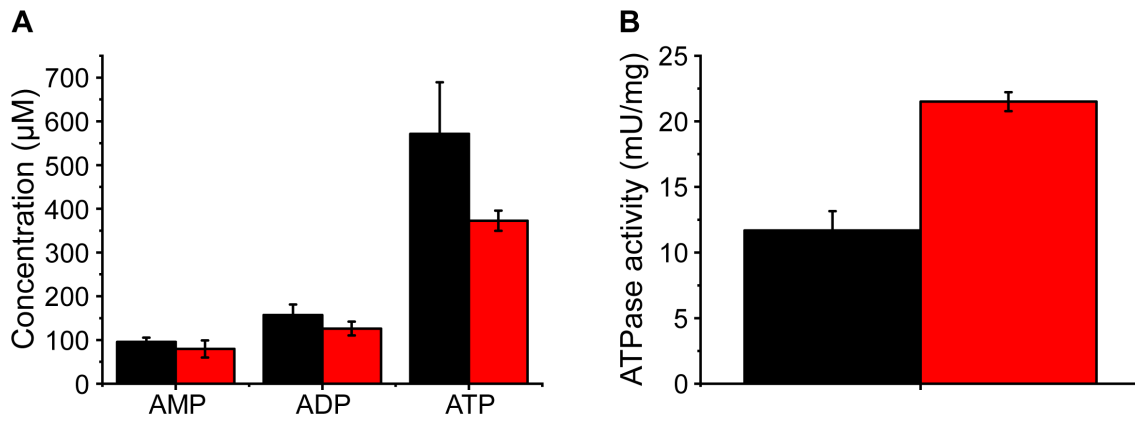


Figure 6.3: A) Intracellular concentrations of AMP, ADP, and ATP in control strain (black) and ATPase strain (red). B) ATPase activity of cell lysate of control strain (black) and ATPase strain (red).

measurements revealed that the concentration of intracellular ATP in the ATPase strain was 34.7% lower than in the control strain (Figure 6.3A). Likewise, the $[\text{ATP}]/[\text{ADP}]$ ratios also dropped in the ATPase strain (2.99 ± 0.28) compared to the control strain (3.74 ± 1.09), while the adenosine energy charge decreased only slightly from 0.78 ± 0.04 (control strain) to 0.76 ± 0.02 (ATPase strain). The only mild reduction of the energy charge is a consequence of the fact that the total adenosine pool $[\text{ATP}] + [\text{ADP}] + [\text{AMP}]$ in the ATPase strain was significantly smaller.

3.4 Effect of ATP Wasting on Growth-Decoupled Product Formation

The issue of decelerated cell growth of the ATPase strain and the associated lower volumetric productivity in an OSF batch process can be circumvented if growth and production phases are decoupled in a two-stage fermentation (TSF). We therefore examined the behavior and performance of the ATPase strain under growth-arrested conditions (mimicking the second [production] phase of a TSF) by using production medium without the addition of a nitrogen source. The medium was inoculated with a tenfold higher cell density compared to the growth-coupled production conditions. Except for a small biomass increase of the control strain at the beginning of the cultivation, the biomass concentration stayed constant or decreased slightly throughout the 32-h cultivation period (Figure 6.4A). ATP wasting under these conditions led to a peak glucose uptake rate of $6.78 \pm 0.28 \text{ mmol g}_{CDW}^{-1} \text{ h}^{-1}$, which stayed in that range until all of the glucose was consumed (Figures 6.2C and 6.4C). The control strain reached a peak value of only

$4.22 \pm 0.60 \text{ mmol g}_{CDW}^{-1} \text{ h}^{-1}$ at the beginning of the cultivation, which then continuously decreased and went down to close to zero after 32 h, although one-fourth of the added glucose was still present in the medium (Figures 6.2C and 6.4A). Considering the mean values within the first 11.75 h of cultivation, the specific glucose uptake rate with ATP wasting was more than twice as high ($6.49 \pm 0.19 \text{ mmol g}_{CDW}^{-1} \text{ h}^{-1}$), than without wasting ($2.82 \pm 0.02 \text{ mmol g}_{CDW}^{-1} \text{ h}^{-1}$; Figure 6.4C). Consequently, the specific productivity of the ATPase strain was more than 142% higher than of the control strain. Interestingly, the productivity of lactate was especially upregulated and more than twelve times higher (Figure 6.4C).

We hypothesize that under nitrogen starvation conditions, the lactate pathway (with a yield of 2 mol ATP per mol glucose) may either have a higher saturation or/and it is preferred for redox balancing, as it requires only one enzymatic step from pyruvate opposed to several enzyme steps needed for the ATP yield optimal pathway (2.5 mol ATP mol⁻¹ glucose) with formation of formate, ethanol, and acetate. The slightly lower ATP yield may pay off due to the reduction in enzyme synthesis costs, which are especially crucial under nitrogen limitation. In contrast to the growth-coupled production conditions, where a higher biomass concentration of the control strain led to a higher volumetric productivity, biomass concentrations in growth-decoupled production conditions were the same for both strains and stayed constant. Thus, the differences in specific glucose uptake rate and specific productivity between the control and the ATPase strain could be directly transferred to the volumetric uptake rate and volumetric productivity (Figures 6.2D, 6.4D).

The yield of the ATPase strain reached the maximum with 1.01 ± 0.08 product C-atoms/glucose C-atoms. In theory, the control strain should also reach the maximal possible yield (Klamt et al., 2018) but stayed with 0.92 ± 0.03 product C-atoms/glucose C-atoms somewhat below the maximum, which could be due to the increase of biomass (and thus a flux of carbon to the biomass) at the beginning of the cultivation as mentioned above. Because not all of the added glucose was consumed by the control strain, the product titer was 27.57% lower than in the ATPase strain (Figure 6.4B,F).

Using the experimentally determined exchange rates and a stoichiometric model of the central metabolism of *E. coli* Hädicke and Klamt, 2017 and the software CellNetAnalyzer (Klamt et al., 2007; von Kamp et al., 2017), we estimated the nongrowth associated ATP maintenance demand in both strains (in the ATPase strain, this value includes the amount of ATP hydrolyzed by the overexpressed ATPase) and obtained a value of $5.34 \text{ mmol ATP g}_{CDW}^{-1} \text{ h}^{-1}$ in the control strain and $12.62 \text{ mmol ATP g}_{CDW}^{-1} \text{ h}^{-1}$ in the AT-

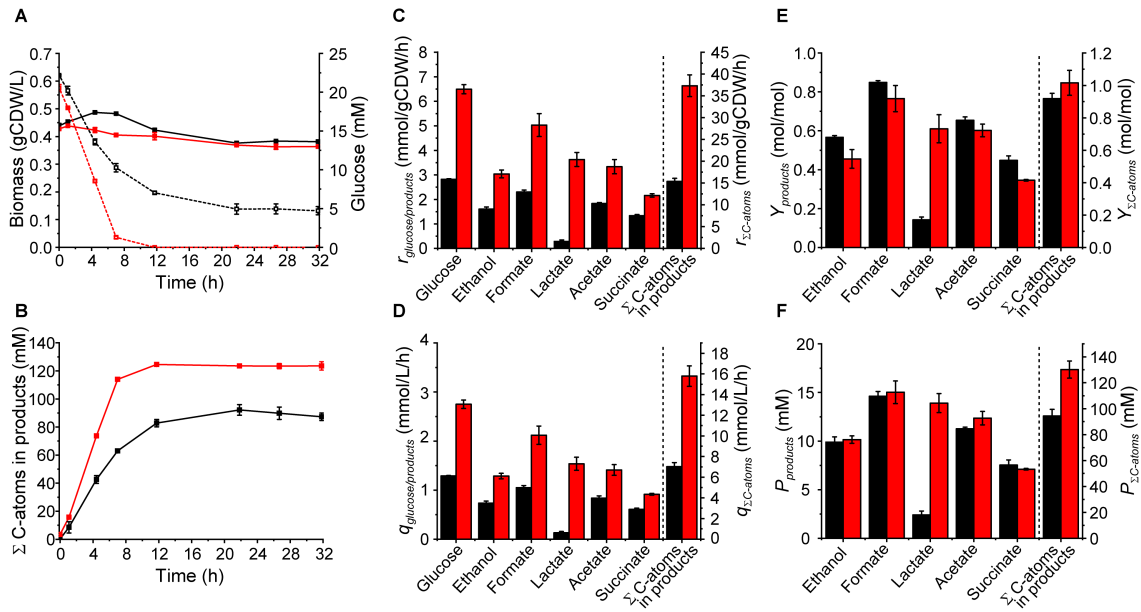


Figure 6.4: Concentration profiles, productivities, yield and titer of the control strain (black) and the ATPase strain (red) under growth-decoupled production conditions. A) Biomass concentration (solid line) and glucose concentration (dashed line) during cultivation, B) product (sum of carbon atoms incorporated in main fermentation products) concentration during cultivation, C) specific glucose uptake rate and specific productivity, D) volumetric glucose uptake rate and volumetric productivity, E) yield, and F) titer.

Pase strain (data not shown). Hence, the expression of the ATPase more than doubles the ATP demand (and metabolic activity) in the nongrowing cells. The factor for the ATP demand is thus well in the range of the measured doubled ATPase activity (Figure 6.3B) in the ATPase strain.

4 Discussion

The goal of the present study was to further examine and further develop the concept of enforced ATP wasting and to broaden its scope for potential applications in metabolic engineering. We studied aspects that have not been investigated before or/and not in the combination as used herein (see Table 1). In particular, i) we developed a new genetic module for dynamic induction of an ATPase; ii) we showed that the F₁-part of the ATPase leads to higher metabolic activity and increased product formation in *E. coli* also under anaerobic conditions; and iii) we tested and proved suitability of ATP wasting also for growth-arrested production which is vital for its use in TSF processes. To keep things simple, in our example process, we considered the total amount of fermentation products as a proxy for the target chemical(s); however, in realistic applications, one may focus on a particular metabolite by deleting pathways to other fermentation products as was done, e.g., for lactate (Hädicke and Klamt, 2015).

Our results with the ATPase strain for growth-coupled product synthesis confirmed earlier results of increased ATP demand in *E. coli* or *L. lactis* (Hädicke et al., 2015; Liu et al., 2016; Koebmann et al., 2002), namely an increase in substrate uptake, specific productivity, product yield, and titer accompanied with a reduced growth rate and biomass yield, causing reduced volumetric productivities. Hence, generalizing results found for aerobic conditions (Koebmann et al., 2002), the flux through glycolysis in *E. coli* wild-type cells is governed by ATP-consuming processes even under anaerobic conditions. However, as the necessary redirection of carbon flux from biomass to ATP formation (coupled with product synthesis) reduced the growth rate of the ATPase strain, less biomass as biocatalyst was present and the volumetric productivity dropped by 25.3% compared to the control strain. Lowered volumetric productivities were also observed in Hädicke et al. (Hädicke et al., 2015) and Liu et al. (Liu et al., 2016), however, Liu et al. (Liu et al., 2016) also reported that fine-tuning of the ATPase activity by selecting a (constitutive) promoter with moderate strength led to a *L. lactis* strain that exhibited even a slightly higher volumetric productivity for acetoin than the control strain. It is important to note that the amount of substrate used, the reference time point, growth inhibition by the accumulated product and other factors determine the relative volumetric productivity of an OSF with ATP wasting against an OSF without ATP wasting. As was already mentioned in Section 3, when we stop the fermentation within the first 3 hours, the volumetric productivity of the ATPase strain would be higher than of the control strain. However, in general, it is very likely that the volumetric productivity of a strain

with enforced ATP wasting will be lower than a strain without increased ATP demand, reflecting the trade-off between high product yield and high volumetric productivity. Here, adjusting the strength of the ATPase activity enables one to adjust product yield and productivity at a desired point with an optimal trade-off. Compared to the approach followed by Liu et al. (Liu et al., 2016), where different strains were constructed each having a static constitutive promoter of a specific strength, our developed system with an inducible ATPase promoter offers a more practical solution for fine tuning the ATPase expression and for studying the influence of different expression levels on cell physiology. Moreover, the inducible promoter now allows applications with dynamic control of the ATPase activity, which will be essential for TSF processes where enforced ATP wasting is desired only in the production phase. For industrial applications, however, external inducers are often too costly, which may favor the use of autonomous switches, e.g., based on quorum sensing (Gupta et al., 2017).

Generally, TSF processes separating growth and production may help to overcome the drawback of reduced volumetric productivity during OSF (Burg et al., 2016; Venayak et al., 2015; David et al., 2016; Nembr et al., 2018). [20,21,39,40] In Klamt et al. (Klamt et al., 2018) we hypothesized that ATP wasting in the production phase could further boost the performance of TSFs. We therefore analyzed the effect of overexpressed ATPase under growth-arrested conditions (mimicking the production phase of a TSF) caused by nitrogen starvation, where the biomass of the ATPase strain and the control strain remain constant. As was shown earlier, the metabolism usually shuts down in nongrowing cells just to cover cellular maintenance, leading to low substrate uptake rates and productivities and thus severely limiting TSF processes (Harder et al., 2018; Chubukov and Sauer, 2014). This can also be observed in the present study: the specific glucose uptake rate of the control strain slows down under nitrogen starvation conditions until it comes almost to a complete rest even though the substrate was not completely consumed. In recent years, there have been several attempts to maintain high metabolic rates in the stationary phase, e.g. by overexpression of PtsI in *E. coli*, which is involved in the regulation of the glucose uptake machinery (Chubukov et al., 2017; Gosset, 2005), by modulating the stringent response program leading to a "high glucose throughput (HGT)" strain (Michalowski et al., 2017), and by directed evolution (Sonderegger et al., 2005). To the best of our knowledge, the glucose uptake rate of $6.49 \text{ mmol g}_{CDW} \text{ h}^{-1}$ we achieved with enforced ATP wasting is the highest reported for nongrowing *E. coli* cells. For example, although not directly comparable (anaerobic conditions in our study and aerobic conditions in the studies mentioned above), the specific glucose uptake rate in the PtsI overex-

pression strain was $\approx 2.5 \text{ mmol g}_{CDW}^{-1} \text{ h}^{-1}$ (Chubukov et al., 2017), $3.27 \text{ mmol g}_{CDW}^{-1} \text{ h}^{-1}$ in the HGT strain (Michalowski et al., 2017), and $1.55 \text{ mmol g}_{CDW}^{-1} \text{ h}^{-1}$ in the evolved strain (Sonderegger et al., 2005). Moreover, the ATPase strain keeps these high rates - in contrast to the control strain - until all of the glucose is completely consumed. The resulting high volumetric productivity during stationary phase could help make TSF processes more efficient as envisioned in the theoretical study of our group (Klamt et al., 2018). In contrast, if yield and titer are more important than productivity (e.g., for high-value products), ATP wasting can be applied in a growth-coupled mode in an OSF as described above.

Our measurements of the adenosine pools in the ATPase and the control strain in the growth-coupled process showed, as expected, reduced ATP levels and a lowered ATP/ADP ratio and thus followed the same trend as in Koebmann et al. (Koebmann et al., 2002). Interestingly, and also consistent with the study of Koebmann et al. (Koebmann et al., 2002), we found that the concentrations of AMP, ADP, and ATP were all lower in the ATPase strain compared to the control strain (Figure 6.3A). As a consequence, energy charges of the ATPase strain (0.76), as well as the control strain (0.78), differed only to a minor extent, meaning that *E. coli* is still able to counterbalance the increased ATP demand caused by the expressed ATPase, at least in terms of energy charge. In future studies, the tipping point between maximizing the glucose uptake rate and overburdening the cells with too high ATP wasting rates should be determined by titrating the expression level of the ATPase. Furthermore, we plan to test the approach for other (heterologous) products and substrates in *E. coli* and other organisms and to fine-tune the ATP wasting level for the respective product.

In summary, we showed that yield and titer, as well as volumetric productivity, can be improved by enforced ATP wasting depending on which cultivation approach (OSF or TSF) is being applied. We believe that this concept can become a general approach in bioprocess and metabolic engineering to construct microbial cell factories with superior performance.

Acknowledgements

The authors thank Peter R. Jensen for providing plasmid pCP41::atpAGD and Trygve Brautaset for providing plasmid pSB-M1g-1-17. This work was supported by the ERC Consolidator Grant "StrainBooster" (grant number 721176), by the Deutsche Forschungsgemeinschaft grant LI 1993/2-1, and by the ERC Starting Grant 715650.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

ATPase, ATP wasting, Escherichia coli, metabolic engineering, nitrogen starvation, one-stage fermentation, two-stage fermentation

References

- Boecker, S., Zahoor, A., Schramm, T., Link, H., & Klamt, S. (2019). Broadening the scope of enforced ATP wasting as a tool for metabolic engineering in *Escherichia coli*. *Biotechnology Journal*, *14*(9), 1800438. <https://doi.org/10.1002/biot.201800438> (cited on page 181).
- Liao, J. C., Mi, L., Pontrelli, S., & Luo, S. (2016). Fuelling the future: Microbial engineering for the production of sustainable biofuels. *Nature Reviews Microbiology*, *14*(5), 288–304. <https://doi.org/10.1038/nrmicro.2016.32> (cited on page 183).
- Becker, J., & Wittmann, C. (2015). Advanced biotechnology: Metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products. *Angewandte Chemie International Edition*, *54*(11), 3328–3350. <https://doi.org/10.1002/anie.201409033> (cited on page 183).
- Lee, S. Y., Kim, H. U., Chae, T. U., Cho, J. S., Kim, J. W., Shin, J. H., Kim, D. I., Ko, Y.-S., Jang, W. D., & Jang, Y.-S. (2019). A comprehensive metabolic map for production of bio-based chemicals. *Nature Catalysis*, *2*(1), 18–33. <https://doi.org/10.1038/s41929-018-0212-4> (cited on page 183).
- Cho, C., Choi, S. Y., Luo, Z. W., & Lee, S. Y. (2015). Recent advances in microbial production of fuels and chemicals using tools and strategies of systems metabolic engineering. *Biotechnology Advances*, *33*(7), 1455–1466. <https://doi.org/10.1016/j.biotechadv.2014.11.006> (cited on page 183).
- Keasling, J. D. (2010). Manufacturing molecules through metabolic engineering. *Science*, *330*(6009), 1355–1358. <https://doi.org/10.1126/science.1193990> (cited on page 183).
- Hädicke, O., & Klamt, S. (2015). Manipulation of the ATP pool as a tool for metabolic engineering. *Biochemical Society Transactions*, *43*(6), 1140–1145. <https://doi.org/10.1042/BST20150141> (cited on pages 183, 184, 197).
- Singh, A., Cher Soh, K., Hatzimanikatis, V., & Gill, R. T. (2011). Manipulating redox and ATP balancing for improved production of succinate in *e. coli*. *Metabolic Engineering*, *13*(1), 76–81. <https://doi.org/10.1016/j.ymben.2010.10.006> (cited on page 183).
- Hädicke, O., Bettenbrock, K., & Klamt, S. (2015). Enforced ATP futile cycling increases specific productivity and yield of anaerobic lactate production in *escherichia coli*. *Biotechnology and Bioengineering*, *112*(10), 2195–2199. <https://doi.org/10.1002/bit.25623> (cited on pages 183–185, 197).
- Bj, K., Hv, W., JI, S., D, N., & Pr, J. (2002). The glycolytic flux in *escherichia coli* is controlled by the demand for ATP. *Journal of Bacteriology*, *184*(14), 3909–3916. <https://doi.org/10.1128/jb.184.14.3909-3916.2002> (cited on page 183).
- Liu, J., Kandasamy, V., Würtz, A., Jensen, P. R., & Solem, C. (2016). Stimulation of acetoin production in metabolically engineered *lactococcus lactis* by increasing ATP demand. *Applied Microbiology and Biotechnology*, *100*(22), 9509–9517. <https://doi.org/10.1007/s00253-016-7687-1> (cited on pages 183–185, 197, 198).
- de Kok, S., Kozak, B. U., Pronk, J. T., & van Maris, A. J. A. (2012). Energy coupling in *saccharomyces cerevisiae*: Selected opportunities for metabolic engineering. *FEMS Yeast Research*, *12*(4), 387–397. <https://doi.org/10.1111/j.1567-1364.2012.00799.x> (cited on page 183).
- Koebmann, B. J., Westerhoff, H. V., Snoep, J. L., Solem, C., Pedersen, M. B., Nilsson, D., Michelsen, O., & Jensen, P. R. (2002). The extent to which ATP demand controls the glycolytic flux depends strongly on the organism and conditions for growth. *Molecular Biology Reports*, *29*(1), 41–45. <https://doi.org/10.1023/A:1020398117281> (cited on pages 183, 185, 188, 197, 199).

