

High-Throughput Screening in Metabolic Engineering for Improved Production Strains

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

Presented by

MAXIMILIAN OLE BAHLS

M.Sc. Biotechnology, ETH Zurich

Born on 30.06.1989

Citizen of Germany

Accepted on the recommendation of

Prof. Dr. Sven Panke (ETH Zurich, Switzerland), examiner

Prof. Dr. Victor de Lorenzo (CSIC, Spain) co-examiner

Prof. Dr. Ralf Takors (University of Stuttgart, Germany), co-examiner

Prof. Dr. Martin Fussenegger (ETH Zurich, Switzerland), co-examiner

There are moments in a match when the ball hits the top of the net and for a split second it can either go forward or fall back.

With a little luck it goes forward, and you win. Or maybe it doesn't, and you lose.

Match Point by Allen Stewart Konigsberg

Abstract

Living cells are endowed with biochemical reaction networks that define the attainable *in vivo* conversions of substrate molecules to reaction products. By analyzing and modifying these reactions, mainly catalyzed by enzymes, the cellular features can be intentionally improved for a particular reaction pathway or even new features are added by recombinant DNA technology. This metabolic engineering champions for example the production of chemicals by microorganisms and lays the foundation for the development of bioprocesses for products naturally not made at economic titers or purities. The vast and ongoing progress made in the fields of systems and synthetic biology underpins novel metabolic engineering strategies and ultimately fuels the development of urgently required environmentally benign bioprocesses for replacing processes utilizing limited fossil resources in unsustainable ways. In this thesis, metabolic engineering of two industrially relevant microorganisms for the showcase production of isopentanol, an alcohol currently derived from petrochemistry is addressed. To this end *Escherichia coli*, arguably the incumbent workhorse in biotechnology, and *Pseudomonas putida*, an up-and-coming microorganism for metabolic engineering, were equipped with a pathway for alcohol production originating from yeast. For both microbes substantial product titers made from glucose as input substrate were found. However, further analysis showed that with *P. putida* also isovaleric acid was made as a major side product. This could be attributed to the versatile metabolism of this bacterium, which is *per se* a feature, but needs further optimization for improved isopentanol titers. To address the side product formation at the process level, a two-stage fed-batch protocol was developed. Limiting oxygen supply during the production phase allowed to improve isopentanol titers as well as curbing the short-chain fatty acid production. Second, with the development of a protocol for constructing sRNA libraries we investigated a potential solution for side product formation at the molecular level. sRNAs allow the knock-down of gene expression by specifically limiting translation of a target mRNA to the corresponding protein. As the precise enzyme(s) responsible for a side reaction are often unknown, for instance here the ones catalyzing the reaction to the unwanted fatty acid, naïvely targeting all potential targets and then screening for improved strain properties is a valid option. To this end, a simple computational workflow was developed for the conversion of genome annotations into DNA oligomers encoding sRNAs against these annotated genes. Due to increasingly cheap on-chip, pooled DNA synthesis such vast libraries can economically be produced. However, as this genetic engineering side is increasingly powerful, screening the vast corresponding strain libraries for improved product titers is to an increasing degree throughput-limiting. Small molecule products, including isopentanol,

are analytically relatively inconspicuous and therefore traditionally require chemical analyses for detection. These analyses, usually by liquid or gas chromatography, quickly become the limiting factor when dealing with large strain libraries. This frequently occurring problem was addressed by the development of a biosensor circuit, allowing the formation of an easily read-out signal as a proxy for the actual alcohol titer. Here, the underlying genetic blueprint consisted of a transcription factor, which is activated by the product molecule of interest and subsequently binds to its cognate promoter for expression of a green fluorescent protein. As there are few transcription factors available encoding this functionality for isopentanol, a related transcription factor (*AlkS*) was chosen as a starting point for engineering this functionality by means of directed evolution. Libraries of *alkS* variants were screened by fluorescence assisted cell sorting and variants encoding the sought after specificity were found. The corresponding biosensor variants were characterized at the single-cell level for the detection of multiple industrially relevant alcohols. Besides, the sensor system was successfully applied to real-life detection of isopentanol produced by microbes. First, a detailed automation-based protocol relying on liquid handling robots was established. Subsequently, this protocol allowed for successful screening of an overexpression pathway library for improved isopentanol titers in an *E. coli* strain equipped with the biosensor circuit. Second, for a *P. putida* isopentanol production strain co-cultivated with an orthogonal *E. coli* biosensor strain, product dependent biosensor output was demonstrated.

In conclusion, this thesis work evaluated two industrially relevant microbes for their potential in isopentanol production and concomitantly developed methods for strain library generation as well as for automated library screening enabled by biosensor circuits.

Zusammenfassung

Lebende Zellen sind mit biochemischen Reaktionsnetzwerken ausgestattet, welche die *in vivo* möglichen Umwandlungen von Substratmolekülen zu Reaktionsprodukten definieren. Durch die Analyse und Modifikation dieser Reaktionen, die hauptsächlich durch Enzyme katalysiert werden, können die zellulären Eigenschaften für einen bestimmten Reaktionsweg gezielt verbessert werden oder sogar neue Eigenschaften durch rekombinante DNS-Technologie hinzugefügt werden. Dieses biotechnologische Fachgebiet des Metabolic Engineering ermöglicht zum Beispiel die Produktion von Chemikalien durch Mikroorganismen und legt damit den Grundstein für die Entwicklung von Bioprozessen für Produkte, die auf natürliche Weise nicht in wirtschaftlichen Produkttitern oder Reinheiten hergestellt werden können. Die enormen Fortschritte in den Bereichen der Systembiologie und der synthetischen Biologie unterstützen neuartige Metabolic Engineering-Strategien und tragen so zu der Entwicklung von dringend benötigten und umweltfreundlichen Bioprozessen bei. Dies ist wichtig, um Produktionsprozesse zu ersetzen, die begrenzte fossile Ressourcen auf nicht nachhaltige Weise nutzen. In dieser Arbeit wird das Metabolic Engineering zweier industriell relevanter Mikroorganismen für die potentielle Produktion von Isopentanol, einem derzeit aus der Petrochemie stammenden Alkohol, untersucht. Zu diesem Zweck wurden *Escherichia coli*, das etablierte Arbeitspferd der Biotechnologie, und *Pseudomonas putida*, ein an Beliebtheit zunehmender Mikroorganismus für das Metabolic Engineering, mit einem aus Hefe stammenden Weg zur Alkoholproduktion ausgestattet. Für beide Mikroben wurden signifikante Produkttiter mit Glukose als Eingangssubstrat gefunden.

Weitere Analysen zeigten jedoch, dass mit *P. putida* auch grössere Mengen an Isovaleriansäure als Nebenprodukt hergestellt wurden. Dies kann möglicherweise auf den vielseitigen Metabolismus dieses Bakteriums zurückgeführt werden, was an sich eine vorteilhafte Besonderheit ist, aber für verbesserte Isopentanol-Titer einer weiteren Optimierung bedarf. Um die Nebenproduktbildung auf Prozessebene zu adressieren, wurde ein zweistufiges Fed-Batch-Protokoll entwickelt. Durch die Limitierung der Sauerstoffzufuhr während der Alkoholproduktion konnten sowohl die Isopentanol-Titer verbessert als auch die Bildung kurzkettiger Fettsäuren verringert werden. Ausserdem wurde mit der Entwicklung eines Protokolls zur Konstruktion von sRNA-Bibliotheken eine mögliche Lösung für die unvorteilhafte Bildung von Nebenprodukten auf molekularer Ebene vorangetrieben. sRNAs ermöglichen eine negative Regulation der Genexpression, indem sie die Translation einer Ziel-mRNA zum entsprechenden Protein begrenzen. Da die Enzyme, welche für eine Nebenreaktion verantwortlich sind, oft nicht im Detail bekannt sind, wie z. B. hier für die

Reaktion zur unerwünschten Fettsäure, ist das naive Anvisieren aller Gene als potentielle Ziele, mit nachfolgendem Screening für verbesserte Stamm-Varianten, eine sinnvolle Option. Zu diesem Zweck wurde ein einfacher computergestützter Arbeitsablauf für die Umwandlung von Genom-Annotationen in DNA-Oligomere entwickelt, die sRNAs gegen alle annotierten Gene kodieren. Durch die zunehmend billige, gemeinsame DNA-Synthese können solche großen Bibliotheken wirtschaftlich hergestellt werden. Da die DNA-Synthese jedoch immer leistungsfähiger wird, ist das Screening der entsprechenden Stammbibliotheken auf verbesserte Produkttiter hin in zunehmendem Maße durchsatzlimitierend. Niedermolekulare Produkte, einschließlich Isopentanol, sind relativ unauffällig und erfordern chemische Analysen zum Nachweis. Diese Analysen, meist mittels Flüssigkeits- oder Gaschromatographie, werden bei großen Produktionsstammbibliotheken schnell zum limitierenden Faktor. Dieses häufig auftretende Problem wurde durch die Entwicklung eines Biosensor-Schaltkreises angegangen, der die Bildung eines leicht auslesbaren Signals als Stellvertreter für den tatsächlichen Alkoholtiter ermöglicht. Hier bestand der zugrundeliegende genetische Bauplan darin, dass ein Transkriptionsfaktor durch das Produktmolekül aktiviert wird und anschließend zur Expression eines fluoreszierenden Proteins führt. Da es wenige Transkriptionsfaktoren gibt, die diese Funktionalität für Isopentanol kodieren, wurde ein naheliegender Transkriptionsfaktor (*AlkS*) als Ausgangspunkt für die Entwicklung dieser Funktionalität durch gerichtete Evolution gewählt. Varianten von *alkS* wurden mittels fluoreszenzunterstützter Zellsortierung ausgewählt und es wurden Varianten gefunden, welche die gesuchte Spezifität kodieren. Die entsprechenden Biosensorvarianten wurden quantitativ für die Detektion mehrerer industriell relevanter Alkohole charakterisiert. Außerdem wurde das Sensorsystem erfolgreich für die Detektion von mikrobiell produziertem Isopentanol eingesetzt. Zunächst wurde ein detailliertes Automatisierungsprotokoll etabliert, das auf der Nutzung von Liquid-Handling-Robotern basiert. Anschließend ermöglichte dieses Protokoll ein erfolgreiches Screening einer DNA-Bibliothek für verbesserte Isopentanol-Titer in einem mit dem Biosensor-Schaltkreis ausgestatteten *E. coli*-Stamm. Zweitens wurde für einen *P. putida* Isopentanol-Produktionsstamm, der zusammen mit einem orthogonalen *E. coli*-Biosensor-Stamm kultiviert wurde, eine produktabhängige Biosensoraktivierung gezeigt.

Zusammenfassend wurden in dieser Arbeit zwei industriell relevante Mikroben auf ihr Potenzial für die Isopentanolproduktion hin untersucht und gleichzeitig Methoden für die Generierung von Stammbibliotheken sowie für das automatisierte Screening mit Hilfe von Biosensor-Stämmen entwickelt.

Contents

Abstract	i
Zusammenfassung	iii
Contents	v
Chapter 1: General Introduction	1
1.1 Engineering microbes for sustainable chemical processes	2
1.2 High-throughput screening of microbe variants	4
1.3 Deep technology and global biofoundries	6
1.4 Scope of the thesis.....	7
Chapter 2: Novel Sensors for Engineering Microbiology	9
2.1 Abstract	10
2.2 Introduction.....	10
2.3 Genetically encoded RNA-based biosensors	14
2.4 Genetically encoded protein-based biosensors.....	18
2.5 Optimization of the biosensor's characteristics.....	28
2.6 Research needs	34
2.7 Concluding remarks	34
2.8 Acknowledgements	34
2.9 Author contributions	35
Chapter 3: Directed Evolution of Biofuel-Responsive Biosensors for Automated Optimization of Branched-Chain Alcohol Biosynthesis	37
3.1 Abstract	38
3.2 Introduction.....	38
3.3 Results	40
3.4 Discussion	54
3.5 Methods.....	58
3.6 Acknowledgements	69
3.7 Author contributions	69
3.8 Supplementary information.....	70

Chapter 4: Short-Chain Alcohol and Fatty Acid Biosynthesis in <i>Pseudomonas putida</i> with Co-Culture Based <i>in vivo</i> Product Detection.....	103
4.1 Abstract.....	104
4.2 Introduction	105
4.3 Results.....	107
4.4 Discussion	116
4.5 Methods	119
4.6 Acknowledgements.....	125
4.7 Author contributions.....	125
4.8 Supplementary information	126
Chapter 5: Towards Genome Wide sRNA-Based Gene Knock Down for Improved <i>P. putida</i> Production Strains.....	137
5.1 Abstract.....	138
5.2 Introduction	138
5.3 Results.....	140
5.4 Discussion	144
5.5 Methods	146
5.6 Acknowledgements.....	148
5.7 Supplementary information	149
Chapter 6: Conclusion and Outlook.....	155
Chapter 7: References	161
Chapter 8: Acknowledgements.....	181
Chapter 9: Curriculum Vitae.....	183

Chapter 1: General Introduction

1.1 Engineering microbes for sustainable chemical processes

The ultimate goal of industrial biotechnology has been, in an artisanal way since millennia, the utilization of living organisms for the synthesis of valuable products from a cheaper or perishable resource. A classic example is the production of alcoholic beverages by yeast fermentation, such as beer from barley or wine from grape juice¹. While the details of the process of alcohol (and CO₂) formation from carbohydrates, catalyzed by complex enzyme cascades, was unknown until relatively recently², the underlying biotechnological concept barely changed over thousands of years. Mainly, a carbon- and nutrient-rich culture medium is provided in a vessel, homogeneously mixed with added microorganisms, and a valuable product is formed within a few hours to days.

While in ancient times microorganisms were most likely available due to their natural presence in certain environments, from the mid-19th century onwards it became clear that different organisms lead to different product qualities, such as lactic “yeast” leading to the formation of lactic acid³, instead of the previously known formation of alcohol. In combination with the ability to isolate pure bacterial cultures (it turned out that lactic acid is actually produced by a bacterium and not by yeast), these findings lead to the first heydays of biotechnology at the start of the last century. For the first time, particular microbes were rationally chosen for a production process, for example *Clostridium* spp. for acetone-butanol-ethanol fermentation⁴. Unfortunately (at least from the perspective of a biotechnologist), a few decades later these bio-processes were rapidly replaced by petrochemical ones^{5, 6}. Despite harsh physicochemical and often environmentally unfriendly conditions, these processes were not only cheaper and more reliable to run, but also lead to the development of new products and industries on a massive scale, e.g. synthetic polymer production. Interestingly, there is renewed interest in developing bio-processes, because they are perceived (often correctly) as environmentally benign, rely on renewable resources and are therefore less carbon-intensive. This renewed interest is largely due to the notion of peak oil, i.e. the issue of reaching the point of maximum oil extraction and eventually running out of cheap fossil resources, as well as global warming caused by excessive usage and burning of these fossil resources^{7, 8}.

As a result, the two examples already mentioned above, microbial alcohol and (poly)lactic acid formation, are once again implemented on an industrial scale, for the formation of fuels and (polylactic acid-based) bioplastics with a reduced carbon footprint⁶. However, instead of relying on a few, naturally-isolated microbes, biotechnology can now build on vast data

repositories obtained from detailed research on the chemistry and physiology of living cells, which can help in improving the living cell as a catalyst.

The current knowledge of cells includes multiple levels of complexity, from the basic DNA level (genomics), via RNA (transcriptomics) and protein levels (proteomics) to the cellular metabolism (metabolomics)⁹. This growing systems-level understanding (systems biology) underpins the design of novel strategies for implementing enzyme cascades¹⁰. Ultimately, such studies can result in bioprocesses for the formation of valuable products that cannot be obtained from natural organisms in an economically viable fashion.

Crucially, instructing cells to change their production behavior is done by writing information in the form of DNA, and in this way new design ideas can nowadays be implemented quickly due to increasingly fast and cheap synthesis of DNA¹¹. Furthermore, for reoccurring tasks such as protein production many pre-tested DNA parts are available, in analogy to parts and blueprints in electrical engineering¹². This engineering strategy based on synthetic DNA is often termed synthetic biology, which is, in the bacterial domain, championing the rational engineering of relatively few, but well-studied microbes, for a particular task of interest. Here, this interest is mainly on the engineering of the metabolism of a microorganism in order to produce an interesting chemical compound from a renewable carbohydrate as efficiently as possible¹³, an interest shared with the field of metabolic engineering¹⁴.

Despite the incredible increase in knowledge about designing and implementing new metabolic pathways and control systems for optimizing pathway performance, a single synthetic biology design rarely leads to the expected outcome¹⁵. In contrast to more established engineering fields, the initial design usually has to be tested in great detail for generating further insight into the system, which is subsequently fed back into the next generation of designs. In fact, this *per se* rational process, known as the design-build-test-learn (DBTL) cycle, has to deal with the fact that the knowledge about the optimal design is still sparse, despite all advances that inform about qualitative interactions. As a consequence, multiple or even random design variations are tested¹⁶. The design space can be quite extensive, but current DNA engineering methods allow the cheap synthesis of millions to billions of varied DNA sequences and therefore make this design space accessible. If the knowledge on how to solve a problem is particularly small, an established method to obtain a working prototype is the simple introduction of small mistakes in an existing but poorly functioning DNA design, followed by evaluating the effect of a large number of such variants for finding incrementally improved variants in a process termed directed evolution¹⁷.

As it turns out, a crucial boundary condition of learning what DNA blueprint to write is represented by the context in which this blueprint is placed. Depending on which of the many available and desirable host microorganisms (or chassis¹⁸) is chosen for engineering, the same DNA construct can lead to different outcomes. This interplay between the single pathway of interest and the – varying – metabolism of the host cell leads to additional uncertainties about the outcomes¹⁹, thus further increasing the need for extensive DBTL cycles.

In summary, despite the best intentions to rationalize the process of designing improved microbes for chemical synthesis, as a matter of fact the process of cell engineering relies to a large extent on some form of directed evolution, sometimes on a small, often on a large scale. This highlights the need for efficient processes to identify improved variants in large populations. Ideally, the information latently contained in such large variant populations (which variations lead to improvements, which do not) can be made available in easily accessible format and used to generate additional knowledge on the system in processes involving artificial intelligence. As a result of these relatively well-established design and build capabilities, high-throughput testing or screening of all microorganism variants, as well as (machine) learning from the results of this screening, has become a focal point in biotechnology^{20, 21}.

1.2 High-throughput screening of microbe variants

Recent advances in synthetic biology and metabolic engineering significantly increased the speed with which new traits and functions can be introduced into microorganisms. Despite constantly improving analysis tools for system-wide metabolic fluxes and the resulting computational support for strain engineering²², such as the *in silico* optimization of product formation by prediction of beneficial gene deletions or novel biosynthetic pathways, the knowledge for directly implementing the optimal pathway in the optimal host microorganism is incomplete²³. As a result, multiple strain variants still have to be tested and iteratively optimized. Screening technology improved to an (ultra)high-throughput state-of-the-art where about 10^8 different design variants can be assessed on the single-cell level or in microfluidic droplets within one work day²⁴. However, such high numbers heavily rely on optical readouts that are necessary to indicate the performance of a variant. For example, fluorogenic dummy molecules allowing optical product detection can often be synthesized for a single reaction step, e.g. in order to improve (or evolve) a single enzyme²⁵. However, this powerful strategy is frequently of limited use for the metabolic engineering of a living

microorganism. First, the problem to solve usually involves many enzyme-catalyzed reactions, with the product of the first reaction being the substrate for the next one and so on. As a result, for optimizing the flux through a metabolic pathway in concert with all enzymes, a potential dummy molecule would have to fit all the intermediate purposes as well until indicating the final product concentration, which is unlikely to occur. Additionally, the overall pathway flux has to be optimized in context of the host cell metabolism. Thus the fluorogenic molecule would have to enter the cell, while not interacting with alternative enzyme pathways, and neither be toxic for the cell. The real product of interest is normally an inconspicuous small molecule that is neither colorful nor fluorescent²¹. For the initial example of alcohol production in yeast, the overall pathway from the carbohydrate (glucose) to alcohol (ethanol) requires more than ten enzyme-catalyzed chemical reactions and ethanol is normally measured with laborious liquid- or gas-chromatographic methods (<10³ measurements per work day)²⁶. In addition, many small molecules including ethanol, are not retained in the microbes but diffuse into the culture medium surrounding them. As a result, the different variants created have to be physically separated from one another in order to be tested without interference, but under conditions that are not harmful for microbial growth. However, separately cultivating large number of cells, for example on standardized microtiter plates, is a laborious undertaking.

A potential solution to this problem is to couple the product of interest to a more accessible readout. A very useful and straight-forward readout is how well the microbes grow, i.e. coupling product formation to growth. This solution could either be directly implemented in the production pathway blueprint²⁷, or achieved with additional engineered DNA circuits that detect the product and conditionally yield a growth advantage²⁸. This advantage could be, for example, the production of an antibiotic resistance marker and selection in a culture medium containing this antibiotic. Unfortunately, measuring growth, for example in terms of colony size on an agar plate or as cell density in liquid culture, requires complicated analyses to detect subtle changes and is only simple for drastic improvements (such as the difference between growth and no-growth). Additionally, the overproduction of most small molecules is to a large extent detrimental to growth as the limited resources are directed towards the product of interest or the product accumulates to such a high concentration that cellular functions are impaired. As a result, the best growing strains could easily be strains that escape the growth-selection²⁹. Alternatively, a gene circuit enabling product detection could be coupled to the formation of fluorescent proteins³⁰. Such fluorescent proteins facilitate the detection of gradual strain improvements, i.e. the more product is made the more

fluorescence is produced, in a comparatively straight-forward manner. Most importantly, fluorescence can be read out quickly and quantitatively, with the corresponding machines being widely available in molecular biology laboratories. However, the number of DNA circuits encoding the capacity of small molecule detection is limited to a few industrially relevant products, though constantly increasing³¹, and their application still requires the physical separation of strain variants in order to prevent false results.

1.3 Deep technology and global biofoundries

As outlined above, the use of microorganisms for the production of chemicals is not a recent idea. Unfortunately, the addition of new target molecules to the list of economically viable production processes is relatively rare. For example, the successful development of a show-case bioprocess for 1,3-propanediol production required about 15 years of work and \$130 million of investment³². However, the preference for environmentally benign processes in combination with building on digital systems biology resources and automated strain development is a relatively recent progress^{7, 33, 34}. This ability might drastically reduce the time required to implement the production of new target molecules. The assembly of digital, big data platforms already lead to “deep” technological innovations, i.e. truly technological innovations that rely on substantial basic research and capital investment as well as machine learning for data analysis³⁵, in other scientific fields and most likely biotechnology could greatly benefit in the future as well³⁶. Most importantly, the digital information has to be integrated into real life DBTL cycles that are powerful enough to cope with the large amount of information to be tested. Here up-and-coming biofoundries, shared high-cost bioengineering hubs, might allow significant improvements on how bioprocess development is carried out^{37, 38}. Roughly, the task of a biofoundry is to build upon software-based pathway design and use automated DNA assembly and prototype testing for strain development. The DBTL cycle is then closed by artificial intelligence-based learning and the new results are fed-back into the pathway design. As a result, biofoundries might turn out to be a crucial piece in the transformation of scientific data (generated in single laboratories) into useful industrial bioprocesses. Besides, the heavy reliance on automated, robotic platforms and data storage in cloud-based systems could be beneficial for adherence to common standards which are otherwise not particularly helpful for lab-scale synthetic biology projects³⁹, but ultimately necessary to utilize the broad knowledge on biological systems engineering on the industrial scale.

1.4 Scope of the thesis

In this doctoral thesis methods of the synthetic biology toolbox were applied to the development of microbial strains for the production and *in situ* detection of branched chain higher alcohols. Concomitantly, an automated screening workflow based on biosensor circuits was developed.

Chapter 2 reviews the types of microbial biosensors currently used and introduces concepts for the development and fine-tuning of sensor specificities by transcription factor engineering. Those concepts are required for the construction of new biosensor systems recognizing industrially relevant compounds that are not yet available from the established biosensor repertoire.

Chapter 3 describes the directed evolution of a positive transcription factor, AlkS, for *in vivo* detection of butanols and related compounds with *Escherichia coli*. Biosensor circuits using AlkS variants were thoroughly characterized and applied for the screening of a combinatorial library for isopentanol production utilizing a robotic liquid handling platform.

Chapter 4 summarizes the development of a bioprocess for isopentanol production with *Pseudomonas putida* KT2440 and compares varied plasmid and genome based pathway expression strategies. Additionally, *in vivo* product detection in a co-culture with an adapted *E. coli* biosensor strain is demonstrated.

Chapter 5 gives an overview of synthetic sRNA expression circuits for the targeted knock down of gene expression in *P. putida* KT2440, e.g. for targeting side-product formation as found for alcohol production in chapter 4. These circuits lay the foundation for designing and implementing a genome-wide sRNA expression plasmid library.

Chapter 2: Novel Sensors for Engineering Microbiology

Maximilian Ole Bahls, Tsvetan Kardashliev, and Sven Panke

This chapter is an updated version of work published in: S.Y. Lee (ed.) Consequences of Microbial Interactions with Hydrocarbons, Oils, and Lipids: Production of Fuels and Chemicals, Handbook of Hydrocarbon and Lipid Microbiology, doi.org/10.1007/978-3-319-50436-0_387 (2017)

2.1 Abstract

The development of sustainable, biocatalytic routes to compounds otherwise derived from petrochemical processes is one of the major objectives in the field of biotechnology. Obtaining suitable microbial strains for this task still depends on the generation of strain variants and the subsequent screening or selection process. While the technical advances in DNA manipulation and synthesis allow rapid generation of millions to billions of metabolic pathway variants for a given product, the knowledge of which variants to generate and how to assess them in a high-throughput manner is lacking behind. The latter problem is increasingly tackled through the use of biosensors, by which product titers are coupled to easily detectable *in vivo* reporters such as fluorescent proteins. This in turn requires an interface where the presence of the desired product can trigger the formation of the reporter. Therefore, this chapter discusses how such an interface can be implemented and reviews the types of genetically encoded sensors available, their construction and applications, and how the specificities of future biosensors could be developed.

2.2 Introduction

Today's economy relies primarily on fossil resources and organic synthesis for its manufacturing needs. This industrial model is intrinsically unsustainable and causes significant burden on human health and the environment, alike. This has led to the concept of "Green Chemistry" which promotes the use of renewable resources and chemical technologies aiming at minimal environmental impact^{40, 41}. More often than not, such "green" manufacturing routes would employ biocatalysts in order to meet the 12 principles of sustainability⁴². These principles include the transition to an economy that relies on renewables as a source of energy and building blocks, and bioconversions as the chassis for industrial chemical production. However, the rate of development of enzyme-driven processes is largely dependent on the pace at which new biocatalysts can be developed. This is in turn to a large extent a function of the speed and accuracy at which individual genetic designs of biocatalysts are assessed. Even though the advances in systems biology of microorganisms (which represent the largest fraction of biocatalysts for industrial biotechnology) increasingly facilitate a rational design of biocatalysts⁴³ or semi-rational designs allow biasing diversification to meaningful outcomes⁴⁴, much of the biocatalyst improvement still rests on the assessment of large numbers of variants generated without exact knowledge of the underlying catalytic mechanisms and screening for the top-performing fraction⁴⁵. Such variant libraries often originate from random or semi-rational

diversity generation, such as random chemical mutagenesis or error-prone gene replication of a known, but inefficiently performing enzyme. The overwhelming fraction of the generated variants is thus comparable or inferior to the original catalyst. Therefore, one factor that limits the acceleration of development of novel biocatalysis schemes is the throughput of existing assay platforms for selecting the top-performing variants^{46, 47}. For example, common screens on solid media or in microtiter plates allow the sampling of 10^4 to 10^5 events per laboratory evolution round, i.e. several orders below the average size of a library generated using state-of-the-art genetic diversity methods^{48, 49}. The advent of microfluidic and single cell analysis techniques has enabled sampling of $>10^7$ events per hour, likely representing the highest currently available throughput for single cell analysis. As a result, such single-cell analysis tools increase the maximum throughput several orders of magnitude and narrow the gap to the library size that can be practically obtained via transformation (10^9 to 10^{10} cfu μg^{-1} DNA⁵⁰).

A key requirement for harnessing the throughput of single cell-based screens is the availability of a fluorescence signal indicative of the performance of individual catalyst variants or genetic designs, as the majority of established laser-based detection methods used in microfluidic technologies relies on the measurement of fluorescence per cell. To this end, genetically encoded biosensors (GEBs) have attracted significant attention for application in fluorescence activated cell sorting (FACS), as they provide the crucial link between catalyst variant performance and the fluorescence signal^{21, 51-53}.

Biosensors can be defined as two component analytical “devices” comprising a module for recognition of an analyte of interest, the “receptor”, and a module for signal output or transduction, the “reporter”. At least one of these modules is derived from a biological system. Antibodies, regulatory proteins or protein binding domains, enzymes, nucleic acids, and even organelles and whole cells have all been used as modules of biosensors^{54, 55}. In the context of metabolic and protein engineering, *in vivo* screening protocols are advantageous as they generally enable higher analytic throughput, particularly in comparison to multistep protocols that require cell lysis and subsequent *in vitro* testing of library members. Therefore, the focus of this chapter is on biosensors that are encoded in their entirety on a plasmid or an organism’s genome and render a functional sensing device upon transcription or translation of the genetic construct. Such biosensors provide a way to link the expression of a fluorescent protein to an intracellular target metabolite concentration while maintaining genotype-phenotype linkage as the cells remain intact throughout the screening process. Note that in principle also sensing of extracellular concentrations is

possible as long as genotype-phenotype linkage is retained, e.g. by suitable forms of compartmentalization⁵⁶. Sensing and signal transduction is achieved by the use of intracellular switches for the metabolite of interest.

Non-coding, functional RNA sequences, regulatory proteins and their cognate DNA sequences or binding proteins and binding domains can all serve to develop sensing devices, making categorization of biosensors non-trivial. Based on the molecular architecture, RNA-based biosensors can be categorized as riboswitch biosensors and RNA-biosensors (composed of RNA mimics of fluorescent proteins, e.g. RmFP⁵⁷ and protein-based biosensors can be grouped into transcription factor-based and fluorescence resonance energy transfer (FRET)-based biosensors. Another way to categorize GEBs is according to their mode of action. Here, riboswitch biosensors and transcription factor biosensors fall into the same category as they require the transcription/translation machinery of the host organism to generate the reporter molecules. FRET and RmFP biosensors comprise a second category for which sensing and signal generations takes place independent of the host's transcription/translation machinery. An overview of receptor and reporter modules, modes of action, key features and characteristics of GEBs from each sub-category is given in Figure 2-1 and recent examples of GEB design and application in metabolic and protein engineering are discussed in more detail in the sections below. At the end of the chapter, we present brief guidelines for construction and troubleshooting of the most versatile, amenable and robust category to date – that of transcription factor biosensors.

