Exploiting the potential of *Pseudomonas putida* as a host for engineered type I polyketide synthases

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Matthias Schmidt

Master of Science in Molekularer und Angewandter Biotechnologie

aus

Siegburg, Deutschland

Berichter:
Univ.-Prof. Dr.-Ing. Lars M. Blank
Prof. Dr. Jay D. Keasling

Tag der mündlichen Prüfung: 14. Dezember 2023

Diese Dissertation ist auf der Internetseite der Universitätsbibliothek verfügbar.

Eidesstattliche Erklärung

Hiermit erkläre ich, Matthias Schmidt, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eignen originären Forschung, generiert wurden.

Hiermit erkläre ich an Eides statt:

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Doktorand dieser Fakultät angefertigt;

2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt;

3. Wenn immer andere, eigene- oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt;

4. Wenn aus anderen, eigenen- oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate;

5. Alle wesentlichen Quellen von Unterstützung wurden benannt;

6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde;

7. Ein Teil oder Teile dieser Arbeit wurden zuvor veröffentlicht und zwar in:

   Schmidt, M.; Pearson, A. N.; Incha, M. R.; Thompson, M. G.; Baidoo, E. E. K.; Kakumanu, R.; Mukhopadhyay, A.; Shih, P. M.; Deutschbauer, A. M.; Blank, L. M.; Keasling, J. D. Nitrogen metabolism in *Pseudomonas putida*: functional analysis using

random barcode transposon sequencing. *Appl Environ Microbiol*. 2022;88(7):e0243021. doi:10.1128/aem.02430-21.

Schmidt, M.; Lee, N.; Zhan, C.; Roberts, J. B.; Nava, A. A.; Keiser, L. S.; Vilchez, A. A.; Chen, Y.; Petzold, C. J.; Haushalter, R. W.; Blank, L. M.; Keasling, J. D. Maximizing heterologous expression of engineered type I polyketide synthases: investigating codon optimization strategies. *ACS Synth Biol*. 2023;12(11):3366-3380. doi: 10.1021/acssynbio.3c00367.

Aachen, den 15.12.2023

Matthias Schmidt

Funding

This research was funded by the DOE Joint BioEnergy Institute (https://www.jbei.org) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research through contract [DE-AC02-05CH11231] between Lawrence Berkeley National Laboratory and the U.S. Department of Energy, by award EE0008926 from the BioEnergy Technologies Office, Office of Energy Efficiency and Renewable Energy, U.S. Department of Energy, and the Philomathia Foundation.

Acknowledgment

First and foremost, I want to thank Prof. Jay D. Keasling. When I first joined his lab back in 2018, I could never have imagined where this journey would take me. Thanks to an ACalNet stipend, I had the opportunity to visit the Keasling Lab in the Bay Area and was exposed to its amazing environment for science and technology. After my 6-month stay, I knew this was the place I wanted to be. Fortunately, I was invited to return in 2019 to work on my master's thesis and, again in 2020, to start my PhD.

All of these incredible experiences would not have been possible without the endless support from Prof. Lars M. Blank. I will forever be grateful for his guidance during my master's degree and now, during my PhD studies. His willingness to accept me as one of his graduate students enabled me to pursue my research abroad.

I want to express my deepest gratitude to my parents, Joanna and Karl Heinz, who have raised and supported me through all these years. You might not always understand why I am so passionate about what I do, but it makes me the happiest person I can possible be. Mama, Papa und Monika ich hoffe, ihr seid stolz auf mich. Ich hab' euch lieb.

While pursuing my dream in the Keasling Lab, something unexpected happened. I met someone truly special who joined our lab in early 2020. Although I did not realize it then, she has become the most important person in my life. I want to thank my fiancée Eunice for being by my side during both the best and most challenging years of my life. Her kindness and love motivate me every day, and I cannot wait for our next chapter in life.

Last but not least, I want to extend my thanks to all my friends and collaborators. My best friends Robin, Eschi, and Kerstin, whom I have known since early high school. My Krefeld Crew: Alex, Mo, Marc, Mike, Malte, and Waldemar. My RWTH Aachen friends Alen, and those Gladbacher Fabian and Robert. I know that I live far away and have been mainly focused on my work, but I promise to reconnect with you all and visit more frequently.

Additionally, I am grateful to the wonderful people I met at JBEI. Elias and Yuzhong, I miss you every day and the lab is not the same without you. Team Putida: Matt, Allie, and Mitch, you have been the best lab mates and mentors I could have hoped for. I also want to thank Namil, Chunjun, Jacob, Alberto, Leah, Luis, Jackie, and Satoshi for their unwavering support as co-workers and friends. A special thanks goes to my undergraduate student, Aaron, who kept my research going while I was not able to and provided significant support in the final stretch of my PhD. Seeing you grow as a young scientist was a great experience, and I am excited to see where your journey takes you next.

Table of contents

Summary

Given their logical architecture, type I polyketide synthases (T1PKSs) are a promising platform for retrobiosynthesis approaches. Entire modules or individual domains can be rationally exchanged, allowing for the customization of the synthesis of almost any desired molecule. Moreover, the immense diversity of PKSs provides an essentially unlimited supply for parts. However, engineering these complex systems presents a different reality, frequently leading to either non-functional proteins or significantly reduced titers. Developing a functional PKS design based on a target molecule requires a lot of experience and can sometimes take several years. Furthermore, the majority of PKS research focuses on a very limited number of well-studied model PKSs.

This work aims at laying out the foundation to explore and characterize novel PKSs in the microbial host *Pseudomonas putida* KT2440. Despite all the advantages of *P. putida*, it can be difficult to prevent degradation of the target molecule. Due to its diverse metabolism, we first studied this omnivorous bacterium extensively using randomly barcoded transposon insertion sequencing (RB-TnSeq). By feeding a wide range of carbon and nitrogen sources to a *P. putida* RB-TnSeq library, we were able to identify hundreds of specific phenotypes and characterized most of its metabolism. These data enabled us to better understand and control *P. putida* as a host for synthetic biology applications.

Next, we utilized a model PKS and two additional heterologous hosts to test various algorithms for codon optimization. In a systematical approach, we successfully identified the most effective strategy and were able to significantly improve protein expression levels. We also investigated the effect of different codon variants on transcript and product levels. In addition, we established a high-throughput testing approach to facilitate the *in vivo* evaluation of PKSs.

The final chapter of this work integrates RB-TnSeq host engineering with a high-throughput testing approach for the development of a δ-lactone producing PKS. We further demonstrate the effectiveness of DNA barcode sequencing (BarSeq) in guiding metabolic engineering efforts to enhance polyketide titers. Moreover, we show that the simplest PKS design, with minimal modifications, proves to be the most promising approach.

Zusammenfassung

Angesichts ihrer rationalen Architektur sind Typ-I-Polyketid-Synthasen (T1PKSs) eine vielversprechende Plattform für Retrobiosynthese-Ansätze. Ganze Module oder einzelne Domänen können gezielt ersetzt werden, was die maßgeschneiderte Synthese fast jeder gewünschten Molekülart ermöglicht. Darüber hinaus bietet die immense Vielfalt der PKSs im Grunde einen unbegrenzten Vorrat an Bausteinen an. Die Modifikation dieser komplexen Systeme zeigt jedoch eine andere Realität, die häufig zu nicht funktionsfähigen Proteinen oder deutlich reduzierten Titern führt. Die Entwicklung eines funktionalen PKS-Designs auf Basis eines Zielmoleküls erfordert viel Erfahrung und manchmal mehrere Jahre. Außerdem konzentriert sich die meiste PKS-Forschung auf eine sehr begrenzte Anzahl von gut untersuchten Modell-PKSs.

Diese Arbeit zielt darauf ab, die Grundlagen für die Erforschung und Charakterisierung neuer PKSs im mikrobiellen Wirt *Pseudomonas putida* KT2440 zu legen. Trotz aller Vorteile von *P. putida* kann es schwierig sein, den Abbau von Zielmolekülen zu verhindern. Aufgrund seines vielfältigen Stoffwechsels haben wir dieses allesfressende Bakterium zunächst ausführlich mit der Sequenzierung von zufällig codierten Transposon Integrationen (RB-TnSeq) untersucht. Durch die Fütterung einer Vielzahl von Kohlen- und Stickstoffquellen an eine *P. putida* RB-TnSeq-Stammbibliothek konnten wir hunderte spezifische Phänotypen identifizieren und den größten Teil seines Stoffwechsels charakterisieren. Diese Daten ermöglichten es uns, *P. putida* als Wirt für synthetische Biologie-Anwendungen besser zu verstehen und zu steuern.

Anschließend nutzten wir eine Modell-PKS und zwei weitere heterologe Wirte, um verschiedene Algorithmen für die Codon-Optimierung zu testen. In einem systematischen Ansatz identifizierten wir erfolgreich die effektivste Strategie und konnten die Proteinausbeute deutlich verbessern. Zusätzlich untersuchten wir auch den Effekt von verschiedenen Codon-Varianten auf Transkript- und Produktlevels. Weiterhin etablierten wir ein Hochdurchsatz-Testverfahren, um die *in vivo* Charakterisierung von PKSs zu erleichtern.

Das abschließende Kapitel dieser Arbeit kombiniert RB-TnSeq-Stammentwicklung mit einem Hochdurchsatz-Testverfahren für die Entwicklung einer δ-Lacton-produzierenden PKS. Zudem demonstrieren wir die Effektivität von DNA Barcode Sequenzierung (BarSeq) bei der rationalen Stammentwicklung und der Steigerung von Polyketid-Titern. Darüber hinaus zeigen wir, dass das einfachste PKS-Design mit minimalen Modifikationen der vielversprechendste Ansatz ist.

List of abbreviations

1,3-DAP	1,3-diaminopropane
1,6-DAH	1,6-diaminohexane
2ABA	2-aminobutyric acid
2mbCoA	2-methylbutyryl-CoA
3ABA	3-aminobutyric acid
3AIBA	3-aminoisobutyric acid
3H24DMPA	(2S,3S)-3-hydroxy-2,4-dimethylpentanoic acid
3H2MPA	(2S,3S)-3-hydroxy-2-methylpentanoic acid
3H4MHA	3-hydroxy-4-methylhexanoic acid
3H4MPA	3-hydroxy-4-methylpentanoic acid
5AVA	5-aminovaleric acid
6AVA	6-aminocaproic acid
ACDH	acyl-CoA dehydrogenase
ACP	acyl carrier protein
ADAD	decarboxylase/agmatine deiminase
AdoCbl	adenosylcobalamin
AI	artificial intelligence
ALE	adaptive laboratory evolution
AMP	adenosine monophosphate
AnsPKS	ansalactam polyketide synthase
APT	arginine:pyruvate transaminase
AST	arginine succinyltransferase
AT	acyl transferase
BarSeq	DNA barcode sequencing
BCAA	branched-chain amino acid
BEDEX	backbone excision-dependent expression
BGC	biosynthetic gene cluster
BHET	bis(2-hydroxyethyl) terephthalate
BHI	brain heart infusion
BKD	branched-chain α-keto acid dehydrogenase
BorPKS	borrelidin polyketide synthase
BpsA	blue-pigment synthase A

Cbl	cobalamin
cDNA	complementary DNA
CF	cystic fibrosis
CLI	command-line interface
CoCoPUTs	Codon and Codon Pair Usage Tables
COG	Cluster of Orthologous Group
COMM	communication domain
cPCR	colony polymerase chain reaction
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DAP-Seq	DNA Affinity Purification and sequencing
DBTL	design, build, test, learn
DH	dehydratase
DIA	data-independent acquisition
DMC	dynamic metabolic control
DMSO	dimethyl sulfoxide
DSB	double-strand break
EAL	ethanolamine ammonia-lyase
EPI	methylmalonyl-CoA epimerase
ER	enoyl reductase
EryPKS	erythromycin polyketide synthase
ESI	electrospray ionization
FDR	false discovery rate
FRT	Flippase Recombination Target
GABA	γ-amino butyric acid
GC	guanine-cytosine
GFP	green fluorescent protein
GOGAT	glutamine oxoglutarate aminotransferase
GRAS	generally recognized as safe
GUI	graphical user interface
HR	homologous recombination
hrca	harmonize relative codon adaptiveness
Hsp	heat-shock protein
HV1	host-vector system safety level 1
ibCoA	isobutyryl-CoA

ibmCoA	isobutyrylmalonyl-CoA
IPTG	isopropyl β-D-1-thiogalactopyranoside
KR	keto reductase
KS	keto synthase
LB	lysogeny broth
LC	liquid chromatography
LipPKS	lipomycin polyketide synthase
M	module
MCM	methylmalonyl-CoA mutase
mCoA	malonyl-CoA
mcu	match codon usage
ML	machine learning
mmCoA	methylmalonyl-CoA
MOPS	3-(N-morpholino)-propanesulfonic acid
MS	mass spectrometry
NGS	next-generation sequencing
NHEJ	nonhomologous end joining
NRPS	non-ribosomal peptide synthase
OD	optical density
OHCU	2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline
PACE	proteobacterial antimicrobial compound efflux
PCA	Principal Component Analysis
PCCase	propionyl-CoA carboxylase
PCP	peptide carrier protein
PCR	polymerase chain reaction
PES	polyethersulfone
PET	poly(ethylene terephthalate)
PHA	polyhydroxyalkanoate
PIA	Pseudomonas Isolation Agar
PKS	polyketide synthase
pmCoA	phenylmalonyl-CoA
PPTase	phosphopantetheinyl transferase
RB	random barcode
RBS	ribosome binding site

RCF	Relative Codon Frequency
RevPKS	reveromycin polyketide synthase
RI	refractive index
RT-qPCR	reverse transcription quantitative polymerase chain reaction
SAGE	serine recombinase-assisted genome engineering
sgRNA	single-guide RNA
ssDNA	single-stranded DNA
SpnPKS	spinosyn polyketide synthase
t-SNE	t-distributed stochastic neighbor embedding
TA	toxin-antitoxin
TCA	tricarboxylic acid
TCEP	2-(carboxyethyl)phosphine
TE	thioesterase
TIS	transposon insertion sequencing
TPA	terephthalic acid
TRY	titer, rate, yield
ubc	use best codon
UV	ultraviolet
WT	wild-type
βKA	β-ketoadipate

List of Figures

List of Tables

Chapter 1

General Introduction

Contributions

This chapter was written by Matthias Schmidt and reviewed by Lars M. Blank and Jay D. Keasling.

1 General introduction

1.1 The genus *Pseudomonas*

Pseudomonads are a type of bacteria characterized by their rod-shaped morphology, presence of a polar flagellum, and a Gram-negative cell wall structure. They exhibit a remarkable metabolic diversity, enabling them to occupy various ecological niches. The *Pseudomonas* genus, which includes 313 identified species, is particularly notable, with *P. aeruginosa*, *P. putida* and *P. syringae* being among its most prominent representatives.[1] Like many other bacterial genera, *Pseudomonas* has undergone several rounds of reclassification over time. With the advent of genome sequencing techniques, extensive phylogenomic studies have been conducted, uncovering misannotations and leading to a revised taxonomy.[2,3] One significant example of this reclassification is the removal of strains that are now classified under the genera *Burkholderia* and *Ralstonia*.[4,5] Distinct from traditional phenotypic or genomic taxonomic methods, Meyer *et al.* introduced an alternative approach to characterizing pseudomonads.[6] They focused on the yellow-green fluorescent pigment called pyoverdine, which is a well-known characteristic of most *Pseudomonas* species.[7] Pyoverdine belongs to a group of molecules known as siderophores, which are specifically designed to scavenge traces of iron(III) during periods of iron-starvation.[8] Depending on their microbial host, these siderophore molecules have significant differences in their chemical structure and are easily distinguishable. As a result, siderotyping has emerged as a valuable method for efficiently grouping 344 strains with reduced time and cost compared to traditional taxonomic approaches.[6]

Undoubtedly, the most extensively studied species among pseudomonads is the opportunistic pathogen *P. aeruginosa*. While it can be found in diverse environments such as soil and water, it predominantly thrives in human-impacted settings, with hospitals posing a major concern.[9] *P. aeruginosa* is labeled as opportunistic because it typically does not affect healthy individuals but rather causes illness and mortality in immunocompromised patients. In the context of human health, it poses particular challenges for individuals with cystic fibrosis (CF). A significant aspect of the problem is *P. aeruginosa*'s extensive antimicrobial resistance, and despite aggressive antibiotic treatments, infections often lead to a decline in lung function or even death.[10] Additionally, its capability to grow across a broad range of temperatures, oxygen concentrations, and carbon sources contributes to the difficulty in treating *P. aeruginosa* infections.[9]

When it comes to industrial relevance, the species *P. putida* has gained a lot of research attention. This can be attributed to its versatile carbon and nitrogen metabolism, its genetic modifiability, and its remarkable tolerance towards organic solvents.[11–13] While several publications have incorrectly portrayed *P. putida* KT2440 as generally recognized as safe (GRAS), it is actually certified as host-vector system safety level 1 (HV1).[14] Nevertheless, extensive research spanning many decades has consistently demonstrated that working with *P. putida* remains highly safe.

Despite its reputation as an industrial workhorse, the utilization of *P. putida* in the biotechnological industry remains relatively limited.[15] The primary industrial application where *P. putida* shows relevance is in the production of polyhydroxyalkanoates (PHAs).[16] PHAs are naturally synthesized by most pseudomonads and serve as carbon and energy storage compounds during carbon excess and nutrient limitation.[17] PHAs offer a promising alternative to petroleum-based plastics, as they exhibit high biodegradability and exert a significantly lower environmental impact.[18] To enhance its industrial viability, *P. putida* can effectively utilize inexpensive renewable feedstocks.[15] In addition to its ability to consume lignin-derived monomers, a recent trend involves the valorization of plastic monomers derived from poly(ethylene terephthalate) (PET).[19,20] The process begins with enzymatic or chemocatalytic depolymerization of PET, resulting in the formation of bis(2-hydroxyethyl) terephthalate (BHET) monomers. These BHET monomers can then be supplied to engineered strains of *P. putida*.[20] Through the expression of specific heterologous genes responsible for BHEP hydrolysis and terephthalic acid (TPA) degradation, these engineered strains are not only capable of utilizing BHET but also producing β-ketoadipate (βKA). Another successful example involves the engineered degradation pathway for the plastic monomer adipic acid.[21] By combining heterologous genes and employing adaptive laboratory evolution (ALE), the final strain of *P. putida* was able to utilize adipic acid as a carbon source and overproduce PHA.

In summary, these studies demonstrate the great potential of harnessing the strengths of *P. putida* to achieve industrial viability in various processes. Alongside the production of bulk chemicals, scientific literature provides numerous examples of successfully producing a diverse array of molecules in this organism such as amino acid derivatives, polyketides, rhamnolipids, and terpenoids.[22] Recently, there has been academic interest in another *Pseudomonas* strain called *P. taiwanensis* VLB120, which exhibits similar industrial potential.[23–25] In comparison to *P. putida*, this particular strain demonstrates exceptional solvent tolerance, has the ability to degrade styrene, and can utilize xylose as a growth substrate.[26]

P. syringae, another representative of the *Pseudomonas* genus, has garnered significant attention in scientific studies primarily due to its pathogenic effects on plants. As one of the most prevalent plant pathogens, it serves as a valuable model organism for exploring interactions between plants and microorganisms, as well as for investigating microbial ecology.[27] Aside from its epiphytic characteristics that enable it to thrive in challenging environments such as exposure to ultraviolet light and dry conditions, *P. syringae* possesses a unique ability to initiate the formation of ice crystals through a specific ice nucleation protein called INA.[28] This capacity for ice nucleation likely evolved to induce frost injuries in plants and create openings on the plant surface.[29] Additionally, some researchers have proposed that *P. syringae* also plays a significant role in the process of rainfall and snowfall formation, thus exerting an influence on global climate patterns.[30]

The genus *Pseudomonas* displays a remarkable metabolic diversity and participates in numerous domains of human health, crop science, and biotechnology. In order to fully comprehend the diversity of pseudomonads and harness their genetic potential, the field of functional genomics has emerged as a critical discipline within molecular biology. As an integral component of genome-wide studies, functional genomics techniques such as transposon insertion sequencing (TIS) have proven to be an indispensable tool for investigating genetic interactions and elucidating the functions of genes.[31]

1.2 Transposon insertion sequencing

The conventional method for studying genetic phenotypes involves creating a single genetic mutation and examining the resulting phenotype under specific conditions. However, a more recent approach involves characterizing mutants within the context of large gene deletion libraries that encompass the entire genome of a microorganism.[32] This recent advancement has been facilitated by the significant progress in next-generation sequencing (NGS) platforms, leading to a substantial reduction in the costs associated with library sequencing.[33,34]

To create genome-wide mutant libraries, researchers systematically generated thousands of single gene deletion strains by individually deleting each non-essential gene.[35,36] These approaches are extremely laborious and time-consuming, and the targeted organism is limited by the availability and efficiency of genetic tools. Furthermore, prior to the advent of NGS, the subsequent phenotype analysis relied on microarrays which assess gene fitness via a fluorescence output. Microarrays not only need customized development for each specific organism but can also be inaccurate and costly.[37,38]

Advancements in genome-wide studies of gene interactions and functions have been characterized by two key elements. The first step involved the application of transposons like Tn5 or *mariner*.[39] Creating a mutant library with randomly inserting transposons shortens the process of mutant library generation significantly.[40] It also requires less expertise regarding the target organism since a simple transformation or conjugation is sufficient.[39] The next step involved the assignment of a random genetic mutation to a specific phenotype and to generate a measurable output. This was accomplished by directly sequencing the entire mutant population, mapping insertion sites, and quantifying the members of the library.

During the late 2000s, van Opijnen *et al*. introduced a method called TnSeq, which outlined the generation of transposon mutant libraries and direct library sequencing.[37] Simultaneously, in that same year, three other methods with a comparable workflow were developed.[41–43] The main distinction among these methods lies in the choice of transposon utilized and the approach to DNA fragmentation.[39] In general, by cultivating these mutant pools under selective conditions, it is possible to predict the essentiality of genes and obtain empirical evidence for their function. For instance, if the number of a specific mutant decreases significantly during the assayed condition, the interrupted gene was most likely essential. It is also possible to detect the increase of a library member, which reveals phenotypes with improved growth. A common example for improved growth phenotypes is the interruption of negative regulators or metabolically demanding gene products.[39] To accurately map the insertion sites and quantify the mutant population, it is necessary to isolate and sequence the transposon locus. In the case of TnSeq, the transposon insertion cassette is flanked by MmeI restriction sites.[37] The type II restriction enzyme MmeI creates unique overhangs which correspond to genomic DNA 20 bp downstream from the recognition site.[44] After DNA extraction and digestion, the resulting gene fragment is isolated and processed for Illumina sequencing by addition of Illumina-specific adaptors and PCR amplification. The resulting sequencing reads contain the 20 bp organism-specific overhang sequence and are quantitatively mapped to the host genome.[37] While transposon mutagenesis offers numerous advantages, it is important to acknowledge that it does not result in a complete in-frame deletion of the entire gene sequence. An issue frequently encountered with transposon mutagenesis is the possibility of incomplete disruption of a gene.[45] For example, if a transposon inserts near the 3' end of a genetic sequence, it may result in false negative phenotypes due to the residual functionality of the gene product. However, follow-up studies have revealed that these positional effects are negligible. The highly efficient process of transposition often leads to the generation of multiple insertion mutants for a given gene and fitness defects can be averaged across multiple

insertion sites. Excluding insertion sites close to the 5' and 3'end of mutated genes showed no significant difference in the resulting fitness values.[37]

Although TnSeq can be scaled, the extensive work and costs associated with sample preparation and repeated insertion mapping make it highly challenging to test multiple conditions simultaneously.[39] To increase sample throughput and cost efficiency, the utilization of genetic barcodes has emerged as a viable solution.[36,46] Based on barcoded deletion libraries, Smith *et al.* developed a new method for barcode analysis through sequencing known as BarSeq.[36,38] Subsequent advancements in BarSeq significantly improved the sample preparation process, enabling the simultaneous investigation of up to 30 different conditions in a single Illumina sequencing lane.[47] Despite the advantages offered by BarSeq, generating barcoded mutant libraries remained a labor-intensive process, requiring the performance of thousands of individual gene deletions.[36]

The most recent approach in the field of TIS is the combination of random barcodes and transposon-site sequencing (RB-TnSeq).[48] This innovative technique merges the generation of transposon mutant libraries from TnSeq with the use of uniquely barcoded single mutants from BarSeq (Figure 1). As a result, RB-TnSeq streamlines library construction and enhances assay throughput, all while keeping costs relatively low. Rather than mapping insertions for each tested condition, the barcodes only need to be assigned to a gene once. Following this initial assignment, the barcoded mutant library can be employed for highly scalable and cost-effective BarSeq analysis in a massively parallel manner.[48] Furthermore, by using millions of random barcodes instead of a limited set of sequence-defined barcodes, transposon mutant libraries are effectively labeled, ensuring comprehensive coverage of non-essential genes.[49] Moreover, the average number of uniquely barcoded strains per gene is approximately 20 which allows to exclude biased insertions as well.[48] A comparison of RB-TnSeq data to other technologies such as TnSeq and high-throughput imaging also showed that the identified phenotypes are highly correlated.[48,50]

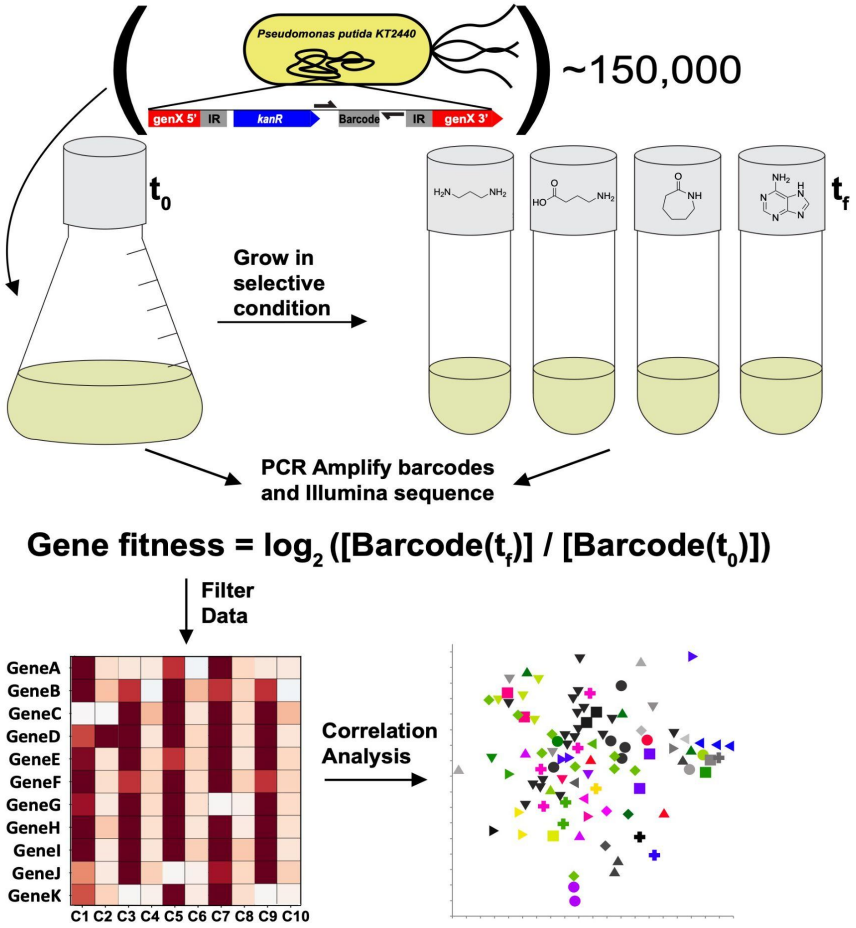


Figure 1: The general workflow of a BarSeq assay based on an RB-TnSeq mutant library. In this approach, uniquely barcoded transposon insertions are used to create a comprehensive library of mutants.[11] Under selective growth conditions, the abundance of certain mutants increases or decreases, which is quantified by the gene fitness value. By filtering the data and performing a correlation analysis, significant fitness phenotypes are identified.

In the past eight years, numerous research studies employing this method have been published, exploring a wide range of microorganisms.[51–54] One notable milestone was a study that conducted a simultaneous analysis of 32 bacteria, leading to the discovery of 12,000 genes with previously unknown functions.[55] Determining the function of a gene holds significance not only in the field of fundamental genetic sciences but also in the field of metabolic engineering. As a widely studied heterologous host, *P. putida* has also accumulated a substantial amount of RB-TnSeq data.[11,12,19] While subject to controversy, studies on RB-TnSeq guided metabolic engineering have demonstrated the utility of such data.[19,56,57] For instance, the efficient production of valerolactam required the identification of the lysine degradation pathway and the previously unknown lactamase OplBA.[56] A singular BarSeq analysis with lysine as the sole source of carbon successfully revealed all the involved genes, verifying

decades of research and even identified previously unknown genes.[52] Using this acquired knowledge, valerolactam titers were enhanced from being undetectable to approximately 90 mg/L.[56] In addition, not only was the lactamase OplBA identified, but also its regulator *oplR* was discovered. In a subsequent study, *oplR* was further developed into a remarkably sensitive lactam biosensor.[58] This serves as an excellent illustration of *P. putida*'s potential as not just a heterologous host but also as a genetic resource for synthetic biology tools. The degradation of lignin-derived compounds is another notable achievement in RB-TnSeq guided metabolic engineering.[19,57] In these studies, the gathered data was leveraged to enhance the production of bisdemethoxycurcumin and improve *P. putida*'s tolerance towards aromatic compounds.[19,57]

One inherent constraint of TnSeq-based techniques is their incapability to generate phenotypes for essential genes. However, this constraint can be overcome by employing Clustered Regularly Interspaced Short Palindromic Repeats interference (CRISPRi).[59] Unlike complete gene disruption, CRISPRi-based techniques modulate gene expression levels, enabling the identification of fitness phenotypes that would otherwise remain undetected using traditional methods.[59]

CRISPRi utilizes a catalytically inactive form of the Cas9 protein known as dCas9, which is navigated to the target DNA strand by a single-guide RNA (sgRNA).[60,61] When bound to the target gene, the dCas9 protein obstructs the RNA polymerase, effectively preventing transcription.[60,61] Numerous studies have used CRISPRi to investigate the impact of gene knockdowns in an arrayed manner. However, individually characterizing these mutants can be time-consuming and labor-intensive. To address this, similar to transposon mutant libraries, pooling the mutants and combining them with NGS enables genome-wide gene fitness studies (CRISPRi-seq).[62,63] Furthermore, the design of sgRNAs can be completely automated, or alternatively, manually restricted to specifically target essential genes only.[63] By restricting sgRNA targets, CRISPRi-seq could be used to complement existing TnSeq datasets, resulting in a comprehensive coverage of the entire genome. Another notable advantage is the ease of transferring CRISPR-seq to new organisms or their derivatives.[63] In contrast, TnSeq libraries need to be completely reconstructed when tested in mutant backgrounds, whereas CRISPRi-seq only requires a simple transformation of the sgRNA library.[37,63] However, a significant drawback of this method is that CRISPRi-based gene knockdowns have an impact on downstream genes.[60,61] When the dCas9 protein binds to an operon, it can interfere with the expression of all the remaining genes within that operon. The analysis of gene fitness is therefore limited to the operon level.[63]

Present investigations utilizing CRISPRi-seq predominantly focus on essential genes involved in cell morphology or antibiotic resistance.[64,65] As per definition, targeting essential genes is a major challenge, especially in the context of metabolic engineering. While CRISPRi is indeed a useful tool to target these genes, researchers are continuously working on advancing and refining other techniques to rationally engineer the metabolism of cells.

1.3 Metabolic engineering

Microbes and their metabolic capabilities have been leveraged by humans for thousands of years. However, the deliberate alteration of microbial metabolism is a relatively recent concept. Before the era of "metabolic engineering", approaches were primarily non-targeted, and microbes were subject to random mutagenesis.[66] Consequently, strains had to undergo screening for the overproduction of the desired metabolite all while lacking any knowledge regarding the underlying genetic modifications.[67] Furthermore, these random mutagenesis approaches often caused unintended alterations of the target metabolism.[68] As a result, there has been a paradigm shift towards adopting a more rational and deliberate approach.[69]

The first major breakthrough was the *in vitro* construction of plasmid DNA, which was soon after followed by the heterologous production of ethanol and insulin.[70–73] However, the beginning of metabolic engineering was marked by its definition in the 1990s as "the improvement of cellular activities by manipulation of enzymatic, transport, and regulatory functions of the cell with the use of recombinant DNA technology".[69] Today, metabolic engineering has evolved into a highly interdisciplinary field that closely collaborates with functional genomics, proteomics, metabolomics, transcriptomics, and synthetic biology.[73]

Metabolic engineering relies on a fundamental concept known as the design, build, test, and learn (DBTL) cycle.[73] The goal of this iterative process is to enhance the production of a desired product by optimizing titers, rates, and yields (TRYs).[73] Although TRYs may hold less importance for high-value molecules, it becomes critical for bio-synthesizing bulk chemicals. To date, only a few examples exist where a petrochemical route has been successfully replaced by a biological alternative.[74]

The DBTL cycle starts with the design of a potential biosynthetic route that leads to the desired product. The pathway design is closely tied to the chosen host and can be a combination of native and non-native reactions.[75] While *Escherichia coli* and *Saccharomyces cerevisiae* are widely used heterologous hosts, they are not necessarily the most suited expression systems.[76] The design process can be significantly accelerated if the host already produces high amounts of the target molecule.[77] Moreover, depending on the availability of synthetic

biology tools for a particular organism, subsequent steps might become extremely challenging. Microbes like *E. coli* and *C. glutamicum* are highly domesticated, and the wealth of relevant literature makes them the preferred procaryotic expression system.[76] However, the rapid development of new gene editing tools and functional genomics techniques facilitates the process of domesticating novel host with superior intrinsic characteristics.[48,78] Regarding industrial applicability, *Halomonas* spp. serves as an excellent example. These halophilic bacteria live in high salt environments and fermentation processes under these conditions are extremely resistant to contaminations.[79] In addition, they naturally accumulate PHAs, making them an ideal host for the industrial production of bioplastics.[80,81] If the chosen host does not provide the required biochemical pathways, missing reactions can be added by the expression of heterologous genes.

In the build stage, synthetic biology techniques, such as genome editing, are employed to introduce non-native enzymes or redirect existing metabolic pathways towards the desired product.[82] These decisions are often guided by experience or supported by computational models.[73,83,84] The traditional method of genome engineering relies on the organism's homologous recombination (HR) system.[73] By introducing homologous DNA into the target host, the native HR system can be exploited to integrate or delete genes from the genome.[85] The efficiency of this process largely depends on the host's HR machinery. While HR is remarkably efficient in some organisms, others require the use of heterologous recombination systems, such as bacteriophage lambda Red.[86,87] Nowadays, HR is primarily used for the deletion of genes or integration of small (< 2 kb) expression cassettes. Since the stable integration of a pathway is typically preferred, heterologous integration systems are the only option.[73] A commonly used tool for the integration of large (> 100 kb) biosynthetic gene clusters (BGCs) into *Streptomyces* are serine recombinases.[88,89] Due to the natural presence of attachment sites (*att*) in the genomes of streptomycetes, these enzymes are readily available for use with these organisms.[90] However, serine recombinases can also be employed in proteobacterial hosts by integrating *att* sites into their genomes.[91] Recent advancements in this field, led to the development of a serine recombinase-assisted genome engineering (SAGE) toolkit.[92] The SAGE system allows for high-throughput and site-specific gene integrations into a wide range of microbial hosts.[92] The basic requirements to utilize SAGE are a strain that contains an *attB* site, a vector containing the corresponding *attP* sequence and the serine recombinase itself. The serine recombinase catalyzes the unidirectional recombination between genomic *attB* and vector *attP* site, and creates two new *att* sites, *attL* and *attR*. Although this process is reversible, it would require another set of heterologous proteins and can, therefore, not be catalyzed

by the recombinase alone.[92] A remarkable feature of the SAGE system is the possibility to recycle the selection marker, enabling the iterative introduction of up to 10 heterologous genes. The SAGE integration vectors contain the integrase ϕC31 *attB* and *attP* site in their backbones. Upon transient expression of ϕC31, the recombination of the *att* sites leads to the excision of the selection marker. Subsequently, the markerless strain regains sensitive to the chosen antibiotic, allowing the iterative use of the same antibiotic marker. Another strategy to remove the selection marker is by flanking the marker gene with Flippase Recombination Target (FRT) sites.[93] These sites are recognized by the tyrosine recombinase Flp.[94] However, FRT sites are homologous to each other, which can lead to unintended recombination events.[92]

One of the most prominent examples of recent developments in genome engineering is the RNA-guided CRISPR-Cas system.[95] The discovery of CRISPR-Cas can be exploited to precisely introduce double-strand breaks (DSBs) into DNA.[96] If employed in a microbial host, DSBs of genomic DNA are lethal, unless the DNA is repaired by HR or nonhomologous end joining (NHEJ).[73] By providing a repair template in the form of double-stranded DNA (dsDNA) or single-stranded DNA (ssDNA), CRISPR-Cas enables the markerless modification of the genome with a theoretical editing efficiency of 100%.[96,97] Like many synthetic biology tools before, CRISPR-Cas genome editing systems have been mainly developed for *E. coli* and *S. cerevisiae*.[98,99] Due to the number of variables (i.e., Cas protein, sgRNA, and HR template), CRISPR-Cas genome engineering is difficult to adapt in non-model organism and usually requires an initial optimization study.[73,100,101] Other limitations, including off-target double-strand breaks (DSBs) and the lack of protospacer adjacent motif (PAM) sequences, further impede the widespread use of CRISPR-Cas gene editing techniques.[99]

While gene editing is a crucial approach to modifying a cell's metabolic flux, it does not always represent the most effective strategy. Regulating the expression of genes is another vital part of metabolic engineering. A common strategy to regulate the expression of a gene involves the engineering of the promoter or ribosomal binding site (RBS) sequence.[73] However, more advanced techniques such as CRISPRi, RNA interference (RNAi), or small stable RNA A-tagging (SsrA) have also been employed.[59,102,103] Furthermore, these techniques enable dynamic control over a cell's metabolism. Contrary to static control systems, dynamic metabolic control (DMC) can be leveraged to adjust the production flux to more advantageous levels during various stages of growth.[104] Although autonomous regulation represents the ideal scenario, inducible promoter systems are a more commonly used tool for DMC. By combing an inducible system with CRISPRi or RNAi, it is possible to enhance the flux

towards central metabolites like malonyl-CoA (mCoA).[103,105] To increase the intracellular concentration of mCoA, it is necessary to target essential genes involved in fatty acid biosynthesis.[106] Deleting these genes or interfering with their functionality in an early growth stage would be lethal for the host organism. Therefore, the optimal strategy is to carefully downregulate the expression or completely inhibit it at a later phase of growth. A rarely employed technique in this context is SsrA-tagging.[102,107,108] SsrA is a naturally occurring RNA molecule that co-translationally appends a C-terminal peptide tag to proteins whose synthesis process has been halted or disrupted.[109] Once a partially synthesized protein is tagged, protease complexes like ClpXP bind to a specific portion of the tag, triggering the protein's rapid degradation.[110] Similarly, when this tag is artificially attached to a target protein, it also induces immediate degradation. By modifying the ClpX binding motif, the degradation efficiency can be regulated based on the presence of the adaptor protein SspB.[102] While this adjustable SspB-dependency has only been demonstrated in *E. coli*, SsrA-tags with reduced degradation efficiency have been successfully implemented to lower the quantity of essential gene products in *P. putida*.[108] Compared to CRISPRi or RNAi, the use of proteolytic degradation tags requires less optimization work as SsrA-tags are notably similar across various bacterial species.[111] The only requirement would be a reliable gene editing method. Moreover, current transcription interference techniques require the maintenance of plasmids which are a known source of metabolic stress.[112]

Previously described methods for metabolic engineering primarily follow a rational design approach. However, in certain scenarios, harnessing the power of natural selection through artificial evolution presents a viable option for optimizing biosynthetic pathways. An emerging trend to improve the metabolic capabilities of a strain or explore specific biological phenomena is ALE.[113] At its most basic level, an ALE experiment involves cultivating a strain under a specific condition and subsequently selecting the mutants that have effectively adapted.[113] Furthermore, the experimental design is typically very minimalistic, generally involving manual propagation of batch cultures, with static transfer volumes and intervals.[114] However, the complexity of an ALE experiment can be significantly increased by dynamically adjusting culture conditions and transfer protocols. To mitigate human error and reduce fatigue, automated platforms have been developed specifically for managing these complex ALE experiments.[115]

Two common applications of ALE include enhancing substrate utilization efficiency and increasing tolerance against inhibitory chemicals.[113] In a study aimed at enabling *P. putida* to metabolize C6 dicarboxylic acids, an additional set of genes was introduced into its

metabolism.[21] Although the engineered strain could already utilize adipic acid as the sole source of carbon, its growth rate remained substantially low. The application of ALE significantly increased growth, while whole-genome sequencing provided insights into the mutations potentially responsible for the improvement. By reverse engineering the initial strain, the identified mutations could be verified.[21] Additionally, this study presents a valuable example of a DBTL cycle with iterative improvements towards the final goal. While the learning process of smaller projects relies on manual data analysis, machine learning (ML) can facilitate the analysis of massive multi-dimensional datasets.[116] However, the potential of ML in the context of metabolic engineering remains largely unexplored and underutilized.[73]

An emerging trend for the design of metabolic pathways are rule-based or ML-based retrobiosynthesis tools.[73] One such tool is ClusterCAD, which combines the power of retrobiosynthesis with the modularity of an enzyme class known as polyketide synthases (PKSs).[117]

1.4 Synthetic biology of polyketide synthases

PKSs are large, multi-domain enzymes that can perform a series of reactions while also catalyzing carbon-carbon condensations. They are primarily recognized for their role in antibiotic production but also contribute to the creation of other essential drugs such as immunosuppressants.[118]

PKSs come in three varieties - type I, type II, and type III. Among these, the non-iterative type I PKS (T1PKS) holds the most promise for synthetic biology applications. A typical T1PKS is comprised of multiple domains, which are part of multiple modules, or even multiple peptides. However, a more streamlined, synthetic T1PKS design could be comprised of only two modules containing a total of six domains (Figure 2). The loading module, for instance, would only need an acyl transferase (AT) domain and an acyl carrier protein (ACP) domain. The AT domain recruits an acyl-CoA and catalyzes the transfer of the acyl chain onto the ACP. With such a loading module, short fatty acyl-CoAs are amongst the most frequent substrates.[119] The subsequent domain would be part of the extension module, starting with a keto synthase (KS) domain. The remaining domains of the extension module would be an AT, an ACP and a thioesterase (TE) domain. The AT within an extension module exclusively recruits mCoA or its derivatives. Once the AT transferred the extender unit to its ACP, the KS domain condenses the now KS-bound starter unit to the ACP-bound extender unit. This reaction leads to the release of carbon dioxide and extends the polyketide chain by two carbons. The final step involves the hydrolysis of the polyketide by the TE domain. For instance, if the loading substrate of the here described T1PKS would be propionyl-CoA and the

elongation step uses mCoA, the final product would be 3-oxopentanoic acid. With a few modifications of the PKS like the incorporation of a keto reductase (KR) domain, the product can be easily diversified. A very similar PKS has been constructed by truncating the lipomycin PKS (LipPKS) from *Streptomyces aureofaciens* Tü 117.[119]

The LipPKS is a non-iterative T1PKS consisting of seven modules, which are divided into four individual peptides.[120] It is not unusual that PKSs are divided into multiple proteins. When assembling larger polyketides, the utilization of N- or C-terminal communication (COMM) domains allows for the expression of smaller, separate proteins. Depending on the recruited starter unit, the resulting fatty acid molecule is 17 carbons long. In the field of PKS engineering, the truncation of a PKS is likely the simplest form of modifying the resulting polyketide. In the case of the LipPKS, the TE domain of the erythromycin PKS (EryPKS) was fused to the C-terminal part of module (M) 1. This modification leads to the premature release of the polyketide following the first extension.[119]

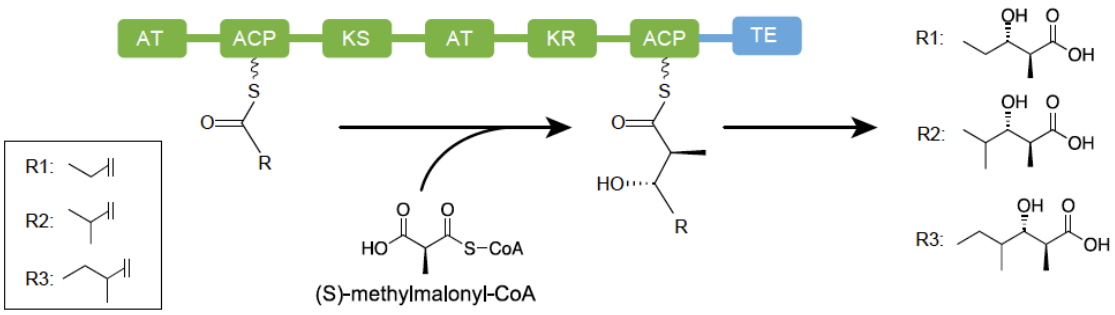


Figure 2: Engineered lipomycin polyketide synthase for the production of short-chain 3-hydroxy acids. The loading substrate is extended with methylmalonyl-CoA and the thioesterase (TE) domain releases the acyl carrier protein (ACP) bound product.[121] R1: propionyl-CoA; R2: isobutyryl-CoA; R3: 2-methylbutyryl-CoA; AT: acyl transferase; KS: keto synthase; KR: keto reductase

In vitro experiments demonstrate the considerable promiscuity of the loading didomain of LipPKS with the highest catalytic efficiency achieved when loading isobutyryl-CoA (ibCoA), 2-methylbutyryl-CoA (2mbCoA) or propionyl-CoA.[119] Therefore, depending on the supplied acyl-CoA, the engineered LipPKS has the capability to produce at least three different short-chain fatty acids. The AT domain of PKSs provides another way to increase product diversity. For LipPKS, the AT of the first extension module exhibited activity only when provided with methylmalonyl-CoA (mmCoA).[119] Other ATs, such as the reveromycin PKS (RevPKS) M4 AT can accept a broad spectrum of mCoA derivatives.[122] In a follow-up study of the truncated LipPKS, its native AT was replaced with the mCoA specific AT from the borrelidin PKS (BorPKS).[123] This modification resulted in the anticipated removal of the methyl group on the α-carbon by incorporating mCoA instead of mmCoA.

Exchanging the AT domain of a PKS is another common engineering strategy to control polyketide chemistry. In a recent study, 12 ATs were swapped into a truncated EryPKS M6 and tested *in vitro* against 13 different mCoA derivatives.[124] Surprisingly, the predicted and *in vivo* observed high specificity of the wild-type (WT) AT could not be confirmed.[125] Instead, the native AT of M6 appeared to be highly promiscuous and accepted side-chains with up to six carbons (hexylmalonyl-CoA). This trend was observed for nearly all of the tested ATs. For instance, the AT of the ansalactam PKS (AnsPKS) M8 is predicted to accept isobutyrylmalonyl-CoA (ibmCoA) only.[126] However, when swapped into EryPKS M6, it demonstrated a broad substrate specificity.[124] Alongside branched substrates, the final AT-swap mutant was also capable of incorporating the aromatic substrate phenylmalonyl-CoA (pmCoA). Therefore, the heterologous AT expanded the range of acceptable mCoA derivatives rather than altering the specificity from mmCoA to ibmCoA. While this finding may seem somewhat contradictory to the goal of rationally programming PKSs, it is unlikely that all these mCoA derivatives would be present in a single system.

An important aspect of domain exchanges is choosing the correct boundaries between donor AT and acceptor PKS.[123] The multidomain structure of PKSs can make it difficult to pinpoint where one domain begins and another ends. However, due to highly conserved amino acid sequences in PKSs, these boundaries can be predicted.[127] AT exchanges based on such predictions, have shown that the selected junctions can have a detrimental effect on the activity of the resulting AT-swap mutant.[123] The effectiveness of a particular junction in one PKS system does not guarantee the same result in another system.[128] Hence, the optimal positions for domain exchanges need to be determined experimentally. Due to challenges in cloning and the constraints of *in vitro* setups, these investigations typically have a low throughput.[123]

A recent study introduced a high-throughput cloning and testing method for domain exchanges.[128] The cloning process is significantly facilitated by employing oligo pools that contain all possible junction sequences and can be used as regular primers. If the donor AT sequence is PCR amplified with such an oligo pool, the resulting PCR product can be utilized to assemble a library of PKS AT-swap mutants.[128] The final library is typically very large (> 5000) and would require the screening of thousands of library members. Traditional approaches determine activity by measuring the production of the expected polyketide, a process that requires time-consuming sample preparation and analytics.[123] Englund *et al.* introduced a novel method by establishing a correlation between solubility and protein activity.[128] In addition, protein solubility was determined by using a biosensor.[128] The biosensor is based on the small heat-shock protein (Hsp) IbpA found in *E. coli*. Upon the presence of insoluble

protein, the translation of *ibpA* mRNA becomes upregulated.[129] By fusing the promoter sequence of *ibpA* to green fluorescence protein (GFP), this upregulation can be transformed into a spectrophotometric output.[128] For instance, if a selected junction leads to an inactive PKS and, therefore, insoluble protein, the cell begins to produce GFP. As a result, the entire junction library can be transformed into the engineered biosensor strain and prescreened by simply measuring *in vivo* fluorescence with a plate reader.[128] Subsequently, library members with a low and high GFP signal were selected and tested *in vitro*. Besides showing a clear correlation between activity and solubility, it was also hypothesized that poorly performing junctions are often located at structurally critical positions.[128] Nevertheless, these results are based on a limited number of PKSs. Thus, a universally applicable strategy to determine domain boundaries is yet to be established.

While exchanges involving the AT domain are amongst the most common, KR domain exchanges and even whole reduction loop swaps have also been performed.[130] The KR domain is responsible for controlling the stereochemistry of the polyketide, and can be engineered to rationally produce a desired enantiopure molecule.[131] Conversely, whole reduction loop swaps are performed to remove the functional group on the β-carbon, thereby yielding aliphatic fatty acid chains.[132] Given the hierarchical order of domains and reactions, PKSs hold an enormous potential for rationally designing their function. Furthermore, PKSs and therefore "parts" are found across nearly all domains of life, exponentially increasing the number of potential products.

1.5 Scope of this study

P. putida has demonstrated to be a robust host for numerous synthetic biology applications, including the expression of PKSs.[22] However, the promise of PKSs, particularly engineered T1PKSs, has yet to be fully explored in this host.

The aim of this study was to examine and enhance the potential of *P. putida* KT2440 as a heterologous host for engineered T1PKSs. Although PKSs are well-established within their native hosts, their expression and productivity in heterologous hosts presents ongoing challenges. To enhance the industrial viability of engineered T1PKSs, the selection of an alternative host with established industrial relevance may be the key to achieving profitable production levels.

Chapter 1 provides a comprehensive overview of the relevant literature, discussing the history and current status of research related to this study. While the focus is primarily on *P. putida*, it is crucial to comprehend the genetic diversity of pseudomonads and how it can be leveraged. The foundation for this is established by the use of TIS techniques. At the forefront of these techniques is RB-TnSeq, which has been instrumental in making informed choices for engineering *P. putida*'s metabolism. Furthermore, the chapter also highlights recent developments in metabolic engineering that form the theoretical foundation for the methods employed in this study. The final segment of this chapter introduces the basics of engineering PKSs.

Chapter 2 describes the used methods and materials for the research conducted. This also includes a list of all the used primers, plasmids, chemicals and generated strains.

Chapter 3.1 presents an RB-TnSeq study about the nitrogen metabolism in *P. putida,* demonstrating the power of BarSeq in this organism. We simultaneously analyzed the metabolism of 52 distinct nitrogen sources, identifying 672 fitness phenotypes than span various parts of metabolisms and also discovering unknown gene functions. Moreover, we validate the reliability of our BarSeq data and characterize some of the identified aminotransferases *in vitro*.

Chapter 3.2 addresses a common issue in synthetic biology: codon optimization. We implement three distinct codon optimization algorithms to assess their impact on the transcript and protein levels of an engineered LipPKS. The most effective algorithm is identified, and we demonstrate increased polyketide titers in three heterologous hosts.

Chapter 3.3 highlights the utility of RB-TnSeq within the field of metabolic engineering. In this section, we examine *P. putida*'s overall suitability as an alternate host for T1PKSs. We also employ RB-TnSeq-guided modifications of *P. putida*'s intracellular CoA pool to elevate the production of polyketide precursors. This engineering approach is then validated by

expressing a modified version of LipPKS and demonstrating the increased synthesis of the anticipated polyketide.

The final chapter of this study is a comprehensive discussion on non-native PKS hosts and their potential contribution towards achieving an industrially relevant engineered T1PKS process.

Chapter 2

Materials and Methods

Partially published as:
Schmidt, M.; Pearson, A. N.; Incha, M. R.; Thompson, M. G.; Baidoo, E. E. K.; Kakumanu, R.; Mukhopadhyay, A.; Shih, P. M.; Deutschbauer, A. M.; Blank, L. M.; Keasling, J. D. Nitrogen metabolism in *Pseudomonas putida*: functional analysis using random barcode transposon sequencing. *Appl Environ Microbiol*. 2022;88(7):e0243021. doi:10.1128/aem.02430-21.

Schmidt, M.; Lee, N.; Zhan, C.; Roberts, J. B.; Nava, A. A.; Keiser, L. S.; Vilchez, A. A.; Chen, Y.; Petzold, C. J.; Haushalter, R. W.; Blank, L. M.; Keasling, J. D. Maximizing heterologous expression of engineered type I polyketide synthases: investigating codon optimization strategies. *ACS Synth Biol*. 2023;12(11):3366-3380. doi: 10.1021/acssynbio.3c00367.

Contributions
This chapter was written by Matthias Schmidt and reviewed by Lars M. Blank and Jay D. Keasling.

2 Materials and Methods

2.1 Chemicals, media and growth conditions

Unless specified differently, all chemicals were obtained from Sigma-Aldrich (USA). Authentic standards for (2S,3S)-3-hydroxy-2,4-dimethylpentanoic acid (3H24DMPA), (2S,3S)-3-hydroxy-2-methylpentanoic acid (3H2MPA), 3-hydroxy-4-methylpentanoic acid (3H4MPA), and 3-hydroxy-4-methylhexanoic acid (3H4MHA) were acquired from Enamine (Ukraine).

General precultures of *E. coli*, *C. glutamicum*, and *P. putida* were grown from a single colony and incubated overnight. We utilized lysogeny broth (LB) medium for *P. putida* and *E. coli* cultures, with temperature settings of 30 and 37 °C respectively.[133] For *C. glutamicum*, we used brain heart infusion (BHI) medium at a temperature of 30 °C. When needed, the medium was supplemented with 50 µg/mL of kanamycin, 30 µg/mL gentamycin, or 100 µg/mL carbenicillin. The conditions for precultures and primary cultures remained the same throughout this study, unless otherwise noted. For primary cultures, 2 mL LB or BHI medium was inoculated with 20 µL preculture and incubated in a 24-well plate (VWR, USA). If the experiment required minimal media, we used 3-(N-morpholino)-propanesulfonic acid-based (MOPS) (Table 1) or M9 minimal media (Table 2).[134] In cases of strains with a genomically inserted selection marker, no antibiotic was added to the cultures. The shaking speed was maintained at 200 rpm. The duration and temperature of the cultivation varied depending on the organism and the requirements of the sample for analyses. Strains containing *lacI* were induced by adding 200 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) to the medium.

Table 1: Final composition of the MOPS minimal medium.

Stock	Component	Final concentration
Salts		
	MOPS	40 mM
	Tricine	4 mM
	K_2HPO_4	1.32 mM
	K_2SO_4	0.29 mM
	$FeSO_4 \cdot 7H_2O$	0.01 mM
	$MgCl_2 \cdot 6H_2O$	0.52 mM
	NaCl	50 mM

	NH_4Cl	9.52 mM
	$CaCl_2$	32.5 µM
	$FeCl_2$	8 µM
Micronutrients		
	$ZnSO_4 \cdot 7H_2O$	0.1 µM
	H_3BO_3	4 µM
	$(NH_4)_6Mo_7O_{24} \cdot 4H_2O$	0.03 µM
	$CoCl_2 \cdot 6H_2O$	0.3 µM
	$MnCl_2 \cdot 2H_2O$	0.8 µM
	$CuSO_4 \cdot 5H_2O$	0.1 µM

Table 2: Final composition of the M9 minimal medium.