- Kim, H.-J., Kwon, Y. D., Lee, S. Y., & Kim, P. (2012). An engineered escherichia coli having a high intracellular level of ATP and enhanced recombinant protein production. *Applied Microbiology and Biotechnology*, 94(4), 1079–1086. <https://doi.org/10.1007/s00253-011-3779-0> (cited on page 183).
- Semkiv, M. V., Dmytruk, K. V., Abbas, C. A., & Sibirny, A. A. (2016). Activation of futile cycles as an approach to increase ethanol yield during glucose fermentation in *saccharomyces cerevisiae*. *Bioengineered*, 7(2), 106–111. <https://doi.org/10.1080/21655979.2016.1148223> (cited on pages 183–185).
- Jensen, P. R., Snoep, J. L., & Westerhoff, H. V. (2003, January 28). *Method of improving the production of biomass or a desired product from a cell* (pat. No. 6511836B1). Retrieved November 2, 2020, from <https://patents.google.com/patent/US6511836B1/en> (cited on page 183).
- Patnaik, R., Roof, W. D., Young, R. F., & Liao, J. C. (1992). Stimulation of glucose catabolism in escherichia coli by a potential futile cycle. *Journal of Bacteriology*, 174(23), 7527–7532. Retrieved November 2, 2020, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC207462/> (cited on page 183).
- Holm, A. K., Blank, L. M., Oldiges, M., Schmid, A., Solem, C., Jensen, P. R., & Vemuri, G. N. (2010). Metabolic and transcriptional response to cofactor perturbations in escherichia coli. *The Journal of Biological Chemistry*, 285(23), 17498–17506. <https://doi.org/10.1074/jbc.M109.095570> (cited on pages 183, 185).
- Chao, Y. P., & Liao, J. C. (1994). Metabolic responses to substrate futile cycling in escherichia coli. *The Journal of Biological Chemistry*, 269(7), 5122–5126 (cited on pages 183, 185).
- Burg, J. M., Cooper, C. B., Ye, Z., Reed, B. R., Moreb, E. A., & Lynch, M. D. (2016). Large-scale bioprocess competitiveness: The potential of dynamic metabolic control in two-stage fermentations. *Current Opinion in Chemical Engineering*, 14, 121–136. <https://doi.org/10.1016/j.coche.2016.09.008> (cited on pages 184, 198).
- Venayak, N., Anesiadis, N., Cluett, W. R., & Mahadevan, R. (2015). Engineering metabolism through dynamic control. *Current Opinion in Biotechnology*, 34, 142–152. <https://doi.org/10.1016/j.copbio.2014.12.022> (cited on pages 184, 185, 198).
- Klamt, S., Mahadevan, R., & Hädicke, O. (2018). When do two-stage processes outperform one-stage processes? *Biotechnology Journal*, 13(2), 1700539. <https://doi.org/10.1002/biot.201700539> (cited on pages 184, 195, 198, 199).
- Harder, B.-J., Bettenbrock, K., & Klamt, S. (2018). Temperature-dependent dynamic control of the TCA cycle increases volumetric productivity of itaconic acid production by *Escherichia coli*. *Biotechnology and Bioengineering*, 115(1), 156–164. <https://doi.org/10.1002/bit.26446> (cited on pages 184, 198).
- Chubukov, V., & Sauer, U. (2014). Environmental dependence of stationary-phase metabolism in *bacillus subtilis* and *escherichia coli*. *Applied and Environmental Microbiology*, 80(9), 2901–2909. <https://doi.org/10.1128/AEM.00061-14> (cited on pages 184, 198).
- Venayak, N., Raj, K., Jaydeep, R., & Mahadevan, R. (2018). An optimized bistable metabolic switch to decouple phenotypic states during anaerobic fermentation. *ACS Synthetic Biology*, 7(12), 2854–2866. <https://doi.org/10.1021/acssynbio.8b00284> (cited on page 185).
- Balzer, S., Kucharova, V., Megerle, J., Lale, R., Brautaset, T., & Valla, S. (2013). A comparative analysis of the properties of regulated promoter systems commonly used for recombinant gene expression

- in *Escherichia coli*. *Microbial Cell Factories*, 12(1), 26. <https://doi.org/10.1186/1475-2859-12-26> (cited on pages 185, 188, 190).
- Clark, D. P. (1989). The fermentation pathways of *Escherichia coli*. *FEMS Microbiology Reviews*, 5(3), 223–234. <https://doi.org/10.1111/j.1574-6968.1989.tb03398.x> (cited on page 185).
- Sambrook, J., & Russell, D. W. (2001). *Molecular cloning: A laboratory manual* (3rd ed). Cold Spring Harbor Laboratory Press. (Cited on page 187).
- Tanaka, S., Lerner, S. A., & Lin, E. C. C. (1967). Replacement of a phosphoenolpyruvate-dependent phosphotransferase by a nicotinamide adenine dinucleotide-linked dehydrogenase for the utilization of mannitol. *Journal of Bacteriology*, 93(2), 642–648. Retrieved November 2, 2020, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC276489/> (cited on page 187).
- Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., & Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science (New York, N.Y.)*, 277(5331), 1453–1462. <https://doi.org/10.1126/science.277.5331.1453> (cited on page 188).
- Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, 97(12), 6640–6645. <https://doi.org/10.1073/pnas.120163297> (cited on page 188).
- Harder, B.-J., Bettenbrock, K., & Klamt, S. (2016). Model-based metabolic engineering enables high yield itaconic acid production by *Escherichia coli*. *Metabolic Engineering*, 38, 29–37. <https://doi.org/10.1016/j.ymben.2016.05.008> (cited on page 188).
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72(1), 248–254. [https://doi.org/10.1016/0003-2697\(76\)90527-3](https://doi.org/10.1016/0003-2697(76)90527-3) (cited on page 188).
- Guder, J. C., Schramm, T., Sander, T., & Link, H. (2017). Time-optimized isotope ratio LC–MS/MS for high-throughput quantification of primary metabolites. *Analytical Chemistry*, 89(3), 1624–1631. <https://doi.org/10.1021/acs.analchem.6b03731> (cited on page 189).
- Hädicke, O., & Klamt, S. (2017). EColiCore2 : A reference network model of the central metabolism of *Escherichia coli* and relationships to its genome-scale parent model. *Scientific Reports*, 7(1), 39647. <https://doi.org/10.1038/srep39647> (cited on page 195).
- Klamt, S., Saez-Rodriguez, J., & Gilles, E. D. (2007). Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC Systems Biology*, 1(1), 2. <https://doi.org/10.1186/1752-0509-1-2> (cited on page 195).
- von Kamp, A., Thiele, S., Hädicke, O., & Klamt, S. (2017). Use of CellNetAnalyzer in biotechnology and metabolic engineering. *Journal of Biotechnology*, 261, 221–228. <https://doi.org/10.1016/j.jbiotec.2017.05.001> (cited on page 195).
- Gupta, A., Reizman, I. M. B., Reisch, C. R., & Prather, K. L. J. (2017). Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nature Biotechnology*, 35(3), 273–279. <https://doi.org/10.1038/nbt.3796> (cited on page 198).
- David, F., Nielsen, J., & Siewers, V. (2016). Flux control at the malonyl-CoA node through hierarchical dynamic pathway regulation in *Saccharomyces cerevisiae*. *ACS Synthetic Biology*, 5(3), 224–233. <https://doi.org/10.1021/acssynbio.5b00161> (cited on page 198).

- Nemr, K., Müller, J. E. N., Joo, J. C., Gawand, P., Choudhary, R., Mendonca, B., Lu, S., Yu, X., Yakunin, A. F., & Mahadevan, R. (2018). Engineering a short, aldolase-based pathway for (r)-1,3-butanediol production in *escherichia coli*. *Metabolic Engineering*, *48*, 13–24. <https://doi.org/10.1016/j.ymben.2018.04.013> (cited on page 198).
- Chubukov, V., Desmarais, J. J., Wang, G., Chan, L. J. G., Baidoo, E. E., Petzold, C. J., Keasling, J. D., & Mukhopadhyay, A. (2017). Engineering glucose metabolism of *escherichia coli* under nitrogen starvation. *npj Systems Biology and Applications*, *3*(1), 16035. <https://doi.org/10.1038/npjbsa.2016.35> (cited on pages 198, 199).
- Gosset, G. (2005). Improvement of *escherichia coli* production strains by modification of the phosphoenolpyruvate:sugar phosphotransferase system. *Microbial Cell Factories*, *4*(1), 14. <https://doi.org/10.1186/1475-2859-4-14> (cited on page 198).
- Michalowski, A., Siemann-Herzberg, M., & Takors, R. (2017). *Escherichia coli* HGT: Engineered for high glucose throughput even under slowly growing or resting conditions. *Metabolic Engineering*, *40*, 93–103. <https://doi.org/10.1016/j.ymben.2017.01.005> (cited on pages 198, 199).
- Sonderegger, M., Schümperli, M., & Sauer, U. (2005). Selection of quiescent *escherichia coli* with high metabolic activity. *Metabolic Engineering*, *7*(1), 4–9. <https://doi.org/10.1016/j.ymben.2004.05.005> (cited on pages 198, 199).

Chapter 7

Deciphering the physiological response of *Escherichia coli* under high ATP demand

Simon Boecker¹, Giulia Slaviero¹, Thorben Schramm, Witold Szymanski, Ralf Steuer, Hannes Link, Steffen Klamt

This chapter was published in: Boecker et al., 2021, *Molecular Systems Biology*, **2021**, 17:e10504, DOI: <https://doi.org/10.15252/msb.202110504>, (license: Creative Commons "Attribution 4.0 International" (CC BY 4.0), <https://creativecommons.org/licenses/by/4.0/>, no modifications). My contributions were metabolite measurements by LC-MS/MS and processing of related data as well as sample preparation for proteomics. Related data is shown in Fig. 7.5, Fig. 7.10, and Fig. 7.11.

¹These authors contributed equally to this work

Abstract

One long-standing question in microbiology is how microbes buffer perturbations in energy metabolism. In this study, we systematically analyzed the impact of different levels of ATP demand in *Escherichia coli* under various conditions (aerobic and anaerobic, with and without cell growth). One key finding is that, under all conditions tested, the glucose uptake increases with rising ATP demand, but only to a critical level beyond which it drops markedly, even below wild-type levels. Focusing on anaerobic growth and using metabolomics and proteomics data in combination with a kinetic model, we show that this biphasic behavior is induced by the dual dependency of the phosphofructokinase on ATP (substrate) and ADP (allosteric activator). This mechanism buffers increased ATP demands by a higher glycolytic flux but, as shown herein, it collapses under very low ATP concentrations. Model analysis also revealed two major rate-controlling steps in the glycolysis under high ATP demand, which could be confirmed experimentally. Our results provide new insights on fundamental mechanisms of bacterial energy metabolism and guide the rational engineering of highly productive cell factories.

1 Introduction

The sugar and energy metabolism of microorganisms has been subject of research for many decades (Jensen and Michelsen, 1992; Kochanowski et al., 2013; Chubukov et al., 2014; Basan et al., 2020). One central goal of these studies is to decipher key principles of cellular metabolism and to uncover regulatory mechanisms that enable microorganisms to adapt to perturbations and varying environments (Chubukov et al., 2014; Bruggeman et al., 2020). The integration of various experimental data, often in combination with mathematical modeling, helps to shed light on global phenomena of microbial metabolism, such as overflow metabolism or diauxie (Bruggeman et al., 2020; Basan et al., 2015; Chen and Nielsen, 2019). However, despite the progress made, a comprehensive understanding of how microbes respond and adapt to perturbations is still lacking in many cases. This also limits our ability to rationally engineer the metabolism of microorganisms for biotechnological applications. One example of directed metabolic interventions to optimize microbial production hosts is the manipulation of the supply of ATP, the energy currency of the cell. Increasing ATP availability can, for example, lead to improved succinate (Zhang et al., 2009; Singh et al., 2011) or recombinant protein (Kim et al., 2012) production. In the opposite direction, artificially enforcing a high turnover ("wasting") of ATP can substantially increase the specific glucose uptake rate and the production rate of certain target compounds (if production of the latter is coupled with ATP formation) (Boecker et al., 2021; Chao and Liao, 1994; Koebmann et al., 2002; Hädicke et al., 2015; Boecker et al., 2019; Zahoor et al., 2020).

Studying the response of the cells to perturbed ATP levels is thus essential not only for understanding fundamental physiological processes but also for guiding metabolic engineering efforts. As one approach, several previous studies investigated the influence of a continuous drain of ATP on the metabolism of *Escherichia coli* (Chao and Liao, 1994; Koebmann et al., 2002; Holm et al., 2010). However, more systematic studies, especially with varying levels of ATP demand under different growth conditions, are still needed to address fundamental questions, for example, to which extent the cells are able to compensate a rising ATP drain by increasing the glucose uptake rate. In particular, it is unknown what the maximal glucose uptake rate is and what happens when the ATP drain is further increased beyond this point.

In this study, we systematically analyzed the consequences of varying levels of ATP turnover in *E. coli* by overexpressing the genes of the ATP-hydrolyzing F₁-subunit of the F_OF₁-ATPase under different conditions (aerobic and anaerobic conditions, cell growth

and growth arrest). As one key result, we found that the glucose uptake rate shows under all conditions a biphasic response curve with respect to increasing ATPase activity, reaching a maximum at a medium ATPase level but dropping markedly when this level is exceeded. Focusing on anaerobic growth, we combined metabolome and proteome data with a kinetic model of *E. coli*'s central metabolism to reveal the underlying mechanism of this behavior. Analysis of the model showed that the dual dependency of the phosphofructokinase on ATP as substrate and ADP as activator can explain the biphasic steady-state response curve of the glycolytic flux. The model also helped to explain unexpected phenomena such as the accumulation of glycolytic metabolites, and it suggested two major rate-controlling steps under high ATP drain, which were confirmed experimentally by overexpressing the genes of the associated metabolic enzymes.

2 Results

2.1 Construction of the ATPase strains with different ATPase expression strengths

As in previous studies (Koeblmann et al., 2002; Boecker et al., 2019; Holm et al., 2010), as ATP-consuming mechanism, we chose the *atpAGD*-encoded F₁-subunit of the F₀F₁-ATP synthase (ATPase) from *E. coli*, which hydrolyzes ATP to ADP and phosphate. We decided to regulate the expression strength via different origins of replication and thus varying copy numbers of the plasmids harboring the ATPase genes. Three different plasmids were constructed: a low copy plasmid (RK2 replicon, LC), a medium copy plasmid (p15A replicon, MC), and a high copy plasmid (pMB1 replicon, HC) (Appendix Table S1). *E. coli* wild-type strain MG1655 was transformed with the three plasmids as well as with the corresponding empty control plasmids leading to the six strains "LC control", "LC ATPase", "MC control", "MC ATPase", "HC control", and "HC ATPase". Expression of *atpAGD* was put under control of the isopropyl β -D-thiogalactopyranoside (IPTG) inducible P_{trc}-promoter, and the same amount of IPTG was used for all strains. Additional controls were the *E. coli* wild-type strain MG1655 without plasmid and without IPTG addition ("WT") and without plasmid but with IPTG addition ("WT + IPTG").

2.2 Anaerobic growth

First, we cultivated all strains anaerobically and monitored the effect of expressing ATPase on growth rate, glucose uptake, and production of the five main fermentation products (ethanol, acetate, formate, lactate, and succinate) in the exponential phase. As expected, the growth rates and biomass yields of the three ATPase strains decreased with increasing expression levels of the ATPase (Fig 7.1A–D; Table 7.1). Furthermore, compared to their control strains, the specific glucose uptake rate of the LC and MC ATPase strain increased by 16.6% to 16.4 mmolGlc/gDW/h and by 17.8% to 17.7 mmolGlc/gDW/h, respectively. Similarly, the specific production rates of fermentation products were elevated, and their cumulated yield increased by 11.4% to 0.93 mol_{productC-atoms} / mol_{GlcC-atoms} in the LC ATPase and by 16.7% to 0.97 mol_{productC-atoms} / mol_{GlcC-atoms} in the MC ATPase strain (Table 7.1). This indicates that a larger portion of the substrate was redirected from biomass to energy production (and thus to the formation of fermentation products) in these two ATPase strains to keep up with the higher cellular ATP demand. However, the data from the HC ATPase strain with highest ATPase level suggest that there is a

maximal ATPase activity, beyond which *E. coli* cannot compensate the ATP drain with even higher glycolytic rates. The specific glucose uptake rate in the HC ATPase strain dropped by 60.3% to 5.1 mmol_{Glc}/gDW/h, which is far below the WT strain level. Apparently, there are limitations in *E. coli*'s metabolism that prevent a further increase in the glycolytic flux needed to counterbalance the high ATPase activity. Using a stoichiometric model of the central metabolism of *E. coli*, we performed metabolic flux analysis, based on the measured exchange rates, to estimate the ATPase flux in the different strains (see Materials and Methods). Importantly, while the ATPase flux increased with higher ATPase abundance from LC to MC ATPase strain, we determined the lowest ATPase flux among all ATPase strains for the HC ATPase strain (Table 7.1). In fact, the markedly reduced substrate uptake rate already implies that the ATP hydrolysis rate must be lower in this strain. However, if less ATP is consumed than in the other two strains, this raises the question why the glucose uptake rate drops so strongly. A hypothesis is that the HC ATPase strain has very low ATP levels, which constrains both the glycolytic and the ATPase flux. This will be further addressed in a later section.

2.3 Anaerobic cultivation under growth arrest

As the next step, we cultivated all strains anaerobically and arrested growth by transferring the cells to a medium without a nitrogen source. Such conditions are of particular interest for biotechnological applications (e.g., in two-stage production processes (Burg et al., 2016; Klamt et al., 2018)). All ATPase strains showed high glycolytic and product exchange rates, while all control strains came to a metabolic halt and barely took up any glucose after 10 h of cultivation (Fig 7.2A–D; Table 7.1). Compared to their corresponding control strains, the LC ATPase strain had a 300% increased glucose uptake rate (8.96 mmol_{Glc}/gDW/h), the MC ATPase strain a 380% increase (10.46 mmol_{Glc}/gDW/h), and the HC ATPase still a 175% increase (8.10 mmol_{Glc}/gDW/h). To our knowledge, the specific glucose uptake rate of the MC ATPase strain is the highest rate ever reported for growth-arrested *E. coli* cells. The HC ATPase strain showed again the lowest glucose uptake rate among the three ATPase strains. While the WT and control strains metabolized most of the glucosecarbon to ethanol, acetate, formate, and succinate, the amount of formed lactate increased with the expression strength of ATPase. The HC ATPase strain converted all glucose almost entirely to lactate (1.74 mol_{Lac}/mol_{Glc}) (Fig 7.2C and D; Table 7.1). As for the case of anaerobic growth, the overall ATPase flux in the HC ATPase strain was lower than in the other ATPase strains.

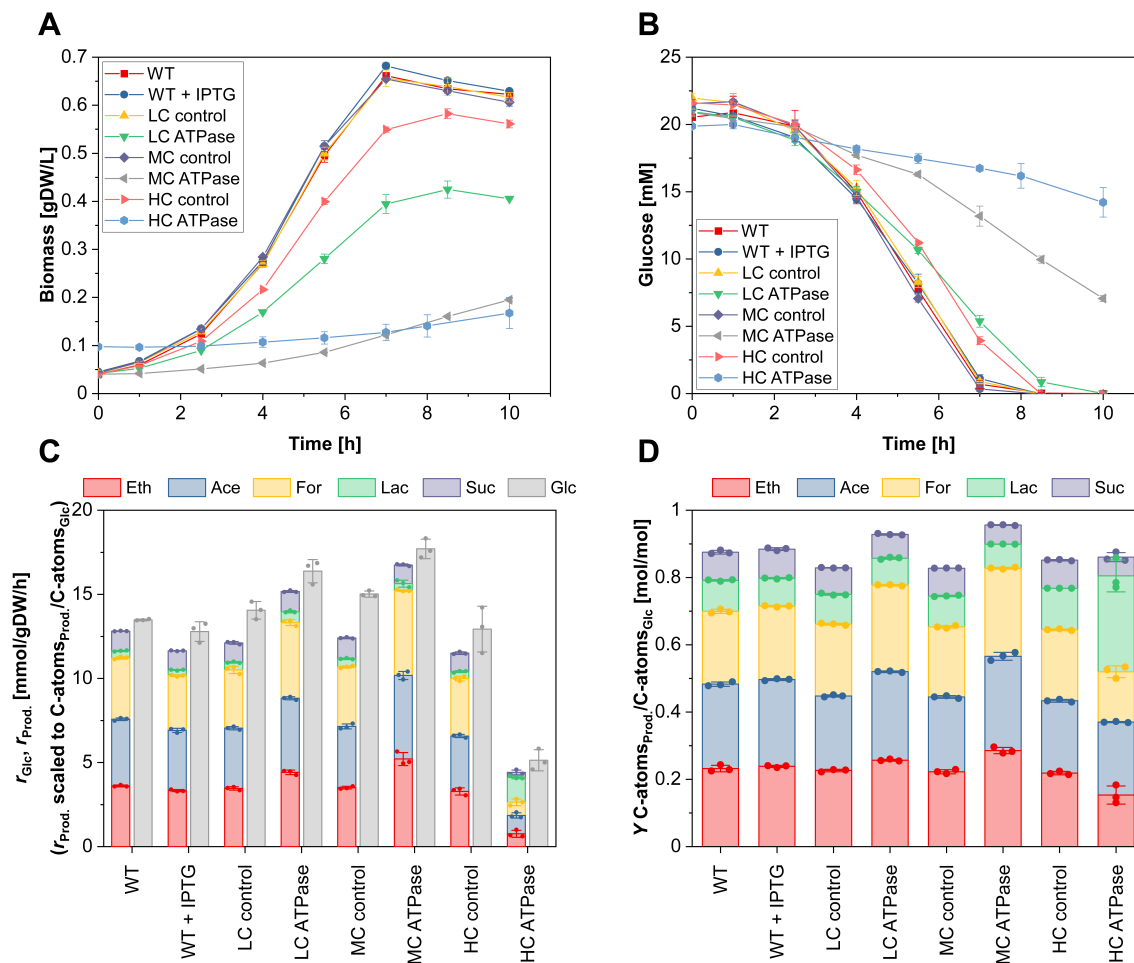


Figure 7.1: Anaerobic growth of the different strains. (A) Time course of biomass concentration. (B) Time course of glucose concentration. (C) Specific glucose (Glc) uptake rate and specific productivity for ethanol (Eth), acetate (Ace), formate (For), lactate (Lac), and succinate (Suc). (D) Yield of Eth, Ace, For, Lac, and Suc. Data information: The reaction rates in (C) were calculated for the exponential phase under assumption of quasi-steady state. As changes in glucose and fermentation products are rather small during cultivation of the HC ATPase strain, a higher initial biomass concentration of 0.1 gDW/l was used for this strain to get more robust data for the calculation of the metabolite exchange rates. The means (A and B) and the means and individual data (C and D) for $n = 3$ biologically independent samples are shown. The error bars represent \pm SD. Source data are available online for this figure.

Table 7.1: Specific growth, glucose uptake, and product synthesis rates and product yields of the tested strains under anaerobic growth with growth (top) and with growth arrest (bottom). Regarding the calculation of the ATPM and the ATPase rates see Materials and Methods. The means \pm SD of $n = 3$ biologically independent samples are shown.