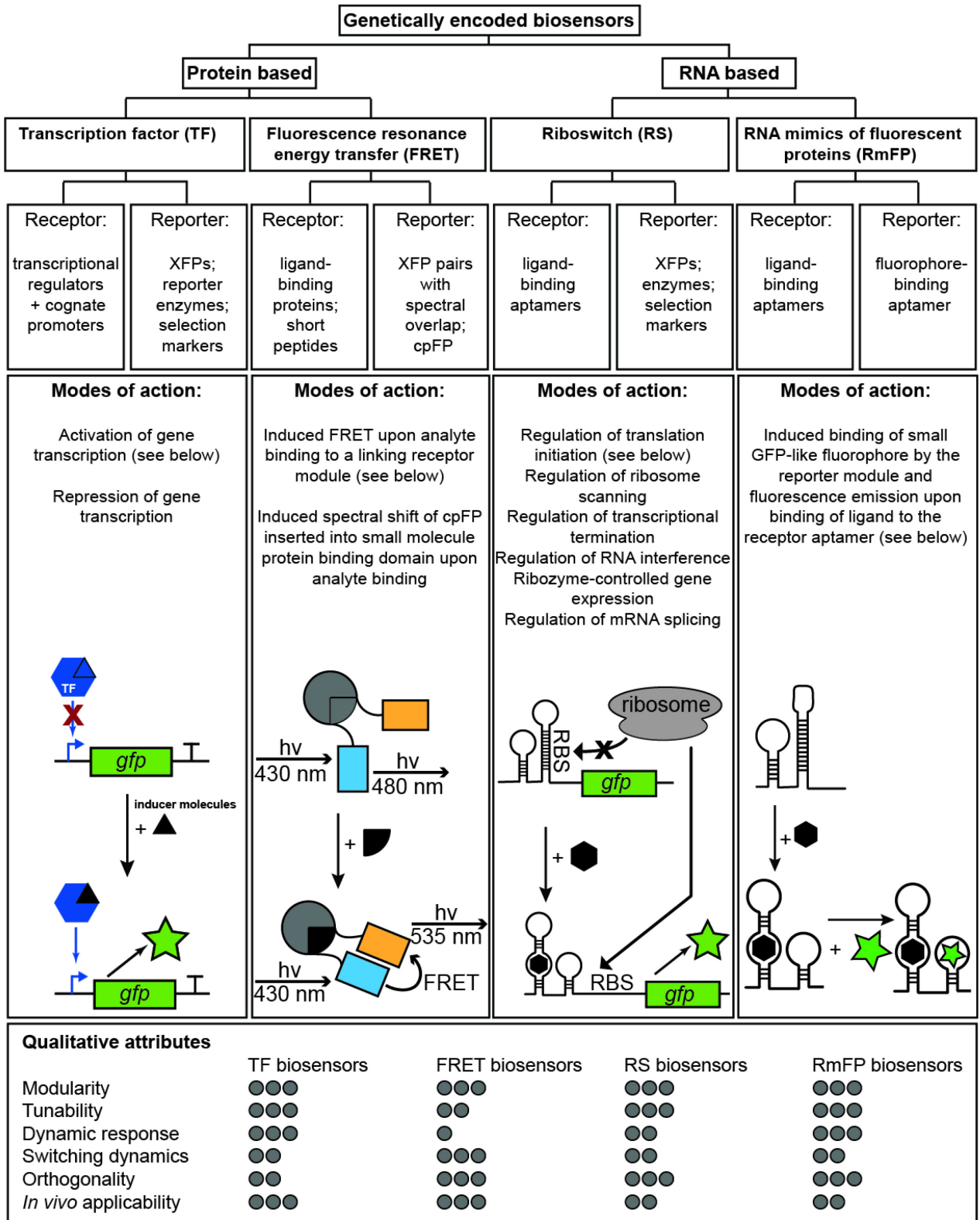


Figure 2-1 Top: Classification and overview of available genetically encoded biosensor types. Middle: Modes of action and schemes for the most common sensor types. Bottom: Ranking of the biosensor classes. cpFP, circularly permuted fluorescent protein; GFP, green fluorescent protein; RBS, ribosome binding site; XFP, fluorescent protein.

2.3 Genetically encoded RNA-based biosensors

Non-coding RNAs (ncRNAs) are abundant and functionally versatile ribonucleic acids that become increasingly important in synthetic biology and applied biology in general⁵⁸. In particular, RNAs which are capable of specifically binding small molecules, i.e. aptamers, or can serve to regulate gene expression, i.e. riboswitches, prove useful for the development of genetically encoded sensing devices^{53, 59}. Such ncRNAs do not only provide functionalities similar to those of regulatory and binding proteins but also have the advantage of comparative ease of RNA modelling, design, manipulation and evolution. In addition, RNA-based regulation of gene expression is rapid as the transcription process is completed prior to the sensing event and enables spatial and also provides temporal and dosage control over gene expression^{60, 61}. In the sections below the recent progress in the field of RNA-based biosensing is summarized including existing approaches for generating and applying such biosensing devices in monitoring of biocatalytic reactions.

Riboswitch biosensors

Riboswitches are short sequences found in the 5'-untranslated region (5'-UTR) of mRNAs that form secondary structures and thus modulate the expression of downstream sequences. Riboswitches consist of an aptamer, i.e. an RNA molecule that can specifically bind a given small molecule, and an additional regulatory component (e.g. a ribozyme, a terminator sequence, or a sequence complementary to a ribosome binding site) which can be conditionally (de)activated upon ligand binding to the appertaining aptamer. The binding of a small molecule to the aptamer leads to conformational changes in the 5'-UTR and consequently changes the state of the regulatory component.

Several mechanisms for ligand-dependent gene regulation by riboswitches exist in bacteria and eukaryotes, some of which exclusively pertain to one domain of life. One main mode of action of riboswitches in prokaryotes is co-transcriptional regulation as a consequence of ligand-dependent formation of a terminator hairpin leading to dissociation of the RNA polymerase^{59, 62}. The inverse mechanism, i.e. ligand-mediated de-stabilization of terminators and subsequent upregulation of gene expression, is also possible⁶³. Riboswitches in bacteria can also function by imposing control over translational regulation. In this case ligand binding to an aptamer upstream of the ribosome binding site (RBS) can either render the regulatory sequence inaccessible to ribosomes or, alternatively, expose the RBS, thereby upregulating gene expression⁶⁴. A less common mechanism in prokaryotes is the ligand-mediated ribozyme activation in which binding of a small molecule

to an aptamer component induces the self-cleavage of the catalytic RNA component and results in rapid mRNA degradation⁶⁵. More diverse modes of action are known in higher organisms⁶⁶.

Indeed, significant effort has been invested to not only understand the molecular mechanism of action of RNA-based regulators but to improve their performance and robustness and to widen their *in vivo* applicability^{59, 67}. The latter has proven to be challenging, particularly for riboswitches that use man-made aptamers, and to date only a handful of those have been adapted to *in vivo* applications (see below). This is counterintuitive given the availability of an *in vitro* aptamer selection technology⁶⁸⁻⁷⁰ and computational methods for modelling and re-design of aptamer specificity^{71, 72}. The difficult transferability of artificially designed riboswitches to *in vivo* setting has been attributed to thermal instability of mRNA structures and RNA misfolding under physiological conditions which has not been sufficiently closely mimicked during *in vitro* aptamer development⁶⁰.

The most successful strategy to identify artificial aptamers with acceptable *in vivo* functionality has been achieved by means of mutagenesis and recombineering of *in vitro* evolved riboswitches and activity screening within a host organism. In this way, an *in vivo* riboswitch biosensor for neomycin has been developed⁷³. A green fluorescent protein (GFP)-based *in vivo* screen for regulated aptamers starting from an *in vitro* selected aptamer pool that bound the aminoglycoside antibiotic identified an aptamer that confers neomycin-dependent control of translation initiation in yeast. It is noteworthy that the *in vivo* identified riboswitch was underrepresented in the original pool of aptamer variants with robust *in vitro* functionality.

Similarly, a riboswitch biosensor that activates protein translation in *Escherichia coli* cells in response to 2,4-dinitrotoluene (DNT) has been engineered⁶⁵. This was achieved by incorporating degenerate bases between an *in vitro* selected trinitrotoluene (TNT) aptamer and the switching component (in this case, an RBS), and carrying out *in vivo* screening that relied on inducible expression of TEV protease and a FRET-substrate to detect riboswitch-upregulated protease expression in the presence of inducer. The isolated riboswitch exhibited a 10-fold relative increase in fluorescence in the presence of DNT. The aforementioned examples strongly suggest that *in vivo* screening is probably an indispensable step when developing riboswitch biosensors for application inside living cells.

At present, most engineered riboswitches still require high effector concentrations for switching, exhibit low dynamic range and high background activity in the absence of ligand⁶⁷.

Nevertheless, steady progress in riboswitch design and innovative uses of naturally occurring aptamers has resulted in the first successful *in vivo* screening applications of riboswitches. An early example describes the use of a RNA switch based on a naturally occurring aptamer for theophylline, a molecule of the xanthine family structurally similar to caffeine, to link theophylline concentrations and GFP expression levels in yeast⁷⁴. Quantitative high-throughput screening of large enzyme libraries, either in clonal cultures or in single cells by FACS, resulted in identification of a caffeine demethylase mutant with 33-fold relative increase of activity and 22-fold improvement of selectivity. Similarly, a screening platform that employs a microfluidic static droplet array and an L-tryptophan riboswitch to analyse intracellular metabolite concentration from single microbial cells was used to isolate microbial strains with up to 145% increased productivity compared to its parental strain⁷⁵. In another study, a riboswitch based on a naturally occurring L-lysine aptamer and a selection module instead of a fluorescent reporter was used to identify aspartate kinase variants with 1.6-fold higher *in vitro* activity relative to the wild-type enzyme⁷⁶. The application of L-lysine riboswitches has been extended to isolation of optimized lysine producer strains^{77, 78}. Another report of riboswitch biosensor application in the context of metabolic engineering describes the use of the theophylline riboswitch biosensor to select strains with higher productivity of the drug methylxanthine⁷⁴.

Riboswitch biosensors based on ribozyme-type regulation have also been used in screening applications. For example, the natural *glmS* ribozyme was used in yeast to select for N-acetylglucosamine producing strains⁷⁹. Another recent example describes the development of an elaborate strategy to identify *Bacillus subtilis* strains with improved vitamin B2 productivity⁵⁶. *B. subtilis* strains that converted cellobiose to vitamin B2 were co-confined with *E. coli* sensor cells inside nL-size alginate beads. Product formation triggered a sequence of reactions in the sensor cells: (1) conversion of B2 into flavin mononucleotide (FMN), (2) binding of FMN by a natural FMN-sensitive RNA riboswitch and (3) self-cleavage of RNA resulting in (4) the synthesis of GFP. The fluorescence intensity was then used to isolate more efficient vitamin B2 producers, while the co-confinement allowed retaining the link between genotype and phenotype.

Taken together, these examples demonstrate the potential and versatility of riboswitches as devices for *in vivo* screening. Given that riboswitch biosensors are still in their infancy, we anticipate that their importance to high-throughput screening will only increase in the years to come.

RNA mimics of fluorescent protein biosensors

An emerging category of genetically encoded sensors makes use of RNA aptamers that bind freely diffusible fluorophores and switch them to a highly fluorescent state⁸⁰. These RmFP aptamers were originally used for tagging and imaging specific RNAs in living cells^{81, 82}. An emerging application of RmFPs is the sensing of intracellular metabolite concentrations. The fluorescent RNA aptamer-fluorophore complex can be converted into a sensor that emits a specific signal only in the presence of a small molecule inducer⁸³. This is achieved by fusing together a fluorophore binding aptamer to a ligand-binding one in a way that only in the presence of a ligand the secondary structure of the fluorophore binding aptamer is correctly assembled, the small fluorophore can be bound and switched to a highly fluorescent state. Sensors that employ the fluorophore binding aptamer “Spinach” (or derivatives thereof) as a reporter module and a different small molecule binding aptamer as receptor module have been built by inserting the latter into a structurally critical stem of the Spinach RNA^{84, 85}. The target-binding aptamer is unstructured in the absence of the target molecule as a critical stem is disrupted thus preventing Spinach from folding and binding the fluorophore. However, when the aptamer binds its target, the correct folding of the critical stem in Spinach leads to fluorescence that can be detected both *in vitro* and in living cells. Sensors that bind S-adenosyl-methionine (SAM), ADP, and other metabolites have been created using this approach⁸⁶. Moreover, these sensors have enabled imaging of the dynamics and turnover of SAM and ADP, cyclic di-GMP and various proteins in living cells⁸⁵. A recent report demonstrates for the first time that *in vivo* detection of enzyme activity is also possible with RmFP biosensors. To this end, the authors used an optimized RmFP sensor for S-adenosyl-L-homocysteine (SAH) to measure methylthioadenosine nucleosidase (MTAN) activity in live *E. coli*, more precisely the increase of SAH levels upon chemical inhibition of MTAN⁸⁷.

While we are yet to witness a “true” screening application, the potential for high-throughput enzymatic assays of these innovative genetically encoded sensors has been implied by the recent developments in the field. Nevertheless, it is noteworthy to mention that technical difficulties similar to those experienced with artificial riboswitch biosensors and FRET sensors (discussed below) can be anticipated along the way to robust *in vivo* RmFP biosensors capable of sensing arrays of chemically diverse compounds. In addition, due to their mode of action, RmFP biosensors appear to be more suited for sensing metabolite dynamics and their use in quantitative screening and identification of microbial producers will likely remain limited.

2.4 Genetically encoded protein-based biosensors

FRET biosensors

FRET-based sensors typically involve a pair of donor and acceptor fluorophores linked by a ligand-binding protein domain such that upon ligand binding a conformational rearrangement is induced causing an alteration in the proximity of the donor and acceptor fluorophores and consequently a measurable FRET change⁸⁸. In the FRET category we also included sensors that are based on single, circularly permuted fluorescent protein (cpFP) fused to binding proteins in a way that ligand binding induces detectable changes in either fluorescence intensity or excitation and emission profiles⁸⁹. A characteristic feature of both of these protein sensor designs is that no transcription or translation event is required after sensing and that the exerted signal is generally reversible on shorter time scales. Despite advantages such as orthogonality to existing metabolic processes of the host organism, high temporal resolution, and relative ease of construction, FRET sensors often suffer from low dynamic ranges and are less suited for monitoring of metabolite accumulation and predominantly are applied to monitoring of intracellular metabolite dynamics, rather than screening for producer strains^{52, 90}.

There is one screening application of a cpFP-based biosensor for hydrogen peroxide, HyPer, available in a novel screen to engineer enzymes for enhanced production of H₂O₂. Cytochrome P450 BM3 variants were expressed in a biosensor strain and, using HyPer's ratiometric signal, variants that generated greater amounts of H₂O₂ than the wild-type enzyme via uncoupling were reported⁹¹.

Transcription factor biosensors

In vivo genetic circuits are broadly reliant on transcription factors as regulatory proteins for controlled protein production⁹² that allow the on-demand activation or repression of gene transcription. However, in contrast to the RNA-based sensors, the sensing process requires the transcription/translation machinery of the cell for reporter activation. For the design of biosensor circuits, the transcription factor of interest is usually constitutively expressed by the cell and upon binding of an inducer molecule, the activated regulator is recruited to its operator in the cognate promoter region, thus activating gene expression (e.g. MalT and maltose⁹³), for example of a reporter. Alternatively, the regulator can be a repressing (as opposed to activating) transcription factor. Here, the transcription factor blocks expression from its cognate promoter, until this repression is relieved by the addition of an inducer molecule (e.g. LacI and allolactose⁹⁴). In addition, the repression can also be activated upon

binding of a specific ‘inducer’ (e.g. MetJ and S-adenosylmethionine⁹⁵). However, if such a system should be used for sensing of an improved product titer due to increased reporter expression, it requires the inversion of the signal from repression to activation, thus making the biosensor circuit more complex to design and implement.

Many regulator-promoter pairs are known from previous studies which facilitates the *a priori* design of a circuit with specificity for the inducer molecule of interest (Figure 2-2a). Even when not fully annotated, regulators can be found in metagenomic libraries for novel but natural target inducers, but also for xenobiotics only recently introduced to nature⁹⁶.

Still, this set of pairs is markedly limited in terms of known compounds for which we can find suitable parts. However, the toolbox of synthetic biology and molecular biology workflows in general give increasingly access to engineered, bespoke regulator specificities as shown in Figure 2-2^{30, 97}. The methods applied for the generation of tailor-made regulators include directed evolution (comprising rounds of *in vitro* mutagenic regulator gene replication followed by *in vivo* selection of improved variants)⁹⁸ but also extensive *in silico* modelling in order to predict regulator sites that correspond to a broadened or novel specificity⁹⁹. Here, a broadened specificity means that the original inducers are still working as inputs, while a

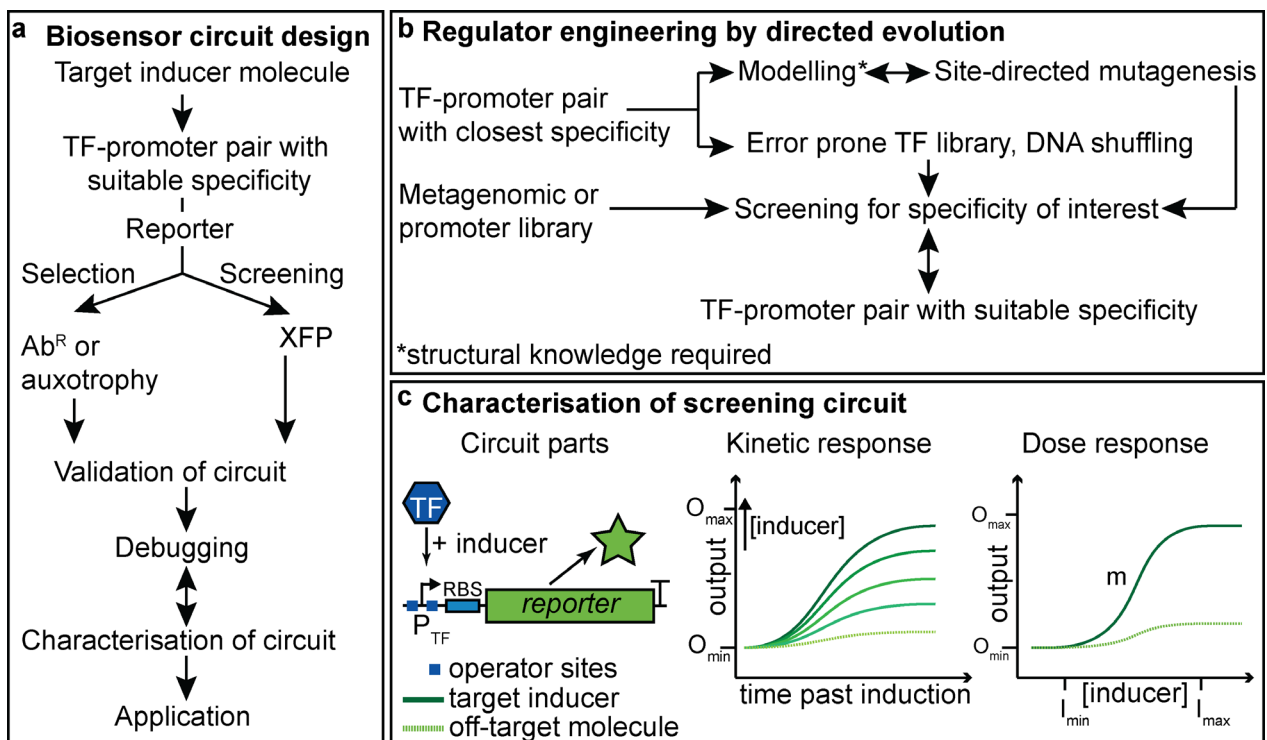


Figure 2-2 a) Biosensor design from target molecule definition to application. b) Regulator engineering in order to obtain suitable target specificity. c) Characterisation of the biosensor circuits in terms of parts, kinetic-, and dose-response. Ab^R, antibiotic resistance; dynamic range of reporter output, $O_{max}-O_{min}$; I_{min} , detection threshold; I_{max} , maximum detection; m , sensitivity; TF, transcription factor XFP, fluorescent protein.

novel specificity would require a switch to the novel inducer including a concomitant loss of function for the original inducer. So far, the model-based strategies are still more an envisioned scenario than an everyday reality in routine lab work, but have a huge potential for simplifying the design of novel biosensors as discussed below.

Common to all circuits, independent of the input source, is an easily accessible, genetically encoded reporter. In general, two options are available: selection and screening strategies. In the case of selection, common reporters driven by the sensor circuit are antibiotic resistance proteins, proteins that complement auxotrophies, or proteins that implement a conditional phenotype based on the conversion of a substrate into a toxic product. All these reporters readily allow the selection of cells in the on-state, i.e. the sensor activated state, as resistant or cured cells in cultures outgrow other cells lacking the induced gene expression, and in the optimal case, exclusively survive under restrictive media conditions. We also include reporter proteins that convert a chromogenic substrate into this category, as identifying activated cells is still straightforward, e.g. by selecting blue colonies due to the expression of β -galactosidase (*lacZ*, e.g.^{100, 101}) on media including the corresponding colourless chromogenic substrate X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside), or by selecting coloured colonies producing pigments that do not require any particular precursor molecule or substrate¹⁰². However, with the colorimetric assays the enrichment of on-state cells is lost.

While the described selection systems are generally helpful if a rather digital on-or-off, switch-like information is sufficient or preferred, in metabolic engineering it is often of more interest to gradually increase a product titer over multiple rounds of improvement. In other words, there is no zero concentration off-state. Here, screening for improved variants has its strength, as it allows a distinct biosensor response depending on the accumulation of the target compound. Traditionally, transcription factor-based systems utilized various reporters including luciferase genes (*luxAB*) or chromogenic substrate conversion via *lacZ* expression, but were almost exclusively replaced by fluorescent proteins (XFP) as also apparent from the previous RNA and FRET-based sensor examples. XFPs have the advantage of being almost self-sufficient reporters, meaning no substrate has to be added to the culture, as only oxygen is required for successful maturation of the chromophore. Crucially, fluorescent proteins have the essential feature of allowing high-throughput screens via FACS with a detection limit principally similar to the one observed for (catalytic) luciferase based assays, which are not routinely implemented at the single cell level¹⁰³. Besides, transcription factor based circuits using XFP variants with reasonably short half-

lives allow the *in vivo* assessment of molecular fluxes over time¹⁰⁴. In particular for engineering of sensor specificities, schemes combining both selection and screening are highly useful as unwanted specificities and constitutive variants can be excluded by a round of negative pre-selection, followed by screening for increasing fluorescence with increasing product concentrations.

For more details and explicit tables of the known transcription factors and their cognate promoters, please refer to one of the numerous excellent reviews in the field (amongst others:^{16, 21, 30, 31, 97, 98, 105-109}). Here, we highlight the origin and construction of various transcription factor based circuits available today and focus on how they were designed, including the engineering of novel transcription factor specificities and applications for screening in metabolic engineering. We conclude with a short section about circuit design and sensor debugging in order to facilitate the development and optimization of future biosensor circuits.

Circuits based on natural transcription factor specificities

For many operons, often belonging to well-studied catabolic and stress-response pathways, the regulatory elements are known. They evolved mainly for the recognition and utilization of alternative nutrient sources under harsh environmental conditions lacking common nutrients, for example for switching to hydrocarbons as energy and carbon source when sugars such as glucose are absent. Biosensor-relevant regulatory elements are also often involved in the sensing of cellular damage, e.g. due to solvent exposure, or of the intracellular redox state. For several decades, such regulators were coupled to reporters including β -galactosidase and luciferases in order to build sensor circuits for the detection of environmental pollutants¹¹⁰.

For multiple biomonitoring tasks regulatory genes with their cognate promoters, in many instances originating from catabolic plasmids of *Pseudomonas putida*, were utilized. An exemplary set of such biosensors was built and characterized in *E. coli* and *P. putida* KT2442 by de Lorenzo and co-workers, including regulators for alkyl- and halobenzoates (based on XylS), alkyl- and halotoluenes (XylR) and salicylates (NahR), coupled to the expression of *luxAB* and *lacZ*¹¹⁰. For the circuit utilizing the transcription factor XylS and its cognate promoter P_m fused to *luxAB*, the minimum detection level of extracellular *m*-toluate was found to be as low as 1 ppm (about 5 to 10 μ M). The same set of regulators was also used as a basis for studies creating novel inducer spectra by engineering the regulator residues, which is discussed in the next section.

NahR was also utilized as a gas phase sensor for naphthalene vapour, by activating luciferase expression from its cognate promoter P_{sal} in *P. putida*¹¹¹. Here the detection limit was found to be in the lower nM range of naphthalene. This study additionally highlights a potential source of error when exploiting circuit parts originating from bio-degradative pathways. Usually such pathways are not highly expressed as long as the preferred carbon sources of the cell are available, hence the activities of the corresponding regulators are down-regulated (catabolite repression). In this study, the addition of 10 μM succinate to the culture medium lead to a 20% decrease in signal output, while at concentrations above 100 μM almost no significant sensor output was detectable.

Similar biosensors were developed for various molecules that are characteristic for hydrocarbon spills and contaminated groundwater samples by van der Meer, Rojo and others^{112, 113}. For instance, various whole-cell biosensors were built on the basis of AlkS, a regulator of an alkane-responsive system in *P. putida* GPo1. The reporters included *gfp* and *luxAB*, with the latter system allowing octane sensing in the nM range¹¹⁴. Additional sensor circuits for the detection of external toxic compounds, such as antibiotics and halogenated aromatics, were for instance based on the regulators TtgR¹¹⁵ and TodS-TodT¹¹⁶ of *P. putida*, respectively.

Intriguingly, many of the sensor specificities described above were published for the detection of spilled xenobiotics, but now might be of high value in future metabolic engineering approaches for the sustainable bio-production of the very same compound classes. Nevertheless, the availability of sensor types for hydrocarbons is far from complete and in particular imperfect for recognizing particular molecular substitution patterns. Additionally, the sensor circuits described were mainly applied for sensing of extracellular molecules, which requires the molecule to be able to translocate across the cell membrane. However, this uptake issue exists in the case of externally added target compounds, but not if the compound is produced intracellularly, for instance in case of the production of medium chain length alkanes^{117, 118}. The expression of uptake systems can be used to lower the detection threshold (I_{min}) for molecules that not efficiently diffuse across the cell membrane. For example, in *E. coli* the heterologous expression of a protocatechuic acid (PCA) permease, *pcaK* of *P. putida*, allowed a 1500-fold lower sensitivity for detecting extracellular PCA with a biosensor circuit based on transcription factor PcaV with $P_{\text{PV}}::\text{sfgfp}$ as biosensor output¹¹⁹.

More recently, the interest in biosensors extended towards applications in metabolic engineering workflows. Here, the sensors are used for the high-throughput *in vivo* screening

of enzyme libraries, i.e. the assessment of millions of similar but slightly different microbe variants producing a target compound. This is in particular valuable for relatively inconspicuous small molecules including amino acids, sugars and various hydrocarbon products such as alkanes or alcohols, which are difficult to detect in a high-throughput fashion with standard chemical analytics.

A primary example for the development of such novel biosensors for FACS based high-throughput screening of metabolites is the development of a LysG-based circuit for amino acids in *Corynebacterium glutamicum*. Here, Eggeling and co-workers fused the cognate promoter region of the transcription factor LysG to the gene for a yellow fluorescent protein (*eyfp*) in order to successfully screen for L-lysine overproducers with cytosolic concentrations in the lower mM range¹²⁰. Similar systems were applied for other amino acids, including L-arginine and L-histidine^{121, 122}.

For the high-throughput assessment of medium chain length alcohols, a biosensor was built in *E. coli* by utilizing BmoR of *Thauera butanivorans* controlling the expression of a *gfp* gene from its cognate promoter. This circuit allowed a linear detection range for butanol from 0.01 to 40 mM. BmoR-based biosensors were applied to proof-of-principle screens for improved n-butanol¹²³ and isobutanol production¹²⁴. Another principle was used by Cheng et al., relying on the competition between a repressor (ArgR), which is activated by arginine and thus repressing GFP expression from its cognate promoter, and the enzyme arginine deaminase (ADI), which converts arginine into citrulline. Here, a more active ADI variant would lead to an increase in fluorescent reporter signal. This was utilized in a FACS-based screen for almost 10^7 ADI variants created by error-prone PCR (epPCR), yielding an improved ADI variant with both higher activity and a lower K_M ¹²⁵. More general sensing approaches include the development of a sensor for NADPH consumption, utilizing a [2Fe-2S]-cluster containing regulator, SoxR, and eYFP under the control of the SoxR-cognate promoter in *E. coli*¹²⁶. In this study, NADPH-dependent alcohol dehydrogenase variant libraries were screened for improved activity, generally pointing to the feasibility of high-throughput screens for various other NADPH-dependent enzymes.

Besides, much data about regulator specificities and pathways that are regulated by molecules of interest is available in online databases. For instance, RegPrecise offers a manually curated database of prokaryotic regulons¹²⁷, Bionemo offers a search function for bio-degradative pathways by molecules of interest, explicitly stating the involved transcription factors when available¹²⁸, and more general pathway databases, e.g. the EAWAG Biocatalysis/Biodegradation Database (<http://eawag-bbd.ethz.ch>, including the

former University of Minnesota Biodegradation/Biocatalysis Database and Pathway Prediction System¹²⁹), allow to search for substrates and the related enzymes, which may indicate where to look for novel regulatory regions of interest.

When no suitable sensor protein is available, metagenomic and promoter libraries were successfully screened for suitable regulator specificities. For example, the Alon promoter library, including several thousand promoter-*gfp* fusions¹³⁰, was mined for a phenylalanine responsive regulator system, which was then applied to a phenylalanine-overproduction screen in *E. coli*¹³¹. Additionally, substrate-induced gene-expression screening (SIGREX) allows the mining of genes from any environmental metagenome by cloning the fragmented DNA upstream of *gfp*, followed by screening for input molecule specific fluorescence^{132, 133}.

While many useful constructs were built with known or mined natural transcription factors, the development of novel sensor systems is still markedly limited by the availability of suitable regulator specificities. One workaround is the *in vivo* conversion of analytes after their formation into compounds for which sensors exist⁹⁷, facilitated by software tools predicting the necessary enzymes for a given analyte¹³⁴. For instance, cocaine was sensed after conversion into benzoic acid (BenR regulator¹³⁵) and 3-hydroxy propionate was sensed after conversion to acrylate (AcuR¹⁰⁴). Concluding, the design of novel biosensor circuits based on transcription factors is a highly promising but ongoing task that has the potential of improving the *in vivo* screening for a great number of novel whole-cell catalysts, if the required specificities are found or engineered.

Circuits Based on Engineered Transcription Factor Specificities

In order to change the specificities of known regulators, the DNA sequence coding for the regulator can be varied and the resulting variants sampled for an altered inducer spectrum (inducers 'a la carte'⁹⁸). This strategy could either follow a random approach, in particular if no information about potentially valuable protein sites, such as the binding pocket, is available, or a (semi-)rational approach targeting particularly interesting residues as determined from crystal structures and homology models.

Studies following the random approach are mainly based on the error-prone *in vitro* regulator-gene replication via PCR, followed by rounds of *in vivo* screening for the novel specificities of interest. This approach frequently yielded regulator variants with the desired novel inducer spectra and altered affinities, in spite of the tremendous amino acid residue space that one can search through. Besides, such experiments provided valuable

information about the location of potential protein sites involved in the inducer and DNA binding.

Classical examples for this strategy include work by Ramos et al., analysing mutants of the XylS regulator originating from the TOL catabolic plasmid of *P. putida* mt-2. Mutants were selected by coupling the regulators to a gene for tetracycline resistance via its cognate P_m promoter, followed by culturing on media containing tetracycline and benzoate analogues which are not natural XylS effectors. The spontaneous mutation rate was increased by adding ethyl methanesulphonate, a mutagenic compound, to the media. As a result, clones were obtained with either constitutive or novel and inducer specific XylS functionality, which were then further analysed for their inducer spectrum with *lacZ* as the reporter^{136, 137}. This enabled the discovery of multiple novel regulator specificities for various benzoate analogues in the low mM range. More recently, variants of XylS and XylS-BenR fusions were constructed by site directed mutagenesis that allowed induction with additional molecules such as acetylsalicylic acid (aspirin)¹³⁸.