Stock	Component	Final concentration
Salts		
	Na_2HPO_4	47.8 mM
	KH_2PO_4	22 mM
	NaCl	3.7 mM
	NH_4Cl	18.7 mM
	$MgSO_4 \cdot 7H_2O$	1 mM
	NH_4Cl	9.52 mM
	$CaCl_2$	0.1 mM

2.2 Plasmid construction

Plasmids generated in this work are listed in Supplemental Table 1. The codon optimized gene sequences of LipM1 fused to EryM6-TE and SpnA were synthesized and cloned into the pBH026 vector backbone by Genscript (USA). The codon optimized genes *accA2* and *pccB* for the synthetic propionyl-CoA carboxylase (PCCase) operon were cloned into pBH027 (Genscript, USA). Primers used (Supplemental Table 3) to generate Gibson assembly parts were designed by the j5 software and purchased from Integrated DNA Technologies (USA).[135,136] All assemblies were performed using the HiFi DNA Assembly Master Mix, following the manufacturer's instructions (NEB, USA). After 1 h at 50 °C, 10 µL assembly mix was transformed into *E. coli* XL1-Blue via 45 s heat shock at 42 °C (Agilent, USA). Cells were then recovered in LB for 1 h at 37 °C and plated on LB agar containing the appropriate

antibiotic. On the next day, single colonies were used to inoculate LB medium with the same antibiotic and grown at 37 °C. After sufficient growth, plasmids were extracted using the QIAprep Spin Miniprep Kit (Qiagen, Germany). Successful assembly was confirmed using the primer-free, whole-plasmid sequencing service by Primordium (USA).

2.3 Strain construction

Strains generated in this work are listed in Supplemental Table 2. In general, in-frame gene deletions were created via HR. Gene integrations were either achieved by HR, Tn7 delivery transposon or serine recombinases.[85,92,137] Electrocompetent *P. putida* or *E. coli* cells were prepared as previously described.[138] Briefly, a single colony of *P. putida* or *E. coli* was cultured overnight in 2 mL LB, and harvested via centrifugation at 16,000 x g. The cells were washed twice in 1 mL 0.3 M sucrose, and resuspended in 200 µL of the same solution. 500 ng of plasmid DNA was added to the resuspension and mixed. 100 µL of this mixture was transferred to a 2 mm gap electroporation cuvette (Bio-Rad Laboratories) and electroporated (25 µF, 200 Ω, 1.8 kV, MicroPulser, Bio-Rad Laboratories). Immediately after, 500 µL of LB was added and mixed, avoiding bubble formation to not affect electroporation efficiency. The cells were then incubated at 30 °C for 2 h for recovery. Finally, 100 µL of the cell suspension was plated on LB agar with the appropriate antibiotic. Integrations and deletions were confirmed via colony PCR (cPCR), respectively. The primers used can be found in Supplemental Table 3. cPCR was performed by boiling (10 min, 100 °C) a single colony in 20 µL of dimethyl sulfoxide (DMSO). 0.5 µL of that boiled suspension were used as a template for PCR with DreamTaq Green PCR Master Mix (Thermo Fisher, USA).

2.3.1 Homologous recombination

P. putida HR was executed as described by Thompson *et al*.[52] Briefly, a vector with a homologous DNA sequence (500-1000 bp), a selection marker, and a counterselection marker was introduced into the target strain via electroporation or conjugation. For conjugation, the vector was introduced into *E. coli* S17λpir. Overnight cultures of both the donor and recipient strains were washed and resuspended in 100 µL of antibiotic-free LB. Equal volumes of these suspensions were mixed, spot-plated on LB agar, and incubated at 30 °C overnight. The next day, the cells were scraped from the agar and resuspended in 1 mL LB. 100 µL of this suspension was spread on Difco (USA) Pseudomonas Isolation Agar (PIA) with either kanamycin or gentamycin. Successful plasmid integration is indicated by single colonies resistant to the respective selection marker. These colonies were then incubated overnight in liquid LB

with kanamycin or gentamycin. 100 µL of a 1:100 dilution was spread on 10% sucrose LB agar plates without NaCl. The following day, single colonies were streaked on both 10% sucrose LB agar and LB agar containing the relevant antibiotic. Colonies growing on sucrose agar but not on the antibiotic plate were selected for cPCR. Primers (Supplemental Table 3) flanking the target site yield a band smaller than the one from the WT when compared.
C. glutamicum HR was executed as outlined by Zhan *et al*.[139] Initially, a BHI overnight culture was inoculated into the electroporation medium, consisting of 37 g BHI powder, 25 g glycine, 10 mL Tween, and 4 g isoniazid. After 5-6 hours of incubation, the cells reached an optical density (OD) of approximately 1 (λ = 600 nm). Post-growth, these cells were washed three times with ice-cold 10% glycerol and then resuspended in 1 mL of the same solution. 80 µL of the competent cells was transferred to ice-cold electroporation cuvettes with a 2 mm gap. 500 ng of plasmid DNA was used for the transformation. After electroporation, the cells were reintroduced to 1 mL of BHI medium and heat-shocked at 46 °C for 5 minutes. The cells recovered at 30 °C for 1-2 hours, then were spread on BHI agar plates supplemented with 25 µg/mL kanamycin and incubated for 2 days. Individual colonies were later streaked onto BHI agar containing 10% sucrose and incubated another 1-2 days at 30 °C to remove the selection marker. The final step was the selection and verification of colonies through cPCR (Supplemental Table 3).

2.3.2 Tn7 integration

Tn7-based integrations into the *P. putida* genome followed a modified protocol from Zobel *et al*.[137] First, overnight cultures of *E. coli* DH5α-λpir pTnS-1, *E. coli* HB101 pRK2013, *E. coli* PIR2 pBG14e, and the recipient *P. putida* strain were washed and combined at a 1:1 ratio. 100 µL of that suspension was spot plated on antibiotic-free LB agar. The following day, 100 µL of a 1:10 dilution of the scraped off cells was spread onto PIA with kanamycin. After an overnight incubation at 30 °C, single colonies were confirmed via cPCR. The integrated pBG14e vectors have an FRT-flanked selection marker, allowing for optional removal of the selection marker using pBBFLP.[21] By electroporating the pBBFLP plasmid into the kanamycin-resistant strain, flippase activity initiates recombination at the FRT sites.[94] To improve excision efficiency, strains were plated on LB agar containing 30 µg/mL tetracycline. Streaking the resulting colonies on both kanamycin LB agar and antibiotic-free agar confirmed the regained kanamycin sensitivity.

2.3.3 Serine recombinase-assisted genome engineering

The integration system comprises the *P. putida* AG5577 and *C. glutamicum* AG6212 "landing pad" strains, as well as the cloning vector library.[92] The *P. putida* landing pad strain contains three distinct polyAttB sequences integrated at three different loci of its genome (PP_4740, PP_2876 and PP_4217/8). *C. glutamicum* AG6212 possesses a single polyAttB sequence that contains all nine *attB* sites at Cgl1777-8. The integration vector library contains the corresponding *attP* sites and the plasmid backbone is flanked by the *attB* and *attP* site of the integrase ϕC31. Similar to the FRT-FLP system, electroporating pGW30 or pALC412 enables selection marker removal. Throughout this study, a modified version of the BxB1 integration vector pJH0204 was used (designated as pBH026). This modified version carries a copy of the *lacI* gene which represses the LlacO1 promoter. In addition to removing the selection marker, excision of the backbone via the ϕC31 integrase leads to removal of the repressor LacI, and the promoter becomes constitutively active.[121]

Integration of pBH026 was achieved by electroporating pBH026 and pGW31 into the target strain. The pGW vector library contains one out of the nine serine recombinases under the control of a constitutive Tac promoter. Furthermore, pGW plasmids cannot replicate in the target strains due to the *E. coli* specific ColE1 origin of replication.[140] Transient expression of the integrase is sufficient to lead to whole plasmid integration during the 2 h recovery period. After recovery, 100 µL of cell culture was plated on LB or BHI agar supplemented with 50 µg/mL kanamycin. Incubation overnight at 30 °C allowed for colony formation. The repression by the genomically integrated *lacI* cassette is usually very strong and requires the removal via ϕC31 excision. Therefore, we usually proceed to the preparation of electrocompetent cells. After preparing the overnight culture for electroporation, pGW30 or pALC412 was transformed into the kanamycin-resistant strain. Recovered cells were then plated on LB or BHI agar with 25 µg/mL apramycin. The pGW30 and pALC412 vectors contain an apramycin selection marker and the temperature-sensitive origin of replication mSFts1.[141] This enables the curation of the excision vectors when cells are grown at 42 °C. However, restreaking the apramycin-resistant cells on antibiotic-free agar is usually sufficient to cure the ϕC31 integrase vector. The resulting colonies are checked for loss of resistance by streaking single colonies on LB agar with and without the corresponding selection marker. Finally, non-resistant cells were selected for cPCR confirmation of the successful integration and backbone excision.

2.4 Plate-based growth assay

Bacterial growth studies were carried out using plate reader kinetic assays, following the methodology described in Thompson *et al*. and Schmidt *et al*.[52,121] Overnight cultures of the bacterial strains were subjected to three washes with nitrogen- or carbon-free MOPS minimal medium. These cultures were then used to inoculate 48-well plates at a ratio of 1:100 (Falcon, USA), with each well containing 500 µL of MOPS medium supplemented with 10 mM of the tested nitrogen or carbon source. To ensure a proper seal, the plates were covered with a gas-permeable microplate adhesive film (VWR, USA). The OD of the cultures was monitored for 24 to 72 hours at 30 °C using a Biotek Synergy H1M plate reader (BioTek, USA) set to fast continuous shaking. The OD was measured at 600 nm. Optionally, red fluorescent protein (RFP) levels were also measured using an excitation wavelength of 535 nm, an emission wavelength of 620 nm, and a gain of 100. Additionally, green fluorescent protein (GFP) levels were measured with an excitation wavelength of 395 nm and an emission wavelength of 509 nm.

2.5 BarSeq assay

BarSeq experiments used the *P. putida* library JBEI-1, following methods described in Thompson *et al*. and Schmidt *et al*.[52,121] 2 mL aliquots of JBEI-1 libraries were thawed on ice, mixed into 25 mL of LB medium with kanamycin, and cultivated at 30 °C until an OD600 of 0.5. Subsequently, three 1-mL aliquots were taken out, pelleted, and stored at −80 °C as initial time points. The libraries underwent three washes in MOPS minimal medium without nitrogen or carbon and were then used to inoculate each experiment at a 1:100 ratio. Tests were carried out in 24-well plates; each well held 2 mL of nitrogen- or carbon-free MOPS minimal medium with 10 mM of the tested nitrogen or carbon source. Plates were cultivated at 30 °C with a shaking speed of 200 rpm. After 24 to 72 h, 1 mL samples were harvested when the cultures seemed dense enough for DNA extraction. These samples were pelleted and kept at −80 °C until DNA was extracted using a DNeasy UltraClean Microbial kit (Qiagen, Germany). BarSeq analysis was conducted as outlined in Wetmore *et al*.[48] Strain fitness is the normalized log2 ratio of the barcode reads in the harvested sample versus the timepoint zero sample. Gene fitness is the weighted average of strain fitness for insertions within 10% to 90% of the gene. The main statistic t value depicts the form of fitness divided by the variance across mutants of the same gene. Statistic t values > |4| and fitness values > |1| were regarded as significant. An in-depth explanation of these calculations can be found in

Wetmore *et al*.[48] All experiments met prior quality metrics. They were performed in biological duplicates, and published fitness data can be accessed at http://fit.genomics.lbl.gov.

2.6 Protein production and purification

The ω-amino acid aminotransferases purification procedure is an adaptation of the method originally described by Yuzawa *et al.*, with the updated version available in Schmidt *et al*.[119,121] In summary, *E. coli* BL21(DE3) cells containing the pET28a vectors carrying PP_0596, PP_2180, or PP_5182 were grown in LB medium with kanamycin. This culture was maintained at 37 °C until the OD600 reached a value of 0.4. The protein expression was induced by adding 250 μM IPTG, after which the cultures were continued at 18 °C for an additional 24 hours.

Following this, cells were harvested by a 20 min centrifugation at 5000 g. The resulting pellet was then resuspended in 30 mL cold wash buffer, consisting of 50 mM sodium phosphate (pH 7.6), 300 mM NaCl, and 10 mM imidazole. Cell lysis was achieved through sonication, applied in 8 cycles of 30 seconds each. Any cellular debris was removed by three centrifugations for 15 min at 8,000 g and 4 °C. The soluble fraction underwent a 1 h incubation with 4 mL of Nickel-NTA agarose beads (Thermo Fisher, USA) at 4 °C. This solution was then run through a Nickel-NTA column, followed by a triple wash using the wash buffer. The targeted protein was subsequently eluted using 12 mL of an elution buffer, which was comprised of 150 mM sodium phosphate buffer (pH 7.6), 50 mM NaCl, and 150 mM imidazole, maintained at 4 °C. The protein's buffer was later swapped for a stock buffer via dialysis, employing the SnakeSkin Dialysis Tubing with 10k MWCO (Thermo Fisher, USA). The stock buffer had the following composition: 100 mM sodium phosphate (pH 7.5), 0.5 mM DTT, and 10% glycerol. The protein solution was concentrated using 30k MWCO Amicon Ultra-15 centrifugal filters (MilliporeSigma, USA) and stored at a temperature of −80 °C.

2.7 *In vitro* aminotransferase characterization

The substrate specificity of the isolated aminotransferases was assessed through *in vitro* production of L-alanine [11]. The *in vitro* reaction was conducted in 100 mM sodium phosphate buffer at pH 9, containing 500 μM pyruvate, 1 mM pyridoxal phosphate (PLP), 5 μM purified enzyme, and 5 mM substrate. Once the substrate was added, the mixture was promptly incubated at 30 °C for 30 minutes. The assay was stopped by heating the solution for a duration of 10 minutes at 100 °C. Subsequent measurements of alanine concentrations were conducted

enzymatically with the Alanine assay kit from Cell Biolabs (USA), following the manufacturer guidelines.

2.8 Codon optimization of engineered polyketide synthase

PKS codon optimizations were performed using the original python script of DNA Chisel with the implemented Kazusa database.[121,142] The target hosts were *E. coli* str. K-12 substr. W3110 (NCBI:txid316407), *C. glutamicum* ATCC13032 (NCBI:txid196627) and *P. putida* KT2440 (NCBI:txid160488). The three algorithms applied are use best codon (ubc), match codon usage (mcu), and harmonize relative codon adaptiveness (hrca). The ubc method replaces every codon with the most abundant codon used for that AA. mcu and hrca harmonize the codon frequency of the optimized gene with the original sequence or the source organism's codon usage, respectively.[142] For the hrca method, the LipPKS part was harmonized with *S. aureofaciens* (NCBI:txid1894) and the EryPKS part with *Saccharopolyspora erythraea* (NCBI:txid1836). To ensure reproducible results, the Numpy random generator was set to 123. A list of all the applied constraints can be found in Table 3. The UniquifyAllKmers constrain leads to less repetitive sequences and lowers overall synthesis difficulty by not repeating the same codon for the same AA. The AvoidHairpin constrain controls the maximum allowed hairpin stem size and ensures a certain distance between the occurrence of hairpins. Hairpins are secondary structures that are formed by ssDNA or mRNA and not only interfere with DNA synthesis but also affect transcription.[143] The AvoidPattern command is used to remove certain nucleotide sequences such as homopolymers or cut sites. EnforceGCContent leads to a set minimum and maximum average GC content in a given window.[142]

The GUI based on DNA Chisel was created using the open-source python package Streamlit and can be reached under https://basebuddy.lbl.gov. Files and scripts used to develop the GUI were deposited on Github (github.com/jbei/basebuddy).

Table 3: Applied codon optimization constrains.

Codon optimization constrain
UniquifyAllKmers(9, include_reverse_complement=True)
AvoidHairpins(stem_size=10, hairpin_window=100)
AvoidPattern("9xA")
AvoidPattern("9xT")
AvoidPattern("6xC")
AvoidPattern("6xG")

```
AvoidPattern("NdeI_site")
AvoidPattern("XhoI_site")
AvoidPattern("SpeI_site")
AvoidPattern("BamHI_site")
AvoidPattern("BsaI_site")
EnforceGCContent(mini=0.3, maxi=0.75, window=50)
EnforceTranslation()
```

2.9 Quantitative reverse transcription PCR

Reverse transcription quantitative PCR (RT-qPCR) was performed as previously described.[121] After a growth period of 24 hours, cells were harvested by centrifuging 1 mL of the cell culture. It was essential to incubate *C. glutamicum* cells with 1 mL of lysozyme (2 mg/mL, Roche, Switzerland) at 30 °C for 10 min to facilitate efficient RNA release post-phenol treatment. Total RNA extraction was accomplished using the RNeasy Plus Universal Mini Kit (Qiagen, Germany). The acquired RNA was subsequently used as the template for the synthesis of complementary DNA (cDNA), using the LunaScript RT SuperMix Kit (NEB, USA). This was followed by the qPCR procedure with the Luna Universal qPCR Master Mix (NEB, USA). The designated PCR mix underwent an initial heating cycle at 95 °C for 1 minute, followed by 40 cycles: each at 95 °C for 15 seconds and 60 °C for 30 seconds. The amplification was captured by a CFX96 Real-Time PCR Detection System (Bio-Rad, USA). The expression ratio resulting from the qPCR was calculated based on the Ct value difference between the target gene and specific housekeeping genes: *rpoD* for *P. putida* and *E. coli*, and *rpoC* for *C. glutamicum*. Primer sequences for both the housekeeping and PKS genes can be found in Supplemental Table 3.

2.10 Phosphopantetheinyl transferase assay

The protocol for indigoidine extraction and quantification has been previously described by Wehrs *et al*.[144] Strains carrying the pMQ80 BpsA vector were grown for 24 h, and harvested by centrifugation. The pellet was then resuspended in the same volume of DMSO. DMSO facilitates the permeation of the cell wall, allowing for the release of the intracellular indigoidine. Afterward, the mixture was vortexed for 10 min at 3000 rpm, followed by centrifugation at maximum speed for 2 minutes. Indigoidine concentration was estimated by

measuring the absorbance of the supernatant at 612 nm with a Biotek Synergy H1M plate reader (BioTek, USA).

2.11 Liquid chromatography and mass spectrometry

This section describes the employed liquid chromatography (LC) and mass spectrometry (MS) techniques to analyze the concentration of substrates, intermediates and products. Furthermore, LCMS-based protein analysis was used to quantify the proteome of recombinant cells.

2.11.1 Caprolactam degradation assay

P. putida KT2440 cannot utilize caprolactam as a growth substrate. However, by opening the lactam ring, cells can access the terminal amino group as a source of nitrogen. This process results in the formation of adipic acid.[11]

To study the degradation of caprolactam in *P. putida* KT2440, WT cells were cultivated in 25 mL MOPS minimal medium with 20 mM glucose as previously described [11]. The nitrogen source was either 10 mM ammonium chloride or 10 mM caprolactam. Triplicates of each condition were grown in 250 mL shake flasks at 30 °C and 200 rpm. The OD and metabolite concentration of the cultures was measured at 1, 6, 12, 24, 48, and 72 hours. Sample preparation for metabolite analysis included a centrifugation at 16,000 *g* for 1 min, and quenching 300 µL supernatant with 300 µL −80 °C cold methanol. After that the mixture was filtered using 3 kDa Amicon Ultra-0.5 centrifugal filters (MilliporeSigma, USA) and stored at −80 °C for further analysis.

To quantify caprolactam and byproducts, we utilized LC separation on a 1260 HPLC system (Agilent Technologies, USA). The separation was conducted using a Kinetex HILIC column, which had dimensions of 100 mm in length, 4.6 mm in internal diameter, and a 2.6 µm particle size (Phenomenex, USA), following the methodology described by Schmidt *et al*.[121] The injection volume was consistently set at 2 µL. The sample tray and column compartment were maintained at 6 °C and 20 °C, respectively.

For the mobile phase, we employed two solvents. Solvent A consisted of 10 mM ammonium formate and 0.2% formic acid dissolved in water. In contrast, Solvent B was a mix of 10 mM ammonium formate and 0.2% formic acid in a solution made up of 90% acetonitrile and 10% water. The gradient began at 90% B, reduced to 70% over 4 minutes, held for 1.5 minutes, then dropped to 40% B in 0.5 minutes. It remained steady at 40% B for 2.5 minutes before increasing to 90% B in 0.5 minutes and maintained for a final 2 minutes. As for the flow rate,

it remained at 0.6 mL/min for 6.5 minutes, increased to 1 mL/min in 0.5 minutes, and was then kept constant at this rate for the subsequent 4 minutes. This sequence resulted in a total run time of 11 minutes.

Post-separation, the 1260 HPLC system's output was channeled to an Agilent Technologies 6520 QTOF-MS, designated for quadrupole time-of-flight mass spectrometric detection. The transition was made via a 1:4 post-column split, guiding the effluent to the electrospray ionization (ESI) ion source of the QTOF-MS. Operating in positive ion mode, the ESI facilitated the conversion of the sample into gas-phase [M+H]+ ions. Instrument settings included a capillary voltage of 3500 V and voltages for the fragmentor, skimmer, and OCT 1 RF of 100 V, 50 V, and 250 V, respectively. The drying gas parameters were set to a temperature of 350 °C, a flow rate of 12 L/min, and a nebulizer pressure of 25 lb/In^2. Tuning was accomplished using the Agilent ESI-Low TOF tuning mix, covering a range of m/z 50 to 1,700. For ensuring high mass accuracy, a reference mass correction was done using purine and HP-0921 solutions (Agilent Technologies, USA). Data was captured in the 50 to 1,100 m/z range. Lastly, to quantify the analytes, we generated a six-point calibration curve from 2-fold serial dilutions starting at 25 µM and descending to 0.78125 µM.

2.11.2 Quantification of polyketide product

In the bacterium *P. putida*, the truncated version of LipPKS is capable of synthesizing enantiopure short-chain 3-hydroxy acids. Specifically, it produces 3H24DMPA, (2S,3S)-3-hydroxy-2,4-dimethylhexanoic acid, and (2S,3S)-3-hydroxy-2,5-dimethylhexanoic acid.[121] By exchanging the AT1 domain, the methyl group on the β-carbon can be removed.[123]

The separation of 3-hydroxy acids was achieved using an Agilent 1260 Infinity II LC System equipped with a Kinetex XB-C18 column (100 mm length, 3 mm internal diameter, 2.6 µm particle size; Phenomenex, USA). This LC separation was carried out at room temperature. The mobile phase was divided into two solvents: solvent A comprised 0.1% formic acid in water, while solvent B contained 0.1% formic acid in methanol. The elution gradient began with 20% of solvent B, gradually increasing to 72.1% over 6.5 minutes. It then ramped up to 95% B over the next 1.3 minutes and remained steady for an additional minute. Following this, the flow rate adjusted to 0.65 mL/min, and the gradient transitioned from 95% to 20% B in 0.2 minutes, then held constant for 1.2 minutes at a flow rate of 0.42 mL/min.

For the detection and quantification of the 3-hydroxy acids, the LC output was channeled into an Agilent InfinityLab LC/MSD iQ single quadrupole mass spectrometer (Agilent Technologies, USA). The detection method employed ESI in the negative-ion mode. The 3-hydroxy

acids were identified by matching both the mass and retention time to those of enantiopure and racemic mixtures of reference standards.

2.11.3 Quantification of sugars and acids

The quantification of sugars and acids was conducted with an Agilent 1100 series LC system (Agilent Technologies, USA) coupled to a refractive index (RI) and ultraviolet (UV) detector (Agilent Technologies, USA). The methodology applied for this analysis was developed and described by Lai *et al*.[145] Prior to analysis, samples were first processed via centrifugation and then filtered using a 0.2 µm polyethersulfone (PES) 96-well filter plate (Pall, USA). Chromatographic separation was achieved using an Agilent Hiplex H column (300 × 7.7 mm, 8 µm particle size, Agilent Technologies, USA) with the column oven maintained at 40°C. The chosen mobile phase was an aqueous solution of 14 mM sulfuric acid, with an isocratic flow rate of 0.4 mL/min over a duration of 50 minutes. The RI detector operated on a positive polarity and was set to a temperature of 40°C. The UV absorbance was monitored at 210 nm. By matching the retention times with those of authentic standards, the analytes were identified, and their concentrations were determined through a calibration curve that ranged from 20 mM to 0.3125 mM.

2.11.4 Proteomics analysis

Proteins were extracted and tryptic peptides generated using previously described proteomic sample preparation methodologies.[121,146] Cell pellets were resuspended in Qiagen P2 Lysis Buffer (Qiagen, Germany) to facilitate cell lysis. After this, proteins were precipitated through the addition of 1 mM NaCl and a 4-fold volume of acetone. This step was followed by two washes using an 80% acetone aqueous solution. The isolated protein pellet was then homogenized by pipette mixing with a solution of 100 mM ammonium bicarbonate mixed with 20% methanol. The DC protein assay (BioRad, USA) was utilized to determine protein concentration. Under ambient conditions, proteins were reduced using 5 mM tris 2-(carboxyethyl)phosphine (TCEP) for half an hour and alkylated in the dark with 10 mM iodoacetamide for the same duration. An overnight enzymatic digestion was carried out with trypsin at a 1:50 trypsin-to-protein ratio.

Resultant peptide samples were subjected to analysis via an Agilent 1290 UHPLC system in tandem with a Thermo Scientific Orbitrap Exploris 480 mass spectrometer.[147] Peptides were channeled onto an Ascentis ES-C18 Column (Sigma Aldrich, USA), and elution was achieved using a gradient over 10 minutes, starting from 98% solvent A (0.1% formic acid in water)

and 2% solvent B (0.1% formic acid in acetonitrile), transitioning to a 65% solvent A and 35% solvent B composition. Once eluted, peptides entered the mass spectrometer in a positive-ion mode and were assessed using data-independent acquisition (DIA) mode. This mode comprised three survey scans spanning from m/z 380 to m/z 985 and 45 MS2 scans with an isolation width of 13.5 m/z.

The raw data from DIA was analyzed via the DIA-NN software suite.[148] The search databases in DIA-NN (in library-free mode) incorporated the latest Uniprot proteome FASTA sequences for relevant microorganisms, in addition to heterologous protein sequences and common proteomic contaminants. DIA-NN automatically determines mass tolerances and optimal mass accuracies based on an initial sample analysis. Each MS run's retention time extraction window was determined using DIA-NN's automated optimization procedure. Protein inference was activated, and quantification was aligned to the Robust LC = High Accuracy strategy. Finalized DIA-NN reports were filtered with a global false discovery rate (FDR) of 0.01 at both the precursor and protein group levels. Quantities of the targeted proteins in the samples were represented graphically using the Top3 method, which considers the mean MS signal response of the three most intense tryptic peptides of each identified protein.[149,150]

2.12 Bioinformatic analysis

Statistical evaluations were conducted using the Python Scipy and Numpy libraries.[151,152] To quantify the number of *P. putida*'s aminotransferases, we employed HMMER version 3.3.2 (hmmer.org) to analyze the *P. putida* KT2440 genome by comparing it to a profile database created from the Pfam HMM files of PF00155, PF00202, PF01063, and PF00266. Hits with E-values greater than 1e-20 were considered significant. However, those labeled as transcriptional regulators were not included in our final set. Furthermore, HMMER was utilized to find potential genes linked to the γ-glutamyl cycle, using the Pfam HMM files of PF00120, PF01266, PF00171, and PF07722, which correspond to the glutamyl-polyamine synthetase, the glutamyl-polyamine oxidase, the aldehyde dehydrogenase, and the amide hydrolase, respectively.

We extracted data on transcription factors and transport-related proteins from the BarSeq dataset using the MiST 3.0 and TransportDB 2.0 databases.[153,154] EggNOGmapper assisted in generating the Clusters of Orthologous Groups (COGs) for *P. putida* KT2440.[155,156] Moreover, our in-depth data examination and suggestions regarding metabolic pathways were largely informed by BioCyc and PaperBlast.[157,158]

The 16S rRNA sequences for phylogenetic analysis were sourced from the RNAcentral database. These sequences were aligned employing the CLUSTAL W technique, as featured in MEGA 11.[159] Following this, we built a phylogenetic tree using the maximum-likelihood approach paired with the Tamura-Nei model.

For Principal Component Analysis (PCA), we derived codon usage information for organisms indexed in Refseq from the Codon and Codon Pair Usage Tables (CoCoPUTs) database. Out of the 235,025 database entries, each entry's codon usage data was adjusted by the total count of codons analyzed. Taxonomic categorizations for each Refseq entry were sourced from the NCBI database, offering a structured approach to taxonomic analysis. To understand the overarching trends in codon usage among different organisms, we employed PCA on the normalized codon datasets. We then highlighted and plotted the two most dominant principal components, which captured the maximum variance, using a scatter plot.

Chapter 3

Results

Chapter 3.1

Nitrogen Metabolism in *Pseudomonas putida*: Functional Analysis using RB-TnSeq

Published as:

Schmidt, M.; Pearson, A. N.; Incha, M. R.; Thompson, M. G.; Baidoo, E. E. K.; Kakumanu, R.; Mukhopadhyay, A.; Shih, P. M.; Deutschbauer, A. M.; Blank, L. M.; Keasling, J. D. Nitrogen metabolism in *Pseudomonas putida*: functional analysis using random barcode transposon sequencing. *Appl Environ Microbiol*. 2022;88(7):e0243021. doi:10.1128/aem.02430-21.

Contributions

Experimental work was divided evenly between M.S. and A.N.P. M.R.I. performed t-SNE analysis and assisted in the writing process. E.E.K.B. and R.K. performed LCMS analysis for the caprolactam degradation assay. A.M.D. provided quantitative sequencing and statistical evaluation of BarSeq data. M.S. and A.N.P prepared the manuscript with the help of M.G.T., M.R.I., A.M., P.M.S., L.M.B. and J.D.K. J.D.K. conceived the project.

3 Results

3.1 Nitrogen Metabolism in *Pseudomonas putida*: Functional Analysis using RB-TnSeq

3.1.1 Abstract

Pseudomonas putida KT2440 is renowned for its versatile and resilient metabolism. However, many genes and proteins enabling these growth characteristics remain unknown. Through pooled mutant fitness assays, we identified genetic phenotypes responsible for assimilating 52 distinct nitrogenous compounds. Additionally, we evaluated *P. putida*'s amino acid biosynthesis by creating single amino acid exclusion conditions. Of these 72 conditions, 672 genes displayed distinct fitness phenotypes, including 100 linked to transcriptional regulation and 112 associated with transport functions. We categorized these conditions into six classes and suggest assimilation pathways for all the tested compounds. In addition, we assayed the activity of three aminotransferases to determine their substrate preference *in vitro*. We also tested the selectivity of five transcriptional regulators, explaining some of the fitness findings and demonstrating their potential as synthetic biosensors. Moreover, we employed a manifold learning technique to make these data more accessible to other researchers.
In this chapter, the utility of RB-TnSeq in *P. putida* is showcased and also complemented by a summary of several decades of metabolic research on this organism.

3.1.2 Introduction

Pseudomonas putida KT2440 and its metabolism have garnered increasing interest for a broad range of applications. As a bioproduction host, *P. putida* can convert aromatic compounds, derived from lignocellulosic feedstocks, into valuable chemicals.[160] Additionally, its metabolic versatility offers a rich source of enzymes for innovative metabolic pathways.[161] However, to engineer such microbes, extensive knowledge of the target host's metabolism is essential, especially when dealing with a metabolic network as complex as in *P. putida* KT2440. In past research, we utilized BarSeq to delve into the degradation of lysine, fatty acid, alcohol, and aromatic compounds in *P. putida*.[12,19,52] Yet, we still lack a detailed *in vivo* functional understanding of many aspects of its nitrogen metabolism.
Being a soil-dwelling bacterium, *P. putida* comes across various organic and inorganic nitrogen sources. Its responses to these environments have been the subject of recent research.

When colonizing the rhizosphere, *P. putida* shows chemotaxis towards plant root exudates, rich in nitrogen compounds like benzylamines, polyamines, pyrrole derivatives, nucleotide derivatives, purines, amino acids, and phenylpropanoids.[162,163] Furthermore, it exhibits chemotaxis towards and degradation of benzoxazinoids, phytotoxic compounds released by maize seedling roots.[164,165] Impressively, *P. putida* not only processes these diverse compounds in nitrogen-rich environments but also adapts to nitrogen-poor conditions. This adaptability is evident through elevated PHA production, suppressed carbon catabolism, and enhanced nitrogen uptake transporter expression.[166]

Considering *P. putida*'s adaptability within the rhizosphere and its growing significance in sustainable chemical bioproduction, it is logical that its nitrogen metabolism is being studied extensively within the metabolic engineering community. The potential of PHAs as a future bioplastic has led to in-depth analyses of *P. putida* strains under nitrogen-limiting growth conditions, aiming to decipher the triggers of PHA synthesis in such environments.[167,168] Furthermore, the bacterium's nitrogen metabolism and regulatory mechanisms have been sources of "parts" for metabolic engineering. A subset of its predicted 39 aminotransferase-activity proteins has been employed in benzylamine derivative and glutaric acid yielding pathways.[169,170] The transcriptional factor regulating the degradation of capro- and valerolactams has evolved into a precise biosensor, promising enhanced lactam production and its use in metabolic pathway design.[58] Moreover, understanding *P. putida*'s processing of nitrogen compounds is key to redirecting metabolic flux towards desired products. For instance, optimizing glutamine, a precursor to the sustainable blue pigment indigoidine, has resulted in better yields via CRISPRi-mediated metabolic adjustments.[171]

Despite its significance in both metabolic engineering and fundamental research, there are existing gaps in our comprehension of *P. putida*'s nitrogen metabolism. Several gene functions, inferred from homology predictions, lack robust *in vivo* functional data. This shortfall can impact the reliability of metabolic models. Additionally, the presence of multiple highly similar enzymes in *P. putida*, each possibly having distinct substrate preferences, highlights the need for functional genomics. Such studies can shed light on specific gene roles and clarify the function of genes with multiple copies.

In this study, we employ BarSeq to analyze *P. putida* KT2440's metabolism of 52 distinct nitrogenous compounds, effectively increasing the available BarSeq data for this bacterium by nearly two-fold. We present evidence supporting numerous recognized nitrogen assimilation pathways and their associated regulatory systems, while also elucidating the roles of genes previously unknown in nitrogen metabolism. Recognizing their significance in recent

metabolic engineering studies, we further assay the substrate specificity of *P. putida*'s aminotransferases and 5-oxoprolinases.[56,169,170] To elucidate the *in vivo* regulation of these enzymes, we delve deeper into understanding their corresponding regulators. This research enhances the viability of *P. putida* both as a primary microbial host and as a source of metabolic engineering components for sustainable chemical synthesis.

3.1.3 Results and Discussion

3.1.3.1 BarSeq reveals the genetic bases of diverse nitrogen metabolisms

Using BarSeq, we determined genes responsible for the utilization of nitrogen from both natural and synthetic compounds. For these tests, we cultured a collection of barcoded transposon insertion mutants in minimal medium supplemented with glucose and different sole sources of nitrogen. Table 4 provides a list of these nitrogen sources along with the corresponding sections where each is discussed in detail. These conditions uncovered 672 genes exhibiting pronounced ($|fitness| > 1$) and statistically significant ($|t| > 5$) fitness phenotypes. The findings comprise 100 transcription factors, 112 transport proteins, and various other enzymes that hold potential for metabolic engineering applications (Figure 3).[153,154] Furthermore, 529 of these genes correspond to proteins that remain unreviewed in the Uniprot database, with 256 not displaying significant phenotypes in prior BarSeq research.[12,172] For a clearer representation of the fitness data, we used the t-distributed stochastic neighbor embedding (t-SNE) manifold learning technique to group genes by their fitness values under the conditions examined.[173] Clusters were labeled according to the condition that produced the most consistent and pronounced fitness score shifts for the genes they contained (Figure 3b). Within this visual representation, it is possible to discern genes that might participate in analogous metabolic pathways. The interactive version of t-SNE, complete with hyperlinks to the Fitness Browser,[55] can be accessed here: ppnitrogentsne.lbl.gov. While it is straightforward to pinpoint genes with unique phenotypes in a single condition, it becomes more complicated to determine the specific role of genes across multiple conditions. This issue is especially pronounced for genes essential in a vast majority of the examined conditions, such as those integral to amino acid synthesis. Unexpectedly, t-SNE managed to differentiate clusters of genes tied to the synthesis of tryptophan, arginine, methionine, and branched-chain amino acids (BCAAs). Though not as thorough as direct condition comparisons, t-SNE offers a valuable graphical interpretation of this varied data collection.

Table 4: Compounds used as nitrogen sources in BarSeq experiments. Nitrogen sources indicated by (N) and amino acid dropout conditions indicated by (-).

Section	Experiments included	
Inorganic nitrogen sources and urea	ammonium (N)	nitrite (N)
	nitrate (N)	urea (N)
Proteinogenic amino acids	L-isoleucine (N)	glycine (N)
	L-leucine (N)	L-histidine (N)
	D-lysine (N)	L-alanine (-)
	L-lysine (N)	L-arginine (-)
	L-methionine (N)	L-asparagine (-)
	L-ornithine (N)	L-aspartate (-)
	L-pipecolic (N)	L-cysteine (-)
	L-proline (N)	L-glutamate (-)
	L-threonine (N)	L-glutamine (-)
	L-phenylalanine (N)	glycine (-)
	L-serine (N)	L-serine (-)
	L-valine (N)	L-histidine (-)
	4-guanidinobutyrate (N)	L-isoleucine (-)
	D-alanine (N)	L-leucine (-)
	L-alanine (N)	L-lysine (-)
	L-arginine (N)	L-methionine (-)
	L-asparagine (N)	L-proline (-)
	L-aspartate (N)	L-threonine (-)
	L-cysteine (N)	L-tyrosine (-)
	L-glutamate (N)	L-tryptophan (-)
	L-glutamine (N)	L-valine (-)
Quaternary amines + ethanolamine	betaine (N)	choline (N)
	carnitine (N)	ethanolamine (N)
Purines and pyrimidines	adenine (N)	hydantoin (N)
	cytosine (N)	uracil (N)
Lactams	butyrolactam (N)	5-oxoproline (N)
	caprolactam (N)	valerolactam (N)
	γ-aminobutyric (N)	
Polyamines, ω-amino acids, and GABA isomers	1,3 diaminopropane (N)	6-aminocaproic acid (N)
	1,6-diaminohexane (N)	β-alanine (N)
	DL-2-aminobutyrate (N)	cadaverine (N)
	DL-3-aminoisobutyrate (N)	putrescine (N)
	3-aminobutyrate (N)	spermidine (N)
	5-aminovalerate (N)	
Not applicable	nicotinic acid (N)	

Figure 3: Global analysis of the *P. putida* KT2440 BarSeq data. (A) Significant genes (|fitness| > 1 and |t | > 5) from all 71 tested conditions sorted by their cluster of orthologous groups (COGs) based on the eggNOG database. "Multiple COGs" indicates that there was more than one COG assigned. (B) Image of the interactive t-SNE visualization (available at https://ppnitrogentsne.lbl.gov) showing the legend, t-SNE clustering (left) and cluster centroids (right). By clicking on a substrate in the legend, the corresponding cluster (left) and centroid (right) is highlighted, opening a list of cluster members and additional information. By clicking the highlighted centroid (right), the user is redirected to the Fitness Browser (https://fit.genomics.lbl.gov), where the fitness data for all significant genes in the condition cluster is shown.

3.1.3.2 Global effectors of nitrogen metabolism

Mutations within the global regulators NtrB (PP_5047) and NtrC (PP_5048) displayed distinct phenotypes across the examined conditions. The NtrBC system in *P. putida* is a two-component enzyme that controls the transcription of multiple nitrogen assimilation genes.[166]

Under conditions of nitrogen deficiency, the activity of the GlnK protein (PP_5234) is modulated by the uridylyltransferase activity of GlnD (PP_1589). After that, NtrB is activated by GlnK, which in turn impacts the phosphorylation status of NtrC. Following this, NtrC adjusts the expression of its associated regulons.[166] Though the mutant library employed for these experiments lacks insertions in *glnK*, the other three components of the signaling cascade are present in the dataset and reside within the same t-SNE-determined cluster (ppnitrogentsne.lbl.gov). The specific reasons why certain nitrogen sources elicit a pronounced NtrC reaction (Figure 4) remain unclear. Generally, NtrC's function is to combat nitrogen deficiency by turning on most of the nitrogen assimilation genes, including transporters.[166]
Establishing a fitness score threshold at > −1.5, we found a minimum of 8 out of 52 nitrogen sources that appear to operate independently of a functional *ntrC* gene (Table 5). The metabolic pathways for these compounds seem less reliant on NtrC activation, hinting at unique regulatory mechanisms or the existence of constitutively active transport and degradation routes for these compounds. The fitness patterns observed in mutants of the extracytoplasmic function sigma factor SigX (PP_2088) appear to align with those of *ntrBC* mutants. In earlier research, it was discovered that *sigX* is somewhat essential for growth using D-lysine as a sole source of carbon.[52] While we could not identify a distinct trend in the fitness data for NtrC, GltBD, another important regulator in nitrogen metabolism, shows a direct correlation between the tested conditions and their subsequent metabolites.

Table 5: Nitrogen source conditions in which *ntrC* fitness is above threshold.

Conditions on which NtrC fitness is >-1.5	
Condition	**NtrC fitness**
L-glutamine	-0.27
putrescine	-0.46
L-asparagine	-0.57
ammonium chloride	-0.66
4-aminobutyric acid	-0.86
4-guanidinobutyric acid	-1.30
L-histidine	-1.31
adenine	-1.43

GltB (PP_5076) and GltD (PP_5075) form the glutamine oxoglutarate aminotransferase (GOGAT) complex in *P. putida*, which is crucial for controlling nitrogen assimilation.[174] Notably, *gltBD* exhibits distinct phenotypes across the different conditions we assayed (Figure 4). On the other hand, the glutamine synthetase (glnA, PP_5046), which is the second component of the GS/GOGAT cycle, lacks insertion mutants in our library. This suggests its

potential essentiality during the library's construction, possibly due to glutamine auxotrophy in these mutants. Figure 4 shows the fitness defects we observed for *gltBD* in nitrogen conditions that either generate free ammonium or do not produce L-glutamate.[175] For instance, in contrast to the L-phenylalanine condition (Table 6), *gltBD* displays significant fitness defects in the nitrate (−6.17) and L-serine (−5.95) conditions.

Table 6: Nitrogen source conditions in which *gltBD* fitness is above threshold.

Conditions on which GltBD fitness is >-1.5	
Condition	**GltBD fitness**
L-isoleucine	0.20
butyrolactam	0.12
L-pyroglutamic acid	0.10
L-valine	0.08
L-phenylalanine	0.04
L-aspartate	-0.07
L-glutamic acid	-0.15
L-proline	-0.16
L-glutamine	-0.32
4-aminobutyric acid	-0.45

An exception to these patterns is the nitrogen source, ammonium chloride. While the fitness defects in *gltBD* under this condition are weak (|fitness| < 2), they remain significant (|t| > 5). Previous studies indicate that with high ammonium concentrations (exceeding 10 mM), its assimilation leverages the glutamate dehydrogenase (GdhA) PP_0675 and GltBD simultaneously.[166,176,177]

In our dataset, the regulatory elements *gacS* (PP_1650) and *gacA* (PP_4099) exhibit pronounced fitness phenotypes. This two-component system shares homology with the extensively researched *barA*/*uvrY* system of *E. coli* and has been recognized as a global regulator of cellular physiology across a variety of organisms.[178,179] The conditions that induced the strongest negative fitness phenotypes for *gacSA* include β-alanine (−1.8), L-alanine (−1.3), spermidine (−2.0), and propanediamine (−7.4). GacSA's regulatory targets are the *rsm* noncoding RNAs, which control the activity of translational repressors, leading to widespread shifts in gene expression.[160,180,181] As a result, GacSA might play a vital role in regulating parts of the metabolism of β-alanine, L-alanine, propanediamine, and spermidine.

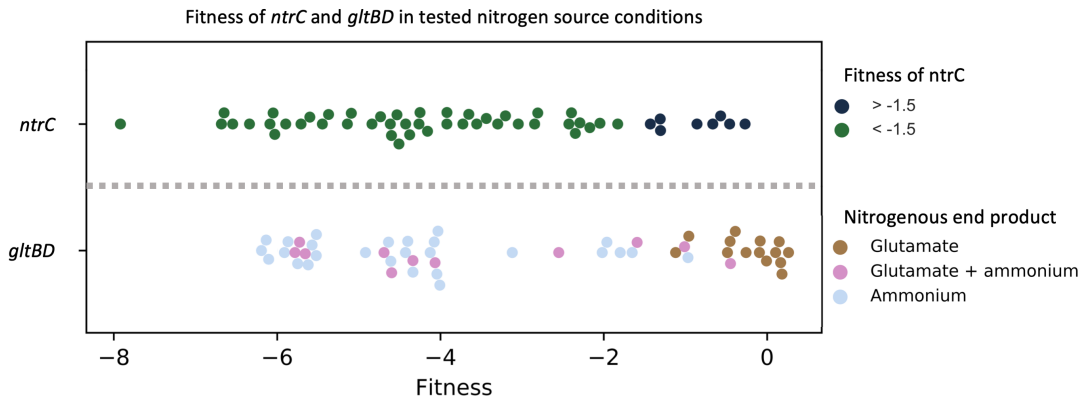


Figure 4: Categorization of fitness phenotypes in global regulators for nitrogen assimilation. Scatterplot shows the average fitness values ($n = 2$) for *ntrC* and *gltBD* in all the tested nitrogen conditions. For *ntrC*, nitrogen conditions are grouped based on whether the fitness phenotype of *ntrC* is < -1.5 (blue) or > -1.5 (green). Conditions where *ntrC* fitness is > -1.5 (green) may be less dependent on *ntrC* activation and are shown in Table 5. *gltBD* phenotypes are sorted based on putative glutamate (brown), glutamate+ammonium (pink), or ammonium (gray) release during nitrogen source utilization. Fitness values for conditions resulting in *gltBD* fitness > -1.5 are shown in Table 6.

3.1.3.3 Inorganic Nitrogen Sources and Urea

P. putida KT2440 exclusively depends on oxygen for its metabolic activities and lacks the ability to utilize other electron acceptors during oxidative phosphorylation.[182] Therefore, oxidized nitrogen species are the only possible assimilation form. While many bacteria favor ammonium as their primary source of inorganic nitrogen,[183] some, like *P. putida*, can also process forms like nitrate and nitrite through specific nitrate reduction mechanisms.[184,185] In the context of pseudomonads, a key regulatory enzyme for this is the two-component system NasS/T (PP_2094 and PP_2093).[185–187] Without oxidized nitrogen sources, NasS and NasT form a complex, leading to the inhibition of nitrate and nitrite reductases. However, in the presence of nitrate or nitrite, NasS separates from the NasS/T complex, allowing the RNA-binding anti-terminator, NasT, to facilitate the complete translation of the nitrate reduction genes (Figure 5) [188].

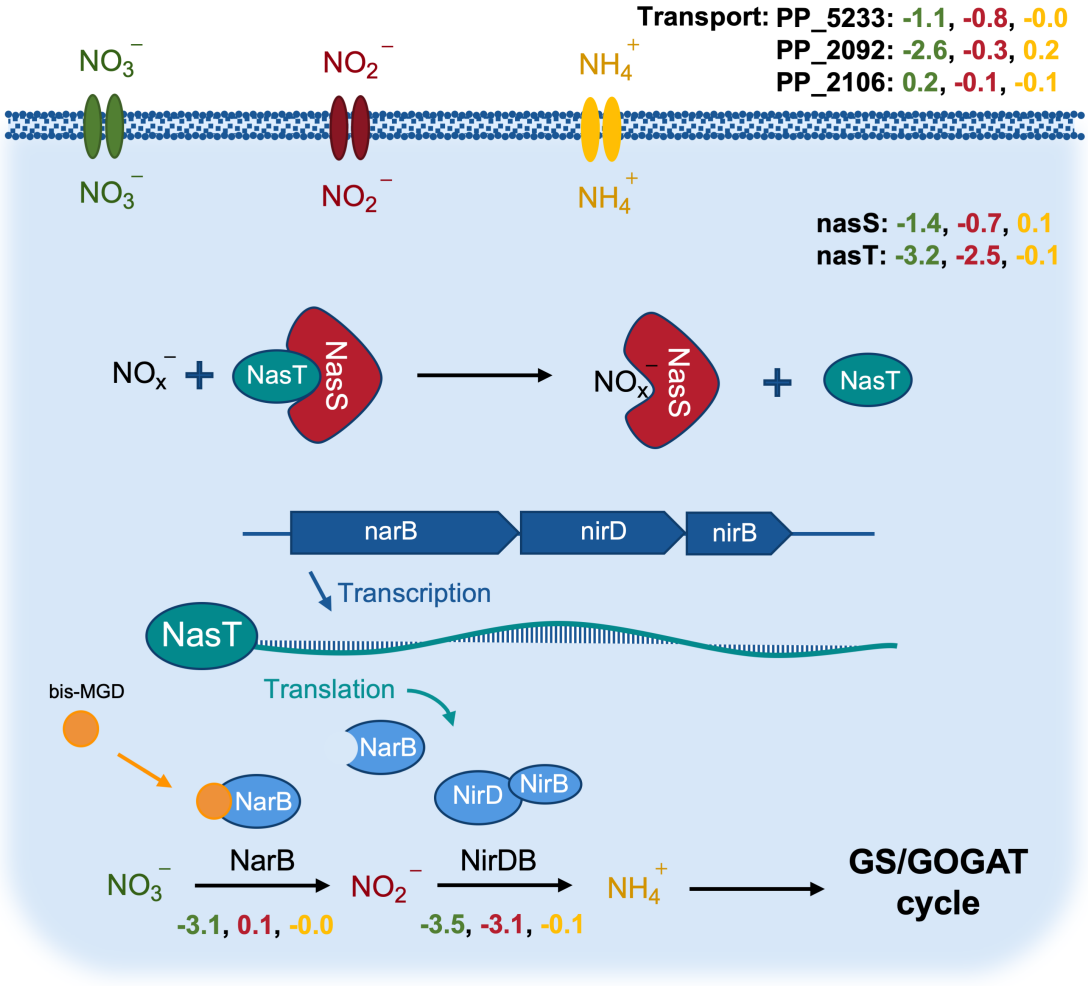


Figure 5: The assimilatory nitrate reduction system in *P. putida* KT2440. Average fitness values ($n = 2$) exhibited in the nitrate (green), nitrite (red), and ammonium (yellow) sole nitrogen source experiments. Shown are putative transporters, action of the NasST regulatory system, the assimilatory pathway, and the role of the bis-molybdopterin guanine dinucleotide (bis-MGD) cofactor. bis-MGD is the required cofactor for NarB.[189] The fitness phenotypes for bis-MGD biosynthesis cluster together with the nitrate phenotypes and can be found in the interactive t-SNE visualization (available at ppnitrogentsne.lbl.gov).

The BarSeq-based growth assay verified the presence of this system in *P. putida*. Under nitrate conditions, the unique growth characteristics of *nasT* and *nasS* mutants are indicated by the significant fitness defects of -3.2 and -1.4, respectively. While studying this system in *Paracoccus denitrificans* PD1222, only the strain lacking *nasT* showed a specific growth phenotype when either nitrate or nitrite was the only nitrogen source.[190]

The reduced fitness of *nasS* in *P. putida* implies that *nasS* plays a role in gene expression of *nasT*. In *Azotobacter vinelandii*, research has indicated that the *nasS* and *nasT* genes have minor overlaps, potentially resulting in the shared translation of neighboring genes.[188,191,192]

However, according to the latest annotations of the *P. putida* KT2440 genome,[193–195] no overlap between *nasS* and *nasT* could be identified. Given the genetic context of the *nasST* sequence, where *nasS* is placed before *nasT*, any alterations in *nasS* could potentially influence the expression of *nasT*.[196]

With BarSeq, we were also able to pinpoint the unique nitrate transporter PP_2092 with a fitness value of -2.6. This gene did not exhibit any noticeable phenotype in the presence of nitrite, hinting that it might be a transporter exclusive to nitrate. Given the absence of any transporters showing strong fitness values under nitrite conditions, there is a possibility that either PP_2092 or some unknown transporters manage nitrite transport. It is also worth noting that, under conditions with pH values below 7.2 and concentrations exceeding 100 µM, nitrite is commonly observed to enter the cell passively through the diffusion of nitrous acid. This could account for the missing fitness data related to this transport process.[197]

In general, the nitrogen source nitrite primarily displays fitness patterns indicating a stress-induced response. Nitrite, in its unbound acidic state, possesses known antimicrobial traits. Its potential impacts cause a range of stresses, from DNA damage and destabilization of the proton motive force to harmful nitrosylation of cofactors and proteins.[198] Our collected fitness data reveal that any disturbance to the recently discovered RES-Xre toxin-antitoxin (TA) system (PP_2433-4) affects growth involving nitrite. Within this mechanism, the toxin (PP_2434) quickly breaks down NAD+ to slow bacterial growth, giving the organism an opportunity to recalibrate its central metabolism. Concurrently, the antitoxin (PP_2433) works to neutralize the toxin, facilitating the replenishment of NAD+ concentrations.[199,200199,200] The notable fitness decline of the antitoxin PP_2433 (-6.4) suggests two possibilities: either the RES-Xre system plays a role in *P. putida*'s reaction to nitrite stress, or nitrite interferes with the cell's redox stability. Any disruption in the redox equilibrium due to the actions of the PP_2434 toxin seems fatal for strains that do not possess a working PP_2433 antitoxin. Additionally, we detected phenotypes linked to various pathways that metabolize NAD(P)H, further indicating a potential depletion of NAD+ under nitrite exposure. For instance, certain genes within the NADH-quinone oxidoreductase operon, spanning PP_4119 to PP_4131, exhibit significant phenotypes when exposed to nitrite, especially PP_4131 (-1.65) and PP_4120 (-1.6).

P. putida seems to use a specific strategy to metabolize glucose for maintaining redox balance under nitrite conditions. Our data suggests that it favors oxidizing glucose to 2-ketogluconate instead of phosphorylating gluconate or directly metabolizing glucose.[201–203] Signs of 2-ketogluconate accumulation in nitrate conditions are evident from distinct fitness characteristics

related to the *kguT* transporter (PP_3377; -1.5) and the NAD(P)H-dependent dehydrogenase *kguD* (PP_3376; -2.65). This method of glucose consumption through 2-ketogluconate yields more NAD(P)+ when compared to the other two strategies. Other nitrogen sources producing similar *kguT* phenotypes include 1,6-hexanediamine (-1.3), caprolactam (-1.55), 2-aminobutyric acid (-1.5), and uracil (-1.2).

The initial steps of the assimilatory nitrogen system involve the enzymes NarB and NirDB. Predictably, the *narB* mutants present a fitness defect solely in nitrate conditions (-3.1). Other notable growth defects in the nitrate condition relate to the production of bis-molybdopterin guanine dinucleotide (bis-MGD), a necessary cofactor for NarB.[189] The mutants for bis-MGD production correlate closely with the nitrate phenotypes, as shown in the interactive t-SNE visualization tool (ppnitrogentsne.lbl.gov). As nitrate gets reduced to nitrite, the nitrite reductase NirDB (PP_1705-6) displays significant growth defects in both nitrate and nitrite conditions (-3.5 and -3.1, respectively). Subsequently, the end product, ammonium, is channeled into the GS/GOGAT cycle either through glutamine synthetase GlnA (no insertions) or glutamate dehydrogenase GdhA (insignificant fitness defect).

In our assays, urea, the simplest organic nitrogen source, triggered significant fitness phenotypes only in *gltBD* (-5.45) and *ntrC* (-1.8). Past transcriptome analysis has revealed that NtrC regulates the expression of the urease operon in *P. putida* (PP_2842-9).[166] Despite this, our study detected no significant phenotypes associated with the urease operon. The existence of a secondary pathway for urea utilization through a urea carboxylase appears unlikely. This is supported by previous research demonstrating that organisms with the carboxylase pathway typically do not possess an urease.[204–207] In addition, our investigation did not identify any annotated urea carboxylase. To achieve a more comprehensive understanding of urea metabolism in *P. putida*, further research is required.