	WT	WT + IPTG	LC control	LC ATPase	MC control	MC ATPase	HC control	HC ATPase
With growth								
μ [h ⁻¹]	0.476 \pm 0.004	0.449 \pm 0.007	0.458 \pm 0.004	0.376 \pm 0.007	0.461 \pm 0.002	0.197 \pm 0.003	0.431 \pm 0.004	0.063 \pm 0.009
r_{GlC} [mmol/gDW/h]	13.48 \pm 0.04	12.78 \pm 0.48	14.05 \pm 0.42	16.38 \pm 0.56	15.02 \pm 0.15	17.7 \pm 0.47	12.93 \pm 1.11	5.13 \pm 0.51
r_{Eth} [mmol/gDW/h]	10.81 \pm 0.14	9.99 \pm 0.15	10.38 \pm 0.24	13.23 \pm 0.33	10.51 \pm 0.19	15.63 \pm 0.94	9.84 \pm 0.52	2.29 \pm 0.51
r_{Ace} [mmol/gDW/h]	11.9 \pm 0.19	10.73 \pm 0.3	10.74 \pm 0.23	13.23 \pm 0.22	10.93 \pm 0.37	14.9 \pm 0.59	9.89 \pm 0.24	3.28 \pm 0.39
r_{For} [mmol/gDW/h]	21.88 \pm 0.37	19.54 \pm 0.26	20.78 \pm 1.07	26.99 \pm 0.86	21.21 \pm 0.3	30.24 \pm 0.25	20.46 \pm 0.66	4.69 \pm 0.96
r_{Lac} [mmol/gDW/h]	0.83 \pm 0.05	0.67 \pm 0.05	0.89 \pm 0.07	1.3 \pm 0.11	0.98 \pm 0.07	0.84 \pm 0.34	0.83 \pm 0.09	2.93 \pm 0.13
r_{Suc} [mmol/gDW/h]	1.77 \pm 0.02	1.72 \pm 0.02	1.73 \pm 0.1	1.3 \pm 0.07	1.85 \pm 0.05	1.66 \pm 0.08	1.65 \pm 0.09	0.43 \pm 0.17
r_{ATPM} [mmol/gDW/h]	4.41 \pm 0.54	3.27 \pm 0.31	3.3 \pm 0.43	14.63 \pm 1.04	3.61 \pm 0.82	29.06 \pm 0.67	2.67 \pm 1.05	6.99 \pm 0.73
r_{ATPase} [mmol/gDW/h]	0.155 \pm 0.005	0.16 \pm 0.001	0.149 \pm 0.002	0.104 \pm 0.003	0.146 \pm 0.001	0.05 \pm 0	0.136 \pm 0.002	0.065 \pm 0.015
Y_{Eth} [mol/mol]	0.696 \pm 0.024	0.716 \pm 0.008	0.678 \pm 0.009	0.77 \pm 0.007	0.668 \pm 0.015	0.857 \pm 0.022	0.655 \pm 0.013	0.46 \pm 0.066
Y_{Ace} [mol/mol]	0.753 \pm 0.016	0.774 \pm 0.007	0.665 \pm 0.008	0.79 \pm 0.006	0.666 \pm 0.01	0.84 \pm 0.029	0.645 \pm 0.01	0.65 \pm 0.004
Y_{For} [mol/mol]	1.29 \pm 0.031	1.309 \pm 0.01	1.28 \pm 0.014	1.542 \pm 0.01	1.251 \pm 0.019	1.572 \pm 0.013	1.268 \pm 0.013	0.899 \pm 0.087
Y_{Lac} [mol/mol]	0.184 \pm 0.004	0.166 \pm 0.003	0.177 \pm 0.005	0.161 \pm 0.005	0.184 \pm 0.003	0.143 \pm 0.001	0.246 \pm 0	0.57 \pm 0.078
Y_{Suc} [mol/mol]	0.125 \pm 0.007	0.13 \pm 0.005	0.118 \pm 0.001	0.106 \pm 0.003	0.124 \pm 0.001	0.085 \pm 0.001	0.125 \pm 0.002	0.083 \pm 0.017
Growth arrest								
μ [h ⁻¹]	~0	~0	~0	~0	~0	~0	~0	~0
r_{GlC} [mmol/gDW/h]	1.9 \pm 0.11	2.02 \pm 0.06	2.34 \pm 0.34	8.96 \pm 0.36	2.18 \pm 0.06	10.46 \pm 0.43	2.95 \pm 0.09	8.1 \pm 0.6
r_{Eth} [mmol/gDW/h]	1.43 \pm 0.21	1.43 \pm 0.04	1.47 \pm 0.04	4.78 \pm 0.35	1.43 \pm 0.09	3.16 \pm 0.4	1.34 \pm 0.06	0.66 \pm 0.14
r_{Ace} [mmol/gDW/h]	1.53 \pm 0.07	1.5 \pm 0.02	1.58 \pm 0.04	4.69 \pm 0.28	1.54 \pm 0.03	2.56 \pm 0.54	1.43 \pm 0.06	0.6 \pm 0.03
r_{For} [mmol/gDW/h]	2.23 \pm 0.18	2.09 \pm 0.03	2.1 \pm 0.1	7.65 \pm 0.47	2.04 \pm 0.07	4.28 \pm 0.82	2.17 \pm 0.12	0.42 \pm 0.03
r_{Lac} [mmol/gDW/h]	0.12 \pm 0.02	0.22 \pm 0.04	0.32 \pm 0.12	5.71 \pm 0.12	0.49 \pm 0.19	11.31 \pm 0.54	1.51 \pm 0.1	13.78 \pm 0.36
r_{Suc} [mmol/gDW/h]	0.75 \pm 0.03	0.71 \pm 0.02	0.95 \pm 0.02	2.2 \pm 0.08	0.89 \pm 0.06	1.95 \pm 0.01	0.94 \pm 0.04	0.35 \pm 0.04
r_{ATPM} [mmol/gDW/h]	4.16 \pm 0.27	4.29 \pm 0.08	4.59 \pm 0.23	18.62 \pm 0.82	4.63 \pm 0.12	19.31 \pm 0.78	5.42 \pm 0.14	15.63 \pm 0.6
r_{ATPase} [mmol/gDW/h]	~0	~0	~0	~0	~0	~0	~0	~0
Y_{Eth} [mol/mol]	0.738 \pm 0.009	0.703 \pm 0.035	0.658 \pm 0.011	0.497 \pm 0.011	0.683 \pm 0.01	0.284 \pm 0.015	0.512 \pm 0.032	0.076 \pm 0.003
Y_{Ace} [mol/mol]	0.84 \pm 0.021	0.758 \pm 0.004	0.704 \pm 0.015	0.667 \pm 0.014	0.709 \pm 0.028	0.259 \pm 0.011	0.406 \pm 0.024	0.092 \pm 0.003
Y_{For} [mol/mol]	1.14 \pm 0.039	0.993 \pm 0.056	0.915 \pm 0.055	0.45 \pm 0.018	0.901 \pm 0.058	0.305 \pm 0.018	0.52 \pm 0.036	0.059 \pm 0.001
Y_{Lac} [mol/mol]	0.063 \pm 0.006	0.084 \pm 0.003	0.128 \pm 0.023	0.728 \pm 0.013	0.221 \pm 0.036	1.177 \pm 0.052	0.734 \pm 0.023	1.74 \pm 0.036
Y_{Suc} [mol/mol]	0.383 \pm 0.026	0.341 \pm 0.007	0.413 \pm 0.003	0.227 \pm 0.002	0.407 \pm 0.006	0.155 \pm 0.005	0.289 \pm 0.008	0.43 \pm 0.002

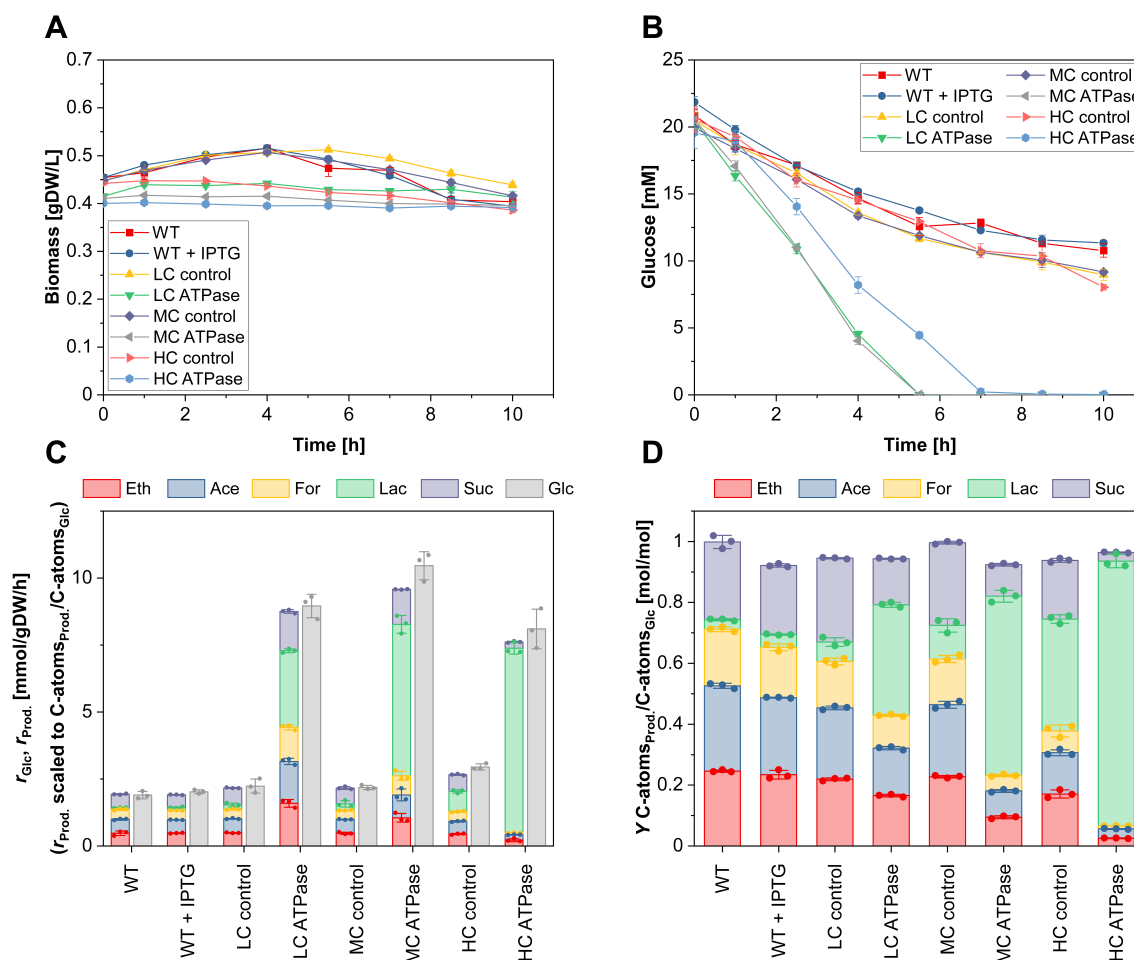


Figure 7.2: Anaerobic cultivation of the different strains under growth arrest. (A) Time course of biomass concentration. (B) Time course of glucose concentration. (C) Specific glucose (Glc) uptake rate and specific productivity for ethanol (Eth), acetate (Ace), formate (For), lactate (Lac), and succinate (Suc). (D) Yield of Eth, Ace, For, Lac, and Suc. Data information: The reaction rates in (C) were calculated from the beginning of cultivation until the last sampling time point where glucose was still present in the medium. The means (A and B) and the means and individual data (C and D) for $n = 3$ biologically independent samples are shown. The error bars represent \pm SD. Note: although no nitrogen source was present in the medium, some minor growth (especially of the control and wild type strains) remained (A), which is a known phenomenon within the first hours of cultivation after nitrogen depletion (Switzer et al., 2020). Source data are available online for this figure.

2.4 Aerobic cultivations

We repeated the same experiments under aerobic conditions (Figs EV 7.8A–D and EV7.9A–D, Appendix Tables S2 and S3). Essentially, the same trends as in the anaerobic cultivations could be observed. While all ATPase strains showed reduced growth rates and biomass yields compared to the control strains, the specific glucose uptake rates increased in the LC and MC ATPase strain by 6.5 and 49.2%, respectively, but dropped in the HC ATPase strain by 39.1% (Fig EV7.8C, Appendix Table S2). The yield of acetate, which is the major overflow byproduct of *E. coli* under aerobic conditions (Wolfe, 2005), increased with the expression strength of the ATPase genes. We also tested the behavior of the ATPase strains under aerobic conditions with growth arrest (Fig EV7.9A–D, Appendix Table S3). While the metabolism of the WT and control strains slowed down after some time (as under anaerobic conditions), all ATPase strains exhibited a very high metabolic activity. Again, a biphasic steady-state response curve of the glucose uptake rate for increasing ATPase levels could be observed. The MC ATPase strain reached the highest glucose uptake rate (10.2 mmol_{Glucose}/gDW/h), which was more than 10-fold higher than in the control strain (and even higher than in the WT with growth). Compared to the anaerobic case, we estimated considerably higher ATPase fluxes for the ATPase strains, which are possible due to the high ATP yield under aerobic conditions.

2.5 Increasing ATP demand induces biphasic response of glucose uptake under all cultivation conditions

The data of the anaerobic and aerobic cultivations with and without growth consistently show a biphasic curve of the steady-state glucose uptake rate as response to the increasing overexpression of the ATPase genes. As summarized in Fig 7.3, for all cultivations, the uptake rate increases from WT over LC to MC ATPase strain and then drops for HC ATPase strain, especially sharply for growing cells. This observation raises the key questions: what causes this biphasic steady-state response, and why does *E. coli* not further increase (or at least maintain a high) glycolytic flux under maximal ATPase expression? For an in-depth analysis of this phenomenon, we focused on the response of the different ATPase strains under anaerobic growth, where the highest specific glucose uptake rates for the LC and MC ATPase strain and the steepest drop of the uptake rate of the HC ATPase strain could be observed.

2.6 Changes at the metabolome and proteome level for increasing ATPase abundance

For a comprehensive metabolic characterization of the different ATPase strains under anaerobic growth, we quantified the intracellular concentration of metabolites (Fig EV7.10A and B; Dataset EV7.8) and determined changes on proteome level (Fig EV7.11; Dataset EV1). ATP, ADP, and AMP concentrations remained relatively constant in the LC and MC ATPase strains compared to the control strains. Only the HC ATPase showed larger changes: the ATP level decreased by 56.0%, while the ADP level increased by 24.6% and the AMP level by almost 1,000% (Fig EV7.10A). These observations are also reflected by the adenylate energy charges, which are high in the WT and LC ATPase strain (0.87 and 0.88, respectively), only slightly lower in the MC ATPase strain (0.80), but significantly reduced in the HC ATPase strain (0.34) (Fig EV7.10A). This indicates that the LC and the MC, but not the HC ATPase strain, can compensate the higher ATP demand by the increased glycolytic flux. The concentrations of several other intracellular metabolites changed also significantly. Generally, we observed that the concentration of metabolites involved in glycolysis (glucose-6-phosphate, fructose-6-phosphate (F6P), fructose-1,6-bisphosphate (FBP), dihydroxyacetone phosphate (DHAP), glyceraldehyde-3-phosphate (G3P), PEP, and pyruvate) consistently increased with expression strength of the ATPase genes, reaching highest values in the HC ATPase strain (Fig EV7.10B). Contrarily, the acetyl-CoA concentration dropped in the HC ATPase strain (Fig EV7.10B). In contrast to metabolite concentrations, at the proteome level, enzymes from the anaerobic core metabolism of *E. coli* were not as clearly up- or downregulated (Fig EV7.11; Dataset EV7.8). In particular, larger changes of glycolytic enzyme levels could not be seen; the abundance levels of these enzymes in the ATPase strains are all within the range of 60% and 150% of the respective levels in the control strains. This indicates that changes in the glycolytic fluxes are mainly induced by allosteric or substrate level regulation rather than by alteration of enzyme levels. As expected, the three subunits of the F1-ATPase were more abundant with increasing copy number of the expression plasmids. The average abundances of the ATPase α -, β -, and γ -subunits were +142% (LC ATPase), +549% (MC ATPase), and +708% (HC ATPase) in comparison to the WT strain (Fig EV7.11; Dataset EV7.8). Among the remaining enzymes from the core carbon metabolism, only the malate dehydrogenase (upregulated), PEP carboxykinase (PCK, upregulated), and the PEP carboxylase (PPC, downregulated) showed stronger changes in the MC and HC ATPase strains compared to their control strains. PCK and PPC are adjacent in the sense

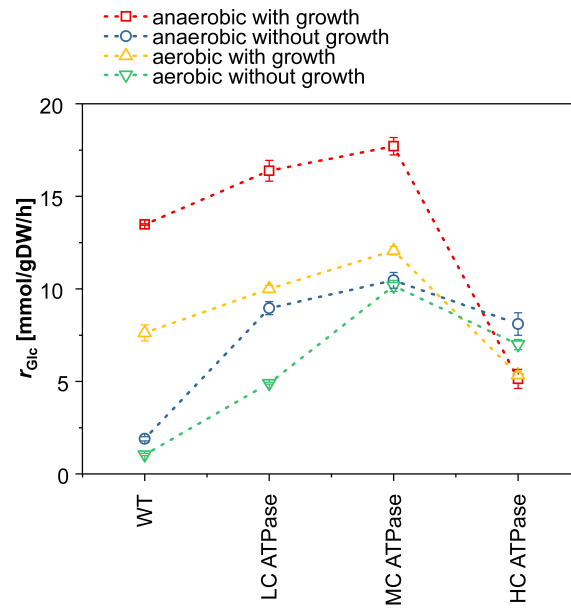


Figure 7.3: For all growth regimes tested, the glucose uptake rate shows a biphasic response curve to increasing ATPase levels. The means for $n = 3$ biologically independent samples are shown. The error bars represent \pm SD.

that PPC catalyzes the carboxylating reaction from PEP to oxaloacetate (with release of a phosphate molecule) and PCK the reaction from oxaloacetate to PEP thereby consuming ATP (see also Fig 5C). Under certain conditions, for example, low ATP or high PEP concentrations, the PCK reaction may also act in the reverse direction, and it seems that the PCK is used to replace the PPC to provide additional ATP. However, despite the fact that also the malate dehydrogenase is upregulated, the overall flux to succinate as final product of this pathway is still relatively low in the HC ATPase strain (Fig 7.1C and D; Table 7.1). Not as prominent but still noticeable was the downregulation of formate acetyltransferase (formerly pyruvate formate lyase; PFL) and of the formate transporter FocA in the HC ATPase strain. While the PFL and the FocA levels were not significantly affected in the LC and MC ATPase strains, in the HC ATPase strain, compared to its control, the levels dropped by 33% and 58%, respectively.

2.7 What limits the glycolytic flux under high ATPase activity - a kinetic modeling approach

With the experimental findings and data at hand, we sought to find a mechanistic explanation for the inability of the HC ATPase strain to sustain a high glycolytic flux, as in the LC and MC ATPase strain, to compensate the high ATP demand. Given the low energy

charge in the HC ATPase strain and the increased ratio between hexose phosphates and FBP, we hypothesized that the kinetics of the phosphofructokinase (PFK), reaction converting F6P to FBP under consumption of (Mg)ATP, might cause the biphasic behavior. This reaction is considered as the committing step of glycolysis and is a major point of regulation in *E. coli* (Fenton and Reinhart, 2009). In *E. coli*, there are two PFKs (PFK1 (encoded in *pfkA*) and PFK2 (*pfkB*)); however, more than 90% of the PFK activity in *E. coli* can be attributed to PFK1 (Kotlarz et al., 1975). This enzyme is allosterically inhibited by phosphoenolpyruvate (PEP) and has a dual dependency on (Mg)ADP and (Mg)ATP (Blangy et al., 1968). On the one hand, high concentrations of the substrate (Mg)ATP are required to obtain a high PFK flux. On the other hand, ADP (as well as GDP) is known to allosterically activate the reaction (Peskov et al., 2008). Indeed, using a kinetic rate law based on convenience kinetics (Liebermeister and Klipp, 2006) and including the effect of the allosteric regulators leads to a biphasic curve when increasing the ADP/(ATP + ADP) ratio (Fig 7.4; cf. also Ref. Peskov et al., 2008). For high ATP levels (low ADP/(ADP + ATP) ratio), ADP as activator is limiting. With increasing ADP levels induced by higher ATPase activity, the PFK rate increases as well, but only to an optimal point beyond which the concentration of the co-substrate ATP becomes limiting. As second step of the glycolysis, the PFK reaction kinetics could thus be the cause for the observed biphasic response of the glucose uptake rate and the low glucose consumption rate in the HC ATPase strain could be a consequence of the low ATP concentration (high ADP/(ATP + ADP) ratio) in this strain.

However, it is not clear whether the biphasic response of the isolated PFK reaction in Fig 7.4 (with fixed concentrations of PEP, FBP, F6P) translates into a biphasic response of the glycolytic flux when the entire central metabolism with all its interactions and feedbacks is taken into account. We therefore constructed a kinetic model of the central fermentative metabolism of *E. coli* for anaerobic growth on glucose (see detailed description in Appendix Supplementary Text; section 1). The model comprises 33 metabolites and 28 reactions and, as shown in Fig 7.5C, accounts for the glycolysis, anaplerotic reactions, relevant parts of the TCA cycle, the major fermentative pathways, a growth reaction, a reaction for nongrowth-associated ATP maintenance (NGAM) demand, and, finally, a reaction for simulating ATP hydrolysis by the ATPase. The latter reaction depends on the ATPase overexpression level. The model also accounts for allosteric regulation of the involved enzymes. As described in detail in the Appendix Supplementary Text, for the first version of the model, we fitted the unknown parameters of the model to the measured growth rate, the substrate uptake rate, and the product exchange fluxes of the wild

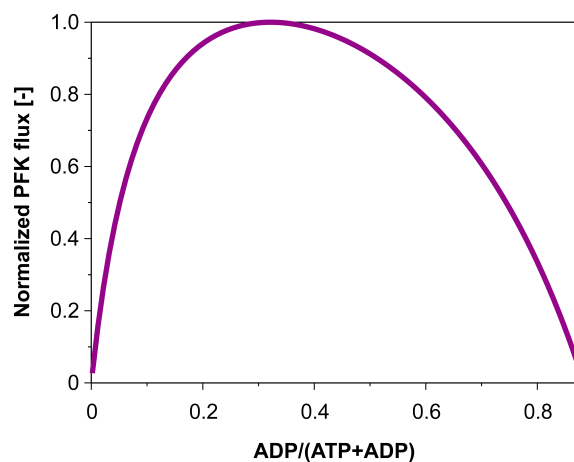


Figure 7.4: Simulation of the PFK rate with varying $ADP/(ATP+ADP)$ ratios. The PFK flux was simulated as single (isolated) reaction with the following fixed metabolite concentrations relevant for the kinetic rate law of the PFK (for the latter see Appendix Supplementary Text, section 1.3): PEP: $0.27 \mu\text{mol/gDW}$; F6P: $0.91 \mu\text{mol/gDW}$; FBP: $9.74 \mu\text{mol/gDW}$; total concentration ATP+ADP: $2.67 \mu\text{mol/gDW}$.

type and the different ATPase strains (given in Table 1) as well as to measured metabolite levels in these strains (see above and Dataset EV7.8). Since the proteomic data indicated only minor or moderate changes in the levels of relevant metabolic enzymes, we used, as an approximation, constant v_{max} values (i.e., constant enzymes concentrations) in all four strains (WT as well as LC, MC, and HC ATPase strains) in the kinetic model simulations. We were able to find a parametrization that gave a reasonable fit with the experimental data (Fig 7.5A and B). In particular, the fitted model could reproduce the observed biphasic response of the glucose uptake along the four strains (first plot in Fig 7.5A) demonstrating that the known metabolic and regulatory interactions contained in the kinetic model are sufficient to generate this behavior.