Several mutant versions of NahR and XylR were generated by epPCR by de Lorenzo and co-workers. In the first case, new-to-nature specificities were found by fusing P_{sal} , the cognate promoter of NahR, to *lacZ*, followed by selection depending on the blue colour of cells cultivated on media containing X-Gal and benzoate, a non-natural inducer of NahR. Upon specificity verification, about two-thirds of the variants were indeed responsive to externally added benzoate in the low mM range, while the other fraction of regulators consisted of constitutively active variants. Interestingly, and as seen in various studies before, none of the novel NahR regulator variants lost its responsiveness to the natural inducer salicylate¹⁰¹. Similar observations were made when variants of XylR for the xenobiotic compound 2,4-dinitrotoluene were generated by epPCR. These variants were selected by fusing the cognate promoter P_u , regulating the upper TOL operon in the presence of *m*-xylene, to (1) a kanamycin resistance gene, (2) *pyrF* in an uracil auxotroph *P. putida* strain (both for selection), (3) *gfp* (for screening), and finally (4) *lacZ* for verification of the induction behaviour. Remarkably, none of the altered amino acid residues were found to form the binding pocket itself. It was reasoned that the residues exchanged are involved in conformational changes upon inducer binding and thus change the signal transmission between different domains of the regulator protein. Moreover, the appealing concept of a multipotent stem form¹³⁹ was introduced to transcription factor engineering, according to which novel specificities arise as a result of the regulator adopting a more promiscuous form with a widened inducer spectrum. This stem form allows the sensing of various novel inducer

molecules, but still includes the natural one(s). Building on such promiscuous regulator variants, ripening towards new specificities could take place via precise, stringent residue exchanges¹⁴⁰.

Another strategy for the creation of new-to-nature regulator variants is the shuffling of DNA fragments generated by digestion of the coding sequences of closely related regulators, which nevertheless differ in their inducer spectrum. For instance, the N-terminal sequence of XylR, being responsible for effector binding, was shuffled with homologous sequences of the regulators DmpR and TbuT. The resulting variants were selected for various novel inducers including bulkier molecules like biphenyls or chemically altered side chains, e.g. nitrotoluenes. The novel regulators were selected by coupling DmpR's cognate promoter P_o to a kanamycin resistance gene, allowing cells cultured with a given inducer and kanamycin to survive only if they harbour a corresponding XylR variant. In order to exclude false positive, constitutive regulator variants, P_o was additionally fused to *sacB*. Expression of *sacB* resulted in a non-viable phenotype for constitutive XylR variants if the whole-cell biosensors are cultivated with sucrose but without inducer¹⁴¹.

The combination of epPCR with subsequent DNA shuffling of the variant hits showed intriguing results as well. Leadbetter and co-workers created novel biosensors based on LuxR, originating from the quorum sensing system of *Vibrio fischeri*, coupled to *gfp* expression via its cognate promoter P_{luxI} . The natural inducer spectrum of LuxR does not include butanoyl homoserine lactone (C4HSL) having a shorter acyl-side chain length than the natural inducers. By rounds of directed evolution via error-prone replication of the regulator, LuxR variants were found that responded with half-maximal activity at a C4HSL concentration of 2.3 μ M, while the wild-type regulator showed no response. The variants created were then subjected to another round of evolution by DNA shuffling, yielding further improved LuxR variants that allowed a half-maximal activity already at 150 nM of added C4HSL¹⁴².

More recently, regulators for biosensors were engineered at specific amino acid sites, which were pre-selected due to their proximity to the inducer binding sites as judged from crystal structures. Here, the regulator variants are usually created via PCR with synthetic, mutagenic DNA oligomers corresponding to the chosen amino acid residues. Cirino and co-workers applied this strategy successfully to the AraC regulator protein, originating from the arabinose operon in *E. coli*. Initially, AraC mutants were created by both epPCR and targeted mutagenesis at four residues of the effector-binding pocket in order to find a molecular reporter specific for D-arabinose, as opposed to the natural inducer L-arabinose.

As a reporter of the biosensor circuit, *gfp* was installed under the control of the cognate promoter P_{BAD} , allowing FACS based screening of the AraC mutant libraries. In order to find suitable novel regulators, a dual screening workflow was used. Every round of positive screening with D-arabinose as the target inducer was followed by a round of deselecting cells that were either constitutively fluorescent or fluorescent in the presence of added L-arabinose. This allowed to create AraC mutants responsive to D-arabinose in the range of 10 nM to 1 mM while L-arabinose concentrations of around 10 mM or higher were required for a recognizable fluorescent output. Interestingly, all characterized AraC variants were found by the targeted approach, while the random, error-prone libraries yielded no highly improved variants¹⁴³. With a similar dual screening strategy, an AraC library, created by simultaneously mutagenizing five residues of the binding pocket, was used to generate a biosensor for mevalonate with a concentration range of 10 to 200 mM of mevalonate added to the culture. Mevalonate is a valuable input molecule for isoprenoid synthesis in *E. coli*, and its production was significantly improved by selecting producer variants with optimized expression levels of a reductase of the corresponding mevalonate pathway with the help of the biosensor system¹⁰⁰. Moreover, AraC based sensors were further developed for triacetic acid lactone (TAL) at low mM concentrations, here with *lacZ* as the reporter. This systems was applied to the selection of an improved 2-pyrone synthase variants converting malonyl-CoA more efficiently into TAL, resulting in variant with about nineteen-fold improvement in k_{cat}/K_M and a twenty-fold improvement of the TAL titer¹⁴⁴.

A more rational engineering strategy is represented by the application of computational design to generate novel specificities, as exemplified by engineering a novel specificity for explosives, including TNT amongst others, into ribose binding protein¹⁴⁵. However, additional experiments indicated that the specific concept applied in that work is not reliably working¹⁴⁶. More recently, Baker et al. computationally designed and successfully expressed proteins in order to specifically bind ligands, namely the steroid digoxigenin¹⁴⁷, with high affinity and selectivity, pointing towards an increasing practicability of *in silico* design of novel specificities for biosensors so far not accessible. The Rosetta modelling protocol¹⁴⁸ was utilized for the creation of model-based, focused variant libraries of PobR, a transcription factor of *Acinetobacter* lacking an experimentally determined structure, for the novel inducer 3,4-dihydroxy benzoate⁹⁹. After FACS-based screening with a sensor circuit consisting of *gfp* under control of the cognate promoter of PobR, variants with activity for 3,4-dihydroxy benzoate in the μM range were isolated. The PobR variants also showed an

increased response with the natural inducer 4-hydroxy benzoate, but no sensitivity for other structurally similar molecules.

In a comprehensive study of novel LacI variants, a repressor from the *lac* operon of *E. coli*⁹⁴, multiple of the aforementioned engineering strategies were utilized¹⁴⁹. This included computational design, targeted residue saturation mutagenesis almost spanning the full protein length, and random epPCR mutagenesis. The LacI variants were screened for induction with four novel inducers, all being saccharides with varying degrees of structural similarity to the natural inducer allolactose. In the first round, non-repressed designs were removed by coupling LacI to a fusion promoter P_{LacO} driving the expression of a porin (TolC). In the presence of colin E1, a toxin, the TolC expression is toxic and thus eliminated undesired LacI variants. Next, FACS-based screens were employed in order to isolate variants induced by the target compounds, facilitated by fusing P_{LacO} to *gfp*. For all compounds, novel LacI variants were found, showing sensor outputs similar to that of the wild-type LacI circuit. Noteworthy, for one target compound, sucralose, the random approach failed to deliver a novel LacI variant, while the successfully computationally designed variant contained four mutations. It was argued that the combinatorial space might have been too large for the epPCR strategy. The specificity of sensor variants with broadened specificity towards gentiobiose and sucralose was further improved by shuffling and combining beneficial mutations followed by additional rounds of FACS screening. This allowed increasing the maximum output significantly while the induction by isopropyl β -D-1-thiogalactopyranoside (a non-metabolisable mimic of allolactose) was drastically reduced and in the case of the gentiobiose-responsive variant almost completely eliminated. Intriguingly, this activity maturation of the intermediate LacI variants for gentiobiose and sucralose is in agreement with the above discussed stem protein intermediate¹⁴⁰, which likely might be required in order to find switched specificities, as opposed to broadened specificities.

2.5 Optimization of the biosensor's characteristics

Besides engineering the structure of the transcription factor and its physical interaction with the inducer molecule of interest, almost all other parts of the biosensor circuit are amenable to optimization for improved specificity and overall circuit behaviour. As a result, debugging of unfavourable sensor features is possible and often necessary. The main characteristics of a whole-cell biosensor are depicted by a dose-response curve (also see Figure 2-2c), indicating how the circuit's output changes with varied input concentrations – ultimately

determining whether a sensor is useful for finding improved microbial producer strains. The dose-response curve shows the fold-change between basal output (subject to undesired activity in the off-state known as “leakage”) and the maximum obtainable output which is also known as the dynamic range, the steepness or slope of the response indicating the sensor sensitivity, and the minimum to maximum inducer concentrations between which meaningful changes in output are observable. In order to evaluate the suitability of the sensor circuit for high-throughput screening assays, in terms of reliability and false-positive rates, a dimensionless Z-factor can be calculated by relating the dynamic range to the data variation¹⁵⁰). Also the response kinetics of the system provides valuable information in order to find the optimal time point for FACS-based screening. In addition, the output kinetics determine whether a given response is short-lived enough to be a useful measure for *in vivo* flux-analysis, which could be tweaked with mRNA- or protein-tags for increased degradation rates. This could be necessary, for instance, when the formation and degradation of intermediates of a production pathway should be detected. If the cellular half-life of the sensor output is significantly longer than that of the transient formation of the input of interest, the signal output would be merely static instead of showing the actual metabolic flux.

The theoretical framework for sensor circuit design and its fine-tuning is significantly expanding with the progress of synthetic biology (for example, see¹⁵¹⁻¹⁵³), but most of the experimental implementations still require substantial design-build-test-learn cycles. Nevertheless, recent studies highlight how the debugging of imperfect sensor systems could be readily achieved with straightforward methods, as soon as the regulator of interest and the reporter are defined.

For an ArsR-based heavy metal biosensor, it was shown that the addition of a second ArsR binding site significantly decreased the leakage of the circuit, as it most likely increased the probability of interaction of the repressor with the RNA polymerase or simply sterically blocked it¹⁵⁴. Adding a binding site downstream of the promoter sequence (serving as a “roadblock”) further decreased the basal expression in this study. Overall, this strategy allowed lowering the basal output while maintaining the inducible control and crucially a high maximum output, thus increasing the fold-change in output upon induction by a factor of approximately five. For a sensor system based on XylR, naturally responding to *m*-xylene and to a lesser extent also to 3-methylbenzyl alcohol, it was shown that rewiring of the circuit parts has an effect on the specificity without changing the transcription factor itself¹⁵⁵. By introducing a positive feedback loop utilizing the P_s promoter for XylR expression and thus

replacing a P_R promoter (negative feedback loop) the sensor became more specific for *m*-xylene. However, a similar positive feedback system with the stronger P_U promoter replacing P_R showed increased output for both *m*-xylene and 3-methylbenzoate but without improving the accuracy of discrimination. This suggests that the combination of the expression of an attenuated regulator with a positive feedback loop can adjust the activation threshold in a way that allows discriminating against the weaker inducer molecule 3-methylbenzyl alcohol. In addition, both the level of regulator expression and the physical localization of expression were shown to influence the biosensor features¹⁵⁶: by utilizing a sensor system consisting of the activator XylS and the cognate promoter P_m driving *gfp* expression, it was shown that both high activator concentration and physical proximity of regulator and target promoter reduce the noise level and thus increase the obtainable dynamic range of the system in *P. putida*. The authors reasoned, supported by modelling, that this observation is caused by little to no expression of the reporter *gfp* as soon as the local concentration of inducer bound to XylS randomly becomes too low to allow binding to the promoter region, thus leading to a large spread of varied GFP concentrations in an induced population of genetically identical cells¹⁵⁶.

Several studies underline how the exchange of standardized circuit parts facilitates the fine-tuning of sensor systems. For instance, a sensor based on DmpR was built for the screening of enzyme libraries for the production of phenol-derived compounds¹⁵⁷. Here the dynamic range was drastically improved by first replacing the existing RBS with an optimized RBS sequence, followed by the strict employment of standard transcriptional terminator sequences. In a similar manner, a set of biosensors for aromatic compounds was optimized¹⁵⁸. While several set ups worked without further requirement for optimization, the unintentional transcription in the off-state of two sensors based on the transcriptional activators XylS and HbpR (the latter originating from the 2-hydroxybiphenyl pathway of *Pseudomonas azelaica*) prohibited a useful dynamic range. Thus it was reasoned that the activators were expressed at too high a concentration, or the RBS strength upstream of the reporter (*gfp*) was too high. In order to debug the system, first four weaker promoters were deployed for the expression of the regulators. This already led to a more suitable dose-response curve for the XylS-based circuit. The HbpR system was further improved by testing several weaker RBS upstream of the reporter gene, which also led to a significant improvement of the fold-change between the on- and off-state. Besides, this is a good showcase for how the availability of tested, standardized parts is facilitating the optimization of sensor-circuits (e.g. Registry of Standard Biological Parts, <http://parts.igem.org/Catalog>).

The suitability of RBS engineering for output optimization was further highlighted by the construction of a *p*-coumaric acid biosensor in *E. coli*, based on the repressor PadR of *B. subtilis* and its cognate promoter P_{PadC} driving *yfp* expression¹⁵⁹. Here, two initially chosen RBSs upstream of PadR were either too weak or too strong in order to allow a useful dynamic range. While the very weak RBS yielded a sensor with high basal output fluorescence (little repressor available), the second RBS tested was so strong that with the addition of 1 mM *p*-coumaric acid the fluorescence did not increase significantly (too much repressor available). However, by utilizing random RBS mutagenesis several working sensor circuits were built with RBSs of intermediate strength, allowing up to a 130-fold change in output fluorescence upon addition of *p*-coumaric acid in the low mM concentration range. Further, an RBS library with about 7000 variants was constructed and analysed by FACS for optimizing the expression strength of both the transcription factor (CdaR) and sensor output (sfGFP) for detection of glucarate. The corresponding data set was subsequently used for deep learning resulting in a model for predicting the dynamic range of the biosensor based on the RBS sequences¹⁶⁰.

Another common issue preventing the application of sensor systems is activation of the sensor circuit by off-target inducers. This can become a severe hindering factor, for example when precursor molecules in the metabolic pathways of interest are similar enough to the product to unwittingly activate the sensor. In order to optimize the inducer spectrum, negative rounds of screening in the presence of the unwanted inducer (if it passes the membrane) can be applied. Here, all variants showing high basal fluorescence, either due to constitutive activation or induction with off-target molecules, are discarded. As mentioned before, a convenient strategy is to apply stringent negative selection first, in order to remove all variants showing off-target effects, due to a toxic phenotype. Then the workflow is continued with screens for the novel specificity of interest. This usually includes iterative rounds of screening for the cells exhibiting the highest fluorescence, as enrichment rounds are subject to cellular noise and thus include many false positive fluorescent cells if on- and off-state are not clear-cut¹⁵⁹. An exemplary system for this dual selection/screening strategy was constructed for the engineering of a choline-inducible and -repressible transcription system, based on the transcriptional repressor BetI of *E. coli*¹⁶¹. The BetI variants were assessed by regulating the expression of a *sfGFP* gene but also of the genes for a herpes simplex virus thymidine kinase (*hsvtk*) and an aminoglycoside-(3')-phosphotransferase (*aph*). While sfGFP served as a gradual indicator of the reporter-output, the latter reporters allowed negative and positive selection, respectively. In the presence of an artificial

nucleoside the expression of hsvTK is toxic, while the expression of APH is required for resistance against kanamycin. Using this sensor circuit under culture condition with and without externally added choline yielded BetI variants that were used for both choline-inducible and choline-repressive promoter systems. Following a positive-negative enrichment FACS strategy allowed to find an L-lysine insensitive variant of the LysG-based biosensor, normally detecting all three natural, basic amino acids. To this end, a LysG library was constructed by structure-guided multi-site-directed saturation mutagenesis and screened positively (high fluorescence) for L-histidine and negatively for L-lysine (low fluorescence)¹⁶². The discussed options and exemplary references for further details on biosensor engineering and debugging are briefly summarized in Table 2-1.

Table 2-1 Optimization and debugging of common issues of transcription factor based genetic biosensor-circuits in microbial hosts for *in vivo* high-throughput screening.

Issue	Possible Solutions	Reference
Sensor specificity		
	iterative rounds of positive screening of variants	161
	negative selection against unwanted inducers or side products of metabolic pathway	143, 162
	DNA shuffling of hit-variants	142
	DNA shuffling of homologous transcription factors	141
	indirect detection via sensing enabling metabolic pathways	135
Sensor output and dynamic range		
	initial negative selection against high basal expression without inducer molecule present	149
	strength of promoter and RBS upstream of the transcription factor or reporter gene	155, 158, 160
	operator multiplication upstream of reporter gene or utilization of the transcription factor as a 'roadblock' for RNA polymerase	154
	alternative carbon source in culture medium in order to avoid catabolite repression	111
	expression of importers for the target molecules (if not intracellularly made)	117

2.6 Research needs

As highlighted above, *in vivo* biosensors can play an important role in high-throughput screening for improved biocatalysts. However, their application is dependent on finding the right sensor specificity and performance characteristic. The required engineering activities are well known, yet still laborious, opening a wide field for the development of rational, computer-based workflows. Such workflows have to take into account the actual *in vivo* conditions in order to avoid implementation problems as discussed for the RNA-based sensors, where *in vitro* selected aptamers often fail to operate satisfactorily *in vivo*. Until this is achieved, directed evolution seems to be a good starting point for finding novel specificities. Finally, a standardized precise and quantitative description of sensor features would facilitate a more rapid exchange of genetic parts and thus the engineering of novel sensor systems.

2.7 Concluding remarks

In this chapter we reviewed the types and features of genetically encoded sensors available today. While RNA- and FRET-based sensors allow rapid detection of target compounds, the engineering of RNA-tools which work under *in vivo* conditions with realistic cytoplasmic conditions including high metabolite concentrations, physiological pH, temperature and so on, is still difficult. However, the accuracy of *in silico* predictions for RNA binding motifs is constantly improving and will hopefully lead to more applications in the years to come. Biosensors based on transcription factors are in a more mature state. However, due to the reliance on the cellular transcription/translation machinery for the production of the reporter protein, their response is somewhat slower. In addition, the predictions for protein folding and protein binding to small molecules are at least as challenging as for RNAs. Fortunately, many interesting regulator specificities are already known from literature and databases today. Intriguingly, many of the systems initially developed for environmental monitoring of hydrocarbon spills might be of considerable interest for metabolic engineering and high-throughput screening for products replacing certain petrochemicals.

2.8 Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 635536 (EmPowerPutida) and 635734 (ROBOX).

2.9 Author contributions

M.O.B. conducted initial literature research regarding transcription factor-based biosensors and designed all figures and tables. T.K. conducted literature research regarding RNA- and FRET-based biosensors. M.O.B., T.K. and S.P. wrote the manuscript, all authors reviewed and approved the manuscript.

Chapter 3: Directed Evolution of Biofuel-Responsive Biosensors for Automated Optimization of Branched-Chain Alcohol Biosynthesis

Maximilian Ole Bahls, Lukas Platz, Gaspar Morgado, Gregor Schmidt, Sven Panke

3.1 Abstract

The biosynthesis of alcohols is an important carbon-neutral alternative to unrenewable, petroleum-derived production. For the development of corresponding microbial production strains an *in vivo* biosensor is highly beneficial, enabling straight forward and automatable screening workflows. However, a very limited number of genetic parts is currently available for the detection of biofuel-related compounds. Here, we built biosensor-circuits by directed evolution of the transcription factor AlkS, heterologously expressed in *Escherichia coli*, for various industrially relevant C₄ and C₅ alcohols. The specificities of the novel biosensor variants were quantitatively characterized by flow cytometry and subsequently applied for *in situ* product detection in two screening applications concerning key steps in short chain alcohol production. First, we screened a site-saturation mutagenesis library of a pyruvate decarboxylase (KivD) for improved whole-cell catalysis of a specific 2-keto acid (4-methyl-2-oxopentanoic) to isopentanol. Second, we screened a combinatorial DNA-library of a 2-isopropylmalate synthase (LeuA), which encoded simultaneous variation of both, expression strength and release of product inhibition, or C-chain length-specific isopentanol detection, produced from glucose via the Ehrlich degradation pathway. For the latter application we created an automated, robotic platform-based workflow with unsupervised machine learning for data clustering of the biosensor outputs. We readily identified significantly improved strain variants, with the best-performing library hit exhibiting an improved isopentanol titer that was 45-fold (5.8 mM) increased over the parental isobutanol production strain. Besides, the specificity of isopentanol over isobutanol formation was improved to 13-fold. We expect the developed biosensor system to be highly useful in future alcohol and biofuel related research as well as in strain optimization for sustainable production of structurally similar chemical compounds.

3.2 Introduction

Synthetic biology and systems metabolic engineering hold the promise of sustainable production of transportation fuels and chemical building blocks from renewable feedstocks^{8, 15, 163}. Such bio-based production processes have the potential to reduce the reliance on limited fossil resources, as well as to critically improve the CO₂ balance of the chemical and transportation industry, as urgently required for tackling anthropogenic climate change¹⁶⁴⁻¹⁶⁶. Biofuels naturally produced by microbial cell factories include ethanol (*Saccharomyces cerevisiae*¹⁶⁷) and n-butanol (in acetone-butanol-ethanol fermentation¹⁶⁸). The production of branched-chain higher alcohols (BCHAs) has become a field of

considerable interest, as BCHAs not only have superior physicochemical properties as compared to ethanol when used for gasoline blends^{169, 170}, but also are closer to the properties required for aviation fuel. Here isopentanol is of particular interest¹⁷¹, as simple subsequent processing allows the production of blending agents for Jet-A1 fuel¹⁷². However, BCHAs are naturally produced in low concentrations and mixtures only¹⁷³, while high-titer bio production of defined BCHA requires metabolically engineered host organisms^{174, 175}. Multiple hosts including *Bacillus subtilis*^{176, 177} and *B. megaterium*¹⁷⁸, *Corynebacterium glutamicum*^{179, 180}, *Escherichia coli*¹⁸¹⁻¹⁸³, *Pseudomonas putida*^{176, 184}, *Synechocystis* PCC 6803¹⁸⁵, and *S. cerevisiae*^{186, 187} were genetically modified in order to produce branched-chain C₄-alcohols such as isobutanol (2-methylpropan-1-ol) or C₅-alcohols such as isopentanol (3-methylbutan-1-ol) and the further reduced compound isoprenol (3-methylbut-3-en-1-ol).

Despite those considerable strain engineering efforts, rapid and cost-efficient screening tools for large strain libraries are still a key limiting factor for achieving economically competitive bioprocesses through design-build-test-learn (DBTL) cycles required for the development of improved microbial strains^{21, 30}. Biosensor-based screening workflows have gained particular interest¹²⁰, as they are not only faster and cheaper to run than traditional chemical analyses for biofuels based on gas- and liquid-chromatography methods but are compatible with commercially available automation and high-throughput equipment as well^{20, 26}. This advantage is of particular importance for the up-and-coming biofoundries, which are providing an integrated infrastructure for accelerating DBTL cycles based on automation, liquid handling robots and artificial intelligence^{37, 38}. In addition, biosensor-based workflows are more environmentally benign by reducing the usage of auxiliary substances such as solvents and carrier gases, thus following the principles of green chemistry¹⁸⁸.

Biosensor capacity is encoded by genetic circuits which allow to report product concentrations *in vivo* and *in situ* by conversion into fluorescent signals, which can easily be measured with established high-throughput assays¹⁰⁶. This conversion is typically based on a transcription factor (TF), which recognizes the product of interest and subsequently triggers expression of a fluorescent protein in a concentration-dependent fashion^{153, 189}. To the best of our knowledge, only a single TF has been described so far for short-chain alcohol recognition in bacteria, the relatively uncharted σ^{54} -dependent BmoR regulator of *Thauera butanovorans*^{190, 191} that can trigger transcription from its cognate promoter P_{BMO}. BmoR/P_{BMO}-based biosensor systems were verified for alcohol-responsiveness and used for growth-based n-butanol titer detection¹²³. A similar system was used for screening of a

small random mutagenesis library of *E. coli* for isobutanol production¹²⁴ as well as for BmoR engineering, resulting in TF variants with reduced butanol affinity¹⁹². Given the diversity of targeted fuel molecules, there is a clear requirement for more biosensors and we chose a well-described TF in order to evolve it for the detection of short chain alcohols and differentiation between isomers such as n-butanol and isobutanol.

To this end, we started from the positive transcription factor AlkS (MalT/LuxR family, 882 amino acid residues) and its cognate promoter P_{alkB} ¹⁹³⁻¹⁹⁵, originating from oil-degrading organisms¹⁹⁶⁻¹⁹⁸ (see also Supplementary Table 3-1). Specifically, we utilized AlkS of *P. putida* GPo1^{197, 199-204}, which is well described for the detection middle-chain n-alkanes (C₅ to C₁₀^{112-114, 117, 205, 206}) and was suggested to recognize the corresponding alcohols to some extent as well²⁰⁷, for setting up an AlkS/ $P_{alkB}::sfGFP$ -based biosensor circuit in *E. coli*. Subsequently, we successfully aimed for engineering the inducer specificity^{98, 208} of AlkS by directed evolution, characterized the resulting biosensor variants in detail, and applied one of the variants in two experiments to improve isopentanol production from a defined substrate or directly from glucose. In order to increase the screening throughput, we concomitantly developed an automated, partially integrated screening workflow based on a robotic liquid handling platform and subsequently evaluated the biosensor data generated by clustering with a straightforward unsupervised machine learning algorithm.

3.3 Results

Design and construction of an AlkS-based biosensor circuit for alcohol detection

In order to convert the presence of alcohol in water into a rapidly detectable signal such as fluorescence, we placed sfGFP expression under the control of P_{alkB} on a medium-copy number plasmid (*ori* pBR322/*rop*). For AlkS expression from a low-copy plasmid (*ori* pSC101), we utilized the natural P_{alkS} promoter and additionally tested three commonly used constitutive promoters of varied strength obtained from the Anderson promoter library of the iGEM biobricks catalogue²⁰⁹ (see Figure 3-1). For 10 mM n-pentanol (C₅, the shortest straight chain functional alkane inducer known for AlkS¹¹⁴) added to the culture medium we found all designs to produce a fluorescence response, indicating a certain flexibility of the system with respect to *alkS* expression.

Still, the biosensor performance depended significantly on the strength of the promoter upstream of *alkS*. The low expression strength promoter (J23117) significantly limited the maximum biosensor output achievable. The circuit architecture with P_{alkS} , which includes a positive-negative feedback loop²⁰³, allowed tight control, as indicated by low basal sfGFP

output, while the maximum induction level at pentanol concentrations above 10 mM was as high as for the circuit with a strong constitutive promoter (J23100). However, the latter design showed a non-desirable high basal output. Generally, the circuit behavior with P_{alkS} was most similar to the one observed with the medium-strength promoter (J23106), but allowed a higher maximum sensor output while maintaining a lower basal expression. Thus, we continued with the $P_{alkS}::alkS/P_{alkB}::sfGFP$ system, allowing the largest fold-change upon induction of around 55-fold after 6 h.

Finally, we checked the homogeneity of the response at single cell level. Upon induction with 10 mM n-pentanol the sfGFP output of the biosensor strain population increased homogenously and almost completely (97%; AlkS, 10 mM n-pentanol), indicating that the biosensor circuit design was sound and that alcohol availability was comparable across the population, i.e. not limited by diffusion across the cell membrane. Additionally, in the absence of AlkS (empty plasmid control) and the presence of n-pentanol no biosensor response was observed, indicating that AlkS is essential for this response. As expected from literature, sensor output was only observed for n-pentanol and no response was found in the presence of the shorter C₄ straight-chain alcohol n-butanol.

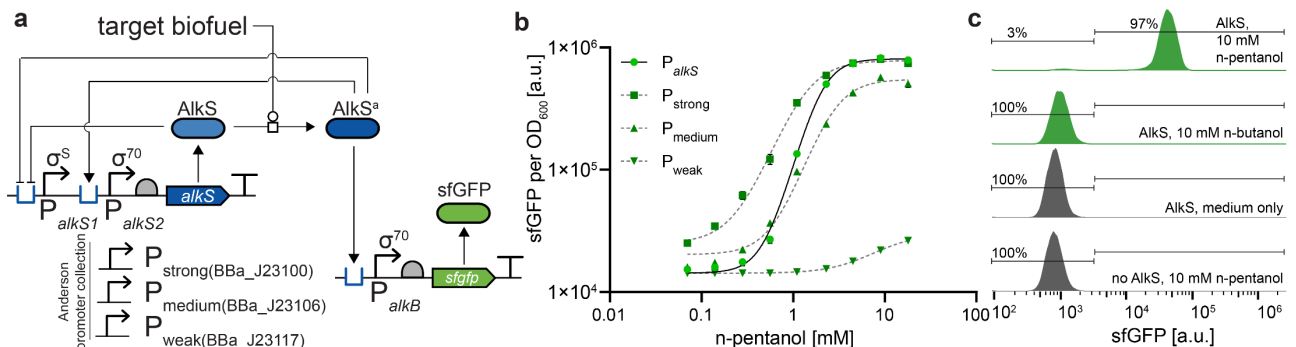


Figure 3-1 Design and characterization of AlkS-based biosensor circuits. a) Genetic circuit map for biofuel detection including a small promoter library for *alkS* expression. b) Dose-response curves of *E. coli* sensor strain with different promoter variants (data points are means with standard deviation, lines indicate Hill fits of means, all $R^2 \geq 0.99$, $n=3$). c) Biosensor response for C₄ (n-butanol) and C₅ (n-pentanol) straight-chain alcohols with the circuit using P_{alkS} upstream of *alkS* (flow cytometer, >35'000 events per population). No AlkS: same plasmid as in AlkS-experiment but without *alkS* gene; medium only: no alcohol added to the medium. AlkS^a, activated AlkS; Bba_x, BioBrick part name; sfGFP, superfolder green fluorescent protein. For simplicity, the sequence containing $P_{alkS1/2}$ is usually referred to as P_{alkS} .

Directed evolution of AlkS and characterization of expanded alcohol detection

Next, we used the $P_{alkS}::alkS/P_{alkB}::sfgfp$ -based biosensor circuit for evolving AlkS for the recognition of smaller and sterically more demanding alcohols than n-pentanol. As shown in Figure 3-2, we generated a variant library by epPCR-based random mutagenesis of the regulator sequence. Based on an alignment with the well-characterized MalT regulator of *E. coli*, the C-terminal domain of AlkS likely contains a helix-turn-helix motif responsible for binding to the cognate P_{alkB} -promoter DNA. This part was excluded from mutagenesis in order to limit unintentional variation in affinity to the promoter DNA (see Supplementary Figure 3-1).

The *alkS* variants obtained were reintroduced in the sensor circuit, resulting in a biosensor strain library with approximately 10^5 members, which we analyzed by flow cytometry and characterized for basic quality aspects. We found that about 29% of the variants retained either wild-type functionality and were inducible with n-pentanol or constitutively in the on-state (Supplementary Figure 3-2). On average, we found in the library immediately after construction about 5.5 amino acid residues exchanged per AlkS variant (7.6 bases exchanged per gene, 28% silent mutations, based on Sanger sequencing of ten *alkS* genes corresponding to a sequence total of approximately 26 kb). In addition, and as expected, we did not find any mutated bases outside of the designed random mutagenesis area. We also observed the bias known for error-prone PCR libraries²¹⁰, i.e. a preference for mutating A or T and overall more transitions than transversions²¹¹.