3.1.3.4 Proteinogenic amino acids

The microbial metabolism of amino acids holds significant relevance for the biotechnological food and bulk chemical industry, prompting extensive research for several decades.[208] Unlike the *Corynebacterium*, *Pseudomonas* species are much less represented in the industrial production of amino acids.[209] Nevertheless, their remarkable tolerance to organic solvents establishes them as excellent hosts for the production of aromatic amino acid-derived compounds. These compounds include p-coumarate, p-hydroxybenzoate, N-acyl amino acids, and aromatic pigments.[210–215] Beyond this, *P. putida* has been engineered to harbor biosynthetic

pathways for other amino acid-derived compounds, such as monoethanolamine, cyanophycin, and indigoidine.[171,216–219]

We conducted a competitive growth assay using RB-TnSeq to evaluate the 20 proteinogenic amino acids and two D-stereoisomers when used as the exclusive nitrogen source. Additionally, our study incorporated RB-TnSeq data from growth assays where L-arginine, L-histidine, L-lysine, and D-lysine served as the only carbon sources, with ammonium chloride supplied as a nitrogen source. Unfortunately, we could not secure fitness data for tyrosine and tryptophan due to inadequate biomass that prevented DNA extraction and RB-TnSeq analysis. Interestingly, while Radkov and Moe previously reported that *P. putida* could not utilize L-stereoisomers of methionine, threonine, or leucine as the sole carbon and nitrogen sources, our findings indicated *P. putida*'s ability to grow using these as exclusive nitrogen sources.[220] The rate at which the carbon skeletons of these three amino acids enter the TCA cycle may be too slow or ineffective for them to serve as the sole source of both carbon and nitrogen.

Microorganisms frequently have multiple degradation pathways for a single amino acid. This is particularly true for L-arginine, for which five distinct degradation routes have been identified in *P. aeruginosa*.[221–223] Our nitrogen RB-TnSeq experiments verified two such pathways, and the carbon source RB-TnSeq data revealed another. While the deiminase pathway, known to exist in *P. aeruginosa*, was not detected in *P. putida* via RB-TnSeq, the *P. putida* arginine deiminase AraA (PP_1001) has been thoroughly studied *in vitro*.[224]

When L-arginine was used as the exclusive nitrogen source, it resulted in notable growth phenotypes in both the prevalent arginine succinyltransferase (AST) pathway (< -2.7) and the arginine decarboxylase/agmatine deiminase route. While L-ornithine can be integrated into the AST pathway, the only significant growth deficiency we noticed under this condition was associated with the dehydrogenase AstD (PP_4478; -2.3).[225,226] Intriguingly, an insertion in the succinyltransferase (PP_4479-80), needed to direct L-ornithine into the AST pathway, showed a positive growth effect (1.5). Given these results, the degradation of L-ornithine via the AST pathway might not be the most efficient. Alternatively, direct deamination by the cyclodeaminase PP_3533 (-0.85) and the transformation of proline to glutamate (PP_4947; -5.15) appear to be more advantageous.

The existence of an arginine decarboxylase/agmatine deiminase (ADAD) pathway in *P. putida* is suggested by the growth deficiency linked to the arginine decarboxylase PP_0567 (-0.75). The amidase family protein PP_2932, displaying a pronounced fitness defect (-1.85) when exposed to arginine, suggests its role in acting on N-carbomylputrescine within the

ADAD pathway. Thus, it might be more appropriately labeled as a N-carbomylputrescine amidase.

When L-arginine functions as the primary carbon source, pronounced growth patterns associated with the arginine:pyruvate transaminase (APT) pathway emerge. The components of this pathway include the transaminase AruH (PP_3721, -1.2), the decarboxylase AruL (PP_3723, -0.9), the dehydrogenase KauB (PP_5278, -1.6), and a guanidinobutyrase (PP_4523, -1.5). Notably, the presumed guanidinobutyrase PP_4523 displays a significant fitness defect when L-arginine is the carbon source, and similarly when 4-guanidinobutyric acid is used as the nitrogen source (-0.95).

The role of AruH in acting on the D- or L-stereoisomer of arginine, or potentially both, remains uncertain due to a minor fitness decline (-1.0) observed in the presumed alanine racemase (Alr) PP_3722 and its positioning within the APT operon. A prior study found that the removal of *alr* resulted in a substantial reduction in the growth rate of *P. putida* when arginine was the only carbon and nitrogen source.[227] This could suggest the existence of another metabolic pathway for arginine that operates through its D-stereoisomer.

The availability of carbon and nitrogen might influence the preference between pathways. Given that the AST pathway needs succinate to transform arginine into glutamate, it is less optimal under conditions with limited carbon. On the other hand, products from the other pathways can seamlessly integrate into the TCA cycle through the degradation of putrescine and γ-amino butyric acid (GABA).[221]

The anticipated histidine-lysine-arginine-ornithine ABC transport system (PP_4483-PP_4486) primarily displays notable phenotypes ($|fitness| > 2$; $|t| > 4$) in the arginine condition, whereas the L-ornithine condition results in more moderate growth deficiencies (-0.44). When analyzed through t-SNE, these genes seem to group with those involved in the AST pathway (ppnitrogentsne.lbl.gov). A separate transporter, PP_5031, appears to be crucial for histidine transport (-0.55). Surprisingly, PP_5031 shows a considerably stronger phenotype when histidine is the sole carbon source (-4.5). This variation might underline the potential limitations of our nitrogen source RB-TnSeq studies in identifying particular transporters.

The catabolism of lysine and alanine reveals an interesting interplay between D- and L-stereoisomers. *P. putida*'s lysine catabolism has been thoroughly investigated, and was comprehensively characterized by Thompson *et al*., who employed both D- and L-lysine as carbon sources in their RB-TnSeq experiments.[52,220,227–232] By combining our nitrogen source data with these previous RB-TnSeq datasets and comparing them to each other, we can determine

the differences between D- and L-stereoisomer catabolism and utilization as carbon or nitrogen source (Figure 6).

When L-lysine is used as a carbon source, both pathways for the two enantiomers seem necessary, as evidenced by similar fitness phenotypes in the genes from both routes (Figure 7). Additionally, when D-lysine serves as the carbon source, both routes exhibit fitness defects as well. However, the L-lysine pathway seems less crucial.[52] The significance of both pathways for both enantiomers could be attributed to a lysine-depended or constitutively expressed *alr* in *P. putida*. The resulting enzyme can then transform the given lysine isomer to its counterpart.[220,233] The same pattern emerges when either lysine enantiomer is utilized as the sole nitrogen source. An exception is the semialdehyde dehydrogenase PP_0213, which presents a more negative fitness value for L-lysine (-0.86). This discrepancy could arise from varied racemase activity under different nitrogen conditions or potential toxicity from accumulating downstream aldehydes.

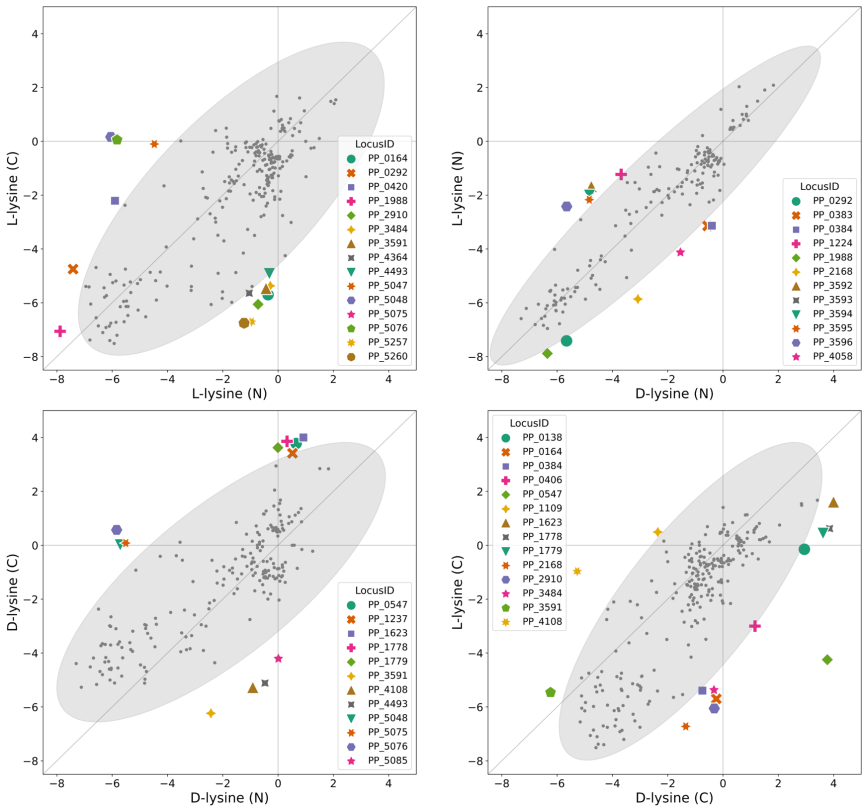


Figure 6: Mahalanobis distance and outlier detection for the comparison of all the tested lysine conditions. Prior to analysis, fitness scores were filtered by t-values ($t > |4|$). Outliers are defined as data points outside of the 95% confidence interval of the distribution, marked by the grey ellipsoid.

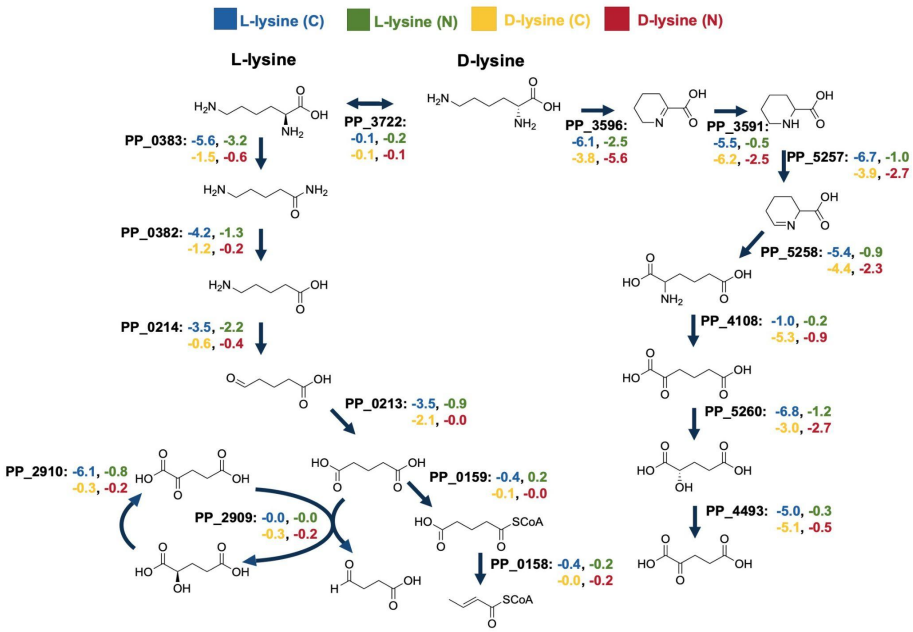


Figure 7: D- and L-lysine degradation pathways and fitness values. Color indicates whether fitness value is from L-lysine carbon source (blue), L-lysine nitrogen source (green), D-lysine carbon source (yellow), or D-lysine nitrogen source (red) experiments.

The primary distinction between the carbon and nitrogen source datasets lies in the genes PP_4493, PP_0213, and *gltBD*. Genes PP_0213 and PP_4493 play a crucial role in channeling the carbon skeleton of D- and L-lysine into the tricarboxylic acid (TCA) cycle, making them vital under carbon-scarce situations. As the products of these genes are fully deaminated, their significance in nitrogen-restricted environments is minimal, with fitness values of -0.41 and -0.39 respectively. The significant phenotype for *gltBD*, with a value of -5.8, can likely be attributed to the extraction of free ammonium during the initial deamination phase.
D-alanine, together with D-glutamate, is crucial for the production of microbial peptidoglycan, which is vital for the growth and survival of bacteria.[234] In *P. putida*, DadX seems to be the primary racemase for alanine.[220,235] Our alanine RB-TnSeq data implies that the conversion process is essential for the survival of *P. putida*, which explains the absence of *dadX* (PP_5269) mutants in our mutant collection. While the *dadX* counterpart *alr* might also have the capability to convert alanine, its negligible growth phenotype indicates that it likely has a limited role in alanine racemization.[220,227,235]
Moreover, when using either L- or D-alanine as the sole nitrogen source, both alanine transaminase homologs do not display any growth deficiencies. Notably, the only pronounced growth defect under these conditions is seen in the deaminating D-amino acid oxidoreductase PP_5270 (DadA). Interestingly, its fitness defect is mainly with L-alanine as the nitrogen

source (-3.6), while for D-alanine, the defect is only mild (-0.65). The fact that D-alanine deamination still relies on PP_5270 is further evidenced by the significant growth defect associated with the transcriptional regulator PP_5271 (-2.05).

The gene PP_5271 shares a significant similarity (93% identity with an E-value of < 9e-88) with the regulator *lrp* from *P. aeruginosa* PA14. In *P. aeruginosa* PA14, *lrp* acts as a transcriptional activator for the *dad* operon, essential for L-alanine breakdown. The regulator is primarily triggered by L-alanine but also shows a modest response to D-alanine and L-valine.[236] A negative fitness defect in a regulator gene such as *lrp* (-2.05) also confirms its role as an activator.[237]

Pseudomonads are renowned for their capacity to break down and produce aromatic compounds.[238] Thus, it is to expect that they have developed a diverse range of aminotransferases that act on aromatics.[239,240] However, our research reveals that L-phenylalanine is the only aromatic amino acid that *P. putida* can utilize as its sole nitrogen source. The pathway leading to the breakdown of L-tryptophan, known as the kynurenine pathway, is not found in *P. putida*, although, it has been reported in the opportunistic pathogen *P. aeruginosa*.[241] The inability of *P. putida* to utilize L-tyrosine as a nitrogen source under our test conditions remains unclear. Transportation does not appear to be a bottleneck, as growth in strains with a L-tyrosine auxotrophy can be restored by the addition of the amino acid.[240] Furthermore, our data indicates that the conversion of phenylalanine into tyrosine by *phhA* (-2.85) might occur prior to deamination. The only aminotransferase with a significant phenotype that we could identify in the L-phenylalanine nitrogen source condition is PP_3590 (-4.0). From the data, we cannot conclude if PP_3590 has a higher substrate specificity for tyrosine, requiring prior conversion of phenylalanine, or if it can act on phenylalanine itself. Furthermore, our data indicates that the conversion of phenylalanine into tyrosine by *phhA* (-2.85) might occur prior to deamination. The only aminotransferase with a significant phenotype that we could identify in the L-phenylalanine nitrogen source condition is PP_3590 (-4.0). From the data, we cannot conclude if PP_3590 has a higher substrate specificity for tyrosine, requiring prior conversion of phenylalanine, or if it can act on phenylalanine itself. The gene product of PP_3590 is labeled as the D-lysine aminotransferase AmaC, even though it has a limited function in D-lysine catabolism.[232]

Our BarSeq assays indicate that PP_3590 possesses a wide substrate range, as fitness defects emerge on various substrates. For instance, it displays a fitness decline of -1.25 on L-pipecolate, showing potential L-2-aminoadipate transaminase activity. Its counterpart, PP_1972, seems to exert minimal influence on the growth of phenylalanine (-0.35) or pipecolate (-0.15).

Earlier research found that a combined deletion of PP_3590 and PP_1972 did not result in an inability to grow without phenylalanine or tyrosine, highlighting the diversity of aromatic aminotransferases in *P. putida*.[240]

There are many intersections between the aromatic amino acid and BCAA metabolism in bacteria.[242] Our findings suggest that the aromatic aminotransferase PP_3590 plays a role in the transamination of L-isoleucine (-1.75) as well. However, the primary transaminase responsible for BCAA degradation and synthesis is PP_3511 (*ilvE*). It exhibits a pronounced fitness defect (< -4) in nearly all examined conditions, including when exposed to ammonium chloride (-5.55). This underlines the significance of PP_3511 in minimal media and its distinct contribution to the BCAA metabolism. Strains deficient in *ilvE* require either gene complementation or the addition of all three BCAAs (valine, isoleucine, leucine) to achieve growth in basic media.[243,244] A comparable pattern is noticed with the histidinol-phosphate aminotransferase PP_0967, which proves crucial for all assessed conditions, except histidine. Interestingly, in a prior study, PP_0967 has been used to catalyze the deamination process of L-phenylalanine to phenylpyruvate.[215]

Many of the pathways involved in the amino acid catabolism in *P. putida* have either been already identified or accurately predicted based on similarity models. Yet, through the use of BarSeq, we produced experimental evidence for the pathways of an additional eight amino acids beyond those highlighted in this discussion. To deepen our knowledge of amino acid metabolism in *P. putida*, we also incorporated drop-out growth assays for every proteinogenic amino acid. By providing 19 out of the 20 protein-forming amino acids, we established conditions where the production of amino acids becomes essential for optimal growth. Due to the sheer amount of functional genomics data for this section, we refer to the Fitness Browser (fit.genomics.lbl.gov) and the interactive t-SNE visualization for more details (ppnitrogentsne.lbl.gov).

3.1.3.5 Quaternary amines and ethanolamine

The nitrogen source choline is a compound with three methyl groups and a positive charge, often found in ionic liquids utilized for breaking down lignin.[245] In their natural environment, most bacteria metabolize choline, using it as a precursor to betaine, a vital substance for osmoprotection.[246] Even though the pathways for synthesizing these compounds are infrequently found in bacteria, the mechanisms for transporting and degrading quaternary amines are widespread in microorganisms.[247,248]

Choline, betaine (also known as trimethylglycine), and carnitine are believed to be metabolized through overlapping routes. Specifically, carnitine and choline enter the betaine metabolism after being broken down by a thiolase and alcohol oxidase, respectively.[249] (Figure 8a). After that, betaine is demethylated to produce glycine. Interestingly, while there is a significant phenotype when using glycine as the sole nitrogen source (< -4), the system that metabolizes glycine (consisting of PP_0986, PP_0988, and PP_0989) is not essential for growth under conditions with quaternary amines (> -0.3). Instead, it is likely that the glycine formed during betaine processing turns into serine through the activity of the hydroxymethyltransferase PP_0322. This is underlined by the pronounced fitness changes observed under the influence of carnitine, choline, and betaine (< -2). While the L-serine dehydratase PP_3144 is crucial for growth when L-serine is the only nitrogen source (-2.3), it is not necessary for growth in the presence of the evaluated quaternary amines. This implies that other serine dehydratases, like PP_0297 or PP_0987, or different deaminating enzymes might be active under these conditions to derive ammonium from serine. Rather than relying on the glycine cleavage system, it is more effective to convert it into serine. This is because the needed methyl-group donor, 5,10-methylenetetrahydrofolate, is produced during the demethylation reaction catalyzed by PP_0310-1.[250]

From the BarSeq assay outcomes and sequence similarity, we have identified the preliminary reactions of carnitine metabolism. The degradation of carnitine starts with the formation of betainyl-CoA which is then further degraded by the thioesterase PP_0301, dehydrogenase PP_0303, and dehydratase PP_0302 (Figure 8a). Previously, the carnitine metabolism in *P. aeruginosa* was thought to involve CoA activation of oxidized carnitine.[251] However, a later enzymatic study of a PP_0303 equivalent indicated its role in acting on oxidized carnitine and acetyl-CoA, leading to the production of acetoacetate and betainyl-CoA.[252]

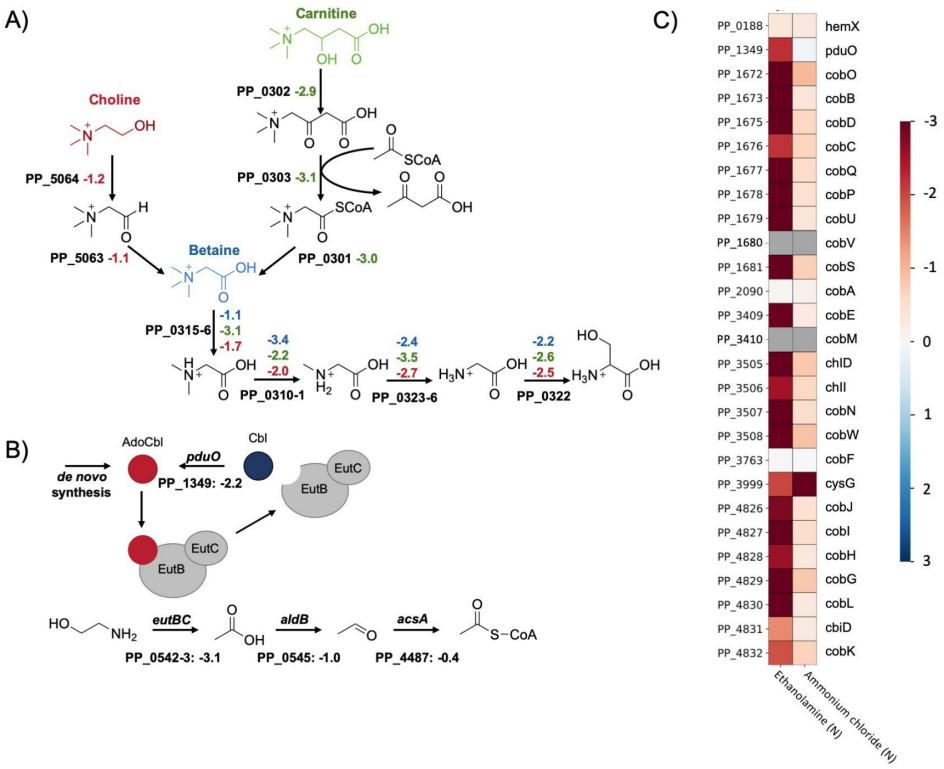


Figure 8: Quaternary amine and ethanolamine degradation in P. putida. A) Putative routes for the quaternary amine catabolism in P. putida KT2440. The figure shows the degradation of choline (red), carnitine (green) and betaine (blue). The corresponding average fitness scores (n=2) are shown next to each gene. B) Ethanolamine degradation pathway, shown with fitness values (n=2) and regeneration of the AdoCbl cofactor. C) Heatmap with average fitness scores (n=2) of genes that are putatively involved in P. putida's adenosylcobalamin biosynthesis. No fitness scores (grey) could be obtained for the genes PP_1680 (cobV) and PP_3410 (cobM).

Additionally, we found fitness defects for transporters and regulators linked with quaternary amine metabolism. While all three quaternary amines induce a pronounced NtrC reaction, which tends to overshadow individual transporters, the choline/betaine/carnitine ABC transporter PP_0294-6 displays fitness decreases in the choline (-1.13) and carnitine (-2.55) conditions. The degradation of betaine and upstream metabolites is controlled by the repressor BetI (PP_5719) and the activators GbdR (PP_0298) and CdhR (PP_0305)).[247,253] For all three nitrogen sources, *gbdR* is a crucial regulator (< -1.0). Conversely, *cdhR* is exclusive to carnitine (-2.5), and *betI* is choline-specific (1.15). Furthermore, the negative and positive fitness values align with the presumed mechanism for these regulators.[253] The considerable fitness defects of PP_0308-9 and their function in this operon remain unclear. A transcriptomic study suggested that they might play a role in promoting filamentous growth.[254]

Ethanolamine is frequently found in nature and is involved in the plant metabolism of choline and serine.[255,256] Hence, it is to be expected that many pseudomonads can utilize it as carbon and nitrogen sources.[257] The degradation of ethanolamine in bacteria can be achieved through acetyl-CoA or via ethanol.[257,258] The initial step in both these pathways is driven by the adenosylcobalamin-dependent ethanolamine ammonia-lyase (EAL) EutBC.[258–260] In the case of *P. putida*, a significant fitness phenotype for *eutBC* (PP_0542-3; -3.1) emerged when using ethanolamine as the sole source of nitrogen (Figure 8b). Additionally, the essentiality of its cofactor adenosylcobalamin (AdoCbl) was also validated. Given that our minimal media lacks supplemented AdoCbl and the absence of distinct phenotypes for the corrinoid-specific transport system PP_0524-5, *P. putida* appears competent in the *de novo* production of AdoCbl.[134,261–263] In fact, we detected particularly strong phenotypes in potential *cob* genes during the ethanolamine growth assay, accompanied by a fitness deficit in the adenosyltransferase PduO (PP_1349; -2.2), which facilitates the adenylation of cobalamin (Cbl) to AdoCbl (Figure 8c).[264] The aerobic synthesis of AdoCbl involves the dephosphorylation of adenosylcobalamin 5'-phosphate without producing Cbl.[264–266] However, studies have indicated that the deamination process of ethanolamine leads to the deadenylation of AdoCbl, consequently producing Cbl.[267] The process of re-adenylating Cbl using PduO is probably more energy-efficient than fully synthesizing AdoCbl.

The subsequent steps of the ethanolamine degradation pathway align closely with the gene cluster recently outlined in *P. aeruginosa*.[257] In addition to *eutBC*, the *eut* operon (PP_0542-6) in *P. putida* includes a transporter (PP_0544; -0.6), an aldehyde dehydrogenase (PP_0545; -0.95), and a transcription factor (PP_0546; -0.95). The negligible phenotype for PP_0544 is likely due to the activation of nonspecific transporters, as evidenced by the significant *ntrC* (-3.95) fitness defect.

3.1.3.6 Purines and pyrimidines

Purines and pyrimidines are chemical structures that are commonly found in nature, with nucleobases, as part of DNA and RNA, being one of their most prevalent forms. Given their high nitrogen content, numerous organisms utilize nucleobases as sources of nitrogen. In our research, we examined the purine base adenine and the pyrimidine bases cytosine and uracil as nitrogen sources using BarSeq.

Much like several other purines, the initial steps in the breakdown of adenine and guanine lead to the mutual intermediate xanthine (Figure 9a). Adenine transitions to xanthine through the intermediate hypoxanthine, whereas guanine is directly transformed into xanthine by the

guanine deaminase PP_4281.[268] The minor fitness defect observed in the adenosine deaminase PP_0591 (-0.45) could suggest its potential activity with adenine. Another possibility is the transformation of adenine to adenosine through adenosine monophosphate (AMP). Yet, the associated adenine phosphoribosyltransferase and nucleotidase homologs display no noticeable fitness defect under these conditions.

The products resulting from the deamination of purine nucleobases, namely hypoxanthine and xanthine, are subsequently transformed by the xanthine dehydrogenase complex *xdhABC* (PP_4178-80) to produce urate. *XdhABC* exhibits a minor fitness defect of -0.3, which might be attributed to the existence of a functionally similar enzyme such as PP_3309 and PP_3310 or PP_4234. In the past, the consensus was that the urate oxidation process produced (S)-allantoin in a single reaction.[269] However, a recent study highlighted that this reaction requires an additional two enzymes, proceeding through the intermediate 5-hydroxyisourate and producing hydrogen peroxide in the process.[270]

As of today, there has not been a report of a functional urate oxidase in *P. putida*. The only predicted urate oxidase, PP_3099, exhibits no growth phenotype across the tested conditions. A possible indirect marker for this reaction could be the pronounced fitness defect observed in the transcription factor PP_2250 (-2.35) and the conserved membrane protein, PP_2251 (-2.25). PP_2251 shares some sequence similarity with the proteobacterial antimicrobial compound efflux (PACE) transporters A1S_1053 (57%; 6e-45) from *Acinetobacter baumannii* and PFL_4585 (64%; 5e-56) from *P. protegens* Pf-5. Therefore, PP_2251 could function as a PACE transporter as well, possibly serving as a defense mechanism against hydrogen peroxide or any other peroxide radicals produced during urate oxidation.[271,272]

The subsequent steps of converting 5-hydroxyisourate to (S)-allantoin exhibited significant fitness defects during BarSeq assays. Initially, 5-hydroxyisourate undergoes hydrolysis catalyzed by PucM (PP_4285; -0.9), resulting in the formation of 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline (OHCU). Subsequently, PucL (PP_4287; -2.7) decarboxylates OHCU, producing the enantiomeric compound (S)-allantoin.[270] The final steps of degrading (S)-allantoin into glyoxylate and urea, are illustrated in Figure 9a.

The catabolism of the pyrimidine base cytosine starts with the extraction of the amine group, catalyzed by the cytosine deaminase CodB (PP_3189; -1.2). Additionally, we verified the importance of the designated cytosine transporter CodA (PP_3187; -1.2). Interestingly, the same deaminase and transporter exhibit fitness defects when in the presence of 3-aminobutyric acid (-1.15). The expression of the *codBA* gene cluster, as well as numerous other

transport mechanisms for nitrogen-containing compounds, is regulated by the global transcription factor NtrC.[166,273]

Since the deamination product of cytosine is uracil, we anticipated similar fitness values across both conditions. Yet, fitness phenotypes for subsequent uracil degradation phenotypes were significantly less pronounced (Figure 9b). For instance, during the initial phase of uracil catabolism, PydAX dehydrogenates uracil to produce 5,6-dihydrouracil. With uracil as the nitrogen source, pydAX (PP_4037-8) exhibits a strong fitness defect of -3.38, while the cytosine condition causes a mild defect of -0.53. This data implies that repeated deaminations of the same nitrogen source sometimes yield less pronounced phenotypes. During periods of nitrogen starvation, limited resources are probably directed more towards pathways that support survival and growth, instead of enzymes responsible for further breaking down degradation by-products.

The enzyme PydB (PP_4036; -2.8) facilitates the ring-opening process of 5,6-dihydrouracil, producing the amide (S)-ureidoglycolate. This amide is then degraded either by *hyuC* (PP_4034; no significant fitness change) or its potential homolog, PP_0614 (no significant fitness change). The following hydrolysis releases free ammonia, carbon dioxide, and β-alanine, with the production of the latter being evidenced by the significant fitness defect of PP_0596 (-6.35). Neither of the genes responsible for hydrolyzing (S)-ureidoglycolate present a fitness defect in these conditions, which might be attributed to their overlapping functionalities. Interestingly, the only notable growth impact related to *hyuC* was observed in the context of the nitrogen source hydantoin.

Hydantoin was identified as a product of allantoin reduction[274] and its derivatives have been employed as anticonvulsants and pesticides.[275,276] Given its xenobiotic nature, *P. putida*'s ability to process hydantoin likely stems from its structural resemblance to 5,6-dihydrouracil. The pyrimidine permease PP_4035 (-1.05) seems to play a role in transporting hydantoin, and the dihydropyrimidinase (PydB) that interacts with 5,6-hydrouracil is probably also responsible for opening its 5-membered ring (-1.15). The final ring-opening reaction results in carbamoylglycine. In experiments using hydantoin as the sole nitrogen source, the amidohydrolase *hyuC* shows a fitness defect of -1.0, marking the most pronounced fitness defect observed for this gene. This might be attributed to PP_0614 (-0.15) not recognizing substrates with a shorter chain length. Bacteria are known to convert R-substituted hydantoins into the corresponding D-amino acid.[277] While the hydantoin and glycine conditions did not show any overlapping phenotypes in their metabolisms, it is likely that carbamoylglycine hydrolysis produces glycine, ammonia, and carbon dioxide.

While we lack data on thymine, we conducted BarSeq assays using a racemic blend of 3-aminoisobutyric acid (3-AIBA) as the nitrogen source. The D-stereoisomer of this mixture is an end product during thymine degradation.[278] Currently, in *P. putida*, no aminotransferases exhibiting D- or L-3-AIBA activities were identified. The observed fitness defects associated with the pyruvate transaminase PP_0596 (-6.3) and dehydrogenase PP_0597 (-1.7) are likely due to the presence of the L-stereoisomer, which is an intermediate in the degradation of valine.[279]

An additional sign that PP_0596 may possess 3-AIBA transaminase activity is the observed fitness defect in the transcriptional regulator *lrp* (-2.15) and the deaminating D-amino acid oxidoreductase PP_5270 (-1.4). As previously mentioned, these genes belong to the *dad* operon, which plays a role in alanine metabolism and can thus indicate pyruvate transaminase activity.[236]

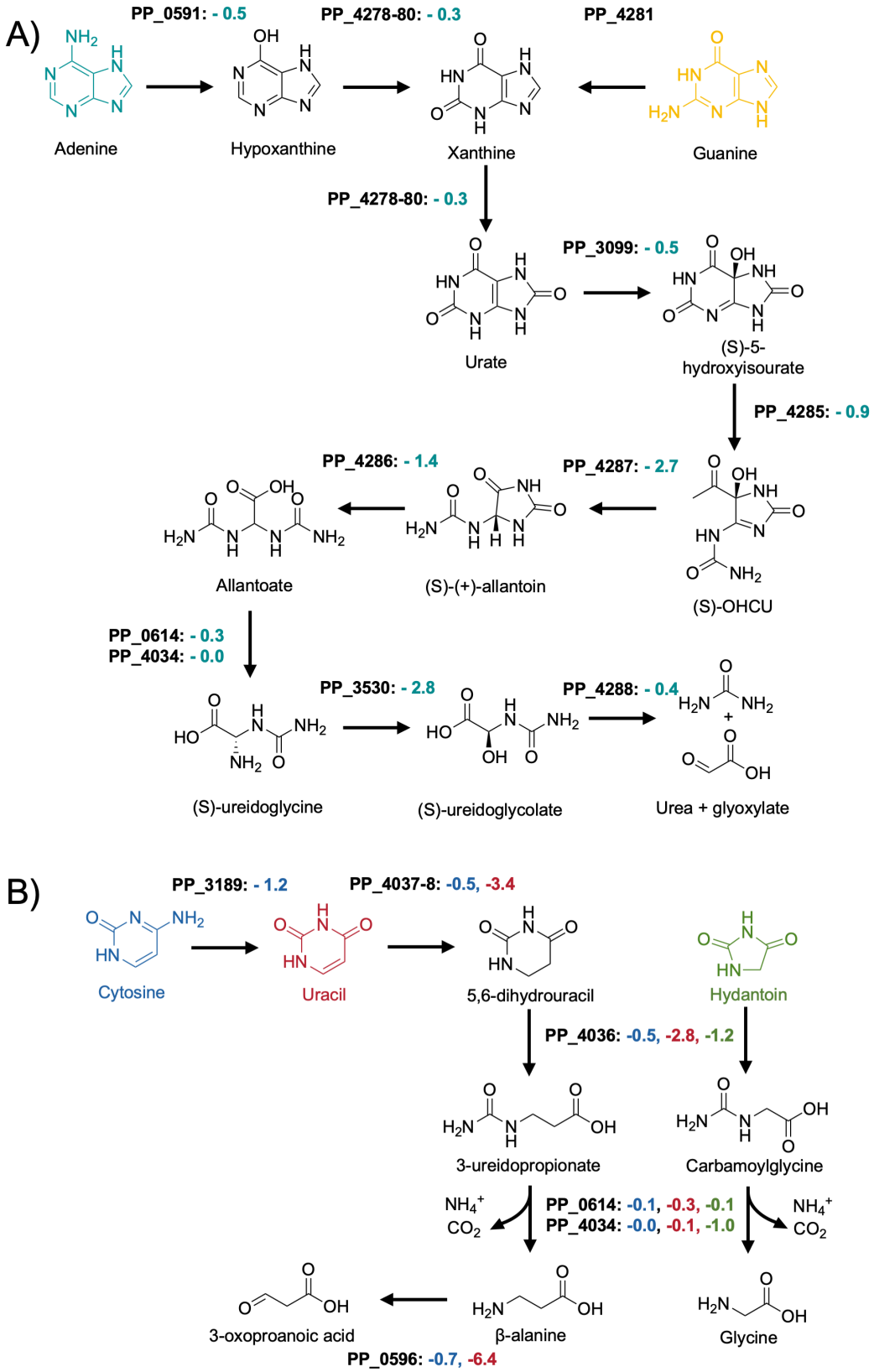


Figure 9: Putative routes for purine (A) and pyrimidine (B) catabolism in *P. putida* KT2440. Shown are the average fitness scores (n=2) for genes involved in adenine (teal), guanine (yellow), cytosine (blue), uracil (red) and hydantoin (green) degradation.

3.1.3.7 Lactams

Lactams are of significant industrial importance and have diverse uses in producing solvents, plastics and pharmaceutical precursors.[280] Given that their production typically relies on petrochemical processes, there has been much interest in research exploring biological methods to produce lactams and related compounds.[280] However, to harness bacteria for lactam production, it is crucial to investigate the metabolic pathways involved with lactams and their precursors.[56]

In our *P. putida* BarSeq experiments, we assayed four distinct lactams as nitrogen sources. This led to the identification of two more lactam hydrolases, PP_2920-2 and PP_4575-7, in addition to the previously known OplBA.[56] Our fitness data suggests that these enzymes play a key role in degrading butyrolactam, 5-oxoproline, valerolactam and caprolactam. When we deleted these hydrolases, *P. putida* could not grow using the corresponding lactam as the sole source of nitrogen. However, growth was restored when we introduced a plasmid with the lactam hydrolase gene (Supplemental Figure 1).

The three components of PP_2920-2 and PP_4575-7 are identified by UniProt as homologs of *pxpABC*, which is a widely distributed prokaryotic 5-oxoprolinase.[281] 5-oxoprolinases were primarily identified in eukaryotes and certain bacteria with a γ-glutamyl cycle. Niehaus *et al.* proposed that prokaryotes needed a mechanism to counteract the natural cyclization of glutamate and glutamine into 5-oxoproline, leading to the discovery of *pxpABC*.[281] Based on our results, it seems that only PP_4575-7 targets 5-oxoproline. *P. putida* has a wide repertoire of CoA ligases and unintended CoA activation and cyclization might be a major burden. The development of additional, more specific oxoprolinases could have led to an evolutionary advantage.

The ω-amino acids GABA and 5-aminovaleric acid (5AVA) are intermediates, showing up in the breakdown processes of arginine, putrescine, and lysine.[52,282,283]

Past research has shown that these C4 and C5 ω-amino acids can undergo cyclization after being activated to an acyl-CoA thioester.[56,280,284] Without a hydrolase, these lactams would end up as metabolic "cul-de-sacs", leading to a decline in available carbon and nitrogen, and therefore, growth. To investigate if these lactamase systems function in this manner, we performed growth assays with ΔPP_4575-7, ΔPP_2920-2, and ΔOplBA using the respective ω-amino acid of each lactam as a nitrogen source (Supplemental Figure 2). Notably, the knockout strains showed minor growth delays and reduced maximal ODs compared to the WT. These findings could support our initial hypothesis for the evolution of these diverse oxoprolinases.

The hydrolases PP_4575-7 and PP_2920-2, responsible for 5-oxoproline and butyrolactam respectively, show a significant fitness defect on their substrates, with an average of -3.82 and -4.45 across their subunits and replicates. The ring-opening reaction of 5-oxoproline yields glutamate, whereas butyrolactam hydrolysis produces the ω-amino acid GABA. The degradation of GABA is discussed in the section dedicated to ω-amino acids.

In earlier studies, OplBA, the lactamase responsible for hydrolyzing valero- and caprolactam, was identified using proteomics data.[56] While it was attempted to identify the lactam hydrolase through BarSeq experiments using valerolactam as the sole carbon source, it was not possible to measure *oplBA* fitness defects. Interestingly, in our observations, *oplBA* mutants displayed strong negative fitness phenotypes when valerolactam was the nitrogen source and positive ones with caprolactam (Figure 10a). The positive fitness outcome for OplBA mutants on caprolactam is quite unexpected, given that previous studies indicated that *oplBA* knockout strains could not grow using caprolactam as their only nitrogen source.[56] This discrepancy might be attributed to the specifics of BarSeq experiments. Library mutants with a functional copy of *oplBA* convert caprolactam into 6-aminocaproic acid (6ACA), which leads to the accumulation of this compound in the supernatant. Consequently, mutants where the *oplBA* genes are deleted can utilize the liberated 6ACA without the need to allocate resources for producing OplBA, especially under nitrogen limited conditions.

The limited utilization of caprolactam, and consequently 6ACA, as a nitrogen source might be attributed to *P. putida*'s inability to process the dicarboxylic end product of 6ACA transamination.[285] We hypothesized that adipic acid was accumulating in the supernatant, which was later confirmed through LCMS analysis. We used caprolactam as the sole nitrogen source and sampled the supernatant at 1, 6, 12, 24, 48, and 72 hours. After 48 hours, caprolactam was undetectable in all our samples. Simultaneously, adipic acid steadily increased at a rate approximately matching biomass growth (Figure 10b). A more detailed discussion about the degradation of 6ACA can be found in the section focusing on ω-amino acids.

A potential reason for the difference in fitness values of *oplBA* between conditions using valerolactam as a carbon and nitrogen source could be the growth dynamics of *P. putida*. While there is no significant data for *oplBA* mutants when valerolactam serves as the carbon source, the aminotransferase targeting 5AVA, PP_0214, exhibits a strong fitness defect (-4.1). It is likely that once valerolactam is converted to 5AVA, it diffuses across the media, effectively being distributed amongst the library members. This distribution would make a functional copy of *oplBA* optional. However, such a distribution would only be possible if *P. putida*'s rate of utilizing 5AVA as a carbon source lags behind its diffusion through the media.

In the experiments using valerolactam as a nitrogen source, mutants with a functional OplBA enzyme might metabolize the 5AVA quicker than it can disperse in the media. This could lead to *oplBA* mutants experiencing nitrogen deficiency, which might account for the observed negative fitness value.

In close proximity to each of the three lactamases are three regulatory genes: *oplR*, PP_2919, and PP_4579. The observed fitness data implies that these transcription factors trigger the expression of their adjacent lactam hydrolase when its specific lactam substrate is present. However, the fitness data does not clarify whether the specificity for the various lactams is due to the transcription factor, the lactam hydrolase, or both. To resolve this, we studied the substrate specificity of the lactam hydrolases using a complementation assay and assessed the specificity of the transcription factors with RFP reporters.

For each of our three hydrolase-deficient strains, we tried to reestablish growth on its respective lactam using plasmid-driven expression of the three distinct lactam hydrolases, leading to a total of 9 strains being assessed. If a lactam hydrolase had broader specificity than its regulator, it should have been able to restore growth in strains with a different lactamase deleted. However, growth was only restored by the plasmid expression of the lactam hydrolases under the anticipated nitrogen source condition, underlining the specificity of these lactam hydrolases (Supplemental Figure 3).

The transcription factor *oplR* was previously confirmed as the regulator of *oplBA* and has been adapted into a set of highly sensitive valero- and caprolactam biosensor plasmids.[58] However, based solely on the fitness data, we cannot conclude that the other neighboring regulators trigger lactamase expression. To determine the function of each transcription factor, we carried out a two-plasmid system assay in *E. coli*.[58] The transcription factor was expressed by an arabinose-inducible promoter in a low-copy vector, while the 200 bp sequence upstream of the lactamase was positioned in front of RFP in a medium-copy plasmid (BBR1). We refer to these plasmids as the sensor and reporter plasmid, respectively. We then varied the induction level of the transcription factor and the concentration of the anticipated lactam ligand in a combinatorial manner. The normalized RFP increased with rising lactam concentrations, and the dynamic range shifted with transcription factor induction (Supplemental Figure 4).

These data suggest that the transcription factors indeed regulate the expression of adjacent lactam hydrolases when exposed to the respective lactams. However, the extent of this induction was minimal, possibly due to these transcription factors not functioning optimally in *E. coli*. Consequently, we chose to examine lactam regulation specificity in *P. putida*.

Unfortunately, the transcription factor plasmids we used are not compatible with *P. putida.* Moreover, substituting the lactam hydrolase with RFP in the genome did not yield the desired dynamic range for our assay (data not shown). As an alternative approach, we tested the ligand specificity of these transcription factors by introducing our reporter plasmid into *P. putida*, relying on the endogenous expression of the transcription factors.

The butyrolactam and pyroglutamate systems both showed significant inducer specificity, with inducing RFP expression only when their anticipated ligand was present. The dynamic range for our reporter plasmid systems was notably broad, demonstrating significant sensitivity in both cases. Interestingly, at the lowest butyrolactam concentration tested (50 μM), the butyrolactam reporter exhibited an RFP signal approximately 900 times greater than the non-induced state. This setup holds promise as a potent biosensor for measuring butyrolactam production in *P. putida*. The OplR reporter plasmid also showed a response to butyrolactam. Prior research demonstrated OplR's high specificity to valero- and caprolactam.[58] However, the maximum tested concentration for butyrolactam was 0.5 mM, and the study was conducted in *E. coli*. In our experiments, we noticed fluorescence when butyrolactam concentrations exceeded approximately 6.25 mM. Yet, as we depend on the endogenous transcription factor native levels, it is unclear if this observation results from OplR reacting to butyrolactam or if there is interference between the PP_2919 and the *oplBA* promoter (Supplemental Figure 4).

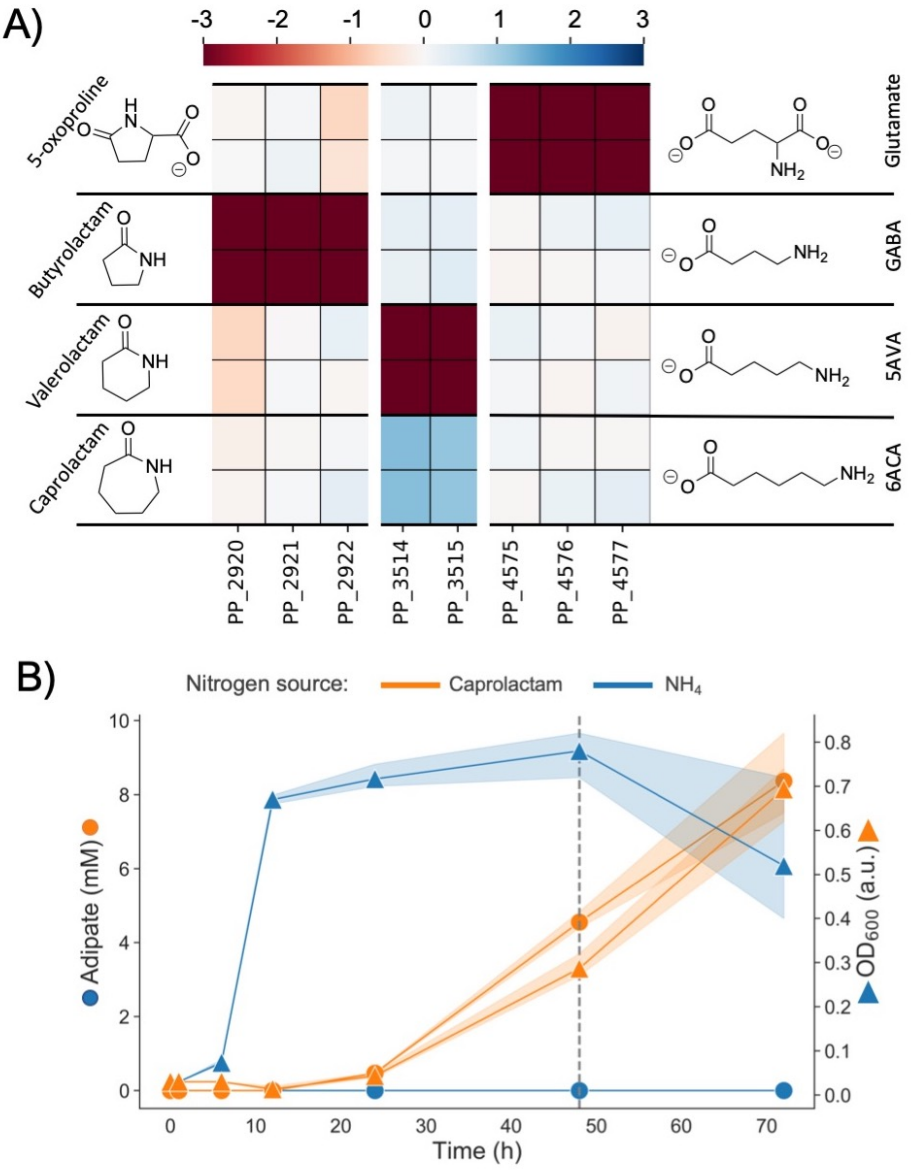


Figure 10: Lactam BarSeq results and LCMS analysis of caprolactam degradation in *P. putida* KT2440. A) Heatmap with fitness scores (n=2) of genes putatively involved in the hydrolysis of caprolactam, valerolactam, butyrolactam and 5-oxoproline. B) LCMS analysis of caprolactam degradation. Wild type cells were grown in MOPS minimal media with ammonium (blue) and caprolactam (orange) as the sole source of nitrogen. Shown is the OD (squares) and concentration of adipate in the supernatant (circles) over a time course of 72 hours. The dashed line marks the time point at which caprolactam was no longer detected in the media.

3.1.3.8 Polyamines, ω-amino acids, and GABA isomers

Polyamines are present in almost all life forms and play a role in various cellular processes, such as gene regulation, stress adaptation, cell development, and membrane homeostasis.[286]

However, not all species use polyamines in the same way, which led to extensive research on their significance in prokaryotes.[287,288]

P. putida detects these molecules using the McpU (PP_1228) chemoreceptor, which directs the cell towards putrescine, cadaverine, and spermidine. When McpU is mutated, the organism struggles to colonize roots.[289] Apart from their biological roles, polyamines are important industrial chemicals. For instance, putrescine is a primary component for nylon production, while propanediamine enhances textile finishes and strengthens materials.[290,291]

The ω-amino acids, which are often degradation products of polyamines, are significant in both biology and industry. They emerge as intermediates in metabolic pathways, like the degradation of proteinogenic amino acids, and some, including GABA, serve as signaling molecules.[292,293] On the industrial site, compounds like 6ACA, 5AVA, and GABA are foundational for the synthesis of various polymers and plastic products.[280]

The three critical transaminases for the utilization of polyamines, ω-amino acids, and GABA isomers seem to be the pyruvate:alanine aminotransferases PP_0596, PP_2180 (SpuC-I), and PP_5182 (SpuC-II). While they present fitness defects for polyamines and amino acids, it is challenging to discern if any of them directly act on polyamines since the tested polyamines transform into ω-amino acids. As a result, we isolated these enzymes to study their substrate specificity in a controlled environment. Interestingly, each of the aminotransferases demonstrated some activity to almost all the tested substrates, even those without a significant fitness phenotype (Table 7).

Table 7: Substrate range of purified pyruvate:alanine aminotransferases. Tested substrates were added in excess to pyruvate. Activity of the aminotransferases on the tested substrates was measured indirectly through the quantification of the co-product alanine. Percent conversion is relative to the expected measurement of alanine if the amine acceptor pyruvate was completely transformed. Error

represents the standard deviation of three replicates. NA indicates that no activity was statistically detected above controls

Substrate Class	Substrate	% conversion by PP_0596	% conversion by PP_2180	% conversion by PP_5182
Polyamines	1,3 diaminopropane	65 ± 0.9 %	95 ± 0.7 %	96 ± 3.5 %
	putrescine	2 ± 0.7 %	109 ± 1.4 %	112 ± 1.8 %
	cadaverine	9 ± 4.2 %	97 ± 1.8 %	103 ± 8.5 %
	1,6 diaminohexane	14 ± 1.5 %	85 ± 4 %	88 ± 0.6 %
	spermidine	101 ± 1.1 %	105 ± 5.8 %	112 ± 1.2 %
ω-amino acids	β-alanine	51 ± 0.5 %	NA	NA
	γ-aminobutyric acid	75 ± 2 %	76 ± 1.5 %	83 ± 4.2 %
	5-aminovaleric acid	70 ± 2.5 %	71 ± 1.4 %	73 ± 0.9 %
	6-aminocaproic acid	78 ± 1.9 %	82 ± 0.8 %	84 ± 1 %
GABA isomers	DL-2-aminobutyric	55 ± 1.3 %	95 ± 1.2 %	100 ± 5.6 %
	3-aminobutyric acid	87 ± 3.1 %	12 ± 1.6 %	16 ± 3.8 %
	DL-3-aminoisobutyric	70 ± 2.4 %	6 ± 1.3 %	12 ± 2.5 %
Proteinogenic amino acids	L-isoleucine	NA	NA	1 ± 0.5 %
	L-leucine	NA	5 ± 0.7 %	NA
	L-lysine	59 ± 49.4 %	61 ± 3.8 %	65 ± 4.9 %
	L-valine	NA	NA	NA

The broad substrate range of these enzymes differs significantly from the specificity shown in the fitness data. For instance, even though PP_2180 or PP_5182 do not display a fitness defect with 1,3-diaminopropane (1,3-DAP), they both exhibit more activity towards this substrate than PP_0596, which does show a significant fitness defect. The discrepancy between the fitness and biochemical data likely stems from the difference in promiscuity between these enzymes and their respective regulators. To further probe our theory, we developed a fluorescent reporter assay.

With a two-plasmid system, using RFP as a readout, we examined combinations of transcription factor and aminotransferase promoters in *E. coli*. Based on our fitness data and recent DNA Affinity Purification and sequencing (DAP-Seq) findings from *P. aeruginosa*, we hypothesized that the transcription factor PP_5268 might control PP_5182.[294] However, when assessing the sequences upstream of PP_5182 or PP_5183 as promoters, we observed no fluorescence signal (data not shown). This could suggest that the transcription factor-promoter systems we analyzed are not compatible in *E. coli*, or perhaps PP_5268 does not regulate PP_5182.

However, we successfully identified the transcription factors that probably regulate the aminotransferases PP_0596 and PP_2180 (Supplemental Figure 5). Located next to PP_0596 is the LysR-type transcription factor PP_0595. The fitness data of this regulator aligns closely

with the phenotypes of PP_0596. When we used the two-plasmid system with this combination, it functioned as expected. Interestingly, PP_0595 seems to activate the promoter in the presence to both 3AIBA and β-alanine (Supplemental Figure 5).

The fitness data of the aminotransferase PP_2180 contrasts with the adjacent MerR-family transcription factor PP_2181. This is often indicative of transcriptional repressors. Initially, when we paired PP_2181 with the region 200 base pairs upstream of PP_2180 as a promoter, there was no significant fluorescence signal (data not shown). Considering PP_2180 being part of an operon, we used the region upstream of PP_2177 as a promoter, which then led to a fluorescence signal in the presence of 1,6-diaminohexane (1,6-DAH) and cadaverine (Supplemental Figure 5). Interestingly, these transcription factors displayed higher specificity than the *in vitro* substrate specificity of the aminotransferases under their control. This shows that in *P. putida*, the functions of PP_0596 and PP_2180 are largely influenced by regulatory mechanisms.

Polyamines can be used as a nitrogen source through two pathways: direct transamination by an aminotransferase or via the γ-glutamylation (Figure 11). Yet, since the product of the γ-glutamylation process is an ω-amino acid, which subsequently becomes a substrate for an aminotransferase, it is uncertain which pathways are utilized solely based on fitness data.

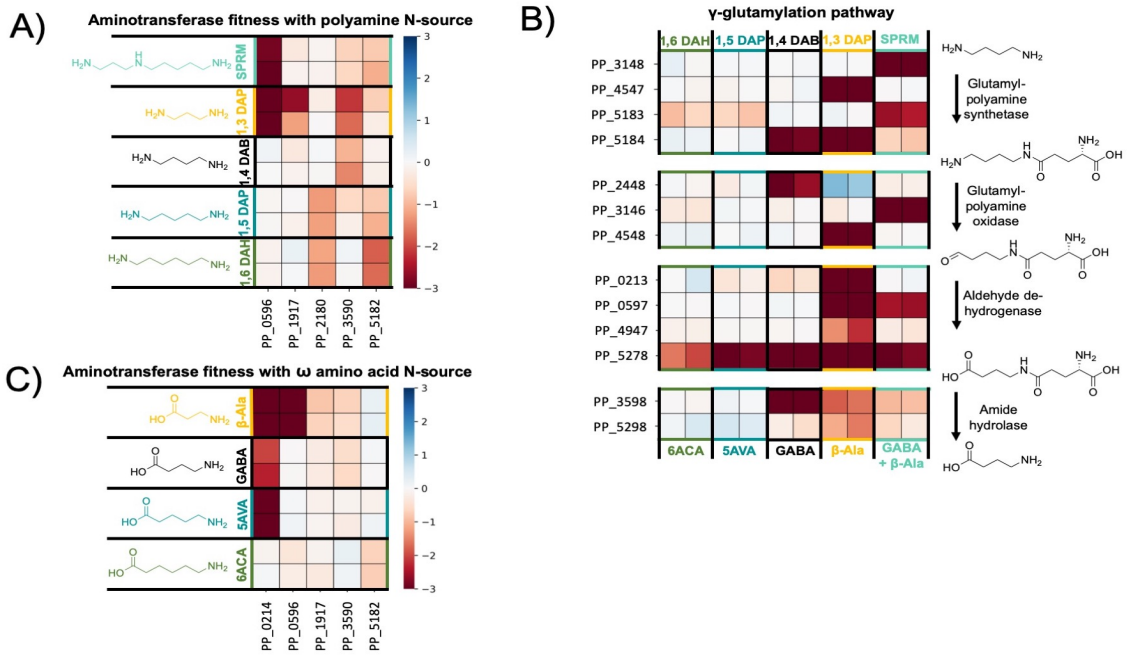


Figure 11: Heatmaps of polyamine and ω-amino acid catabolism in *P. putida*. For each enzyme class shown, HMMER was used to identify all putative genes present in *P. putida* corresponding to that pFam in *P. putida*. Data was filtered to find genes in each class with fitness values >-1.0 and t > |4| in the nitrogen sources shown. A) Heatmap with fitness values (n=2) for genes putatively involved

in polyamine transamination in *P. putida* KT2440. B) Heatmap with fitness values (n=2) for genes putatively involved in polyamine gamma-glutamylation in *P. putida* KT2440. C) Heatmap with fitness values (n=2) for genes putatively involved in omega-amino acid transamination in *P. putida* KT2440. Spermidine (SPRM) = teal; 1,3 diaminopropane (1,3 DAP) = yellow; beta-alanine (BAL) = yellow; putrescine/1,4-diaminobutane (1,4 DAB) = black; gamma-aminobutyric acid (GABA) = black; cadaverine/1,5-diaminopentane (1,5 DAP) = blue; 5-aminovalerate (5AVA) = blue; 1,6-diaminohexane (1,6 DAH) = green; 6-aminocaproic acid (6ACA) = green.