Next, we used the model to simulate the steady-state response curves of glucose uptake rate, energy charge, and ATPase flux when increasing the ATPase level (represented by v_{max} of the ATPase reaction) continuously from 0 (WT) to a maximal value of 85 mmol/gDW/h (Fig 7.6A). Starting with a low ATPase level, the ATPase flux increases causing a higher ADP concentration (lower energy charge) which, as discussed in Fig 7.4, enhances the PFK and thereby the glycolytic flux. However, further increasing the abundance of ATPase beyond a critical point reduces the ATP concentration, now limiting the PFK flux and thus the glucose uptake rate (Fig 7.6A). These model simulations also confirmed that the rate of ATP hydrolysis by the ATPase is lowest in the HC ATPase strain, despite the fact that it has the highest ATPase abundance (i.e., the highest v_{max} ;

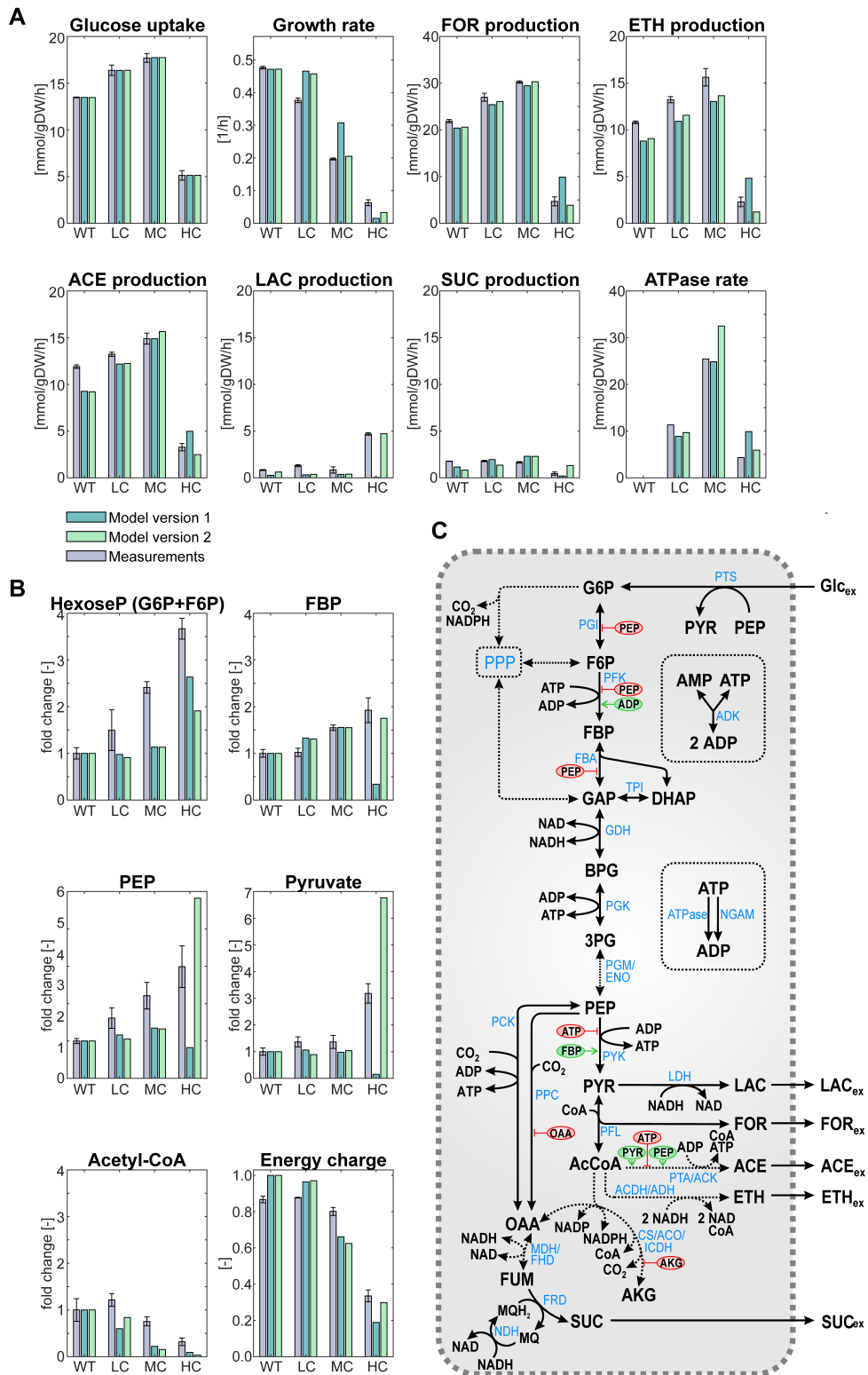


Figure 7.5: The figure caption is on the next page.

Figure 7.5: Comparison of the simulations of the kinetic model version 1 and version 2 with experimental data of the different strains under anaerobic growth conditions. (A) Comparison of the simulations of the kinetic model version 1 and version 2 with the measured exchange fluxes. (B) Comparison of the simulations of the kinetic model version 1 and version 2 with the measured metabolite concentrations. (C) Metabolic map containing most relevant reactions (blue) and metabolites (black) of *E. coli* under anaerobic conditions. Dashed arrows indicate (lumped) reactions with several enzymes involved. Allosteric regulations of the respective reactions are marked in red (inhibition) or green (activation). The kinetic model (versions 1 and 2) covers almost all of the shown reactions and regulations; a map directly related to the model is shown in Appendix Fig S1. Abbreviations of metabolites and reaction names: Glcex: external glucose (substrate); G6P: D-glucose-6-phosphate; F6P: D-fructose-6-phosphate; FBP: fructose-1,6-bisphosphate; GAP: D-glyceraldehyde-3-phosphate; DHAP: dihydroxyacetone phosphate; BPG: 1,3-bisphospho-D-glycerate; 3PG: 3-phosphoglycerate; PEP: phosphoenol-pyruvate; PYR: pyruvate; AcCoA: acetyl coenzyme A; CoA: coenzyme A; AKG: α -ketoglutarate; OAA: oxaloacetate; FUM: fumarate; SUC: succinate; FOR: formate; LAC: lactate; ACE: acetate; ETH: ethanol; ATP: adenosine triphosphate; ADP: adenosine diphosphate; NAD: oxidized nicotinamide adenine dinucleotide; NADP: oxidized nicotinamide adenine dinucleotide phosphate; NADH: reduced nicotinamide adenine dinucleotide; NADPH: reduced nicotinamide adenine dinucleotide phosphate; CO₂: carbon dioxide; MQH₂: menaquinol; MQ: menaquinone. PTS: phosphotransferase system; PGI: glucose-6-phosphate isomerase; PFK: phosphofruktokinase; FBA: fructose-bisphosphate aldolase; TPI: triose-phosphate isomerase; GHD: glyceraldehyde-3-phosphate dehydrogenase; PGK: phosphoglycerate kinase; PGM: phosphoglycerate mutase; ENO: enolase; PYK: pyruvate kinase; PFL: pyruvate formate-lyase (also known as formate acetyltransferase); LDH: lactate dehydrogenase; PTA: phosphate acetyltransferase; ACK: acetate kinase; ACDH: acetaldehyde-CoA dehydrogenase; ADH: alcohol dehydrogenase; PCK: phosphoenolpyruvate carboxykinase; PPC: phosphoenolpyruvate carboxylase; CS: citrate synthase; ACO: aconitate hydratase A/B; ICDH: isocitrate dehydrogenase; MDH: malate dehydrogenase; FHD: fumarase; FRD: fumarate reductase; NDH: NADH dehydrogenase; ADK: adenylate kinase; NGAM: ATP consumption for non-growth-associated maintenance; ATPase: ATP hydrolysis by F1-ATPase in the ATPase strains. Data information: For the measurements, the means for $n = 3$ biologically independent samples are shown (A and B) and the error bars represent \pm SD.

ATPase). Available ATP is rapidly consumed by the large amounts of ATPase in the HC ATPase strain, which keeps the concentration of ATP at a very low level limiting in turn both the PFK as well as the ATPase flux. To further test our hypothesis regarding the PFK mechanism and its effect on the glucose uptake under increasing ATP demand, we utilized the model to evaluate how the steady-state response curves in Fig 7.6A change if we remove the allosteric regulation of PFK by ADP (the respective ADP-dependent term in the kinetic rate law was fixed (clamped) to its wild-type value in all simulations). Indeed, the biphasic response of the glucose uptake disappears and a monotonic decrease in glucose uptake, upon enhanced ATP demand, can be seen (Fig 7.6B). Taken together, our modeling results strongly support our hypothesis that the dual dependency of the PFK on ADP and ATP causes the observed biphasic response in the glycolytic flux. As a consequence, the PFK mechanism enables the cell to buffer moderate ATPase activities (and to compensate moderate increases in ATP demand) by an instantaneous adjustment of the glycolytic flux, but it collapses under very high ATP demand as in the HC ATPase strain.

While the kinetic model (version 1) reproduced most experimental data in the different strains reasonably well, especially the exchange rates (Fig 7.5A), some deviations can be noted in the metabolite concentrations. The measured accumulation of hexose phosphates (glucose-6-phosphate, F6P) is correctly reflected by the model, again indicating that the PFK activity is the limiting factor in the HC ATPase strain. However, the measured high concentrations of the glycolytic metabolites FBP, PEP, and pyruvate in the HC ATPase strain are not captured in the simulation results. For example, the measured concentrations of FBP and pyruvate increased markedly in the HC ATPase strain, but the opposite behavior (strong decrease) was displayed by the model. Even after several rounds of parameter fitting, we could not find a parametrization that leads to a better reproduction of the qualitative trends of the metabolomics data. These discrepancies between data and model simulations suggested that there are missing regulatory elements in our kinetic model, and we, therefore, introduced two major changes (model version 2; Appendix Supplementary Text). As a first change, because of (a) the low concentration of acetyl-CoA, (b) the accumulation of pyruvate, (c) the increased lactate yield, and (d) the observed reduced PFL levels in the HC ATPase strain, we introduced a term in the kinetics of the PFL reaction, which ensures a lower abundance of PFL (and thus a decreased v_{max}) under low ATP concentrations. As a possible mechanistic explanation, we hypothesized that, at low energy charges, the large and costly PFL (consisting of 759 amino acids) is replaced by cheaper pathways (e.g., via lactate fermentation), al-

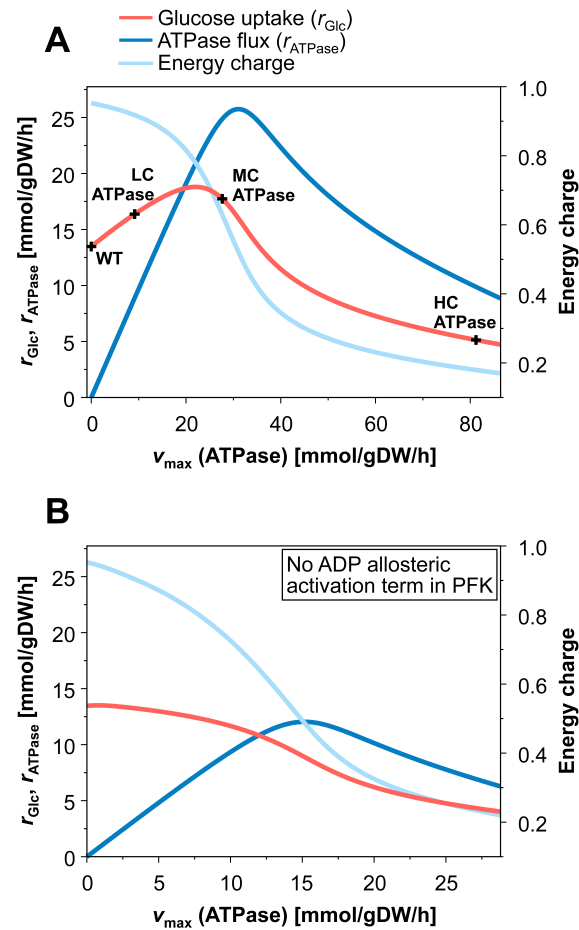


Figure 7.6: Simulation of the steady-state response curves of glucose uptake rate, energy charge, and ATPase flux under anaerobic growth for increasing ATPase activities. (A) Simulations with the standard kinetic model (version 1). The measured glucose uptake rates of the different strains are indicated (cf. with Fig 3). (B) Simulations with the kinetic model (version 1) as in (A), but without ADP activation term in the PFK kinetics.

though this may partially reduce the ATP yield (see also Discussion). However, even with introduction of a downregulation of the PFL under low ADP concentrations, the model was still not able to reflect the high concentration levels of PEP and FBP in the HC ATPase strain. In fact, with (i) the observed high concentrations of PEP and ADP (the substrates for the pyruvate kinase (PYK) reaction), (ii) the high FBP level (allosteric activator of pyruvate kinase), and (iii) the highly negative standard Gibbs free energy change of this reaction ($\Delta_r G^{\circ} = -21.78$ kJ/mol) (Park et al., 2016), one would expect a higher PYK flux and thus a decrease of the PEP level, which contradicts the measured high PEP concentration in the HC ATPase strain (Fig 7.5B). As a second change in the model, we therefore introduced a term inhibiting PYK (flux) under high pyruvate concentrations. Although such an allosteric inhibition of PYK by pyruvate is not known, it has been reported that alanine, which is directly produced from pyruvate, may act as inhibitor for PYK (Taber et al., 1998), and we observed higher concentrations of alanine in all ATPase strains (Dataset EV7.8). With these two changes, the resulting model version 2 was now able to reproduce the qualitative trends in the fluxes and metabolite concentrations in all three ATPase strains (Fig 7.5A and B). With this, we can summarize our understanding for the observed phenomena as follows: increasing the ATPase activity reduces the growth rate in all strains and elevates the glycolytic flux in the LC and MC ATPase strains due to (moderately) increased ADP levels, which enhances the activity of PFK. However, in the HC ATPase strain, there is a sharp decrease of ATP, a substrate of the PFK, which now limits the PFK and thereby the glycolytic flux. As consequence of low ATP levels, the PFL abundance is reduced, which lowers the flux from pyruvate to acetyl-CoA. Thus, pyruvate accumulates, which leads to higher lactate production rates and inhibits (probably indirectly) PYK activity. Consequently, the levels of PEP increase further, which further slows down the PFK flux due to negative inhibition, and thus, the concentrations of hexose phosphates increase. Finally, high PEP concentrations propagate also upward to other glycolytic intermediates (DHAP, G3P, FBP). Since PEP also inhibits fructose-bisphosphate aldolase, this further contributes to FBP accumulation. We note that the steady-state response curves shown for model version 1 in Fig 7.6 are not affected by the model changes and look very similar for model version 2.

2.8 Using Monte Carlo sampling of kinetic parameters to assess the robustness of model predictions

The kinetic model (version 2) of *E. coli*'s central metabolism constructed in the previous section is relatively large and comprises more than 100 unknown parameters. To assess the robustness of model predictions with respect to kinetic parameters, we employed Monte Carlo analysis as previously described in (Murabito et al., 2014). Briefly, we used this method to sample Michaelis–Menten parameters over two orders of magnitude while preserving the metabolic steady state to which the original parameters were fitted. For each sampled set of kinetic parameters, systems properties can be computed. Here, we focused on the flux control coefficients (FCCs), which are known from metabolic control analysis and quantify the relative change of a steady-state metabolic flux when changing the enzyme level, that is, when changing the v_{\max} , of other reactions (Sauro, 2013). The resulting distributions of FCCs obtained from the Monte Carlo analysis allow us to assess the control properties with different (but consistent) parametrizations and consequently to analyze the uncertainty of these global characteristics in our deterministic model. For details of the method, we refer to (Murabito et al., 2014) and the Appendix Supplementary Text (section 4). For the wild-type metabolic steady state, we found predominantly narrow distributions of FCCs, indicating a low sensitivity of these FCCs against parameter variations (Appendix Fig S3). FCCs with broader distributions were typically sign-dominant. That is, while the numerical value of the FCCs varies as a function of kinetic parameters, the sign of the FCCs remains either positive or negative, respectively, indicating robust qualitative control properties and hence robust predictions based on the FCCs. Importantly, the FCC distribution of the NGAM reaction on PTS was narrow with predominantly positive sign (>95% of sampled instances), confirming that the increase of glucose uptake in the wild type as response to higher ATP demand is a robust feature of the model. Using instead the HC ATPase strain (with its high ATP demand) as reference metabolic (steady) state, most FCCs still show narrow distributions, but the fraction of enzymes with a broader distribution of FCCs increased (Appendix Fig S4). The moderately increased sensitivity to parameter variations can be interpreted as reduced robustness of the HC ATPase metabolic state. In particular, the enzymes PYK, PFL, and PFK exhibited broad but sign-dominant (positive) distributions of their respective FCCs on PTS and glycolytic flux, indicating that the sign (positive control) of these enzymes on the glycolytic flux in this metabolic state is largely independent of the choice of kinetic parameters. Furthermore, the FCCs of the NGAM and

ATPase reactions on PTS and glycolytic flux are here predominantly negative. Together with the predominantly positive FCCs of NGAM on PTS and glycolytic flux in the wild type, this confirms the bi-phasic response of the glycolytic flux to increased ATP demand as a robust feature of the model. The full results of the Monte Carlo analysis are shown and discussed in Appendix Supplementary Text.

2.9 Using model predictions to obtain higher glycolytic fluxes in the HC ATPase strain

Motivated by the Monte Carlo analysis from the previous section, in a final step, we aimed to use the kinetic model (version 2) to make experimentally testable predictions and, in this way, to further verify our reasoning of the low glycolytic flux in the HC ATPase strain. We hypothesized that we could enhance the glycolytic flux by overexpressing genes of enzymes catalyzing reactions with the highest metabolic control. We computed the FCCs of all glycolytic reactions in the kinetic model. Since FCCs are valid only for small changes in the enzyme level, we also computed the resulting steady-state glycolytic flux in the model when doubling the enzyme level, corresponding to doubled v_{\max} values of the respective reactions. The results can be found in the Appendix Supplementary Text (section 2.3 and Fig S2). Consistent with our previous reasoning on the experimental findings and consistent with the results from the Monte Carlo analysis, we found that the PFL and the PFK have the highest control on the glycolytic flux in the HC ATPase strain. We therefore overexpressed the genes of these two enzymes to test whether they indeed represent bottlenecks. We constructed four variants of the HC ATPase strain: one overexpressing the PFK-encoding gene *pfkA*, one overexpressing the PFL-encoding gene *pflB*, one overexpressing both genes, and one overexpressing the phosphoglycerate kinase (PGK) encoding gene *pgk* (in all cases additionally to the *atpAGD* operon on the high copy plasmid). Overexpression of *pgk* was chosen as a control since the model predicted an FCC close to zero for the PGK reaction. Thus, in contrast to the other three strains, PGK overexpression should not have a major influence on the glycolytic flux of the HC ATPase strain. The four variants were grown under anaerobic conditions and compared to the HC ATPase strain.

As shown in Fig 7.7A–D and Table 7.2, the overexpression of *pfkA* and *pflB* indeed had significant effects on the specific glucose uptake rate as well as on the composition of the fermentation products, while we observed no significant changes for overexpression of *pgk* compared to the HC ATPase strain. In particular, the glucose uptake rate

increased by 46.4% (*pfkA*), 48.0% (*pflB*), and 53.4% (*pfkA+pflB*) compared to the HC ATPase strain and remained almost constant for the *pgk* overexpressing strain. As expected and qualitatively predicted by the model, due to the enhancement of the PFL flux, the lactate yield is largely reduced in the *pflB* and *pfkA+pflB* overexpressing strains. The lactate yield dropped also in the *pfkA* overexpressing strain, although to a lesser extent. Here, it is likely that the higher overall glycolytic flux, enabled by higher PFK activity, increases ATP supply and thus leads to less downregulation of PFL. Therefore, more carbon is redirected to acetate, ethanol, and formate and less to lactate also in the *pfkA* overexpressing strain. We also observed that the growth rate was increased, at most in the two *pfkA* overexpressing strains and less in the strain overproducing (only) PFL. This might again be related to the relative high protein costs of PFL. Although the glucose uptake rates of the *pfkA* and *pflB* overexpression strains were still below WT level, the data in Table 7.2 confirm that PFK and PFL are indeed limiting (bottleneck) reactions for the glycolytic rate in the HC ATPase strain, which can be partly overcome by overexpression of the corresponding genes.