This sensor-strain library was cultivated in separate batches supplemented with 10 mM of either n-butanol, isobutanol, or isopentanol. The most fluorescent fractions of the corresponding populations were enriched by three consecutive rounds of cultivation in the presence of alcohol followed by fluorescence assisted cell sorting (FACS, for details on populations and gating see Supplementary Figure 3-3). Next, we verified that sorted populations were indeed enriched in alcohol-inducible biosensor variants and excluded variants that showed high basal sfGFP output (exemplarily shown for the sorting conducted with n-butanol, Figure 3-2b).

From the epPCR library, we readily found eight different AlkS variants for n-butanol, three variants for isopentanol and a single variant for isobutanol. On average, after enrichment 3.5 amino acid residues had been exchanged per AlkS variant. As we did not expect all random mutations to be beneficial or optimal for expanded alcohol detection, we continued by site-saturation mutagenesis in the parental AlkS of sites that had been identified in

several AlkS variants. Mutations and selected target-sites are summarized in Supplementary Table 3-2.

For the n-butanol screen, a mutation at position L401 was present in two independent AlkS variants, and we diversified this position with degenerate codons (NNK). Fifteen out of the 20 possible amino acids were readily retrieved (missing I, M, N, Q, and W), and all but two (F, P) of the residues found improved the biosensor response to n-butanol. In particular, AlkS L401G (abbreviated as AlkS “A”) showed the most improved sensor output, while also maintaining a low basal output level in the absence of inducer (Supplementary Figure 3-4 and 3-5). Applying the same diversification strategy to positions K183 and Q410 individually, which were mutated in addition to L401 in the error-prone variant with fewest mutations (variant five for n-butanol in Supplementary Table 3-2), did not result in AlkS variants with an improved induction profile for n-butanol further validating the importance of L401. For isopentanol and isobutanol, all variants found showed either an A375V/T or S379P mutation, and following the same strategy as explained above we confirmed mutation S379P (“B”) to be most beneficial for induction with both isopentanol and isobutanol, while site directed mutagenesis at position A375 lead to similarly improved induction with isopentanol but no significant improvement for isobutanol.

In this first library, we identified AlkS mutants that worked well for the short chain n-butanol and the branched chain isopentanol, but only to a smaller extent for the branched isobutanol. Hence, we conducted two additional rounds of evolution for isobutanol. The second round, now based on error-prone replication of the AlkS S379P and again followed by site saturation mutagenesis, lead to the further improved variant AlkS T336 R355H S379P (“C”). Finally, the third round, based on epPCR of the gene of variant C, led to the identification of two variants: “D” (AlkS T336 R355H S379P R397G), which showed an improved biosensor output for *sec*-butanol (butan-2-ol, see below) and “D_{e-p}” (AlkS D17E I19V R110G V171L T336 R355H V376L S379P R397G), which was a variant obtained directly from the screen with an improved dynamic range for isobutanol, but for which the improved dynamic range could not be recovered by transfer of single mutations to paternal variants.

Next, we tested the new biosensor strains synthesizing the defined AlkS variants (parent, A, B, C, D_{e-p}) against a diverse panel of C₃ to C₅ alcohols to characterize the inducer specificity (see Figure 3-2c and Supplementary Figure 3-6). As expected, the parental AlkS biosensor showed sfGFP output only for n-pentanol, with a change in fluorescence of about 85-fold. For the AlkS A variant, we found strong induction with n-butanol (64-fold) and isopentanol (about 30-fold). However, no induction with the other alcohols including isobutanol was

observed, a feature utilized later for the screening of strain libraries. For AlkS B we observed strong induction with isopentanol (50-fold) as well as n-butanol (40-fold) and *sec*-butanol (28-fold). Besides, this AlkS variant showed first improved inducibility with isobutanol (about 3-fold). For the AlkS mutants based on additional rounds of directed evolution, C and D_{e-p}, the output with isobutanol improved further to 5-fold and 10-fold, respectively. The fold-change increased up to 30-fold for AlkS D_{e-p} when 100 mM isopropanol was used as the inducer (please note that for the other alcohol concentrations the toxicity limit was significantly below 100 mM, see Supplementary Figure 3-7). We completed the biosensor characterization by determining the full transfer functions (fit to Hill curves, all $R^2 > 0.93$, Figure 3-2d, full data in Supplementary Figure 3-7 to 3-10) for the best-in-class biosensor and its cognate alcohol (Table 3-1).

In summary, the biosensor variants showed maximum fold-changes between 20- and 80-fold with EC₅₀ concentrations in the range of 5 to 30 mM and Z' scores >0.7 (indicating excellent properties for screening^{123, 150}). The fold-changes in sensor output were smaller for alcohols with shorter C-chains as well as sterically more demanding methyl groups or C-C double bonds. This reduction resulted mainly from increased basal sensor output of the corresponding AlkS variants, while the maximum output stayed similar to, or even increased above, the maximum wild-type sensor output with n-pentanol (Supplementary Figure 3-9).

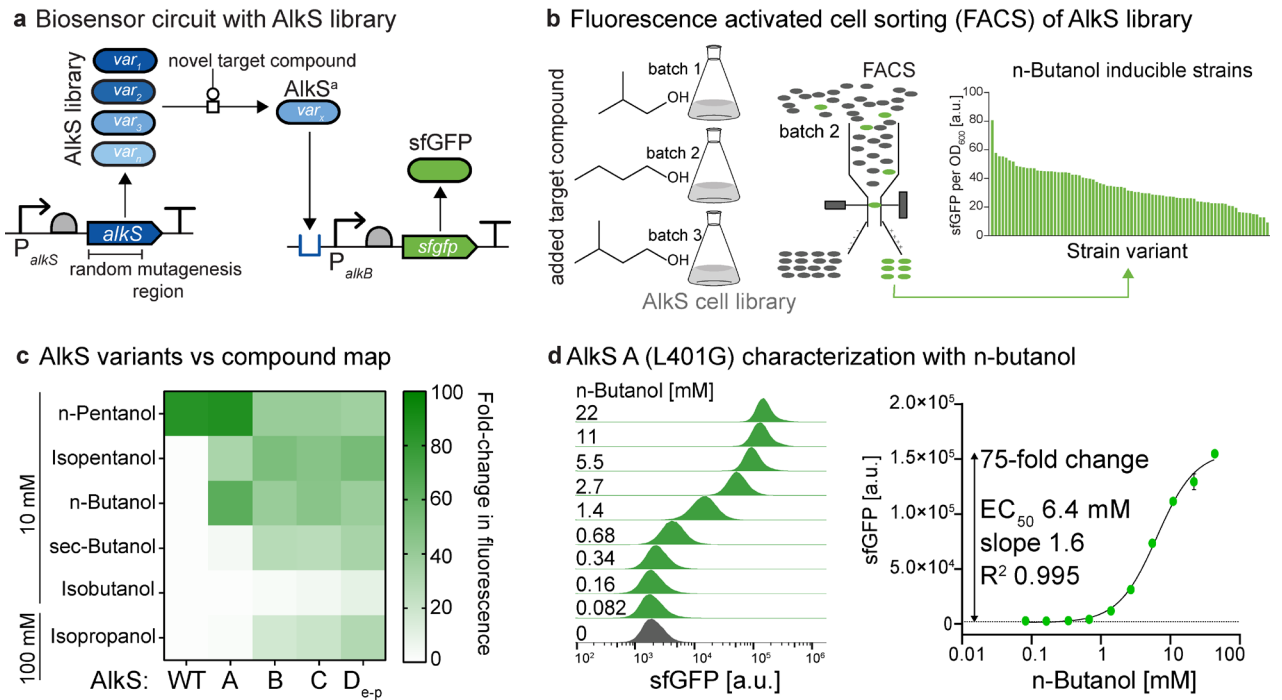


Figure 3-2 Directed evolution of transcription factor AlkS for small- and branched chain alcohol detection and biosensor characterization a) Scheme of the biosensor genetic circuit, including the part of AlkS that experienced random mutagenesis (amino acids 2-821). b) Representative FACS-based enrichment of alcohol responsive AlkS mutants (batch 1, isobutanol; batch 2, n-butanol (data shown); batch 3, isopentanol). c) Overview of alcohol induction profiles of AlkS variants (flow cytometer, >100'000 events per population, mean of median fluorescence, n=3). d) Representative single-cell dose-response of a biosensor strain with AlkS A and the corresponding transfer function with n-butanol (>100'000 events per population, line indicates Hill fit of means of median fluorescence, error-bars represent standard deviations, dotted line indicates mean basal sfGFP fluorescence, n=3). Abbreviations: AlkS^a, activated AlkS regulator.

Table 3-1 Biosensor performance characteristics for evolved AlkS variants and the corresponding target compounds. EC₅₀: half maximal effective concentration; n: slope (or Hill coefficient, a measure of sensitivity); † alcohol concentration at which the sensor response value for fold-change was determined, growth was severely affected at higher alcohol concentrations tested; Z': effectiveness of the circuit for high-throughput screens¹⁵⁰, generally, Z'-values above 0.5 are considered sufficient for an effective screening assay¹²³, also, the concentration at which the Z'-value was computed is given. For dose-response curves and parameter calculation see Supplementary Figure 3-9.

Compound	AlkS mutant (abbreviation)	EC ₅₀ [mM]	n	Max. fold-change, concentration [mM] [†]	Z', conc. [mM]
n-Pentanol	wild-type (WT)	3.9	1.8	86, 18.4	0.98, 9.2
n-Butanol	L401G (A)	6.4	1.6	75, 43.8	0.94, 11
Isopentanol	S379P (B)	9.8	0.95	29, 9.2	0.84, 4.6
Isoprenol	S379P (B)	7.2	1.3	21, 39.4	0.95, 39.4
sec-Butanol	(B) + T336S, R355H, R397G (D)	28.9	0.86	48, 87.6	0.94, 43.8
Isobutanol	(D) + D17E, I19V, R110G, V171L (D _{e-p})	12.4	2.0	19, 43.2	0.71, 43.2

Biosensor-based isopentanol detection in whole-cell catalysis

As a next step towards actual library screening, we verified the applicability of the biosensor circuit for detecting alcohols produced *in vivo*, instead of external supplementation of the inducer. Isopentanol was chosen as the target product for its outstanding physico-chemical properties suitable for bio-fuel related applications. For actual isopentanol biosynthesis, the commonly used Ehrlich degradation pathway was utilized¹⁷³. This pathway catalyzes the transformation of various 2-keto acids, precursors in amino acid biosynthesis, to products such as isopentanol and isobutanol¹⁷⁴. In order to limit false-positive biosensor activation from expected mixed iso-alcohol products, AlkS variant A (L401G) was used in the following experiments, as it responded well to isopentanol, but hardly to the structural isomer isobutanol (Figure 3-2c and Supplementary Figure 3-11).

To keep the experiment simple, we initially added only two enzymes to *E. coli*, specifically the enzymes for the final steps of Ehrlich degradation (see Figure 3-3): KivD (an α -ketoisovalerate decarboxylase of *Lactococcus lactis*) and YqhD (a NADPH-dependent aldehyde reductase of *E. coli*). To enable isopentanol formation, the corresponding 2-keto acid substrate 4-methyl-2-oxopentanoic acid was supplied to the culture medium (0.2% (v/v)). As expected, the expression of the two enzymes, from a bicistronic operon under the control of an L-arabinose inducible promoter, allowed whole-cell conversion of the 2-keto acid to isopentanol, while no product formation was observed in the absence of the operon (empty pSEVA631 plasmid control in Figure 3-3b).

Previously, the highest specific activity of KivD had been found *in vitro* for the substrate 2-ketoisovalerate (leading to isobutanol), with a reduced relative activity for 4-methyl-2-oxopentanoic acid (leading to isopentanol) of about 23%²¹², and the specificity of KivD for larger alcohols, such as n-pentanol, was shown to be improved for less bulky amino acid residues at position V461²¹³, an essential residue in catalytic function^{182, 213}. This suggested that mutagenesis at site 461 might lead to a KivD variant with improved properties for isopentanol production. To this end, a KivD V461X site-saturation mutagenesis (SSM) library was generated, re-introduced into the biosensor strain, and supplied with 4-methyl-2-oxopentanoic acid. The resulting library produced a variety of GFP expression levels, indicating a variety of isopentanol-levels, and we used this to validate the correlation between strain-produced isopentanol level as determined by chemical analysis and the GFP level resulting from the triggering of the biosensor circuit in the same cell. Over the whole range of product titers we found a good correlation between biosensor output and actual isopentanol concentration ($R^2 = 0.98$, see Figure 3-3c). The maximum isopentanol

concentration, found for wild-type KivD V461V, was 4.4 mM or 27% (mol/mol) 2-keto acid conversion and resulted in a nine-fold increase in biosensor fluorescence output. Besides, all strain populations showed homogeneously expressed sfGFP in flow cytometer analysis after whole-cell catalysis. The single cell data set also showed excellent agreement with plate reader-based measurements as used here in the following library screening workflow ($R^2 > 0.99$, see Supplementary Figure 3-12).

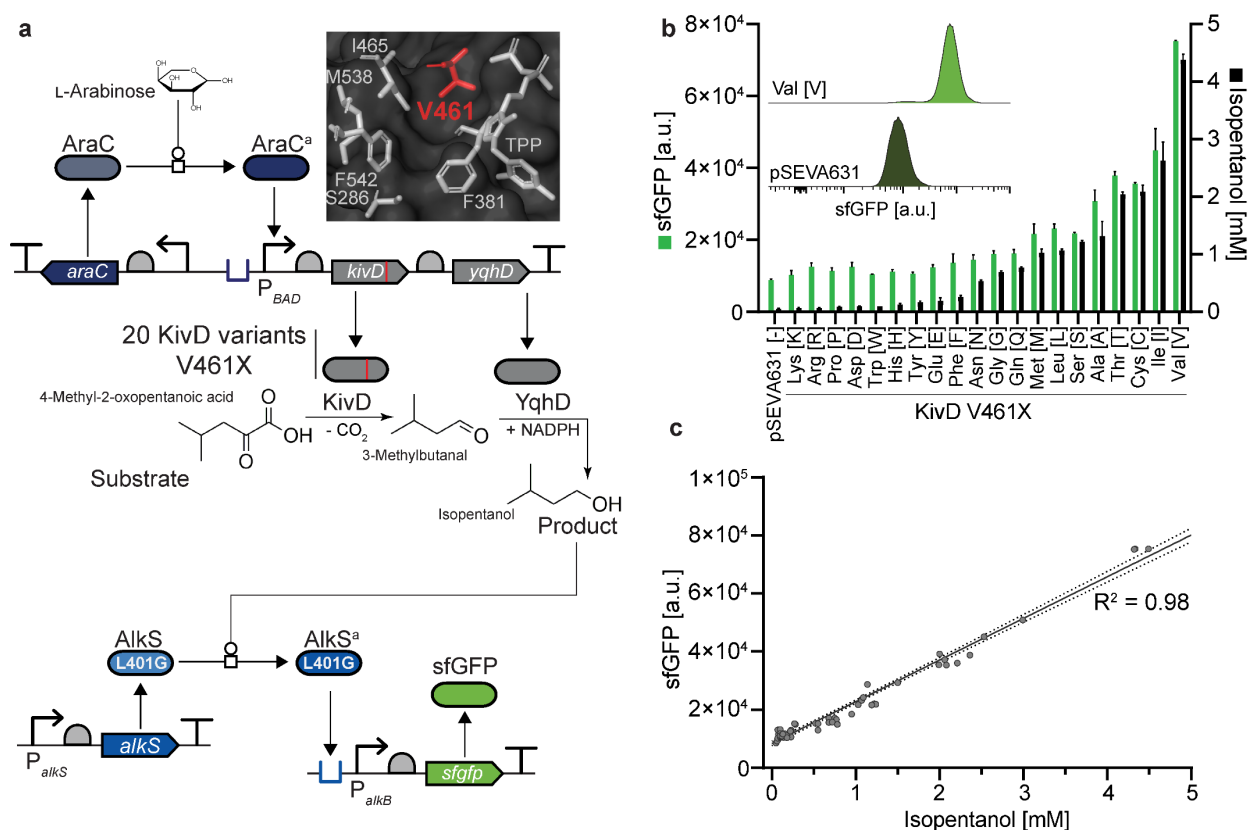


Figure 3-3 Biosensor verification for production and *in situ* detection of isopentanol. a) Genetic circuits for varied whole-cell transformation of a 2-keto acid substrate (4-methyl-2-oxopentanoic acid) to isopentanol by overexpression of a KivD V461X library and YqhD in combination with biosensor-based product detection. b) KivD variants ranked by isopentanol titer with the corresponding biosensor fluorescence output. The inset shows the fluorescence of representative populations for the best performing KivD V461 (Val [V]) and the negative control (pSEVA631). All flow cytometer data >100'000 events per population, means of medians with standard deviation, isopentanol titers are means with standard deviation, all n=3. c) Correlation of isopentanol product concentrations and the corresponding median biosensor fluorescence outputs (line is best-fit linear regression, dotted lines indicate the 95% confidence bands). AlkS^a, activated AlkS regulator; AraC^a, activated AraC regulator; KivD, α -ketoisovalerate decarboxylase (visualized with KdcA structure, PDB 2VBZ), sfGFP, superfolder green fluorescent protein; TPP, thiamine pyrophosphate; YqhD, alcohol dehydrogenase.

Automation-based screening of a combinatorial library for optimized BCHA biosynthesis

In order to efficiently produce branched-chain higher alcohols from glucose via the full Ehrlich degradation pathway, enzymes AlsS (α -acetolactate synthase, from *B. subtilis*) IlvC (keto-acid reductoisomerase) and IlvD (dihydroxy-acid dehydratase, both from *P. putida*) have to be overproduced in addition to KivD and YqhD¹⁸⁴. These five enzymes correspond to the pathway from pyruvate to isobutanol (Figure 3-4, brown and grey), but at the same time constitute the upper part of the isopentanol pathway. In order to adjust the metabolic flux towards isopentanol, enzymes LeuA (2-isopropylmalate synthase), LeuB (3-isopropylmalate dehydrogenase), LeuC (3-isopropylmalate dehydratase large subunit), and LeuD (3-isopropylmalate dehydratase small subunit, all from *E. coli*'s leucine biosynthesis pathway) have to be overproduced as well (Figure 3-4, blue). As a result, isobutanol is an unavoidable side product in isopentanol production (and vice versa, due to the essentiality of *leuABCD*²¹⁴) in this commonly applied production route. However, the AlkS A (L401G) variant allows selective detection of isopentanol without activation by isobutanol (see Figure 3-4c) and thus the detection of strains with an improved isopentanol yield per given substrate $Y_{P/S}$.

For the implementation of the proposed production-detection system, we started by stable, site-specific integration²¹⁵ of the isobutanol production operon into the genome of *E. coli* BW25113 $\Delta ilvE$, which lacks the gene for a branched-chain-amino-acid aminotransferase. This knock-out improves BCHA production¹⁸³ by eliminating side reactions otherwise competing with KivD for shared 2-keto acid substrates (also see Figure 3-4). As expected, when growing on production medium the $\Delta ilvE$ strain allowed improved isobutanol production (4.3 mM) as compared to the parental strain BW25113 and a common cloning strain (1.4 mM and 1.7 mM, respectively, see Supplementary Figure 3-13). In addition to the isobutanol production operon, overexpression of *leuABCD* is required in order to increase the metabolic flux towards isopentanol and thus away from isobutanol as described above. For this, we cloned the *leuABCD* operon on a broad-host range plasmid (*ori* pBBR1) under the control of the L-arabinose inducible *araC/P_{BAD}* system. Furthermore, the expression of the first gene in an operon can have a substantial influence on the expression of the operon²¹⁶, so we targeted the translation efficiency of the first gene in the operon, *leuA*, whose gene product also catalyzes the first committed step of isopentanol production. The *leuA* RBS strength was varied with a sixteen-membered degenerate library (NGGANG²¹⁷), spanning a predicted translation initiation range (TIR) of 89 to 15023 arbitrary units. As parental LeuA is known to be limited by feed-back inhibition¹⁸³, we additionally randomized

the LeuA amino acid position G462, a residue crucial for the release of product inhibition²¹⁸ (feedback release (FBR)). To this end, a degenerate NNK sequence coding for all 20 natural amino acids (and a stop codon) was used. A plasmid-based pool of this combinatorial library was combined with the isobutanol production host equipped with the biosensor system (Figure 3-5a).

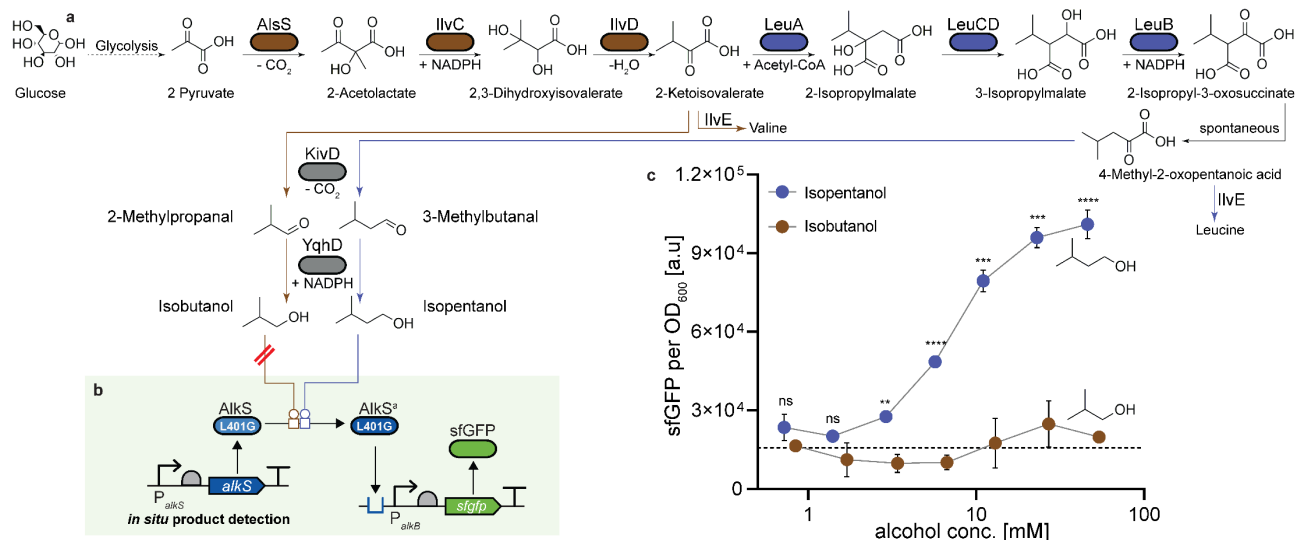


Figure 3-4 Branched-chain higher alcohol biosynthesis via Ehrlich degradation and biosensor-based product detection. a) Reactions and corresponding enzymes required for isobutanol and isopentanol biosynthesis. b) Specific product detection *in situ* with AlkS L401G-based biosensor system. c) Biosensor response for externally added isopentanol (blue, target product) and isobutanol (brown, expected side product). Data points are means with standard deviation, dotted line indicates basal sensor output without alcohol added ($n=3$). Abbreviations: AlkS^a, activated AlkS regulator; AlsS, acetolactate synthase; AraC^a, activated AraC regulator; IlvC, ketol-acid reductoisomerase; IlvD, dihydroxy-acid dehydratase; KivD, α -ketoisovalerate decarboxylase; LeuA, 2-isopropylmalate synthase; LeuB, 3-isopropylmalate dehydrogenase; LeuC, 3-isopropylmalate dehydratase (large subunit); LeuD, 3-isopropylmalate dehydratase (small subunit); sfGFP, superfolder green fluorescent protein; YqhD, alcohol dehydrogenase. Multiple unpaired t-tests with P-values (biosensor output for isobutanol vs isopentanol): ns \geq 0.05, ** 0.001 to 0.01, *** 0.0001 to 0.001, **** $<$ 0.0001.

For screening of the isopentanol production library, we developed an automated workflow using liquid handling robots and 384-well plates, which allow economic storage and handling of the variants in arrayed format. As the product molecules of interest diffuse readily across microbial membranes, this workflow also allowed for side-stepping false positive sensor activation, which we had observed for one-pot cultivations (Supplementary Figure 3-14). The workflow allowed reliable recovery of an isopentanol production strain from a mock library (Supplementary Figure 3-15). Briefly, the robotic system picked individual bacterial colonies (colony forming units, cfu) and transferred them into liquid cultures of 60 μ L volume (throughput limiting step, about 200 cfu h^{-1}). After 16 h of incubation, these precultures were replicated into production medium by the robot. Finally, isopentanol formation was assessed

by fluorescence output of the biosensor after 20 h production time. The detailed screening schedule and the integrated automation steps are summarized in Supplementary Figure 3-16 and 3-17. The raw biosensor data obtained from the plate reader was subsequently analyzed with a k-means²¹⁹ unsupervised machine learning algorithm in order to obtain a cluster of improved strain variants, comparison of which can inform on the logic of potential production improvement, as opposed to looking at a single top performer.

For the isopentanol production library, a total of 4320 cfu were screened (see Supplementary Figure 3-15). This represents a thirteen-fold theoretical oversampling of the strain library. However, we found that only about 50% of the initial library cfu contained the correctly assembled combinatorial DNA fragment (16/32, by colony PCR) thus reducing the actual oversampling value to six-fold, with a 58% likelihood to have screened the full library. Based on the biosensor data, three distinct strain clusters were generated (Figure 3-5b and Supplementary Figure 3-18). For all clusters arbitrarily selected variants were re-cultured in a deep-well plate fermentation system²²⁰ and evaluated in terms of biosensor output as well as BCHA titers produced. For all selected strains a clear positive correlation between isopentanol production and biosensor output was found ($R^2 = 0.94$, Figure 3-5c). As expected from the interdependency of the isobutanol and isopentanol production pathways, this coincided with a clear negative correlation of biosensor outputs and isobutanol concentrations ($R^2 = 0.95$).

For two of the sensor output clusters (no. 1 and 2, cluster sizes of 277 and 3976, respectively) neither increased biosensor output nor improved isopentanol production was observed (0/6 variants), as was expected from the low mean cluster values for biosensor output. Accordingly, for two of these six variants *leuA* was completely absent, in three variants frameshift mutations or deletions had occurred, while the last variant had a very low predicted TIR of 100 a.u. For the variants obtained from the cluster with highest mean sensor output (no. 3, cluster size of 355), about 90% (15/17 variants) showed increased biosensor output as well as improved isopentanol titers. In this cluster, we mainly found high predicted TIR values for *leuA* translation, while no clear pattern for the G462X sites was found except that we could not retrieve any variant with wild-type LeuA G462 (Supplementary Figure 3-19). Also, we verified by NGS of the full plasmid library pool that the best performing DNA sequences were not overrepresented before screening (Supplementary Figure 3-20). From this result we concluded that for efficient isopentanol production [i] elimination of the feedback inhibition of *leuA* is required and [ii] high expression strength for *leuA* is required at the same time.

Starting from the parental isobutanol strain producing 0.13 mM isopentanol and 8.2 mM isobutanol, the best-performing isopentanol variant “isoC₅OH_{MAX}” obtained from cluster no. 3 (RBS with AGGAGG consensus motif and a predicted TIR 15023 a.u., G462F) produced a titer of 5.8 mM isopentanol, with 0.4 mM of isobutanol as a side product. This represents a 45-fold improvement in product titer, with a molar ratio of isopentanol over isobutanol of about 13-fold. The plasmid harboring the best performing *leuABCD* operon found in the screen was re-introduced in the parental isobutanol production strain (without the biosensor system). Alternatively, we inserted into this strain one of two control plasmids, one carrying the best performing *leuABCD* operon but with a weak RBS in front of *leuA* (predicted TIR of 89 a.u.) or with the high-performing RBS but without a FBR mutation in *LeuA* (G462G). When growing on production medium (5 g L⁻¹ glucose, 1.25 g L⁻¹ yeast extract) the optimized strain produced 8.3 mM of isopentanol and 0.28 mM of isobutanol, while the two control strains produced significantly less isopentanol (0.95 mM or 1.0 mM, respectively), confirming that the combination of both high expression strength as well the FBR mutation are required for high isopentanol production (see Figure 3-5d).

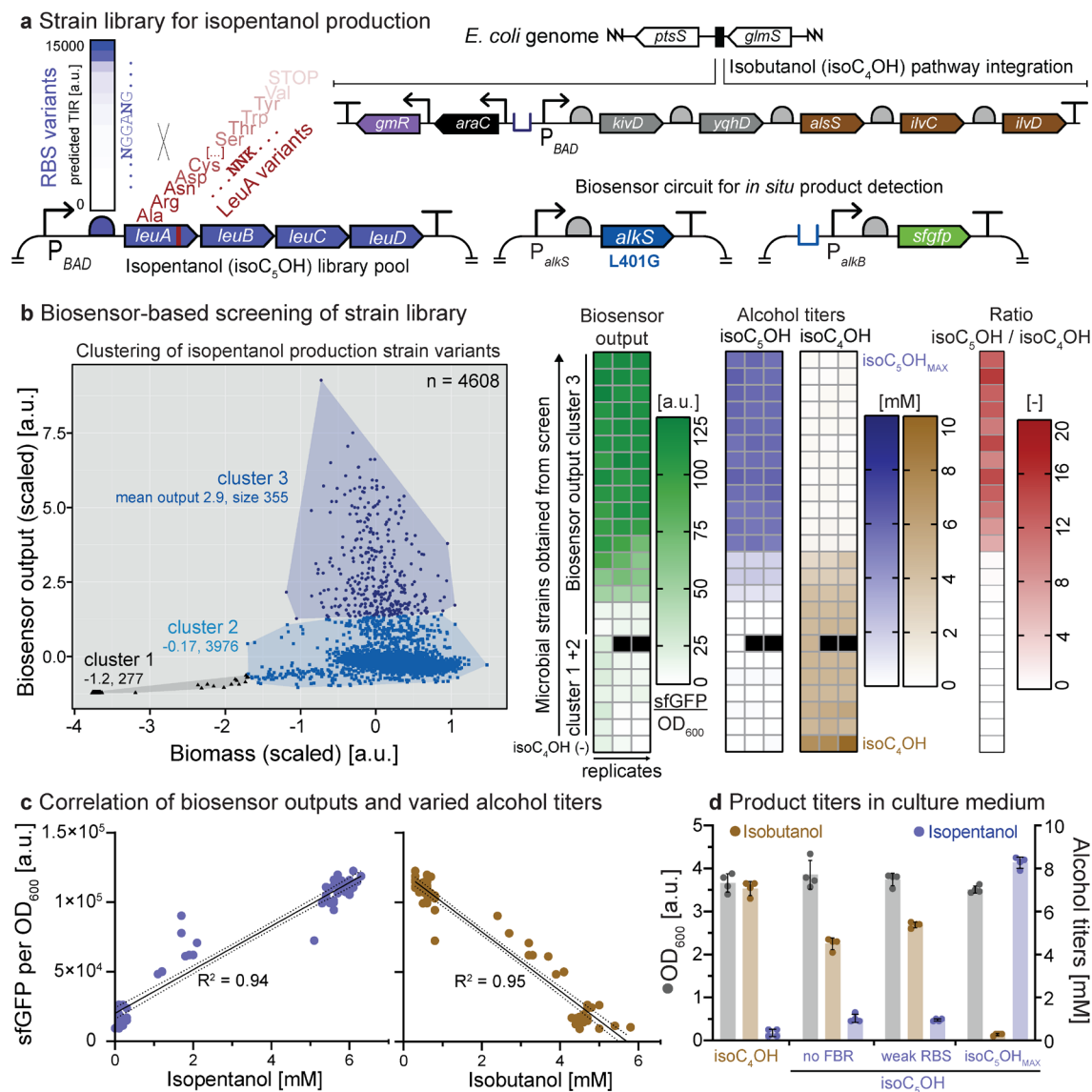


Figure 3-5 Biosensor-based high-throughput library screen for improved isopentanol production strains. a) Overview of combinatorial library for *leuA* expression, catalyzing the first committed step of isopentanol production, in an isobutanol production strain background equipped with the AlkS L401G biosensor system. Sixteen ribosomal binding site (RBS) variants were combined with 20 variants of *leuA* stemming from site-saturation mutagenesis at feed-back release (FBR) site G462X. b) Clustering of scaled biosensor output and biomass, based on a k-means unsupervised learning algorithm. The corresponding data set was obtained from an automated screening workflow with 4320 library variants picked plus additional blank and control wells (culture volume of 60 μ L, total 4608 wells). Variants of each cluster were re-assessed in terms of biosensor output and alcohol titers produced (three variants each for clusters one and two, 17 variants for cluster three; culture volume of 750 μ L, black wells indicate no growth, $n=3$). c) Correlation of alcohol titers with biosensor fluorescence outputs for the microbial strains obtained from the screen (line is best-fit linear regression, dotted lines indicate the 95% confidence bands). d) Verification of switch from isobutanol to isopentanol production with culture supernatants of the top-performing isopentanol strain (“isoC₅OH_{MAX}”, TIR of 15023 a.u. and G462F) and control strains either lacking the FBR mutation (TIR of 15023 a.u. and G462G, “no FBR”) or the strong RBS (TIR of 89 a.u. and G462F, “weak RBS”) in comparison to the parental isobutanol production strain (“isoC₄OH”) after 20 h cultivation at 30°C (single data points with bar representing their mean and error-bars indicating standard deviations, $n=4$). Transcription initiation rate (TIR) predicted according to²²¹. Abbreviations: as in Figure 3-4. GmR, gentamycin resistance (gentamicin-3-acetyltransferase); RBS, ribosome binding site.