In the nitrogen source conditions of 1,6-DAH and cadaverine, we observed mild fitness defects (< -1.0) for the aminotransferases PP_5182 and PP_2180. As mentioned earlier, both of these aminotransferases react with these substrates *in vitro*. Moreover, the regulator PP_2181 specifically responds to 1,6-DAH and cadaverine, hinting that these molecules might primarily be substrates for PP_2180. Deleting either aminotransferase on its own does not significantly impair growth on these substrates, likely because their functions overlap.

In strains where both PP_5182 and PP_2180 are deleted, we observed minimal growth when cadaverine or 1,6-DAH are the only nitrogen sources (Supplemental Figure 6). This minimal growth could be attributed to the activity of another aminotransferase, such as PP_0596, or activity of the γ-glutamylation pathway. Interestingly, two enzymes from this pathway, the glutamyl-polyamine synthetase PP_5183 and the aldehyde dehydrogenase PP_5278, display minor fitness defects in the cadaverine and 1,6-DAH nitrogen source conditions. However, it seems unlikely that the γ-glutamylation pathway has a major role in the utilization of cadaverine and 1,6-DAH.

On the other hand, the γ-glutamylation pathway seems to be critical for direct transamination and processing of putrescine, 1,3-DAP, and spermidine. Each of these conditions exhibits significant fitness phenotypes at every stage of the pathway. While it is possible that putrescine, 1,3-DAP, and spermidine might also undergo direct deamination by aminotransferases, this seems less likely for putrescine, which does not display strong fitness phenotypes for any of the aminotransferases. The primary glutamyl-polyamine synthetases for 1,3-DAP seem to be PP_4547 (-5.6) and PP_5184 (-4.5); the latter also plays a role in putrescine degradation (-2.9). Intriguingly, even though spermidine contains 1,3 DAP and putrescine, different expected glutamyl-polyamine synthetases, PP_5183 (-2.5) and PP_3149 (-3.2), seem crucial for its degradation.

Each substrate appears to be paired with a unique glutamyl-polyamine oxidase: PP_2448 works on putrescine (-2.9), PP_4548 is effective on 1,3-DAP (-5.6), and PP_3146 targets spermidine (-4.6). Regarding the next step that involves an aldehyde dehydrogenase, the

fitness information is not as definitive. All three substrates introduce noticeable fitness defects (< -3) in PP_5278. However, both 1,3-DAP and spermidine cause defects in multiple other aldehyde dehydrogenases. Specifically, PP_0597 shows a -5.1 defect with 1,3-DAP and -2.6 with spermidine. Furthermore, 1,3-DAP is responsible for fitness defects of -5.5 in PP_0213 and -1.8 in PP_4947.

The concluding enzyme in the γ-glutamylation pathway is the amide hydrolase. The only amide hydrolase we could identify with a fitness defect was PP_3598 (-3.8) in the putrescine condition. Both spermidine and 1,3-DAP led to minor fitness issues for PP_3598 and its counterpart, PP_5298, probably because of their overlapping functions. This amide hydrolase releases an ω-amino acid, which subsequently undergoes transamination to liberate additional nitrogen.

ω-amino acids, whether supplemented directly or produced from polyamine degradation, are broken down through transamination to an aldehyde and then oxidized to form a dicarboxylic acid. 6ACA is the least effectively metabolized ω-amino acid we examined, usually showing negligible to zero growth until approximately 72 hours (data not shown). This slow growth can be attributed to a mix of regulatory challenges and the buildup of its byproduct, adipic acid, which *P. putida* cannot further metabolize (Figure 10b). Although their phenotypes are mild (> -1), the *spuC* homologs PP_5182 and PP_2180 are amongst the most likely aminotransferases involved in the degradation of 6ACA.

When either of these aminotransferases is removed individually, growth on 6ACA is not completely eliminated. However, a strain with both knocked out cannot grow on 6ACA (data not shown). Assuming that the aminotransferases we evaluated *in vitro* exhibit similar substrate flexibility *in vivo* and considering that regulatory challenges may influence the low growth rate on 6ACA, we anticipate that introducing these aminotransferases via plasmid expression would enhance growth. To validate our hypothesis, we reintroduced functional gene copies to knockout mutants of PP_0596, PP_5182, and PP_2180 by using the arabinose-responsive pBADT plasmid containing the relevant gene (Supplemental Figure 7). Unexpectedly, the control strain, which is the WT *P. putida* carrying the pBADT-RFP, could not grow under these specific culture conditions. Our theory is that the stress of upkeeping the plasmid, coupled with inadequate levels of the kanamycin-resistance protein, causes cell death in the early stages when 6ACA serves as the only nitrogen source.[295] However, all of our functionally complemented strains displayed growth on 6ACA. This suggests that regulatory factors are indeed contributing to the slow processing of 6ACA. Moreover, expressing any of the three

aminotransferases, which have confirmed activity on 6ACA, outside of their native regulation can enhance growth.

The catabolism of 5AVA, a byproduct in lysine degradation, is well-established, with fitness data from carbon source RB-TnSeq experiments documented earlier.[52] Initially, 5AVA undergoes transamination by the aminotransferase PP_0214 (-4.1). Subsequently, the resulting aldehyde is transformed to glutarate through the action of the succinate semialdehyde dehydrogenase PP_0213 (-2.1). Glutarate can then follow one of two pathways: it can either be degraded via a glutaryl-CoA intermediate or be converted to succinate. When 5AVA serves as both the carbon and nitrogen source, significant fitness defects are observed for the enzymes involved in the transformation to succinate, specifically glutarate hydroxylase (-1.4) and L-2-hydroxyglutarate oxidase (-2.2). It seems that the conversion to glutaryl-CoA is primarily relevant when 5AVA serves as a carbon source. This is evident from the fitness defects observed for the CoA ligase PP_0159 (-1.3) and the glutaryl-CoA dehydrogenase PP_0158 (-1.7). Conversely, when 5AVA acts as a nitrogen source, this pathway seems less critical, as both PP_0159 and PP_0158 show a slightly positive fitness value of 0.2. This can probably be attributed to the fact that nitrogen has already been liberated and there is an overlapping functionality in the subsequent degradation of glutarate.

The transcription factor PP_0595 is specific to β-alanine and so is the aminotransferase PP_0596 it regulates. Our research also indicates its essential role in the utilization of uracil, 3-aminobutyric acid (3ABA), and 3AIBA (Supplemental Figure 8). When β-alanine is used as a nitrogen source, both genes show a strong fitness defect (-7.8). Located within the same operon is the enzyme methylmalonate-semialdehyde dehydrogenase PP_0597 (-5.9), which facilitates the transformation of malonate-semialdehyde to acetyl-CoA. The dependence on PP_0596, *davT*, and *gltBD* under the β-alanine condition might suggest the existence of a complex metabolic pathway for a more effective β-alanine metabolism. When exposed to β-alanine, DavT might drive the reverse reaction (from succinate-semialdehyde and glutamate to 4-ABA and 2-oxoglutarate), aiming to decrease the intracellular alanine levels generated by PP_0596.

In *P. putida*, GABA can function as the exclusive source for both carbon and nitrogen.[296] Given GABA's abundant presence in root exudates, *P. putida* has evolved to have sensitive and precise receptors that recognize GABA, leading to chemotaxis.[296,297] In contrast to *E. coli*, *P. putida*'s genes responsible for the degradation and transport of GABA are not grouped within a single operon.[229,298] This makes the identification of unique phenotypes challenging due to the largely unspecific fitness defects observed. Nonetheless, the primary

aminotransferase that acts on GABA seems to be DavT (PP_0214; -2.15), which results in the production of succinate-semialdehyde. This compound is subsequently transformed to succinate by the action of DavD.[229] The fitness defect associated with *davD* (-0.45) is not statistically significant (with |t|<4). This hints at the possibility that other dehydrogenases, such as PP_2488 (no fitness data) or PP_3151 (-0.2; with |t|<4), might also catalyze this reaction. Further, our *in vitro* assay indicated that a minimum of three other aminotransferases (namely, PP_0596, PP_2180, and PP_5182) can use GABA as a substrate (Table 7). Even though the second aminotransferase identified by BarSeq, PP_3590, displays a non-significant phenotype (-0.55; with |t| = 3.7), it shows a strong fitness defect when exposed to the DL-2-aminobutyric acid (2ABA) condition (-0.95).

2ABA is a non-natural amino acid used in pharmaceutical production and has been heterologously produced in *E. coli*.[299,300] The only biological function of 2ABA is the replacement of L-cysteine in the synthesis of glutathione, resulting in the formation of ophthalmic acid.[301] From a biosynthetic perspective, 2ABA can be produced during isoleucine biosynthesis from threonine, with the transamination of 2-oxobutyrate by the enzyme IlvE.[300] Given the critical nature of *ilvE* under minimal media conditions, its role *in P. putida* remains unconfirmed using RB-TnSeq. Nevertheless, we successfully identified the D-amino acid oxidoreductase PP_5270, which has a fitness score of -2.75. The gene product of PP_5270 acts on the alanine generated during the transamination of pyruvate.

While no significant phenotype is observed for PP_0596 under this condition, our studies have confirmed that PP_0596 is capable of transaminating 2ABA (Table 7). Moreover, the transamination product, 2-oxobutyrate, serves as an intermediate in the biosynthesis of isoleucine. Rather than generating isoleucine, 2-oxobutyrate seems to be channeled directly into the isoleucine degradation process via the branched-chain α-keto acid dehydrogenase (BKD) complex, as indicated by its fitness score of -1.54. This decarboxylation reaction is further confirmed by the essentiality of the methylcitrate cycle, which involves the genes PP_2334-6 (-1.5). The methylcitrate cycle plays a critical role in metabolizing propionyl-CoA, which can exclusively be produced from 2-oxobutyrate through the BKD reaction.

The final GABA isomer we evaluated was 3ABA, previously discussed in the context of pyrimidine metabolism. 3ABA stands out among several GABA isomers that can trigger plant immunity, and its agricultural application is aimed at preventing crop blight.[302] While the most efficient degradation pathway for 3ABA seems to involve the pyruvate transaminase, PP_0596 (-5.25), the cotranscribed gene, PP_0597 (malonyl semialdehyde dehydrogenase), appears to be detrimental when utilizing 3-ABA (+2.85). This suggests that PP_0597 is not

the ideal enzyme for further processing of its downstream product, acetoacetate. Additionally, the reaction facilitated by PP_0597 is closely related to that of PP_4667. This similarity raises the possibility that the expression of PP_0597 could cause a metabolic imbalance in valine metabolism.

Chapter 3.2

Maximizing Heterologous Expression of Engineered Type I Polyketide Synthases: Investigating Codon Optimization Strategies

Published as:

Schmidt, M.; Lee, N.; Zhan, C.; Roberts, J. B.; Nava, A. A.; Keiser, L. S.; Vilchez, A. A.; Chen, Y.; Petzold, C. J.; Haushalter, R. W.; Blank, L. M.; Keasling, J. D. Maximizing heterologous expression of engineered type I polyketide synthases: investigating codon optimization strategies. *ACS Synth Biol*. 2023;12(11):3366-3380. doi: 10.1021/acssynbio.3c00367.

Contributions

M.S. generated all *P. putida* and *E. coli* strains and performed most of the experiments with the assistance of A.A.V. C.Z. performed *C. glutamicum* engineering. The SAGE system and the strains AG5577 and AG6212 were provided by Adam Guss. The pBH vectors were provided by R.W.H. N.L. performed RT-qPCR and prepared figures with the help of M.S. LCMS analysis was performed by M.S. with the assistance of L.S.K. Y.C. and C.J.P. performed proteomics analysis. PCA analysis and online tool development was performed by A.A.N. and J.B.R. with the assistance of M.S. The manuscript was prepared by M.S. with the help of N.L., R.W.H., L.M.B. and J.D.K. J.D.K. conceived the project.

3.2 Maximizing Heterologous Expression of Engineered Type I Polyketide Synthases: Investigating Codon Optimization Strategies

3.2.1 Abstract

Type I polyketide synthases (T1PKSs) offer vast possibilities for the biosynthesis of unique chemicals. However, despite recent advancements in this field, expressing PKSs heterologously continues to pose challenges. Codon optimization is most often used as a primary measurement to enhance heterologous gene expression. Inappropriate codon optimization strategies, however, can have adversely effects on protein synthesis and production outcomes. In our research, we examined 11 codon variants of an engineered T1PKS, assessing their impact on heterologous expression in *Corynebacterium glutamicum*, *Escherichia coli*, and *Pseudomonas putida*. The top-performing codon variants showed at least a 50-fold increase in PKS protein levels, also paving the way for the synthesis of a novel polyketide in each host organism. Additionally, we have introduced a complimentary online tool at (base-buddy.lbl.gov), providing transparent and customizable codon optimization based on the latest codon usage data. In this chapter, the importance of codon optimization is emphasized, while it also sets the foundation for high-throughput assembly and evaluation of PKSs in alternative hosts.

3.2.2 Introduction

T1PKSs are enzymes predominantly found in bacteria and fungi, playing a crucial role in producing secondary metabolites. Throughout history, humans have leveraged the medicinal value of these metabolites, leading to the discovery of many essential drugs.[118] Furthermore, the modular architecture of T1PKSs enables the iterative assembly of long carbon chains, while also providing the flexibility for the optional incorporation of functional groups.[303] The theoretical design space offered by this modularity has attracted significant research attention, and T1PKSs have been successfully reprogrammed for the production of unnatural polyketides.[132,304,305]

However, one of the biggest challenges in PKS engineering is the native host itself. Most of the discovered PKSs originate from the genus *Streptomyces*, a GC-rich, gram-positive and filamentous bacterium. While certain streptomycetes have been highly domesticated and optimized for the production of larger, high value molecules, their efficacy for the production of industrially relevant bulk chemicals is limited.[306] To date, little progress has been made in exploring alternative PKS hosts. Besides the required genetic modifications for PKS

expression, polyketide titers in non-native hosts are often very low, which can usually be traced back to poor precursor availability or low protein levels.[307,308] Common strategies for improving polyketide titers include the supplementation of the media, the exchange of promoters, or codon optimization.[304] However, the importance of the latter one is often neglected or underestimated.

The host's codon preference usually has a distinct pattern and can be summarized in codon usage tables.[309] The choice of specific codons can impact transcription and translation rates, and is also involved in expression control mechanisms. Well-studied examples include the rare *Streptomyces* codon TTA and the use of alternative start codons such as GTG or TTG.[310,311]

Codon optimization represents a strategy to address these variations in codon preferences during heterologous gene expression. This approach involves selectively substituting specific codons while preserving the amino acid sequence of the protein. There are three commonly used strategies for this purpose: (i) replacing the original codons with the most frequently used codon of the targeted host,[312] (ii) matching the codon frequency of the targeted host[313] and (iii) harmonizing the codon frequency of the targeted host with the codon frequency of the native host.[314] After codon optimization, the final nucleotide sequence will likely be significantly different from the original sequence, and many researchers have to rely on DNA synthesis services to synthesize the codon-optimized gene. Although convenient, the optimization algorithms offered by most synthesis services are not publicly available and their functionality is very limited. The recently published open-source codon optimization tool DNA Chisel offers an easy way to apply these aforementioned methods and further customize the resulting nucleotide sequence.[142]

In this study, we investigated the expression and activity of an engineered T1PKS with different codon variants in the three hosts *Corynebacterium glutamicum, Escherichia coli*, and *Pseudomonas putida*. *E. coli* and *C. glutamicum* are well-established industrial hosts for the large-scale production of proteins and small molecules, while *P. putida* has shown enormous potential for the valorization of renewable feedstocks.[57,315,316] By targeting three heterologous hosts and applying the three most common gene optimizations methods, we designed 9 codon variants of the engineered PKS (Figure 12). Furthermore, as a conventional approach, we obtained and cloned the native sequences from *Streptomyces aureofaciens* Tü117 and *Saccharopolyspora erythraea* NRRL2338, and also tested the effect of two different start codons, GTG and ATG. Due to the large size of PKSs and the metabolic burden of plasmid maintenance, we developed a backbone excision-dependent expression (BEDEX) system to

facilitate the cloning process and enable constitutive expression in the heterologous hosts. Also, we hypothesized that the strong repression of BEDEX vectors would decrease the occurrence of mutations during DNA assembly. We further confirmed the universal functionality of BEDEX vectors in our selected hosts and applied the BEDEX system to heterologously express the 11 codon variants of the engineered PKS. To characterize our codon variants *in vivo*, we measured the PKS protein and transcript levels and demonstrated the production of an unnatural polyketide in each host.

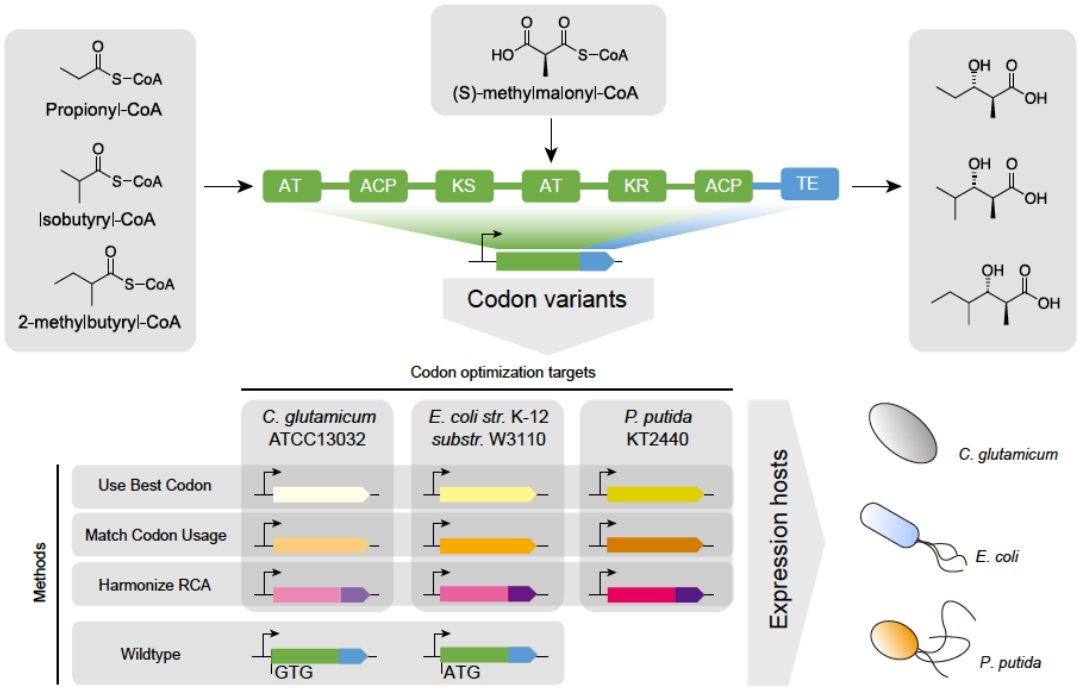


Figure 12: Engineered polyketide synthase with applied codon optimization strategies and targeted heterologous hosts. The loading module and module 1 originates from the lipomycin polyketide synthase (LipPKS) from *Streptomyces aureofaciens* Tü117 (green). The thioesterase domain (TE) originates from the erythromycin PKS (EryPKS) from *Saccharopolyspora erythraea* NRRL2338 (blue). By fusing these two parts together, the engineered PKS design yields a variety of short-chain 3-hydroxy acids. The gene sequence of the reprogrammed LipPKS was codon optimized using the DNA Chisel algorithms for 'Use Best Codon' (yellow), 'Match Codon Usage' (orange) and 'Harmonize RCA' (red/purple). All algorithms preserve the amino acid sequence of the protein. The 'Use Best Codon' method replaces each codon with the most frequently used codon. 'Match Codon Usage' matches the codon frequency of the original codon sequence with the codon usage of the targeted host. 'Harmonize RCA' applies and matches the codon frequency of the targeted host with the codon usage of the native host. The 'Harmonize RCA' algorithm required to codon optimize the LipPKS and EryPKS parts separately. Codon optimizations targeted the three heterologous hosts *C. glutamicum, E. coli*, and *P. putida*.

3.2.3 Results

3.2.3.1 Comparative analysis on codon usage patterns of host organisms

To gain a deep understanding of codon usage across our target host organisms, we examined the codon usage trends in three native PKS hosts, our chosen heterologous host species, and other extensively researched organisms for reference. A phylogenetic analysis based on 16S rRNA sequences of these species revealed that *C. glutamicum* has the closest relation to *Streptomyces*, while *E. coli* and *P. putida* share a closer relationship with one another (Figure 13a). Based on the phylogenetic insights, *C. glutamicum* emerges as the most fitting heterologous host for genes derived from streptomycetes. However, a significant difference in codon usage could exist, given that the GC-content in *Streptomyces* surpasses that in *C. glutamicum* by approximately 20%.

A PCA was conducted on the codon usage data of the aforementioned organisms (Figure 13b). The native PKS hosts, namely *S. aureofaciens*, *S. erythraea*, and *S. coelicolor*, are closely grouped together. In contrast, the evaluated heterologous hosts, *C. glutamicum*, *E. coli*, and *P. putida*, are located farther apart. It is interesting to note that even though *C. glutamicum* is from the same phylum, it does not align with the other actinobacteria. Additionally, the notably higher GC-content in *P. putida* (> 61%) does not seem to make its codon usage more analogous to streptomycetes.

Figure 13c presents the Relative Codon Frequency per corresponding amino acid (RCF) for *C. glutamicum*, *S. aureofaciens*, *E. coli*, and *P. putida*. The RCF in *S. aureofaciens* reveals ten pronounced peaks (> 0.8), with *P. putida* being the only other organism reflecting a similar peak for the TGC codon. The elevated GC-content in *S. aureofaciens* (> 72%) likely results in these pronounced peaks, especially for amino acids represented by just two codons. On the other hand, the proteobacterium *E. coli* generally favors codons with a lower GC-content but appears to be more accommodating of diverse codons as well.

In summary, the data seems to indicate that there is not a perfect industrially relevant host that mirrors the codon preference of typical native PKS hosts. Utilizing PCA for codon usage might offer valuable insights into predicting the success rate of expressing the WT nucleotide sequence in a chosen heterologous host. To get a deeper insight into this hypothesis, we carried out codon optimization on a modified PKS derived from *S. aureofaciens* Tü117 and compared expression levels with the WT nucleotide sequence within our chosen heterologous hosts.

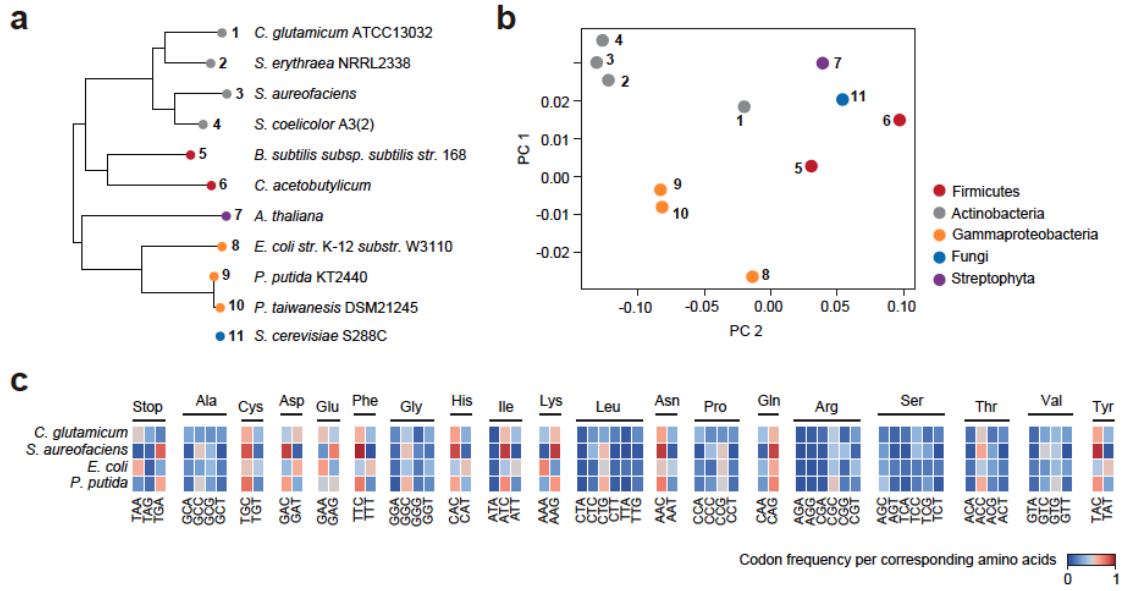


Figure 13: Global analysis of codon usage preferences between different species. (a) 16S rRNA sequence-based phylogenetic analysis of targeted hosts, native polyketide synthase hosts and outgroup references. *Saccharomyces cerevisiae* S288C was excluded. (b) Principal component analysis (PCA) of codon usage tables associated with the same species from the phylogenetic analysis. The first two principal components (PCs), accounting for the highest explained variance, were selected and visualized. (c) Comparison of the codon frequency per corresponding amino acid between *C. glutamicum*, *S. aureofaciens*, *E. coli*, and *P. putida*. The amino acids methionine and tryptophan are encoded by a singular codon and were excluded from this analysis.

3.2.3.2 Codon optimization of engineered lipomycin polyketide synthase

Our synthetic PKS design followed the strategy of Yuzawa *et al.*[119] Briefly, we removed the initial 59 N-terminal amino acids from the LipPKS and cut off M1 post the ACP1. Subsequently, we connected the residual protein to the EryPKS M6 TE, including the interdomain linker between EryPKS ACP6 and TE6 (Figure 12). This N-terminal truncation enhances protein expression, and the TE processes the compound following the mmCoA extension.[317] Notably, the loading domain of LipPKS demonstrates high substrate flexibility, accommodating starter units like ibCoA, 2mbCoA, isovaleryl-CoA, and propionyl-CoA.[119] Given its relatively compact size coupled with its broad substrate specificity, this engineered PKS is an optimal choice for heterologous expression.

Although this design has been proven to produce short-chain 3-hydroxy acids *in vivo*, its effectiveness in our selected heterologous hosts remains unassessed. To heighten the likelihood of achieving functional expression of our engineered PKS, we utilized three codon optimization strategies: ubc, mcu, and hrca through DNA Chisel. We then examined their

impact on the transcription and translation of the target PKS gene.[142] The ubc method aligns with the previously outlined codon optimization technique (i), while mcu and hrca follow the principles of strategies (ii) and (iii) respectively.[142,312–314] Moreover, we incorporated the WT nucleotide sequence for each part of the PKS and tested the influence of the two distinct start codons, GTG and ATG. Comprehensive details of the settings used for codon optimization are shown in Table 3.

Given that the utilized algorithms belong to a Python-based toolkit with a command-line interface (CLI), we developed an intuitive graphical user interface (GUI) to simplify the usage of these open-source codon optimization tools. Our codon optimization tool is openly accessible, demands no prior CLI expertise, and can be visited at https://basebuddy.lbl.gov. Beyond the pre-integrated codon usage database from Kazusa,[309] we have also incorporated the newest version of the CoCoPUTs database.[318] When compared to the Kazusa database, CoCoPUTs contain more recent sequencing data and also supply codon usage tables for a wider selection of organisms.[319]

3.2.3.3 Development of a backbone-excision dependent expression system

To improve the precision and reliability of our findings in *C. glutamicum* and *P. putida*, we employed SAGE to insert our codon variants into the genomes of these hosts.[92,320] The strains AG5577 and AG6212, which are derivatives of *P. putida* and *C. glutamicum* respectively, contain nine distinct *attB* sites within their genomes (personal communication, Adam Guss). These genomic "landing pads" allow for efficient and accurate gene integrations through the SAGE method (Figure 14a).

We also aimed to enhance the utility of the SAGE system by combining it with BEDEX vectors (Figure 14a). The BEDEX vectors primarily comprise two critical components: the LlacO1 promoter and the removable *lacI* gene. Within the SAGE system, the incorporated plasmid backbone is eliminated through the transient expression of the ΦC31 integrase. By discarding the *lacI* gene from the host genome, the *LlacO1* promoter is no longer repressed and becomes constitutively active. To verify the effectiveness of the BEDEX vectors across all three host organisms, we used RFP as an easily measurable indicator (Figure 14b). While we integrated the *rfp* gene into the genomes of *P. putida* AG5577 and *C. glutamicum* AG6212, we used plasmid-driven expression in *E. coli* BL21. Given the presence of the ColE1 origin of replication, the SAGE system's vector collection is not compatible with engineering *E. coli*.

In *C. glutamicum* AG6212, the detected RFP signal was significantly lower than in RFP-expressing *E. coli* and *P. putida* cells. The control sample from *C. glutamicum* showed a normalized RFP signal of 174 ± 4, which is roughly half of the value observed in the uninduced and unexcised integration of the BEDEX vector (data not shown). Nonetheless, either inducing or excising the backbone led to a substantial surge in RFP signal, increasing it by as much as 4-fold.

In *P. putida* AG5577, expression was significantly repressed when *lacI* was present. Induction using 200 µM IPTG resulted in a 44-fold increase in the normalized RFP signal. On the other hand, excising the backbone caused the signal to increase by 213-fold, with the expression appearing unaltered by the presence of the inducer. In the case of *E. coli*, the introduction of IPTG led to a 105-fold increase of the normalized RFP signal, while the non-induced state showed a comparably modest signal, measuring at 1548 ± 287.

These findings validate the effective incorporation of the target gene into the genomes of *C. glutamicum* and *P. putida*. Furthermore, by removing *lacI* from the plasmid backbone, we managed to activate expression without the need to induce the *LlacO1* promoter.

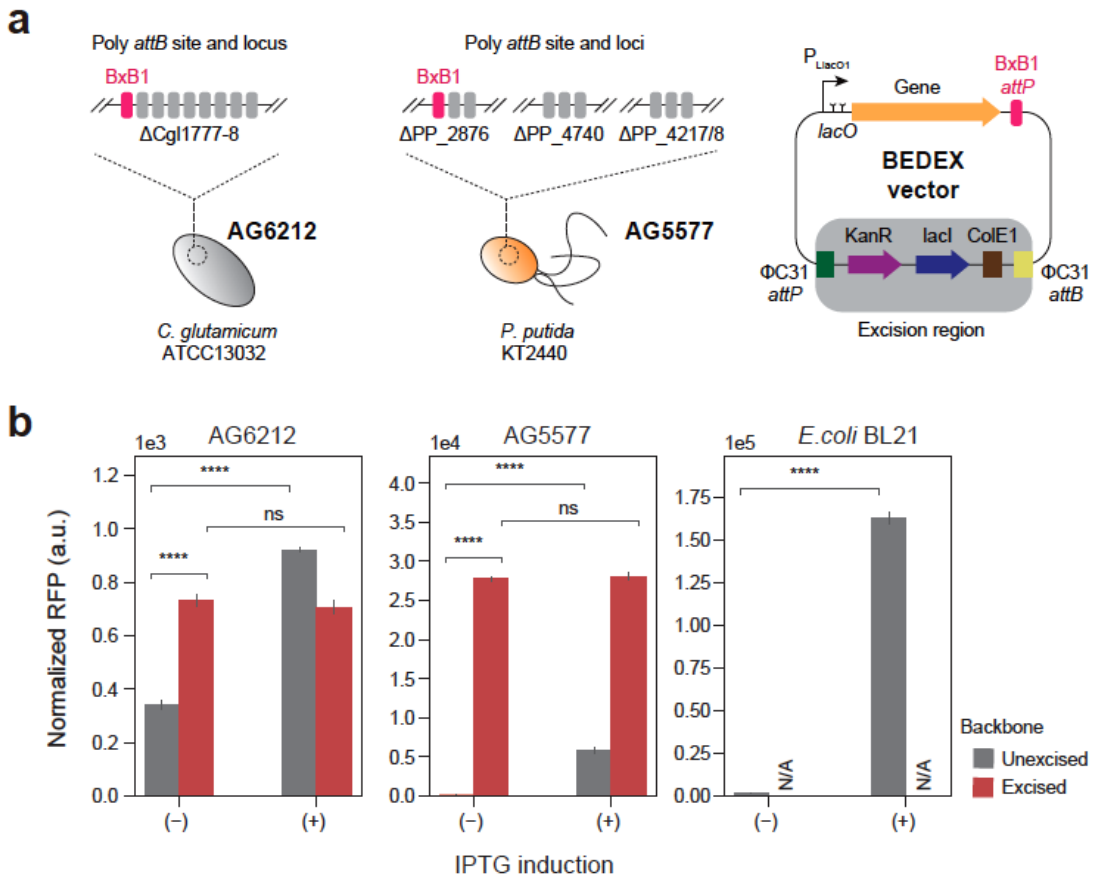


Figure 14: Extending the SAGE system with backbone excision-dependent expression (BEDEX) vectors. (a) *C. glutamicum* AG6212 and *P. putida* AG5577 contain a total of 9 unique *attB* sites each.

Heterologous serine recombinases catalyze the integration of vectors containing the corresponding *attP* site. Expressing the integrase ΦC31 removes the integrated vector backbone and allows for selection marker recycling. Excising the backbone of BEDEX vectors also removes the repressor LacI. (b) *C. glutamicum* AG6212, *P. putida* AG5577, and *E. coli* BL21 containing BEDEX vector carrying RFP. In *E. coli* BL21, BEDEX vectors are non-integrative and maintained by replication. Induction or excision of the BEDEX vector backbone leads to a significant increase in RFP levels (n = 3).

3.2.3.4 Quantification of heterologous protein and transcript

The 11 codon variants of LipPKS, carried by the BEDEX vectors, were transferred into *C. glutamicum* AG6212*, E. coli* BL21, and *P. putida* AG5577. In *C. glutamicum* AG6212 and *P. putida* AG5577, the LipPKS gene was expressed genomically, while in *E. coli* BL21, it was expressed from plasmid DNA. Figure 15a displays the protein levels of the modified LipPKS.

In *C. glutamicum* samples, the LipPKS counts were typically lower by 1 to 2 orders of magnitude compared to those in *E. coli* and *P. putida*. Yet, when focusing on protein abundance, the differences among the organisms become smaller. Depending on the host organism and the codon variant used, LipPKS abundance fluctuates between 0 and 3.9 ± 1.5%. Interestingly, relative to the optimized genes, the WT nucleotide sequences were considerably less effective: their protein levels were either not detectable or vastly lower than those of the codon-optimized versions. However, a mutation changing the initial GTG start codon to ATG led to, on average, a threefold increase in peptide counts.

Typically, PKS genes that are codon-optimized for a specific organism yield the highest PKS levels when expressed in that designated organism. For instance, PKS genes optimized for *P. putida* (Pp) and subsequently expressed in *P. putida* showed a LipPKS abundance of 1.1 ± 0.4% across all three codon optimization strategies. In contrast, when PKS genes optimized for *E. coli* (Ec) were expressed in *P. putida*, the LipPKS abundance dropped to just 0.2 ± 0.1%. This disparity was even more significant in *E. coli*: genes codon-optimized specifically for *E. coli* generated, on average, 22 times more LipPKS than those optimized for *P. putida* and *C. glutamicum*.

However, an exception to this trend was observed with the Cg_hrca codon variant expressed in *P. putida*. Even though it was optimized for *C. glutamicum*, it demonstrated the highest peptide abundance in the host *P. putida*. This anomaly could be attributed to the significant similarity among the hrca optimized PKS genes (Supplemental Figure 9).

Another significant distinction observed was the relative standard deviation between plasmid-based expression and genomic expression. The variation ranged dramatically, with plasmid-

based expression having a deviation as high as 121%, while genomic expression had a much smaller deviation, with deviations as low as 2%. This information implies that expressing LipPKS from a plasmid might result in more inconsistent data, potentially making it less repeatable across different experiments or replicates.

Besides assessing LipPKS abundance, we also measured IbpA levels in *E. coli* and *P. putida* samples (Figure 15a). IbpA, a small Hsp, is known to be expressed when there is insoluble protein present in the cell. Elevated IbpA levels often indicate misfolded proteins and can provide insights into the performance of heterologously produced proteins in these hosts.[321] In *E. coli*, a strong Pearson correlation of 0.94 (n = 33) was found between IbpA and LipPKS levels. This correlation was absent in *P. putida*. Notably, in *P. putida*, the expression of the WT_ATG variant led to the highest IbpA signal, even though the LipPKS production was relatively low at 0.09 ± 0.02%. This suggests that the lack of codon usage similarity might lead to not just lower protein levels but also protein insolubility.

So far, the impact of codon usage on gene expression, whether at the transcriptional or translational level, remains unclear, especially when considering multiple species.[322,323] To examine the relationship between the transcript and the target protein, we measured the LipPKS transcript level using RT-qPCR. We derived transcript levels from the difference in Ct value between the target transcript and a housekeeping gene (Supplemental Table 3). Amongst the organisms studied, *P. putida* exhibited the most pronounced correlation between transcript and protein levels with a coefficient (R = 0.67). For the other two organisms, the correlation was weaker, not surpassing an absolute value of 0.3 (Supplemental Figure 10). Typically, a low transcript presence corresponded with reduced protein levels. However, in *C. glutamicum*, codon variants Cg_ubc, Cg_mcu, and Ec_ubc were exceptions. We could not pinpoint any transcripts for them, yet significant LipPKS peptide quantities were present. This discrepancy also applied to Ec_mcu and Ec_hrca variants in *P. putida*. Excluding the Cg_mcu variant in *C. glutamicum*, all these outliers were activated at inoculation, which might have influenced mRNA concentrations.

An interesting pattern we observed was the elevated transcript levels associated with the hrca codon optimization, irrespective of the targeted host (Figure 15b). This was particularly pronounced in *E. coli*, where the hrca codon optimizations resulted in a 1266 ± 111% relative transcript amount, in contrast to the ubc optimizations which produced 378 ± 96% relative transcript. The increased transcript levels for Pp_hrca and Cg_hrca codon variants did not correlate with increased protein quantities. These observations might result from the fact that

codon frequency may influence translation speeds, as demonstrated by the heightened LipPKS counts in the host-specific variants of hrca optimizations.

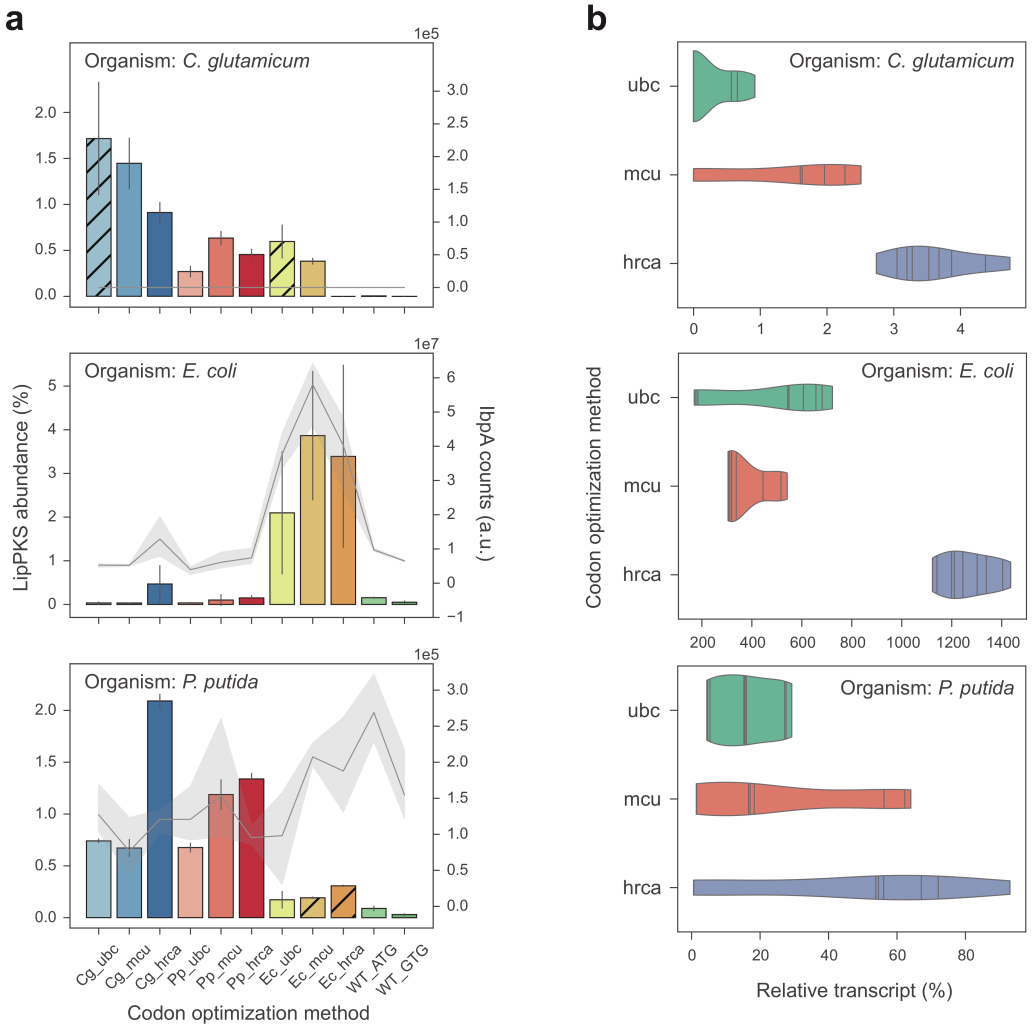


Figure 15: Relative abundance of LipPKS peptides and transcript in heterologous hosts expressing different codon variants of LipPKS. (a) Calculated protein abundance (n = 3) is the relative intensity of the top 3 peptides that correspond to the target protein divided by the intensity of all proteins detected. Hatched bars indicate strains with non-excisable backbones. Levels of the insolubility marker IbpA are represented by a gray line (n = 3). IbpA is not present in *C. glutamicum*. (b) Each violin represents the distribution of the relative transcript amount for all target host optimizations with the same optimization strategy (n = 9). Relative transcript was calculated from the Ct value difference between target transcript and housekeeping gene. The lines within the violins are individual data points. Target host optimization: Cg = *C. glutamicum*; Ec = *E. coli*; Pp = *P. putida*; Optimization strategy: ubc = use best codon; mcu = match codon usage; hrca = harmonize relative codon adaptiveness

3.2.3.5 Production of an unnatural polyketide by engineered PKS

The PKS protein production does not require any genetic alterations in the selected hosts. However, to validate the functionality of our engineered PKS and its ability to produce the intended polyketide, additional pathways had to be expressed. For effective polyketide synthesis in *E. coli*, we introduced the PKS vectors into the readily accessible *E. coli* K207-3 strain.[324] This particular strain has the essential mmCoA pathway and phosphopantetheinyl transferase (PPTase), ensuring the availability of the extender substrate and the activation of the ACP domains. The broad substrate range of the LipPKS loading didomain allows it to be non-selective, accepting propionyl-CoA as an alternate starter unit. This eliminates the need to express an ibCoA pathway.[119]

Owing to *P. putida*'s inherent broad-specificity PPTase and its capability to utilize the ibCoA precursor valine as a carbon source, the only necessary modification is the integration of an mmCoA pathway.[325,326] As previously documented, the mmCoA mutase (MCM) and epimerase (EPI) sourced from *Sorangium cellulosum* So ce56 function sufficiently in *P. putida.*[327] Thus, we incorporated the unmodified operon, under control of the tac promoter, into the MR11 *attB* site of *P. putida* AG5577. In contrast to the BEDEX vector, the chosen vector backbone does not possess the *lacI* gene, thereby not requiring backbone excision (Figure 16a). The modified strain was named *P. putida* AG5577mm.

The third host, *C. glutamicum*, is unable to use valine as a carbon source, and it is unlikely that an ibCoA pathway exists. Nevertheless, prior research has identified the presence of both mmCoA and propionyl-CoA in this organism.[139,328] However, it is uncertain if *C. glutamicum* contains a type I PKS-compatible PPTase. To address this, we used traditional cloning methods to integrate the PPTase *sfp* into *C. glutamicum* AG6212.[139] To enhance the likelihood of polyketide synthesis, we also introduced the heterologous CoA ligase *CCL4* and the ketoisovalerate decarboxylase *kivd*. Leveraging the innate aldehyde dehydrogenase activity, these two genes might collectively enable the production of the desired LipPKS starter unit, ibCoA.[329–331] We further increased the likelihood of polyketide production by adding isobutyric and propionic acid to the culture medium. Lastly, we knocked out the gene *prpDBC2* to boost flux through mmCoA, which reduces flux through the methyl citrate cycle (Figure 16b). The resulting modified strain was named *C. glutamicum* AG6212cz.

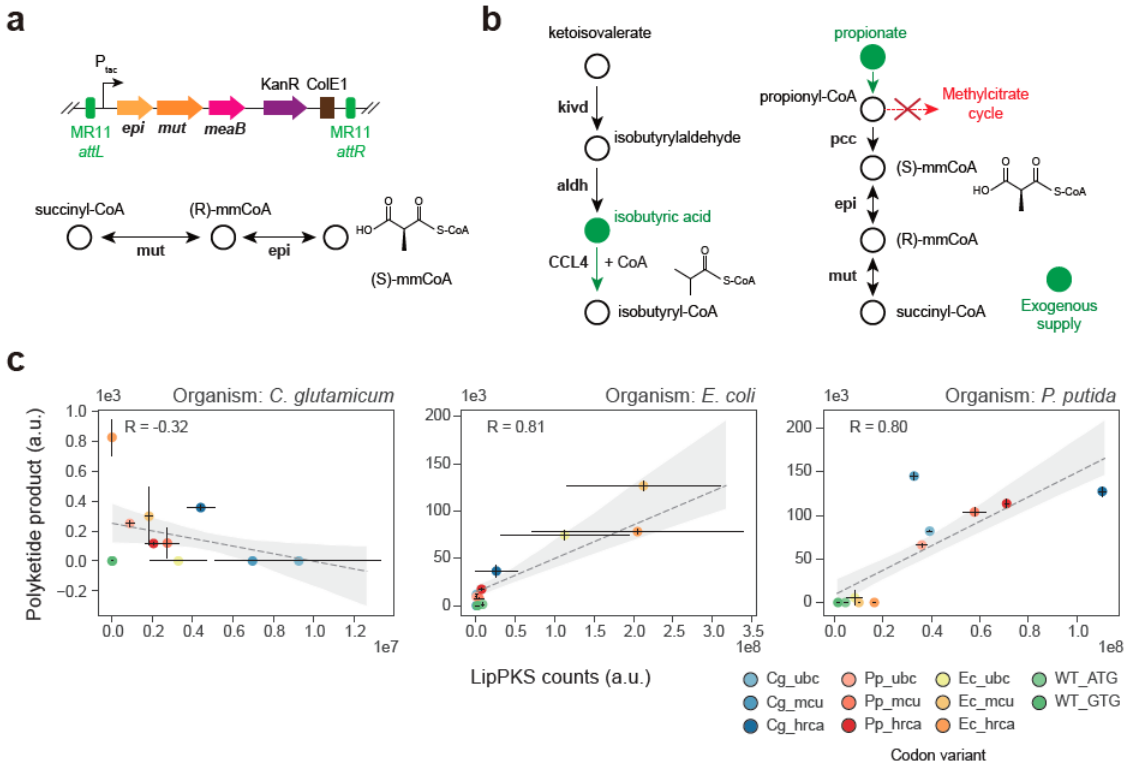


Figure 16: Production of an unnatural polyketide by engineered heterologous hosts. (a) Serine recombinase-assisted integration of the methylmalonyl-CoA mutase (mut) and epimerase (epi) pathway into *P. putida* AG5577. mmCoA: methylmalonyl-CoA (b) Engineered pathway for the production of the LipPKS loading substrate isobutyryl-CoA by the heterologous enzymes Kivd (ketoisovalerate decarboxylase) and CCL4 (2-methylpropanoate--CoA ligase) in *C. glutamicum* AG6212. Isobutyric and propionic acids were added to the *C. glutamicum* production medium (green). aldh: endogenous aldehyde dehydrogenase activity. (c) Regression plot of the LipPKS counts and polyketide production levels in the engineered hosts *C. glutamicum* AG6212cz, *E. coli* K207-3, and *P. putida* AG5577mm (n = 3). The calculation of the Pearson correlation included every replicate as a single data point. Target host optimization: Cg = *C. glutamicum*; Ec = *E. coli*; Pp = *P. putida*; Optimization strategy: ubc = use best codon; mcu = match codon usage; hrca = harmonize relative codon adaptiveness

Figure 16c shows the polyketide production in all three modified hosts. Authenticity of the produced polyketide was validated using a standard (Supplemental Figure 11). For *in vivo* synthesis of the anticipated polyketide, the host organisms must generate one of the essential starter units: propionyl-CoA, ibCoA, or 2mbCoA. It should be mentioned that the LipPKS loading domain has a strong preference for ibCoA as its substrate.[119] Hence, in *C. glutamicum* AG6212cz and *P. putida* AG5577mm, our primary emphasis was on synthesizing 3H24DMPA. Given *E. coli* K207-3's lack of a BKD complex, the only possible product it can

produce is 3H2MPA. An in-depth reaction scheme of the polyketide synthesis process is shown in Figure 2.

In both *P. putida* and *E. coli*, we observed a significant Pearson correlation coefficient for the protein and product levels, with values of 0.80 and 0.81, respectively. The Cg_mcu codon variant led to the highest product titer in *P. putida*. In *E. coli*, the highest product titer was achieved with the Ec_mcu codon variant. Surprisingly, our *in vivo* findings in *P. putida* provide evidence for the extended mCoA product (Supplemental Figure 11). This observation is in contrast with the previously reported and predicted mmCoA specificity of LipPKS AT1.[119]

In *C. glutamicum*, polyketide production was minimal, almost to the point of being undetectable (Supplemental Figure 10). This was surprising given the substantial levels of LipPKS present. Furthermore, no obvious correlation could be observed between LipPKS peptides and the polyketide product (R = -0.32). In this chapter, not only did we assess the quantities of LipPKS in this host, but we also evaluated the peptide counts for Sfp, Kivd, and CCL4. However, we could only detect peptides related to Sfp and Kivd. Peptides corresponding to CCL4 were absent (Supplemental Figure 12). As a result, it is possible that this strain lacks the CoA ligase needed for the CoA activation of the starter unit, isobutyric acid.

3.2.4 Discussion

In this chapter, we determined the optimal codon optimization strategy for an engineered T1PKS for heterologous expression in *C. glutamicum*, *E. coli*, and *P. putida*. Additionally, we shed light on the correlation between codon variation, protein concentration, transcript amount, and product yield.

We employed the Python library DNA Chisel, which enabled us to perform highly customizable and transparent codon optimizations of our target gene. The extent of customization played an important role in finding the right balance between the optimization objective and the challenges posed by synthesis. For example, by applying the UniquifyAllKmers constraint, we were able to considerably decrease repetitive sequences, a known challenge in the chemical synthesis of DNA.[332] Consequently, the efficiency of codon optimization might be limited by the inherent constraints of existing DNA synthesis methods.

In our research, the Kazusa codon usage database seemed to be sufficient. However, for future studies involving more unusual PKSs or heterologous hosts, this might become a constraint. While the ubc and mcu methods require only the codon usage of the targeted host, the hrca method additionally needs the codon usage table from the source organism. Yet, our newly

introduced online tool integrates the CoCoPUTs database, enhancing the accuracy of the codon usage table and expanding the range of accessible organisms.

Common problems of codon optimizations are reproducibility, transparency and the used codon usage database, particularly with commercial optimization tools. These choices can significantly impact the outcome of the optimized gene. For instance, the Kazusa's codon table for *S. aureofaciens* (NCBI:txid1894) lists just 80 CDSs. In contrast, the CoCoPUTs table for the same species contains 37,337 CDSs. This significant difference can likely be attributed to the advancements in genome sequencing, leading to more precise bacterial genotyping.[333]

Since we created 11 codon variants of a large gene (> 7 kb), we tried to avoid plasmid-based expression systems. Besides known complications such as inconsistencies in plasmid stability, varying copy numbers, and growth anomalies, there was also a potential for intercompatibility issues across our different organisms.[92,334–339] The SAGE system has emerged as an indispensable tool in our genetic engineering efforts, especially when reproducibility and throughput capabilities were bottlenecks. This system facilitates stable, multiplexed integrations of large genes into the desired host genomes. It is highly efficient and demands minimal expertise of the targeted microbe. Used in combination with BEDEX vectors, this system allows for precise control of gene expression and can enhance the efficiency of plasmid assembly by reducing the cellular burden from strong constitutive promoters or possibly harmful gene products.[112] Our studies further established that BEDEX vectors function universally across our selected hosts. However, the RFP levels detected in *C. glutamicum* were substantially lower when compared to *E. coli* or *P. putida*. Several factors could account for this, such as an incompatible RBS or promoter.[340,341] Still, there was a significant difference between the excised and unexcised vector backbone. As a result, the promoter *LlacO1* must have become constitutively active in this host.

To extend our evaluation of the BEDEX system, we utilized it in the construction of our LipPKS expression strains. The results during the construction process mirrored what we had observed with RFP. However, we had challenges in removing the backbone from specific codon variants. Highly incompatible codon variants seemed to have caused growth deficits, favoring the selection of *lacI* positive strains. To investigate this phenomenon, we expressed the codon variants Pp_mcu and Ec_mcu using an inducible vector system in *P. putida*. As previously demonstrated, the codon variant optimized for Pp_mcu resulted in a substantial amount of LipPKS peptides and their associated product. In contrast, the Ec_mcu variant only yielded the corresponding peptides. When expressing these LipPKS variants, protein levels for each were alike. However, the concentration of the insolubility marker IbpA was

significantly elevated in *P. putida* expressing the Ec_mcu LipPKS (Supplemental Figure 13). Additionally, no product formation was observable with the Ec_mcu codon variant. This might suggest that an incompatible codon variant leads to the production of insoluble or improperly folded proteins.

The observed variations in LipPKS protein quantities emphasize that codon usage bias plays an important role in the heterologous expression of PKSs. When compared to the WT codon version, our top-performing codon variants showed significant increase, at least 50-fold, in PKS protein concentrations, which was necessary for the production of the associated polyketide. Yet, the reasons for these differences remain somewhat unclear.

In *P. putida*, our RT-qPCR analyses might show a potential relationship between codon preference and transcript abundance. In contrast, the data from *E. coli* and *C. glutamicum* fail to demonstrate such a correlation. For these two latter organisms, it is possible that codon frequency exerts a more pronounced effect on protein levels. Existing research provides supporting arguments for both scenarios.[322,323] A widely accepted dogma is that infrequent codons impede the translation machinery, while commonly occurring codons facilitate translation.[322] Conversely, some studies showed that the effects observed are predominantly a result of differences in transcriptional efficacy.[323,342] Besides a role in modulating eukaryotic chromatin, there is a hypothesis suggesting that non-optimal gene sequences might facilitate the attachment of unidentified intragenic transcription factors.[342] Additionally, it is important to note that, in prokaryotes, transcription and translation processes occur concurrently, without spatial separation.[343] This simultaneous occurrence makes it particularly difficult to pinpoint whether the observed variations are based on transcriptional inefficiencies or translational problems in these organisms.

While our research could not conclusively address this ongoing debate, it did provide a solution for codon optimization of PKSs. Based on our findings, the mcu method seems to yield the highest protein and product concentrations. In contrast, the hrca method led to the most elevated transcript levels. Hence, the selection of the best-suited strategy largely depends on the primary objective or constraints of the targeted host.