3 Discussion

In this study, we systematically analyzed the consequences of increasing ATP demand on the physiology of the *E. coli* wild-type strain MG1655 under various conditions (aerobic/anaerobic, with/ without cell growth). On the one hand, this study was curiosity-driven to explore maximal physiological capabilities of *E. coli* and to investigate how this bacterium responds to situations of high ATP demand (which may be relevant under challenging environmental conditions, e.g., under osmotic, acidic, or toxin-induced stress). On the other hand, this study aimed to deliver new insights toward the use and potential of enforced ATP wasting as a metabolic engineering strategy. We collected a comprehensive dataset on metabolic fluxes, metabolite concentrations, and protein abundances. In order to integrate these data with our current knowledge of the complex metabolism and its regulation in *E. coli*, we constructed a kinetic model that enabled us to modulate the ATP maintenance reaction and thus simulating different levels of ATP wasting in the cells. The key findings of this study can be summarized as follows. First, in all conditions tested, there is a biphasic steady-state response curve of the glucose uptake rate with respect to increasing

ATPase activity. There is a maximum uptake rate at a medium ATPase level, and the glucose uptake rate drops markedly beyond this level. Second, the model indicates that

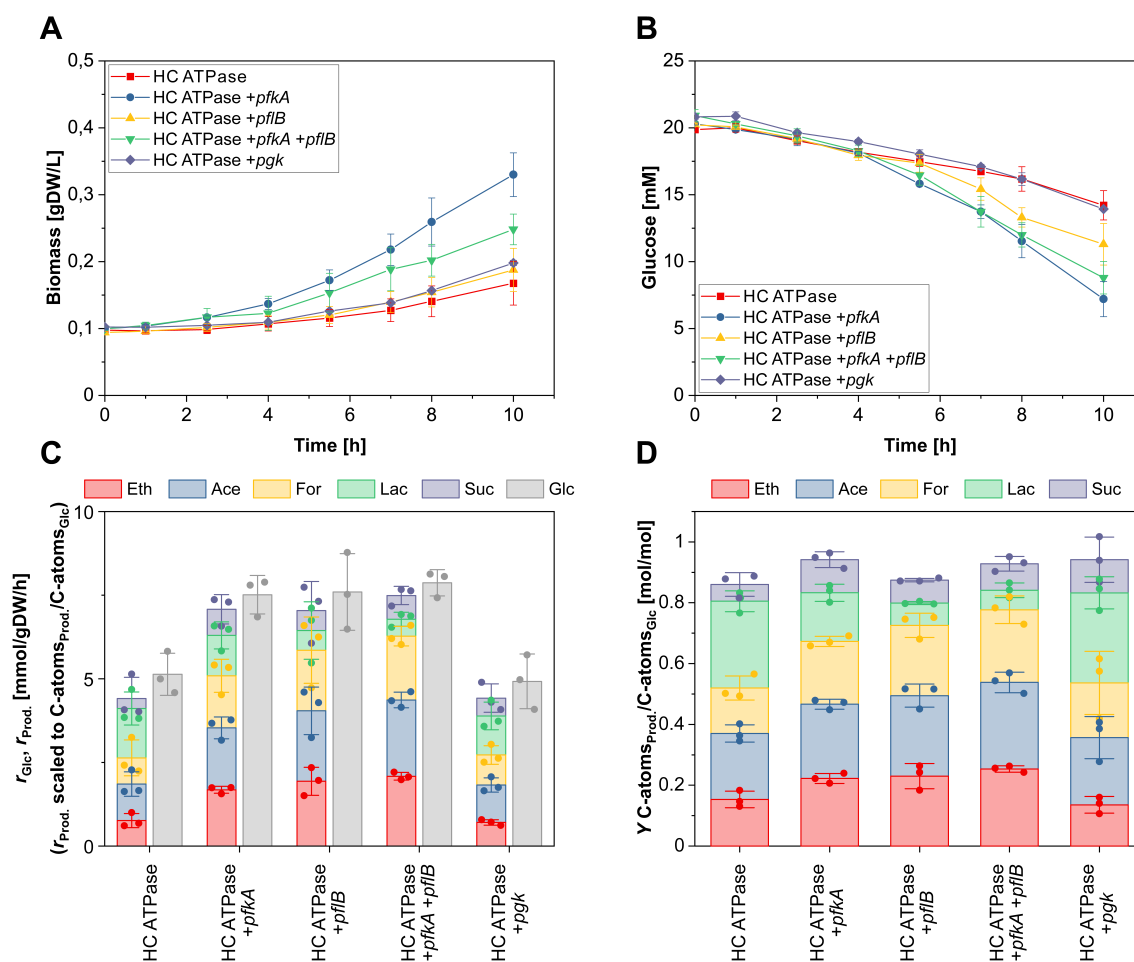


Figure 7.7: Anaerobic growth of the five HC ATPase strain variants. (A) Time course of biomass concentration. (B) Time course of glucose concentration. (C) Specific glucose (Glc) uptake rate and specific productivity for ethanol (Eth), acetate (Ace), formate (For), lactate (Lac), and succinate (Suc). (D) Yield of Eth, Ace, For, Lac, and Suc. Data information: The reaction rates in (C) were calculated for the exponential phase under assumption of quasi-steady state. The means (A and B) and the means and individual data (C and D) for $n = 3$ biologically independent samples are shown. The error bars represent \pm SD. Source data are available online for this figure.

Table 7.2: Growth rates, specific glucose uptake and product synthesis rates, and product yields of the five HC ATPase variants during anaerobic growth. For the calculations of the ATPM and the ATPase rates, see Methods. The means \pm SD of $n = 3$ biologically independent samples are shown. P values for a two-sample t -test are given with respect to the HC ATPase strain. $P > 0.05$ was considered as not significant (n.s.).

	HC ATPase	HC ATPase + <i>pkk</i>	HC ATPase + <i>pkB</i>	HC ATPase + <i>pkA</i> + <i>pkB</i>	HC ATPase + <i>pkk</i>
μ [h ⁻¹]	0.063 \pm 0.009	0.134 \pm 0.013 ($P=0.003$)	0.079 \pm 0.009 ($P=0.161$, n.s.)	0.098 \pm 0.006 ($P=0.011$)	0.074 \pm 0.003 ($P=0.203$, n.s.)
$r_{Glucose}$ [mmol/gDW/h]	5.13 \pm 0.51	7.51 \pm 0.47 ($P=0.009$)	7.59 \pm 0.94 ($P=0.031$)	7.87 \pm 0.32 ($P=0.003$)	4.92 \pm 0.67 ($P=0.743$, n.s.)
r_{Eth} [mmol/gDW/h]	2.29 \pm 0.51	5.05 \pm 0.26 ($P=0.002$)	5.81 \pm 1.02 ($P=0.012$)	6.26 \pm 0.28 ($P=0.001$)	2.12 \pm 0.21 ($P=0.686$, n.s.)
r_{Acet} [mmol/gDW/h]	3.28 \pm 0.39	5.56 \pm 0.54 ($P=0.009$)	6.32 \pm 0.79 ($P=0.008$)	6.84 \pm 0.29 ($P=0.001$)	3.36 \pm 0.35 ($P=0.835$, n.s.)
r_{Form} [mmol/gDW/h]	4.69 \pm 0.96	9.32 \pm 0.84 ($P=0.007$)	10.88 \pm 1.38 ($P=0.006$)	11.44 \pm 0.33 ($P=0.001$)	5.38 \pm 0.32 ($P=0.385$, n.s.)
r_{Lac} [mmol/gDW/h]	2.95 \pm 0.31	2.41 \pm 0.15 ($P=0.019$)	1.18 \pm 0.22 ($P=0.001$)	1.01 \pm 0.35 ($P=0.008$)	2.33 \pm 0.22 ($P=0.027$)
r_{Suc} [mmol/gDW/h]	0.45 \pm 0.17	1.17 \pm 0.05 ($P=0.004$)	0.89 \pm 0.03 ($P=0.024$)	1.07 \pm 0.08 ($P=0.010$)	0.80 \pm 0.04 ($P=0.047$)
r_{ATPM} (calculated) [mmol/gDW/h]	6.99 \pm 0.73	9.26 \pm 0.23 ($P=0.014$)	12.67 \pm 2.02 ($P=0.020$)	12.97 \pm 0.39 ($P=0.001$)	5.47 \pm 0.34 ($P=0.055$, n.s.)
r_{ATPase} (calculated) [mmol/gDW/h]	4.32	6.59	10.00	10.30	2.80
Y_{Eth}	0.065 \pm 0.015	0.099 \pm 0.005 ($P=0.038$)	0.053 \pm 0.009 ($P=0.382$, n.s.)	0.068 \pm 0.004 ($P=0.794$, n.s.)	0.075 \pm 0.005 ($P=0.404$, n.s.)
Y_{Acet}	0.460 \pm 0.066	0.667 \pm 0.040 ($P=0.019$)	0.690 \pm 0.101 ($P=0.055$, n.s.)	0.761 \pm 0.026 ($P=0.004$)	0.407 \pm 0.067 ($P=0.471$, n.s.)
Y_{Form}	0.650 \pm 0.004	0.732 \pm 0.030 ($P=0.018$)	0.794 \pm 0.025 ($P=0.001$)	0.834 \pm 0.057 ($P=0.008$)	0.663 \pm 0.105 ($P=0.873$, n.s.)
Y_{Lac}	0.899 \pm 0.087	1.239 \pm 0.039 ($P=0.007$)	1.387 \pm 0.016 ($P=0.001$)	1.434 \pm 0.057 ($P=0.002$)	1.078 \pm 0.170 ($P=0.254$, n.s.)
Y_{Suc}	0.570 \pm 0.078	0.320 \pm 0.039 ($P=0.015$)	0.147 \pm 0.062 ($P=0.004$)	0.128 \pm 0.070 ($P=0.004$)	0.592 \pm 0.084 ($P=0.796$, n.s.)
Y_{Suc} [mol/mol]	0.083 \pm 0.017	0.163 \pm 0.003 ($P=0.003$)	0.113 \pm 0.011 ($P=0.103$, n.s.)	0.131 \pm 0.002 ($P=0.016$)	0.168 \pm 0.033 ($P=0.037$)

the PFK reaction with its dual dependency on ADP/ATP causes this biphasic behavior. The PFK mechanism is known to buffer increased ATP demands by a higher glycolytic flux (due to elevated ADP levels) but, as shown herein, it collapses under high ATP demands with low ATP concentrations. Third, the metabolomics data under anaerobic conditions show an increasing accumulation of glycolytic metabolites reaching highest values at maximal ATPase level. This behavior cannot be explained with current knowledge, and we postulate that there are unknown regulatory mechanisms for PYK (presumably allosteric regulation by pyruvate or alanine) and PFL (enzyme-level regulation). Finally, we validated the model predictions that PFK and PFL are rate-limiting in the HC ATPase strain and found that overexpressing the genes of these enzymes indeed restores some of the glycolytic capacity. The kinetic model played an important role in this study to identify and analyze potential mechanisms in the metabolism of *E. coli* that led to the observed phenotypes under high ATP demand. It has to be noted that the model is relatively large and comprises more than 100 unknown parameters, many of which will not be uniquely identifiable, despite fitting the model against a considerable set of data. However, the model is based on established biological knowledge of *E. coli*'s central metabolism, it is able to reproduce measurements of the different strains reasonably well and it gave predictions that could be successfully verified. Moreover, the results of the Monte Carlo sampling of kinetic parameters showed that key properties of the kinetic model and its predictions are robust over a wide range of parameter variations. Hence, despite potential parameter identifiability issues, the model could demonstrate its predictive power and thus represents a solid and plausible basis that supports our hypotheses and explains major findings of this study. However, as is true for every model, we can neither prove its correctness nor that other models with alternative mechanisms may exist that reproduce the observed phenomena equally well.

Our results are in several aspects consistent with current knowledge on the physiology of energy metabolism in *E. coli* but, at the same time, indicate gaps in our understanding. For example, the increasing levels of FBP in the LC and MC ATPase strains are consistent with the linear correlation between glycolytic flux and FBP concentration that has been shown for *E. coli* under various conditions and carbon sources (Kochanowski et al., 2013; Kotte et al., 2010). Based on these earlier findings, it was proposed that FBP - presumably indirectly via the transcription factor Cra (Bley Folly et al., 2018) - acts as a general flux-sensing metabolite for *E. coli*. However, among all strains, the HC ATPase strain had the highest level of intracellular FBP but by far the lowest glycolytic flux. Thus, the generalization of FBP being a flux-sensor might not to be true for extreme metabolic per-

turbations as in the HC ATPase strain. Our data also confirm previous work on overflow metabolism and proteome allocation phenomena in *E. coli*. In particular, the observed increase in acetate formation under aerobic conditions in the LC and MC ATPase strains is likely a consequence of proteome reallocation from respiratory pathways (with high ATP yield but also high protein costs) toward overflow metabolism. This consequently results in lower ATP yields but enables higher glycolytic fluxes (and thereby higher total ATP synthesis rates), due to reduced protein costs (Chen and Nielsen, 2019). The observed downregulation of PFL in the HC ATPase strain under anaerobic conditions could be a strategy to optimize proteome allocation, here as response to low ATP concentrations. Again, lactate excretion (yielding 2 mol ATP per mol of glucose) seems disadvantageous compared to the PFL reaction in combination with the formation of acetate, ethanol, and formate resulting in an ATP yield of 2.5 mol ATP per mol of glucose. However, since PFL is a rather large enzyme (759 amino acids vs. 343 amino acids in the average essential protein in *E. coli* (Gong et al., 2008), under the low ATP levels in the HC ATPase strain, it could be more cost-efficient for *E. coli* (in terms of ATP demand) to use the lactate pathway. The lactate dehydrogenase appears to be constitutively available under anaerobic conditions, and the pathway becomes active with rising pyruvate concentrations. Under growth-arrested conditions with a limited nitrogen source, which further limits the proteome pool, the effect is even more drastic. Here, almost the entire carbon is converted into lactate in the HC ATPase strain (yield of 1.74 mol_{Lac}/molGlc, Fig 7.2; Table 7.1), and increased lactate yields are also observed in the LC and MC ATPase strains. Another interesting insight from the proteomic data in the HC ATPase strain is the replacement of PPC with PCK. Due to the low ATP level, PCK may run in the direction of oxaloacetate and ATP synthesis, thus increasing the ATP yield compared to the sole use of PPC. While overexpression of the PCK genes has been used to increase the ATP yield in *E. coli* strains (Zhang et al., 2009; Chao and Liao, 1993; Kwon et al., 2008; Aslan et al., 2017; Kyselova et al., 2018), we are not aware of a previous report showing that *E. coli* naturally switches to PCK to enhance ATP supply. Identifying the regulatory mechanisms that enable this switch is an interesting aspect of future work. When analyzing proteome (re)allocation, we also need to consider the effect of the ATPase overproduction on the proteome pool. Especially in the HC ATPase strain, the expression of the ATPase genes consumes cellular resources such as amino acids and may thereby reduce the overall capacity to synthesize other proteins, including metabolic enzymes. Hence, in addition to the discussed low ATP concentrations, changes in the proteome composition (such as the reduced PFL abundance) may also be induced by the synthesis costs

of the ATPase subunits. However, there are several evidences that the activity of the ATPase is the dominant factor, rather than the costs of its synthesis. First of all, compared to the WT (where the three components of the F₁-subunit are used to build the F₀F₁-ATP synthase), the proteomic data indicated a moderate averaged 8.1-fold increase of the three F₁-ATPase components in the HC ATPase strain, which is still a comparably small fraction of the overall proteome pool. Moreover, we see a 6.5-fold increase of ATPase abundance already in the MC ATPase strain; hence, there is effectively only a 25% increase of the ATPase level in the HC strain compared to the MC strain. With this relatively small change, it appears unlikely that the drastic change in the glycolytic flux and ATP levels between these two strains is mainly caused by a reduction of available resources for protein synthesis. Furthermore, except for the mentioned major changes, other enzymes of central metabolic steps, especially in the glycolysis, show a relatively constant level in the proteomic data and appear thus to be affected to a minor extent only. Another evidence in this direction is the fact that the additional overexpression of the *pfkA* or/and of the *pflB* gene (but not of the *pgk* gene) in the HC ATPase strain increased the glycolytic flux, as predicted, although this will even further reduce the available proteome pool for other enzymes. To further demonstrate that the lack of ATP and not proteome burden causes the low glycolytic flux in the HC ATPase strain, we cultivated the latter and its control strain anaerobically as before (with glucose as main substrate), but this time with addition of fumarate enabling the strains to gain more ATP (via fumarate respiration). As also suggested by our model, the glycolytic flux in the HC ATPase strain should then increase due to the higher ATP levels fueling the PFK reaction while we do not expect significant changes in the HC control as it is not ATP-limited. In fact, with addition of fumarate, the glucose uptake rate even decreased slightly in the HC control strain, but increased markedly in the HC ATPase strain by more than 100% almost reaching the level of the control strain (Fig EV7.12 and Appendix Table S4). The extra amount of glucose was almost completely converted to lactate while the gained ATP was directly consumed by the ATPase as reflected by a high ATPase flux. This result is another strong indicator that it is the low ATP level in the HC ATPase strain rather than ATPase synthesis costs that prevents higher glycolytic fluxes. The finding that overexpression of the PFK- and PFL-encoding genes may largely increase the glycolytic flux and, especially in the case of PFK, also the growth rate in the HC ATPase strain under anaerobic conditions corroborates, on the one hand, our hypothesis on physiological constraints in this strain but, on the other hand, demonstrates that *E. coli*, as could be expected for the extreme perturbation in the HC ATPase strain, is not for all

conditions primed to adjust an optimal expression pattern maximizing its growth rate (cf. Bruggeman et al., 2020). Clearly, adaptive laboratory evolution may yield strains that adapt to these physiological changes resulting, for example, in an upregulation of PFK.

Our results are also of high relevance for metabolic engineering strategies that harness the concept of enforced ATP wasting for strain optimization. Several previous works have already demonstrated the potential of increased ATP turnover as a strategy to maximize substrate uptake and product synthesis rates (Boecker et al., 2021; Chao and Liao, 1994; Koebmann et al., 2002; Hädicke et al., 2015; Boecker et al., 2019; Zahoor et al., 2020; Liu et al., 2016). However, our study is the first showing that there is an optimal level of ATPase expression, at which the specific glucose uptake rate and metabolic activity reaches a maximum. To use the full potential of ATP wasting, it will be important to find the precise maximum (which will be specific for the production organism, the respective substrate– product combination, and the chosen conditions) and to properly adjust the optimal level of the ATPase. For each condition tested, the MC ATPase strain showed the highest glucose uptake rate (which are, to the best of our knowledge, for some cultivation conditions, the highest ever reported so far). Our data also reveal that the highest relative increase can be seen for the growth-arrested cultivations (+1,016% for aerobic and +380% for anaerobic conditions; Fig 7.3). Moreover, the relative drop in the metabolic activity of the HC ATPase strain is less severe than in the cases with growth indicating a higher robustness against maximal ATPase levels under these conditions. With these findings, we anticipate that the highest potential of enforced ATP wasting lies in the optimization of twostage (or even three-stage (Boecker et al., 2021)) processes, in which ATP wasting may greatly boost the activity of the cells in the (growth-arrested) production phase (Burg et al., 2016; Klamt et al., 2018).

4 Material and Methods

4.1 Strains and plasmid construction

All strains, plasmids, and primers used in this study are summarized in Appendix Table S1. *E. coli* NEB 5-alpha competent cells (New England Biolabs, # C2987U) were used for all cloning techniques and plasmid propagation. Standard molecular cloning techniques followed protocols described earlier (Sambrook and Russell, 2001). The ATPase encoding genes *atpAGD* were amplified from plasmid pCP41::atpAGD (Koebmann et al., 2002) by

polymerase chain reaction (PCR) using the Q5 Hot Start High-Fidelity DNA Polymerase (New England Biolabs, # M0493L) and the primer pair *atpAGD_mono_fw/atpAGD_mono_rv* as described in (Boecker et al., 2019). To construct plasmids pSB58.6 and pSB62.6, *gfp-mut3* was cut out from pSB-T1g and pSB-T2g (Balzer et al, 2013) using restriction enzymes NdeI (New England Biolabs, # R0111S) and BamHI-HF (New England Biolabs, # R3136S). The *atpAGD* PCR-amplicon was digested with the same enzymes and ligated into the plasmid backbones of pSB-T1g and pSB-T2g using T4 DNA Ligase (New England Biolabs, # M0202S), yielding plasmids pSB58.6 and pSB62.6, respectively. To construct plasmid pSB66.1, the pMB1 replicon was cut out from pSB62.6 using restriction enzymes AscI (New England Biolabs, # R0558S) and SpeI-HF (New England Biolabs, # R3133S). The p15A replicon was amplified by PCR from plasmid pZA31-luc (Lutz, 1997) using primer pair *p15A_SpeI_fw/p15A_AscI_rv*. The amplicon was digested with AscI and SpeI-HF and ligated into the AscI/SpeI-HF digested plasmid pSB62.6. To construct the control plasmids pSB60.1, pSB64.1, and pSB68.1, *atpAGD* was cut out from pSB58.6, pSB62.6, and pSB66.1 using restriction enzymes NdeI and BamHI-HF. The 5'-overhangs were filled-in using the Klenow Fragment (Thermo Scientific, # EP0054) and the blunt-ended DNA fragments were self-ligated. The *E. coli* wild-type strain MG1655 (Blattner et al., 1997) was transformed with the ATPase expression and control plasmids, generating three ATPase expression strains (low, medium, and high) and three control strains (low, medium, and high) with varying plasmid copy numbers. For additional expression of *pfkA*, *pflB*, and *pgk* together with *atpAGD* in the high copy plasmid pSB62.6, the plasmid was linearized by PCR with the primer pair *pSB73.4_Gibson_fw/atpD_rv*. The genes encoding *pfkA*, *pflB*, and *pgk* were amplified from the genomic DNA of *E. coli* MG1655 by PCR using the primer pairs *pfkA_rbs_fw/pfkA_rbs_rv*, *pflB_rbs_fw/pflB_rbs_rv*, and *pgk_rbs_fw/pgk_rbs_fw*, respectively. A ribosomal binding site (rbs) from the natural ATPase operon of *E. coli* MG1655 (between *atpD* and *atpC*) was inserted into each forward primer to allow polycistronic expression of *atpAGD* and the respective gene. The linearized plasmid pSB62.6 and the DNA fragments harboring the amplified genes were ligated by Gibson assembly, yielding plasmids pSB84.3, pSB85.3, and pSB88.11 (Appendix Table S1). For coexpression of *pfkA* and *pflB* together with *atpAGD* in the high copy plasmid, pSB84.3 was linearized by PCR with the primer pair *pSB73.4_Gibson_fw/pfkA_rv*. *pflB* was amplified from the genomic DNA of *E. coli* MG1655 by PCR using the primer pair *pflB_rbs2_fw/pflB_rbs_rv*. The same rbs as used above was inserted into the forward primer to allow polycistronic expression of *atpAGD*, *pfkA*, and *pflB* from a single operon. The *pflB*-harboring DNA fragment and the linearized plasmid pSB84.3 were ligated by

Gibson assembly, yielding plasmid pSB86.4 (Appendix Table S1).