3.4 Discussion

Microbial strain engineering is a key priority for enabling the sustainable production of chemical building blocks and biofuels. With ever more powerful bioengineering and DNA synthesis methods for the creation of extensive strain libraries, biosensor-based screening of such libraries becomes a focal point of synthetic biology research^{21, 26}. Mainly, biosensors can facilitate more economical and higher throughput assessments of strain library members in comparison to established chemical analysis methods.

Previously, AlkS-based biosensors were utilized for *in vivo* detection of n-alkanes^{118, 205} ($\geq C_5$), mainly in order to detect environmental hazards including oil spills¹¹². Here we applied the AlkS system to the sustainable bio-production of similar compounds and thus for a potential replacement of oil-reliant processes. Directed evolution of the AlkS transcription factor allowed for the construction of genetic circuits coding for *in situ* detection of various BCHA, a compound group that was previously not recognized by wild-type AlkS. The approach of initial error-prone library screening followed by SDM at potential hot-spot sites revealed two single residues of the total 882 amino acids constituting AlkS which are amenable for significantly expanding the inducer spectrum. Specifically, residue L401 was mutated for recognition of n-butanol as well as isopentanol and S379 for alcohols such as 2-butanol and isobutanol, while also serving as the parental template for further improved AlkS mutants harboring several mutations. All of those compounds are part of the OECD list of high production volume chemicals²²² and thus produced on the kilo-ton scale in at least one country or region. Besides, they have multiple target markets, including those for biofuels, solvents and lubricants, as well as for flavor and fragrance related products¹⁷¹. None of these hot-spot sites had been found previously in an effort made for improved hexane and pentane detection with AlkS variants²⁰⁵.

As no protein structure is available for AlkS and only very few related regulator structures of the MalT-family were solved, it is difficult to directly map the mutated amino acids to a specific function or mechanism. However, similar observations for engineering non-natural effector responses were made with mutants of XylS¹⁴¹ and HbpR²⁰⁸, two transcription factors also originating from *Pseudomonas* spp. In these studies the detailed underlying functional principles remained elusive as well. Nevertheless, a link between broadened inducer spectrum and increased basal output might be in accordance with a recently proposed double autoinhibition mechanism of signal transduction ATPases²²³ (see Supplementary Figure 3-21 for a corresponding AlkS structural model based on homology²²⁴). This signal transduction mechanism includes a tightly controlled state, in which the sensor domain

covers the core nucleotide-binding oligomerization domain (NOD). Upon first, low affinity inducer binding a more open and responsive, but partially leaky, state is formed, with reduced autoinhibitory interaction between NOD and sensor domain. Next, high-affinity inducer binding allows an ADP to ATP exchange in the NOD, which triggers formation of the active, multimeric regulator. Such a mechanism would suggest that single, specific mutations in the sensor domain sequence might allow for a changed inducer spectrum while maintaining the tightly controlled, autoinhibition state with basal output levels similarly low as the one found for the wild-type regulator, as seen in the present work for the n-butanol-enabling mutation L401G. However, less-specific mutations in the NOD-sensor domain interface could lead to a regulator structure with its equilibrium state shifted towards the high-affinity binding state (i.e. skipping the low-affinity binding state) and thus allowing induction with additional inducers otherwise not recognized, at the cost of a partially increased basal activity, as observed here for isobutanol-detecting AlkS variants.

When applying the new biosensor to facilitate screening activities, we found in a KivD V461X library screen that the targeted amino acid position is indeed crucial for 4-methyl-2-oxopentanoic acid conversion, as expected from the location of the residue in the active site pocket²²⁵. The wild-type valine residue performed best in whole-cell catalysis experiments for isopentanol production, while for other substrates larger than α -ketoisovalerate, the optimal KivD substrate *in vitro*²¹², beneficial mutants such as V461A had been found previously, for example for improved n-pentanol production²¹³. Here, most likely such improvements were not observable as the specific substrate was supplied at a relatively high concentration and was thus not in competition with other high-concentration substrates originating from a common overexpression pathway. Crucially, this library screen confirmed that the biosensor system reliably ranks strain variants in agreement with the actual product titers achieved, while the corresponding fluorescence measurements took seconds to be carried out instead of several hours for gas chromatography analysis.

For the biosensor system using AlkS L401G, significantly different sensor outputs between isobutanol and isopentanol as well as the structural isomers isobutanol and n-butanol were found, a feature that had not been found previously despite efforts to engineer BmoR¹⁹². This inducer-specificity is an essential feature in order to detect strains that produce isopentanol at high yields. The most common metabolic engineering approach for branched-chain higher alcohols utilizes the Ehrlich degradation pathway¹⁷⁵. As a result, isobutanol is an essentially unavoidable side product in all such metabolic engineering efforts, e.g. in *E. coli*-based BCHA production from cellulose and protein rich waste streams²²⁶, which

leads to false positive biosensor output unless the system is specific for the C₅ product. Here, we demonstrated a successful screen for improved isopentanol production without adverse false-positive sensor activation, despite starting from a parental strain producing isobutanol from a genome-integrated overexpression pathway.

In the screened combinatorial library for isopentanol production, we found high *leuA* RBS strength to be crucial for product formation, while for the FBR site G462X no clear pattern was observed. This is consistent with a previous study in which the LeuA G462D expression level was found to be crucial for isopentanol formation¹⁸³. Specifically, for the best performing strain obtained from the screen we found the consensus AGGAGG Shine-Dalgarno sequence, which also had the highest predicted TIR in the strain library (15023 a.u.). Besides, at the FBR site-saturation position we found a G462F mutation. The corresponding isopentanol production plasmid was re-introduced in the original isobutanol production strain, now without the biosensor system. When growing on production medium in 24 square-well plates (5 g L⁻¹ glucose and 1.25 g L⁻¹ yeast extract) with oxygen transfer rates similar to those in shake flasks²²⁷, titers of 8.3 mM isopentanol (equivalent to 0.73 g L⁻¹) and 0.28 mM isobutanol were produced after 20 h of cultivation with glucose being fully depleted. For control strains either having a weak RBS or no FBR mutation, reduced isopentanol titers of 0.95 mM and 1.0 mM were found, respectively, while isobutanol titers were significantly higher with 5.4 mM and 4.5 mM, respectively. Assuming a contribution of the yeast extract as determined previously¹⁸³, the maximum isopentanol titer found corresponds to a product yield on glucose of about 0.14 g_{isoC₅OH} g_{Glc}⁻¹. Those results are on par with state-of-the-art metabolic engineering efforts for isopentanol production (0.81 g L⁻¹ and 0.13 g_{isoC₅OH} g_{Glc}⁻¹ for 5 g L⁻¹ glucose and 5 g L⁻¹ yeast extract) with an *E. coli* strain harboring a synthetic RBS for expression of LeuA (G462D) but also harboring seven additional genomic gene knock-outs as well as requiring three plasmids for pathway overexpression¹⁸³.

The number of strain variants tested in this screening campaign of more than 4000 cfu is, to the best of our knowledge, amongst the largest published for biofuel-related research and could easily be extended due to the automated workflow. Previously, a two-step screen of 960 cfu of an RBS library for improved n-butanol production had been successfully implemented with a BmoR/P_{BMO}::*tetR* sensor strain¹²³. This was achieved by overexpression of the last two steps of the Ehrlich degradation pathway, with RBS libraries for the corresponding two enzymes, in combination with supplementation of the intermediate keto-acid substrate to the culture medium and additional genome engineering in order to obtain

a single product with limited side-product formation. Next, the sensor strain culture was mixed with spent medium for product detection by increased sensor strain growth in the presence of tetracyclin. For isobutanol production from glucose in defined medium, 50 strain variants of a random genome-mutagenesis library were screened also by overexpressing the last two steps of the Ehrlich degradation pathway together with a BmoR/P_{BMO}::*gfp* biosensor system for *in situ* product detection¹²⁴. In this screen, changes in specific fluorescence output of up to 1.4-fold were found, as compared to the parental strain. Besides, the two screens described relied on 96-well plates and manual liquid handling.

Here, we directly equipped production strains with the biosensor system for *in situ* product detection, without the need for additional genome engineering efforts for avoidance of sensor-confusing side product formation, while also leaving undefined yeast extract in the culture medium composition. The combined production/sensing strategy streamlines the workflow and also prevents false-positives from carry-over as found in the two-step growth-based biosensor assay. We used 384-well plates for all screening steps in order to increase throughput and reduce material consumption, while also using fixed metal tips and a single replicator for liquid handling on the robotic system. Overall, this workflow significantly reduced the amount of plastic waste generated, while also requiring no toxic or harmful solvents for robotics operation and fluorescence-based titer readout (in contrast to traditional biofuel production strain evaluation). Besides, the optical measurements for the whole library took few minutes with about 1 min plate⁻¹, while GC-FID methods for fermentation product analysis usually need a similar time span for a single sample²²⁸. The biosensor-based readout data sets were directly utilized for simple unsupervised learning and data clustering as desirable for automated, efficient and unbiased pattern identification in screening for high-producing strains³³. These data clusters could readily be utilized for full computational integration of library evaluation in DBTL cycles in biofoundries for the accelerated development of industrially relevant microbial strains^{20, 37}. The data clusters are also suitable for visualization of library variant distributions that simple rankings of specific fluorescence might miss due to normalization to biomass, while also allowing principle component analysis if multidimensional readouts are used³⁴.

In conclusion, we highlighted the evolvability of the transcription factor AlkS for the recognition of industrially relevant compounds and biofuels. The resultant biosensor circuits were applied successfully in different strain backgrounds including cloning strains as well as genetically more stable strains suitable for process upscale²²⁹. Finally, we applied the thoroughly characterized biosensor circuits to automated screening of, in terms of variant

size as well as side products, challenging microbial strain libraries for isopentanol production. The developed biosensor system could be readily employed for additional tasks such as dynamic feedback control and time-of-requirement expression of alcohol efflux pumps²³⁰. The AlkS variants found most likely also allow for the detection of other compounds available from biosynthesis such as (gaseous) n-butane²³¹ or additional industrially relevant alcohols derived from Ehrlich degradation, e.g. phenylalcohols²³², as well as other molecule classes with a relatively straight C-chain, e.g. ketones²⁷. Besides, AlkS is known to be functional in other industrially relevant host microorganisms such as *P. putida*²⁰⁶. As the majority of the molecules mentioned readily diffuse across microbial membranes, the developed sensor strains should be useful for screening of several other product-molecule sources as well, for instance in co-cultivation schemes with organisms that are less amenable to intricate genetic circuit engineering.

3.5 Methods

Materials and chemicals

If not stated otherwise, all chemicals including short DNA oligomers were purchased from Sigma Aldrich (Buchs, Switzerland), DNA isolation and purification kits from Macherey-Nagel (Dueren, Germany) or Zymo Research (Irvine, CA, USA), enzymes from New England Biolabs (Ipswich, MA, USA) and DNA fragments from IDT (Coralville, IA, USA). PCRs were carried out using high-fidelity Phusion or Q5 polymerases and the sequences of DNA parts were verified by Sanger sequencing at Microsynth (Balgach, Switzerland).

Bacterial strains, media, and cultivation conditions

In general, all cultivation and cloning procedures followed standard molecular biology protocols²³³ and suppliers' recommendations. A red fluorescent derivative of *E. coli* NEB 10 β (the strains used in this study are summarized in Supplementary Table 3-2) was constructed by replacing the allose operon (*alsRBACEK*) with the gene for a constitutively expressed red fluorescent protein (λ P_L::*mCherry*, see Supplementary Figure 3-22) by λ -red based recombineering²³⁴. The resulting *E. coli* 10 β (*mCherry*) strain was used for all DNA cloning procedures unless stated otherwise, AlkS library screening and biosensor characterization. *E. coli* DH5 α was used for KivD library screening experiments and *E. coli* BW25113 for all other alcohol production experiments. *E. coli* strains were transformed with DNA by electroporation with a micropulser (Bio-Rad, Hercules, CA, USA) at 1.8 kV and cuvettes with a 1 mm gap (Cell Projects, Harrietsham, UK). For cloning procedures and precultures Lysogeny broth²³³ (LB) was used (Becton Dickinson, Franklin

Lakes, NJ, USA). If not stated otherwise, all other experiments were conducted in M9 mineral medium²³³ (pH 7.4 adjusted with NaOH) supplemented with 4 g L⁻¹ glucose, 20 mg L⁻¹ thiamine and 100 µL L⁻¹ of trace element solution US*¹⁹⁴. For *E. coli* 10β(*mCherry*) a leucine-isoleucine-valine (LIV) amino acid mix²¹⁷ was added. For combinatorial library screening and alcohol production in 96-well or 24-well plates, M9-based production medium was prepared with 4 g L⁻¹ glucose and 1.25 g L⁻¹ yeast extract or 5 g L⁻¹ glucose and 1.25 g L⁻¹ yeast extract, respectively. For solid media plates 15 g L⁻¹ agar was added (Applichem, Darmstadt, Germany). Antibiotics were used, where appropriate, at the following concentrations: 100 µg mL⁻¹ carbenicillin, 25 µg mL⁻¹ chloramphenicol, 50 µg mL⁻¹ kanamycin, 10 µg mL⁻¹ gentamycin, and 50 µg mL⁻¹ streptomycin. LB cultures were routinely cultivated at 37°C, 220 rpm, and mineral medium cultures at 30°C, 220 rpm. Culture volumes occupied 20% of the nominal volume for cultivation tubes and 10% for Erlenmeyer flasks unless stated otherwise. Cultivations in deep well plates (96-well or 24-well format) were carried out in a miniaturized fermentation system²²⁷ (EnzyScreen, Hemsteede, The Netherlands) with a shaking amplitude of 50 mm at 300 rpm (LT-X, Kuhner, Birsfelden, Switzerland). For induction and alcohol production experiments plates were sealed with aluminium foil or clear peelable heat seals, or screw-cap tubes and flasks were used in order to prevent alcohol evaporation. Cells were kept at 4°C for short term storage and as cryostocks with 20% (v/v) glycerol at -80°C otherwise.

Construction of sensor plasmids

All plasmids used in this study are summarized in Supplementary Table 3-2. DNA oligomers and additional DNA sequences synthesized are listed in Supplementary Table 3-3 and 3-5, respectively. Selected plasmid maps are shown in Supplementary Figure 3-23.

To construct the sensor-input plasmid, a low-copy number plasmid harboring the gene for the transcriptional regulator AlkS under its natural promoter P_{alkS} was constructed by PCR amplification of the corresponding region with primers AL1 and AL2 from pESM7¹⁹³ and ligation into pCK01²³⁵ using EcoRI and HindIII sites, resulting in pCK01_ P_{alkS} _alkS (pSC101 *ori*, Cm^R). Constitutive promoter variants (sequences obtained from Anderson promoter collection²⁰⁹) for *alkS* expression were inserted by two-fragment isothermal assembly (ITA) with synthesized DNA fragment BBa_J231xx_alkS and a pCK01_ P_{alkS} _alkS fragment PCR-amplified with primers pCK01_J231XX_alkS_fwd and pCK01_J231XX_alkS_rev. A negative control plasmid pCK01_empty_cargo was created by EagI and SacI double digest of pCK01_ P_{alkS} _alkS and self-ligation of the plasmid backbone, which removed a 3.4 kb

DNA stretch including $P_{alkS}::alkS$. In general, template plasmid DNA in PCR-based procedures was digested with DpnI when appropriate.

The sensor output plasmid was constructed by exchanging the origin of replication of pSEVA231¹⁹⁵ for that of pBR322 obtained from pRK793²³⁶ by restriction with *Ascl* and *FseI* and subsequent ligation. The pBR322 fragment was PCR amplified with primers pBR322_for-*Ascl* and pBR322_rev-*FseI*. Next, a promoterless version of sfGFP (incl. RBS) was inserted within *EcoRI* and *PacI* sites of the multiple cloning site. The corresponding DNA fragment containing sfGFP was PCR amplified from pET28_sfGFP (Sandia Biotech, Albuquerque, NM, USA) with SG_fwd and SG_rev prior to restriction and ligation. The P_{alkB} sequence was inserted upstream of *sfgfp* by ligation within *SpeI* and *EcoRI* sites, resulting in plasmid pAB_ P_{alkB} _sfgfp_KanR (pBR322/*rop ori*, KanR). The DNA fragment containing the promoter sequence was obtained as the long DNA oligomer PalkB, made double stranded with the oligonucleotide PalkB_rev and Q5 DNA polymerase treatment before restriction. For screening experiments with BW25113 $\Delta ilvE^{214}$, the KanR resistance of pAB_ P_{alkB} _sfgfp_KanR was changed to AmpR by two-fragment ITA with PCR fragments of pAB_ P_{alkB} _sfgfp_KanR (with primers pSEVA_abR_exchange_bb_fwd and pSEVA_abR_exchange_bb_rev) and pSEVA181 (with primers pSEVA_abR_exchange_AbR_fwd and pSEVA_abR_exchange_AbR_rev).

AlkS libraries: Mutagenesis, screening, and analysis

AlkS variants were constructed by error-prone PCR (epPCR) amplification of *alkS* with primers AlkS_fwd_Gibson and AlkS_rev_Gibson, which generated a fragment that excluded the start codon as well as a putative C-terminal DNA binding domain (see Supplementary Figure 3-1). The epPCR was run with Taq polymerase, unbalanced concentrations of dNTPs²³⁷ (0.35 mM dATP, 0.4 mM dCTP, 0.2 mM dGTP, 1.35 mM dTTP), 0.05 mM Mn^{2+} , 2.95 mM Mg^{2+} and 60 ng of pCK01_ P_{alkS} _alkS template DNA per 50 μ L reaction for 30 cycles. The template DNA was linearized with *MscI* and *NcoI* prior to PCR (leading to the loss of a 38 bp fragment in the gene for CmR and thus preventing simple regeneration of the source vector by religation). The *alkS* plasmid-library was constructed by two-fragment ITA with the epPCR *alkS* part and a PCR-amplified pCK01_ P_{alkS} _alkS backbone fragment (with primers AlkS_fwd_Gibson_backbone and AlkS_rev_Gibson_backbone, here the template DNA was linearized with *AatII* and *EcoRV*, $\Delta 48$ bp in *alkS*). The *alkS* plasmid library was used to transform *E. coli* 10 β (*mCherry*) and plasmid DNA was purified from pooled cfu. The library consisted of approximately $8.1 \cdot 10^4$ variants stored at -20°C . For FACS-based screening, *E. coli* 10 β (*mCherry*) [pAB_ P_{alkB} _sfgfp_KanR] was transformed

with the *alkS* library and recovered for 1 h in SOC medium²³³ (1.5×10^5 re-transformants). Subsequently, LB with antibiotics was added to the culture for additional 2 h followed by centrifugation (3500 rcf, 4°C, 10 min). The cell pellet was resuspended in M9 medium with 0.2 g L⁻¹ glucose and cultured for 16 h. This culture was diluted ten-fold in fresh M9 medium with 4 g L⁻¹ glucose, split into separate cultures and incubated for 3 h before target alcohols were added at 10 mM for additional 3 h of incubation. Next, the cells were washed twice with PBS (pH 7.4; Thermo Fisher Scientific, Waltham, MA, USA) and diluted to an OD₆₀₀ of 0.02 (BioPhotometer D30, Eppendorf). For two rounds of enrichment (Becton Dickinson Influx Sorter), cells were sorted into LB and after 2 h of incubation the medium was changed by centrifugation and subsequent re-suspension in defined medium with 0.2 g L⁻¹ glucose, subsequently repeating the workflow as described above. During the third enrichment sorting, cells were spotted on LB agar plates in 384-well format (Omintray single well plates, Thermo Fisher Scientific). Briefly, the gates for selecting alive cells were based on perpendicular forward scattering (FSCPerp) vs side scattering (polygon-shaped gate, selecting the overall cell population) as well as FSCPerp vs trigger pulse width (rectangle-shaped gate, selecting single cells). For enrichment of functional biosensor variants, gating was based on the sfGFP fluorescence signal (excitation wavelength of 488 nm and emission measured with a 530/30 nm band pass filter), while mCherry was used as a biomass marker (561 nm and 610/20 nm band pass filter). In each screening round approximately 10⁷ events were recorded (about 100-fold oversampling of library sizes) and the most fluorescent cells were enriched (<0.2% of total population), followed by increasingly restrictive gating (see Supplementary Figure 3-3 for details).

Single cfu of the sorted libraries were re-grown in LB (500 µL in deep well plates, 10 h) and subsequently diluted in defined medium (2.5% (v/v) inoculum, 750 µL, 14 h). For fluorescence measurements, cells were re-inoculated in defined medium (4% (v/v) inoculum, 500 µL) and after 2 h of incubation 250 µL of fresh medium with or without 0.3% (v/v) of alcohols was added (i.e. 0.1% (v/v) in the culture). After 6 h of induction, 200 µL samples were taken in micro-titer plates (Greiner Cellstar, Kremsmuenster, Austria) and the optical density was measured at 600 nm (OD₆₀₀) as well as green fluorescence at 488 nm with a 530/20 nm filter (sfGFP) and red fluorescence at 579 nm with a 616/20 nm filter (mCherry). If necessary, samples were diluted in PBS to fit the linear measurement range of the plate reader (Tecan M1000 Pro, Maennedorf, Switzerland).

The *alkS* sequences of the input plasmid of the top-performing biosensor constructs were subsequently sequenced including the upstream RBS and promoter region (nine *alkS*

variants for n-butanol, six for isopentanol, four for isobutanol, as well as the parental wild-type). For Sanger sequencing of each *alkS* variant four primers were used (pAlkS_fwd_1, pAlkS_fwd_2, pAlkS_fwd_3, and pAlkS_fwd_4). The *alkS* sequences were aligned with the ones of the natural OCT plasmid (NCBI GenBank AJ245436) and five additional expression plasmids containing the *alkS* sequence (AJ299427, AF118921, AJ302087, AJ414668, and AJ302086) using Clustal Omega 1.2.4²³⁸ (via EMBL-EBI web service). Note that in the parental *alkS* sequence we found bases C1228A (Q410K), C1392G (T464T), and T2568C (H856H) mutated as compared to the *alkS* sequence of the OCT plasmid (AJ245436).

For PCR-based single site-directed *alkS* mutagenesis at amino acid residues K183, A375, S377, S379, L401, and Q410, primers with NNK codons at the corresponding DNA positions were used (AlkS_QC_residue_NNK_fwd and AlkS_QC_residue_NNK_rev, e.g. AlkS_QC_K183_NNK_fwd and AlkS_QC_K183_NNK_rev for residue K183) and applied to template plasmid pCK01_P_{alkS}_alkS. *E. coli* 10β(*mCherry*) [pAB_P_{alkB}_sfGFP_KanR] was transformed with the individual libraries and single cfu were transferred to deep well plates and evaluated as described above (>90 cfu per site). For isobutanol, two additional rounds of evolution were conducted as described above but with more restrictive FACS-gating conditions and a negative pre-enrichment round to remove variants with a high basal biosensor output (Supplementary Figure 3-3). The second epPCR library was created based on pCK01_P_{alkS}_alkS_T1135C (coding for S379P, size of 1.2*10⁵ variants and 1.6*10⁵ re-transformants). Site-directed mutagenesis and variant evaluation was carried out at positions K115+A117, T336, R355, and T591 using primers AlkS_QC_residue_NNK_fwd and AlkS_QC_residue_NNK_rev with pCK01_P_{alkS}_alkS_T1135C as template. The third epPCR library was created based on pCK01_P_{alkS}_alkS_A1006T_G1064A_T1135C (coding for T336S, R355H, S379P, 5.1*10⁵ variants and 5.8*10⁶ re-transformants). Amino acid specific site-directed mutagenesis and variant evaluation was carried out as described above for D17E, I19V, R110G, V171L, V376L, and R397G using primers AlkS_residue_fwd and AlkS_residue_rev with template pCK01_P_{alkS}_alkS_A1006T_G1064A_T1135C.

Single-cell biosensor characterization and AlkS structure modeling

In general, the selected biosensor variants were cultivated in deep well plates and analyzed in a plate reader as described above. For initial biosensor circuit characterization the cells were incubated with alcohols for 6 h and in all other experiments for 16 h. Additional single cell data was obtained from diluted cultures (OD₆₀₀ of 0.05 in PBS) with a Beckman Coulter Cytoflex S (Brea, CA, USA) or a Becton Dickinson LSRFortessa for fitting of biosensor transfer functions. The sfGFP fluorescence was measured at an excitation wavelength of

488 nm and emission with a 525/40 nm or 530/30 nm band pass filter, respectively. All cultivations were carried out in biological triplicates unless stated otherwise. Gating was based on forward scattering area (FSC-A) vs side scattering area (SSC-A, polygon-shaped gate, selecting the overall cell population) as well as side scattering height (SSC-H) vs SSC-A (rectangle-shaped gate, selecting single cells). Samples for alcohol analysis were taken before addition to the biosensor culture and after 16 h of incubation. For all alcohols we observed neither chemical impurities nor significant conversion or degradation during the cultivations (Supplementary Figure 3-10). For single cell data evaluation and visualization FlowJo software was used (FlowJo LLC, Ashland, OR, USA). Transfer functions were fitted with a non-linear regression model as four-parameter dose-response curves (Prism 8.4.2, GraphPad, San Diego, CA, USA) and Z' values as described previously¹⁵⁰ (more details are available in Supplementary Figure 3-9). The AlkS mutations corresponding to the characterized biosensor variants were mapped to a structural model based on the crystal structure of a signal transduction ATPase of *Pyrococcus horikoshii* (UniProtKB O58663, 3.40 Å resolution) created with the Phyre2 web tool²²⁴ (72% sequence coverage with 100% confidence).

KivD SSM library screening for isopentanol production by whole-cell conversion

The parental plasmid p631_araC_P_{BAD}_kivDyqhD was created in two steps from pIPO2¹⁸⁴. First, an araC_P_{BAD}_kivD_yqhD_alsS_ilvC_ilvD fragment was PCR-amplified from template pIPO2 with primers p_AvrII_araC_isoalcohol_forward and p_NotI_araC_isoalcohol_reverse, digested with AvrII and NotI and subsequently ligated into the multiple cloning site of pSEVA631 cut open with the same enzymes, resulting in p631_araC_P_{BAD}_isobutanol. Second, the alsS_ilvC_ilvD part was removed by PCR amplification of a corresponding DNA fragment with primers pkivD_yqhD_NheI_fwd and pkivD_yqhD_NheI_rev using p631_araC_P_{BAD}_isobutanol as template. This fragment was subsequently cut with NheI and self-ligated, resulting in p631_araC_P_{BAD}_kivDyqhD. For PCR-based SSM at KivD amino acid position V461 primers pkivD_V461X_fwd and pkivD_V461X_rev were used (NNK codon) with plasmid p631_araC_P_{BAD}_kivDyqhD as template. For V461A and V461D mutations primers kivD_V461A_fwd and kivD_V461A_rev or kivD_V461D_fwd and kivD_V461D_rev were used, respectively. The corresponding 20 plasmids encoding the full SSM library as well as an empty plasmid control (pSEVA631) were used for transformation of the sensor strain *E. coli* DH5α [pCK01_P_{alkS}_alkS_L401G; pAB_P_{alkB}_sfGfp_KanR]. The screening was carried out as described for single-cell biosensor characterization, but 0.2% (v/v) substrate (4-methyl-2-oxopentanoic acid) and

0.2% (w/v) L-arabinose were added concomitantly to the cultures instead of alcohol (percentages represent final amounts in the culture). Samples for optical measurements and product analysis were taken after 20 h of incubation. The crystal structure of KdcA (PDB 2VBF) was used in order to visualize the conserved active center of KivD¹⁸⁵ with PyMOL 3.2.3 (Schrodinger, New York, NY, USA).

Isoalcohol pathways and combinatorial library for isopentanol production

For genomic pathway integration, an *araC_P_{BAD}_kivD_yqhD_alsS_ilvC_ilvD* fragment was amplified from pIPO2 with primers p_araC_ISO2_fwd_NheI and p_araC_ISO2_rev_NotI and integrated into a pBG fragment via restriction and ligation. The pBG fragment was amplified from pBG14g²¹⁵ with primers pTn7_upstream_NotI_fwd and pTn7_rev_NheI. The resulting pBG_araC_P_{BAD}_isobutanol plasmid was co-transformed with pTNS2 for genomic integration in three different *E. coli* strains: 10β(*mCherry*), BW25113, and BW25113 Δ*ilvE*²¹⁴. For *E. coli* BW25113 Δ*ilvE* the KanR resistance cassette was removed with pCP20 as described previously²³⁹. The correct insertion of the isobutanol pathway as well as KanR removal were verified by whole genome NGS (Novogene, Cambridge, UK). *E. coli* PIR2 was used for pBG plasmid (*ori* R6K) cloning and maintenance.

For creation of p221_araC_P_{BAD}_leuABCD the *leuABCD* operon of *E. coli* BW25113 was amplified with primers pleuABCD_pBAD_fwd and pleuABCD_MCS_rev from genomic DNA, *araC_P_{BAD}* (incl. RBS) was amplified from pNG413.1¹⁸⁴ with primers pBAD_MCS_fwd and pBAD_leuABCD_rev, and pSEVA221 was digested with KpnI and Sall. The corresponding DNA parts were assembled by three-fragment ITA.