Moreover, our study provides additional evidence for the differences in protein production when utilizing ATG versus GTG start codons, aligning with previous research results.[344] The use of GTG start codons appears to be more dominant in organisms with high GC content. Meanwhile, the percentage of GTG initiation codons decreases to below 25% in organisms with a GC content under 65%.[345] For the heterologous expression of PKSs, the G to A

mutation can be easily introduced via site-directed mutagenesis, offering an economically feasible strategy to improve expression.

Besides *E. coli*, our study found the other two tested hosts, *C. glutamicum* and *P. putida*, to be suitable for expressing engineered T1PKSs. The functionality of the PKS and the synthesis of the anticipated 3-hydroxy acid product were confirmed across all three host organisms. Despite the significant protein levels in *C. glutamicum*, the low product concentrations are potentially a result of the poor supply of ibCoA. The loading didomain of LipPKS shows a strong affinity for ibCoA. Still, it does exhibit a broader specificity by also accepting other CoA derivatives like 2mbCoA, isovaleryl-CoA, and propionyl-CoA.[119] However, when LipPKS uses propionyl-CoA, its catalytic efficiency decreases to roughly one-fourth compared to its activity with ibCoA. To further optimize the output, there is a clear need to assess intracellular acyl-CoA concentrations and to refine the precursor pathways in *C. glutamicum*. This more detailed investigation of the cell metabolism will ensure a better understanding of the underlying reasons for the low polyketide titers.

Even though we identified an optimal strategy for codon optimization of a T1PKS, the reason for its success remains unclear. Gaining a comprehensive understanding might require the codon optimization of a diverse set of genes across a wide range of hosts, a study that would be very costly given the current price point of DNA synthesis. Moreover, we should consider that the methods we identified might not be the pinnacle of codon optimization techniques. The rapid development in artificial intelligence (AI) and ML shows promising results.[346] These emerging technologies could be instrumental in redefining the process of codon optimization for heterologous gene sequences in the future.

Chapter 3.3

From Type I Polyketide Synthases to Functional Genomics: Optimizing δ-lactone Production in *Pseudomonas putida*

Contributions

M.S. generated all strains and performed all experimental work with the assistance of A. A.V. A.N.P. assisted in earlier stages during the cloning process. M.G.T. assisted with the project outline and BarSeq experiments. A.M.D. provided quantitative sequencing and statistical evaluation of BarSeq data. The Tn7 integration system was provided by Sebastian Köbbing and L.M.B. The pBH vectors were provided by R.W.H. Results will be published in a scientific journal and the manuscript is in preparation. M.S. analyzed all results. The manuscript is being written by M.S. with the help of L.M.B. and J.D.K. J.D.K. conceived the project.

3.3 From Type I Polyketide Synthases to Functional Genomics: Optimizing δ-lactone Production in *Pseudomonas putida*

3.3.1 Abstract

Lactones are volatile compounds used to enhance flavors in food or add fragrances to cosmetics. Currently, production largely depends on petrochemical processes, which are associated with the generation of toxic waste and other environmental concerns.

The heterologous *de novo* biosynthesis of δ-lactones in microbes remains highly challenging and most existing synthetic pathways utilize exogenous fatty acids as a starting material. Despite our limited understanding of the native δ-lactone pathways in plants, PKSs have been proposed as a potential route for producing these valuable compounds.

In this chapter, we detail the optimization of *P. putida* as a non-native host for PKSs and the subsequent production of δ-lactones using an engineered type I PKS. By enhancing the expression of a standalone version of this PKS and ensuring the production of the C5 and C6 precursors, we establish a sufficient supply of the first extension intermediate. Finally, by fusing a linker domain to the first extension module and introducing a second extension module to the optimized strain, we achieved the production of the desired C7 and C8 δ-lactones.

3.3.2 Introduction

The efficient transformation of plant biomass into valuable products stands as one of the foremost challenges in modern biotechnology. Plant-derived biomass primarily comprises three elements: lignin, hemicellulose, and cellulose. While hemicellulose and cellulose possess relatively simple structures, lignin is inherently more complex, being a diverse blend of predominantly aromatic polymers.[347] Consequently, when lignin undergoes hydrolysis, it results in a mixed aqueous solution with saccharides, organic acids, ketones, furfurals, and a variety of other chemicals.[348,349]

A simple and convenient way to valorize the hydrolysate is the use of microbes, however, some of the lignin-derived compounds are toxic and only a few tolerant microorganisms are known or well enough studied. A prominent representative in this context is the genus of Gram-negative bacteria known as *Pseudomonas*. Specifically, for industrial applications, the omnivorous soil bacterium *P. putida* has gained significant interest. Due to its unique metabolism, *P. putida* can utilize an extensive array of carbon sources, including lignin-derived monomers.[350] Additionally, with its relatively high GC content (61.5%), *P. putida* emerges as an optimal host for heterologous pathway expression sourced from GC-rich bacteria.[219] A

highly versatile approach is provided by the expression of PKSs, especially by the non-iterative T1PKS.

Over the years, the modularity of T1PKS has been a compelling target for engineering, allowing for the synthesis of a diverse range of natural products.[123,124,132,303] However, the expression of PKSs in non-model hosts, such as *P. putida*, remains in its infancy. The yields are frequently suboptimal, and given the vast diversity of polyketides, formulating a one-size-fits-all strategy for enhancement proves challenging.[219,327,351–354] Given the substantial size and intricacy of PKSs, their expression is likely to be a significant bottleneck.[121] While this constraint was partially lifted by codon optimization, another pivotal aspect to consider is the availability of precursors.[121] Consequently, refining both the host's metabolic pathways and the composition of the culture medium emerges as key optimization strategies.

The characteristic of *P. putida* to use a broad range of molecules as a carbon source is also one of its greatest disadvantages. Most of the time, the bacterium possesses pathways to convert target compounds into intermediates or fully degrade the desired product.[56] As a result, the use of *P. putida* as a production platform often requires an intense study of the metabolic network and the application of genetic engineering. Recent developments in TnSeq are providing a promising approach to overcome this problem. By combining TnSeq with RB, it is possible to discover whole biodegradation pathways in a single experiment.[11,12,48,172]

In this section, our objective was to harness the capabilities of *P. putida* KT2440, using it as a foundation for engineered T1PKSs. The initial phase of introducing a non-traditional PKS host comprised of verifying the functionality of the heterologous PKS. The selected T1PKS, derived from the lipomycin synthase of *S. aureofaciens* Tü117, had been previously used to produce an industrially relevant 3-hydroxy acid.[123] The identification of the expected product in *P. putida* suggested that this host meets the fundamental requirements for PKS expression.[121] However, to acquire a deeper understanding of the host's phosphopantetheinylation capacities, we employed a colorimetric assay to measure endogenous PPTase activity. This facilitated a comparison between the native PPTase and strains expressing an additional heterologous PPTase.[325,355]

After that our focus shifted to precursor availability. The starter units for the evaluated PKS are typical intermediates found in the BCAA catabolism.[119] To rationally engineer the host, relevant genes were identified by feeding each BCAA as the sole carbon source to an RB-TnSeq mutant library. Once identified, downstream genes for precursor accumulation were eliminated, and growth deficiency was verified through plate-based growth assays.

Finally, we assessed the impact of these metabolic engineering efforts by measuring the production of the expected polyketide under varying conditions. This optimized strain was then employed to express the complete PKS pathway, resulting in the production of the final δ-lactone.

3.3.3 Results and Discussion

3.3.3.1 Phosphopantetheinyl transferase activity and methylmalonyl-CoA metabolism

The basic requirement for the functional expression of a PKS is the activation of the cis-ACP domain via phosphopantetheinylation. While it is known that *P. putida* possesses a broad-specificity PPTase, it remains unclear whether the addition of a heterologous PPTase improves T1PKS activity.[325,355] To address this, we first integrated the PPTase gene *sfp* from *Bacillus subtilis* into the *P. putida* genome using a mini-Tn7 delivery transposon system.[137] However, instead of measuring the polyketide production directly, we opted to use the colorimetric blue-pigment synthase A (BpsA) assay to determine *in vivo* PPTase activity.[355] The non-ribosomal peptide synthase (NRPS) BpsA converts two molecules of L-glutamine into the blue pigment indigoidine.[356] The activation of its peptide carrier protein (PCP) domain by a PPTase is the rate-limiting step of indigoidine synthesis and can serve as an indirect measure of PPTase activity.[355]

Indigoidine production in WT *P. putida* and *P. putida* Tn7::pBG14e-Sfp is shown in Figure 17b. Both strains that were expressing the *bpsA* gene produced a significant amount of the blue pigment, indigoidine. The WT PPTase activity resulted in a normalized indigoidine signal of 0.85 □ 0.15. In contrast, the strain with the additional PPTase Sfp showed an approximately 2-fold higher indigoidine level. Thus, by integrating a copy of *sfp* into the host genome, we were able to improve indigoidine production. In order to determine the effects of this modification on an actual T1PKS, we used an AT-swapped version of the LipPKS that accepts mCoA rather than mmCoA.[123]

Production of the expected polyketide 3H4MPA could be observed in both strains. However, there was no significant difference in their production levels. The endogenous PPTase activity in *P. putida* seems sufficient to activate the AT domain of the expressed T1PKS protein. As we prefer a minimalistic approach and aim to avoid unnecessary modifications of the metabolism, we chose not to include an additional PPTase in our strains.

The subsequent step focused on the availability of the common PKS extender unit, mmCoA. The mmCoA mutase (MCM) and epimerase (EPI) pathway from *S. cellulosum* has been

successfully expressed in *P. putida.*[121,327] However, the functional expression of the alternative PCCase pathway, which comprises of the *accA2* and *pccB* gene, has not been reported in this organism.

In this study, we were able to clone and integrate this highly toxic gene using BEDEX vectors.[121] By including the repressor *lacI* into the backbone of the integrative vector and placing the target gene under the control of a *lac* promoter, toxicity or other burdensome effects can be mitigated. Once the backbone is excised, the *lac* promoter becomes constitutively active.[121] In addition to utilizing BEDEX vectors, we also codon-optimized the *accA2* and *pccB* genes from *S. coelicolor* using the hrca method.[121]

As a result, we confirmed the functional expression of this PCCase in *P. putida* by detecting the mmCoA-extended polyketide (Figure 17c). However, while the MCM/EPI pathway led to an approximately 30% higher polyketide titer, the culture conditions we used were not optimized for the PCCase pathway. A standard approach to maximize mmCoA production using this pathway involves supplementing the medium with propionate.[307]

Overall, *P. putida* has demonstrated its suitability as a host for the general assessment of engineered T1PKSs. The presence of a highly promiscuous PPTase, combined with the functional expression of both mmCoA pathways, enables the production of a wide range of polyketides. Nonetheless, polyketide titers remain low and one of the reasons is most likely the availability of the precursor acyl-CoAs. The LipPKS loading pathway accepts branched acyl-CoAs that are intermediates in *P. putida*'s native BCAA metabolism.[119] Therefore, we decided to utilize BarSeq as a tool to guide our engineering efforts for the accumulation of the preferred LipPKS starter units ibCoA and 2mbCoA.[48]

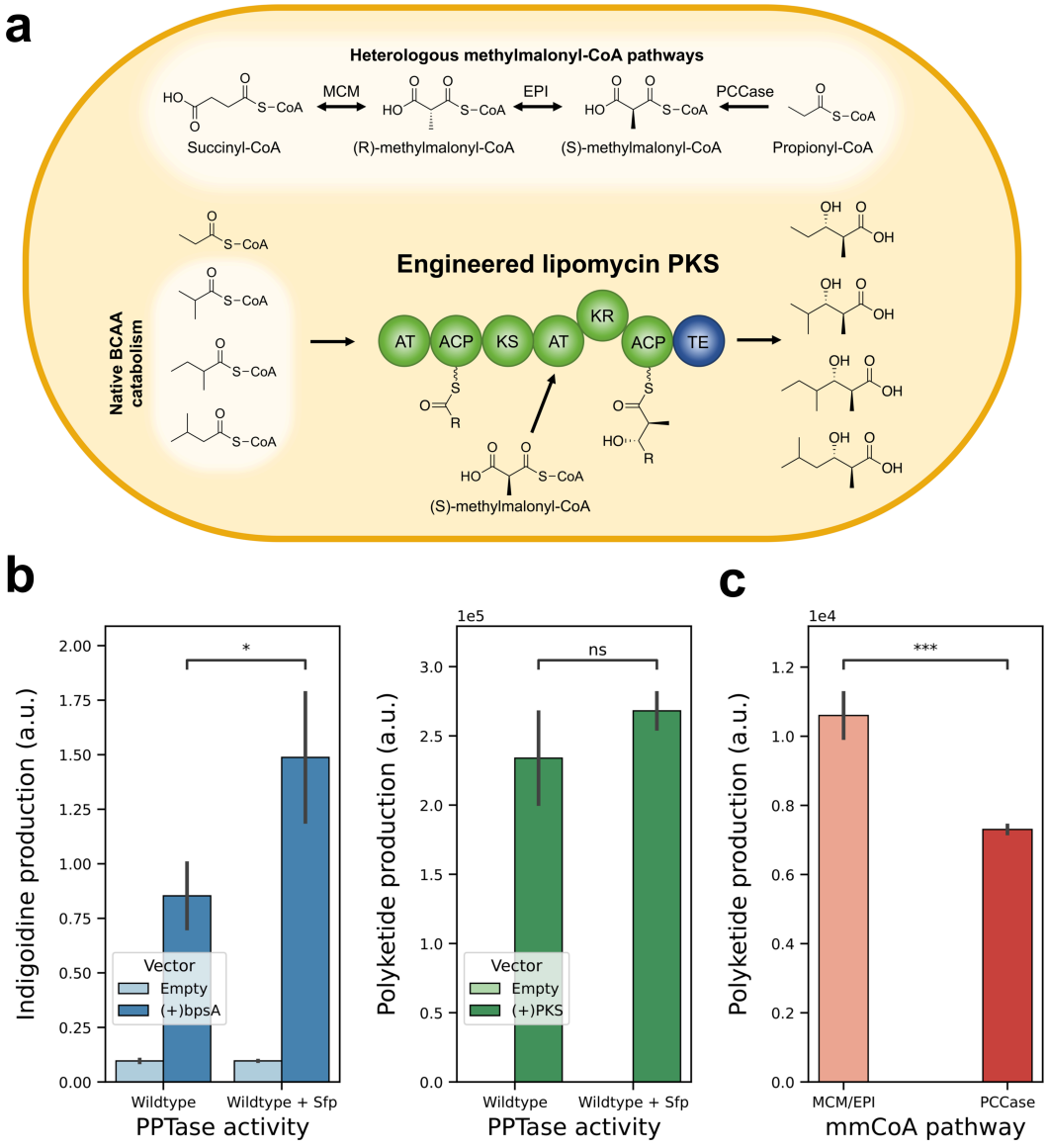


Figure 17: Engineered lipomycin polyketide synthase (LipPKS) expression in *Pseudomonas putida*. (a) General design of the engineered LipPKS pathway and its required precursors. The tested methylmalonyl-CoA (mmCoA) pathways comprised of either the mmCoA mutase and epimerase (MCM/EPI) from *Sorangium cellulosum* So ce56 or the propionyl-CoA carboxylase (PCCase) complex from *Streptomyces coelicolor*. The preferred loading substrates are located in *P. putida*'s native BCAA catabolism. The polyketide product is an enantiopure 3-hydroxy acid, which varies based on the substrate loaded. (b) The lue-pigment synthase A (BpsA) assay is used for determining *in vivo* phosphopantetheinyl transferase (PPTase) activity. Adding the heterologous PPTase Sfp did not result in increased polyketide titers. (c) Comparison between the MCM/EPI and PCCase pathway in *P. putida*. Under the tested conditions, the MCM/EPI resulted in a slight increase of the polyketide product. AT: acyl transferase; ACP: acyl carrier protein; KS: keto synthase; KR: keto reductase; TE: thioesterase

3.3.3.2 BarSeq analysis of the branched-chain amino acid metabolism in *P. putida*

To identify the genes responsible for the degradation of the precursor acyl-CoAs, we fed L-leucine, L-valine and L-isoleucine as the sole carbon source to a *P. putida* RB-TnSeq mutant library (Figure 18). To ensure an even better coverage we included the corresponding 2-oxo-acids for these BCAAs as well. The resulting differences in mutant abundancies can be measured and used to calculate the fitness values for conditionally essential genes.[48] While the BCAA metabolism has been characterized previously by using them as a nitrogen source, the use as a carbon source has not been reported yet.[11] The addition of carbon source experiments can increase the significance for fitness data downstream from the deamination reaction and facilitates the identification phenotypes for specific transporters.[11]

The strongest transporter phenotypes we identified were for the genes PP_0878-PP_0881 and PP_1137-PP_1141, especially in the isoleucine and leucine condition. While we could not identify a specific valine transporter, the valine carbon source, valine nitrogen source and 3-oxobutanoic acid condition caused interesting growth phenotypes for the conserved proteins of unknown function PP_0642 and PP_5452. In the valine carbon source condition, PP_0642 and PP_5452 have positive fitness values of 1.9 and 1.7, respectively. In contrast, the valine nitrogen source and 3-methyl-2-oxobutanoic acid condition resulted in similarly strong but negative fitness values. It is likely that PP_0642 and PP_5452 are either involved in sensing or transporting valine related metabolites.

The degradation of BCAAs seems to be initiated by the aminotransferase IlvE, which shows a significant fitness defect of -4.5 across all three tested substrates. The subsequent step, which is catalyzed by the BKD complex (PP_4401-PP_4404), is shared across all the BCAA and 2-oxoacid conditions. This reaction also results in the targeted acyl-CoAs, ibCoA and 2mbCoA. Therefore, preventing the reaction downstream of the BKD complex should accumulate the desired LipPKS starter units.

Using BarSeq, we identified three acyl-CoA dehydrogenases (ACDHs), namely PP_4064, PP_3492 and PP_2216, potentially responsible for the degradation of ibCoA and 2mbCoA. The most essential ACDHs for ibCoA degradation appear to be PP_4064 (-2.0) and PP_3492 (-2.1), while 2mbCoA only requires PP_2216 (-4.4). To verify these results, we created a combinatorial library of deletion mutants for PP_4064, PP_3492 and PP_2216 and conducted growth experiments with BCAAs as the sole source of carbon (Supplemental Figure 15).

The identified ACDHs for isovaleryl-CoA and 2mbCoA seem to be specifically essential for growth on L-leucine and L-isoleucine, respectively. Although indicated by BarSeq, the ACDH PP_3492 did not cause a growth defect when L-valine is used as the sole carbon

source. However, the second identified ACDH, PP_4064, results in a strong growth defect. A combination of both deletions is still not sufficient to fully abolish growth on L-valine, and even the triple deletion strain ΔPP_2216ΔPP_3492ΔPP_4064 could not prevent growth with L-valine as the sole carbon source. Nonetheless, ΔPP_2216ΔPP_3492ΔPP_4064 might be our best-performing strain, as it cannot efficiently utilize any of the BCAAs as growth substrates, leading to a potential accumulation of the targeted acyl-CoAs. We have also opted to proceed with the strain ΔPP_3492ΔPP_4064 due to its already minimal growth on L-valine. While we successfully identified all the relevant ACDHs involved in the metabolism of BCAAs, one known limitation of BarSeq experiments is their ability to identify phenotypes for redundant genes.[48] It is, therefore, important to verify BarSeq results by classical mutant growth experiments, especially if redundancy of gene functions is expected.

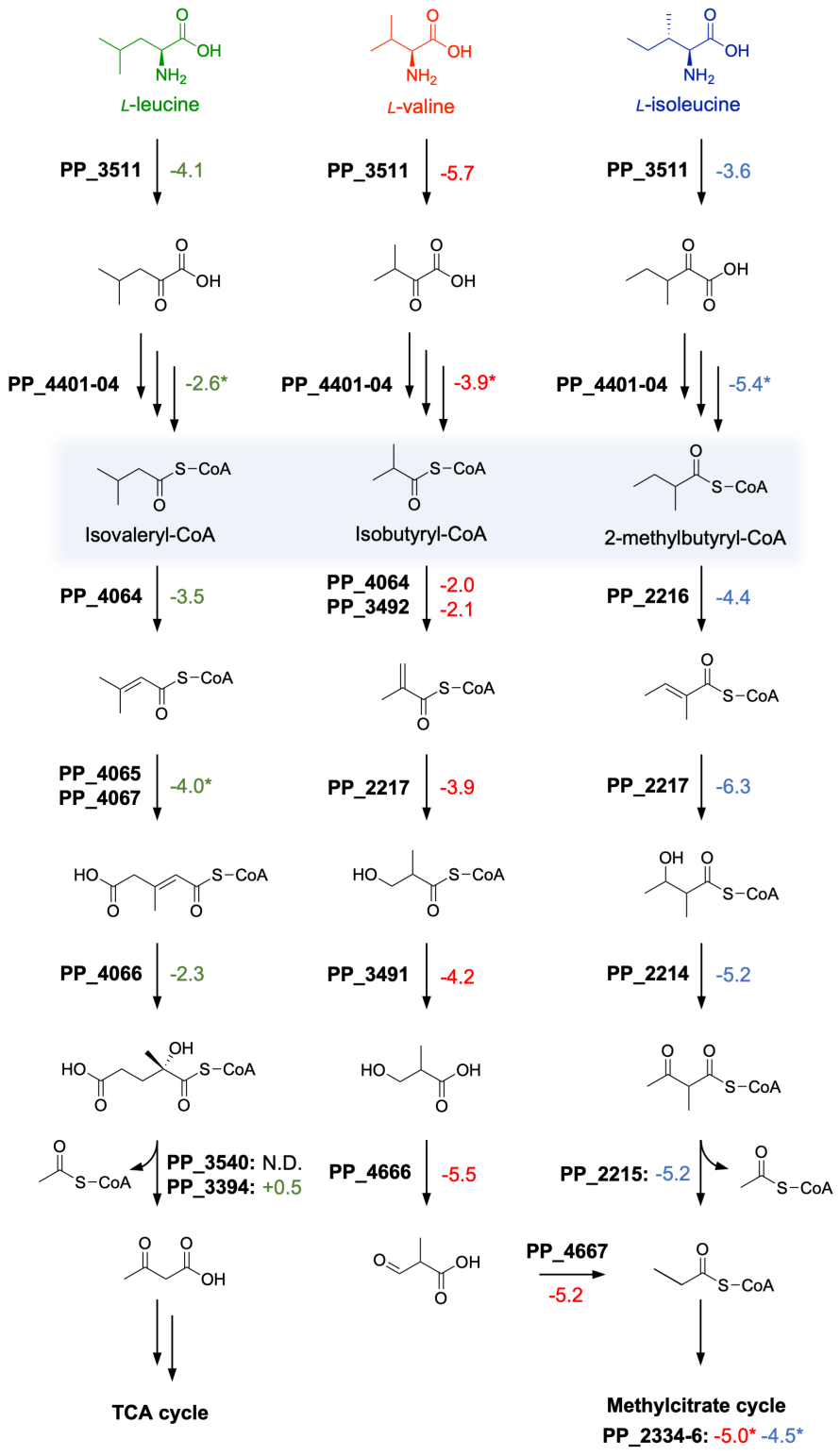


Figure 18: BarSeq analysis of the branched-chain amino acid (BCAA) degradation in *Pseudomonas putida*. Fitness defects were considered significant when the fitness value was > |1| and the t-value was > 5. Shown are the fitness values (n=2) associated with growth on L-leucine (green), L-valine (red) and L-isoleucine (blue) as the sole source of carbon. Fitness values marked with an asterisk (*) represent averages across multiple genes.

3.3.3.3 Improved polyketide production by engineered host

The initially tested version of the LipPKS contains an mmCoA-specific AT which will add a methyl group to the α-carbon position of the polyketide. After a second extension, this methyl group will shift to the γ-carbon. Since the final target molecule is a δ-lactone with a carbon chain at the δ-position, an additional sidechain at the γ-position is not desired (Figure 19). Consequently, the WT-AT of the first extension module of LipPKS needs to be exchanged for an mCoA-accepting AT. Based on previous results, we chose the highly promiscuous ansPKS M8 (AnsM8) AT and the mCoA-specific AT from the BorPKS M1 (BorM1).[124] Moreover, the AT-exchange used the updated domain boundaries (us3/ds44) as described by Englund *et al.*[128] Additionally, we included the WT-AT due to its recently reported *in vivo* activity with mCoA.[121]

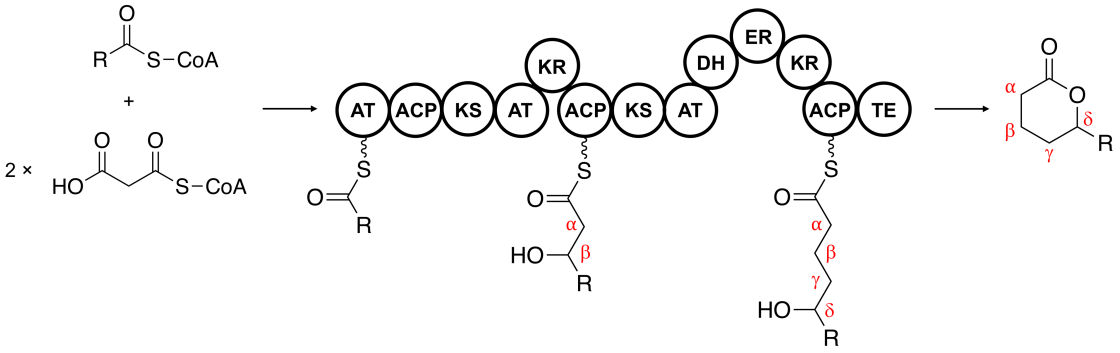


Figure 19: The general design of a δ-lactone producing polyketide synthase. Carbon positions of the growing polyketide chain are highlighted by red letters. AT: acyl transferase; ACP: acyl carrier protein; KS: keto synthase; KR: keto reductase; DH: dehydratase; ER: enoyl reductase; TE: thioesterase

Our best-performing AT-swap mutant contains the AnsM8-AT, resulting in an over 3-fold increase in 3H4MPA (ibCoA + mCoA) compared to the WT-AT (Figure 20a). Over a 96-hour period, the titers of 3H4MPA also seem to consistently increase. Furthermore, the ratio of 3H4MPA to 3H4MHA (2mbCoA + mCoA) appears to align with the *in vitro* kinetics for the LipPKS loading didomain, which shows a higher preference for ibCoA [119]. Interestingly, this product ratio does not seem to apply to the WT-AT. The equal product titers for 3H4MPA and 3H4MHA could result from the very low affinity for mCoA, leading to a stalling extension.

To our surprise, the AT from BorM1 did not result in any detectable product. This might be due to a non-compatible exchange junction. The BorM1 AT-exchange in LipPKS has been successfully performed previously.[123] However, rather than using the us3/ds44 junction, the

originally determined optimal junction us1/ds44 was used. Junction boundaries are most likely not universally applicable to every acceptor and donor pair and might require individual assessment.

Polyketide production in our combinatorial ACDH deficient library is shown in Figure 20b. The polyketide titer for the ibCoA-derived 3-hydroxy acid is below the WT background in every engineered strain. The deletion of PP_4064 resulted in an approximately 5-fold reduction of 3H4MPA. However, the results for the 2mbCoA-derived polyketide improved significantly. 3H4MHA production in the double mutant of PP_2216 and PP_4064 increased by 81 ± 2 %. The additional knockout of PP_3492 did not lead to any further improvements.

The increased titers for the 2mbCoA-derived polyketide indicate that the pathway upstream of the ACDHs must be still active. The reactions catalyzed by IlvE and the BKD complex are shared among the three BCAAs (Figure 18). Consequently, the accumulating ibCoA could be hydrolyzed by endogenous type-II TEs. In a recent *P. putida* study, the production of propionic acid via propionyl-CoA was almost completely abolished by deleting the genes for the TEs TesB and PP_4975.[357] By screening previously published proteomics data for the presence of any of the identified TEs, we were able to confirm that PP_4975, PP_2308 and TesB are highly expressed in *P. putida* (Supplemental Figure 14).[121] To test our hypothesis, we deleted the TE genes PP_4975 and *tesB* and measured the production of the corresponding acids in ΔPP_2216ΔPP_3492ΔPP_4064. Surprisingly, we were unable to detect any formation of isobutyric acid, whereas 2-methylbutyric acid production increased from non-detectable in WT to 328 ± 1 mg/L in ΔPP_2216ΔPP_3492ΔPP_4064 (data not shown). The deletion of *tesB* resulted in a decrease of about 60% in 2-methylbutyric acid levels. The additional deletion of PP_4975 had no significant effect on 2-methylbutyric acid production. If our hypothesis is correct, and the corresponding acid is produced by endogenous TE activity, we should observe an increase in the 2mbCoA-derived polyketide 3H4MHA.

In a final attempt, we used our potentially 2mbCoA-accumulating strain and tested the effect of the deletion of *tesB* on polyketide production. Unexpectedly, 3H4MHA titers were reduced to almost non-detectable levels when we included the deletion of *tesB* (data not shown).

While we succeeded in reducing 2-methylbutyric acid accumulation, it did not lead to the desired increase of the target product. It is likely that *P. putida* maintains tight control over its CoA pool and employs TEs to uphold its intricate CoA homeostasis.[357] By eliminating a major TE like TesB, the use of free CoA might be restricted to more essential metabolic functions. Furthermore, BarSeq data indicates that disrupting either *tesB* or PP_2213 (acyl-CoA synthase) has a significant fitness defect when 2-methylbutyric acid serves as the sole

carbon source. The accumulation of an acyl-CoA without an active TE might be toxic to the cell, and can lead to severe growth defects.[139]

The fact that we detect increasing amounts of 2-methylbutyric acid in our engineered strains correlates well with the improved 3H4MHA titers. However, it remains unclear why the deletion of the identified ACDHs does not increase 3H4MPA levels. The growth defect observed when using L-valine as the carbon source is similar to that seen with L-isoleucine. Therefore, *P. putida* might utilize the previously identified protein of unknown function, PP_0642, to transport the ibCoA precursor, 3-methyl-2-oxobutanoic acid, outside the cell. This hypothesis is further supported by the described growth phenotypes we observed in the valine carbon source, valine nitrogen source and 3-methyl-2-oxobutanoic acid condition. While PP_0642 has a positive fitness value for the valine carbon source condition, it becomes negative for the valine nitrogen source condition. High levels of intracellular 3-methyl-2-oxobutanoic acid might activate the expression of PP_0642, which is detrimental when *P. putida* relies on valine as a carbon source. Deletion of this membrane protein could improve flux towards ibCoA.

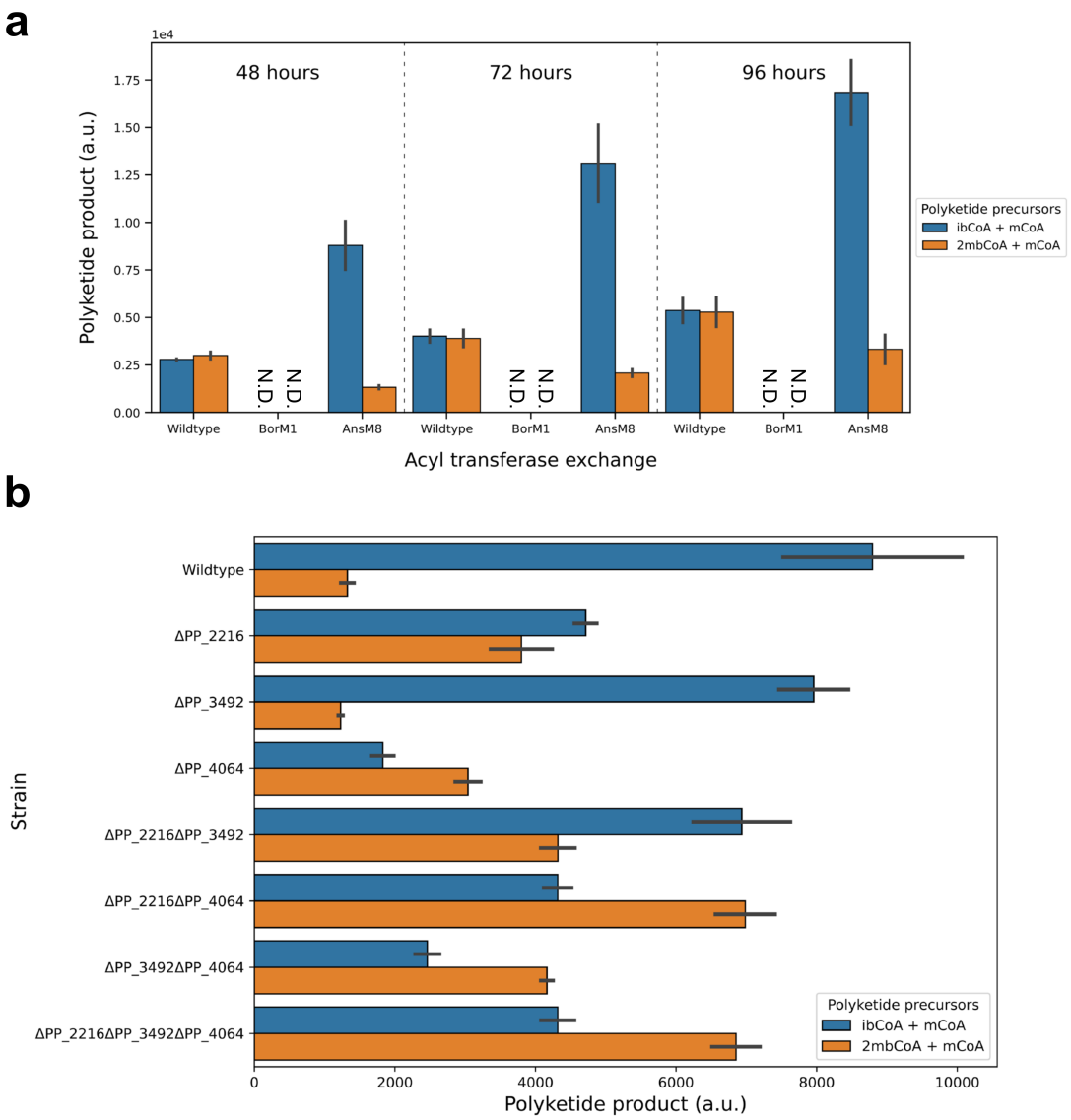


Figure 20: Polyketide production in the rationally engineered host *Pseudomonas putida*. (a) Time course of polyketide production with acyl transferase-exchanged lipomycin polyketide synthase (LipPKS) mutants. The production of the polyketide 3-hydroxy-4-methylpentanoic acid (blue) is achieved by condensing isobutyryl-CoA (ibCoA) and malonyl-CoA (mCoA). The precursors for 3-hydroxy-4-methylhexanoic acid (orange) are 2-methylbutyryl-CoA (2mbCoA) and mCoA. (b) Polyketide production in strains deficient for BarSeq-identified acyl-CoA dehydrogenases PP_2216, PP_3492, or PP_4064. A hatched bar indicates the presence of isobutyric acid or 2-methylbutyric acid, respectively.

3.3.3.4 δ-lactone producing polyketide synthase in optimized host

After attempting to optimize the production of the first extension intermediate, we aimed to introduce a second extension module with a fully reducing loop. This design would lead to the desired δ-lactone product (Figure 19). The expression of PKSs in non-native hosts seems

to benefit from dividing modules into multiple peptides rather than expressing a large, single peptide.[358] Therefore, we chose to extend our synthetic PKS with the standalone module SpnB from the spinosyn PKS (SpnPKS) pathway.[359] The only modification needed was to replace the C-terminal COMM domain of the SpnB protein with a TE domain. The TE from EryPKS should be a suitable candidate, as it is known not only to release the ACP-bound polyketide but also lactonizing it.[124]

The next step is to ensure that the first and second module interact with each other. This is achieved by linker regions or COMM domains in between the N-terminal and C-terminal parts of the proteins. The N-terminus of SpnB already contains a COMM domain to interact with its native partner, SpnA. Consequently, we fused the SpnA COMM domain to the C-terminal ACP of our engineered LipPKS (Figure 21a). Besides this design (D1), we also included two other designs, D2 and D3, that could potentially result in the production of a δ-lactone.

Prior to attempting the production of our target molecules, we conducted a degradation assay with the chemically synthesized standards for ethylvalerolactone and isopropylvalerolactone (Figure 21b). As to be expected with a fragrant molecule, most of the supplemented δ-lactones (> 55 %) evaporated. The addition of *P. putida* cells did not result in a significantly higher reduction. The only noticeable difference was the disappearance of the 5-hydroxy acid degradation product of ethylvalerolactone. While we were able to detect the expected mass for 5-hydroxyheptanoic acid in the cell-free sample, none was detected in the inoculated media. This can be attributed to *P. putida* further degrading the spontaneously delactonized product. Furthermore, we also tested various media and cultivation formats (data not shown). Besides our standard culture condition, which uses LB medium and 24-well plates, we decided to include TB medium and cultivation in glass tubes. Surprisingly, the only condition that led to the detection of our target compounds was the combination of TB medium and glass tubes. Among the three PKS designs, only D1 led to detectable amounts of ethylvalerolactone and isopropylvalerolactone (Figure 21c). It is unclear why the other two designs did not lead to the expected δ-lactones. One possibility is an incompatible codon usage of the WT spnA parts.[121] While the fusion of the small linker sequence might have little to no effect on the protein structure, the hybrid design D2 could potentially have resulted in misfolded protein. Consequently, the native WT *spnA* sequence (D3) might not have been expressed at all. In general, it seems that minimal modifications to an existing PKS system have higher success rates than more aggressive designs. Even after two decades of PKS research, we still do not fully understand the underlying reasons for successful or unsuccessful PKS designs.

Differences as small as a single amino acid during domain swaps have been shown to result in inactive proteins.[128] Due to this intricacy of engineering PKSs, more intense research of the fundamental science of protein folding and expression is required.

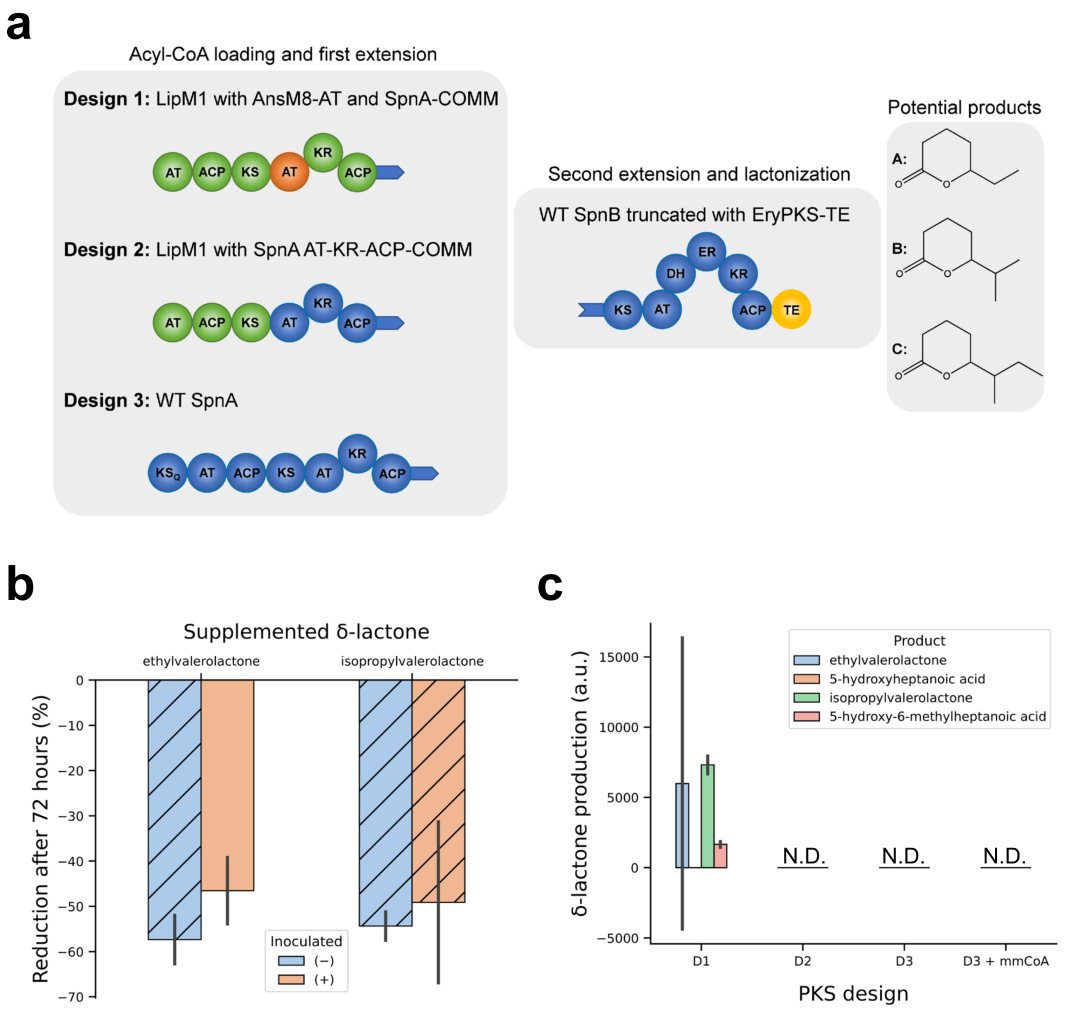


Figure 21: δ-lactone production and polyketide synthase (PKS) design in *Pseudomonas putida*. (a) The three tested designs for the first extension resulting in the acyl carrier protein-bound substrate for the second extension module (SpnB). The module SpnB from *Sacchapolyspora spinosa* was truncated using the erythromycin PKS (EryPKS) thioesterase (TE) from *S. erythraea*. The most likely products are ethylvalerolactone (A), isopropylvalerolactone (B) and 1-methylpropylvalerolactone (C). (b) 72-hour degradation assay for ethylvalerolactone and isopropylvalerolactone in minimal media with glucose. Media was either kept sterile (light blue) or inoculated with *P. putida* cells (light orange). A hatched bar indicates the presence of the expected mass for the corresponding 5-hydroxy acid (data not shown). (c) δ-lactone production by the tested PKS designs (D1, D2, and D3) in *P. putida*. The PKS design 3 (D3) utilizes methylmalonyl-CoA (mmCoA) as a loading substrate and requires the integration of a mmCoA pathway. AT: acyl transferase; ACP: acyl carrier protein; KS: keto synthase; KR: keto reductase; DH: dehydratase; ER: enoyl reductase; TE: thioesterase; KS_Q: decarboxylating keto synthase; COMM: linker sequence

Our future steps for this project will involve testing the δ-lactone-producing PKS in our engineered hosts. This might uncover bottlenecks beyond the production of the first polyketide intermediate. A major concern is the codon variant of the SpnB module. Although the codon usage and GC-content of the native host *S. spinosa* and *P. putida* are similar, we have previously demonstrated that even slight differences in codon usage can negatively impact protein expression.[121] If we do not observe an increase in the 2mbCoA-derived δ-lactone, it may be due to limited expression of the second extension module. Analyzing protein levels through proteomics can provide valuable insights into the underlying cause of low product titers.
As a final step, we will produce the target δ-lactones from a biomass hydrolysate. This will demonstrate *P. putida*'s potential role in a circular bioeconomy.

Chapter 4

General discussion

Contributions

This chapter was written by Matthias Schmidt and reviewed by Lars M. Blank and Jay D. Keasling.

4 General discussion

4.1 *Pseudomonas putida* as a host for synthetic biology applications

P. putida is often described as a workhorse for industrial biotechnology. However, despite a few examples such as the production of rhamnolipids, PHAs, and 2,5-furandicarboxylic acid, the biotechnological use of *P. putida* is still in its infancy.[15] This can either be attributed to the unfavorable characteristics of *P. putida* or to the biotechnological industry itself.

Traditional hosts like *S. cerevisiae* and *E. coli* have the advantage of an already developed infrastructure and decades of intense, industrial research. While *P. putida* has accumulated a significant amount of scientific literature, it lacks a compelling reason to replace well-established hosts. However, the primary advantage of *P. putida* lies in its potential role in a circular bioeconomy.[15] The use of alternative feedstocks, especially those that humans cannot consume, can be valorized by appropriate microbial hosts. Among these alternative carbon sources, lignocellulose stands out as the most prominent representative. The global annual production of this biomass is around 180 billion tons, of which only 3% are converted into value-added compounds.[360]

Lignocellulose comprises three components: cellulose, hemicellulose, and lignin. While cellulose and hemicellulose are primarily C5 and C6 sugars that most microbes can metabolize, lignin is a complex mixture of aromatic compounds.[360] To fully utilize the carbon stored in lignocellulose, *P. putida* is an exceptional choice due to its extensive metabolism for phenolic compounds.[361] The production of PHAs from lignin hydrolysate has been recently demonstrated, and titers can be further improved with a few genetic modifications.[361]

Another major component of lignocellulose is xylose.[360] Natively, *P. putida* does not possess the required pathway to metabolize this C5 sugar. However, due to its great genetic accessibility, it is possible to enable growth on xylose through the expression of heterologous genes.[362] Thus, a strain that also utilizes xylose could be significantly more efficient in converting biomass hydrolysates.

Beyond lignocellulosic feedstocks, hydrolysates derived from plastic depolymerization are becoming more relevant.[363] Instead of a cost-intense and hazardous chemical recycling process, biological upcycling using engineered *P. putida* strains is an environment friendly alternative.[20]

Overall, it is very likely that *P. putida* will play a major role in the future of a circular bioeconomy. However, it will also be crucial to expand the number of potential products beyond current target molecules.

While *P. putida*'s omnivorous characteristics are beneficial for the use of alternative feedstocks, it can be very challenging to prevent product or precursor degradation. As demonstrated by our research, the use of RB-TnSeq is a powerful tool to rationally engineer *P. putida.*[56,172] Furthermore, we have generated and published a substantial amount of BarSeq data for this organism, which can be freely accessed via the fitness browser (fit.genomics.lbl.gov).[11,12,48,172]

In these datasets we cover most of *P. putida*'s alcohol, fatty acid, aromatic, and nitrogen metabolism. The chances of finding significant phenotypes for a target product or structurally similar molecules are relatively high. However, the fitness browser can only give a hint about the full spectrum of the deposited BarSeq data and manual post-processing is still an important factor. This can make it more difficult for other researchers to interpret BarSeq results. The implementation of these data into major genome database collections such as BioCyc would facilitate its accessibility.[157]

The reliability of functional analyses using RB-TnSeq is mostly influenced by gene redundancy and essentiality. The presence of genes that are similar in function usually results in minor fitness defects and can be easily overlooked. The complementation of BarSeq with omics technologies can be a way to facilitate the identification of redundant genes.[56] Furthermore, the combination of sole carbon source and nitrogen source assays has been shown to improve BarSeq coverage, especially in the case of transporters.[11]

By definition, RB-TnSeq libraries can only be prepared for non-essential genes.[48] However, some conditions involve genes that are part of the central metabolism of a cell. As a result, BarSeq assays never cover the entire genome of an organism. An interesting approach to solve this problem could be the complementation of BarSeq with CRISPRi-seq.[63] The CRISPRi-seq library would need to contain only sgRNA for essential genes, which would facilitate library construction and lower overall costs.

All in all, *P. putida* has proven to be an extremely versatile host, and the wealth of available literature for this microbe makes it an excellent candidate for synthetic biology applications.[219] However, despite multiple decades of research, we have not fully exploited its potential for the expression of heterologous pathways. A very recent development is the use of *P. putida* as a host for engineered T1PKSs.[121] Yet, the promise of engineering T1PKSs has not been fulfilled.

4.2 The future of polyketide synthase engineering

Due to their logical design, T1PKSs have emerged as a compelling target for retrobiosynthesis approaches. Individual domains, or even entire modules, can be systematically exchanged to tailor the synthesis of almost any desired molecule.[117] In addition, the vast diversity of PKSs delivers a virtually limitless number of parts.[118] However, the reality of engineering these intricate systems presents a contrasting perspective, often resulting in non-functional proteins or extremely low titers.[132,364] It takes a lot of experience and sometimes several years to get from a target molecule to a functioning PKS design. Furthermore, most PKS research focuses on a very limited number of well-studied model PKSs.[124,305]

The highest success rate seems to be achieved by truncating large PKSs after the first or second extension of the loaded substrate.[119,305] Choosing a PKS module that already performs all of the required modifications to the α-carbon residue will further facilitate the build process. In general, the closer the native PKS system is to the final design, the higher the chances are for success.

To test a synthetic PKS system, it must to be expressed in a suitable host. Ideally, researcher can exploit the native host, allowing the engineered PKS to be characterized *in vivo*. However, many of the source organisms are either inaccessible or lack the necessary tools for an *in vivo* characterization. As a result, researchers often turn to heterologous expression. Among these heterologous hosts, *E. coli* is undoubtedly the most popular choice.[306] While *in vivo* testing in *E. coli* has its limitations due to missing PKS extenders, its protein expression machinery is very robust and widely used for recombinant protein production.[365] The *in vitro* characterization of a PKS system can offer certain advantages over *in vivo* studies. For instance, the PKS would operate completely independent of the host metabolism, and all required cofactors and substrates can be added directly to the reaction mixture. This can be particularly useful if the pathway for the required mCoA derivative is unknown or difficult to express.[124] Furthermore, *in vitro* studies make it possible to determine enzyme kinetics and identify the optimal starter and extender units.[119] However, purifying enzymes as large as PKSs can be very challenging and time-consuming. Therefore, prioritizing the introduction of a PKS pathway into a heterologous host and evaluating PKS engineering efforts *in vivo* is advisable.

Regrettably, PKS engineering often involves significant amount of trial and error. Even minor modifications, such as a single amino acid difference during domain exchanges, can render the protein inactive.[128] Consequently, having high testing throughput, which allows for the evaluation of many designs simultaneously, is extremely beneficial. The combination of the SAGE system with BEDEX vectors, as developed in this work, appears to provide the

required throughput.[92,121] Fast, efficient and stable integrations into the target host genome significantly facilitate the testing phase, which allows to screen hundreds of designs. Moreover, BEDEX vectors not only offer precise control over gene expression but also seem to improve cloning efficiency. The repression of the *LlacO1* promoter in the presence of LacI is strong enough to almost entirely prevent leaky expression of the target gene.[366] This is most likely caused by the strong constitutive *lacIq* promoter driving the expression of the *lacI* gene and the presence of two *lacO* binding motifs in the promoter sequence of *LlacO1*. The BEDEX principle also made it possible to clone and integrate the highly toxic PCCase complex, *accA2* and *pccB*. Prior to the development of BEDEX vectors, it was not possible to clone this operon without the cloning host *E. coli* mutating or excising the related genes.

Nevertheless, even with all these measures in place, most PKS designs are abandoned before evaluation. The expression efficiency of PKS is typically very low, and introducing unnatural boundaries can further reduce expression levels.[123] In our work, we managed to significantly enhance expression efficiency by employing codon optimization before heterologous PKS expression.[121] By comparing strains containing different codon versions of an engineered PKS, we confirmed the detrimental effects of codon usage on protein, transcript, and product levels. Additionally, we established *P. putida* and *C. glutamicum* as alternative PKS hosts. Although both hosts were suitable for the heterologous expression of engineered PKSs, *P. putida* appears to be superior due to its rapid growth rate, genetic accessibility, diverse metabolism, and ease of handling.[22] The potential application of this host in a circular bioeconomy positions it as a future-proof choice.[360] Furthermore, the amount of available BarSeq data for *P. putida* streamlines essential modifications to its metabolism, such as preventing product degradation.[11,12,56,172]

Once a basic PKS design is well-established within a heterologous host and demonstrates the expected activity, the system can be further characterized by domain exchanges.[128] For example, the exchange of the AT domain can drastically increase the number of possible products.[124] Other less utilized strategies include full reduction loop swaps or KR exchanges.[132,367] The presented approach might be very slow in the beginning but will significantly increase the amount of usable PKS modules over time, eventually leading to an arsenal of functional PKSs. The final step will involve to orchestrate these modules and rationally combine them using appropriate linker domains.[368]

Ultimately, we need to further investigate the basics of PKS engineering before we can fully leverage the potential of a PKS production platform. Codon optimization and a high-throughput *in vivo* testing system will set a strong foundation for future exploration of novel PKSs.

4.3 Conclusion and Outlook

This work focused on establishing the host *P. putida* KT2440 as a chassis for testing novel PKSs. A major challenge, because alternative hosts are scarce and the enormous size of these enzymes results in high metabolic burden. Furthermore, despite all the advantages of *P. putida*, it can be difficult to tame. Due to the diverse metabolism of *P. putida* and its ability to degrade a wide range of carbon and nitrogen sources, we first studied this omnivorous bacterium extensively using RB-TnSeq. By feeding a wide range of substrates to a barcoded transposon mutant library, we were able to identify hundreds of specific phenotypes and characterized most of its carbon and nitrogen metabolism. This knowledge enabled us to better understand and control *P. putida* as a host for synthetic biology applications.

A highly controversial topic, not just in the field of PKS engineering, is the optimization of codons for the heterologous expression of genes. In a systematic approach, we utilized a model PKS and three heterologous hosts to test various algorithms for codon optimization. This study underscored the importance of selecting the appropriate codon optimization tool, suggesting that commercial tools might not always be the best choice. The development of a free online tool that implements the tested algorithms should assist the synthetic biology community in standardizing codon optimization and reporting. Many scientific publications do not specify the algorithm used, potentially leading to less reproducible results due to the significant effects of codon usage. We further established a high-throughput testing approach to facilitate the *in vivo* evaluation of PKSs.

The final chapter of this work integrated RB-TnSeq host engineering with a high-throughput testing approach for the development of δ-lactone producing PKSs. We demonstrated the effectiveness of BarSeq in guiding metabolic engineering efforts and enhancing polyketide titers. Moreover, we showed that the most simplistic PKS design, with minimal modifications, proved to be the most effective approach.

The tools and testing procedures developed in this study will deepen our understanding of alternative PKS hosts and aid in the development of PKS production platforms for synthesizing unnatural polyketides. Broadening the spectrum of potential and meaningful products through PKS engineering might increase the interest in the industrial applications of these captivating enzymes.

5 References

1. Parte AC, Sardà Carbasse J, Meier-Kolthoff JP, Reimer LC, Göker M. List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol*. 2020;70(11):5607-5612. doi:10.1099/ijsem.0.004332

2. Anzai Y, Kim H, Park JY, Wakabayashi H, Oyaizu H. Phylogenetic affiliation of the pseudomonads based on 16S rRNA sequence. *Int J Syst Evol Microbiol*. 2000;50 Pt 4:1563-1589. doi:10.1099/00207713-50-4-1563

3. Nikolaidis M, Mossialos D, Oliver SG, Amoutzias GD. Comparative snalysis of the core proteomes among the *Pseudomonas* major evolutionary groups reveals species-specific adaptations for Pseudomonas aeruginosa and *Pseudomonas chlororaphis*. *Diversity (Basel)*. 2020;12(8):289. doi:10.3390/d12080289

4. Yabuuchi E, Kosako Y, Oyaizu H, Yano I, Hotta H, Hashimoto Y, Ezaki T, Arakawa M. Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia* cepacia (Palleroni and Holmes 1981) comb. nov. *Microbiol Immunol*. 1992;36(12):1251-1275.

5. Yabuuchi E, Kosako Y, Yano I, Hotta H, Nishiuchi Y. Transfer of two *Burkholderia* and an *Alcaligenes* species to *Ralstonia* gen. Nov.: proposal of *Ralstonia pickettii* (Ralston, Palleroni and Doudoroff 1973) comb. Nov., *Ralstonia solanacearum* (Smith 1896) comb. Nov. and *Ralstonia eutropha* (Davis 1969) comb. Nov. *Microbiol Immunol*. 1995;39(11):897-904.