4.2 Media and cultivation conditions

All liquid and solid media used for cultivation of the ATPase and control strains contained kanamycin (except for "WT" and "WT + IPTG") with a final concentration of 50 $\mu\text{g/ml}$. For growth assays, cells were freshly transformed with the corresponding plasmid and plated on LB₀ agar plates (10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl, 15 g/l agar). A single colony was picked and used to inoculate 5 ml of LB₀ medium. The medium was incubated at 37°C and 150 rpm for 5 h. For aerobic cultivation, cells were diluted 1:500 into chemically defined medium (MM: 4 g/l glucose, 34 mM NaH₂PO₄, 64 mM K₂HPO₄, 20 mM (NH₄)₂SO₄, 1 μM Fe(SO₄)₄, 300 μM MgSO₄, 1 μM ZnCl₂, 10 μM CaCl₂, adapted from (Tanaka et al., 1967), containing 0.01 mM of IPTG (except for "WT") and cultivated at 37°C and 250 rpm overnight. The cells were centrifuged at 5,000 g, washed, and used to inoculate 25 ml of fresh MM (containing 0.01 mM IPTG, except for "WT") to an optical density at 420 nm (OD₄₂₀) of 0.2 (0.4 for the HC ATPase strain). The cells were cultivated in 250-ml shake flasks with three baffles at 37°C and 250 rpm. For anaerobic cultivation, cells from the LB₀-culture were diluted 1:100 into MM (containing 0.01 mM IPTG, except for "WT") and cultivated at 37°C without shaking overnight. The cells were centrifuged at 5,000 g, washed, and used to inoculate fresh MM (containing 0.01 mM IPTG, except for "WT") to an OD₄₂₀ of 0.2 (0.4 for the HC ATPase and *pfkA*, *pflB*, or *pgk* co-expressing strains). The medium was filled into 5-ml screw-cap glass vials (completely filled to the top), and the vials were incubated at 37°C without shaking. For every time point, new vials were opened to guarantee anaerobic conditions. For cultivation of growth-arrested cells, the same procedures for aerobic and anaerobic cultivation were followed as described above, but MM without added (NH₄)₂SO₄ and an initial OD₄₂₀ of 2.0 were used for cultivation. For anaerobic cultivations in medium containing additionally fumarate, the cells were cultivated as described above, but MM supplemented with 20 mM of fumarate was used for the overnight and main cultures. Cell growth was monitored measuring the OD₄₂₀ and using a factor of 0.22 to convert one OD₄₂₀ unit to gram dry weight per liter (gDW/l). All cultivations were performed in biological triplicates, if not stated otherwise.

4.3 Analytical methods

Extracellular glucose, ethanol, acetate, formate, succinate, lactate, pyruvate, and fumarate in the medium were quantified as described earlier (Boecker et al., 2019). Oxalate was quantified by the same method but was not secreted in significant amounts by the strains. For quantification of intracellular metabolites (except pyruvate), cells (0.5 mg of biomass, from mid-exponential growth phase, growth conditions as described above) were applied to filter disks (Merck Millipore, # HVLP02500) under constant nitrogen flow to keep anaerobic conditions. The medium was removed by suction filtration, and the filter disks were immediately transferred to 1 ml of a -20°C cold acetonitrile/methanol/water (40:40:20) quenching solution. After incubation at -20°C for at least 30 min, the samples were shaken vigorously, and 500 μ l of the mixture was centrifuged at 17,000 g and -9°C for 15 min. Next, 400 μ l of the supernatant was kept at -80°C until metabolite quantification. Extracts were mixed with a ¹³C-labeled internal standard in a 1:1 ratio and analyzed by liquid chromatography-tandem mass spectrometry, which was performed as previously described (Guder et al., 2017) using an Agilent 6495 triple quadrupole mass spectrometer (Agilent Technologies). The ratio of ¹²C and ¹³C peak heights was used to quantify metabolites. ¹²C/¹³C ratios were normalized to OD at the time point of sampling. Absolute ATP, ADP, and AMP concentrations were determined with the ¹³C internal standard and authentic standards (Guder et al., 2017). A specific cell volume of 2 μ l/mg was used to calculate the cell volume. The intracellular adenosine energy charge was calculated with the formula $([ATP] + 0.5[ADP])/([ATP] + [ADP] + [AMP])$. For quantification of intracellular pyruvate, cell extracts were prepared as described above, but 1 mg of biomass and 2 ml of the quenching solution were used. 1.7 ml of the extract were centrifuged at 17,000 g and -9°C for 15 min and 1.5 ml of the supernatant transferred to a new test tube. The solvents were evaporated in a speed-vac and the residues dissolved in 55 μ l of H₂O. Absolute pyruvate concentrations were determined using the pyruvic acid assay kit (Megazyme, # K-PYRUV) and normalized to OD at the time point of sampling. For proteomics analysis, cells were cultivated as described above. 1×10^9 cells were harvested by centrifugation (3 min, 4°C, 17,000 g), the supernatant discarded and the cells resuspended in 2 ml of ice-cold PBS buffer. The cells were centrifuged again (3 min, 4°C, 17,000 g) and the PBS-washing step was repeated twice. Cell pellets were stored at -80°C until further analysis. The pelletized cells were lysed in 400 μ l of 2% sodium lauroyl sarcosinate (SLS) in 100 mM ammonium bicarbonate by heat (20 min, 90°C) and sonication. After 10 min of centrifugation at 17,000 g, the protein concentration in the supernatant was determined with a bicinchoninic acid (BCA)-based

protein assay kit (Thermo Fisher, # 23252). 7.5 μl of 0.2 M tris(carboxylethyl)phosphine in 100 mM ammonium carbonate were added to 300 μl of the supernatant. The solution was incubated for 15 min at 90°C. 7.5 μl of 74 mg/ml iodoacetamide were added to the cooled-off samples and incubated for 30 min at 25°C under shaking of 500 rpm. Using 2% SLS in 100 mM ammonium bicarbonate, 200 μl aliquots of the samples were prepared containing a total protein mass of 50 μg . 600 μl of 100 mM ammonium bicarbonate and 8.5 μl of 0.1 $\mu\text{g}/\mu\text{l}$ porcine trypsin were added to the samples for incubation overnight at 30°C. 5% trifluoroacetic acid was added to the samples to a final concentration of 1.5%. After incubation for 10 min at room temperature, the samples were centrifuged at 17,000 g for 10 min at 4°C. The supernatant was used for solid phase extraction of the peptides using C18-columns (Macherey-Nagel). Peptides were analyzed using a Q-Exactive Plus mass spectrometer connected to an Ultimate 3000 RSLC nano and a nanospray flex ion source (Thermo Scientific). The analytical setting was reported in detail previously (Donati et al, 2021). In short, peptide separation was performed on a reverse-phase HPLC column (75 $\mu\text{m} \times 42 \text{ cm}$) packed in-house with C18 resin (2.4 μm , Dr. Maisch GmbH, Germany). The following separating gradient was used: 96% solvent A (0.15% formic acid) and 4% solvent B (99.85% acetonitrile, 0.15% formic acid) to 30% solvent B over 60 min at a flow rate of 300 nl/min. The data acquisition mode was set with the following parameters: 1 MS scan at a resolution of 70,000 with 50 ms max. ion injection fill time, MS/MS at 17,500 scans of the 10 most intense ions with 50 ms max. fill time. Label-free quantification (LFQ) of the data was performed using Progenesis QIP (Waters) and MASCOT (v2.5, Matrix Science) for spectrum/peptide identification. Progenesis outputs were further processed with SafeQuant (Glatter et al., 2012; Ahrné et al., 2016).

4.4 Calculation of growth rate, specific exchange rates, and yields

For experiments with growth, the growth rate (μ) for the exponential phase was determined by plotting the natural logarithm of the biomass concentrations of each sampled time point (within the exponential growth period) against the cultivation time. The slope of the linear regression equals μ . Specific uptake and excretion rates for the exponential phase in growth-coupled experiments were determined with the formula:

$$r_M = \mu(c_{M,e} - c_{M,s}) / (c_{X,e} - c_{X,s}) \text{ [mmol/gDW/h]}$$

where μ is the growth rate, $c_{M,e}$ and $c_{M,s}$ represent the end and start concentrations of the respective metabolite M (mmol/l glucose, ethanol, acetate, formate, lactate, succi-

nate, pyruvate, or fumarate), and $c_{X,e}$ and $c_{X,s}$ represent the end and start concentrations of the biomass (gDW/l). In experiments with growth arrest, where the biomass concentration remained nearly constant, the specific rates are calculated as:

$$r_M = (c_{M,e} - c_{M,s}) / X_{Av} / \Delta t \text{ [mmol/gDW/h]}$$

where X_{Av} is the average biomass concentration (gDW/l), and $\Delta t = t_e - t_s$ the length of the time period (difference of end and start time). This procedure was used for each of the three replicates from which then the mean and the standard deviation was calculated for each rate.

Metabolite yields were determined by plotting Δc_M (mmol/l) against Δc_{Glc} (mmol/l) for every sampled time point of the exponential growth period (for growth-coupled cultivation) or of the indicated time period (of growth-arrested cultivation). The slope of the linear regression equals the yield of the respective metabolite. Biomass yields were determined by plotting Δc_X (gDW/l) against Δc_{Glc} (g/l) for every sampled time point of the exponential growth period (for growth-coupled cultivation). The slope of the linear regression equals the biomass yield.

4.5 Statistical analysis

Unless stated otherwise, P-values for comparisons between different strains were calculated using an unpaired two-sample t-test with the software OriginPro (version 2020b, OriginLab Corporation). Statistical details of the individual experiments can be found in the captions of the respective tables and figures.

4.6 Metabolic flux analysis to determine ATP turnover rates through the ATPase

Using a stoichiometric model of the central metabolism of *E. coli* (83 reactions and 54 internal metabolites; adapted from (Hädicke and Klamt, 2017) and the MATLAB (MathWorks, version R2020b) toolbox CellNetAnalyzer (Klamt et al., 2007; von Kamp et al., 2017), metabolic flux analysis based on the experimentally determined exchange rates was performed to estimate the ATPase flux in the different strains. In stoichiometric models, unspecific ATP consumption (which includes the non-growth-associated maintenance (NGAM) demand of ATP) is usually represented by an "ATPM" pseudo reaction hydrolyzing ATP. In the ATPase strains, the estimated flux through this reaction

in the stoichiometric model contains both the NGAM demand as well as the actual ATPase flux and the latter can thus be calculated as the difference of the calculated ATPM flux in the ATPase strains and the calculated ATPM flux in the corresponding control strains. Herein, it was assumed that, after consideration of the measured growth rate and exchange fluxes, the remaining degrees in the network were used by the cell to produce a maximum amount of ATP (which is accounted for by maximizing the ATPM flux under the given constraints). When performing these calculations, it happens (especially for anaerobic conditions) that the experimentally determined rates contradict each other (e.g., due to linear dependencies). In those cases, CellNetAnalyzer can be used to minimally adjust the measured rates to obtain a consistent scenario ("Check feasibility" function) before the ATPM flux is maximized. The stoichiometric model together with a detailed description of the calculations is provided on GitHub (see Data availability).

4.7 Kinetic model

The two versions for the kinetic model were implemented and simulated with COPASI (Hoops et al., 2006) and are described in detail in the Appendix Supplementary Text. The model files are also provided on GitHub (see Data availability).

Data availability

The kinetic models (provided in COPASI and SBML format) as well as the stoichiometric model used for calculating the ATPase fluxes (provided as CellNetAnalyzer project and as SBML file) are available under the following GitHub repository:

https://github.com/klamt-lab/Models_E.coli_High_ATP_Demand

Metabolomics MS data: Edmond Repository [Dyld9hM3KIMXqRg2]

<https://edmond.mpdl.mpg.de/imeji/collection/Dyld9hM3KIMXqRg2>

Proteomics MS data: MassIVE Repository MSV000088475

<https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?accession=MSV000088475>

Acknowledgements

We are grateful to Timo Glatter for his support and input concerning proteomics experiments. This work was funded by the European Research Council (721176).

Author contributions

SK conceived and supervised the study. SB performed plasmid and strain construction, cultivations, external metabolite quantification, sampling for proteomics and metabolomics analysis, and determination of metabolites fluxes. GS constructed the kinetic model and carried out the simulations. RS helped in model construction and the Monte Carlo-based sampling of kinetic parameters. TS and HL generated the metabolomics data. TS and WS generated the proteomic data. SB, GS, and SK analyzed the data and wrote the manuscript. SB and GS generated the figures. All authors discussed and approved the content of the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

Keywords and Subject Categories

Keywords ATP homeostasis; central metabolism; glycolysis; kinetic model; metabolic engineering

Subject Categories Metabolism; Microbiology, Virology & Host Pathogen Interaction

Expanded view

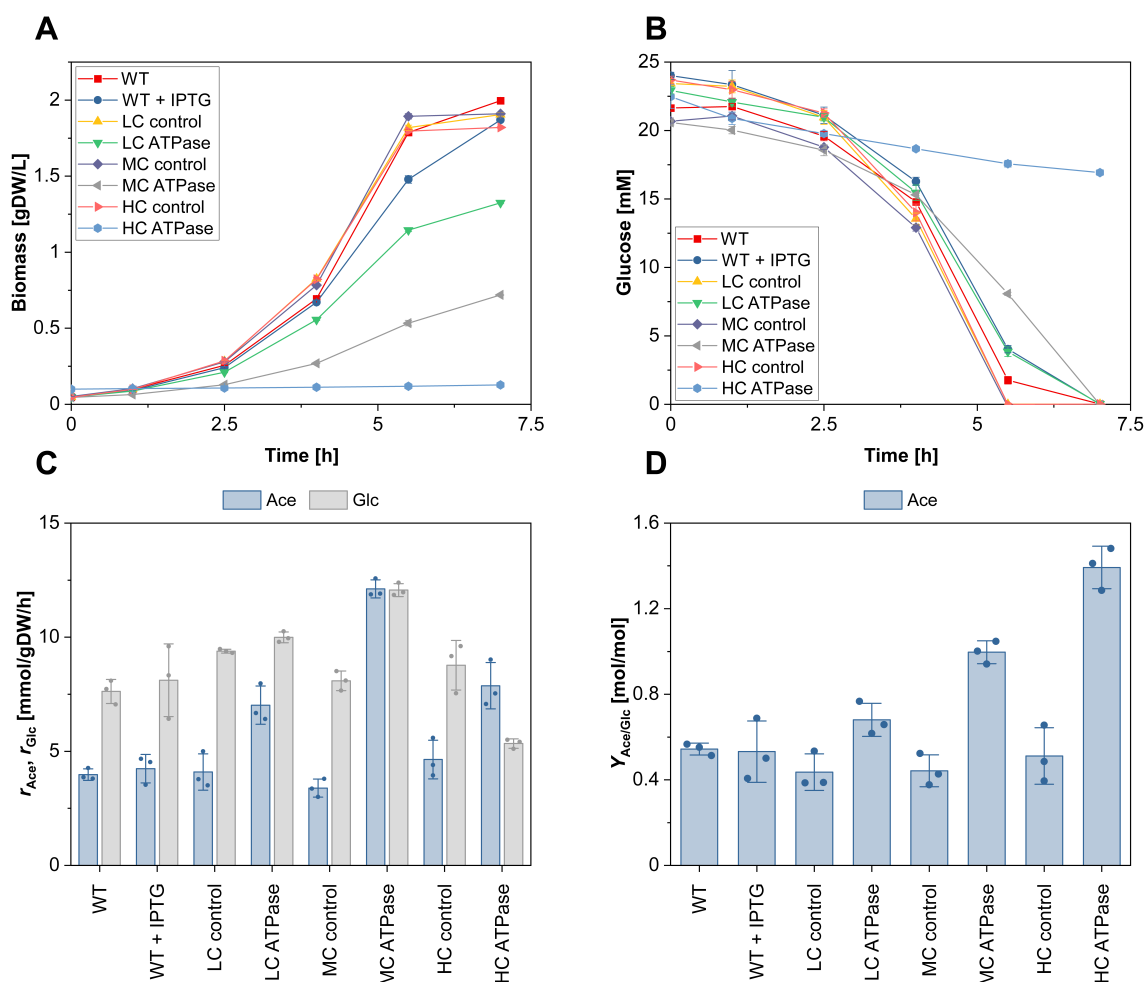


Fig. EV 7.8: Aerobic growth of the different ATPase strains. (A) Time course of biomass concentration. (B) Time course of glucose concentration. (C) Specific glucose (Glc) uptake rate and specific productivity for acetate (Ace). (D) Yield of Ace. Data information: The reaction rates in (C) were calculated for the exponential phase under assumption of quasi-steady state. Because changes in glucose and acetate are rather small during cultivation of the HC ATPase strain, a higher initial biomass concentration of 0.1 gDW/l was used for this strain to get data that are more robust for calculating the metabolite exchange rates. The means (A and B) and the means and individual data (C and D) for $n = 3$ biologically independent samples are shown. The error bars represent \pm SD. Source data are available online for this figure.

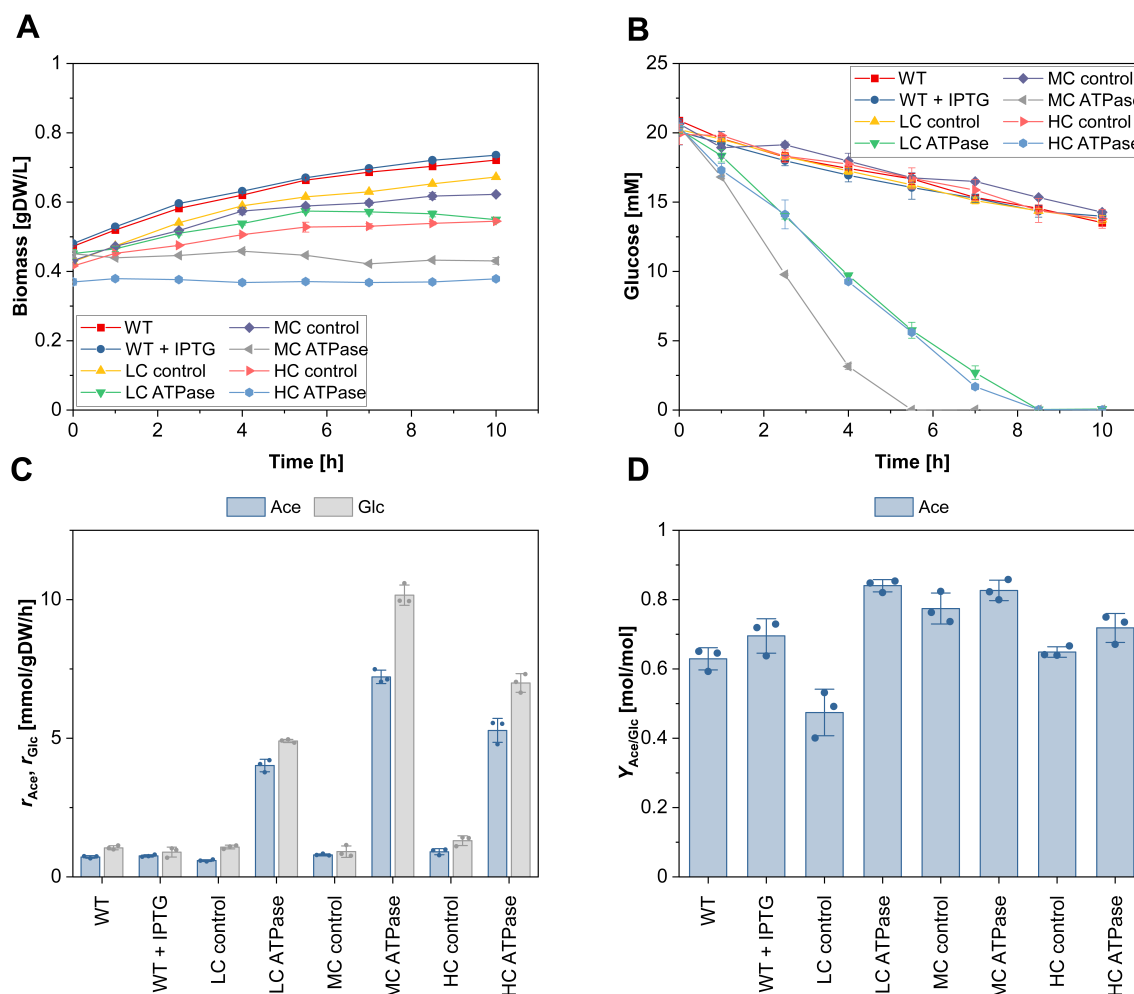


Fig. EV 7.9: Aerobic cultivation of the different strains under growth arrest. (A) Time course of biomass concentration. (B) Time course of glucose concentration. (C) Specific glucose (Glc) uptake rate and specific productivity for acetate (Ace). (D) Yield of Ace. Data information: The reaction rates in (C) were calculated from the beginning of cultivation until the last sampling time point where glucose was still present in the medium. The means (A and B) and the means and individual data (C and D) for $n = 3$ biologically independent samples are shown. The error bars represent \pm SD. Note: although no nitrogen source was present in the medium, some minor growth (especially of the control and wild-type strains) remained (A), which is a known phenomenon within the first hours of cultivation after nitrogen depletion Switzer et al., 2020. Source data are available online for this figure.