For combinatorial *leuA* library generation by restriction and ligation, first p221_araC_P_{BAD}_leuABCD template DNA was amplified with primers p221_araC_Pbad_blunt_fwd and p221_araC_Pbad_KpnI_rev. The synthetic combinatorial *leuA* library fragment with an NGGANG in the RBS consensus motif and a degenerate NNK codon for position V462 (1.5 kb; GeneArt, Thermo Fischer Scientific) was amplified with primers pleuA_GeneArt_KpnI_short_fwd and pGeneART_leuA_lib_rev. The primers for blunt-end ligation were phosphorylated with T4 polynucleotide kinase prior to PCR (p221_araC_Pbad_blunt_fwd and pGeneART_leuA_lib_rev). The two fragments were separately digested with KpnI and subsequently combined by ligation. This library DNA was used for transformation of *E. coli* 10β(*mCherry*), cfu were pooled, plasmid DNA isolated, and then stored as a plasmid library pool (1.0*10⁵ transformants). The distribution of the

individual degenerate library parts in the plasmid pool was checked by NGS (Novogene, Cambridge, UK; Supplementary Figure 3-20).

The plasmid library pool was used for transformation of *E. coli* BW25113 $\Delta ilvE$ gISO2 [pCK01_ P_{alkS} _ $alkS$ _L401G pAB_ P_{alkB} _ $sfgfp$ _AmpR] and the resulting *E. coli* BW25113 $\Delta ilvE$ gISO2 biosensor strain library cultured on agar plates (total of 6.0×10^4 transformants). The correct assembly of library plasmids was checked by cPCR with primers pIP_seq_araC_rev and pPBAD_rev with *Taq* DNA polymerase. The corresponding PCR product (634 bp) was cut with KpnI (260 bp + 374 bp) in order to distinguish the library plasmid from the template plasmid (no KpnI site). Translation initiation rates for RBS were predicted per web-based RBS Calculator²²¹ accessed via DeNovoDNA (State College, PA, USA.).

Automation-based high-throughput screening of combinatorial library

Single cfu of the *E. coli* BW25113 $\Delta ilvE$ gISO2 biosensor strain library (described above) were first cultured in 384-well master plates in LB medium (30°C, 350 rpm, 16 h) and subsequently replicated in M9-based production medium additionally supplemented with 0.2% (w/v) L-arabinose for induction and actual screening (30°C, 350 rpm, 20 h). Plates were closed with peelable, clear heat seals to prevent evaporation (PlateLoc thermal sealer; Agilent, Santa Clara, CA, USA). All steps run on a customized automation-platform based on two EVO200 robots (Tecan) are described in more detail below. Biosensor measurements after the final incubation were carried out with a plate reader as described above. The obtained raw data was directly used for k-means clustering with R²⁴⁰ (version 4.0.3) and the number of clusters determined by plotting the ratio of “within sum of squares/between sum of squares”²⁴¹ (see Supplementary Figure 3-18). Briefly, the algorithm of Hartigan and Wong was used²¹⁹, with three clusters (cluster=3) and twenty random starts (nstart = 20). For ggplot2-based data visualization, the factoextra R package was used. Strain variants obtained from all clusters were re-grown from the master cultures on LB agar plates, single cfu cultured in deep well plates (n=3) and subsequently diluted in M9 production medium with 0.2% (w/v) L-arabinose as described above. After 24 h cell densities and biosensor outputs were measured as well as samples taken for metabolite analysis. The full workflow is shown in Supplementary Figure 3-16. For Sanger sequencing of the degenerate library sites a corresponding DNA fragment was amplified by colonyPCR with primers pIP_seq_araC_rev and pGeneART_leuA_lib_rev and subsequently sequenced with primers pBAD-for and pFP14_leuABCD_2.

Automated colony picking

For robotic handling, agar plates were prepared in Omnitray single well plates with a lid. For reliable colony picking, the agar inside the plates had to be perfectly level and without a meniscus at the edges. This was achieved by adhering to the following protocol. LB agar was autoclaved and allowed to cool to approximately 60 °C before antibiotics were added. An aliquot of 40 mL of was added to each plate using a serological pipette. The plates were covered with lids and left for one hour at room temperature to allow the agar to solidify. To dry the plates, they were left without lid under a laminar flow hood for 30 minutes. Plates were stored inside a plastic bag at 4°C until further use. The *E. coli* BW25113 $\Delta ilvE$ gISO2 biosensor strain library was plated at an average density of 280 cfu per plate, corresponding to 2.5 cfu per cm². Colony picking was performed on a Tecan EVO 200 robotic platform, controlled by the EVOware standard software (Tecan). Inputs to the colony picking method were provided using a custom Excel input sheet. The platform was equipped with a Pickolo colony picking module (SciRobotics Ltd., Kfar Saba, Israel) consisting of a light table, camera and software. The Pickolo module was used to take images of the agar plates and identify the positions of the colonies. A custom image processing script, implemented in Fiji²⁴², was used to correct the images of the agar plates for uneven illumination before identification of colonies. To ensure picking of single colonies, the minimum distance between colonies was set to approximately 2 mm. The robotic platform was equipped with a liquid displacement pipetting system using ultrapure water as system liquid. The pipetting system featured eight stainless steel fixed tips connected to 2500 μ L dilutors, and two tip wash stations. Before pipetting media or picking of colonies, the tips were decontaminated as follows. The tips were rinsed from the inside (3 mL) and outside (4 mL) in wash station one. Volumes of 2500 μ L (when pipetting media) or 200 μ L (when picking colonies) of 2% sodium hypochlorite solution were aspirated from a 100 mL trough (Tecan) into each tip and incubated for ten seconds. The hypochlorite solution was then dispensed back into the 100 mL trough. To clear the tips of residual hypochlorite solution they were rinsed from the inside (3 mL) and outside (4 mL) in wash station two. To contain the microbe containing aerosols that emerge during use of the wash stations, the wash stations were equipped with a custom made lid. The lid was manufactured from acrylic sheets using a laser cutter. Colonies were picked into clear 384-well F-bottom plates (Greiner). Each well of the plates was automatically filled with 40 μ L of LB media before colony picking started. For picking, 20 μ L of media was aspirated into a freshly decontaminated tip, the tip was brought in contact with the colony, and the cells transferred to the 384-well plate by dispensing 20 μ L. To ensure

that no cross-contamination occurred during colony picking process, 24 wells in each plate were not inoculated with a colony and served as negative control wells. The platform allowed picking of about 200 colonies per hour, and up to 1440 colonies were picked without user interaction in a single run. All method files, the design files of the wash station lid, the Pickolo colony picking profile (*.prf file) and the Excel input sheet are available online from github (<https://github.com/gregorwschmidt/tecan-colonypickingandcounting>).

Automated replication of libraries into fresh media

Replication of the picked libraries into fresh media was performed on a Tecan EVO 200 robotic platform, controlled by the EVOware standard software. The robotic platform was equipped with a liquid displacement pipetting system using ultrapure water as system liquid. The pipetting system featured eight teflon-coated fixed tips connected to four 1000 μ L and four 50 μ L dilutors, and two tip wash stations. Before pipetting media, the tips were decontaminated with 2% sodium hypochlorite as described above. The platform was additionally equipped with a PCR cycler (Biometra TRobot, Analytik Jena, Jena, Germany). Transfer of master cultures into screening cultures was achieved using a 96-pin replicator (Scinomix, Earth City, MO, USA). First, all wells of the screening plates (black 384-well clear F-bottom plates, Greiner) were automatically filled with 60 μ L production media using the pipetting system of the robot. Then, cells were transferred by dipping the 96-pin replicator first into the master plate and then into the screening plates using the robotic manipulator arm of the system. To replicate a single 384-well plate four distinct pin replication steps were necessary. Between pin replication steps, the 96-pin replicator was decontaminated as follows. Using the robotic manipulator arm of the system, the pin replicator was dipped into an Omnitray plate filled with 70 mL of 2% hypochlorite solution. The pin replicator was moved inside the hypochlorite solution for 20 seconds to facilitate washing. Two more wash steps in ultrapure water and 95% ethanol were performed in the same fashion. Then, the pin replicator was placed inside a skirted 96-well PCR plate (FrameStar, 4titude Ltd., Surrey, UK) which was located in the PCR cycler and preheated to 80°C. The pin replicator was left to dry inside the heated PCR plate for at least three minutes, before it was used for the next replication step. Replication of four 384-well master plates into four 384-well screening plates with the automated method took 1.8 hours. All method files and the Excel input sheet are available online from github (<https://github.com/gregorwschmidt/tecan-libraryreplication>). Details and images of the crucial automation steps are available in Supplementary Figure 3-17.

Isopentanol production strains and cultivation in deep well plates

For curing of the biosensor plasmids (but not the p221_ *araC*_P_{BAD}_ *leuABCD*_{max} plasmid variant with AGGAGG consensus motif and V462F), the best performing isopentanol strain obtained from the screen was cultivated in LB supplemented with kanamycin at 42°C for 16 h. The plasmid DNA was isolated and used for transformation of *E. coli* 10β(*mCherry*). The re-isolated p221_ *araC*_P_{BAD}_ *leuABCD*_{max} plasmid was subsequently used as the template for PCR-based mutagenesis, introducing either a weak RBS (plasmid p221_ *araC*_P_{BAD}_ *leuABCD*_{weakRBS}) or back-mutation to V462V at the FBR site (p221_ *araC*_P_{BAD}_ *leuABCD*_{noFBR}). For the corresponding PCRs, primers *leuA*_weak_RBS_fwd and *leuA*_weak_RBS_rev or *leuA*_G462_fwd and *leuA*_G462_rev were used, respectively. Next, the three different p221_ *araC*_P_{BAD}_ *leuABCD* variants were used for separate transformation of *E. coli* BW25113 Δ *ilvE* gISO2. The resulting strains as well as the parental isobutanol production strain were cultivated in 2 mL of LB in deep well plates in 24-square well format for 10 h and re-inoculated in 2 mL M9-based production medium (5% (v/v) inoculum) for 16 h (n=4). These precultures were diluted to an OD₆₀₀ of 0.2 in 2 mL of the same medium but supplemented with 0.2% (w/v) L-arabinose and the plates sealed with aluminum foil. Samples for metabolite and glucose analysis were taken after 20 h of cultivation.

Metabolite and glucose analysis

For quantitative alcohol determination in deep well plate experiments in 96-well format 700 μ L of each sample were spun down (21,130 rcf, 4°C, 10 min) and 500 μ L of the culture supernatant were diluted in 500 μ L water. In experiments in 24-well format, 2 mL of each sample were spun down and 1 mL of undiluted supernatant used. Next, 100 μ L of internal standard solution (1% (v/v) n-pentanol in water) was added and the sample briefly vortexed, except for qualitative alcohol analysis experiments during biosensor characterization (Supplementary Figure 3-10). Subsequently 500 μ L methyl tert-butyl ether (SupraSolv, Merck, Darmstadt, Germany) was added and the mixture thoroughly vortexed. If necessary, samples were briefly centrifuged for complete phase separation. For gas chromatography (GC) analysis, 400 μ L of the organic phase were transferred into glass vials closed with PTFE snap-caps (Wicom, Maienfeld, Switzerland). The GC system consisted of a 6890N GC-FID with a 7683 auto-sampler and a DB-WAX UI capillary column (30 m with 0.25 mm diameter; all Agilent) with helium as the carrier gas (2 mL min⁻¹; PanGas, Dagmersellen, Switzerland). Samples were injected in splitless mode with a volume of 1 μ L. The temperature of the injector and the FID were maintained constant at 225°C and 250°C,

respectively. The oven temperature was initially held at 40°C for five minutes. Next, the temperature was increased to 120°C at a rate of 15°C min⁻¹, followed by heating to 230°C at a rate of 50°C min⁻¹ and held constant for four minutes. For glucose analysis, 500 µL undiluted supernatant samples in snap cap vials were analyzed on an HPLC system (Agilent 1200, 10 µL injection per sample) with a refractive index detector and a Metab-AAC column (30 cm x 7.8 mm, 10 µm particle size; Inera, Dueren, Germany) at 40°C using an isocratic protocol (5 mM H₂SO₄ as mobile phase, flow of 600 µL min⁻¹). Concentrations were determined by extrapolation from standard curves based on samples of diluted pure compounds.

3.6 Acknowledgements

We thank M. Otto and F. Pfeifer for help with metabolite analytics; D. Gerngross for help with automated colony picking; T. Roberts for critically reading the manuscript; S. Schmitt for help with NGS data analysis; R. Nitschel, R. Takors and Ingenza Inc. for providing plasmids pIP02 and pNG413.1; V. de Lorenzo and the members of his laboratory for providing pSEVA plasmids; S. Koebbing and L. Blank for providing pBG14g; T. Lopes and the D-BSSE Single Cell Facility for excellent assistance with FACS experiments. We acknowledge financial support from The European Union Project EmPowerPutida (grant ID 635536) for M.O.B and are indebted to the whole project consortium for most valuable feedback and discussions.

3.7 Author contributions

M.O.B. conceived and carried out the experimental work as well as data analysis. L.P. contributed to biosensor characterization and alcohol production (supervised by M.O.B and S.P, summarized in a written MSc Thesis: Platz, Lukas (2019) Optimization of biosensor-based high-throughput screening systems for short-chain alcohols, Universität Heidelberg). G. M. constructed initial biosensor plasmids and cloning strains. G. W. S developed the automation-based screening platform. M.O.B., G. W. S. and S.P. wrote the manuscript, all authors reviewed and approved the manuscript. S.P. provided the project outline, supervision and resources.

3.8 Supplementary information

Supplementary Table 3-1 Three well-described AlkS transcriptional regulators of various alkane degrading micro-organisms obtained from UniProtKB-Swiss-Prot (accessed 6. April 2021). 23 additional entries based on un-reviewed, computationally analysed entries are not shown.

Organism	Residues (uniprot#)	Identity	DNA source	References
<i>Pseudomonas putida</i> GPo1 also known as <i>P. oleovorans</i>	882 (P17051)	100%	OCT plasmid	193, 194 This study
<i>Pseudomonas putida</i> P1	883 (Q9L4M7)	84%	genomic	195, 204
<i>Alcanivorax borkumensis</i>	872 (Q0VKZ4)	31%	genomic	196

AlkS variant	Amino acids exchanged	Occurrence	Comment	Round of directed evolution
WT	Q410K	n/a	as compared to uniprot ref. P17051	
1	A463T, G753R	1x		1
2	V11D, D40E, <u>L401S</u> , L469H	1x		
3	M103K, A117V, S379P, L542Q	2x		
4	R50K, A375T, F538L, Q767R	1x		
5	<u>K183I</u> , <u>L401S</u>	1x	mutation upstream of start codon, additional MGARIA peptide	
6	Y302F, S377G, S627G, V653I	1x		
7	I31V, K451N, K595R	1x		
8	K183R, E267V, I300V, Y307F, Q390L	1x		
Target alcohol				
nC4-OH				
1	P10L, C47S, A375V, E504Q	2x		1
2	M103K, <u>A117V</u> , <u>S379P</u> , L542Q	3x	same as #3 above	
3	R50K, <u>A375T</u> , F538L, Q767R	1x	same as #4 above	
isoC5-OH				
1	M103K, <u>A117V</u> , <u>S379P</u> , L542Q	4x	same as #3 and #2 above (for nC4-OH and isoC5-OH, respectively)	2
1	<u>R355H</u> , (<u>S379P</u>), <u>T591A</u>	1x		
2	<u>E102G</u> , <u>K115T</u> , <u>T336S</u> , (<u>S379P</u>)	1x		
3	(<u>S379P</u>), V513M, E567G, L580V	1x	high basal fluorescence output, discarded	3
1	<u>D17E</u> , I19V, R110G, V171L, (<u>T336S</u>), (<u>R355H</u>), V376L, (<u>S379P</u>), R397G	1x		

Supplementary Table 3-2 Overview of AlkS variants found in FACS-based screening of error-prone *alkS* strain libraries. Verified hot-spot sites are indicated in red, amino acids targeted by site-directed mutagenesis are underlined. For round two and three, parentheses indicate the parental mutations. Mutation Q410K of wild-type (WT) AlkS is not listed again for mutant variants. nC4-OH, n-butanol; isoC5-OH, isopentanol; isoC4-OH, isobutanol.

Supplementary Table 3-2 Overview of strains and plasmids used in this study. AmpR, ampicillin resistance; AprR, apramycin resistance, CmR, chloramphenicol resistance; GmR, gentamycin resistance; KanR, kanamycin resistance; StrepR, streptomycin resistance.

Strain or plasmid name	Relevant characteristics	Source/Reference
Strains		
<i>Escherichia coli</i> NEB 10 β	$\Delta(ara-leu)$ 7697 <i>araD139 fhuA</i> $\Delta lacX74 galK16 galE15 e14-$ $\phi 80dlacZ\Delta M15 recA1 relA1 endA1$ <i>nupG rpsL (StrR) rph spoT1</i> $\Delta(mrrhsdRMS-mcrBC)$	New England Biolabs, #C3019I
<i>E. coli</i> 10 β (<i>mCherry</i>)	As above, but <i>allose</i> operon replaced with constitutive <i>mCherry</i> expression cassette, $\Delta alsRBACEK::\lambda P_L-mCherry$	This study
<i>E. coli</i> PIR2	F ⁻ $\Delta(argF-lac)169 rpoS(Am) robA1$ <i>creC510 hsdR514 endA recA1</i> <i>uidA(\Delta M1ul)::pir⁺</i>	Invitrogen
<i>E. coli</i> DH5 α	F ⁻ $\Phi 80lacZ\Delta M15 \Delta(lacZYA-argF)$ <i>U169 recA1 endA1 hsdR17(rk⁻, mk⁺)</i> <i>phoA supE44 thi-1 gyrA96 relA1 λ</i>	Invitrogen
<i>E. coli</i> BW25113	KEIO collection, <i>rrnB3 \Delta lacZ4787</i> <i>hsdR514 \Delta(araBAD)567</i> $\Delta(rhaBAD)568 rph-1$	²¹⁴
<i>E. coli</i> BW25113 $\Delta ilvE$	As above, but $\Delta ilvE$	²¹⁴
<i>E. coli</i> BW25113 $\Delta ilvE$ gISO2	As above, but <i>araC</i> $P_{BAD_kivD_yqhD_alsS_ilvC_ilvD}$ integrated downstream of <i>glmS</i>	This study
Plasmids		
pCK01	Expression vector, pSC101, CmR	²³⁵

Continues on next page

Strain or plasmid name	Relevant characteristics	Source/Reference
pESM7	Expression vector containing P_{alkS_alkS}	193
pSEVA231	MCS, pBBR1, KanR	195
pRK793	pBR322/ <i>rop</i> , AmpR	236
pCK01_ P_{alkS_alkS}	P_{alkS_alkS} , pSC101, CmR	This study, addgene ID 166502 (with AlkS L401G)
pCK01_ P_{weak_alkS}	P_{J23117} , <i>alkS</i> , pSC101, CmR	This study
pCK01_ P_{medium_alkS}	P_{J23106} , <i>alkS</i> , pSC101, CmR	This study
pCK01_ P_{strong_alkS}	P_{J23100} , <i>alkS</i> , pSC101, CmR	This study
pCK01_empty_cargo	pSC101, CmR	This study
pET28_sfGFP	Expression vector containing <i>sfgfp</i> , pBR322, KanR	Sandia Biotech
pAB_ P_{alkB_sfgfp} _KanR	P_{alkB_sfgfp} , pBR322/ <i>rop</i> , KanR	This study, addgene ID 166503
pSEVA181	MCS, pUC, AmpR	195
pAB_ P_{alkB_sfgfp} _AmpR	P_{alkB_sfgfp} , pBR322/ <i>rop</i> , AmpR	This study
pNG413.1	<i>araC</i> _ P_{BAD_lacZ} , pBBR1, AprR	184
PIP02	<i>araC</i> _ $P_{BAD_kivD_yqhD_alsS_ilvC_ilvD}$, pBBR1, AprR	184
pSEVA221	MCS, RK2, KanR	195
p221_ <i>araC</i> _ $P_{BAD_leuABCD}$	<i>araC</i> _ $P_{BAD_leuABCD}$, RK2, KanR	This study

Continues on next page

Strain or plasmid name	Relevant characteristics	Source/Reference
p221_ <i>araC</i> _P _{BAD} _ <i>leuABCD</i> _{max}	As above, but strong RBS for <i>leuA</i> , LeuA G462F	This study
p221_ <i>araC</i> _P _{BAD} _ <i>leuABCD</i> _{weakRB}	As above, but weak RBS for <i>leuA</i> , LeuA G462F	This study
p221_ <i>araC</i> _P _{BAD} _ <i>leuABCD</i> _{noFBR}	As above, but strong RBS for <i>leuA</i> , LeuA G462G	This study
pSEVA631	MCS, pBBR1, GmR	195
p631_ <i>araC</i> _P _{BAD} _ <i>kivDyqh</i> <i>D</i>	<i>araC</i> _P _{BAD} _ <i>kivD_yqhD</i> , GmR	pBBR1, This study
p631_ <i>araC</i> _P _{BAD} _ <i>isobutan</i> <i>ol</i>	<i>araC</i> _P _{BAD} _ <i>kivD_yqhD_alsS_ilvC_il</i> <i>vD</i> , pBBR1, GmR	This study
pCP20	FLP1 recombinase, λ cl857 ⁺ , λ p _R Rep ^{ts} , AmpR, CmR	239
pBG14g	Tn7L/R, R6K, KanR, GmR	215
pBG_ <i>araC</i> _P _{BAD} _ <i>isobutano</i> <i>I</i>	Tn7L/R, <i>araC</i> _P _{BAD} _ <i>kivD_yqhD_alsS_ilvC_il</i> <i>vD</i> , R6K, KanR, GmR	This study
pTNS2	Tn7 transposase, R6K; AmpR	243

Supplementary Table 3-3 DNA oligomers used in this study.

Primer	Sequence (5' to 3')
AL1	GCGAATTCCTCGACCAACTGCTGCACG
AL2	CGCAAGCTTTACCAGCGCCAACAACAAC
SG_fwd	ATGGAATTCACACTTTACTTTTTAATTAG CAGGAGATTTAACATATGAGCAAAGGAGA AGAACTTTTC
SG_rev	CCTTTAATTAATTAGGATCCTTTGTAGAG CTCATC
pBR322_for-AscI	ATATATGGCGCGCCCCCGCCGCATCCATA CCGC
pBR322_rev-FseI	ATATATGGCCGGCCCCGTAGAAAAGATCA AAGG
PalkB	AAGACTAGTCTGGTTTTTCCAGCAGACGA CGGAGCAAAAACCTACCCGTAGGTGTAGTT GGCGCAAGCGTCCGATTAGCTCAGGTTTT AAGATGGAATTCACACATCGCCGACATCG TGCC
PalkB_rev	GGCACGATGTCGGCGATG
pCK01_J231XX alkS_fwd	ACCGTTGCTCTTGCGTTCG
pCK01_J231XX alkS_rev	CTTCTCGATCTTGCGCCTGG
pSEVA_abR_exchange_bb_fwd	GTCCTTTTCCGCTGCATAACCCTG
pSEVA_abR_exchange_bb_rev	GCCCGGCGGCAACC
pSEVA_abR_exchange_AbR_fwd	GGTTGCCGCGGGC
pSEVA_abR_exchange_AbR_rev	CAGGGTTATGCAGCGGAAAAGGAC

Continues on next page

Primer	Sequence (5' to 3')
AlkS_fwd_Gibson	GGCGCGAGAATAGCATAATG
AlkS_rev_Gibson	CAGCAATCTGCTTCCGTGTC
AlkS_fwd_Gibson_backbone	GACACGGAAGCAGATTGCTG
AlkS_rev_Gibson_backbone	CATTATGCTATTCTCGCGCC
pAlkS_fwd_1	CTGAAAGCACCGAAATGA
pAlkS_fwd_2	CCGGCCAAGCATATTTCA
pAlkS_fwd_3	AAATGCTTTTCGAGTGAGG
pAlkS_fwd_4	GGCATAACGGGCATAAATAA
AlkS_QC_K183_NNK_fwd	GTTCTCGCAGCTTNNKCTTGCAGGC
AlkS_QC_K183_NNK_rev	CTGCAAGMNNAAGCTGCGAGAACCC
AlkS_QC_A375_NNK_fwd	GACTGTCGCTGGNNKGTGAGCATGTCTG
AlkS_QC_A375_NNK_rev	CATGCTGACMNNCCAGCGACAGTCATTTCG
AlkS_QC_S377_NNK_fwd	CGCTGGGCAGTCNNKATGTCTGAGAG
AlkS_QC_S377_NNK_rev	CTCAGACATMNGACTGCCAGCGAC
AlkS_QC_S379_NNK_fwd	GCAGTCAGCATGNNKGAGAGAATAATTTT AGATTTGTCAATTTTCG
AlkS_QC_S379_NNK_rev	TATTCTCTCMNNCATGCTGACTGCCAGC G
AlkS_QC_L401_NNK_fwd	GAGACAGTGGCTGNNKGAGCTGCCGAAGC
AlkS_QC_L401_NNK_rev	GCTTCGGCAGCTCMNNCAGCCACTGTCTC
AlkS_QC_Q410_NNK_fwd	CAGGCCTGGCACNNKAAACCCATAGTGC
AlkS_QC_Q410_NNK_rev	CTATGGGTTTMNNGTGCCAGGCCTGCTTC
AlkS_QC_K115_A117_NNK_fwd	GTGAGCNNKCCTNNKCTCTTGCGAGACCT TG

Continues on next page

Primer	Sequence (5' to 3')
AlkS_QC_K115_A117_NNK_rev	CAAGAGMNNAGGMNNGCTCACACCCTCTC TCAC
AlkS_QC_T336_NNK_fwd	GAAATTACTTGGNNKGAAAATCCTGC
AlkS_QC_T336_NNK_rev	GATTTTCMNNCCAAGTAATTTCTCGTAG
AlkS_QC_R355_NNK_fwd	CTGGCATTGGNNKAGAGGTGAATACCAG
AlkS_QC_R355_NNK_rev	GTATTCACCTCTMNNCCAATGCCAGAAAG C
AlkS_QC_T591_NNK_fwd	CTGCTTGGACNNKTCAGAAG
AlkS_QC_T591_NNK_rev	GCTTTTCTTCTGAMNGTCCAAG
AlkS_D17E_rev	TTTGTTCCGCTCCGACCTTAGCGACCGGG AAATC
AlkS_D17E_fwd	AGCGGAACAAATTACGACTCTAGTAAGTG CCAAAGTTCATAGTTGC
AlkS_I19V_rev	TCGTAACTTGATCCGCTCCGACCTTAGCG ACCGGG
AlkS_I19V_fwd	TCAAGTTACGACTCTAGTAAGTGCCAAAG TTCATAGTTGCATATATCGGCC
AlkS_R110G_rev	CCTCTCCCACACGGCTGAATTTTACCATT TCGAAAGTTTCAAGCTG
AlkS_R110G_fwd	TGTGGGAGAGGGTGTGAGCAAGCCTGCGC TCTTG
AlkS_V171L_rev	CTGCAAGTGCAAACCTGATATTTTTTGGT GTATTTAACATAAAC
AlkS_V171L_fwd	TGCACTTGCAGGCAATACAATAAAAGGGT TCTCGC
AlkS_V376L_rev (w/o S379P)	TGCTGAGTGCCCAGCGACAGTCATTCGCC C

Continues on next page

Primer	Sequence (5' to 3')
AlkS_V376L_fwd (w/o S379P)	GGCACTCAGCATGTCTGAGAGAATAATTT TAGATTTGTCATTTTCGTC
AlkS_R397G_rev	ACTGTCCCAGCGCATCTATTTTCGCCCTGA CGAAATGAC
AlkS_R397G_fwd	GCTGGGACAGTGGCTGTTAGAGCTGCCGA AG
pleuABCD_pBAD_fwd	GTTATAAAAACATATGAGCCAGCAAGTCA TTATTTTCG
pleuABCD_MCS_rev	CTTGCATGCCTGCAGGTTAATTCATAAAC GCAGGTTGTTTTGC
pBAD_MCS_fwd	GAATTCGAGCTCGGTACGTAGTCAATAAA CCGGTG
pBAD_leuABCD_rev	GACTTGCTGGCTCATATGTTTTTATAACC TCCTTAG
p_AvrII_araC_isoalcohol_forward	ATACCTAGGGCTTCCGGTAGTCAATAAAC CGGTGGATC
p_AvrII_araC_isoalcohol_reverse	ATAGCGGCCGCGAGAATAGGAACTTCGAA CTGCAGGTC
pkivD_yqhD_NheI_fwd	ATAGCTAGCGTCGTGACTGGGAAAACCCT GGCG
pkivD_yqhD_NheI_rev	ATAGCTAGCGGCTTAGCGGGCGGCTTCG
pkivD_V461X_fwd	GATGGTTATACANNKGAAAGAGAAATTCA TGGACCAAATCAAAG
pkivD_V461X_rev	GTCCATGAATTTCTCTTTCMNNTGTATAA CCATCATTATTG

Continues on next page

Primer	Sequence (5' to 3')
kivD_V461A_fwd	GATGGTTATACAGCGGAAAGAGAAATTCA TGGACCAAATCAAAG
kivD_V461A_rev	GTCCATGAATTTCTCTTTCCGCTGTATAA CCATCATTATTG
kivD_V461D_fwd	GATGGTTATACAGATGAAAGAGAAATTCA TGGACCAAATCAAAG
kivD_V461D_rev	GTCCATGAATTTCTCTTTCATCTGTATAA CCATCATTATTG
p_araC_ISO2_fwd_NheI	AAAATAGCTAGCCGGTCACACTGCTTCCG GTAG
p_araC_ISO2_rev_NotI	AAAATAGCGGCCGCTAGGAACTTCGAACT GCAGGTCG
pTn7_upstream_NotI_fwd	TCCTCTAGAGTCGACCTGCAGGC
pTn7_rev_NheI	AAAATAGCTAGCAGACGTCTTGACATAAG CCTGTTC
pleuA_GeneArt_KpnI_short_fwd	GGCCTCTGAGTCGGTACCATTTCGAGCTCT A
pGeneART_leuA_lib_rev	CCGACGCCGTGGAAGC
p221_araC_Pbad_blunt_fwd	CCTGGCTACCGATATTGTCGAGTCATCTG CC
p221_araC_Pbad_KpnI_rev	CGAATGGTACCTCCCAAAAAACGGGTAT GGAGAACAG
pip_seq_araC_rev	CGAGTATCCCGGCAGC
pPBAD_rev	CGCAGCGAGCTAACGCACATAC
pBAD-for	ATGCCATAGCATTTTTTATCC

Continues on next page

Primer	Sequence (5' to 3')
pFP14_leuABCD_2	GTTAGTTAGCCAGATTTGTAATATG
leuA_weak_RBS_fwd	GCTCTATGGATGTTATAAAAACATATGAG CCAGCAAGTC
leuA_weak_RBS_rev	CTCATATGTTTTTATAACATCCATAGAGC TCGAATGGTACC
leuA_G462_fwd	CACCTGACCCAGCGCATCTTTACCG
leuA_G462_rev	TGCGCTGGGTCAGGTGGATATCGTC

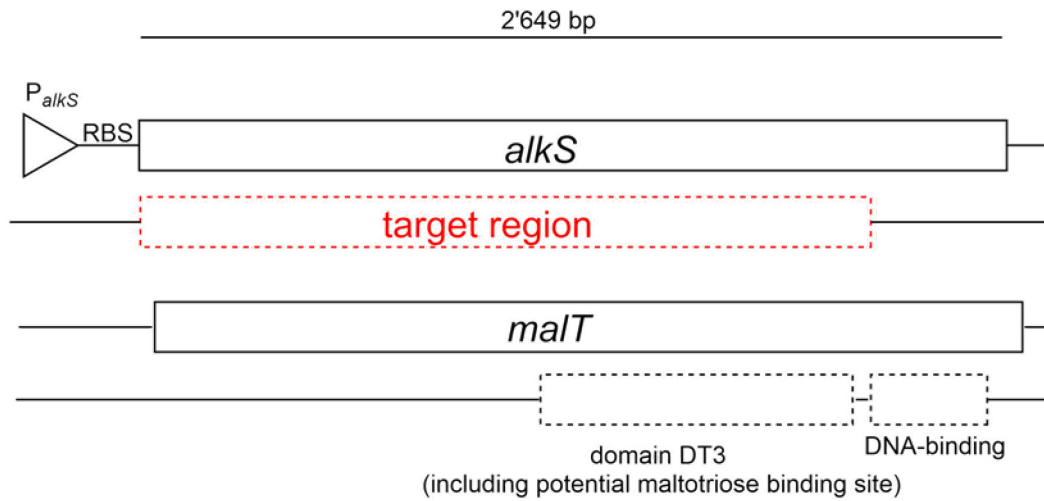
Supplementary Table 3-4 DNA parts synthesized for this study.