6. Meyer J-M, Geoffroy VA, Baida N, Gardan L, Izard D, Lemanceau P, Achouak W, Palleroni NJ. Siderophore typing, a powerful tool for the identification of fluorescent and nonfluorescent pseudomonads. *Appl Environ Microbiol*. 2002;68(6):2745-2753. doi:10.1128/AEM.68.6.2745-2753.2002

7. Elliott RP. Some properties of pyoverdine, the water-soluble fluorescent pigment of the pseudomonads. *Appl Microbiol*. 1958;6(4):241-246. doi:10.1128/am.6.4.241-246.1958

8. Neilands JB. Microbial envelope proteins related to iron. *Annu Rev Microbiol*. 1982;36:285-309. doi:10.1146/annurev.mi.36.100182.001441

9. Diggle SP, Whiteley M. Microbe Profile: *Pseudomonas aeruginosa*: opportunistic pathogen and lab rat. *Microbiology (Reading, Engl)*. 2020;166(1):30-33. doi:10.1099/mic.0.000860

10. Elborn JS. Cystic fibrosis. *Lancet*. 2016;388(10059):2519-2531. doi:10.1016/S0140-6736(16)00576-6

11. Schmidt M, Pearson AN, Incha MR, Thompson MG, Baidoo EEK, Kakumanu R, Mukhopadhyay A, Shih PM, Deutschbauer AM, Blank LM, Keasling JD. Nitrogen metabolism in *Pseudomonas putida*: functional analysis using random barcode transposon sequencing. *Appl Environ Microbiol*. 2022;88(7):e0243021. doi:10.1128/aem.02430-21

12. Thompson MG, Incha MR, Pearson AN, Schmidt M, Sharpless WA, Eiben CB, Cruz-Morales P, Blake-Hedges JM, Liu Y, Adams CA, Haushalter RW, Krishna RN, Lichtner P, Blank LM, Mukhopadhyay A, Deutschbauer AM, Shih PM, Keasling JD. Fatty acid and alcohol metabolism in *Pseudomonas putida*: functional analysis using random barcode transposon sequencing. *Appl Environ Microbiol*. 2020;86(21):e01665-20. doi:10.1128/AEM.01665-20

13. Kieboom J, Dennis JJ, de Bont JA, Zylstra GJ. Identification and molecular characterization of an efflux pump involved in *Pseudomonas putida* S12 solvent tolerance. *J Biol Chem*. 1998;273(1):85-91. doi:10.1074/jbc.273.1.85

14. Kampers LFC, Volkers RJM, Martins Dos Santos VAP. *Pseudomonas putida* KT2440 is HV1 certified, not GRAS. *Microb Biotechnol*. 2019;12(5):845-848. doi:10.1111/1751-7915.13443

15. Weimer A, Kohlstedt M, Volke DC, Nikel PI, Wittmann C. Industrial biotechnology of *Pseudomonas putida*: advances and prospects. *Appl Microbiol Biotechnol*. 2020;104(18):7745-7766. doi:10.1007/s00253-020-10811-9

16. Poltronieri P, Kumar P. Polyhydroxyalkanoates (phas) in industrial applications. In: Martínez LMT, Kharissova OV, Kharisov BI, eds. *Handbook of Ecomaterials*. Springer International Publishing; 2017:1-30. doi:10.1007/978-3-319-48281-1_70-2

17. Poblete-Castro I, Becker J, Dohnt K, dos Santos VM, Wittmann C. Industrial biotechnology of *Pseudomonas putida* and related species. *Appl Microbiol Biotechnol*. 2012;93(6):2279-2290. doi:10.1007/s00253-012-3928-0

18. Mozejko-Ciesielska J, Szacherska K, Marciniak P. *Pseudomonas* species as producers of eco-friendly polyhydroxyalkanoates. *J Polym Environ*. 2019;27(6):1151-1166. doi:10.1007/s10924-019-01422-1

19. Incha MR, Thompson MG, Blake-Hedges JM, Pearson AN, Schmidt M, Deutschbauer A, Keasling JD. Leveraging host metabolism for bisdemethoxycurcumin production in *Pseudomonas putida*. *BioRxiv*. Published online August 31, 2019. doi:10.1101/753889

20. Werner AZ, Clare R, Mand TD, Pardo I, Ramirez KJ, Haugen SJ, Bratti F, Dexter GN, Elmore JR, Huenemann JD, Peabody GL, Johnson CW, Rorrer NA, Salvachúa D, Guss AM, Beckham GT. Tandem chemical deconstruction and biological upcycling of poly(ethylene terephthalate) to β-ketoadipic acid by *Pseudomonas putida* KT2440. *Metab Eng*. 2021;67:250-261. doi:10.1016/j.ymben.2021.07.005

21. Ackermann YS, Li W-J, Op de Hipt L, Niehoff P-J, Casey W, Polen T, Köbbing S, Ballerstedt H, Wynands B, O'Connor K, Blank LM, Wierckx N. Engineering adipic acid metabolism in *Pseudomonas putida*. *Metab Eng*. 2021;67:29-40. doi:10.1016/j.ymben.2021.05.001

22. Loeschcke A, Thies S. Engineering of natural product biosynthesis in *Pseudomonas putida*. *Curr Opin Biotechnol*. 2020;65:213-224. doi:10.1016/j.copbio.2020.03.007

23. Köhler KAK, Rückert C, Schatschneider S, Vorhölter F-J, Szczepanowski R, Blank LM, Niehaus K, Goesmann A, Pühler A, Kalinowski J, Schmid A. Complete genome

sequence of *Pseudomonas* sp. strain VLB120 a solvent tolerant, styrene degrading bacterium, isolated from forest soil. *J Biotechnol*. 2013;168(4):729-730. doi:10.1016/j.jbiotec.2013.10.016

24. Wynands B, Lenzen C, Otto M, Koch F, Blank LM, Wierckx N. Metabolic engineering of *Pseudomonas taiwanensis* VLB120 with minimal genomic modifications for high-yield phenol production. *Metab Eng*. 2018;47:121-133. doi:10.1016/j.ymben.2018.03.011

25. Otto M, Wynands B, Lenzen C, Filbig M, Blank LM, Wierckx N. Rational engineering of phenylalanine accumulation in *Pseudomonas taiwanensis* to enable high-yield production of trans-cinnamate. *Front Bioeng Biotechnol*. 2019;7:312. doi:10.3389/fbioe.2019.00312

26. Köhler KAK, Blank LM, Frick O, Schmid A. D-Xylose assimilation via the Weimberg pathway by solvent-tolerant *Pseudomonas taiwanensis* VLB120. *Environ Microbiol*. 2015;17(1):156-170. doi:10.1111/1462-2920.12537

27. Xin X-F, Kvitko B, He SY. *Pseudomonas syringae*: what it takes to be a pathogen. *Nat Rev Microbiol*. 2018;16(5):316-328. doi:10.1038/nrmicro.2018.17

28. Hirano SS, Upper CD. Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae* - a pathogen, ice nucleus, and epiphyte. *Microbiol Mol Biol Rev*. 2000;64(3):624-653. doi:10.1128/mmbr.64.3.624-653.2000

29. Lindow SE, Arny DC, Upper CD. Bacterial ice nucleation: a factor in frost injury to plants. *Plant Physiol*. 1982;70(4):1084-1089. doi:10.1104/pp.70.4.1084

30. Morris CE, Monteil CL, Berge O. The life history of *Pseudomonas syringae*: linking agriculture to earth system processes. *Annu Rev Phytopathol*. 2013;51:85-104. doi:10.1146/annurev-phyto-082712-102402

31. Auffray C, Imbeaud S, Roux-Rouquié M, Hood L. From functional genomics to systems biology: concepts and practices. *C R Biol*. 2003;326(10-11):879-892. doi:10.1016/j.crvi.2003.09.033

32. Kobras CM, Fenton AK, Sheppard SK. Next-generation microbiology: from comparative genomics to gene function. *Genome Biol*. 2021;22(1):123. doi:10.1186/s13059-021-02344-9

33. Mardis ER. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet*. 2008;9:387-402. doi:10.1146/annurev.genom.9.081307.164359

34. Schuster SC. Next-generation sequencing transforms today's biology. *Nat Methods*. 2008;5(1):16-18. doi:10.1038/nmeth1156

35. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*. 2006;2:2006.0008. doi:10.1038/msb4100050

36. Giaever G, Chu AM, Ni L, Connelly C, Riles L, Véronneau S, Dow S, Lucau-Danila A, Anderson K, André B, Arkin AP, Astromoff A, El-Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro S, Curtiss M, Davis K, Deutschbauer A, Entian K-D, Flaherty P, Foury F, Garfinkel DJ, Gerstein M, Gotte D, Güldener U, Hegemann JH, Hempel S, Herman Z, Jaramillo DF, Kelly DE, Kelly SL, Kötter P, LaBonte D, Lamb DC, Lan N, Liang H, Liao H, Liu L, Luo C, Lussier M, Mao R, Menard P, Ooi SL, Revuelta JL, Roberts CJ, Rose M, Ross-Macdonald P, Scherens B, Schimmack G, Shafer B, Shoemaker DD, Sookhai-Mahadeo S, Storms RK, Strathern JN, Valle G, Voet M, Volckaert G, Wang C, Ward TR, Wilhelmy J, Winzeler EA, Yang Y, Yen G, Youngman E, Yu K, Bussey H, Boeke JD, Snyder M, Philippsen P, Davis RW, Johnston M. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature*. 2002;418(6896):387-391. doi:10.1038/nature00935

37. van Opijnen T, Bodi KL, Camilli A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat Methods*. 2009;6(10):767-772. doi:10.1038/nmeth.1377

38. Smith AM, Heisler LE, Mellor J, Kaper F, Thompson MJ, Chee M, Roth FP, Giaever G, Nislow C. Quantitative phenotyping via deep barcode sequencing. *Genome Res*. 2009;19(10):1836-1842. doi:10.1101/gr.093955.109

39. Cain AK, Barquist L, Goodman AL, Paulsen IT, Parkhill J, van Opijnen T. A decade of advances in transposon-insertion sequencing. *Nat Rev Genet*. 2020;21(9):526-540. doi:10.1038/s41576-020-0244-x

40. Lampe DJ, Churchill ME, Robertson HM. A purified *mariner* transposase is sufficient to mediate transposition in vitro. *EMBO J*. 1996;15(19):5470-5479. doi:10.1002/j.1460-2075.1996.tb00930.x

41. Goodman AL, McNulty NP, Zhao Y, Leip D, Mitra RD, Lozupone CA, Knight R, Gordon JI. Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe*. 2009;6(3):279-289. doi:10.1016/j.chom.2009.08.003

42. Langridge GC, Phan M-D, Turner DJ, Perkins TT, Parts L, Haase J, Charles I, Maskell DJ, Peters SE, Dougan G, Wain J, Parkhill J, Turner AK. Simultaneous assay of every *Salmonella Typhi* gene using one million transposon mutants. *Genome Res*. 2009;19(12):2308-2316. doi:10.1101/gr.097097.109

43. Gawronski JD, Wong SMS, Giannoukos G, Ward DV, Akerley BJ. Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for Haemophilus genes required in the lung. *Proc Natl Acad Sci USA*. 2009;106(38):16422-16427. doi:10.1073/pnas.0906627106

44. Morgan RD, Dwinell EA, Bhatia TK, Lang EM, Luyten YA. The MmeI family: type II restriction-modification enzymes that employ single-strand modification for host protection. *Nucleic Acids Res*. 2009;37(15):5208-5221. doi:10.1093/nar/gkp534

45. Jacobs MA, Alwood A, Thaipisuttikul I, Spencer D, Haugen E, Ernst S, Will O, Kaul R, Raymond C, Levy R, Chun-Rong L, Guenthner D, Bovee D, Olson MV, Manoil C. Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci USA*. 2003;100(24):14339-14344. doi:10.1073/pnas.2036282100

46. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, Mittmann M, Pai C, Rebischung C, Revuelta JL, Riles L, Roberts CJ, Ross-MacDonald P, Scherens B, Snyder M, Sookhai-Mahadeo S, Storms RK, Véronneau S, Voet M, Volckaert G, Ward TR, Wysocki R, Yen GS, Yu K, Zimmermann K, Philippsen P, Johnston M, Davis RW. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science*. 1999;285(5429):901-906. doi:10.1126/science.285.5429.901

47. Robinson DG, Chen W, Storey JD, Gresham D. Design and analysis of Bar-seq experiments. *G3 (Bethesda)*. 2014;4(1):11-18. doi:10.1534/g3.113.008565

48. Wetmore KM, Price MN, Waters RJ, Lamson JS, He J, Hoover CA, Blow MJ, Bristow J, Butland G, Arkin AP, Deutschbauer A. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. *MBio*. 2015;6(3):e00306-15. doi:10.1128/mBio.00306-15

49. Deutschbauer A, Price MN, Wetmore KM, Shao W, Baumohl JK, Xu Z, Nguyen M, Tamse R, Davis RW, Arkin AP. Evidence-based annotation of gene function in *Shewanella oneidensis* MR-1 using genome-wide fitness profiling across 121 conditions. *PLoS Genet*. 2011;7(11):e1002385. doi:10.1371/journal.pgen.1002385

50. Nichols RJ, Sen S, Choo YJ, Beltrao P, Zietek M, Chaba R, Lee S, Kazmierczak KM, Lee KJ, Wong A, Shales M, Lovett S, Winkler ME, Krogan NJ, Typas A, Gross CA. Phenotypic landscape of a bacterial cell. *Cell*. 2011;144(1):143-156. doi:10.1016/j.cell.2010.11.052

51. Bosak T, Schubotz F, de Santiago-Torio A, Kuehl JV, Carlson HK, Watson N, Daye M, Summons RE, Arkin AP, Deutschbauer AM. System-wide adaptations of *Desulfovibrio alaskensis* G20 to phosphate-limited conditions. *PLoS ONE*. 2016;11(12):e0168719. doi:10.1371/journal.pone.0168719

52. Thompson MG, Blake-Hedges JM, Cruz-Morales P, Barajas JF, Curran SC, Eiben CB, Harris NC, Benites VT, Gin JW, Sharpless WA, Twigg FF, Skyrud W, Krishna RN, Pereira JH, Baidoo EEK, Petzold CJ, Adams PD, Arkin AP, Deutschbauer AM, Keasling JD. Massively parallel fitness profiling reveals multiple novel enzymes in *Pseudomonas putida* lysine metabolism. *MBio*. 2019;10(3). doi:10.1128/mBio.02577-18

53. Liu H, Shiver AL, Price MN, Carlson HK, Trotter VV, Chen Y, Escalante V, Ray J, Hern KE, Petzold CJ, Turnbaugh PJ, Huang KC, Arkin AP, Deutschbauer AM. Functional genetics of human gut commensal *Bacteroides thetaiotaomicron* reveals metabolic requirements for growth across environments. *Cell Rep*. 2021;34(9):108789. doi:10.1016/j.celrep.2021.108789

54. Thorgersen MP, Xue J, Majumder ELW, Trotter VV, Ge X, Poole FL, Owens TK, Lui LM, Nielsen TN, Arkin AP, Deutschbauer AM, Siuzdak G, Adams MWW. Deciphering microbial metal toxicity responses via random bar code transposon site sequencing and activity-based metabolomics. *Appl Environ Microbiol*. 2021;87(21):e0103721. doi:10.1128/AEM.01037-21

55. Price MN, Wetmore KM, Waters RJ, Callaghan M, Ray J, Liu H, Kuehl JV, Melnyk RA, Lamson JS, Suh Y, Carlson HK, Esquivel Z, Sadeeshkumar H, Chakraborty R, Zane GM, Rubin BE, Wall JD, Visel A, Bristow J, Blow MJ, Arkin AP, Deutschbauer AM. Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature*. 2018;557(7706):503-509. doi:10.1038/s41586-018-0124-0

56. Thompson MG, Valencia LE, Blake-Hedges JM, Cruz-Morales P, Velasquez AE, Pearson AN, Sermeno LN, Sharpless WA, Benites VT, Chen Y, Baidoo EEK, Petzold CJ, Deutschbauer AM, Keasling JD. Omics-driven identification and elimination of valerolactam catabolism in *Pseudomonas putida* KT2440 for increased product titer. *Metab Eng Commun*. 2019;9:e00098. doi:10.1016/j.mec.2019.e00098

57. Borchert AJ, Bleem A, Beckham GT. RB-TnSeq identifies genetic targets for improved tolerance of *Pseudomonas putida* towards compounds relevant to lignin conversion. *Metab Eng*. 2023;77:208-218. doi:10.1016/j.ymben.2023.04.007

58. Thompson MG, Pearson AN, Barajas JF, Cruz-Morales P, Sedaghatian N, Costello Z, Garber ME, Incha MR, Valencia LE, Baidoo EEK, Martin HG, Mukhopadhyay A, Keasling JD. Identification, characterization, and application of a highly sensitive lactam biosensor from *Pseudomonas putida*. *ACS Synth Biol*. 2020;9(1):53-62. doi:10.1021/acssynbio.9b00292

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

60. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res*. 2013;41(15):7429-7437. doi:10.1093/nar/gkt520

61. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 2013;152(5):1173-1183. doi:10.1016/j.cell.2013.02.022

62. Liu X, Gallay C, Kjos M, Domenech A, Slager J, van Kessel SP, Knoops K, Sorg RA, Zhang J-R, Veening J-W. High-throughput CRISPRi phenotyping identifies new essential genes in *Streptococcus pneumoniae*. *Mol Syst Biol*. 2017;13(5):931. doi:10.15252/msb.20167449

63. de Bakker V, Liu X, Bravo AM, Veening J-W. CRISPRi-seq for genome-wide fitness quantification in bacteria. *Nat Protoc*. 2022;17(2):252-281. doi:10.1038/s41596-021-00639-6

64. Jiang W, Oikonomou P, Tavazoie S. Comprehensive genome-wide perturbations via CRISPR adaptation reveal complex genetics of antibiotic sensitivity. *Cell*. 2020;180(5):1002-1017.e31. doi:10.1016/j.cell.2020.02.007

65. de Wet TJ, Winkler KR, Mhlanga M, Mizrahi V, Warner DF. Arrayed CRISPRi and quantitative imaging describe the morphotypic landscape of essential mycobacterial genes. *eLife*. 2020;9. doi:10.7554/eLife.60083

66. Kisumi M, Komatsubara S, Chibata I. Enhancement of isoleucine hydroxamate-mediated growth inhibition and improvement of isoleucine-producing strains of *Serratia marcescens*. *Appl Environ Microbiol*. 1977;34(6):647-653. doi:10.1128/aem.34.6.647-653.1977

67. Park JH, Lee SY. Towards systems metabolic engineering of microorganisms for amino acid production. *Curr Opin Biotechnol*. 2008;19(5):454-460. doi:10.1016/j.copbio.2008.08.007

68. Schrumpf B, Eggeling L, Sahm H. Isolation and prominent characteristics of an L-lysine hyperproducing strain of *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol*. 1992;37(5). doi:10.1007/BF00240726

69. Bailey JE. Toward a science of metabolic engineering. *Science*. 1991;252(5013):1668-1675. doi:10.1126/science.2047876

70. Cohen SN, Chang AC, Boyer HW, Helling RB. Construction of biologically functional bacterial plasmids in vitro. *Proc Natl Acad Sci USA*. 1973;70(11):3240-3244. doi:10.1073/pnas.70.11.3240

71. Bräu B, Sahm H. Cloning and expression of the structural gene for pyruvate decarboxylase of *Zymomonas mobilis* in *Escherichia coli*. *Arch Microbiol*. 1986;144(3):296-301. doi:10.1007/BF00410966

72. Chang CN, Rey M, Bochner B, Heyneker H, Gray G. High-level secretion of human growth hormone by *Escherichia coli*. *Gene*. 1987;55(2-3):189-196. doi:10.1016/0378-1119(87)90279-4

73. Volk MJ, Tran VG, Tan S-I, Mishra S, Fatma Z, Boob A, Li H, Xue P, Martin TA, Zhao H. Metabolic engineering: methodologies and applications. *Chem Rev*. 2023;123(9):5521-5570. doi:10.1021/acs.chemrev.2c00403

74. Zeng A-P, Biebl H. Bulk chemicals from biotechnology: the case of 1,3-propanediol production and the new trends. *Adv Biochem Eng Biotechnol*. 2002;74:239-259. doi:10.1007/3-540-45736-4_11

75. Ikeda M, Katsumata R. Metabolic engineering to produce tyrosine or phenylalanine in a tryptophan-producing *Corynebacterium glutamicum* Strain. *Appl Environ Microbiol*. 1992;58(3):781-785. doi:10.1128/aem.58.3.781-785.1992

76. Kim J, Salvador M, Saunders E, González J, Avignone-Rossa C, Jiménez JI. Properties of alternative microbial hosts used in synthetic biology: towards the design of a modular chassis. *Essays Biochem*. 2016;60(4):303-313. doi:10.1042/EBC20160015

77. Katsumata R, Ikeda M. Hyperproduction of tryptophan in *Corynebacterium glutamicum* by pathway engineering. *Nat Biotechnol*. 1993;11(8):921-925. doi:10.1038/nbt0893-921

78. Blombach B, Grünberger A, Centler F, Wierckx N, Schmid J. Exploiting unconventional prokaryotic hosts for industrial biotechnology. *Trends Biotechnol*. 2022;40(4):385-397. doi:10.1016/j.tibtech.2021.08.003

79. Tan D, Xue Y-S, Aibaidula G, Chen G-Q. Unsterile and continuous production of polyhydroxybutyrate by *Halomonas* TD01. *Bioresour Technol*. 2011;102(17):8130-8136. doi:10.1016/j.biortech.2011.05.068

80. Yin J, Chen J-C, Wu Q, Chen G-Q. Halophiles, coming stars for industrial biotechnology. *Biotechnol Adv*. 2015;33(7):1433-1442. doi:10.1016/j.biotechadv.2014.10.008

81. Quillaguamán J, Guzmán H, Van-Thuoc D, Hatti-Kaul R. Synthesis and production of polyhydroxyalkanoates by halophiles: current potential and future prospects. *Appl Microbiol Biotechnol*. 2010;85(6):1687-1696. doi:10.1007/s00253-009-2397-6

82. Stephanopoulos G, Vallino JJ. Network rigidity and metabolic engineering in metabolite overproduction. *Science*. 1991;252(5013):1675-1681. doi:10.1126/science.1904627

83. Marx A, de Graaf AA, Wiechert W, Eggeling L, Sahm H. Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing. *Biotechnology and Bioengineering*. Published online January 20, 1996.

84. Burgard AP, Pharkya P, Maranas CD. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng*. 2003;84(6):647-657. doi:10.1002/bit.10803

85. Capecchi MR. Altering the genome by homologous recombination. *Science*. 1989;244(4910):1288-1292. doi:10.1126/science.2660260

86. Güldener U, Heck S, Fielder T, Beinhauer J, Hegemann JH. A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res*. 1996;24(13):2519-2524. doi:10.1093/nar/24.13.2519

87. Murphy KC. Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. *J Bacteriol*. 1998;180(8):2063-2071. doi:10.1128/JB.180.8.2063-2071.1998

88. Phelan RM, Sachs D, Petkiewicz SJ, Barajas JF, Blake-Hedges JM, Thompson MG, Reider Apel A, Rasor BJ, Katz L, Keasling JD. Development of next generation synthetic biology tools for use in *Streptomyces venezuelae*. *ACS Synth Biol*. 2017;6(1):159-166. doi:10.1021/acssynbio.6b00202

89. Wang W, Zheng G, Lu Y. Recent advances in strategies for the cloning of natural product biosynthetic gene clusters. *Front Bioeng Biotechnol*. 2021;9:692797. doi:10.3389/fbioe.2021.692797

90. Kuhstoss S, Rao RN. Analysis of the integration function of the streptomycete bacteriophage φC31. *J Mol Biol*. 1991;222(4):897-908. doi:10.1016/0022-2836(91)90584-S

91. Morita K, Morimura K, Fusada N, Komatsu M, Ikeda H, Hirano N, Takahashi H. Site-specific genome integration in alphaproteobacteria mediated by TG1 integrase. *Appl Microbiol Biotechnol*. 2012;93(1):295-304. doi:10.1007/s00253-011-3545-3

92. Elmore JR, Dexter GN, Baldino H, Huenemann JD, Francis R, Peabody GL, Martinez-Baird J, Riley LA, Simmons T, Coleman-Derr D, Guss AM, Egbert RG. High-throughput genetic engineering of nonmodel and undomesticated bacteria via iterative site-specific genome integration. *Sci Adv*. 2023;9(10):eade1285. doi:10.1126/sciadv.ade1285

93. Köbbing S. Entwicklung von Werkzeugen der synthetischen Biologie für Pseudomonas putida. *RWTH Aachen University*. Published online 2020. doi:10.18154/rwth-2020-09053

94. de las Heras A, Carreño CA, de Lorenzo V. Stable implantation of orthogonal sensor circuits in Gram-negative bacteria for environmental release. *Environ Microbiol*. 2008;10(12):3305-3316. doi:10.1111/j.1462-2920.2008.01722.x

95. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 2014;346(6213):1258096. doi:10.1126/science.1258096

96. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096):816-821. doi:10.1126/science.1225829

97. Li Y, Lin Z, Huang C, Zhang Y, Wang Z, Tang Y-J, Chen T, Zhao X. Metabolic engineering of *Escherichia coli* using CRISPR-Cas9 meditated genome editing. *Metab Eng*. 2015;31:13-21. doi:10.1016/j.ymben.2015.06.006

98. Lian J, HamediRad M, Zhao H. Advancing metabolic engineering of *Saccharomyces cerevisiae* using the CRISPR/Cas system. *Biotechnol J*. 2018;13(9):e1700601. doi:10.1002/biot.201700601

99. Dong H, Cui Y, Zhang D. CRISPR/Cas technologies and their applications in *Escherichia coli*. *Front Bioeng Biotechnol*. 2021;9:762676. doi:10.3389/fbioe.2021.762676

100. Aparicio T, de Lorenzo V, Martínez-García E. CRISPR/Cas9-based counterselection boosts recombineering efficiency in *Pseudomonas putida*. *Biotechnol J*. 2018;13(5):e1700161. doi:10.1002/biot.201700161

101. Peng F, Wang X, Sun Y, Dong G, Yang Y, Liu X, Bai Z. Efficient gene editing in *Corynebacterium glutamicum* using the CRISPR/Cas9 system. *Microb Cell Fact*. 2017;16(1):201. doi:10.1186/s12934-017-0814-6

102. McGinness KE, Baker TA, Sauer RT. Engineering controllable protein degradation. *Mol Cell*. 2006;22(5):701-707. doi:10.1016/j.molcel.2006.04.027

103. Yang Y, Lin Y, Li L, Linhardt RJ, Yan Y. Regulating malonyl-CoA metabolism via synthetic antisense RNAs for enhanced biosynthesis of natural products. *Metab Eng*. 2015;29:217-226. doi:10.1016/j.ymben.2015.03.018

104. Hartline CJ, Schmitz AC, Han Y, Zhang F. Dynamic control in metabolic engineering: Theories, tools, and applications. *Metab Eng*. 2021;63:126-140. doi:10.1016/j.ymben.2020.08.015

105. Wu J, Du G, Chen J, Zhou J. Enhancing flavonoid production by systematically tuning the central metabolic pathways based on a CRISPR interference system in *Escherichia coli*. *Sci Rep*. 2015;5:13477. doi:10.1038/srep13477

106. Milke L, Marienhagen J. Engineering intracellular malonyl-CoA availability in microbial hosts and its impact on polyketide and fatty acid synthesis. *Appl Microbiol Biotechnol*. 2020;104(14):6057-6065. doi:10.1007/s00253-020-10643-7

107. Lei Y, Chen W, Xiang L, Wu J, Zhen Z, Jin J-M, Liang C, Tang S-Y. Engineering an SspB-mediated degron for novel controllable protein degradation. *Metab Eng*. 2022;74:150-159. doi:10.1016/j.ymben.2022.10.013

108. Li J, Fu J, Yue C, Shang Y, Ye B-C. Highly efficient biosynthesis of protocatechuic acid via recombinant *Pseudomonas putida* KT2440. *J Agric Food Chem*. Published online June 27, 2023. doi:10.1021/acs.jafc.3c01511

109. Karzai AW, Roche ED, Sauer RT. The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue. *Nat Struct Biol*. 2000;7(6):449-455. doi:10.1038/75843

110. Sauer RT, Bolon DN, Burton BM, Burton RE, Flynn JM, Grant RA, Hersch GL, Joshi SA, Kenniston JA, Levchenko I, Neher SB, Oakes ESC, Siddiqui SM, Wah DA, Baker TA. Sculpting the proteome with AAA(+) proteases and disassembly machines. *Cell*. 2004;119(1):9-18. doi:10.1016/j.cell.2004.09.020

111. Flynn JM, Levchenko I, Seidel M, Wickner SH, Sauer RT, Baker TA. Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. *Proc Natl Acad Sci USA*. 2001;98(19):10584-10589. doi:10.1073/pnas.191375298

112. Silva F, Queiroz JA, Domingues FC. Evaluating metabolic stress and plasmid stability in plasmid DNA production by *Escherichia coli*. *Biotechnol Adv*. 2012;30(3):691-708. doi:10.1016/j.biotechadv.2011.12.005

113. Sandberg TE, Salazar MJ, Weng LL, Palsson BO, Feist AM. The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. *Metab Eng*. 2019;56:1-16. doi:10.1016/j.ymben.2019.08.004

114. Lenski RE, Rose MR, Simpson SC, Tadler SC. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am Nat*. 1991;138(6):1315. doi:10.1086/285289

115. Sandberg TE, Lloyd CJ, Palsson BO, Feist AM. Laboratory evolution to alternating substrate environments yields distinct phenotypic and genetic adaptive strategies. *Appl Environ Microbiol*. 2017;83(13):e00410-17. doi:10.1128/AEM.00410-17

116. Radivojević T, Costello Z, Workman K, Garcia Martin H. A machine learning automated recommendation tool for synthetic biology. *Nat Commun*. 2020;11(1):4879. doi:10.1038/s41467-020-18008-4

117. Tao XB, LaFrance S, Xing Y, Nava AA, Martin HG, Keasling JD, Backman TWH. ClusterCAD 2.0: an updated computational platform for chimeric type I polyketide

synthase and nonribosomal peptide synthetase design. *Nucleic Acids Res*. 2023;51(D1):D532-D538. doi:10.1093/nar/gkac1075

118. Weissman KJ. Genetic engineering of modular PKSs: from combinatorial biosynthesis to synthetic biology. *Nat Prod Rep*. 2016;33(2):203-230. doi:10.1039/c5np00109a

119. Yuzawa S, Eng CH, Katz L, Keasling JD. Broad substrate specificity of the loading didomain of the lipomycin polyketide synthase. *Biochemistry*. 2013;52(22):3791-3793. doi:10.1021/bi400520t

120. Bihlmaier C, Welle E, Hofmann C, Welzel K, Vente A, Breitling E, Müller M, Glaser S, Bechthold A. Biosynthetic gene cluster for the polyenoyltetramic acid alpha-lipomycin. *Antimicrob Agents Chemother*. 2006;50(6):2113-2121. doi:10.1128/AAC.00007-06

121. Schmidt M, Lee N, Zhan C, Roberts JB, Nava AA, Keiser L, Vilchez A, Chen Y, Petzold CJ, Haushalter RW, Blank LM, Keasling JD. Maximizing heterologous expression of engineered type I polyketide synthases: investigating codon optimization strategies. *ACS Synth Biol*. 2023;12(11):3366-3380. doi: 10.1021/acssynbio.3c00367

122. Takahashi S, Toyoda A, Sekiyama Y, Takagi H, Nogawa T, Uramoto M, Suzuki R, Koshino H, Kumano T, Panthee S, Dairi T, Ishikawa J, Ikeda H, Sakaki Y, Osada H. Reveromycin A biosynthesis uses RevG and RevJ for stereospecific spiroacetal formation. *Nat Chem Biol*. 2011;7(7):461-468. doi:10.1038/nchembio.583

123. Yuzawa S, Deng K, Wang G, Baidoo EEK, Northen TR, Adams PD, Katz L, Keasling JD. Comprehensive in vitro analysis of acyltransferase domain exchanges in modular polyketide synthases and its application for short-chain ketone production. *ACS Synth Biol*. 2017;6(1):139-147. doi:10.1021/acssynbio.6b00176

124. Englund E, Schmidt M, Nava AA, Lechner A, Deng K, Jocic R, Lin Y, Roberts J, Benites VT, Kakumanu R, Gin JW, Chen Y, Liu Y, Petzold CJ, Baidoo EEK, Northen TR, Adams PD, Katz L, Yuzawa S, Keasling JD. Expanding extender substrate selection for unnatural polyketide biosynthesis by acyltransferase domain exchange within a modular polyketide synthase. *J Am Chem Soc*. 2023;145(16):8822-8832. doi:10.1021/jacs.2c11027

125. Khosla C, Tang Y, Chen AY, Schnarr NA, Cane DE. Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu Rev Biochem*. 2007;76:195-221. doi:10.1146/annurev.biochem.76.053105.093515

126. Wilson MC, Nam S-J, Gulder TAM, Kauffman CA, Jensen PR, Fenical W, Moore BS. Structure and biosynthesis of the marine streptomycete ansamycin ansalactam A and its distinctive branched chain polyketide extender unit. *J Am Chem Soc*. 2011;133(6):1971-1977. doi:10.1021/ja109226s

127. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res*. 2011;39(Web Server issue):W339-46. doi:10.1093/nar/gkr466

128. Englund E, Schmidt M, Nava AA, Klass S, Keiser L, Dan Q, Katz L, Yuzawa S, Keasling JD. Biosensor guided polyketide synthases engineering for optimization of domain exchange boundaries. *Nat Commun*. 2023;14(1):4871. doi:10.1038/s41467-023-40464-x

129. Miwa T, Chadani Y, Taguchi H. *Escherichia coli* small heat shock protein IbpA is an aggregation-sensor that self-regulates its own expression at posttranscriptional levels. *Mol Microbiol*. 2021;115(1):142-156. doi:10.1111/mmi.14606

130. Klaus M, Grininger M. Engineering strategies for rational polyketide synthase design. *Nat Prod Rep*. 2018;35(10):1070-1081. doi:10.1039/c8np00030a

131. Zheng J, Piasecki SK, Keatinge-Clay AT. Structural studies of an A2-type modular polyketide synthase ketoreductase reveal features controlling α-substituent stereochemistry. *ACS Chem Biol*. 2013;8(9):1964-1971. doi:10.1021/cb400161g

132. Hagen A, Poust S, Rond T de, Fortman JL, Katz L, Petzold CJ, Keasling JD. Engineering a polyketide synthase for in vitro production of adipic acid. *ACS Synth Biol*. 2016;5(1):21-27. doi:10.1021/acssynbio.5b00153

133. Bertani G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol*. 1951;62(3):293-300. doi:10.1128/jb.62.3.293-300.1951

134. LaBauve AE, Wargo MJ. Growth and laboratory maintenance of *Pseudomonas aeruginosa*. *Curr Protoc Microbiol*. 2012;Chapter 6:Unit 6E.1. doi:10.1002/9780471729259.mc06e01s25

135. Hillson NJ, Rosengarten RD, Keasling JD. j5 DNA assembly design automation software. *ACS Synth Biol*. 2012;1(1):14-21. doi:10.1021/sb2000116

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

137. Zobel S, Benedetti I, Eisenbach L, de Lorenzo V, Wierckx N, Blank LM. Tn7-based device for calibrated heterologous gene expression in *Pseudomonas putida*. *ACS Synth Biol*. 2015;4(12):1341-1351. doi:10.1021/acssynbio.5b00058

138. Choi K-H, Kumar A, Schweizer HP. A 10-min method for preparation of highly electrocompetent *Pseudomonas aeruginosa* cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *J Microbiol Methods*. 2006;64(3):391-397. doi:10.1016/j.mimet.2005.06.001

139. Zhan C, Lee N, Lan G, Dan Q, Cowan A, Wang Z, Baidoo EEK, Kakumanu R, Luckie B, Kuo RC, McCauley J, Liu Y, Valencia L, Haushalter RW, Keasling JD. Improved polyketide production in *C. glutamicum* by preventing propionate-induced growth inhibition. *Nat Metab*. Published online July 13, 2023. doi:10.1038/s42255-023-00830-x

140. Hershfield V, Boyer HW, Yanofsky C, Lovett MA, Helinski DR. Plasmid ColEl as a molecular vehicle for cloning and amplification of DNA. *Proc Natl Acad Sci USA*. 1974;71(9):3455-3459. doi:10.1073/pnas.71.9.3455

141. Silo-Suh LA, Elmore B, Ohman DE, Suh S-J. Isolation, characterization, and utilization of a temperature-sensitive allele of a *Pseudomonas* replicon. *J Microbiol Methods*. 2009;78(3):319-324. doi:10.1016/j.mimet.2009.07.002

142. Zulkower V, Rosser S. DNA Chisel, a versatile sequence optimizer. *Bioinformatics*. 2020;36(16):4508-4509. doi:10.1093/bioinformatics/btaa558

143. Bikard D, Loot C, Baharoglu Z, Mazel D. Folded DNA in action: hairpin formation and biological functions in prokaryotes. *Microbiol Mol Biol Rev*. 2010;74(4):570-588. doi:10.1128/MMBR.00026-10

144. Wehrs M, Prahl J-P, Moon J, Li Y, Tanjore D, Keasling JD, Pray T, Mukhopadhyay A. Production efficiency of the bacterial non-ribosomal peptide indigoidine relies on the respiratory metabolic state in *S. cerevisiae*. *Microb Cell Fact*. 2018;17(1):193. doi:10.1186/s12934-018-1045-1

145. Lai B, Plan M, Hodson M, Krömer J. Simultaneous determination of sugars, carboxylates, alcohols and aldehydes from fermentations by high performance liquid chromatography. *Fermentation*. 2016;2(1):6. doi:10.3390/fermentation2010006

146. Chen Y, Gin J, Petzold CJ. Alkaline-SDS cell lysis of microbes with acetone protein precipitation for proteomic sample preparation in 96-well plate format V.1 . protocols.io https://doi.org/10.17504/protocols.io.6qpvr6xjpvmk/v1. Published January 18, 2023. Accessed June 13, 2023. https://www.protocols.io/view/alkaline-sds-cell-lysis-of-microbes-with-acetone-p-6qpvr6xjpvmk/v1

147. Chen Y, Gin J, Petzold CJ. Discovery proteomic (DIA) LC-MS/MS data acquisition and analysis v1. *protocols.io*. Published online March 17, 2022. doi:10.17504/protocols.io.e6nvwk1z7vmk/v1

148. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods*. 2020;17(1):41-44. doi:10.1038/s41592-019-0638-x

149. Ahrné E, Molzahn L, Glatter T, Schmidt A. Critical assessment of proteome-wide label-free absolute abundance estimation strategies. *Proteomics*. 2013;13(17):2567-2578. doi:10.1002/pmic.201300135

150. Silva JC, Gorenstein MV, Li G-Z, Vissers JPC, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol Cell Proteomics*. 2006;5(1):144-156. doi:10.1074/mcp.M500230-MCP200

151. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat İ, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P, SciPy 1.0 Contributors. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020;17(3):261-272. doi:10.1038/s41592-019-0686-2

152. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, van Kerkwijk MH,

Brett M, Haldane A, Del Río JF, Wiebe M, Peterson P, Gérard-Marchant P, Sheppard K, Reddy T, Weckesser W, Abbasi H, Gohlke C, Oliphant TE. Array programming with NumPy. *Nature*. 2020;585(7825):357-362. doi:10.1038/s41586-020-2649-2

153. Gumerov VM, Ortega DR, Adebali O, Ulrich LE, Zhulin IB. MiST 3.0: an updated microbial signal transduction database with an emphasis on chemosensory systems. *Nucleic Acids Res*. 2020;48(D1):D459-D464. doi:10.1093/nar/gkz988

154. Elbourne LDH, Tetu SG, Hassan KA, Paulsen IT. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res*. 2017;45(D1):D320-D324. doi:10.1093/nar/gkw1068

155. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, von Mering C, Bork P. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res*. 2019;47(D1):D309-D314. doi:10.1093/nar/gky1085

156. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol*. 2021;38(12):5825-5829. doi:10.1093/molbev/msab293

157. Karp PD, Billington R, Caspi R, Fulcher CA, Latendresse M, Kothari A, Keseler IM, Krummenacker M, Midford PE, Ong Q, Ong WK, Paley SM, Subhraveti P. The BioCyc collection of microbial genomes and metabolic pathways. *Brief Bioinformatics*. 2019;20(4):1085-1093. doi:10.1093/bib/bbx085

158. Price MN, Arkin AP. PaperBLAST: Text mining papers for information about homologs. *mSystems*. 2017;2(4). doi:10.1128/mSystems.00039-17

159. Tamura K, Stecher G, Kumar S. MEGA11: Molecular evolutionary genetics analysis version 11. *Mol Biol Evol*. 2021;38(7):3022-3027. doi:10.1093/molbev/msab120

160. Eng T, Banerjee D, Lau AK, Bowden E, Herbert RA, Trinh J, Prahl J-P, Deutschbauer A, Tanjore D, Mukhopadhyay A. Engineering *Pseudomonas putida* for efficient aromatic conversion to bioproduct using high throughput screening in a bioreactor. *Metab Eng*. 2021;66:229-238. doi:10.1016/j.ymben.2021.04.015

161. Mehrer CR, Rand JM, Incha MR, Cook TB, Demir B, Motagamwala AH, Kim D, Dumesic JA, Pfleger BF. Growth-coupled bioconversion of levulinic acid to butanone. *Metab Eng*. 2019;55:92-101. doi:10.1016/j.ymben.2019.06.003

162. López-Farfán D, Reyes-Darias JA, Matilla MA, Krell T. Concentration dependent effect of plant root exudates on the chemosensory systems of *Pseudomonas putida* KT2440. *Front Microbiol*. 2019;10:78. doi:10.3389/fmicb.2019.00078

163. Fernández M, Morel B, Corral-Lugo A, Krell T. Identification of a chemoreceptor that specifically mediates chemotaxis toward metabolizable purine derivatives. *Mol Microbiol*. 2016;99(1):34-42. doi:10.1111/mmi.13215

164. Neal AL, Ahmad S, Gordon-Weeks R, Ton J. Benzoxazinoids in root exudates of maize attract *Pseudomonas putida* to the rhizosphere. *PLoS ONE*. 2012;7(4):e35498. doi:10.1371/journal.pone.0035498

165. Neal AL, Ton J. Systemic defense priming by *Pseudomonas putida* KT2440 in maize depends on benzoxazinoid exudation from the roots. *Plant Signal Behav*. 2013;8(1):e22655. doi:10.4161/psb.22655

166. Hervás AB, Canosa I, Santero E. Transcriptome analysis of *Pseudomonas putida* in response to nitrogen availability. *J Bacteriol*. 2008;190(1):416-420. doi:10.1128/JB.01230-07

167. Poblete-Castro I, Escapa IF, Jäger C, Puchalka J, Lam CMC, Schomburg D, Prieto MA, Martins dos Santos VAP. The metabolic response of *P. putida* KT2442 producing high levels of polyhydroxyalkanoate under single- and multiple-nutrient-limited growth: highlights from a multi-level omics approach. *Microb Cell Fact*. 2012;11:34. doi:10.1186/1475-2859-11-34

168. Mozejko-Ciesielska J, Dabrowska D, Szalewska-Palasz A, Ciesielski S. Medium-chain-length polyhydroxyalkanoates synthesis by *Pseudomonas putida* KT2440 relA/spoT mutant: bioprocess characterization and transcriptome analysis. *AMB Express*. 2017;7(1):92. doi:10.1186/s13568-017-0396-z

169. Galman JL, Slabu I, Weise NJ, Iglesias C, Parmeggiani F, Lloyd RC, Turner NJ. Biocatalytic transamination with near-stoichiometric inexpensive amine donors mediated by bifunctional mono- and di-amine transaminases. *Green Chem*. 2017;19(2):361-366. doi:10.1039/C6GC02102F

170. Han T, Kim GB, Lee SY. Glutaric acid production by systems metabolic engineering of an l-lysine-overproducing *Corynebacterium glutamicum. Proc Natl Acad Sci USA*. 2020;117(48):30328-30334. doi:10.1073/pnas.2017483117

171. Banerjee D, Eng T, Lau AK, Sasaki Y, Wang B, Chen Y, Prahl J-P, Singan VR, Herbert RA, Liu Y, Tanjore D, Petzold CJ, Keasling JD, Mukhopadhyay A. Genome-scale metabolic rewiring improves titers rates and yields of the non-native product indigoidine at scale. *Nat Commun*. 2020;11(1):5385. doi:10.1038/s41467-020-19171-4

172. Incha MR, Thompson MG, Blake-Hedges JM, Liu Y, Pearson AN, Schmidt M, Gin JW, Petzold CJ, Deutschbauer AM, Keasling JD. Leveraging host metabolism for bisdemethoxycurcumin production in *Pseudomonas putida*. *Metab Eng Commun*. 2020;10:e00119. doi:10.1016/j.mec.2019.e00119

173. van der Maaten L, Hinton G. Visualizing data using t-SNE. *Journal of Machine Learning Research*. Published online 2008.

174. Brown PR, Tata R. Growth of *Pseudomonas aeruginosa* mutants lacking glutamate synthase activity. *J Bacteriol*. 1981;147(1):193-197. doi:10.1128/jb.147.1.193-197.1981

175. Sonawane AM, Röhm KH. A functional *gltB* gene is essential for utilization of acidic amino acids and expression of periplasmic glutaminase/asparaginase (PGA) by

Pseudomonas putida KT2440. *Mol Genet Genomics*. 2004;271(1):33-39. doi:10.1007/s00438-003-0951-y

176. Eberl L, Ammendola A, Rothballer MH, Givskov M, Sternberg C, Kilstrup M, Schleifer KH, Molin S. Inactivation of *gltB* abolishes expression of the assimilatory nitrate reductase gene (*nasB*) in *Pseudomonas putida* KT2442. *J Bacteriol*. 2000;182(12):3368-3376. doi:10.1128/JB.182.12.3368-3376.2000

177. Reitzer L. Nitrogen assimilation and global regulation in *Escherichia coli*. *Annu Rev Microbiol*. 2003;57:155-176. doi:10.1146/annurev.micro.57.030502.090820

178. Venturi V. Control of *rpoS* transcription in *Escherichia coli* and *Pseudomonas*: why so different? *Mol Microbiol*. 2003;49(1):1-9. doi:10.1046/j.1365-2958.2003.03547.x

179. Lapouge K, Schubert M, Allain FH-T, Haas D. Gac/Rsm signal transduction pathway of gamma-proteobacteria: from RNA recognition to regulation of social behaviour. *Mol Microbiol*. 2008;67(2):241-253. doi:10.1111/j.1365-2958.2007.06042.x

180. Martínez-Gil M, Ramos-González MI, Espinosa-Urgel M. Roles of cyclic Di-GMP and the Gac system in transcriptional control of the genes coding for the *Pseudomonas putida* adhesins LapA and LapF. *J Bacteriol*. 2014;196(8):1484-1495. doi:10.1128/JB.01287-13

181. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IMJ, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DMD, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Chiara E Catenazzi M, Chang S, Neil Cooley R, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 2008;456(7218):53-59. doi:10.1038/nature07517

182. Demling P, Ankenbauer A, Klein B, Noack S, Tiso T, Takors R, Blank LM. *Pseudomonas putida* KT2440 endures temporary oxygen limitations. *Biotechnol Bioeng*. 2021;118(12):4735-4750. doi:10.1002/bit.27938

183. Merrick MJ, Edwards RA. Nitrogen control in bacteria. *Microbiol Rev*. 1995;59(4):604-622.

184. Pino C, Olmo-Mira F, Cabello P, Martínez-Luque M, Castillo F, Roldán MD, Moreno-Vivián C. The assimilatory nitrate reduction system of the phototrophic bacterium *Rhodobacter capsulatus* E1F1. *Biochem Soc Trans*. 2006;34(Pt 1):127-129. doi:10.1042/BST0340127

185. Luque-Almagro VM, Gates AJ, Moreno-Vivián C, Ferguson SJ, Richardson DJ, Roldán MD. Bacterial nitrate assimilation: gene distribution and regulation. *Biochem Soc Trans*. 2011;39(6):1838-1843. doi:10.1042/BST20110688

186. Caballero A, Esteve-Núñez A, Zylstra GJ, Ramos JL. Assimilation of nitrogen from nitrite and trinitrotoluene in *Pseudomonas putida* JLR11. *J Bacteriol*. 2005;187(1):396-399. doi:10.1128/JB.187.1.396-399.2005

187. Romeo A, Sonnleitner E, Sorger-Domenigg T, Nakano M, Eisenhaber B, Bläsi U. Transcriptional regulation of nitrate assimilation in *Pseudomonas aeruginosa* occurs via transcriptional antitermination within the nirBD-PA1779-cobA operon. *Microbiology (Reading, Engl)*. 2012;158(Pt 6):1543-1552. doi:10.1099/mic.0.053850-0

188. Wang B, Rensing C, Pierson LS, Zhao H, Kennedy C. Translational coupling of nasST expression in *Azotobacter vinelandii* prevents overexpression of the nasT gene. *FEMS Microbiol Lett*. 2014;361(2):123-130. doi:10.1111/1574-6968.12621

189. Nicholas DJD, Nason A. Molybdenum and nitrate reductase II. Molybdenum as a constituent of nitrate reductase. *Journal of Biological Chemistry*. Published online 1954.

190. Luque-Almagro VM, Lyall VJ, Ferguson SJ, Roldán MD, Richardson DJ, Gates AJ. Nitrogen oxyanion-dependent dissociation of a two-component complex that regulates bacterial nitrate assimilation. *J Biol Chem*. 2013;288(41):29692-29702. doi:10.1074/jbc.M113.459032

191. Merino E, Jensen RA, Yanofsky C. Evolution of bacterial *trp* operons and their regulation. *Curr Opin Microbiol*. 2008;11(2):78-86. doi:10.1016/j.mib.2008.02.005

192. Levin-Karp A, Barenholz U, Bareia T, Dayagi M, Zelcbuch L, Antonovsky N, Noor E, Milo R. Quantifying translational coupling in *E. coli* synthetic operons using RBS modulation and fluorescent reporters. *ACS Synth Biol*. 2013;2(6):327-336. doi:10.1021/sb400002n

193. Nelson KE, Weinel C, Paulsen IT, Dodson RJ, Hilbert H, Martins dos Santos VA, Fouts DE, Gill SR, Pop M, Holmes M, Brinkac L, Beanan M, DeBoy RT, Daugherty S, Kolonay J, Madupu R, Nelson W, White O, Peterson J, Khouri H, Hance I, Chris Lee P, Holtzapple E, Scanlan D, Tran K, Moazzez A, Utterback T, Rizzo M, Lee K, Kosack D, Moestl D, Wedler H, Lauber J, Stjepandic D, Hoheisel J, Straetz M, Heim S, Kiewitz C, Eisen JA, Timmis KN, Düsterhöft A, Tümmler B, Fraser CM. Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ Microbiol*. 2002;4(12):799-808. doi:10.1046/j.1462-2920.2002.00366.x

194. Regenhardt D, Heuer H, Heim S, Fernandez DU, Strömpl C, Moore ERB, Timmis KN. Pedigree and taxonomic credentials of *Pseudomonas putida* strain KT2440. *Environ Microbiol*. 2002;4(12):912-915. doi:10.1046/j.1462-2920.2002.00368.x

195. Weinel C, Nelson KE, Tümmler B. Global features of the *Pseudomonas putida* KT2440 genome sequence. *Environ Microbiol*. 2002;4(12):809-818. doi:10.1046/j.1462-2920.2002.00331.x

196. Hutchison CA, Merryman C, Sun L, Assad-Garcia N, Richter RA, Smith HO, Glass JI. Polar effects of transposon insertion into a minimal bacterial genome. *J Bacteriol*. 2019;201(19). doi:10.1128/JB.00185-19

197. Maeda S, Okamura M, Kobayashi M, Omata T. Nitrite-specific active transport system of the cyanobacterium *Synechococcus* sp. strain PCC 7942. *J Bacteriol*. 1998;180(24):6761-6763. doi:10.1128/JB.180.24.6761-6763.1998

198. Gao S-H, Fan L, Yuan Z, Bond PL. The concentration-determined and population-specific antimicrobial effects of free nitrous acid on *Pseudomonas aeruginosa* PAO1. *Appl Microbiol Biotechnol*. 2015;99(5):2305-2312. doi:10.1007/s00253-014-6211-8

199. Kusumawardhani H, van Dijk D, Hosseini R, de Winde JH. Novel toxin-antitoxin module SlvT-SlvA regulates megaplasmid stability and incites solvent tolerance in *Pseudomonas putida* S12. *Appl Environ Microbiol*. 2020;86(13). doi:10.1128/AEM.00686-20

200. Skjerning RB, Senissar M, Winther KS, Gerdes K, Brodersen DE. The RES domain toxins of RES-Xre toxin-antitoxin modules induce cell stasis by degrading NAD+. *Mol Microbiol*. 2019;111(1):221-236. doi:10.1111/mmi.14150

201. Fuhrer T, Fischer E, Sauer U. Experimental identification and quantification of glucose metabolism in seven bacterial species. *J Bacteriol*. 2005;187(5):1581-1590. doi:10.1128/JB.187.5.1581-1590.2005

202. Nikel PI, Fuhrer T, Chavarría M, Sánchez-Pascuala A, Sauer U, de Lorenzo V. Reconfiguration of metabolic fluxes in *Pseudomonas putida* as a response to sub-lethal oxidative stress. *ISME J*. 2021;15(6):1751-1766. doi:10.1038/s41396-020-00884-9

203. del Castillo T, Ramos JL, Rodríguez-Herva JJ, Fuhrer T, Sauer U, Duque E. Convergent peripheral pathways catalyze initial glucose catabolism in *Pseudomonas putida*: genomic and flux analysis. *J Bacteriol*. 2007;189(14):5142-5152. doi:10.1128/JB.00203-07

204. Roon RJ, Levenberg B. An adenosine triphosphate-dependent, avidin-sensitive enzymatic cleavage of urea in yeast and green algae. *J Biol Chem*. 1968;243(19):5213-5215.

205. Roon RJ, Levenberg B. Urea amidolyase. I. Properties of the enzyme from *Candida utilis*. *J Biol Chem*. 1972;247(13):4107-4113.

206. Mackay EM, Pateman JA. The regulation of urease activity in *Aspergillus nidulans*. *Biochem Genet*. 1982;20(7-8):763-776. doi:10.1007/BF00483972

207. Kanamori T, Kanou N, Atomi H, Imanaka T. Enzymatic characterization of a prokaryotic urea carboxylase. *J Bacteriol*. 2004;186(9):2532-2539. doi:10.1128/JB.186.9.2532-2539.2004

208. Scott E, Peter F, Sanders J. Biomass in the manufacture of industrial products--the use of proteins and amino acids. *Appl Microbiol Biotechnol*. 2007;75(4):751-762. doi:10.1007/s00253-007-0932-x

209. Hermann T. Industrial production of amino acids by coryneform bacteria. *J Biotechnol*. 2003;104(1-3):155-172. doi:10.1016/S0168-1656(03)00149-4

210. Nijkamp K, Westerhof RGM, Ballerstedt H, de Bont JAM, Wery J. Optimization of the solvent-tolerant *Pseudomonas putida* S12 as host for the production of p-coumarate from glucose. *Appl Microbiol Biotechnol*. 2007;74(3):617-624. doi:10.1007/s00253-006-0703-0

211. Craig JW, Chang F-Y, Kim JH, Obiajulu SC, Brady SF. Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse proteobacteria. *Appl Environ Microbiol*. 2010;76(5):1633-1641. doi:10.1128/AEM.02169-09

212. Verhoef S, Ballerstedt H, Volkers RJM, de Winde JH, Ruijssenaars HJ. Comparative transcriptomics and proteomics of p-hydroxybenzoate producing *Pseudomonas putida* S12: novel responses and implications for strain improvement. *Appl Microbiol Biotechnol*. 2010;87(2):679-690. doi:10.1007/s00253-010-2626-z

213. Xing X, Jiang P. Recombinant bacteria for producing deoxyviolacein and uses thereof. Published online July 28, 2011.