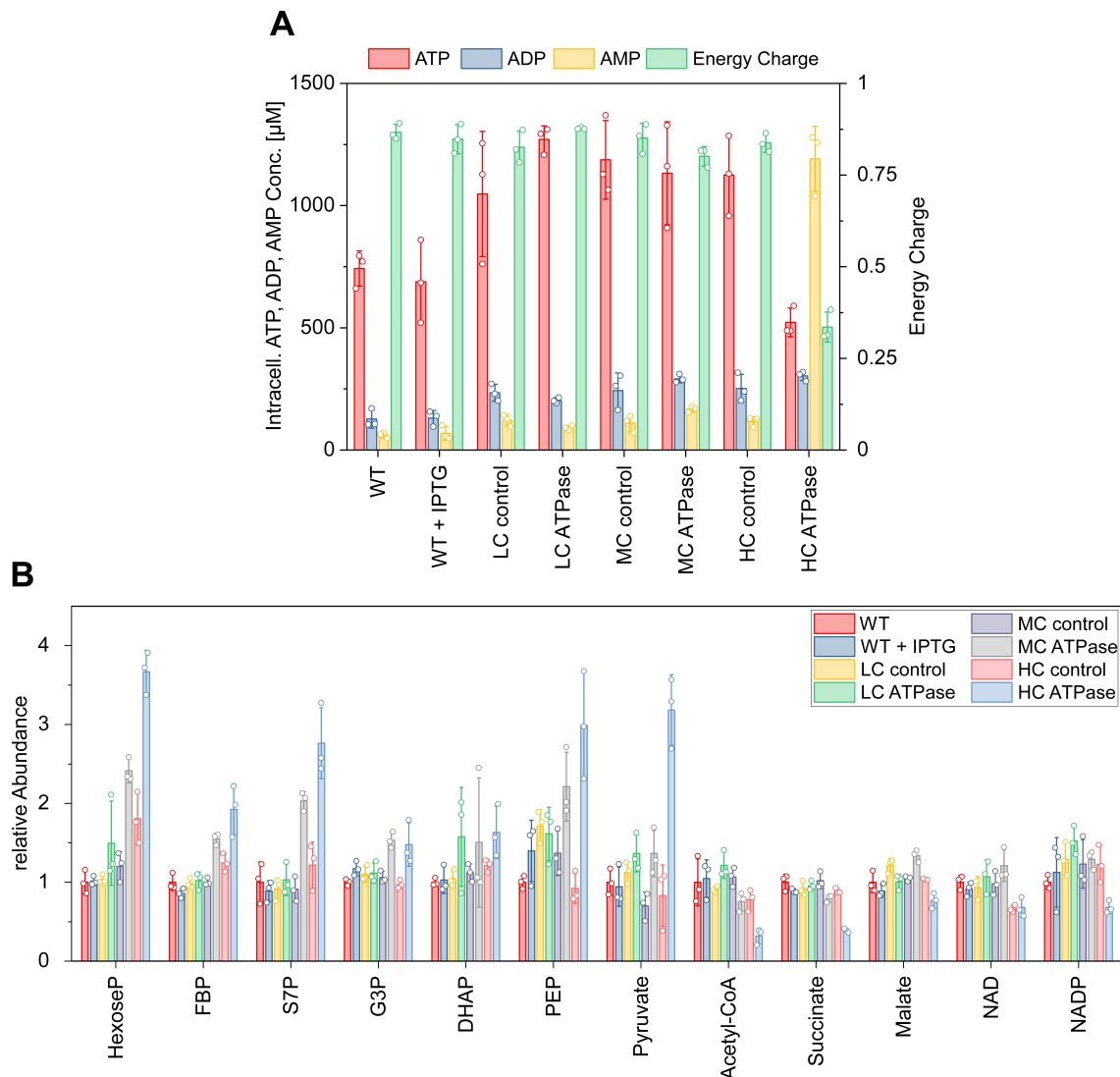


Fig. EV 7.10: Intracellular metabolite concentrations of the different strains during anaerobic growth. (A) Absolute intracellular ATP, ADP, and AMP concentrations and energy charge. (B) Relative intracellular metabolite concentrations from core metabolism. Data information: The means and individual data of $n = 3$ biologically independent samples are shown and the error bars represent \pm SD.

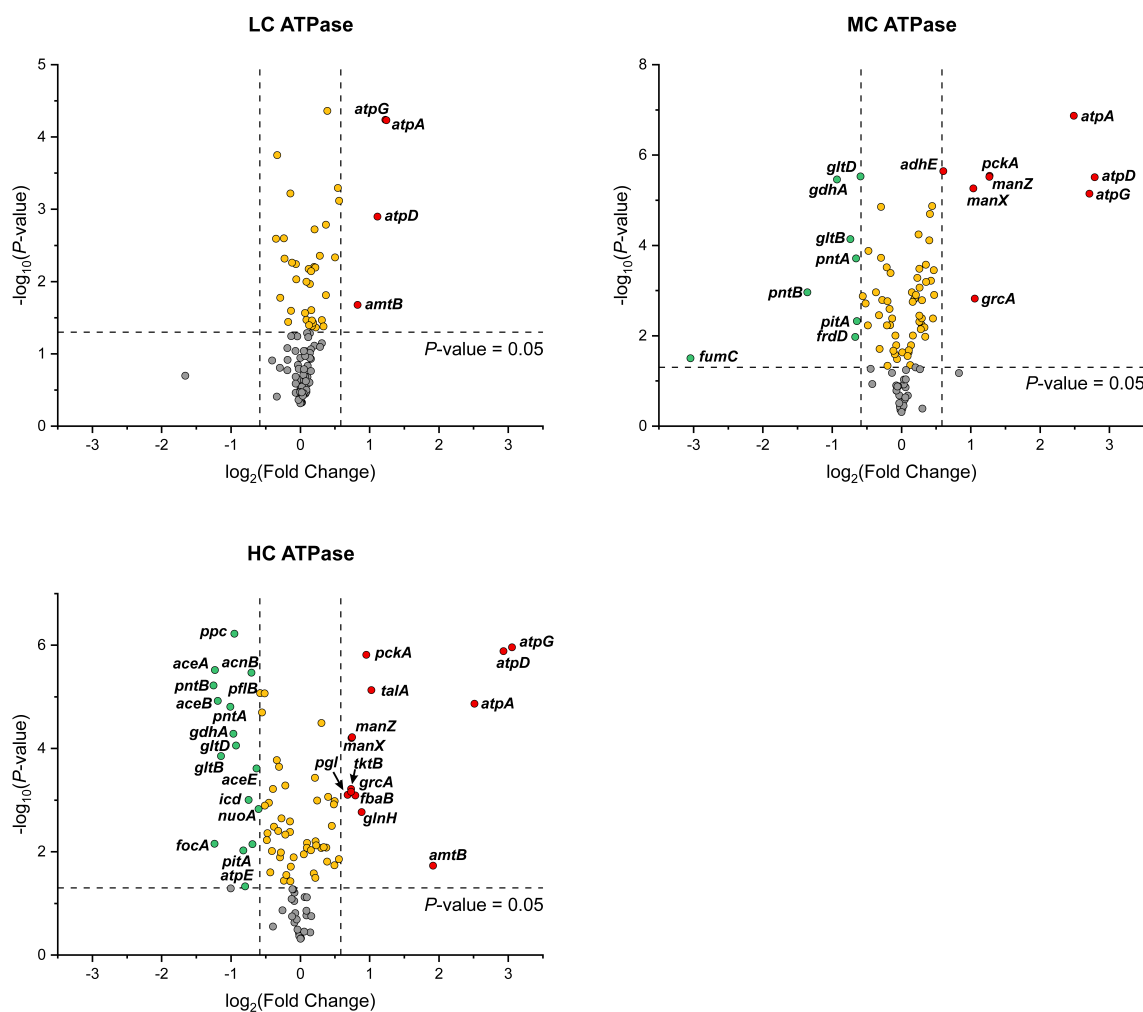


Fig. EV 7.11: Changes on proteome level (only proteins from core metabolism are shown) of the three ATPase strains in regard to the corresponding control strains under anaerobic cultivation conditions with growth. Significantly (P value < 0.05) upregulated (upregulation > 1.5 -fold) proteins are depicted in red, significantly (P value < 0.05) downregulated (downregulation > 0.33 -fold) proteins are depicted in green. Proteins with a significant down- or upregulation but below the thresholds of 1.5-fold up- or 0.33-fold downregulation are depicted in yellow, proteins with no significant change (P value > 0.05) are depicted in grey. The gene names of the corresponding proteins are given for significantly down- or upregulated proteins. P values were calculated for a two-sample t -test from $n = 3$ biologically independent samples.

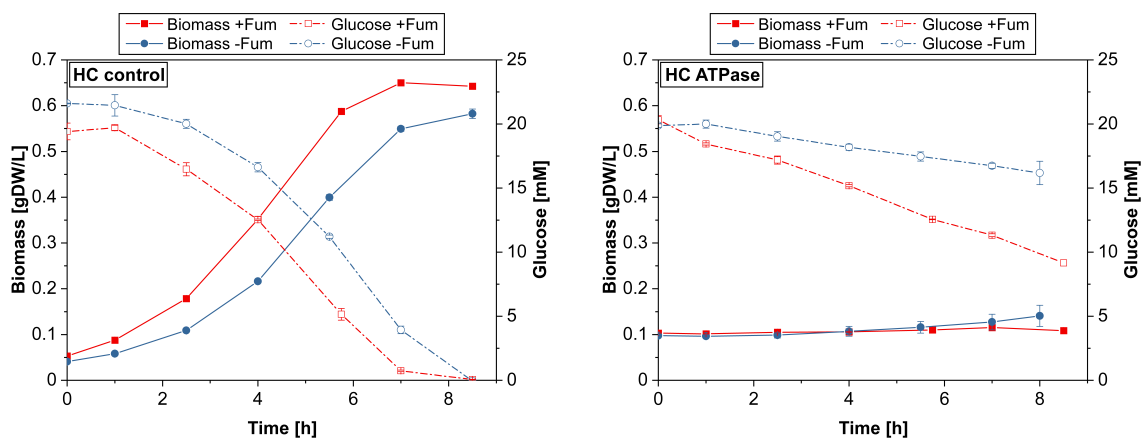


Fig. EV 7.12: Anaerobic growth of the HC control (left) and HC ATPase (right) strains with (red) and without (blue) fumarate addition. The average time courses of biomass and glucose concentrations of $n = 3$ (-Fum) and $n = 2$ (+Fum) biologically independent samples are shown. The error bars represent \pm SD. See also Appendix Table S4 for the determined specific rates. Source data are available online for this figure.

References

- Boecker, S., Harder, B.-J., Kutscha, R., Pflügl, S., & Klamt, S. (2021). Increasing ATP turnover boosts productivity of 2,3-butanediol synthesis in *Escherichia coli*. *Microbial Cell Factories*, *20*(1), 63. <https://doi.org/10.1186/s12934-021-01554-x> (cited on pages 205, 207, 232).
- Jensen, P. R., & Michelsen, O. (1992). Carbon and energy metabolism of atp mutants of *Escherichia coli*. *Journal of Bacteriology*, *174*(23), 7635–7641. <https://doi.org/10.1128/jb.174.23.7635-7641.1992> (cited on page 207).
- Kochanowski, K., Volkmer, B., Gerosa, L., Haverkorn van Rijsewijk, B. R., Schmidt, A., & Heinemann, M. (2013). Functioning of a metabolic flux sensor in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(3), 1130–1135. <https://doi.org/10.1073/pnas.1202582110> (cited on pages 207, 229).
- Chubukov, V., Gerosa, L., Kochanowski, K., & Sauer, U. (2014). Coordination of microbial metabolism. *Nature Reviews Microbiology*, *12*(5), 327–340. <https://doi.org/10.1038/nrmicro3238> (cited on page 207).
- Basan, M., Honda, T., Christodoulou, D., Hörl, M., Chang, Y.-F., Leoncini, E., Mukherjee, A., Okano, H., Taylor, B. R., Silverman, J. M., Sanchez, C., Williamson, J. R., Paulsson, J., Hwa, T., & Sauer, U. (2020). A universal trade-off between growth and lag in fluctuating environments. *Nature*, *584*(7821), 470–474. <https://doi.org/10.1038/s41586-020-2505-4> (cited on page 207).
- Bruggeman, F. J., Planqué, R., Molenaar, D., & Teusink, B. (2020). Searching for principles of microbial physiology. *FEMS microbiology reviews*, *44*(6), 821–844. <https://doi.org/10.1093/femsre/fuaa034> (cited on pages 207, 232).
- Basan, M., Hui, S., Okano, H., Zhang, Z., Shen, Y., Williamson, J. R., & Hwa, T. (2015). Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature*, *528*(7580), 99–104. <https://doi.org/10.1038/nature15765> (cited on page 207).
- Chen, Y., & Nielsen, J. (2019). Energy metabolism controls phenotypes by protein efficiency and allocation. *Proceedings of the National Academy of Sciences*, *116*(35), 17592–17597. <https://doi.org/10.1073/pnas.1906569116> (cited on pages 207, 230).
- Zhang, X., Jantama, K., Moore, J. C., Jarboe, L. R., Shanmugam, K. T., & Ingram, L. O. (2009). Metabolic evolution of energy-conserving pathways for succinate production in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(48), 20180–20185. <https://doi.org/10.1073/pnas.0905396106> (cited on pages 207, 230).
- Singh, A., Cher Soh, K., Hatzimanikatis, V., & Gill, R. T. (2011). Manipulating redox and ATP balancing for improved production of succinate in *E. coli*. *Metabolic Engineering*, *13*(1), 76–81. <https://doi.org/10.1016/j.ymben.2010.10.006> (cited on page 207).
- Kim, H.-J., Kwon, Y. D., Lee, S. Y., & Kim, P. (2012). An engineered *Escherichia coli* having a high intracellular level of ATP and enhanced recombinant protein production. *Applied Microbiology and Biotechnology*, *94*(4), 1079–1086. <https://doi.org/10.1007/s00253-011-3779-0> (cited on page 207).
- Chao, Y. P., & Liao, J. C. (1994). Metabolic responses to substrate futile cycling in *Escherichia coli*. *The Journal of Biological Chemistry*, *269*(7), 5122–5126 (cited on pages 207, 232).
- Koebmann, B. J., Westerhoff, H. V., Snoep, J. L., Solem, C., Pedersen, M. B., Nilsson, D., Michelsen, O., & Jensen, P. R. (2002). The extent to which ATP demand controls the glycolytic flux depends

- strongly on the organism and conditions for growth. *Molecular Biology Reports*, 29(1), 41–45. <https://doi.org/10.1023/A:1020398117281> (cited on pages 207, 209, 232).
- Hädicke, O., Bettenbrock, K., & Klamt, S. (2015). Enforced ATP futile cycling increases specific productivity and yield of anaerobic lactate production in *Escherichia coli*. *Biotechnology and Bioengineering*, 112(10), 2195–2199. <https://doi.org/10.1002/bit.25623> (cited on pages 207, 232).
- Boecker, S., Zahoor, A., Schramm, T., Link, H., & Klamt, S. (2019). Broadening the scope of enforced ATP wasting as a tool for metabolic engineering in *Escherichia coli*. *Biotechnology Journal*, 14(9), 1800438. <https://doi.org/10.1002/biot.201800438> (cited on pages 207, 209, 232, 233, 235).
- Zahoor, A., Messerschmidt, K., Boecker, S., & Klamt, S. (2020). ATPase-based implementation of enforced ATP wasting in *Saccharomyces cerevisiae* for improved ethanol production. *Biotechnology for Biofuels*, 13(1), 185. <https://doi.org/10.1186/s13068-020-01822-9> (cited on pages 207, 232).
- Holm, A. K., Blank, L. M., Oldiges, M., Schmid, A., Solem, C., Jensen, P. R., & Vemuri, G. N. (2010). Metabolic and transcriptional response to cofactor perturbations in *Escherichia coli*. *The Journal of Biological Chemistry*, 285(23), 17498–17506. <https://doi.org/10.1074/jbc.M109.095570> (cited on pages 207, 209).
- Burg, J. M., Cooper, C. B., Ye, Z., Reed, B. R., Moreb, E. A., & Lynch, M. D. (2016). Large-scale bioprocess competitiveness: The potential of dynamic metabolic control in two-stage fermentations. *Current Opinion in Chemical Engineering*, 14, 121–136. <https://doi.org/10.1016/j.coche.2016.09.008> (cited on pages 210, 232).
- Klamt, S., Mahadevan, R., & Hädicke, O. (2018). When do two-stage processes outperform one-stage processes? *Biotechnology Journal*, 13(2), 1700539. <https://doi.org/10.1002/biot.201700539> (cited on pages 210, 232).
- Switzer, A., Burchell, L., McQuail, J., & Wigneshweraraj, S. (2020). The adaptive response to long-term nitrogen starvation in *Escherichia coli* requires the breakdown of allantoin. *Journal of Bacteriology*, 202(17), e00172–20. <https://doi.org/10.1128/JB.00172-20> (cited on pages 213, 241).
- Fenton, A. W., & Reinhart, G. D. (2009). Disentangling the web of allosteric communication in a homotetramer: Heterotropic inhibition in phosphofructokinase from *Escherichia coli*. *Biochemistry*, 48(51), 12323–12328. <https://doi.org/10.1021/bi901456p> (cited on page 217).
- Kotlarz, D., Garreau, H., & Buc, H. (1975). Regulation of the amount and of the activity of phosphofructokinases and pyruvate kinases in *Escherichia coli*. *Biochimica Et Biophysica Acta*, 381(2), 257–268. [https://doi.org/10.1016/0304-4165\(75\)90232-9](https://doi.org/10.1016/0304-4165(75)90232-9) (cited on page 217).
- Blangy, D., Buc, H., & Monod, J. (1968). Kinetics of the allosteric interactions of phosphofructokinase from *Escherichia coli*. *Journal of Molecular Biology*, 31(1), 13–35. [https://doi.org/10.1016/0022-2836\(68\)90051-X](https://doi.org/10.1016/0022-2836(68)90051-X) (cited on page 217).
- Peskov, K., Goryanin, I., & Demin, O. (2008). Kinetic model of phosphofructokinase-1 from *Escherichia coli*. *Journal of Bioinformatics and Computational Biology*, 6(4), 843–867. <https://doi.org/10.1142/S0219720008003643> (cited on page 217).
- Liebermeister, W., & Klipp, E. (2006). Bringing metabolic networks to life: Convenience rate law and thermodynamic constraints. *Theoretical Biology & Medical Modelling*, 3, 41. <https://doi.org/10.1186/1742-4682-3-41> (cited on page 217).

- Park, J. O., Rubin, S. A., Xu, Y.-F., Amador-Noguez, D., Fan, J., Shlomi, T., & Rabinowitz, J. D. (2016). Metabolite concentrations, fluxes and free energies imply efficient enzyme usage. *Nature Chemical Biology*, *12*(7), 482–489. <https://doi.org/10.1038/nchembio.2077> (cited on page 223).
- Taber, R. L., Campbell, A., & Spencer, S. (1998). A simple experiment demonstrating the allosteric regulation of yeast pyruvate kinase. *Biochemical Education*, *26*(1), 73–76. [https://doi.org/10.1016/S0307-4412\(97\)00117-9](https://doi.org/10.1016/S0307-4412(97)00117-9) (cited on page 223).
- Murabito, E., Verma, M., Bekker, M., Bellomo, D., Westerhoff, H. V., Teusink, B., & Steuer, R. (2014). Monte-carlo modeling of the central carbon metabolism of lactococcus lactis: Insights into metabolic regulation. *PLOS ONE*, *9*(9), e106453. <https://doi.org/10.1371/journal.pone.0106453> (cited on page 224).
- Sauro, H. M. (2013). *Systems biology: An introduction to metabolic control analysis* (1st edition, version 1.01). Ambrosius Publishing. (Cited on page 224).
- Kotte, O., Zaugg, J. B., & Heinemann, M. (2010). Bacterial adaptation through distributed sensing of metabolic fluxes. *Molecular Systems Biology*, *6*, 355. <https://doi.org/10.1038/msb.2010.10> (cited on page 229).
- Bley Folly, B., Ortega, A. D., Hubmann, G., Bonsing-Vedelaar, S., Wijma, H. J., van der Meulen, P., Miliars-Argentis, A., & Heinemann, M. (2018). Assessment of the interaction between the flux-signaling metabolite fructose-1,6-bisphosphate and the bacterial transcription factors CggR and cra. *Molecular Microbiology*, *109*(3), 278–290. <https://doi.org/10.1111/mmi.14008> (cited on page 229).
- Gong, X., Fan, S., Bilderbeck, A., Li, M., Pang, H., & Tao, S. (2008). Comparative analysis of essential genes and nonessential genes in escherichia coli k12. *Molecular genetics and genomics: MGG*, *279*(1), 87–94. <https://doi.org/10.1007/s00438-007-0298-x> (cited on page 230).
- Chao, Y. P., & Liao, J. C. (1993). Alteration of growth yield by overexpression of phosphoenolpyruvate carboxylase and phosphoenolpyruvate carboxykinase in escherichia coli. *Applied and Environmental Microbiology*, *59*(12), 4261–4265. <https://doi.org/10.1128/aem.59.12.4261-4265.1993> (cited on page 230).
- Kwon, Y.-D., Lee, S. Y., & Kim, P. (2008). A physiology study of escherichia coli overexpressing phosphoenolpyruvate carboxykinase. *Bioscience, Biotechnology, and Biochemistry*, *72*(4), 1138–1141. <https://doi.org/10.1271/bbb.70831> (cited on page 230).
- Aslan, S., Noor, E., & Bar-Even, A. (2017). Holistic bioengineering: Rewiring central metabolism for enhanced bioproduction. *Biochemical Journal*, *474*(23), 3935–3950. <https://doi.org/10.1042/BCJ20170377> (cited on page 230).
- Kyselova, L., Kreitmayer, D., Kremling, A., & Bettenbrock, K. (2018). Type and capacity of glucose transport influences succinate yield in two-stage cultivations. *Microbial Cell Factories*, *17*(1), 132. <https://doi.org/10.1186/s12934-018-0980-1> (cited on page 230).
- Liu, J., Kandasamy, V., Würtz, A., Jensen, P. R., & Solem, C. (2016). Stimulation of acetoin production in metabolically engineered lactococcus lactis by increasing ATP demand. *Applied Microbiology and Biotechnology*, *100*(22), 9509–9517. <https://doi.org/10.1007/s00253-016-7687-1> (cited on page 232).
- Sambrook, J., & Russell, D. W. (2001). *Molecular cloning: A laboratory manual* (3rd ed). Cold Spring Harbor Laboratory Press. (Cited on page 232).
- Lutz, R. (1997). Independent and tight regulation of transcriptional units in escherichia coli via the LacR/o, the TetR/o and AraC/i1-i2 regulatory elements. *Nucleic Acids Research*, *25*(6), 1203–1210. <https://doi.org/10.1093/nar/25.6.1203> (cited on page 233).

- Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., & Shao, Y. (1997). The complete genome sequence of escherichia coli k-12. *Science (New York, N.Y.)*, 277(5331), 1453–1462. <https://doi.org/10.1126/science.277.5331.1453> (cited on page 233).
- Tanaka, S., Lerner, S. A., & Lin, E. C. C. (1967). Replacement of a phosphoenolpyruvate-dependent phosphotransferase by a nicotinamide adenine dinucleotide-linked dehydrogenase for the utilization of mannitol. *Journal of Bacteriology*, 93(2), 642–648. Retrieved November 2, 2020, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC276489/> (cited on page 234).
- Guder, J. C., Schramm, T., Sander, T., & Link, H. (2017). Time-optimized isotope ratio LC–MS/MS for high-throughput quantification of primary metabolites. *Analytical Chemistry*, 89(3), 1624–1631. <https://doi.org/10.1021/acs.analchem.6b03731> (cited on page 235).
- Glatter, T., Ludwig, C., Ahrné, E., Aebersold, R., Heck, A. J. R., & Schmidt, A. (2012). Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem lys-c/trypsin proteolysis over trypsin digestion. *Journal of Proteome Research*, 11(11), 5145–5156. <https://doi.org/10.1021/pr300273g> (cited on page 236).
- Ahrné, E., Glatter, T., Viganò, C., Schubert, C. v., Nigg, E. A., & Schmidt, A. (2016). Evaluation and improvement of quantification accuracy in isobaric mass tag-based protein quantification experiments. *Journal of Proteome Research*, 15(8), 2537–2547. <https://doi.org/10.1021/acs.jproteome.6b00066> (cited on page 236).
- Hädicke, O., & Klamt, S. (2017). EColiCore2 : A reference network model of the central metabolism of escherichia coli and relationships to its genome-scale parent model. *Scientific Reports*, 7(1), 39647. <https://doi.org/10.1038/srep39647> (cited on page 237).
- Klamt, S., Saez-Rodriguez, J., & Gilles, E. D. (2007). Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC Systems Biology*, 1(1), 2. <https://doi.org/10.1186/1752-0509-1-2> (cited on page 237).
- von Kamp, A., Thiele, S., Hädicke, O., & Klamt, S. (2017). Use of CellNetAnalyzer in biotechnology and metabolic engineering. *Journal of Biotechnology*, 261, 221–228. <https://doi.org/10.1016/j.jbiotec.2017.05.001> (cited on page 237).
- Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., Singhal, M., Xu, L., Mendes, P., & Kummer, U. (2006). COPASI—a COMplex PATHway SIMulator. *Bioinformatics (Oxford, England)*, 22(24), 3067–3074. <https://doi.org/10.1093/bioinformatics/btl485> (cited on page 238).

Conclusion and outlook

Thorben Schramm

Key findings

Chapter 3: We systematically investigated in-source modifications during FI-MS with authentic metabolite standards spiked into a metabolite extract sample. We found that a standard produced on average 68 significant m/z features. Using a network approach and considering in-source fragmentation, we were able to explain 49 % of the significant features. We showed that extensive in-source modifications occur during FI-MS. Yet, many of these modifications can be explained. Our data is a useful resource helping to avoid misannotations of metabolites.

Chapter 4: We created a large library of *argG* mutants in *E. coli* and presented a high-throughput method to enrich temperature-sensitive mutants in the library. Our method combined a TIMER-based single cell growth rate reporter and fluorescence activated cell sorting (FACS) to sort slow or non-growing cells. 90 % of randomly isolated strains from the enriched population were temperature-sensitive showing a high efficiency of our enrichment approach. We further showed that temperature-sensitive *argG* can be used as a metabolic valve to dynamically arrest growth and, simultaneously, overproduce citrulline in a two-stage bioprocess. In 1 L-bioreactors, we produced 3 g/L citrulline within 45 h.

Chapter 5: Using a customized barcoded CRISPR/Cas9 method, we created a pooled library of 15,120 *E. coli* strains. Each strain had a single amino acid substitution in one of 346 essential genes. We showed that competitive growth assays coupled to deep sequencing of plasmid-borne barcodes can be used to identify temperature-sensitive strains on a large scale. Growth analysis of 92 isolated temperature-sensitive strains revealed different degrees of temperature-sensitivity among the strains and a gradual response of growth to the temperature. By FI-MS, we show that 42 temperature-sensitive enzymes, out of 80 isolated strains with mutant enzymes, function as metabolic valves that introduce metabolic bottlenecks at 42 °C. The data also revealed that metabolic bottlenecks can cause secondary bottlenecks. Further, we used seven strains to decouple cell growth from the overproduction of chemicals. A temperature-sensitive strain carrying a mutation in DNA replication, *dnaX*(Leu289Gln), allowed us to control growth in an arginine overproduction strain creating a two-stage bioprocess. Our results show a great potential of temperature-sensitivity as tool in metabolic engineering. With 92 identified temperature-sensitive mutants, we also provide a valuable resource for future studies in *E. coli*.

Chapter 6: The concept of *enforced ATP wasting* was tested in a strain overexpressing ATPase. In the ATPase strain, glucose uptake rates were the second highest reported for growth-arrested strains under anaerobic conditions. Higher rates were only achieved in a follow-up study (see Chapter 7). The consumption of glucose and the production of fermentation products remained high until glucose was depleted from the medium. These results showed that *enforced ATP wasting* by ATPase overexpression can improve product titers and yields as well as the production rate.

Chapter 7: Following up on Chapter 6, different expression strengths of ATPase in *E. coli* were tested in growing and non-growing cells under aerobic and anaerobic conditions. With increasing ATPase expression levels, the glucose uptake rates increased. However, at very high expression levels, the uptake rate decreased sharply and was lower than in a wild type strain. Phosphofructokinase caused this effect since it is activated by ADP and has ATP as substrate. Further, we showed that the PfkA and PflB catalysed reactions were rate limiting steps under very high ATPase expression levels. These results revealed how glycolytic flux is adjusted to increases in ATP demand and what the limits to this compensatory mechanism are. The study further showed that *enforced ATP wasting* is very effective to increase metabolic activity under growth arrest.

Open questions

How can we extend the number of explainable in-source modifications during FI-MS?

High-throughput metabolomics methods like FI-MS are very useful in metabolic engineering attempts that rely on testing hundreds or thousands of engineered strains. Our systematic analysis of in-source modifications can help avoid misannotations of metabolites during FI-MS (Chapter 3) and, thus, can improve the method. Although, in-source modifications are abundant and easily detected during FI-MS, they are not exclusive to FI-MS. Therefore, our findings about in-source modifications are also valuable to other mass spectrometry methods that use ESI. With our network approach, we could already explain 49 % of the significant peaks using a set of 51 reference mass differences from literature. An open question is how can the other 51 % of the significant peaks be explained?

Extending the list of reference peaks can explain more peaks. However, systematically adding all possible carbohydrates with a mass below 1,100 Da leads to a combinatorial explosion. This poses the risk that a m/z value in the list matches with a measured m/z difference by chance and does not reflect the actual in-source modification. For example, a modification with an adduct that is not soluble or has not been observed in biological cells is unrealistic. This problem is even more severe if a network approach is used. Then, sequential combinations could potentially cover all possible masses. Thus, we require constraints for the reference list that permit only chemically feasible and likely modifications.

A possible approach is to focus on how metabolites interact directly with each other during ESI and form new compounds. For this analysis, measuring pairwise combinations of metabolite standards could help identifying common chemical synthesis reactions during ESI. Using the information about such common synthesis reactions, the number of peaks in our reference list can then be increased.

Another approach to increase the number of explainable peaks is to make more use of isotope-labelled metabolites, either as pure standards or as uniformly-labelled metabolite extract sample. This can help identifying the origin of an in-source modification. For example, in our FI-MS study, we observed that G3P had a large number of significant m/z features. However it is unclear whether these peaks derive from in-source fragmentations, chemical reactions with itself or its fragments, chemical reactions with other metabolites, or other in-source modifications. Measuring G3P as pure standards as well as spiked-into fully ^{13}C -, ^{15}N -, and [$^{13}\text{C}+^{15}\text{N}$]-labelled metabolite extract samples would provide information about the origin of an adduct.

Which mutations induce temperature-sensitivity?

An algorithm that predicts temperature-sensitive mutations with high accuracy would be very useful for metabolic engineering but also for studying essential genes. So far, only few prediction tools are available, and the here used TSPred algorithm (Tan et al., 2014; Varadarajan et al., 1996) still requires the testing of many predicted mutants. A challenge is that in order to refine the prediction tools, we need many more temperature-sensitive mutants to identify better design laws or to train machine-learning algorithms. In this, we need not only many more mutants for a single gene but for many genes. Ideally, the mutants have very few mutations making it easier to deconvolute the impact of a mutation to the stability of a protein. High number of mutations or missing sequencing data is also the reason why existing libraries of temperature-sensitive mutants

(Ben-Aroya et al., 2008; Li et al., 2011; Kofoed et al., 2015) are only of limited use to learn design rules of temperature-sensitive mutations.

In our study on two-stage citrulline production (Chapter 4), we described a high-throughput method to enriched temperature-sensitive *argG* mutants using a flow cytometry approach. Out of ca. 81.000 sorted cells, we randomly isolated 90 mutants. 81 of them were temperature-sensitive (90 %). These results indicate that the absolute number of possible temperature-sensitive variants of a protein could be very high. Considering how huge the design space of proteins is, finding a temperature-sensitive mutant is still challenging though. We also observed that the temperature-sensitive *ArgG* variants, which we sequenced, had between 2 and 11 mutations. Some of these mutations were synonymous. In case of multiple non-synonymous mutations, it requires tedious follow-up work to clarify, which of the mutations contribute to temperature-sensitivity. In future studies, both, the total number of mutations and the number of synonymous mutations could be reduced in our approach by either optimizing the error-prone-PCR further or by substituting it with a more targeted cloning method using oligonucleotide pools. Our enrichment approach could then provide the required throughput to identify many more temperature-sensitive mutants of single genes and support studies about design principles of temperature-sensitivity.

In another study on temperature-sensitivity (Chapter 5), we used a barcoded CRISPR/Cas9 method to encode mutations for single amino acid substitutions into the genome of *E. coli*. Using a tool to predict temperature-sensitive mutations (Tan et al., 2014; Varadarajan et al., 1996), we were able to reduce the design space and cover 346 essential genes. We found that only a small fraction of the predicted mutations induced temperature-sensitivity (1,045 out of 15,120). Thus, the results of our study could be used to improve the prediction tool. Alternatively, the CRISPR/Cas9 genome editing method can also be used for site saturation mutagenesis of a gene (Garst et al., 2017). In future studies, our approach to identify temperature-sensitive mutants could be combined with site saturation mutagenesis to study temperature-sensitivity in an unbiased way.

What makes microbial strains stop overproducing under growth arrest?

In Chapter 4, we observed that *argG* knockout strains are metabolically active for at least 32 hours under growth arrest and produced citrulline. However, the production rate decreased over time. As described in Chapter 5, the production rate during the two-stage

arginine overproduction using a temperature-sensitive *dnaX*(Leu289Gln) strain also declined over time, and similar effects were observed in Chapter 6 and in another study as well (Harder et al., 2018). The dynamics of the production rate declines were different in these experiments indicating a high specificity to the individual overproduction strain.

However, taken together, the production rates can initially be high under growth arrest but within one or two days *E. coli* cells stop overproducing chemicals. So far, it is unclear what the reason is for the decline of the production rates under growth arrest. Possible explanations are cell ageing, regulatory responses, product toxicity, or inhibitions due to fermentation products like acetate.

Our results in Chapter 6 and Chapter 7 showed that *enforced ATP wasting* is very effective in increasing the metabolic activity under growth arrest by nitrogen limitation. An open question is whether *enforced ATP wasting* can also reduce or even stop the decline of the production rate. By feeding glucose to the medium throughout an experiment or by reducing the initial biomass, it would be possible to test how long strains with *enforced ATP wasting* keep their production rates at a high level.

Another open question is whether *enforced ATP wasting* can improve the production rates in strains that are growth-arrested by other means than nitrogen limitation or in strains, in which the product formation is not directly linked to ATP regeneration. In future studies, the here presented two-stage overproduction strains with temperature-sensitive mutations could be coupled with *enforced ATP wasting* to study further applications of the concept.

Concluding remarks

Only with advanced bioprocesses that are economically competitive, we will achieve a transition to a sustainable bioeconomy. The field of metabolic engineering is thus focussing on developing microbial overproduction strains that enable economically viable bioprocesses. Since the production of chemicals by microbes faces a trade-off between product and biomass formation, a promising approach to create improved bioprocesses is to decouple cell growth and production.

In the here presented work, we investigated temperature-sensitivity as a tool to dynamically control cell growth and metabolism. We presented two different approaches to identify temperature-sensitive mutants in high-throughput. Analysis of the temperature-sensitive mutants revealed that many of them allowed us to dynamically control cell growth and metabolism. With some of these mutants, we created two-stage bioprocesses

and overproduced chemicals under growth arrest. However, the production strains require further optimization and engineering to reach commercialisation.

A very promising approach to improve production rates, is the *enforced ATP wasting* that we further studied in this work. In particular, *enforced ATP wasting* performed very well in increasing production rates under growth arrest. Future applications could integrate temperature-sensitive mutants and *enforced ATP wasting* to further advance two-stage bioprocesses.

Fast mass spectrometry-based metabolomics methods play a crucial role in metabolic engineering since they allow us to analyse many engineered microbial strains. FI-MS is a particularly fast method. Our findings about in-source modifications during FI-MS can be used to further improve the method. In the future, FI-MS could emerge as a crucial technology to screen large strain libraries in metabolic engineering.

References

- Tan, K. P., Khare, S., Varadarajan, R., & Madhusudhan, M. S. (2014). TSpred: A web server for the rational design of temperature-sensitive mutants. *Nucleic Acids Research*, *42*, W277–W284. <https://doi.org/10.1093/nar/gku319> (cited on pages 252, 253).
- Varadarajan, R., Nagarajaram, H. A., & Ramakrishnan, C. (1996). A procedure for the prediction of temperature-sensitive mutants of a globular protein based solely on the amino acid sequence. *Proceedings of the National Academy of Sciences*, *93*(24), 13908–13913. <https://doi.org/10.1073/pnas.93.24.13908> (cited on pages 252, 253).
- Ben-Aroya, S., Coombes, C., Kwok, T., O'Donnell, K. A., Boeke, J. D., & Hieter, P. (2008). Toward a comprehensive temperature-sensitive mutant repository of the essential genes of *Saccharomyces cerevisiae*. *Molecular Cell*, *30*(2), 248–258. <https://doi.org/10.1016/j.molcel.2008.02.021> (cited on page 253).
- Li, Z., Vizeacoumar, F. J., Bahr, S., Li, J., Warringer, J., Vizeacoumar, F. S., Min, R., VanderSluis, B., Bellay, J., DeVit, M., Fleming, J. A., Stephens, A., Haase, J., Lin, Z.-Y., Baryshnikova, A., Lu, H., Yan, Z., Jin, K., Barker, S., ... Boone, C. (2011). Systematic exploration of essential yeast gene function with temperature-sensitive mutants. *Nature Biotechnology*, *29*(4), 361–367. <https://doi.org/10.1038/nbt.1832> (cited on page 253).
- Kofoed, M., Milbury, K. L., Chiang, J. H., Sinha, S., Ben-Aroya, S., Giaever, G., Nislow, C., Hieter, P., & Stirling, P. C. (2015). An updated collection of sequence barcoded temperature-sensitive alleles of yeast essential genes. *5*(9), 1879–1887. <https://doi.org/10.1534/g3.115.019174> (cited on page 253).
- Garst, A. D., Bassalo, M. C., Pines, G., Lynch, S. A., Halweg-Edwards, A. L., Liu, R., Liang, L., Wang, Z., Zeitoun, R., Alexander, W. G., & Gill, R. T. (2017). Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nature Biotechnology*, *35*(1), 48–55. <https://doi.org/10.1038/nbt.3718> (cited on page 253).
- Harder, B.-J., Bettenbrock, K., & Klamt, S. (2018). Temperature-dependent dynamic control of the TCA cycle increases volumetric productivity of itaconic acid production by *Escherichia coli*. *Biotechnology and Bioengineering*, *115*(1), 156–164. <https://doi.org/10.1002/bit.26446> (cited on page 254).

Acknowledgements

On a Monday in October 2015, I started a three month-lasting internship in the laboratory of a young principal investigator at the Max Planck Institute in Marburg. Who would have thought that this marked the beginning of a year-long journey that included an internship, a Master Thesis, and the pursuit of a PhD? The time in Marburg and Tübingen shaped not only my professional but also my private life profoundly. Looking back the past seven years, I would do it all over again - without hesitation! I thank everyone that made these years so enjoyable.

In particular, I thank you, **Hannes!** Under your supervision, I became a scientist. Throughout the years that we worked together, I could always trust on your support. Your guidance allowed me to develop my own scientific mind without loosing focus on the essential aspects of our work.

Further, I thank **Lotte Søgaard-Andersen** and **Ilka Bischofs-Pfeifer**, who were members of my thesis advisory committee. I thank **Lennart Randau**, **Martin Thanbichler**, **Lars-Oliver Essen**, and **Victor Sourjik**, who agreed to examine this thesis.

Thank you **Ale**, **Chris**, **Chun-Ying**, **Dominik**, **Duřica**, **Felicia**, **Janhavi**, **Johanna**, **Martin**, **Matic**, **Michaela**, **Michelle**, **Peter**, **Sevalli**, **Stefano**, **Timur**, and **Vanessa!** As former and current members of the AG Link, you made and make my time in the group unforgettable and fun!

Special thanks are going to **Niklas**, **Paul**, and **Andi**. Our weekly meetings for dinner and playing cards formed friendships that I never want to miss!

I thank **Steffen Klamt**, **Simon Böcker**, and **Giulia Slaviero** for our great scientific

Chapter 7

collaboration and discussions at Tegernsee. I thank **Jörg Stülke**, **Robert Warneke**, **Janek Meißner**, and **Gottfried Uden** for our great collaborations. I thank **Timo Glatter**, **Witold Szymanski**, and **Jörg Kahnt** for support with proteomics. I thank **Janina Geißert** and the NCCT team for support with NGS.

Further, I thank **David**, **Manuela**, **Ali**, **Jens**, **Sarah**, **Melissa**, **Manuel**, **Dirk**, and **Silke** for your help with bureaucracy and IT.

Special thanks are going to my family and, especially, to my parents, **Anke** and **Thomas**, my grandparents, **Wolfgang**, **Christine**, **Uve**, and **Eva**, and my sister, **Finnja**. You enabled me to pursue my studies, and I am ever grateful for your believe in me!

I thank you, **Anjeli**, for being with me on this past journey and on many more to come!

Abgrenzung der Eigenleistung

Chapter 2 This chapter was published in: Schramm and Link, 2021, *BIOspektrum*, volume 27, 2021, pages 34–36, DOI: <https://doi.org/10.1007/s12268-021-1538-0>, (*non-peer-reviewed review*). My contribution to this review was in share with Prof. Dr.-Ing. Hannes P. Link the conceptualization, the writing, and the preparation of figures.

Chapter 3 Under revision at *ACS Analytical Chemistry*, 2022. Following the CRediT authorship contribution statements, my contributions included the conceptualization, investigation, software, visualization, writing - original draft. Andreas Verhülsdonk and I handled the metabolite standards. I prepared metabolite extract samples and performed measurements by FI-MS. I was involved in the development of the processing software of FI-MS raw data. I was involved in the preparation of figures. I co-wrote the paper. Niklas Farke and Thorben Schramm contributed equally.

Chapter 4 This chapter was published in: Schramm et al., 2020, *Metabolic Engineering*, 2020, 60, pp. 14-24, DOI: <https://doi.org/10.1016/j.ymben.2020.03.004>. Following the CRediT authorship contribution statements, my contributions included the conceptualization, investigation, visualization, project administration, writing - original draft, writing - review & editing, and formal analysis. I constructed plasmids and the mutant *argG* library, engineered *E. coli* strains, and performed growth and production experiments. Martin Lempp and I performed the bioreactor experiments together. Martin Lempp performed the off-gas analysis of the bioreactor experiments. Silvia Gonzáles Sierra performed cell sorting by flow cytometry (FACS). I performed all metabolite measurements by LC-MS/MS and by a glucose assay kit. I purified enzymes and performed activity assays. I prepared samples for proteomics. Timo Glatter performed proteomics measurements and processed the related data. I analysed the data, prepared figures, and co-wrote the paper.

Chapter 5 Manuscript in preparation. Following the CRediT authorship contribution statements, my contributions included the conceptualization, investigation, visualization, project administration, software, writing - original draft, formal analysis. I adapted the barcoded multiplex CRISPR/Cas9 system. I designed the temperature-sensitive *E. coli* strain library

Chapter 7

using my own software. I cloned and constructed the temperature-sensitive *E. coli* strain library. I performed cultivation and overproduction experiments. Vanessa Pahl constructed the *dnaX*(Leu289Gln)-based arginine overproduction strain, performed the arginine overproduction two-stage experiment, and measured the arginine levels by LC-MS/MS. I prepared samples for next generation sequencing (NGS). I measured metabolite concentrations by FI-MS and LC-MS/MS. I analysed the data, created the figures, and co-wrote the manuscript.

Chapter 6 This chapter was published in: Boecker et al., 2019, *Biotechnology Journal*, **2019**, 14, 1800438, pp. 1-9, DOI: <https://doi.org/10.1002/biot.201800438>. My contributions were metabolite measurements by LC-MS/MS and processing of related data. Related data is shown in Fig. 6.3.A.

Chapter 7 This chapter was published in: Boecker et al., 2021, *Molecular Systems Biology*, **2021**, 17:e10504, DOI: <https://doi.org/10.15252/msb.202110504>. My contributions were metabolite measurements by LC-MS/MS and processing of related data as well as sample preparation for proteomics. Related data is shown in Fig. 7.5, Fig. 7.10, and Fig. 7.11.

Folgendes Manuskript (Chapter 3) ist Bestandteil einer weiteren Qualifikationsarbeit:

Schramm, T.*, **Farke, N.***, **Verhülsdonk, A.**, **Link, H.**, A network approach identifies in-source modifications of primary metabolites during flow-injection mass spectrometry, **2022**, *under revision at ACS Analytical Chemistry* (*contributed equally)

Name des Mitverfassers: Niklas Farke

Titel der wissenschaftlichen Qualifikationsarbeit: Inference of functional metabolite-protein interactions from metabolic control analysis

Qualifikationsziel: Promotion, Doktorgrad der Naturwissenschaften (Dr. rer. nat.)

Ort, Datum

Niklas Farke
(Mitverfasser)

(weiter auf der nächsten Seite)

Ort, Datum

Thorben Schramm

Ich, Prof. Dr.-Ing. Hannes Link, bin mit der Abfassung der Dissertation als kumulative Dissertation einverstanden und bestätige die vorstehenden Angaben:

Ort, Datum

Prof. Dr.-Ing. Hannes P. Link
(Erstbetreuer)

Erklärung

Ich versichere, dass ich meine Dissertation mit dem Titel "Entkopplung von Wachstum und Überproduktion von Chemikalien in *Escherichia coli*" selbstständig ohne unerlaubte Hilfe angefertigt und 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

Thorben Schramm