	Sequence (5' to 3')
BBa_J23100_alkS	CAAGGCCAGGCGCAAGATCGAGAAGGCCAAGGCTCAGGTGC GTGCCAAGGTTGAGCACCCGTTCTTGACGGCTAGCTCAGTC CTAGGTACAGTGCTAGCCCAGTCAGCTGGCGCGAGAATAGC ATAATGAAAATAATAATAAATAATGATTTCCCGGTTCGCTAA GGTCGGAGCGGATCAAATTACGACTCTAGTAAGTGCCAAAG TTCATAGTTGCATATATCGGCCAAGATTGAGTATCGCGGAT GGAGCCGCTCCCAGAGTATGCCTTTACAGAGCCCCACCTGG ATATGGGAAAACCGTTGCTCTTGCGTTTCGAGTGGC
BBa_J23106_alkS	CAAGGCCAGGCGCAAGATCGAGAAGGCCAAGGCTCAGGTGC GTGCCAAGGTTGAGCACCCGTTCTTTACGGCTAGCTCAGTC CTAGGTATAGTGCTAGCCCAGTCAGCTGGCGCGAGAATAGC ATAATGAAAATAATAATAAATAATGATTTCCCGGTTCGCTAA GGTCGGAGCGGATCAAATTACGACTCTAGTAAGTGCCAAAG TTCATAGTTGCATATATCGGCCAAGATTGAGTATCGCGGAT GGAGCCGCTCCCAGAGTATGCCTTTACAGAGCCCCACCTGG ATATGGGAAAACCGTTGCTCTTGCGTTTCGAGTGGC
BBa_J23117_alkS	CAAGGCCAGGCGCAAGATCGAGAAGGCCAAGGCTCAGGTGC GTGCCAAGGTTGAGCACCCGTTCTTGACAGCTAGCTCAGTC CTAGGGATTGTGCTAGCCCAGTCAGCTGGCGCGAGAATAGC ATAATGAAAATAATAATAAATAATGATTTCCCGGTTCGCTAA GGTCGGAGCGGATCAAATTACGACTCTAGTAAGTGCCAAAG TTCATAGTTGCATATATCGGCCAAGATTGAGTATCGCGGAT GGAGCCGCTCCCAGAGTATGCCTTTACAGAGCCCCACCTGG ATATGGGAAAACCGTTGCTCTTGCGTTTCGAGTGGC
Combinatorial_library_1e uA	GCGTAACAAGGCAATGCTGGAAGGCCTCTGAGTCATTCGAG CTCTANGGANGTTATAAAAACATATGAGCCAGCAAGTCATT ATTTTCGATACCACATTGCGCGACGGTGAACAGGCGTTACA GGCAAGCTTGAGTGTGAAAGAAAAACTGCAAATTGCGCTGG CCCTTGAGCGTATGGGTGTTGACGTGATGGAAGTCGGTTTC CCCGTCTCTTCGCCGGGCGATTTTGAATCGGTGCAAACCAT CGCCCGCCAGGTTAAAAACAGCCGCGTATGTGCGTTAGCTC

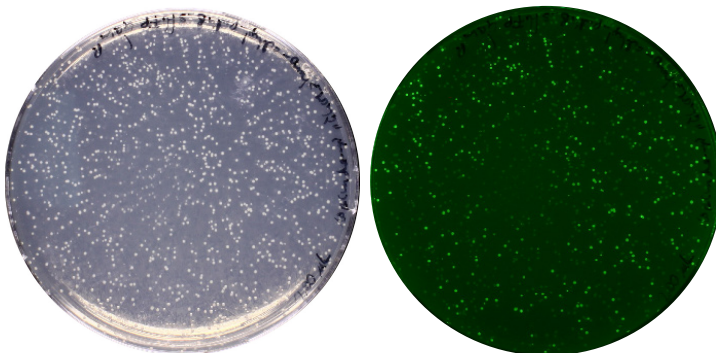
Continues on next page

Sequence (5' to 3')

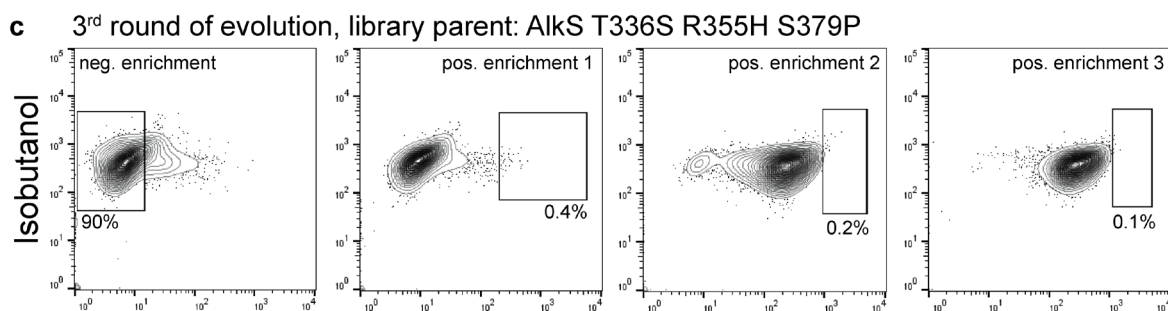
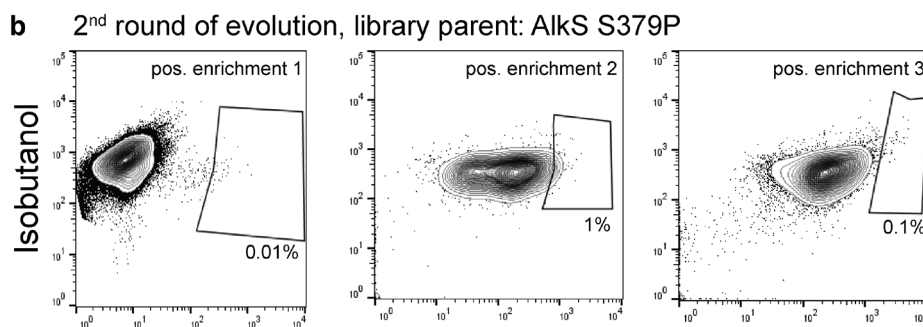
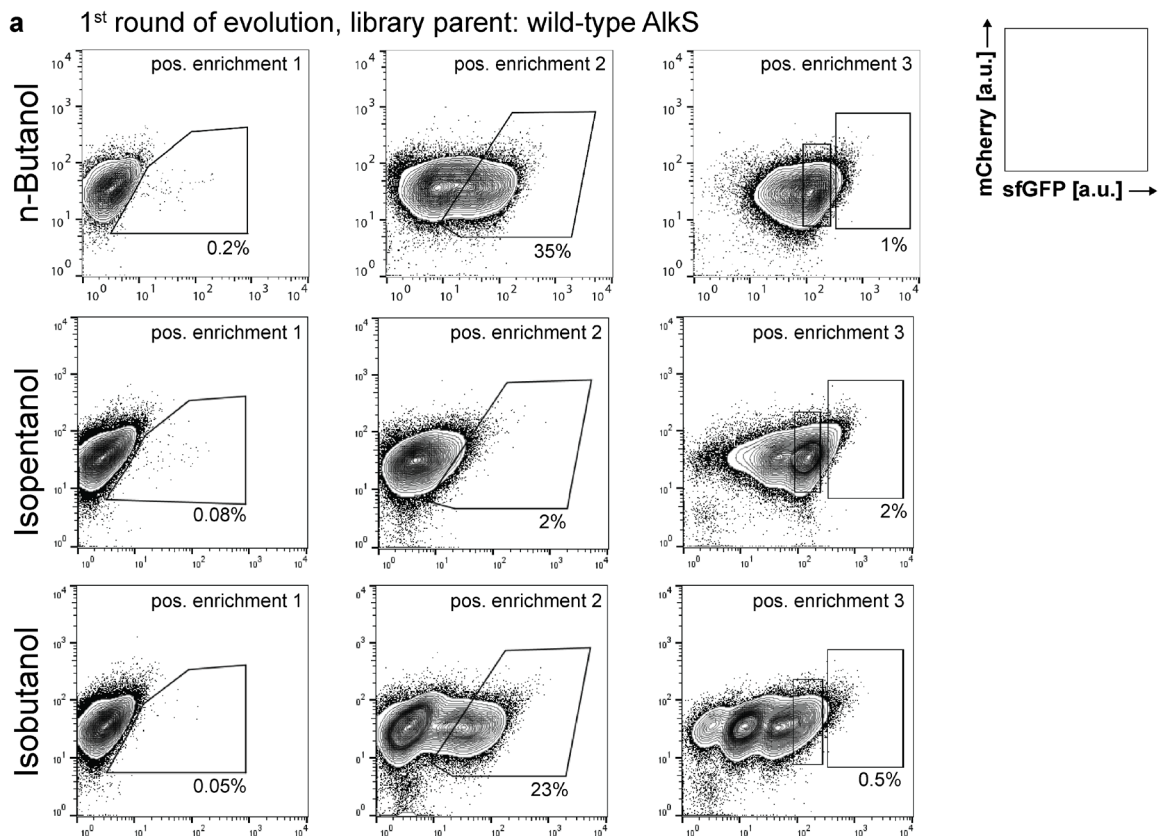
GCTGCGTGGAAAAAGATATCGACGTGGCGGCCGAATCCCTG
AAAGTCGCCGAAGCCTTCCGTATTCATACCTTTATTGCCAC
TTCGCCAATGCACATCGCCACCAAGCTGCGCAGCACGCTGG
ACGAGGTGATCGAACGCGCTATCTATATGGTGAAACGCGCC
CGTAATTACACCGATGATGTTGAATTTTCTTGCGAAGATGC
CGGGCGTACACCCATTGCCGATCTGGCGCGAGTGGTCGAAG
CGGCGATTAATGCCGGTGCCACCACCATCAACATTCCGGAC
ACCGTGGGCTACACCATGCCGTTTGAGTTCGCCGGAATCAT
CAGCGGCCTGTATGAACGCGTGCCTAACATCGACAAAGCCA
TTATCTCCGTACATACCCACGACGATTTGGGCCTGGCGGTC
GGAAACTCACTGGCGGCGGTACATGCCGGTGCACGCCAGGT
GGAAGGCGCAATGAACGGGATCGGCGAGCGTGCCGGAAACT
GTTCCCTGGAAGAAGTCATCATGGCGATCAAAGTTCGTAAG
GATATTCTCAACGTCCACACCGCCATTAATCACCAGGAGAT
ATGGCGCACCAGCCAGTTAGTTAGCCAGATTTGTAATATGC
CGATCCCGGCAAACAAAGCCATTGTTGGCAGCGGCGCATTC
GCACACTCCTCCGGTATACACCAGGATGGCGTGCTGAAAAA
CCGCGAAAACCTACGAAATCATGACACCAGAATCTATTGGTC
TGAACCAAATCCAGCTGAATCTGACCTCTCGTTCGGGGCGT
GCGGCGGTGAAACATCGCATGGATGAGATGGGGTATAAAGA
AAGTGAATATAATTTAGACAATTTGTACGATGCTTTCCTGA
AGCTGGCGGACAAAAAGGTCAGGTGTTTGATTACGATCTG
GAGGCGCTGGCCTTCATCGGTAAGCAGCAAGAAGAGCCGGA
GCATTTCCGTCTGGATTACTTCAGCGTGCAGTCTGGCTCTA
ACGATATCGCCACCGCCCGCGTCAAACCTGGCCTGTGGCGAA
GAAGTCAAAGCAGAAGCCGCCAACGGTAACGGTCCGGTCGA
TGCCGTCTATCAGGCAATTAACCGCATCACTGAATATAACG
TCGAACTGGTGAAATACAGCCTGACCGCAAAGGCCACGGT
AAAGATGCGCTGNNKCAGGTGGATATCGTCGCTAACTACAA
CGGTGCGCGCTTCCACGGCGTCGG



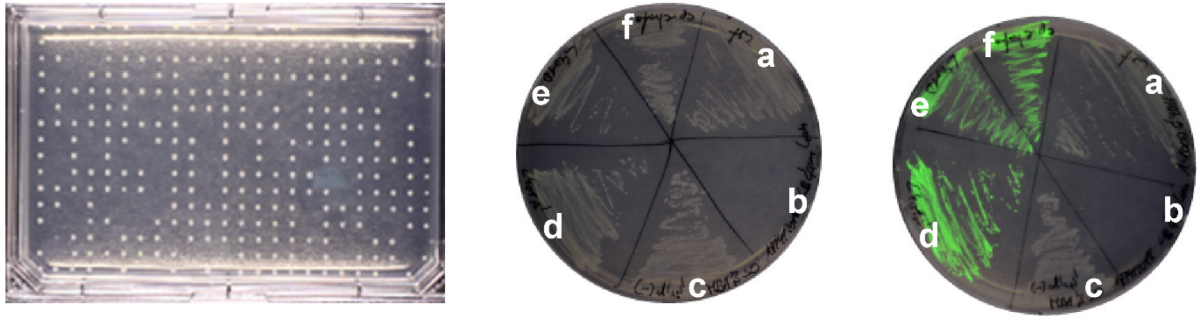
Supplementary Figure 3-1 Target region for error-prone PCR mutagenesis of transcription factor gene *alkS*. For comparison, we include the sequence of the gene for the maltose binding protein, *malt*, for whose gene product the functions of some regions have been clarified²⁴⁴. In order to avoid unintentional variation in expression strength as well as DNA-binding, the regulatory region upstream (and including) the start codon as well as a putative C-terminal DNA-binding region (helix-turn-helix motif) were excluded, respectively. Amino acid sequence alignments of *alkS* and *malt* with BLASTp and COBALT tools, accessed via the National Center for Biotechnology Information web page.



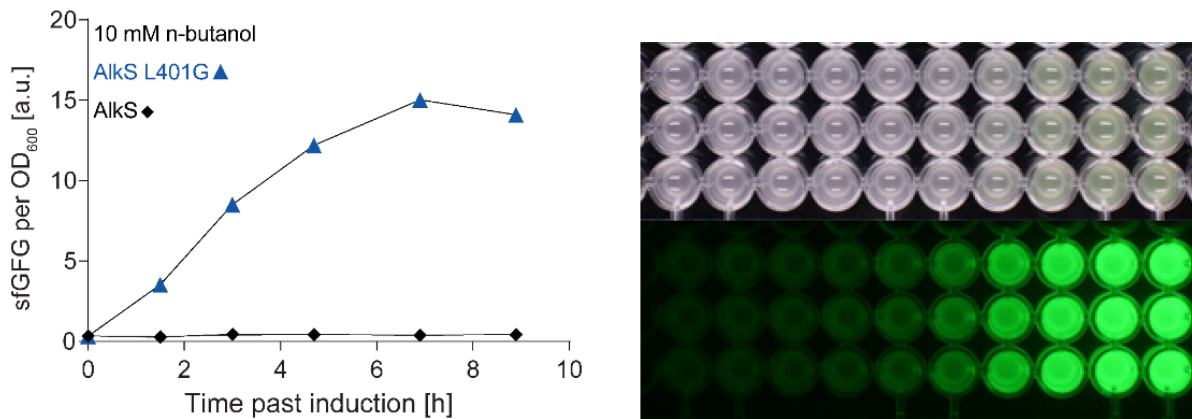
Supplementary Figure 3-2 Error-prone PCR mutagenesis library in *E. coli* 10 β (*mCherry*) [pCK01_*P_{alkS}*_*alkS* pAB_*P_{alkB}*_*sfgfp*_KanR] biosensor strain. Cultivation with M9 glucose medium supplemented with 0.1% (v/v) n-pentanol (16 h, 37°C). Green fluorescence indicates functional *alkS* expression, either because of proper induction of an AlkS variant by n-pentanol or because a constitutively active AlkS variant had been generated during mutagenesis (about 29% of total cfu, 333/1140). Left: Bright field image with UV cut filter, right: 470 nm excitation and 545 nm peak filter (Dark Hood DH-50, Biostep, Burkhardtsdorf, Germany).



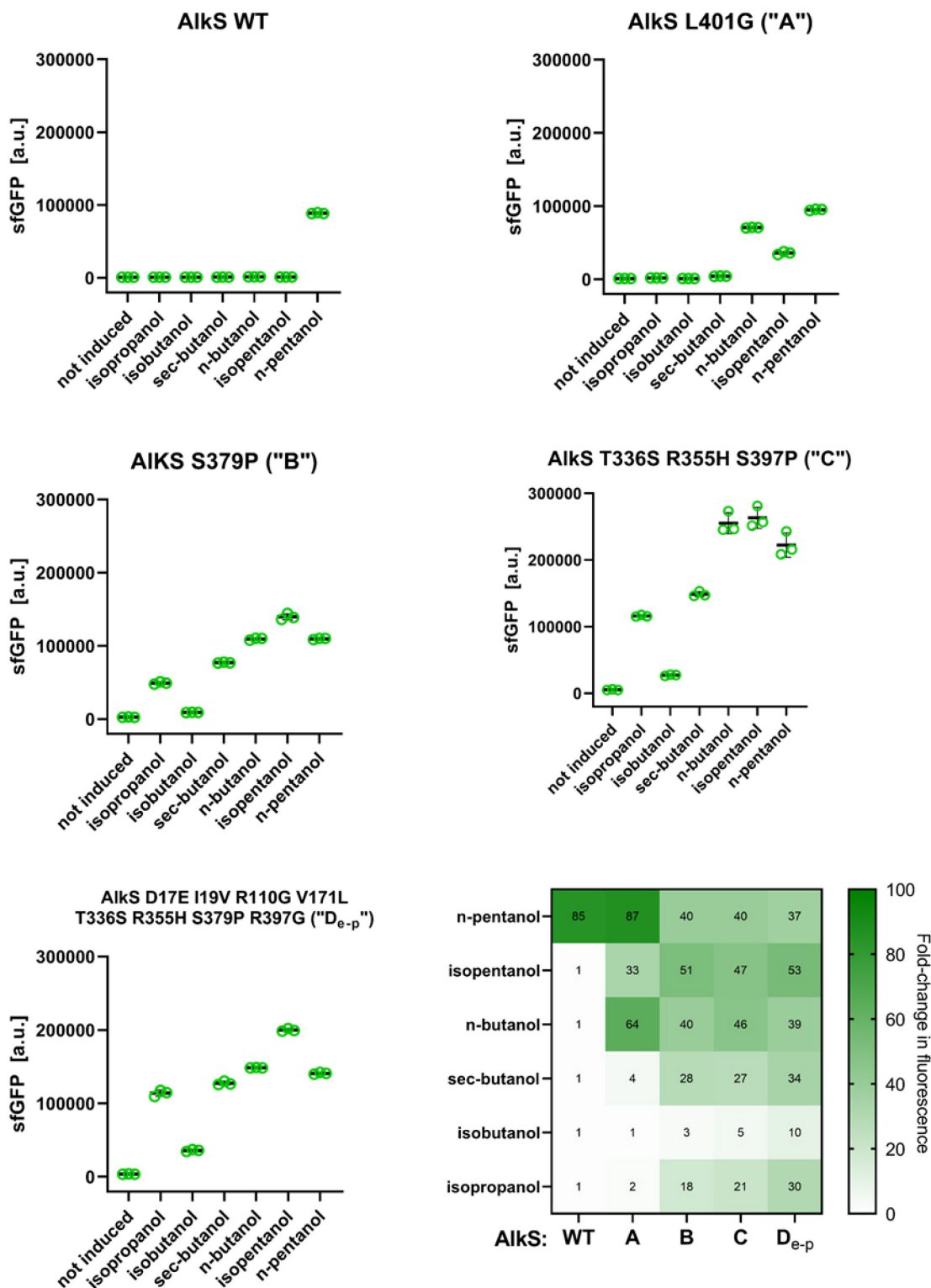
Supplementary Figure 3-3 FACS-based screening of AlkS libraries in *E. coli* 10β(*mCherry*) [pCK01_alkS_alkS pAB_alkB_sfgfp_KanR] for improved inducibility with C4- and C5-alcohols. Between enrichment rounds, cells were re-grown and re-induced with 10 mM alcohol. In enrichment round three, cells were spotted onto LB agar plates (384 events per plate). The sfGFP fluorescence was taken as sensor output (excitation at 488 nm, emission with 530/30 nm band pass filter), mCherry was used as a biomass marker only (561 nm, 610/20 nm). Sorting gate labels indicate the fraction of the total population. Unlabeled gates (column “pos. enrichment 3”, rows one, two and three) had viability control purposes only. 2% contour plots with outliers are shown, created with FlowJo software (FlowJo LLC, Ashland, OR, USA).



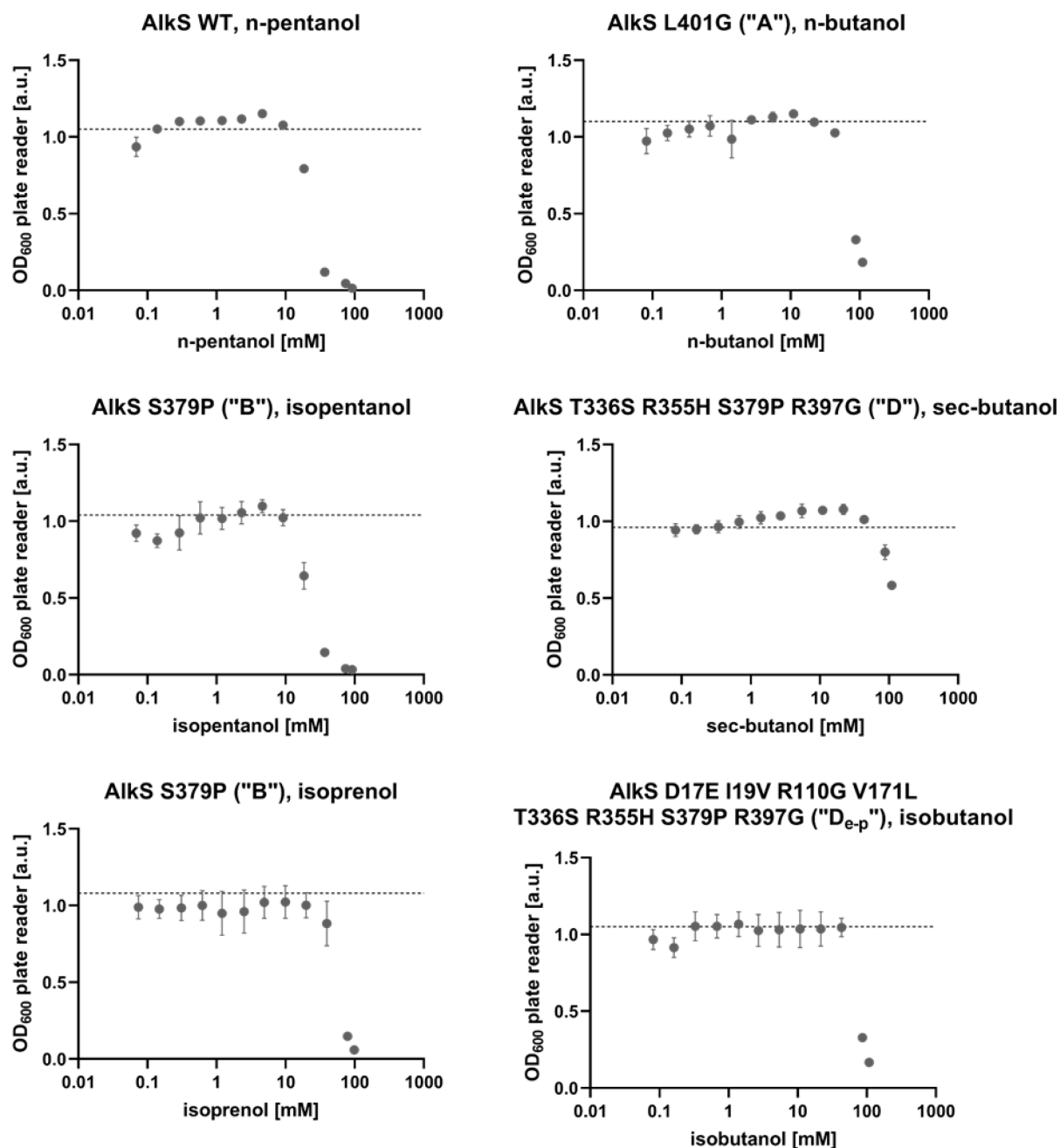
Supplementary Figure 3-4 FACS-based screening of AlkS libraries and representative hit verification. Left: OmniTray LB agar plate with re-grown *E. coli* 10 β (*mCherry*) [pCK01_ *P_{alkS}*_ *alkS* pAB_ *P_{alkB}*_ *sfgfp*_ KanR] single-cells spotted in 384-plate format during FACS screen. Middle and Right: Verification of AlkS position L401 for n-butanol recognition. a) wild-type AlkS, b) no cells, c) sensor circuit without AlkS, d) AlkS L401S (site-directed mutagenesis), e) AlkS L401D (site-directed mutagenesis), f) re-transformed AlkS L183I L401S (original screening hit). Middle culture plate is LB agar only, right is LB supplemented with 0.1% (v/v) n-butanol. Overlay of bright field image with UV cut filter and fluorescence image with 470 nm excitation and 545 nm peak filter for middle and right picture (round agar plates; Dark Hood DH-50).



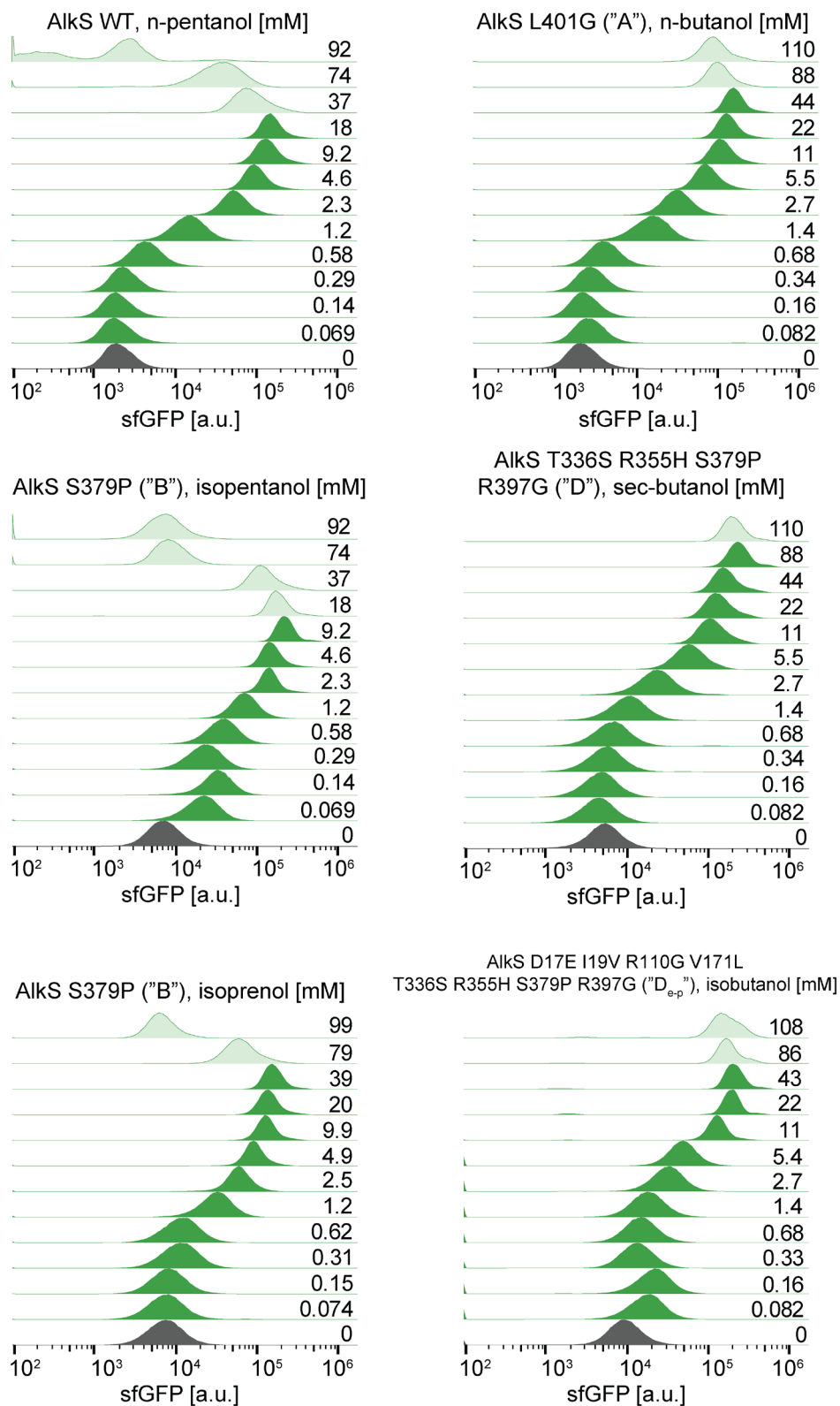
Supplementary Figure 3-5 Representative fluorescence outputs of *E. coli* 10 β (*mCherry*) [pCK01_ *P_{alkS}*_ *alkS* pAB_ *P_{alkB}*_ *sfgfp*_ KanR] biosensor strains. Left: Response-kinetics for 10 mM n-butanol with wild-type AlkS (black diamonds) or AlkS L401G (blue triangles). M9 medium with 4 g L⁻¹ glucose, 37°C, 25 mL culture volume in 250 mL Erlenmeyer flasks with lid (n=1). Right: Dose-response of AlkS L401G biosensor strain in deep well plates sealed with aluminium foil (750 μ L culture volume, 8 h of incubation, n=3, 200 μ L sample volume). Leftmost wells are cultures without alcohol added, others are serial dilution of 10 mM n-butanol (rightmost well, decreasing n-butanol concentration to the left). Top, bright field image with UV cut filter; bottom, fluorescence image with 470 nm excitation and 545 nm peak filter (Dark Hood DH-50).