214. Schmitz S, Nies S, Wierckx N, Blank LM, Rosenbaum MA. Engineering mediator-based electroactivity in the obligate aerobic bacterium *Pseudomonas putida* KT2440. *Front Microbiol*. 2015;6:284. doi:10.3389/fmicb.2015.00284

215. Molina-Santiago C, Cordero BF, Daddaoua A, Udaondo Z, Manzano J, Valdivia M, Segura A, Ramos J-L, Duque E. *Pseudomonas putida* as a platform for the synthesis of aromatic compounds. *Microbiology (Reading, Engl)*. 2016;162(9):1535-1543. doi:10.1099/mic.0.000333

216. Aboulmagd E, Voss I, Oppermann-Sanio FB, Steinbüchel A. Heterologous expression of cyanophycin synthetase and cyanophycin synthesis in the industrial relevant bacteria *Corynebacterium glutamicum* and *Ralstonia eutropha* and in *Pseudomonas putida*. *Biomacromolecules*. 2001;2(4):1338-1342. doi:10.1021/bm010075a

217. Voss I, Diniz SC, Aboulmagd E, Steinbüchel A. Identification of the *Anabaena* sp. strain PCC7120 cyanophycin synthetase as suitable enzyme for production of cyanophycin in gram-negative bacteria like *Pseudomonas putida* and *Ralstonia eutropha*. *Biomacromolecules*. 2004;5(4):1588-1595. doi:10.1021/bm049861g

218. Wiefel L, Bröker A, Steinbüchel A. Synthesis of a citrulline-rich cyanophycin by use of *Pseudomonas putida* ATCC 4359. *Appl Microbiol Biotechnol*. 2011;90(5):1755-1762. doi:10.1007/s00253-011-3224-4

219. Loeschcke A, Thies S. *Pseudomonas putida*-a versatile host for the production of natural products. *Appl Microbiol Biotechnol*. 2015;99(15):6197-6214. doi:10.1007/s00253-015-6745-4

220. Radkov AD, Moe LA. Amino acid racemization in *Pseudomonas putida* KT2440. *J Bacteriol*. 2013;195(22):5016-5024. doi:10.1128/JB.00761-13

221. Mercenier A, Simon JP, Haas D, Stalon V. Catabolism of L-arginine by *Pseudomonas aeruginosa*. *J Gen Microbiol*. 1980;116(2):381-389. doi:10.1099/00221287-116-2-381

222. Stalon V, Vander Wauven C, Momin P, Legrain C. Catabolism of arginine, citrulline and ornithine by *Pseudomonas* and related bacteria. *J Gen Microbiol*. 1987;133(9):2487-2495. doi:10.1099/00221287-133-9-2487

223. Li C, Lu C-D. Arginine racemization by coupled catabolic and anabolic dehydrogenases. *Proc Natl Acad Sci USA*. 2009;106(3):906-911. doi:10.1073/pnas.0808269106

224. Patil MD, Rathod VP, Bihade UR, Banerjee UC. Purification and characterization of arginine deiminase from *Pseudomonas putida*: Structural insights of the differential affinities of l-arginine analogues. *J Biosci Bioeng*. 2019;127(2):129-137. doi:10.1016/j.jbiosc.2018.07.021

225. Vander Wauven C, Jann A, Haas D, Leisinger T, Stalon V. N2-succinylornithine in ornithine catabolism of *Pseudomonas aeruginosa*. *Arch Microbiol*. 1988;150(4):400-404. doi:10.1007/BF00408314

226. Tricot C, Vander Wauven C, Wattiez R, Falmagne P, Stalon V. Purification and properties of a succinyltransferase from *Pseudomonas aeruginosa* specific for both arginine and ornithine. *Eur J Biochem*. 1994;224(3):853-861. doi:10.1111/j.1432-1033.1994.00853.x

227. Radkov AD, Moe LA. A Broad spectrum racemase in *Pseudomonas putida* KT2440 plays a key role in amino acid catabolism. *Front Microbiol*. 2018;9:1343. doi:10.3389/fmicb.2018.01343

228. Chang YF, Adams E. Induction of separate catabolic pathways for L- and D-lysine in *Pseudomonas putida*. *Biochem Biophys Res Commun*. 1971;45(3):570-577. doi:10.1016/0006-291x(71)90455-4

229. Espinosa-Urgel M, Ramos JL. Expression of a *Pseudomonas putida* aminotransferase involved in lysine catabolism is induced in the rhizosphere. *Appl Environ Microbiol*. 2001;67(11):5219-5224. doi:10.1128/AEM.67.11.5219-5224.2001

230. Revelles O, Espinosa-Urgel M, Molin S, Ramos JL. The davDT operon of *Pseudomonas putida*, involved in lysine catabolism, is induced in response to the pathway intermediate delta-aminovaleric acid. *J Bacteriol*. 2004;186(11):3439-3446. doi:10.1128/JB.186.11.3439-3446.2004

231. Revelles O, Espinosa-Urgel M, Fuhrer T, Sauer U, Ramos JL. Multiple and interconnected pathways for L-lysine catabolism in *Pseudomonas putida* KT2440. *J Bacteriol*. 2005;187(21):7500-7510. doi:10.1128/JB.187.21.7500-7510.2005

232. Revelles O, Wittich R-M, Ramos JL. Identification of the initial steps in D-lysine catabolism in *Pseudomonas putida*. *J Bacteriol*. 2007;189(7):2787-2792. doi:10.1128/JB.01538-06

233. Lambert MP, Neuhaus FC. Factors affecting the level of alanine racemase in *Escherichia coli*. *J Bacteriol*. 1972;109(3):1156-1161. doi:10.1128/jb.109.3.1156-1161.1972

234. Radkov AD, Moe LA. Bacterial synthesis of D-amino acids. *Appl Microbiol Biotechnol*. 2014;98(12):5363-5374. doi:10.1007/s00253-014-5726-3

235. Duque E, Daddaoua A, Cordero BF, De la Torre J, Antonia Molina-Henares M, Ramos J-L. Identification and elucidation of in vivo function of two alanine racemases from *Pseudomonas putida* KT2440. *Environ Microbiol Rep*. 2017;9(5):581-588. doi:10.1111/1758-2229.12576

236. Boulette ML, Baynham PJ, Jorth PA, Kukavica-Ibrulj I, Longoria A, Barrera K, Levesque RC, Whiteley M. Characterization of alanine catabolism in *Pseudomonas aeruginosa* and its importance for proliferation in vivo. *J Bacteriol*. 2009;191(20):6329-6334. doi:10.1128/JB.00817-09

237. Zhi J, Mathew E, Freundlich M. Lrp binds to two regions in the *dadAX* promoter region of *Escherichia coli* to repress and activate transcription directly. *Mol Microbiol*. 1999;32(1):29-40. doi:10.1046/j.1365-2958.1999.01314.x

238. Schwanemann T, Otto M, Wierckx N, Wynands B. *Pseudomonas* as versatile aromatics cell factory. *Biotechnol J*. 2020;15(11):e1900569. doi:10.1002/biot.201900569

239. Whitaker RJ, Gaines CG, Jensen RA. A multispecific quintet of aromatic aminotransferases that overlap different biochemical pathways in *Pseudomonas aeruginosa*. *J Biol Chem*. 1982;257(22):13550-13556. doi:10.1016/S0021-9258(18)33482-3

240. Molina-Henares MA, García-Salamanca A, Molina-Henares AJ, de la Torre J, Herrera MC, Ramos JL, Duque E. Functional analysis of aromatic biosynthetic pathways in *Pseudomonas putida* KT2440. *Microb Biotechnol*. 2009;2(1):91-100. doi:10.1111/j.1751-7915.2008.00062.x

241. Bortolotti P, Hennart B, Thieffry C, Jausions G, Faure E, Grandjean T, Thepaut M, Dessein R, Allorge D, Guery BP, Faure K, Kipnis E, Toussaint B, Le Gouellec A. Tryptophan catabolism in *Pseudomonas aeruginosa* and potential for inter-kingdom relationship. *BMC Microbiol*. 2016;16(1):137. doi:10.1186/s12866-016-0756-x

242. Pittard J, Yang J. Biosynthesis of the aromatic amino acids. *Ecosal Plus*. 2008;3(1). doi:10.1128/ecosalplus.3.6.1.8

243. Nogales J, Palsson BØ, Thiele I. A genome-scale metabolic reconstruction of *Pseudomonas putida* KT2440: iJN746 as a cell factory. *BMC Syst Biol*. 2008;2:79. doi:10.1186/1752-0509-2-79

244. Leprince A, de Lorenzo V, Völler P, van Passel MWJ, Martins dos Santos VAP. Random and cyclical deletion of large DNA segments in the genome of *Pseudomonas putida*. *Environ Microbiol*. 2012;14(6):1444-1453. doi:10.1111/j.1462-2920.2012.02730.x

245. Gadilohar BL, Shankarling GS. Choline based ionic liquids and their applications in organic transformation. *J Mol Liq*. 2017;227:234-261. doi:10.1016/j.molliq.2016.11.136

246. Wood JM. Bacterial osmoregulation: a paradigm for the study of cellular homeostasis. *Annu Rev Microbiol*. 2011;65:215-238. doi:10.1146/annurev-micro-090110-102815

247. Wargo MJ. Homeostasis and catabolism of choline and glycine betaine: lessons from *Pseudomonas aeruginosa*. *Appl Environ Microbiol*. 2013;79(7):2112-2120. doi:10.1128/AEM.03565-12

248. Geiger O, López-Lara IM, Sohlenkamp C. Phosphatidylcholine biosynthesis and function in bacteria. *Biochim Biophys Acta*. 2013;1831(3):503-513. doi:10.1016/j.bbalip.2012.08.009

249. Belda E, van Heck RGA, José Lopez-Sanchez M, Cruveiller S, Barbe V, Fraser C, Klenk H-P, Petersen J, Morgat A, Nikel PI, Vallenet D, Rouy Z, Sekowska A, Martins Dos Santos VAP, de Lorenzo V, Danchin A, Médigue C. The revisited genome of *Pseudomonas putida* KT2440 enlightens its value as a robust metabolic chassis. *Environ Microbiol*. 2016;18(10):3403-3424. doi:10.1111/1462-2920.13230

250. Leys D, Basran J, Scrutton NS. Channelling and formation of "active" formaldehyde in dimethylglycine oxidase. *EMBO J*. 2003;22(16):4038-4048. doi:10.1093/emboj/cdg395

251. Wargo MJ, Hogan DA. Identification of genes required for *Pseudomonas aeruginosa* carnitine catabolism. *Microbiology (Reading, Engl)*. 2009;155(Pt 7):2411-2419. doi:10.1099/mic.0.028787-0

252. Bastard K, Smith AAT, Vergne-Vaxelaire C, Perret A, Zaparucha A, De Melo-Minardi R, Mariage A, Boutard M, Debard A, Lechaplais C, Pelle C, Pellouin V, Perchat N, Petit J-L, Kreimeyer A, Medigue C, Weissenbach J, Artiguenave F, De Berardinis V, Vallenet D, Salanoubat M. Revealing the hidden functional diversity of an enzyme family. *Nat Chem Biol*. 2014;10(1):42-49. doi:10.1038/nchembio.1387

253. Meadows JA, Wargo MJ. Transcriptional regulation of carnitine catabolism in *Pseudomonas aeruginosa* by CdhR. *mSphere*. 2018;3(1). doi:10.1128/mSphere.00480-17

254. Lee S, Takahashi Y, Oura H, Suzuki-Minakuchi C, Okada K, Yamane H, Nomura N, Nojiri H. Effects of carbazole-degradative plasmid pCAR1 on biofilm morphology in *Pseudomonas putida* KT2440. *Environ Microbiol Rep*. 2016;8(2):261-271. doi:10.1111/1758-2229.12376

255. Nuccio ML, Russell BL, Nolte KD, Rathinasabapathi B, Gage DA, Hanson AD. The endogenous choline supply limits glycine betaine synthesis in transgenic tobacco expressing choline monooxygenase. *Plant J*. 1998;16(4):487-496. doi:10.1046/j.1365-313x.1998.00316.x

256. Rontein D, Nishida I, Tashiro G, Yoshioka K, Wu WI, Voelker DR, Basset G, Hanson AD. Plants synthesize ethanolamine by direct decarboxylation of serine using a pyridoxal phosphate enzyme. *J Biol Chem*. 2001;276(38):35523-35529. doi:10.1074/jbc.M106038200

257. Lundgren BR, Sarwar Z, Pinto A, Ganley JG, Nomura CT. Ethanolamine catabolism in *Pseudomonas aeruginosa* PAO1 is regulated by the enhancer-binding protein EatR (PA4021) and the alternative sigma factor RpoN. *J Bacteriol*. 2016;198(17):2318-2329. doi:10.1128/JB.00357-16

258. Roof DM, Roth JR. Ethanolamine utilization in *Salmonella typhimurium*. *J Bacteriol*. 1988;170(9):3855-3863. doi:10.1128/jb.170.9.3855-3863.1988

259. Chang GW, Chang JT. Evidence for the B12-dependent enzyme ethanolamine deaminase in *Salmonella*. *Nature*. 1975;254(5496):150-151. doi:10.1038/254150a0

260. Scarlett FA, Turner JM. Microbial metabolism of amino alcohols. Ethanolamine catabolism mediated by coenzyme B12-dependent ethanolamine ammonia-lyase in *Escherichia coli* and *Klebsiella aerogenes*. *J Gen Microbiol*. 1976;95(1):173-176. doi:10.1099/00221287-95-1-173

261. Cameron B, Briggs K, Pridmore S, Brefort G, Crouzet J. Cloning and analysis of genes involved in coenzyme B12 biosynthesis in *Pseudomonas denitrificans*. *J Bacteriol*. 1989;171(1):547-557. doi:10.1128/jb.171.1.547-557.1989

262. Heller K, Mann BJ, Kadner RJ. Cloning and expression of the gene for the vitamin B12 receptor protein in the outer membrane of *Escherichia coli*. *J Bacteriol*. 1985;161(3):896-903. doi:10.1128/jb.161.3.896-903.1985

263. Cadieux N, Bradbeer C, Reeger-Schneider E, Köster W, Mohanty AK, Wiener MC, Kadner RJ. Identification of the periplasmic cobalamin-binding protein BtuF of *Escherichia coli*. *J Bacteriol*. 2002;184(3):706-717. doi:10.1128/JB.184.3.706-717.2002

264. Zayas CL, Escalante-Semerena JC. Reassessment of the late steps of coenzyme B12 synthesis in *Salmonella enterica*: evidence that dephosphorylation of adenosylcobalamin-5'-phosphate by the CobC phosphatase is the last step of the pathway. *J Bacteriol*. 2007;189(6):2210-2218. doi:10.1128/JB.01665-06

265. Warren MJ, Raux E, Schubert HL, Escalante-Semerena JC. The biosynthesis of adenosylcobalamin (vitamin B12). *Nat Prod Rep*. 2002;19(4):390-412. doi:10.1039/b108967f

266. Martens JH, Barg H, Warren MJ, Jahn D. Microbial production of vitamin B12. *Appl Microbiol Biotechnol*. 2002;58(3):275-285. doi:10.1007/s00253-001-0902-7

267. Mori K, Bando R, Hieda N, Toraya T. Identification of a reactivating factor for adenosylcobalamin-dependent ethanolamine ammonia lyase. *J Bacteriol*. 2004;186(20):6845-6854. doi:10.1128/JB.186.20.6845-6854.2004

268. Vogels GD, Van der Drift C. Degradation of purines and pyrimidines by microorganisms. *Bacteriol Rev*. 1976;40(2):403-468.

269. Tipton PA. Urate to allantoin, specifically (S)-allantoin. *Nat Chem Biol*. 2006;2(3):124-125. doi:10.1038/nchembio0306-124

270. Ramazzina I, Folli C, Secchi A, Berni R, Percudani R. Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. *Nat Chem Biol*. 2006;2(3):144-148. doi:10.1038/nchembio768

271. Hassan KA, Liu Q, Henderson PJF, Paulsen IT. Homologs of the *Acinetobacter baumannii* AceI transporter represent a new family of bacterial multidrug efflux systems. *MBio*. 2015;6(1). doi:10.1128/mBio.01982-14

272. Juttukonda LJ, Green ER, Lonergan ZR, Heffern MC, Chang CJ, Skaar EP. *Acinetobacter baumannii* OxyR regulates the transcriptional response to hydrogen peroxide. *Infect Immun*. 2019;87(1). doi:10.1128/IAI.00413-18

273. Zimmer DP, Soupene E, Lee HL, Wendisch VF, Khodursky AB, Peter BJ, Bender RA, Kustu S. Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: scavenging as a defense against nitrogen limitation. *Proc Natl Acad Sci USA*. 2000;97(26):14674-14679. doi:10.1073/pnas.97.26.14674

274. Ware E. The chemistry of the hydantoins. *Chem Rev*. 1950;46(3):403-470. doi:10.1021/cr60145a001

275. Shorvon S. *Status Epilepticus*. Cambridge University Press; 1994. doi:10.1017/CBO9780511526930

276. Drauz K, Grayson I, Kleemann A, Krimmer H-P, Leuchtenberger W, Weckbecker C. Amino Acids. In: *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH Verlag GmbH & Co. KGaA; 2007. doi:10.1002/14356007.a02_057.pub2

277. Ogawa J, Kaimura T, Yamada H, Shimizu S. Evaluation of pyrimidine- and hydantoin-degrading enzyme activities in aerobic bacteria. *FEMS Microbiol Lett*. 1994;122(1-2):55-60. doi:10.1111/j.1574-6968.1994.tb07143.x

278. Solem E. The absolute configuration of beta-aminoisobutyric acid formed by degradation of thymine in man. *Clin Chim Acta*. 1974;53(2):183-190. doi:10.1016/0009-8981(74)90097-7

279. Massey LK, Sokatch JR, Conrad RS. Branched-chain amino acid catabolism in bacteria. *Bacteriol Rev*. 1976;40(1):42-54.

280. Zhang J, Barajas JF, Burdu M, Wang G, Baidoo EE, Keasling JD. Application of an Acyl-CoA ligase from *Streptomyces aizunensis* for lactam biosynthesis. *ACS Synth Biol*. 2017;6(5):884-890. doi:10.1021/acssynbio.6b00372

281. Niehaus TD, Elbadawi-Sidhu M, de Crécy-Lagard V, Fiehn O, Hanson AD. Discovery of a widespread prokaryotic 5-oxoprolinase that was hiding in plain sight. *J Biol Chem*. 2017;292(39):16360-16367. doi:10.1074/jbc.M117.805028

282. Miller DL, Rodwell VW. Metabolism of basic amino acids in *Pseudomonas putida*. Intermediates in L-arginine catabolism. *J Biol Chem*. 1971;246(16):5053-5058.

283. Bandounas L, Ballerstedt H, de Winde JH, Ruijssenaars HJ. Redundancy in putrescine catabolism in solvent tolerant *Pseudomonas putida* S12. *J Biotechnol*. 2011;154(1):1-10. doi:10.1016/j.jbiotec.2011.04.005

284. Chae TU, Ko Y-S, Hwang K-S, Lee SY. Metabolic engineering of *Escherichia coli* for the production of four-, five- and six-carbon lactams. *Metab Eng*. 2017;41:82-91. doi:10.1016/j.ymben.2017.04.001

285. Niu W, Willett H, Mueller J, He X, Kramer L, Ma B, Guo J. Direct biosynthesis of adipic acid from lignin-derived aromatics using engineered *Pseudomonas putida* KT2440. *Metab Eng*. 2020;59:151-161. doi:10.1016/j.ymben.2020.02.006.

286. Tabor CW, Tabor H. Polyamines in microorganisms. *Microbiol Rev*. 1985;49(1):81-99.

287. Oota M, Tsai AY-L, Aoki D, Matsushita Y, Toyoda S, Fukushima K, Saeki K, Toda K, Perfus-Barbeoch L, Favery B, Ishikawa H, Sawa S. Identification of naturally occurring polyamines as root-knot nematode attractants. *Mol Plant*. 2020;13(4):658-665. doi:10.1016/j.molp.2019.12.010

288. Johnson L, Mulcahy H, Kanevets U, Shi Y, Lewenza S. Surface-localized spermidine protects the *Pseudomonas aeruginosa* outer membrane from antibiotic treatment and oxidative stress. *J Bacteriol*. 2012;194(4):813-826. doi:10.1128/JB.05230-11

289. Corral-Lugo A, De la Torre J, Matilla MA, Fernández M, Morel B, Espinosa-Urgel M, Krell T. Assessment of the contribution of chemoreceptor-based signalling to biofilm formation. *Environ Microbiol*. 2016;18(10):3355-3372. doi:10.1111/1462-2920.13170

290. Gaymans RJ, Van Utteren TEC, Van Den Berg JWA, Schuyer J. Preparation and some properties of nylon 46. *J Polym Sci Polym Chem Ed*. 1977;15(3):537-545. doi:10.1002/pol.1977.170150303

291. Xu F, Zhong L, Xu Y, Zhang C, Zhang F, Zhang G. Highly efficient flame-retardant and soft cotton fabric prepared by a novel reactive flame retardant. *Cellulose*. 2019;26(6):4225-4240. doi:10.1007/s10570-019-02374-4

292. Roberts MR. Does GABA act as a signal in plants?: Hints from molecular studies. *Plant Signal Behav*. 2007;2(5):408-409. doi:10.4161/psb.2.5.4335

293. Bouché N, Fromm H. GABA in plants: just a metabolite? *Trends Plant Sci*. 2004;9(3):110-115. doi:10.1016/j.tplants.2004.01.006

294. Trouillon J, Ragno M, Simon V, Attrée I, Elsen S. Transcription inhibitors with XRE DNA-binding and cupin signal-sensing domains drive metabolic diversification in *Pseudomonas*. *mSystems*. 2021;6(1). doi:10.1128/mSystems.00753-20

295. Bulger RJ, Roosen-Runge U. Bactericidal activity of the ampicillin/kanamycin combination against *Escherichia coli*, *Enterobacter klebsiella* and *Proteus*. *Am J Med Sci*. 1969;258(1):7-13. doi:10.1097/00000441-196907000-00002

296. Reyes-Darias JA, García V, Rico-Jiménez M, Corral-Lugo A, Lesouhaitier O, Juárez-Hernández D, Yang Y, Bi S, Feuilloley M, Muñoz-Rojas J, Sourjik V, Krell T. Specific gamma-aminobutyrate chemotaxis in pseudomonads with different lifestyle. *Mol Microbiol*. 2015;97(3):488-501. doi:10.1111/mmi.13045

297. Kinnersley AM, Turano FJ. Gamma aminobutyric acid (GABA) and plant responses to stress. *CRC Crit Rev Plant Sci*. 2000;19(6):479-509. doi:10.1080/07352680091139277

298. Bartsch K, von Johnn-Marteville A, Schulz A. Molecular analysis of two genes of the *Escherichia coli* gab cluster: nucleotide sequence of the glutamate:succinic semialdehyde transaminase gene (*gabT*) and characterization of the succinic semialdehyde dehydrogenase gene (*gabD*). *J Bacteriol*. 1990;172(12):7035-7042. doi:10.1128/jb.172.12.7035-7042.1990

299. Weber N, Hatsch A, Labagnere L, Heider H. Production of (S)-2-aminobutyric acid and (S)-2-aminobutanol in *Saccharomyces cerevisiae*. *Microb Cell Fact*. 2017;16(1):51. doi:10.1186/s12934-017-0667-z

300. Zhang K, Li H, Cho KM, Liao JC. Expanding metabolism for total biosynthesis of the nonnatural amino acid L-homoalanine. *Proc Natl Acad Sci USA*. 2010;107(14):6234-6239. doi:10.1073/pnas.0912903107

301. Soga T, Baran R, Suematsu M, Ueno Y, Ikeda S, Sakurakawa T, Kakazu Y, Ishikawa T, Robert M, Nishioka T, Tomita M. Differential metabolomics reveals ophthalmic acid as an oxidative stress biomarker indicating hepatic glutathione consumption. *J Biol Chem*. 2006;281(24):16768-16776. doi:10.1074/jbc.M601876200

302. Cohen YR. β-aminobutyric acid-induced resistance against plant pathogens. *Plant Dis*. 2002;86(5):448-457. doi:10.1094/PDIS.2002.86.5.448

303. Donadio S, Staver MJ, McAlpine JB, Swanson SJ, Katz L. Modular organization of genes required for complex polyketide biosynthesis. *Science*. 1991;252(5006):675-679. doi:10.1126/science.2024119

304. Yuzawa S, Mirsiaghi M, Jocic R, Fujii T, Masson F, Benites VT, Baidoo EEK, Sundstrom E, Tanjore D, Pray TR, George A, Davis RW, Gladden JM, Simmons BA, Katz L, Keasling JD. Short-chain ketone production by engineered polyketide synthases in *Streptomyces albus*. *Nat Commun*. 2018;9(1):4569. doi:10.1038/s41467-018-07040-0

305. Miyazawa T, Fitzgerald BJ, Keatinge-Clay AT. Preparative production of an enantiomeric pair by engineered polyketide synthases. *Chem Commun*. 2021;57(70):8762-8765. doi:10.1039/d1cc03073f

306. Pfeifer BA, Khosla C. Biosynthesis of polyketides in heterologous hosts. *Microbiol Mol Biol Rev*. 2001;65(1):106-118. doi:10.1128/MMBR.65.1.106-118.2001

307. Pfeifer BA, Admiraal SJ, Gramajo H, Cane DE, Khosla C. Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of *E. coli*. *Science (New York, NY)*. 2001;291(5509):1790-1792.

308. Ongley SE, Bian X, Neilan BA, Müller R. Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Nat Prod Rep*. 2013;30(8):1121-1138. doi:10.1039/c3np70034h

309. Nakamura Y, Gojobori T, Ikemura T. Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res*. 2000;28(1):292. doi:10.1093/nar/28.1.292

310. Leskiw BK, Bibb MJ, Chater KF. The use of a rare codon specifically during development? *Mol Microbiol*. 1991;5(12):2861-2867. doi:10.1111/j.1365-2958.1991.tb01845.x

311. Belinky F, Rogozin IB, Koonin EV. Selection on start codons in prokaryotes and potential compensatory nucleotide substitutions. *Sci Rep*. 2017;7(1):12422. doi:10.1038/s41598-017-12619-6

312. Sharp PM, Li WH. The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res*. 1987;15(3):1281-1295. doi:10.1093/nar/15.3.1281

313. Hale RS, Thompson G. Codon optimization of the gene encoding a domain from human type 1 neurofibromin protein results in a threefold improvement in expression level in *Escherichia coli*. *Protein Expr Purif*. 1998;12(2):185-188. doi:10.1006/prep.1997.0825

314. Claassens NJ, Siliakus MF, Spaans SK, Creutzburg SCA, Nijsse B, Schaap PJ, Quax TEF, van der Oost J. Improving heterologous membrane protein production in *Escherichia coli* by combining transcriptional tuning and codon usage algorithms. *PLoS ONE*. 2017;12(9):e0184355. doi:10.1371/journal.pone.0184355

315. Eggeling L, Bott M. A giant market and a powerful metabolism: L-lysine provided by *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol*. 2015;99(8):3387-3394. doi:10.1007/s00253-015-6508-2

316. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol*. 2014;5:172. doi:10.3389/fmicb.2014.00172

317. Yuzawa S, Bailey CB, Fujii T, Jocic R, Barajas JF, Benites VT, Baidoo EEK, Chen Y, Petzold CJ, Katz L, Keasling JD. Heterologous gene expression of N-terminally truncated variants of LipPks1 suggests a functionally critical structural motif in the N-terminus of modular molyketide synthase. *ACS Chem Biol*. 2017;12(11):2725-2729. doi:10.1021/acschembio.7b00714

318. Athey J, Alexaki A, Osipova E, Rostovtsev A, Santana-Quintero LV, Katneni U, Simonyan V, Kimchi-Sarfaty C. A new and updated resource for codon usage tables. *BMC Bioinformatics*. 2017;18(1):391. doi:10.1186/s12859-017-1793-7

319. Holcomb DD, Alexaki A, Katneni U, Kimchi-Sarfaty C. The Kazusa codon usage database, CoCoPUTs, and the value of up-to-date codon usage statistics. *Infect Genet Evol*. 2019;73:266-268. doi:10.1016/j.meegid.2019.05.010

320. Elmore JR, Furches A, Wolff GN, Gorday K, Guss AM. Development of a high efficiency integration system and promoter library for rapid modification of *Pseudomonas putida* KT2440. *Metab Eng Commun*. 2017;5:1-8. doi:10.1016/j.meteno.2017.04.001

321. Lethanh H, Neubauer P, Hoffmann F. The small heat-shock proteins IbpA and IbpB reduce the stress load of recombinant *Escherichia coli* and delay degradation of inclusion bodies. *Microb Cell Fact*. 2005;4(1):6. doi:10.1186/1475-2859-4-6

322. Sørensen MA, Kurland CG, Pedersen S. Codon usage determines translation rate in *Escherichia coli*. *J Mol Biol*. 1989;207(2):365-377. doi:10.1016/0022-2836(89)90260-x

323. Zhou Z, Dang Y, Zhou M, Li L, Yu C-H, Fu J, Chen S, Liu Y. Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proc Natl Acad Sci USA*. 2016;113(41):E6117-E6125. doi:10.1073/pnas.1606724113

324. Murli S, Kennedy J, Dayem LC, Carney JR, Kealey JT. Metabolic engineering of *Escherichia coli* for improved 6-deoxyerythronolide B production. *J Ind Microbiol Biotechnol*. 2003;30(8):500-509. doi:10.1007/s10295-003-0073-x

325. Gross F, Gottschalk D, Müller R. Posttranslational modification of myxobacterial carrier protein domains in *Pseudomonas* sp. by an intrinsic phosphopantetheinyl transferase. *Appl Microbiol Biotechnol*. 2005;68(1):66-74. doi:10.1007/s00253-004-1836-7

326. Marshall VD, Sokatch JR. Regulation of valine catabolism in *Pseudomonas putida*. *J Bacteriol*. 1972;110(3):1073-1081. doi:10.1128/jb.110.3.1073-1081.1972

327. Gross F, Ring MW, Perlova O, Fu J, Schneider S, Gerth K, Kuhlmann S, Stewart AF, Zhang Y, Müller R. Metabolic engineering of *Pseudomonas putida* for methylmalonyl-CoA biosynthesis to enable complex heterologous secondary metabolite formation. *Chem Biol*. 2006;13(12):1253-1264. doi:10.1016/j.chembiol.2006.09.014

328. Botella L, Lindley ND, Eggeling L. Formation and metabolism of methylmalonyl coenzyme A in *Corynebacterium glutamicum*. *J Bacteriol*. 2009;191(8):2899-2901. doi:10.1128/JB.01756-08

329. Zhang K, Woodruff AP, Xiong M, Zhou J, Dhande YK. A synthetic metabolic pathway for production of the platform chemical isobutyric acid. *ChemSusChem*. 2011;4(8):1068-1070. doi:10.1002/cssc.201100045

330. Xu H, Zhang F, Liu B, Huhman DV, Sumner LW, Dixon RA, Wang G. Characterization of the formation of branched short-chain fatty acid:CoAs for bitter acid biosynthesis in hop glandular trichomes. *Mol Plant*. 2013;6(4):1301-1317. doi:10.1093/mp/sst004

331. Lang K, Zierow J, Buehler K, Schmid A. Metabolic engineering of *Pseudomonas* sp. strain VLB120 as platform biocatalyst for the production of isobutyric acid and other secondary metabolites. *Microb Cell Fact*. 2014;13:2. doi:10.1186/1475-2859-13-2

332. Palluk S, Arlow DH, de Rond T, Barthel S, Kang JS, Bector R, Baghdassarian HM, Truong AN, Kim PW, Singh AK, Hillson NJ, Keasling JD. De novo DNA synthesis using polymerase-nucleotide conjugates. *Nat Biotechnol*. 2018;36(7):645-650. doi:10.1038/nbt.4173

333. Muir P, Li S, Lou S, Wang D, Spakowicz DJ, Salichos L, Zhang J, Weinstock GM, Isaacs F, Rozowsky J, Gerstein M. The real cost of sequencing: scaling computation to keep pace with data generation. *Genome Biol*. 2016;17:53. doi:10.1186/s13059-016-0917-0

334. Troeschel SC, Thies S, Link O, Real CI, Knops K, Wilhelm S, Rosenau F, Jaeger K-E. Novel broad host range shuttle vectors for expression in *Escherichia coli*, *Bacillus subtilis* and *Pseudomonas putida*. *J Biotechnol*. 2012;161(2):71-79. doi:10.1016/j.jbiotec.2012.02.020

335. Lin-Chao S, Bremer H. Effect of the bacterial growth rate on replication control of plasmid pBR322 in *Escherichia coli*. *Mol Gen Genet*. 1986;203(1):143-149. doi:10.1007/BF00330395

336. San Millan A, Toll-Riera M, Qi Q, Betts A, Hopkinson RJ, McCullagh J, MacLean RC. Integrative analysis of fitness and metabolic effects of plasmids in *Pseudomonas aeruginosa* PAO1. *ISME J*. 2018;12(12):3014-3024. doi:10.1038/s41396-018-0224-8

337. Pena-Gonzalez A, Rodriguez-R LM, Marston CK, Gee JE, Gulvik CA, Kolton CB, Saile E, Frace M, Hoffmaster AR, Konstantinidis KT. Genomic characterization and copy number variation of *Bacillus anthracis* plasmids pXO1 and pXO2 in a historical collection of 412 strains. *mSystems*. 2018;3(4). doi:10.1128/mSystems.00065-18

338. Mi J, Sydow A, Schempp F, Becher D, Schewe H, Schrader J, Buchhaupt M. Investigation of plasmid-induced growth defect in *Pseudomonas putida*. *J Biotechnol*. 2016;231:167-173. doi:10.1016/j.jbiotec.2016.06.001

339. Anthony Mason C, Bailey JamesE. Effects of plasmid presence on growth and enzyme activity of *Escherichia coli* DH5? *Appl Microbiol Biotechnol*. 1989;32(1). doi:10.1007/BF00164823

340. Kang M-K, Lee J, Um Y, Lee TS, Bott M, Park SJ, Woo HM. Synthetic biology platform of CoryneBrick vectors for gene expression in *Corynebacterium glutamicum* and its application to xylose utilization. *Appl Microbiol Biotechnol*. 2014;98(13):5991-6002. doi:10.1007/s00253-014-5714-7

341. Li M, Chen J, Wang Y, Liu J, Huang J, Chen N, Zheng P, Sun J. Efficient multiplex gene repression by CRISPR-dCpf1 in *Corynebacterium glutamicum*. *Front Bioeng Biotechnol*. 2020;8:357. doi:10.3389/fbioe.2020.00357

342. Zhao F, Zhou Z, Dang Y, Na H, Adam C, Lipzen A, Ng V, Grigoriev IV, Liu Y. Genome-wide role of codon usage on transcription and identification of potential regulators. *Proc Natl Acad Sci USA*. 2021;118(6). doi:10.1073/pnas.2022590118

343. Webster MW, Weixlbaumer A. The intricate relationship between transcription and translation. *Proc Natl Acad Sci USA*. 2021;118(21). doi:10.1073/pnas.2106284118

344. Reddy P, Peterkofsky A, McKenney K. Translational efficiency of the *Escherichia coli* adenylate cyclase gene: mutating the UUG initiation codon to GUG or AUG results in increased gene expression. *Proc Natl Acad Sci USA*. 1985;82(17):5656-5660. doi:10.1073/pnas.82.17.5656

345. Villegas A, Kropinski AM. An analysis of initiation codon utilization in the domain bacteria - concerns about the quality of bacterial genome annotation. *Microbiology (Reading, Engl)*. 2008;154(Pt 9):2559-2661. doi:10.1099/mic.0.2008/021360-0

346. Constant DA, Gutierrez JM, Sastry AV, Viazzo R, Smith NR, Hossain J, Spencer DA, Carter H, Ventura AB, Louie MTM, Kohnert C, Consbruck R, Bennett J, Crawford KA, Sutton JM, Morrison A, Steiger AK, Jackson KA, Stanton JT, Abdulhaqq S, Hannum G, Meier J, Weinstock M, Gander M. Deep learning-based codon optimization with large-scale synonymous variant datasets enables generalized tunable protein expression. *BioRxiv*. Published online February 12, 2023. doi:10.1101/2023.02.11.528149

347. Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W. Lignin biosynthesis and structure. *Plant Physiol*. 2010;153(3):895-905. doi:10.1104/pp.110.155119

348. Gullón P, Romaní A, Vila C, Garrote G, Parajó JC. Potential of hydrothermal treatments in lignocellulose biorefineries. *Biofuels, Bioprod Bioref*. 2012;6(2):219-232. doi:10.1002/bbb.339

349. Toor SS, Rosendahl L, Rudolf A. Hydrothermal liquefaction of biomass: A review of subcritical water technologies. *Energy*. 2011;36(5):2328-2342. doi:10.1016/j.energy.2011.03.013

350. Kumar P, Maharjan A, Jun H-B, Kim BS. Bioconversion of lignin and its derivatives into polyhydroxyalkanoates: Challenges and opportunities. *Biotechnol Appl Biochem*. 2019;66(2):153-162. doi:10.1002/bab.1720

351. Chai Y, Shan S, Weissman KJ, Hu S, Zhang Y, Müller R. Heterologous expression and genetic engineering of the tubulysin biosynthetic gene cluster using Red/ET recombineering and inactivation mutagenesis. *Chem Biol*. 2012;19(3):361-371. doi:10.1016/j.chembiol.2012.01.007

352. Dudnik A, Bigler L, Dudler R. Heterologous expression of a *Photorhabdus luminescens* syrbactin-like gene cluster results in production of the potent proteasome inhibitor glidobactin A. *Microbiol Res*. 2013;168(2):73-76. doi:10.1016/j.micres.2012.09.006

353. Gross F, Luniak N, Perlova O, Gaitatzis N, Jenke-Kodama H, Gerth K, Gottschalk D, Dittmann E, Müller R. Bacterial type III polyketide synthases: phylogenetic analysis and potential for the production of novel secondary metabolites by heterologous expression in pseudomonads. *Arch Microbiol*. 2006;185(1):28-38. doi:10.1007/s00203-005-0059-3

354. Wenzel SC, Gross F, Zhang Y, Fu J, Stewart AF, Müller R. Heterologous expression of a myxobacterial natural products assembly line in pseudomonads via red/ET recombineering. *Chem Biol*. 2005;12(3):349-356. doi:10.1016/j.chembiol.2004.12.012

355. Owen JG, Copp JN, Ackerley DF. Rapid and flexible biochemical assays for evaluating 4'-phosphopantetheinyl transferase activity. *Biochem J*. 2011;436(3):709-717. doi:10.1042/BJ20110321

356. Takahashi H, Kumagai T, Kitani K, Mori M, Matoba Y, Sugiyama M. Cloning and characterization of a *Streptomyces* single module type non-ribosomal peptide synthetase catalyzing a blue pigment synthesis. *J Biol Chem*. 2007;282(12):9073-9081. doi:10.1074/jbc.M611319200

357. Ma C, Shi Y, Mu Q, Li R, Xue Y, Yu B. Unravelling the thioesterases responsible for propionate formation in engineered *Pseudomonas putida* KT2440. *Microb Biotechnol*. 2021;14(3):1237-1242. doi:10.1111/1751-7915.13804

358. Zhang J, Bista R, Miyazawa T, Keatinge-Clay AT. Boosting titers of engineered triketide and tetraketide synthases to record levels through T7 promoter tuning. *Metab Eng*. 2023;78:93-98. doi:10.1016/j.ymben.2023.05.008

359. Waldron C, Matsushima P, Rosteck PR, Broughton MC, Turner J, Madduri K, Crawford KP, Merlo DJ, Baltz RH. Cloning and analysis of the spinosad biosynthetic gene cluster of *Saccharopolyspora spinosa*. *Chem Biol*. 2001;8(5):487-499. doi:10.1016/s1074-5521(01)00029-1

360. Rajesh Banu J, Preethi, Kavitha S, Tyagi VK, Gunasekaran M, Karthikeyan OP, Kumar G. Lignocellulosic biomass based biorefinery: A successful platform towards circular bioeconomy. *Fuel*. 2021;302:121086. doi:10.1016/j.fuel.2021.121086

361. Salvachúa D, Rydzak T, Auwae R, De Capite A, Black BA, Bouvier JT, Cleveland NS, Elmore JR, Huenemann JD, Katahira R, Michener WE, Peterson DJ, Rohrer H, Vardon DR, Beckham GT, Guss AM. Metabolic engineering of *Pseudomonas putida* for increased polyhydroxyalkanoate production from lignin. *Microb Biotechnol*. 2020;13(1):290-298. doi:10.1111/1751-7915.13481

362. Bator I, Wittgens A, Rosenau F, Tiso T, Blank LM. Comparison of three xylose pathways in *Pseudomonas putida* KT2440 for the synthesis of valuable products. *Front Bioeng Biotechnol*. 2019;7:480. doi:10.3389/fbioe.2019.00480

363. Tiso T, Narancic T, Wei R, Pollet E, Beagan N, Schröder K, Honak A, Jiang M, Kenny ST, Wierckx N, Perrin R, Avérous L, Zimmermann W, O'Connor K, Blank LM. Towards bio-upcycling of polyethylene terephthalate. *Metab Eng*. 2021;66:167-178. doi:10.1016/j.ymben.2021.03.011

364. Zargar A, Lal R, Valencia L, Wang J, Backman TWH, Cruz-Morales P, Kothari A, Werts M, Wong AR, Bailey CB, Loubat A, Liu Y, Chen Y, Chang S, Benites VT, Hernández AC, Barajas JF, Thompson MG, Barcelos C, Anayah R, Martin HG, Mukhopadhyay A, Petzold CJ, Baidoo EEK, Katz L, Keasling JD. Chemoinformatic-guided engineering of polyketide synthases. *J Am Chem Soc*. 2020;142(22):9896-9901. doi:10.1021/jacs.0c02549

365. Jeong H, Kim HJ, Lee SJ. Complete genome sequence of *Escherichia coli* strain BL21. *Genome Announc*. 2015;3(2). doi:10.1128/genomeA.00134-15

366. Cook TB, Rand JM, Nurani W, Courtney DK, Liu SA, Pfleger BF. Genetic tools for reliable gene expression and recombineering in *Pseudomonas putida*. *J Ind Microbiol Biotechnol*. 2018;45(7):517-527. doi:10.1007/s10295-017-2001-5

367. Eng CH, Yuzawa S, Wang G, Baidoo EEK, Katz L, Keasling JD. Alteration of polyketide stereochemistry from anti to syn by a ketoreductase domain exchange in a type I modular polyketide synthase subunit. *Biochemistry*. 2016;55(12):1677-1680. doi:10.1021/acs.biochem.6b00129

368. Keatinge-Clay A, Ray K, Lutgens J, Bista R, Zhang J, Desai R, Hirsch M, Miyazawa T, Cordova A. Assessing and harnessing updated polyketide synthase modules through combinatorial engineering. *Res Sq*. Published online July 28, 2023. doi:10.21203/rs.3.rs-3157617/v1

369. Shanks RMQ, Caiazza NC, Hinsa SM, Toutain CM, O'Toole GA. *Saccharomyces cerevisiae*-based molecular tool kit for manipulation of genes from gram-negative bacteria. *Appl Environ Microbiol*. 2006;72(7):5027-5036. doi:10.1128/AEM.00682-06

370. Rosenberg AH, Lade BN, Chui DS, Lin SW, Dunn JJ, Studier FW. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene*. 1987;56(1):125-135. doi:10.1016/0378-1119(87)90165-X

371. Bi C, Su P, Müller J, Yeh Y-C, Chhabra SR, Beller HR, Singer SW, Hillson NJ. Development of a broad-host synthetic biology toolbox for *Ralstonia eutropha* and its application to engineering hydrocarbon biofuel production. *Microb Cell Fact*. 2013;12:107. doi:10.1186/1475-2859-12-107

372. Lee TS, Krupa RA, Zhang F, Hajimorad M, Holtz WJ, Prasad N, Lee SK, Keasling JD. BglBrick vectors and datasheets: A synthetic biology platform for gene expression. *J Biol Eng*. 2011;5:12. doi:10.1186/1754-1611-5-12

373. Pearson AN, Thompson MG, Kirkpatrick LD, Ho C, Vuu KM, Waldburger LM, Keasling JD, Shih PM. The pGinger family of expression plasmids. *Microbiol Spectr*. 2023;11(3):e0037323. doi:10.1128/spectrum.00373-23

374. Schäfer A, Tauch A, Jäger W, Kalinowski J, Thierbach G, Pühler A. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene*. 1994;145(1):69-73. doi:10.1016/0378-1119(94)90324-7

375. Choi K-H, Gaynor JB, White KG, Lopez C, Bosio CM, Karkhoff-Schweizer RR, Schweizer HP. A Tn7-based broad-range bacterial cloning and expression system. *Nat Methods*. 2005;2(6):443-448. doi:10.1038/nmeth765

376. Figurski DH, Helinski DR. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc Natl Acad Sci USA*. 1979;76(4):1648-1652. doi:10.1073/pnas.76.4.1648

377. Boyer HW, Roulland-Dussoix D. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J Mol Biol*. 1969;41(3):459-472. doi:10.1016/0022-2836(69)90288-5

378. Bagdasarian M, Lurz R, Rückert B, Franklin FC, Bagdasarian MM, Frey J, Timmis KN. Specific-purpose plasmid cloning vectors. II. Broad host range, high copy

number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene*. 1981;16(1-3):237-247. doi:10.1016/0378-1119(81)90080-9

6 Supplemental Information

Supplemental Table 1: Plasmids used in this study.

Plasmid	Description	Reference
pMQ30g	Suicide vector for allelic replacement with Gm^r, SacB	[369]
pMQ30g ΔPP_0596		This work.[11]
pMQ30g ΔPP_2180		This work.[11]
pMQ30g ΔPP_5182		This work.[11]
pMQ30g ΔPP_4575-7		This work.[11]
pMQ30g ΔPP_2920-2		This work.[11]
pMQ30g ΔPP_3514-5		[56]
pMQ30k	Suicide vector for allelic replacement with Kan^r, SacB	[369]
pMQ30k ΔPP_2216		This work.
pMQ30k ΔPP_3492		This work.
pMQ30k ΔPP_4064		This work.
pMQ30k ΔPP_4975		This work.
pMQ30k ΔtesB		This work.
pMQ80	pRO1600/ColE1 vector with Gm^r, P_{BAD}/*araC*, and URA3 marker	[369]
pMQ80 BpsA	Derivative of pMQ80 with gene for BpsA and P_{BAD}/*araC* replaced by P_{gapdh}	This work.
pMQ80 SY172	Derivative of pMQ80 with LipPKS gene by Yuzawa *et al.*[123] and P_{BAD}/*araC* replaced by P_{gapdh}	This work.
pMQ80 SY176	Derivative of pMQ80 with BorM1-AT swapped LipPKS gene by Yuzawa *et al.*[123] and P_{BAD}/*araC* replaced by P_{gapdh}	This work.
pET28a	Protein expression vector with ColE1 origin, N-terminal 6xHis-tag and T7lac promoter	[370]
pET28a PP_0596		This work.[11]
pET28a PP_2180		This work.[11]
pET28a PP_5182		This work.[11]
pBADT	Broad host-range vector with BBR1 origin, arabinose inducible promoter, and kanamycin resistance marker	[371]
pBADT PP_0596	Complementation vector	This work.[11]
pBADT PP_2180	Complementation vector	This work.[11]
pBADT PP_5182	Complementation vector	This work.[11]
pBADT PP_4575-7	Complementation vector	This work.[11]
pBADT PP_2920-2	Complementation vector	This work.[11]

pBADT PP_3514-5	Complementation vector	[56]
pBBR1k	pBADT derived broad host-range expression vector with *araC* and P_{BAD} removed	[58]
pBBR1k PP_2920p-RFP	Reporter plasmid with 200 bp upstream of corresponding gene driving RFP expression	This work.[11]
pBBR1k PP_4578p-RFP	Reporter plasmid with 200 bp upstream of corresponding gene driving RFP expression	This work.[11]
pBBR1k PP_3515p-RFP	Reporter plasmid with 200 bp upstream of corresponding gene driving RFP expression	This work.[11]
pBBR1k PP_0596p-RFP	Reporter plasmid with 200 bp upstream of corresponding gene driving RFP expression	This work.[11]
pBBR1k PP_2177p-RFP	Reporter plasmid with 200 bp upstream of corresponding gene driving RFP expression	This work.[11]
pBbS8a	pSC101 origin with carbenicillin resistance and arabinose inducible promoter	[372]
pBbS8a PP_0595		This work.[11]
pBbS8a PP_2181		This work.[11]
pJH204	BxB1 integration vector containing BxB1 *attP* site. Kanamycin selection marker and ColE1 origin flanked by ΦC31 *attB* and *attP* site.	[92]
pBH026	Derivative of BxB1 integration vector pJH204 with promoter *LlacO1* repressed by LacI. Kanamycin selection marker, ColE1 origin and *lacI* are flanked by ΦC31 *attB* and *attP* site.	This work.[121]
pBH026 RFP	BxB1 integration vector with *rfp* gene	This work.[121]
pBH026 LipPKS-Cg_ubc	BxB1 integration vector with ubc codon optimized LipM1+TE gene for expression in *C. glutamicum*	This work.[121]
pBH026 LipPKS-Cg_mcu	BxB1 integration vector with mcu codon optimized LipM1+TE gene for expression in *C. glutamicum*	This work.[121]
pBH026 LipPKS-Cg_hrca	BxB1 integration vector with hrca codon optimized LipM1+TE gene for expression in *C. glutamicum*	This work.[121]

pBH026 LipPKS-Pp_ubc	BxB1 integration vector with ubc codon optimized LipM1+TE gene for expression in *P. putida*	This work.[121]
pBH026 LipPKS-Pp_mcu	BxB1 integration vector with mcu codon optimized LipM1+TE gene for expression in *P. putida*	This work.[121]
pBH026 LipPKS-Pp_hrca	BxB1 integration vector with hrca codon optimized LipM1+TE gene for expression in *P. putida*	This work.[121]
pBH026 LipPKS-Ec_ubc	BxB1 integration vector with ubc codon optimized LipM1+TE gene for expression in *E. coli*	This work.[121]
pBH026 LipPKS-Ec_mcu	BxB1 integration vector with mcu codon optimized LipM1+TE gene for expression in *E. coli*	This work.[121]
pBH026 LipPKS-Ec_hrca	BxB1 integration vector with hrca codon optimized LipM1+TE gene for expression in *E. coli*	This work.[121]
pAN001	Derivative of pBH026 with additional *URA3* marker for yeast assembly.	This work.[121]
pAN001 LipPKS-WT_ATG	BxB1 integration vector with wild-type LipM1+EryM6-TE nucleotide sequence and ATG start codon	This work.[121]
pAN001 LipPKS-WT_GTG	BxB1 integration vector with wild-type LipM1+EryM6-TE nucleotide sequence and GTG start codon	This work.[121]
pJH209	MR11 integration vector containing MR11 *attP* site. Kanamycin selection marker and ColE1 origin are flanked by ΦC31 *attB* and *attP* site.	[92]
pJH209 Sc_mmCoA	MR11 integration vector with mmCoA pathway from *S. cellulosum* So56 under the control of P_{tac}	This work.[121]
pGW30	ColE1 vector with apramycin marker expressing ΦC31 integrase under the control of P_{tac}	[92]
pALC412	Derivative of pGW30 with optimized expression of ΦC31 integrase for use in *C. glutamicum*	Adam Guss, personal communication.

pGW31	ColE1 vector with apramycin marker expressing BxB1 integrase under the control of P_{tac}	[92]
pGW36	ColE1 vector with apramycin marker expressing MR11 integrase under the control of P_{tac}	[92]
pGingerBG-NahR	BBR1 vector with gentamicin marker containing the P_{sal}/NahR inducible system	[373]
pGingerBG-NahR LipPKS-Pp_mcu	Derivative of pGingerBG-NahR carrying LipPKS-Pp_mcu	This work.[121]
pGingerBG-NahR LipPKS-Ec_mcu	Derivative of pGingerBG-NahR carrying LipPKS-Ec_mcu	This work.[121]
pK18	Suicide vector for allelic replacement with Kan^r, SacB	[374]
pK18 ΔprpDBC2	Suicide vector for in-frame deletion of *prpDBC2*	[139]
pK18 ΔCgl0605::kivd-CCL4	Suicide vector for replacement of Cgl0605 with *kivd* from *Lactococcus lactis* and *CCL4* from *Humulus lupulus*	This work.[121].
pK18 ΔCgl1016::sfp	Suicide vector for replacement of Cgl1016 with *sfp* from *Bacillus subtilis*	[139]
pBH027	Derivative of pBH026 with RV *attP* site	Robert Haushalter, personal communication.
pBH027 PCCase-Pp_hrca	Derivative of pBH027 for the integration of the hrca codon optimized genes *accA2* and *pccB* from *S. coelicolor*	This work.
pBH026 LipPKS(BorM1-AT)-Pp_mcu	Derivative of pBH026 LipPKS-Pp_mcu with BorM1 AT-exchange	This work.
pBH026 LipPKS(AnsM8-AT)-Pp_mcu	Derivative of pBH026 LipPKS-Pp_mcu with AnsM8 AT-exchange	This work.
pBH026 D1	Derivative of pBH026 LipPKS(AnsM8-AT)-Pp_mcu with TE domain replaced by SpnA-COMM	This work.
pBH026 D2	Derivative of pBH026 LipPKS-Pp_mcu with AT-KR-ACP replaced by SpnA AT-KR-ACP	This work.
pBH026 D3	Derivative of pBH026 containing wild-type SpnA from *S. spinosa*	This work.
pJH216	R4 integration vector containing R4 *attP* site, Gm^r, and ColA origin.	[92]

pJH216t SpnB-TE	Derivative of pJH216 containing SpnB from *S. spinosa* under the control of P_{tac}	This work.
pTnS-1	oriR6K vector with Amp^r and t*nSABC+D* operon	[375]
pRK2013	RK2-derived plasmid with ColE1, Kan^r and *tra+mob+*	[376]
pBG14e-FRT	Tn7 integration vector with oriR6K, Kan^r, FRT sites, *Tn7L* and *Tn7R* extremes and synthetic promoter 14e driving GFP	Sebastian Köbbing, personal communication.
pBG14e-Sfp-Pp_mcu	Derivative of pBG14e carrying Pp_mcu codon optimized *sfp* from *Bacillus subtilits*	This work.

Supplemental Table 2: Strains used in this study. Strains carrying a plasmid or express genes via SAGE or Tn7 integrations are not listed.

Strain	Description	Reference
E. coli XL1 Blue		Agilent
E. coli BL21(DE3)		NEB
E. coli K207-3		[324]
E. coli HB101		[377]
E. coli PIR2		Thermo Fisher Scientific
E. coli DH5α-λpir		[137]
P. putida KT2440	Wild-type	[378]
ΔPP_0596		This work.[11]
ΔPP_2180		This work.[11]
ΔPP_5182		This work.[11]
ΔPP_4575-7		This work.[11]
ΔPP_2920-2		This work.[11]
ΔPP_3514-5		[56]
AG5577	Derivative of *P. putida* KT2440 with three triple *attB* sites replacing PP_4740 and PP_2876 and integrated between PP_4217/PP_4218, respectively.	Adam Guss, personal communication.
ΔPP_2216		This work.
ΔPP_3492		This work.
ΔPP_4064		This work.
ΔPP_2216 ΔPP_3492		This work.
ΔPP_2216 ΔPP_4064		This work.
ΔPP_3492 ΔPP_4064		This work.
ΔPP_2216 ΔPP_3492 ΔPP_4064		This work.

ΔPP_2216 ΔPP_3492 ΔPP_4064 ΔtesB		This work.
ΔPP_2216 ΔPP_3492 ΔPP_4064 ΔPP_4975		This work.
ΔPP_2216 ΔPP_3492 ΔPP_4064 ΔtesB ΔPP_4975		This work.
AG5577mm	MR11 integrated pJH209 Sc_mmCoA for the production of mmCoA.	This work.[121]
AG6212	Derivative of *C. glutamicum* ATCC13032 with poly *attB* site replacing Cgl1777-8	Adam Guss, personal communication.
AG6212cz	Derivative of AG6212 with the following modifications: ΔprpDBC2ΔCgl0605::kivd-CCL4ΔCgl1016::sfp	This work.[121]

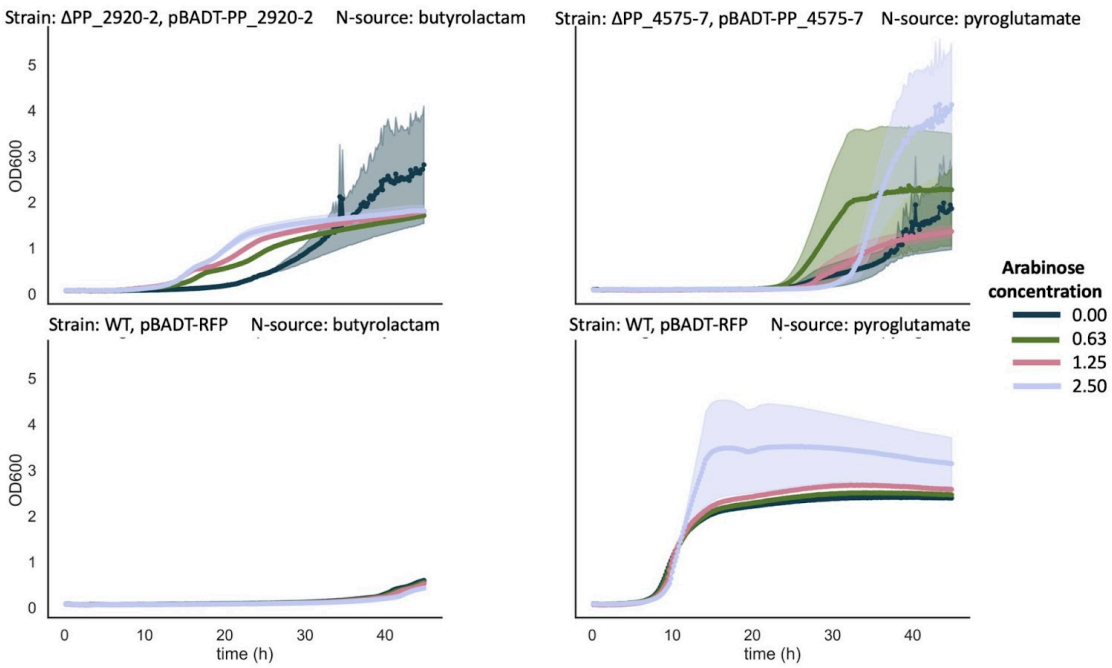
Supplemental Table 3: Primers used in this study.

Primer	Sequence	Description
For primers used in Chapter 3.1, the reader is referred to Schmidt *et al.* (2022).		
For primers used in Chapter 3.2, the reader is referred to Schmidt *et al.* (2023).		
Primers used in Chapter 3.3		
cPCR-PP_2216_F	CTTCTGGCGCACGTATTCT	Confirmation of PP_2216 deletion
cPCR-PP_2216_R	GGATGTCCAACAGGATGGTT	Confirmation of PP_2216 deletion
cPCR-PP_3492_F	CCTGATCATCGACAAGGACAAG	Confirmation of PP_3492 deletion
cPCR-PP_3492_R	TGGATGATGCGCGGTATTG	Confirmation of PP_3492 deletion
cPCR-PP_4064_F	ACGAGTGGATTGACAGTGAC	Confirmation of PP_4064 deletion
cPCR-PP_4064_R	CAGCAGGCGGTCGATAC	Confirmation of PP_4064 deletion
cPCR-tesB_F	GCCTGCGTTGTGCTTTC	Confirmation of tesB deletion
cPCR-tesB_R	GCCCAGCATCACCTTCA	Confirmation of tesB deletion
cPCR-PP_4975_F	CATCAGCGCAGGCATTTC	Confirmation of PP_4975 deletion
cPCR-PP_4975_R	CCTGATCATCGGCAACCT	Confirmation of PP_4975 deletion
cPCR-PCCase_F	GCGAAGAATTGGGCTACGA	Anneals to 3’ end of target gene.

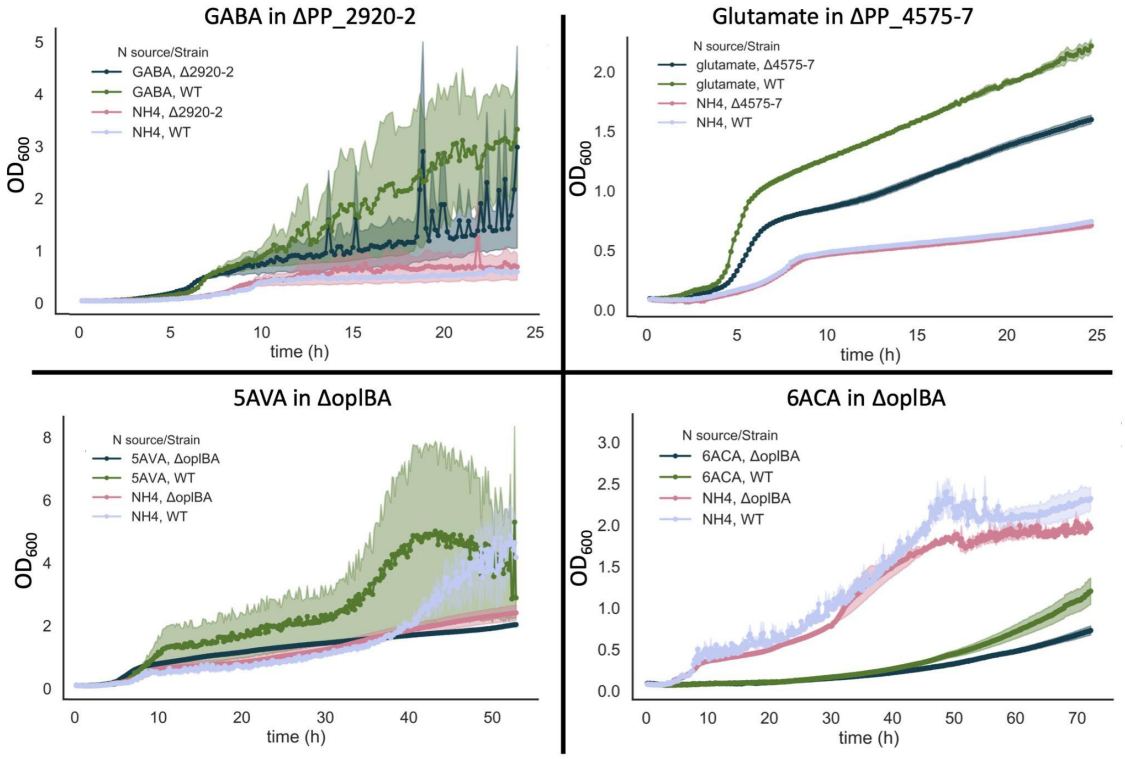
		Confirmation of integration into the host genome
cPCR-PCCase_R	ACGCCTGAGGTAGAGTTGT	Anneals to 5’ end of target gene. Confirmation of integration into the host genome
cPCR-SpnB_R	CACCAACGACGACGAACA	Anneals to 5’ end of target gene. Confirmation of integration into the host genome
cPCR-SpnA-Pp_mcu_F	CGTGATCGAGACTGGTGATG	Anneals to 3’ end of target gene. Confirmation of integration into the host genome
cPCR-SpnA-Pp_mcu_R	ATGGCCGACGTTGGTTT	Anneals to 5’ end of target gene. Confirmation of integration into the host genome
gPCR-pMQ30k-F	TCTAAGAAAC-CATTATTATCATGACATTAACC	Amplification of plasmid backbone
gPCR-pMQ30k-R	AAAACTG-TATTATAAGTAAATGCATGTATACTAAAC	Amplification of plasmid backbone
gPCR-PP_2216-up_F	GTTTAGTATA-CATGCATTTACTTATAATACAGTTTTCACCGGCAC-GCCAGGCCG	Amplification of upstream homology
gPCR-PP_2216-up_R	TGGTTCACCTCAG-TACAGGCATTCAACAGC	Amplification of upstream homology
gPCR-PP_2216-do_F	TGTTGAATGCCTG-TACTGAGGTGAAC-CATGTGAAGGAGCAGACTG-CATGGCATTCG	Amplification of downstream homology
gPCR-PP_2216-do_R	GGTTAATGTCATGATAA-TAATGGTTTCTTA-GAGCTCCGGGGTTCCCTCGC	Amplification of downstream homology
gPCR-PP_3492-up_F	GTTTAGTATA-CATGCATTTACTTATAATACAGTTTT-GCCGGTCTCAATGCCCTGACCC	Amplification of upstream homology
gPCR-PP_3492-up_R	CATCTGTG-TACTCCCGGGCTCAC	Amplification of upstream homology
gPCR-PP_3492-do_F	CCTGTGAGCCCGGGAG-TACACAGATGTAAC-GCAAAGGCAAC-GCGGTTTTTGG	Amplification of downstream homology

gPCR-PP_3492-do_R	GGTTAATGTCATGATAA-TAATGGTTTCTTAGAG-CAGCTTCGCTGCAGCTGC	Amplification of downstream homology
gPCR-PP_4064-up_F	GTTTAGTATA-CATGCATTTACTTATAATAC AGTTTTGGATGTTGTAG-TGACTGAGCGTGG	Amplification of upstream homology
gPCR-PP_4064-up_R	CATGCTGGGGCACCTTCTT-GTTCT	Amplification of upstream homology
gPCR-PP_4064-do_F	CAAGAACAA-GAAGGTGCCCCAG-CATGTGAACATAAGGGAC-GGGCGCAC	Amplification of downstream homology
gPCR-PP_4064-do_R	GGTTAATGTCATGATAA-TAATGGTTTCTTAGA-CAGGGTGGTGCCGAACAG	Amplification of downstream homology
gPCR-tesB-up_F	GTTTAGTATA-CATGCATTTACTTATAATAC AGTTTTAAACTG-CAAGGCGGTAAACCGGAGC	Amplification of upstream homology
gPCR-tesB-up_R	CATCGGCGGTTCTCCGTAG-GAGC	Amplification of upstream homology
gPCR-tesB-do_F	CTCGCTCCTACGGAGAAC-CGCCGATGTGAGCCTGGGC GAGGTGC	Amplification of downstream homology
gPCR-tesB-do_R	GGTTAATGTCATGATAA-TAATGGTTTCTTA-GATCTCCGTATCAGAG-CAAAGAATGGGG	Amplification of downstream homology
gPCR-PP_4975-up_F	GTTTAGTATA-CATGCATTTACTTATAATAC AGTTTTGCCCTTGCCCAC-GTCACCG	Amplification of upstream homology
gPCR-PP_4975-up_R	CATCTGTACTCCTGAC-CTGCTTGGC	Amplification of upstream homology
gPCR-PP_4975-do_F	AGCCAAGCAGGTCAGGAG-TACAGATGTAGACCCTGGTT GCCTGTGCCG	Amplification of downstream homology
gPCR-PP_4975-do_R	GGTTAATGTCATGATAA-TAATGGTTTCTTAGAACCTT-GGGCGTGCTGATCG	Amplification of downstream homology
gPCR-pBH026 LipPKS-Pp_mcu_F	TCCAACCCGGTAAGACAC-GACTTATCGCCACTGG-CAGCA	Amplification of plasmid backbone part 1
gPCR-pBH026 LipPKS-Pp_mcu_R	CCAGGATCACGTGGGCGTT-GGTGC	Amplification of plasmid backbone part1
gPCR-pBH026 LipPKS-Pp_mcu_F	GACAGCACCAC-CGCTAACACCG	Amplification of plasmid backbone part 2

gPCR-pBH026 LipPKS-Pp_mcu_R	GGCGA-TAAGTCGTGTCTTAC-CGGGTTGGACTCAAGACG	Amplification of plas-mid backbone part 2
gPCR-AnsM8-AT-Pp_mcu_F	CGGCACCAACGCCCAC-GTGATCCTG-GAACAGGCGCCCGAGCCG	Amplification of AnsM8-AT
gPCR-AnsM8-AT-Pp_mcu_R	TCTCCGGTGTTAGCGGTGGT GCTGTCCAACCAG-TAGCGCTGGCGC	Amplification of AnsM8-AT
gPCR-BorM1-AT-Pp_mcu_F	CGGCACCAACGCCCAC-GTGATCCTGGAAGAAC-CGCCAGCGGAAGACG	Amplification of BorM1-AT
gPCR-BorM1-AT-Pp_mcu_R	TCTCCGGTGTTAGCGGTGGT GCTGTCCAGCCAGTAG-TGCTGATGCTGG	Amplification of BorM1-AT
gPCR-pJH216t_F	CGGCAACAGCTGACCAA-GCTTCCATTCAGGTCGAGC	Amplification of plas-mid backbone
gPCR-pJH216t_R	CGTAACTGGTGGTCAC-CGTCAT-ATGTATATCTCCTTCTTAAA TTGTTATCCGC	Amplification of plas-mid backbone
gPCR_spnB-TE_F	AGAAGGAGATATACAT-ATGACGGTGACCACCAG-TTACG	Amplification of spnB fused to EryPKS TE
gPCR_spnB-TE_R	AATGGAAGCTT-GGTCAGCTGTTGCCGCCACC	Amplification of spnB fused to EryPKS TE
gPCR-pBH026 LipPKS-Pp_mcu_F	TCTCGGCAGGAG-CAAGGTGAGATGACAGGA-GATCCTGCC	Amplification of plas-mid backbone part 1.1 for D1 and D2
gPCR-pBH026 LipPKS-Pp_mcu_R	CAGGATCAC-GTGGGCGTTGG	Amplification of plas-mid backbone part 1.1 for D1
gPCR-pBH026 LipPKS-Pp_mcu_R	GCTCGGCCTTCAGGTACG	Amplification of plas-mid backbone part 1.2 for D2
gPCR-pBH026 LipPKS-Pp_mcu_F	CGAACGCGAGTCCTGAG-GATCCAAACTCGAG-TAAGGATCTCCAGG	Amplification of plas-mid backbone part 2
gPCR-pBH026 LipPKS-Pp_mcu_R	GTCATCTCACCTT-GCTCCTGCCGA-GAAAGTATCC	Amplification of plas-mid backbone part 2
gPCR-SpnA(AT-KR-ACP)_F	GCGGCACCAACGCCCAC-GTGATCCTG-GAACAGCCCCCGGGAGTGC	Amplification of SpnA(AT-KR-ACP) and SpnA-COMM
gPCR-SpnA(AT-KR-ACP)_R	CGAGTTTGGATCCTCAG-GACTCGCGTTCGCCC	Amplification of SpnA(AT-KR-ACP)
gPCR-SpnA-COMM_F	CATCGCCGCGTAC-CTGAAGGCCGAGCTGCTTCC CGAGTCCGCAGGAGCA	Amplification of SpnA-COMM

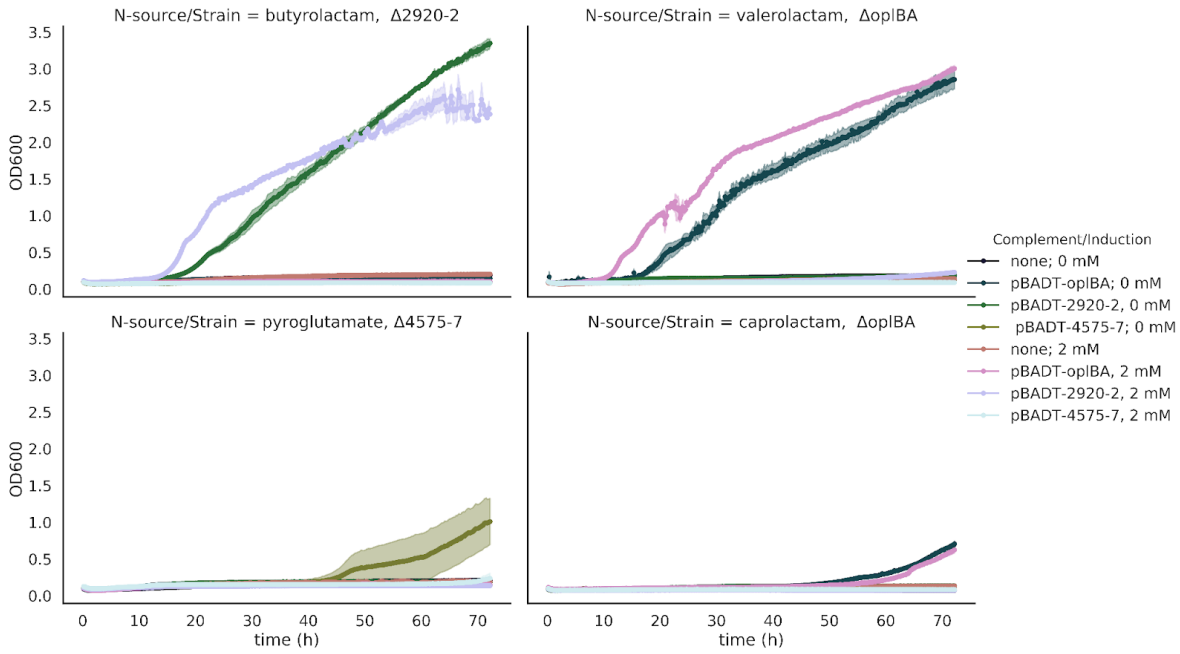


Supplemental Figure 1: Complementation assays of lactamase knockouts. Knockout strains were complemented with arabinose induced expression of the lactamase from a pBADT kanamycin resistant backbone. Wild type controls carried RFP expressed from the same backbone. It appeared that leaky expression of the lactamase was sufficient to restore growth in the deletion strains.

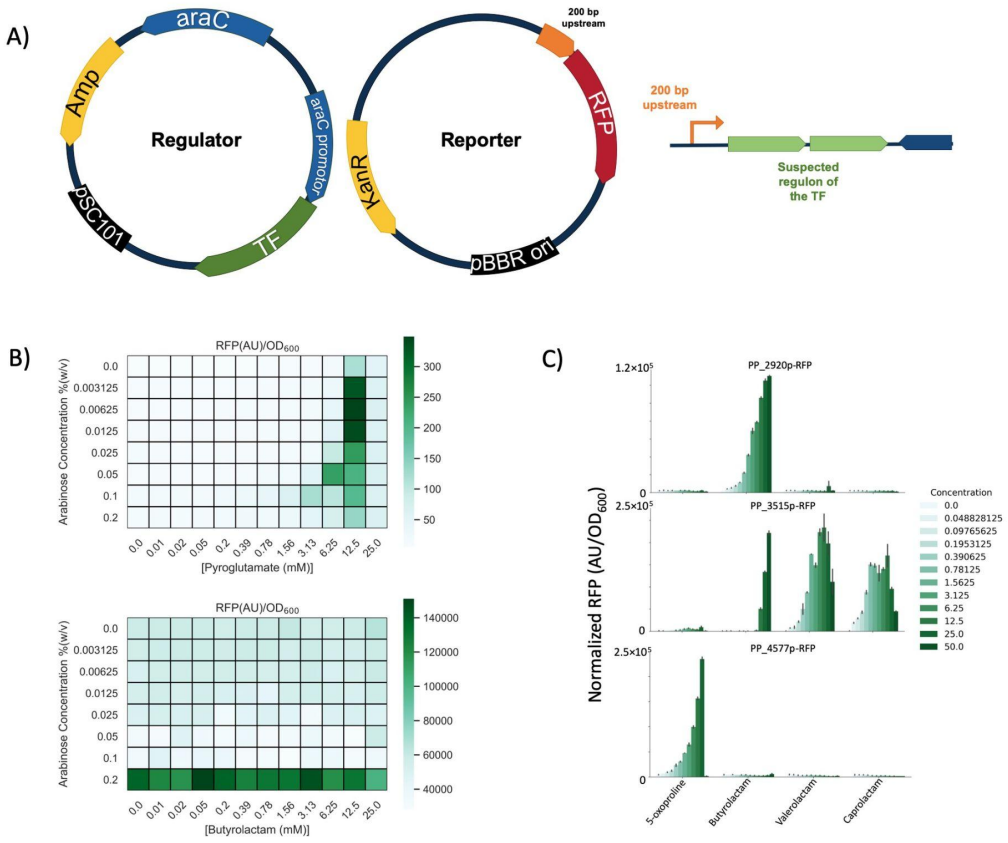


Supplemental Figure 2: Growth curves of lactamase knockout strains with the ω-amino acid corresponding to the lactam substrate of the deleted lactamase as a nitrogen source. Growth with

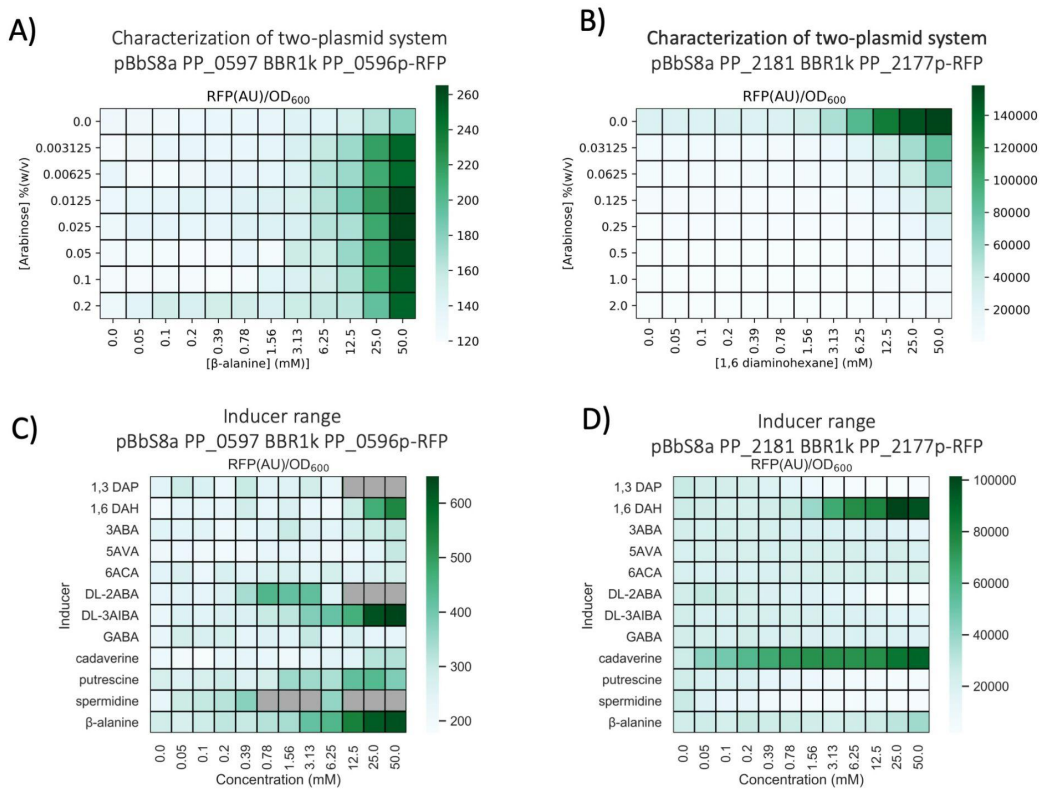
ammonium chloride as a nitrogen source appears the same in the wild-type and knockout strains, while growth with the ω-amino acid appears slightly hindered in the lactamase knockout strains.



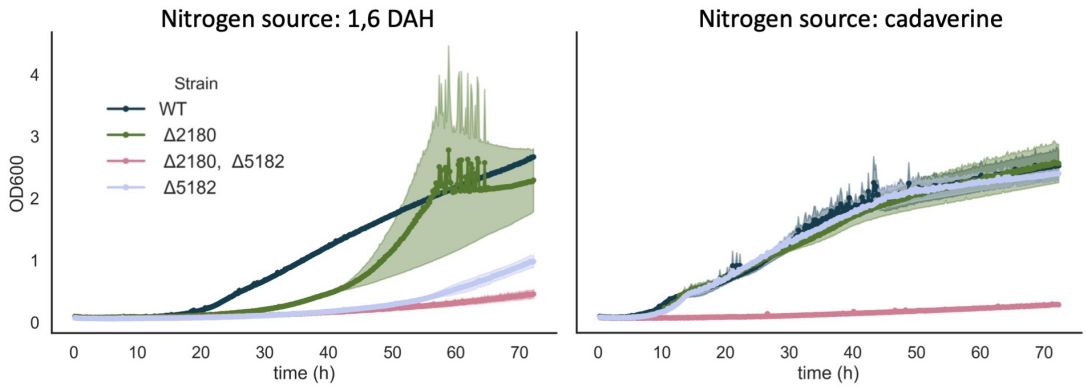
Supplemental Figure 3: Complementation assay of lactamase knockouts with plasmid-based expression of three different lactamases. For each lactam substrate, deletion strains of the native lactamase could only have growth restored with plasmid-based expression of the lactamase that was deleted, indicating high specificity of the lactamases.



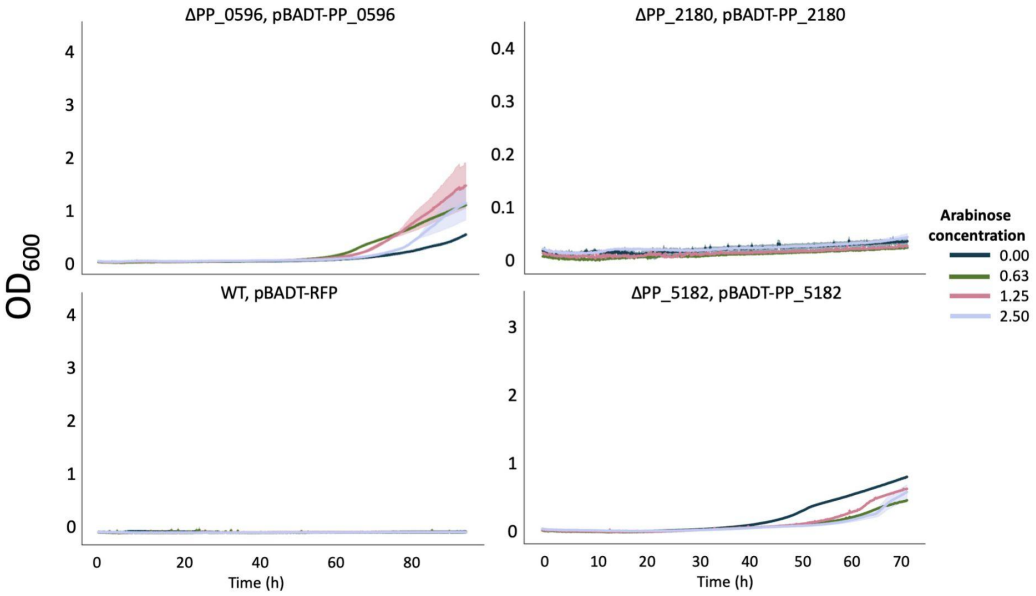
Supplemental Figure 4: Transcription factor assay for *P. putida* lactamases. A) Genetic layout of two-plasmid biosensor systems. B) Initial two-plasmid test in *E. coli* XL1 blue, with varying expression of transcription factor and concentration of inducer. C) Reporter plasmid assay in *P. putida.* Two-fold dilutions of valerolactam, caprolactam, 5-oxoproline, and butyrolactam were tested at a concentration range from 0 to 50 mM (n = 3). Lactams were added to MOPS minimal medium with 20 mM glucose and 10 mM ammonium.



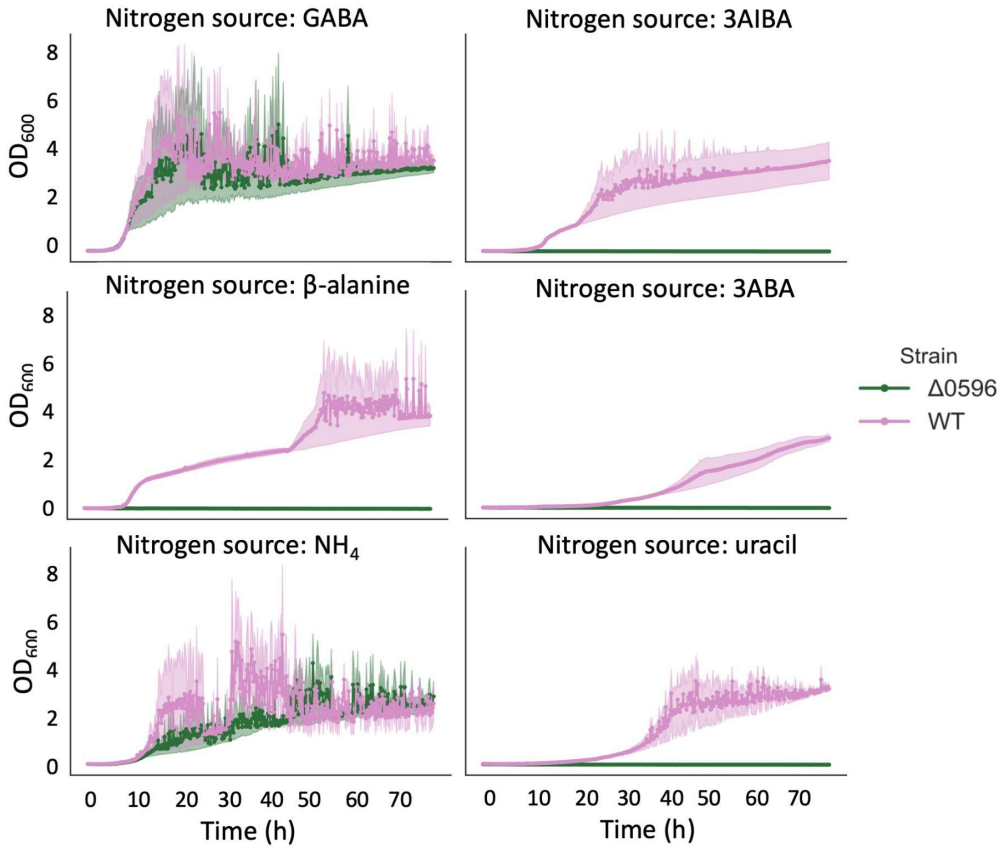
Supplemental Figure 5: Identification of potential transaminase transcription factors. A) Two-plasmid biosensor system with the transcription factor PP_0597 and the promoter of PP_0596. B) Two-plasmid biosensor system with the transcription factor PP_2181 and the promoter of PP_2177 C) Inducer range of the pBbS8a PP_0597 BBR1k PP_0596p RFP system. Expression of PP_0596 was induced with 0.0125 wt% arabinose. Since RFP signal is normalized by OD, low OD can artificially inflate the apparent response. Gray squares indicate at the end of the 24-hour assay, final OD was less than double the starting OD. It appears that the presence of this biosensor system increases the toxicity of some of the inducers tested. This could be due to cross talk with native *E. coli* regulation. D) Inducer range of the pBbS8a PP_2181 BBR1k PP_2177p RFP system. Expression of PP_2181 relied on promoter leakiness, as was indicated optimal in part B.



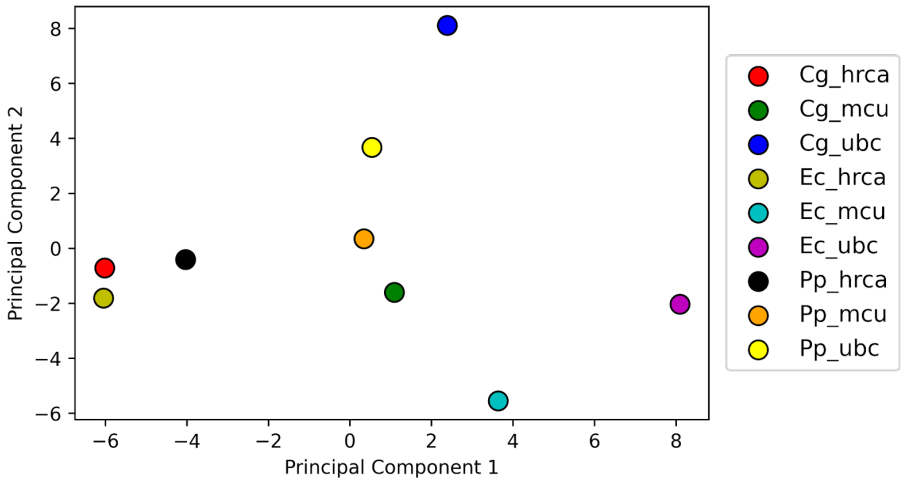
Supplemental Figure 6: Growth of PP_2180 and PP_5182 deletion strains on the diamines cadaverine and 1,6-DAH. Shown are the WT (blue), ΔPP_2180 (green), ΔPP_2180 + ΔPP_5182 (pink), and ΔPP_5182 (grey) strains.



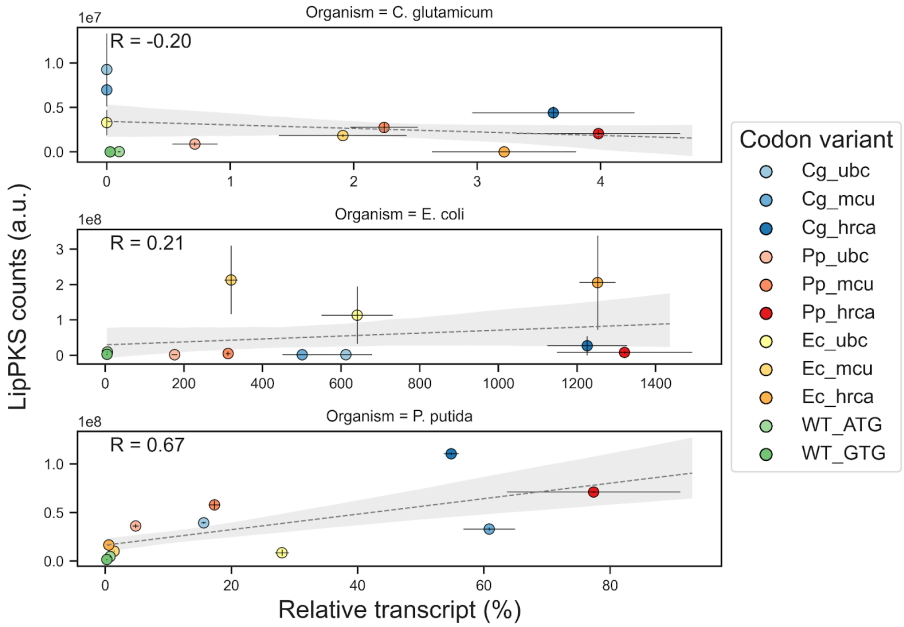
Supplemental Figure 7: Complemented aminotransferase knockout strains grown with 6ACA as a sole nitrogen source.



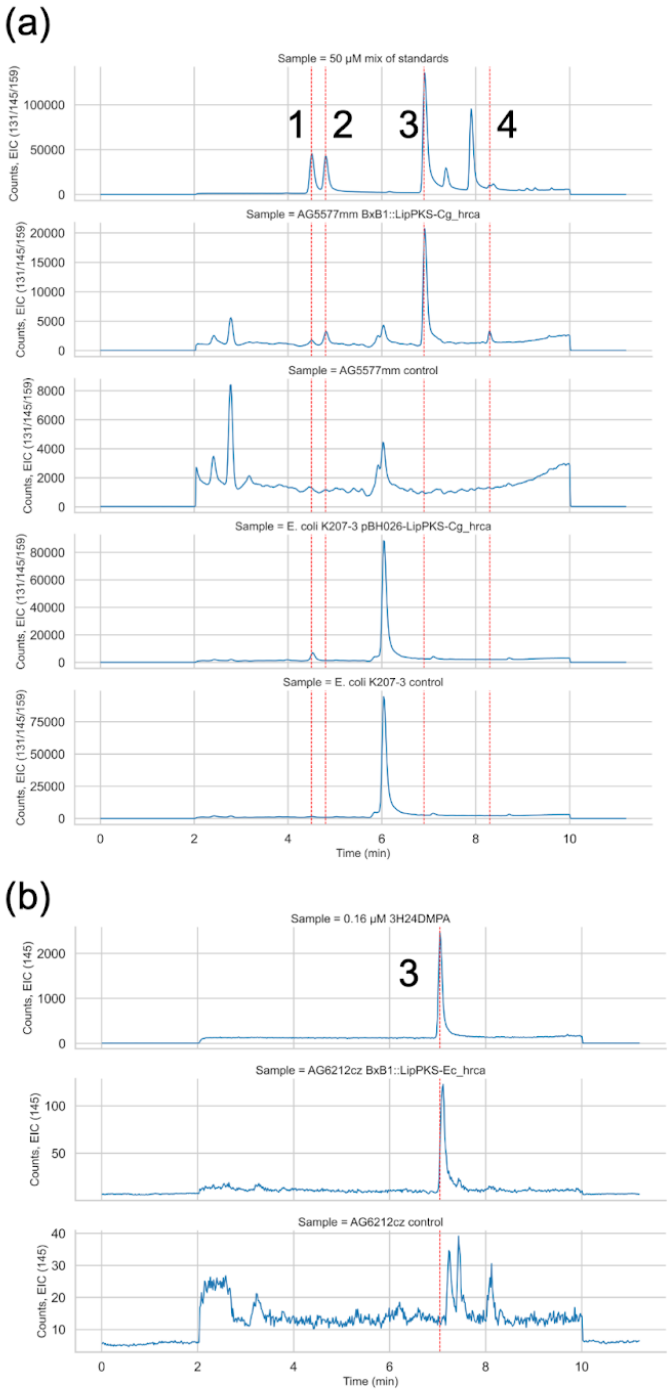
Supplemental Figure 8: Comparison of growth of ΔPP_0596 (pink) vs WT (green) on various nitrogen sources. PP_0596 appeared necessary for utilization of uracil, 3ABA, 3AIBA, and β-alanine.



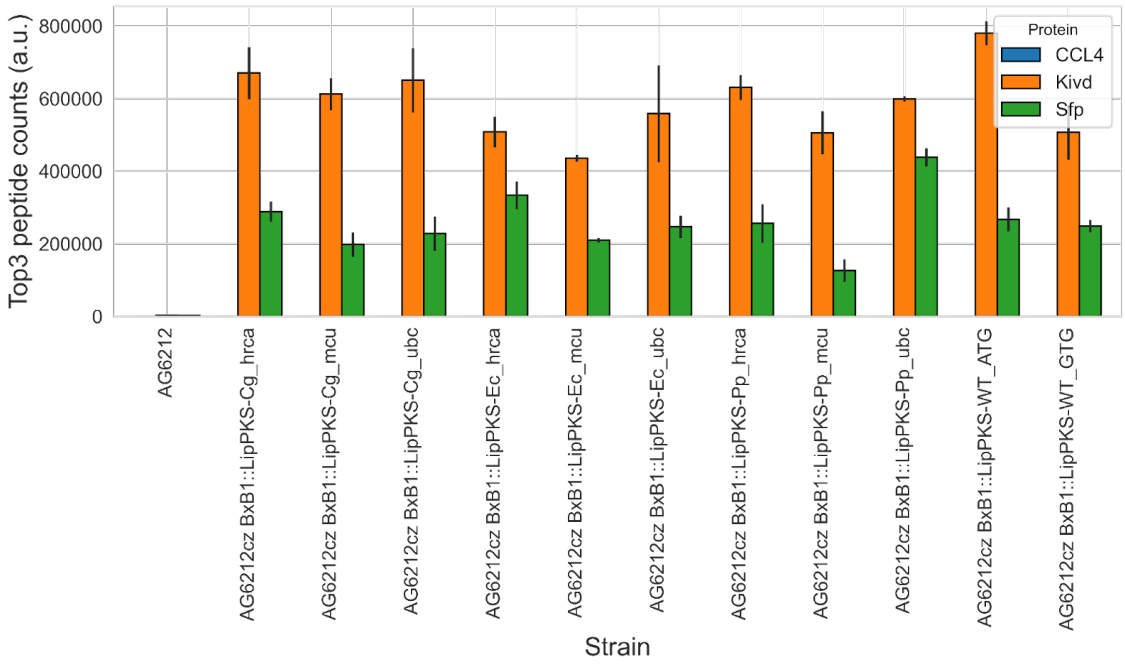
Supplemental Figure 9: Principal component analysis of optimized LipPKS sequences. A high similarity in codon usage leads to clustering of the respective nucleotide sequences. The hrca optimized genes are in close proximity to each other.



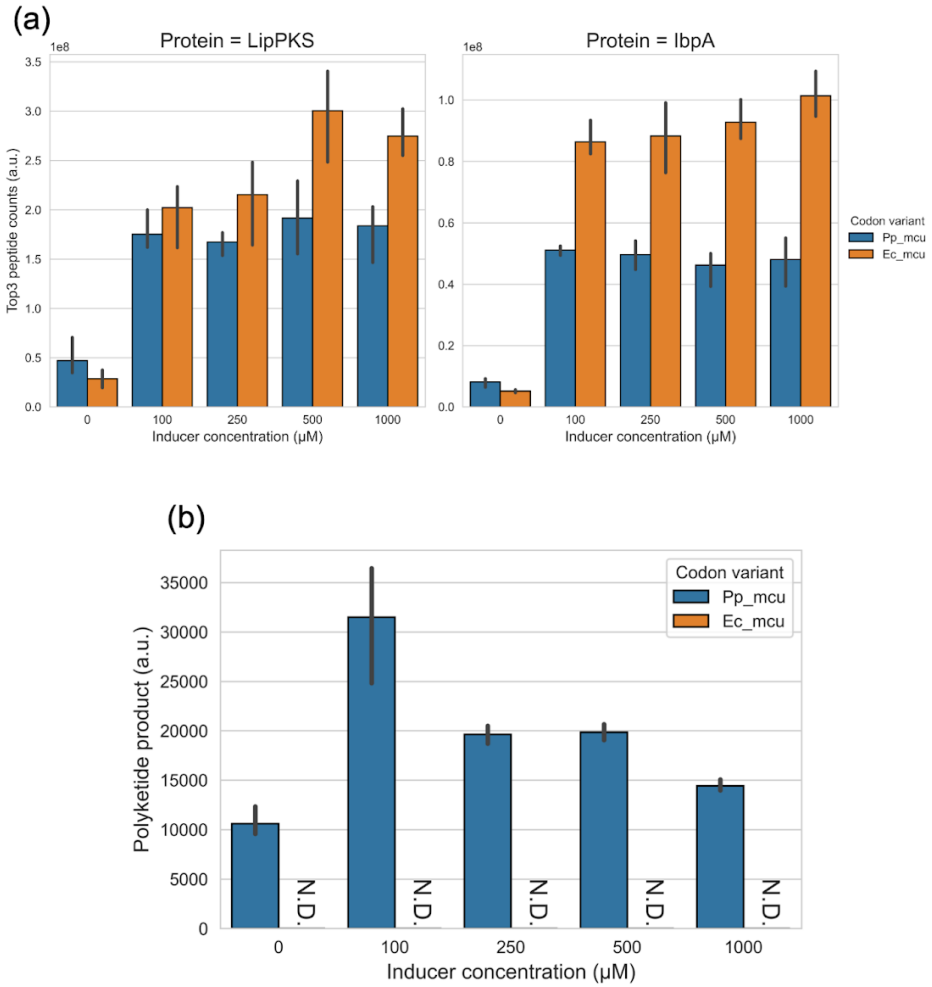
Supplemental Figure 10: Regression plot of LipPKS counts and relative transcript for *C. glutamicum*, *E. coli*, and *P. putida* (n = 3). Only *P. putida* showed a slight correlation between these two parameters.



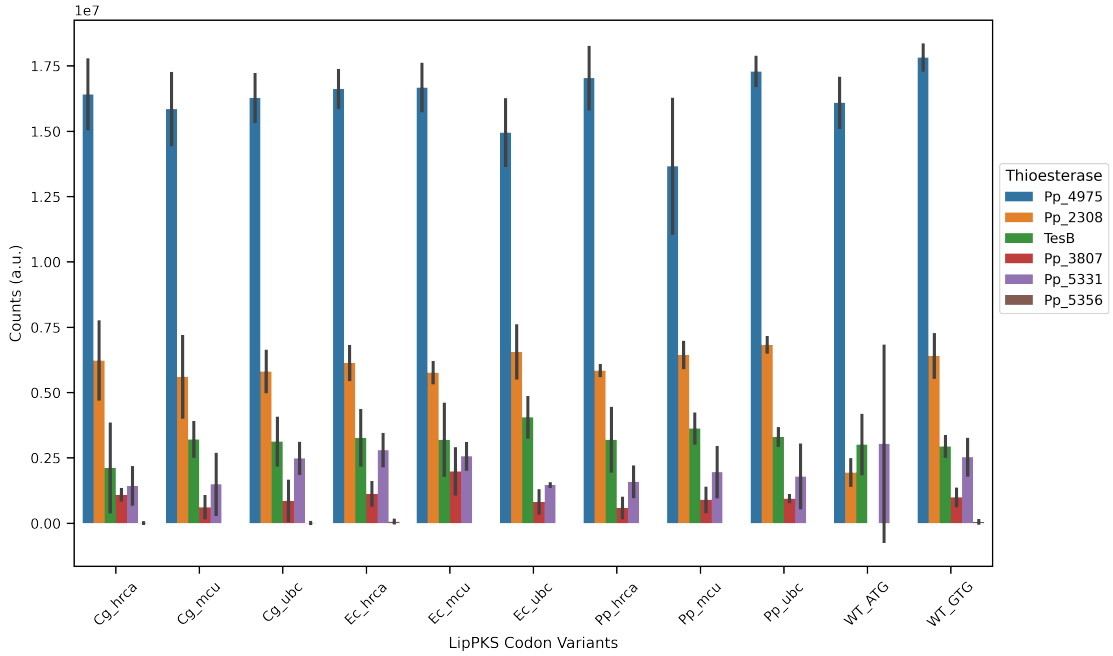
Supplemental Figure 11: LC-MS chromatograms showing the production of unnatural polyketides in *C. glutamicum* (a), *E. coli* and *P. putida* (b). Red lines are indicating the retention time of the authentic standard. Additional peaks for the authentic standards are most likely caused by racemic mixtures of the 3-hydroxy acids. Standards for (2S,3S)-3-hydroxy-2,4-dimethylpentanoic acid (3H24DMPA) and (2S,3S)-3-hydroxy-2-methylpentanoic acid (3H2MPA) are enantiopure. Standards for 3-hydroxy-4-methylpentanoic acid (3H4MPA) and 3-hydroxy-2,4-dimethylhexanoic acid (3H24DMHA) are racemic mixtures. 1 = 3H2MPA; 2 = 3H4MPA; 3 = 3H24DMPA; 4 = 3H24DMHA



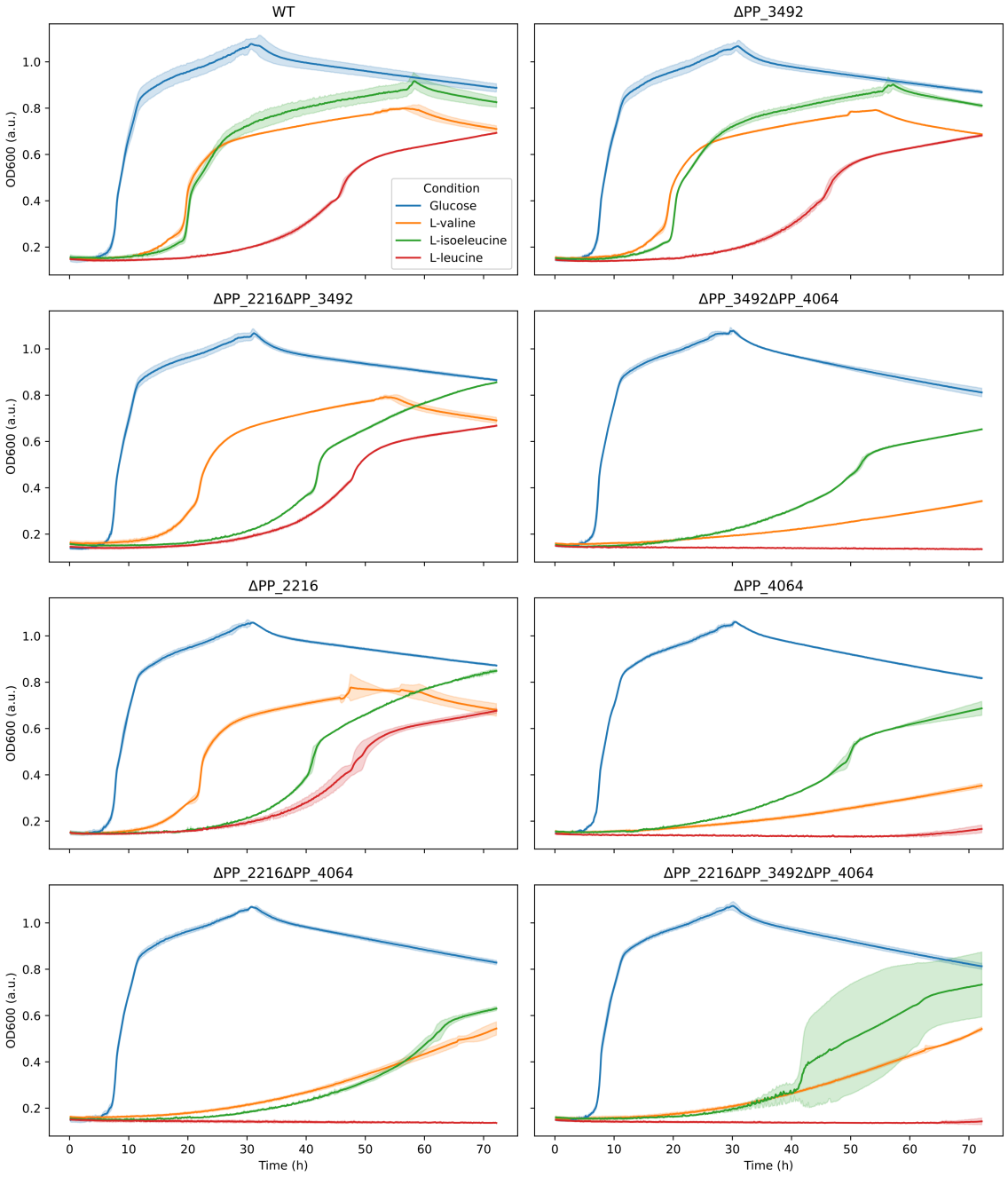
Supplemental Figure 12: Detection of peptides for supplementary pathways in the *C. glutamicum* ATCC 13032 derivative AG6212cz. The corresponding peptides for the phosphopantetheinyl transferase (PPTase) Sfp and ketoisovalerate decarboxylase, Kivd, were detected in all AG6212cz strains. CCL4 peptides could not be detected.



Supplemental Figure 13: Expression of the LipPKS codon variants Pp_mcu and Ec_mcu from the inducible vector system pGingerBG-NahR in *P. putida*. The inducer salicylate was added at the time of inoculation. (a) LipPKS protein levels were comparable, while IbpA levels were significantly higher for Ec_mcu. (b) LC-MS analysis of the same samples (n = 3) revealed no production of the corresponding polyketide for the Ec_mcu variant.



Supplemental Figure 14: Peptide counts for detected thioesterases in PKS-expressing *P. putida* strains. Peptide counts were extracted from a previously published proteomics dataset [121].



Supplemental Figure 15: Plate-based growth assay with BCAAs as carbon sources. Mutant growth assay using glucose (blue), L-valine (yellow), L-isoleucine (green), or L-leucine (red) as the sole source of carbon.

Curriculum vitae

Personal data

Name:	Matthias Schmidt
Born:	April 30, 1993, in Siegburg (NRW), Germany
Nationality:	German

Education

since 2020	PhD student at the Institute of Applied Microbiology of the RWTH Aachen University (Aachen, Germany)
2016-2020	Master of Science in Molecular and Applied Biotechnology at RWTH Aachen University (Aachen, Germany)
2012-2016	Bachelor of Science in Chemistry and Biotechnology at Hochschule Niederrhein University of Applied Sciences (Krefeld, Germany)
2004-2012	Anno Gymnasium Siegburg (Siegburg, Germany) for the acquisition of Abitur (higher education entrance qualification)

Work experience

since 2020	Visiting scholar at the University of California, Berkeley (Berkeley, United States) and the Joint BioEnergy Institute (Emeryville, United States)
2019	Visiting scholar at the Joint BioEnergy Institute (Emeryville, United States)
2018	Visiting scholar at the Joint BioEnergy Institute (Emeryville, United States)
2017-2018	Research assistant at the Institute of Applied Microbiology of the RWTH Aachen University (Aachen, Germany)

Awards and Grants

2023	ME15 Travel Grant to attend the Metabolic Engineering 15 conference in Singapore (Singapore)
2022	Pseudomonas 2022 Travel Grant to attend the Pseudomonas 2022 conference in Atlanta (Atlanta, United States)
2018	ACalNet Travel Grant to visit the Joint BioEnergy Institute in Emeryville (Emeryville, United States)

Publications

Englund, E., Schmidt, M., Nava, A. A., Klass, S., Keiser, L., Dan, Q., Katz, L., Yuzawa, S., and Keasling, J. D. (2023) Biosensor guided polyketide synthases engineering for optimization of domain exchange boundaries. *Nat. Commun. 14*, 4871.

Englund, E., Schmidt, M., Nava, A. A., Lechner, A., Deng, K., Jocic, R., Lin, Y., Roberts, J., Benites, V. T., Kakumanu, R., Gin, J. W., Chen, Y., Liu, Y., Petzold, C. J., Baidoo, E. E. K., Northen, T. R., Adams, P. D., Katz, L., Yuzawa, S., and Keasling, J. D. (2023) Expanding extender substrate selection for unnatural polyketide biosynthesis by acyltransferase domain exchange within a modular polyketide synthase. *J. Am. Chem. Soc. 145*, 8822–8832.

Pearson, A. N., Incha, M. R., Ho, C., Schmidt, M., Roberts, J. B., Nava, A. A., and Keasling, J. D. (2023) Characterization and diversification of AraC/XylS family regulators guided by transposon sequencing. *BioRxiv*.

Schmidt, M., Lee, N., Zhan, C., Roberts, J. B., Nava, A. A., Keiser, L., Vilchez, A., Chen, Y., Petzold, C. J., Haushalter, R. W., Blank, L. M., and Keasling, J. D. (2023) Maximizing heterologous expression of engineered type I polyketide synthases: investigating codon optimization strategies. *ACS Synth Biol. 12,* 3366-3380.

Valencia, L. E., Incha, M. R., Schmidt, M., Pearson, A. N., Thompson, M. G., Roberts, J. B., Mehling, M., Yin, K., Sun, N., Oka, A., Shih, P. M., Blank, L. M., Gladden, J., and Keasling, J. D. (2022) Engineering *Pseudomonas putida* KT2440 for chain length tailored free fatty acid and oleochemical production. *Commun. Biol. 5*, 1363.

Schmidt, M., Pearson, A. N., Incha, M. R., Thompson, M. G., Baidoo, E. E. K., Kakumanu, R., Mukhopadhyay, A., Shih, P. M., Deutschbauer, A. M., Blank, L. M., and Keasling, J. D. (2022) Nitrogen metabolism in *Pseudomonas putida*: functional analysis using random barcode transposon sequencing. *Appl. Environ. Microbiol. 88*, e0243021.

Thompson, M. G., Incha, M. R., Pearson, A. N., Schmidt, M., Sharpless, W. A., Eiben, C. B., Cruz-Morales, P., Blake-Hedges, J. M., Liu, Y., Adams, C. A., Haushalter, R. W., Krishna, R. N., Lichtner, P., Blank, L. M., Mukhopadhyay, A., Deutschbauer, A. M., Shih, P. M., and Keasling, J. D. (2020) Fatty acid and alcohol metabolism in *Pseudomonas putida*: functional analysis using random barcode transposon sequencing. *Appl. Environ. Microbiol. 86*, e01665-20.

Incha, M. R., Thompson, M. G., Blake-Hedges, J. M., Liu, Y., Pearson, A. N., Schmidt, M., Gin, J. W., Petzold, C. J., Deutschbauer, A. M., and Keasling, J. D. (2020) Leveraging host metabolism for bisdemethoxycurcumin production in *Pseudomonas putida*. *Metab. Eng. Commun. 10*, e00119.

Eck, A., Schmidt, M., Hamer, S., Ruff, A. J., Förster, J., Schwaneberg, U., Blank, L. M., Wiechert, W., and Oldiges, M. (2018) Improved microscale cultivation of *Pichia pastoris* for clonal screening. *Fungal Biol. Biotechnol. 5*, 8.

Poster presentations

Englund, E., Schmidt, M., Nava, A. A., Klass, S., Keiser, L., Dan, Q., Katz, L., Yuzawa, S., and Keasling, J. D. (May 16-19, 2022). Incorporation of unusual extender substrates into polyketides by acyltransferase swapping. 3rd Synthetic Biology of Natural Products Conference. Cancun, Mexico.

Schmidt, M.; Thompson, M. G.; Pearson A. N.; Incha, M. R.; Vilchez, A. A.; Blake-Hedges, J. M.; Haushalter, R. W.; Baidoo, E. E. K; Amer, B.; Petzold, C. J.; Chen, Y.; Deutschbauer, A. M.; Blank, L. M.; Keasling, J. D. (April 18-23, 2022). Exploiting the potential of *Pseudomonas putida* as a host for engineered type I polyketide synthases. Pseudomonas 2022. Atlanta, GA, United States.

Schmidt, M.; Thompson, M. G.; Pearson A. N.; Incha, M. R.; Vilchez, A. A.; Deutschbauer, A. M.; Yuzawa, S.; Blank, L. M.; Keasling, J. D. (June 11-15, 2023). Exploiting the potential of *Pseudomonas putida* as a host for engineered type I polyketide synthases. Metabolic Engineering 15. Singapore.