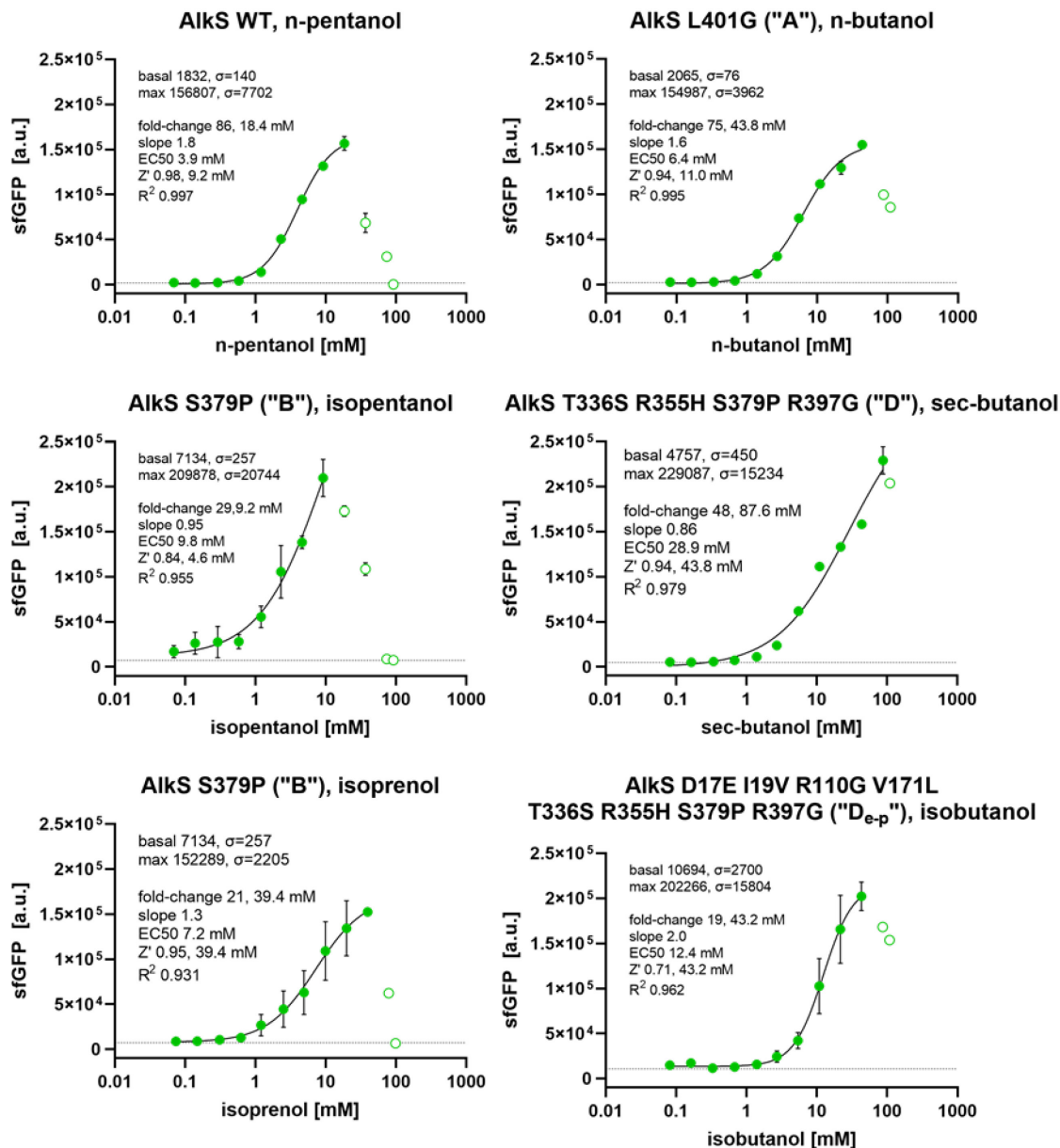
Supplementary Figure 3-6 Biosensor strain *E. coli* 10β(*mCherry*) [pCK01_*P_{alkS}*_alkS pAB_*P_{alkB}*_sfGFP_KanR] variants and their response to a panel of alcohols in deep well plates. All compound concentrations are 10 mM, except for isopropanol (100 mM). M9 medium with 4 g L⁻¹ glucose, 30°C, 750 μL culture volume, plates sealed with aluminum foil, 16 h of incubation. Data points are medians with mean (black bar) and standard deviation (n=3) based on single cell measurements (Beckman Coulter Cytoflex S, >50'000 events per population).



Supplementary Figure 3-7 Best-in-class *E. coli* 10β(*mCherry*) [pCK01_P_{alkS}_alkS pAB_P_{alkB}_sfgfp_KanR] biosensor strain variants and their response to a concentration gradient of the corresponding alcohol. M9 medium with 4 g L⁻¹ glucose, 30°C, 750 μL culture volume, plates sealed with aluminum foil, 16 h of incubation. Biomass measured at 600 nm (OD₆₀₀, 200 μL sample volume, Tecan M1000 Pro). Data points are means with standard deviation (n=3), dotted lines represents mean OD₆₀₀ values for cultures without alcohols added.



Supplementary Figure 3-8 Best-in-class *E. coli* $10\beta(mCherry)$ [pCK01_P_{alkS}_alkS pAB_P_{alkB}_sfGFP_KanR] biosensor strain variants and their response to a concentration gradient of the corresponding alcohol. M9 medium with 4 g L⁻¹ glucose, 30°C, 750 μ L culture volume, plates sealed with aluminum foil, 16 h of incubation. Representative single cell fluorescence output data (Cytoflex S, >100'000 events per population, n=1). Grey populations show basal fluorescence for cultures without alcohols added, light green populations were excluded from dose-response curve fitting due to severely harmed growth (see Supplementary Figure 3-7 and 3-8).



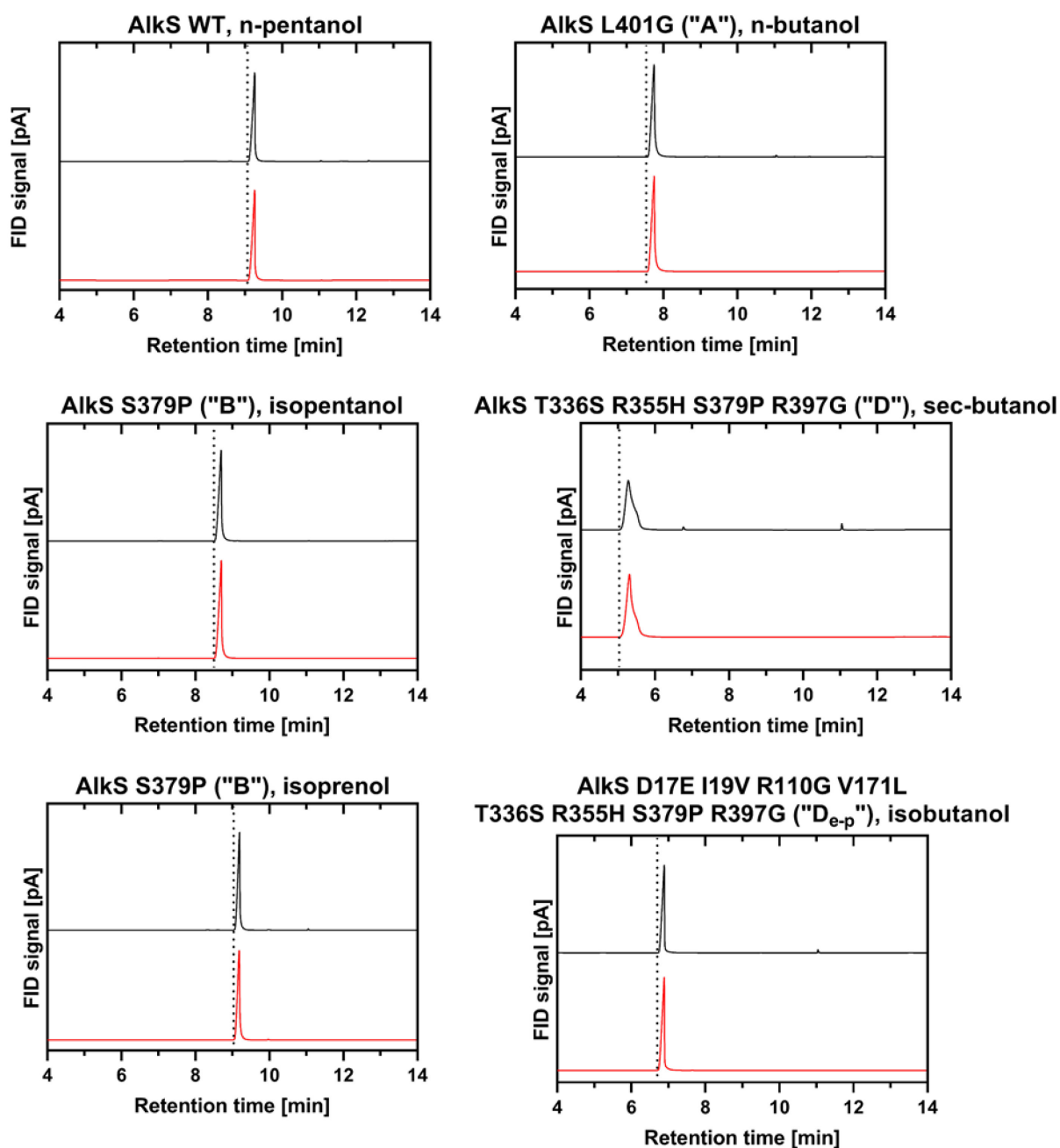
$$y = sfGFP = sfGFP_{basal} + \frac{c_x^n (sfGFP_{max} - sfGFP_{basal})}{c_x^n + EC_{50}^n}$$

$$Z' = 1 - \frac{3\sigma_{sfGFP_{opt}} + 3\sigma_{sfGFP_{basal}}}{|sfGFP_{opt} - sfGFP_{basal}|}$$

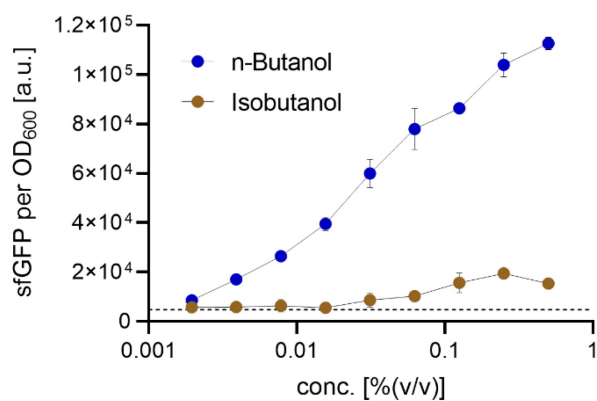
$$fold-change = \frac{sfGFP_{max}}{sfGFP_{basal}}$$

- (1) **Supplementary Figure 3-9** Dose-response curves of best-in-class *E. coli*
- (2) 10 β (*mCherry*) [pCK01_P_{alkS}_alkS pAB_P_{alkB}_sfGFP_KanR] biosensor variants
- (3) and the corresponding alcohols. M9 medium with 4 g L⁻¹ glucose, 30°C, 750 μ L culture

volume, plates sealed with aluminum foil, 16 h of incubation. Data points are means of medians with standard deviations (Cytoflex S, >100'000 events per population, n=3). Full circles indicate data points used for Hill fit^(eq. 1), open circles were excluded due to severely harmed growth (see Supplementary Figure 3-7 and 3-8). Solid lines indicate Hill fit, dotted lines represent mean basal fluorescence values for cultures without alcohols added. c_x , alcohol concentration; EC₅₀, half maximal effective concentration; σ , standard deviation; sfGFP, superfolder green fluorescent protein; slope, sensitivity n (Hill coefficient). Z'^(eq. 2) indicates effectiveness of screens, please note that optimal (opt) and maximal (max) sfGFP values might differ depending on the standard deviation. Fold-changes in biosensor output were calculated according to ^(eq. 3). Hill-fits with GraphPad Prism 8.4.2, "[agonist] vs. response, variable slope (four parameters)" equation.

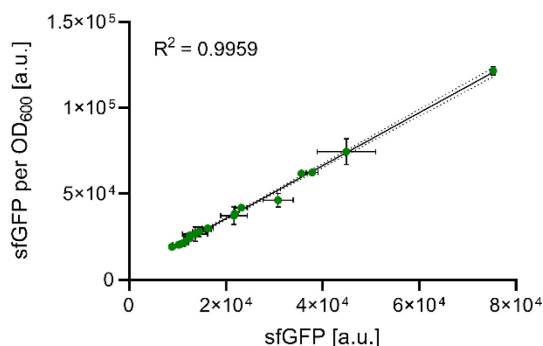
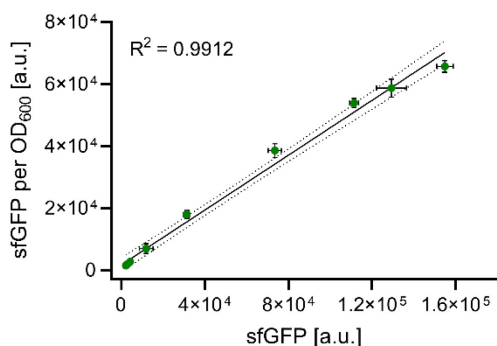


Supplementary Figure 3-10 GC-FID analysis of chemical compounds used for induction experiments in culture medium before (red, bottom) and after *E. coli* 10 β (*mCherry*) [pCK01_alkS_alkS pAB_PalkB_sfgfp_KanR] biosensor strain cultivation (black, top). M9 medium with 4 g L⁻¹ glucose, 30°C, 750 μ L culture volume, plates sealed with aluminum foil, 16 h of incubation. Dotted lines indicate the left peak-basis before strain cultivation. All samples were extracted with methyl *tert*-butyl ether, splitless injection, DB-WAX capillary column (Agilent).

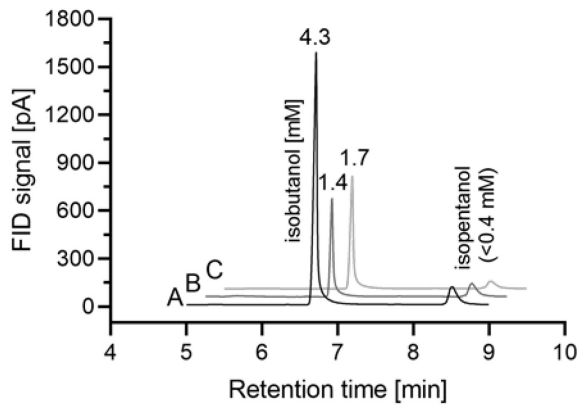


Supplementary Figure 3-11

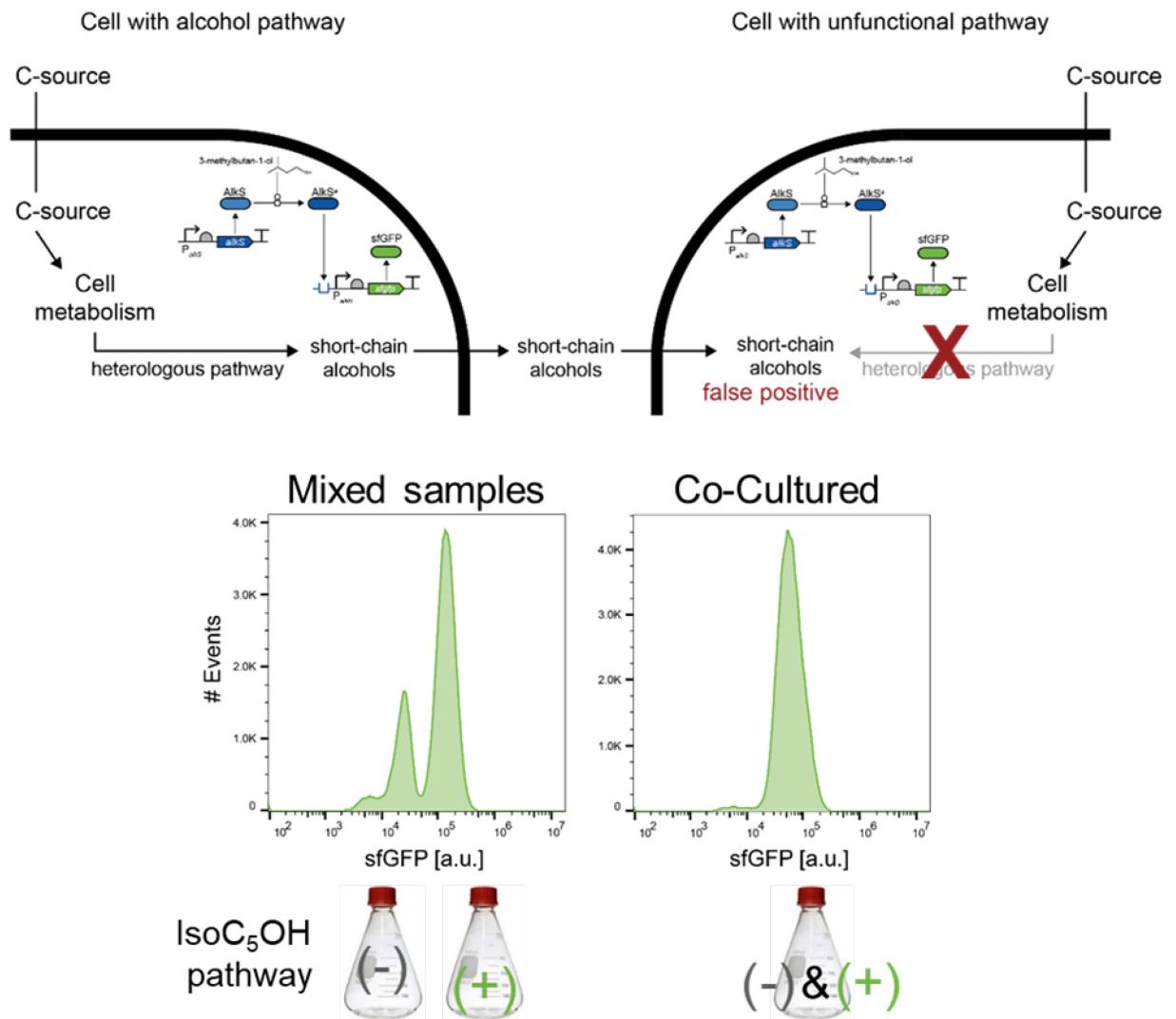
Fluorescence output of biosensor strain *E. coli* DH5 α [pCK01_P_{alkS}_alkS pAB_P_{alkB}_sfgfp_KanR] with AlkS L401G for varied concentrations of the structural butanol isomers isobutanol (brown) and n-butanol (blue). M9 medium with 4 g L⁻¹ glucose, 30°C, 750 μ L culture volume, plates sealed with aluminum foil, 16 h of incubation. Data points are means with standard deviation (n=3), dotted line represents mean basal fluorescence value for cultures without alcohol added (Tecan M1000 Pro).



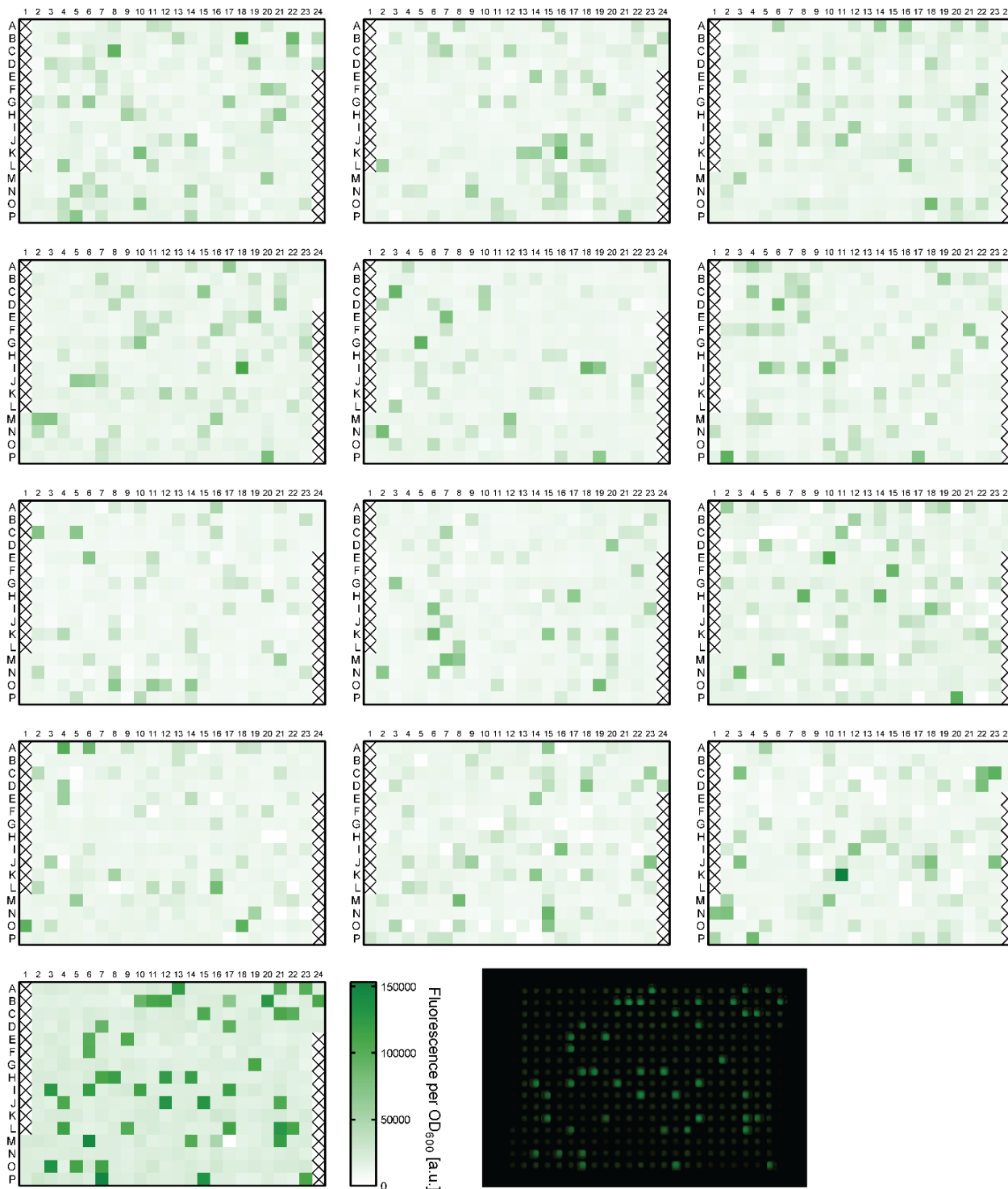
Supplementary Figure 3-12 Correlation of biosensor output data obtained from single cell (sfGFP, Cytoflex S) or plate reader (sfGFP per OD₆₀₀, Tecan M1000 Pro) measurements. Left: *E. coli* 10 β (mCherry) [pCK01_P_{alkS}_alkS pAB_P_{alkB}_sfgfp_KanR] biosensor strain response to varied, supplemented n-butanol concentrations, right: *E. coli* DH5 α [pCK01_P_{alkS}_alkS pAB_P_{alkB}_sfgfp_KanR] biosensor strain output in KivD site-saturation mutagenesis experiments for isopentanol production by whole-cell catalysis. Data points are mean values with standard deviations (n=3). Solid line is best-fit linear regression, dotted lines indicate the 95% confidence bands (GraphPad Prism 8.4.2).



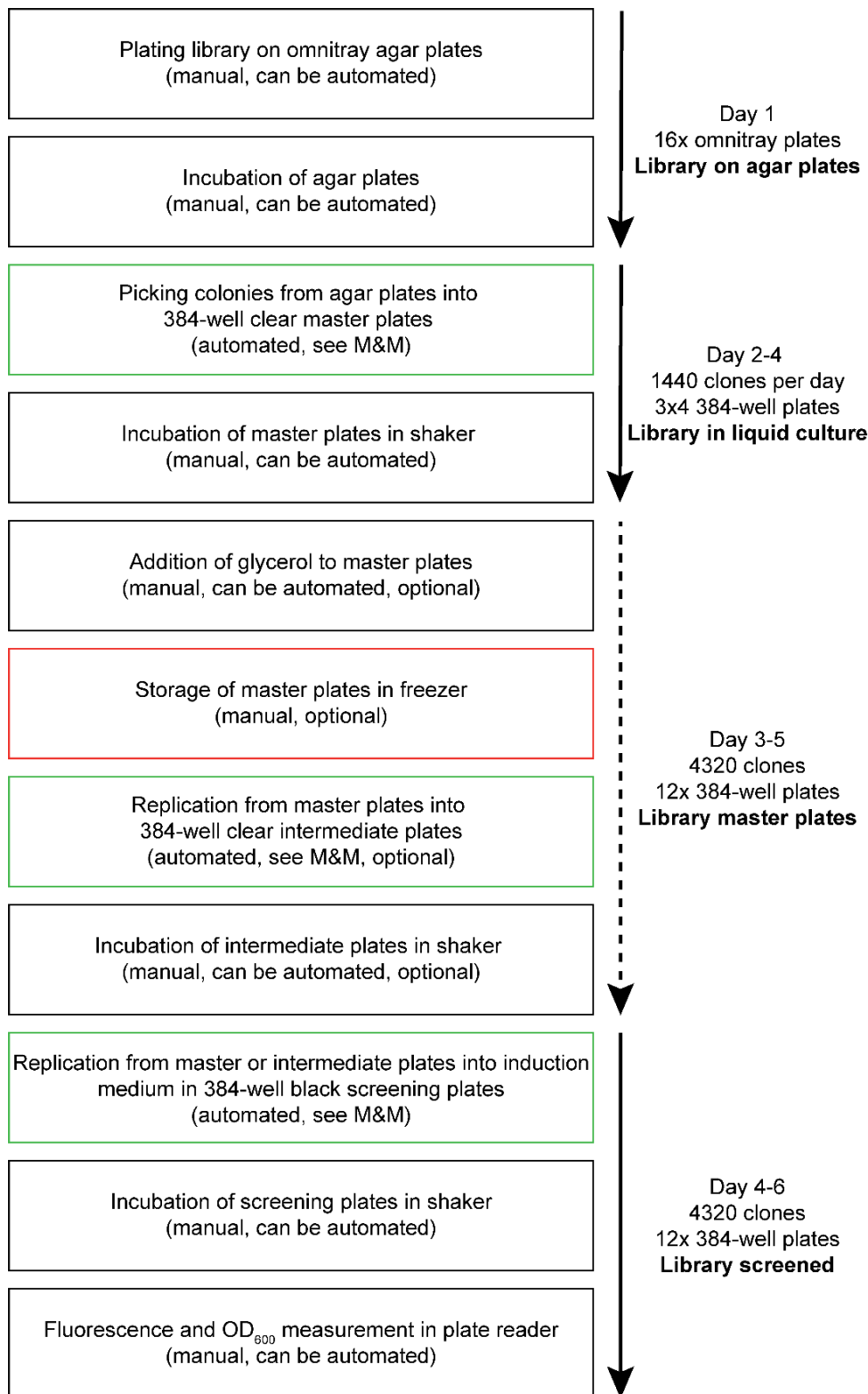
Supplementary Figure 3-13 GC-FID analysis of culture supernatants after alcohol production. The isobutanol pathway was stably integrated in the genome downstream of *glmS* (per Tn7 transposon) in the following *E. coli* strain backgrounds: A) BW25113 $\Delta ilvE$, B) BW25113, C) NEB 10 β (*mCherry*). Isobutanol production experiments were conducted in M9 medium with 5 g L⁻¹ glucose and 5 g L⁻¹ yeast extract, 30°C, 10 mL culture volume, 100 mL Erlenmeyer flasks with lid, 24 h of incubation.



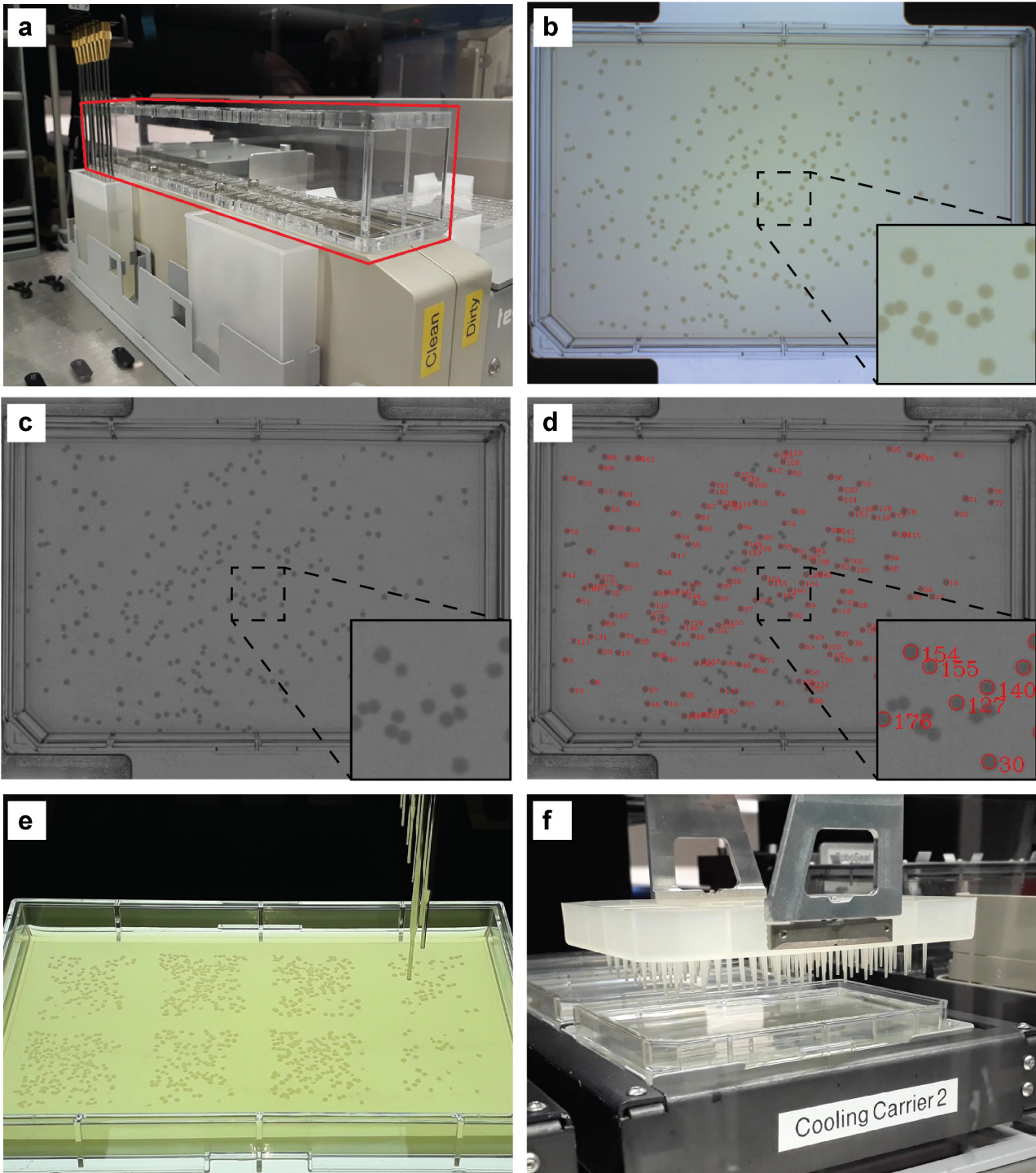
Supplementary Figure 3-14 Occurrence of false positives in one pot assays for short chain alcohol production and concomitant biosensor-based *in situ* product detection. Top: Schematic overview of potential product crosstalk. A short-chain alcohol of interest is produced by a microbial cell (left) equipped with the corresponding alcohol pathway. The alcohol passively diffuses across cell membranes and reaches the cytoplasm of a non-producing cell (right), where it activates the AlkS L401G-based biosensor circuitry and leads to a false-positive sensor activation. Bottom: Two different *E. coli* DH5α [pCK01_P_{alkS}_alkS pAB_P_{alkB}_sfGFP_KanR] biosensor strain variants without (-) or with (+) a plasmid-based isopentanol (isoC₅OH) production pathway were cultivated separately in two flasks (left) or in a single flask as a co-culture (right). After alcohol production, the cell populations were analyzed at the single cell level (Cytotflex S). For the strain variants cultivated separately, the culture samples were mixed just before flow cytometer analysis (“mixed samples”) yielding two clearly distinct populations in terms of biosensor output (sfGFP). For the co-culture, only a single, highly fluorescent population was observed. M9 medium with 5 g L⁻¹ glucose and 1.25 g L⁻¹ yeast extract, 30°C, 10 mL culture volume, 100 mL Erlenmeyer flasks with lid, 16 h of incubation.



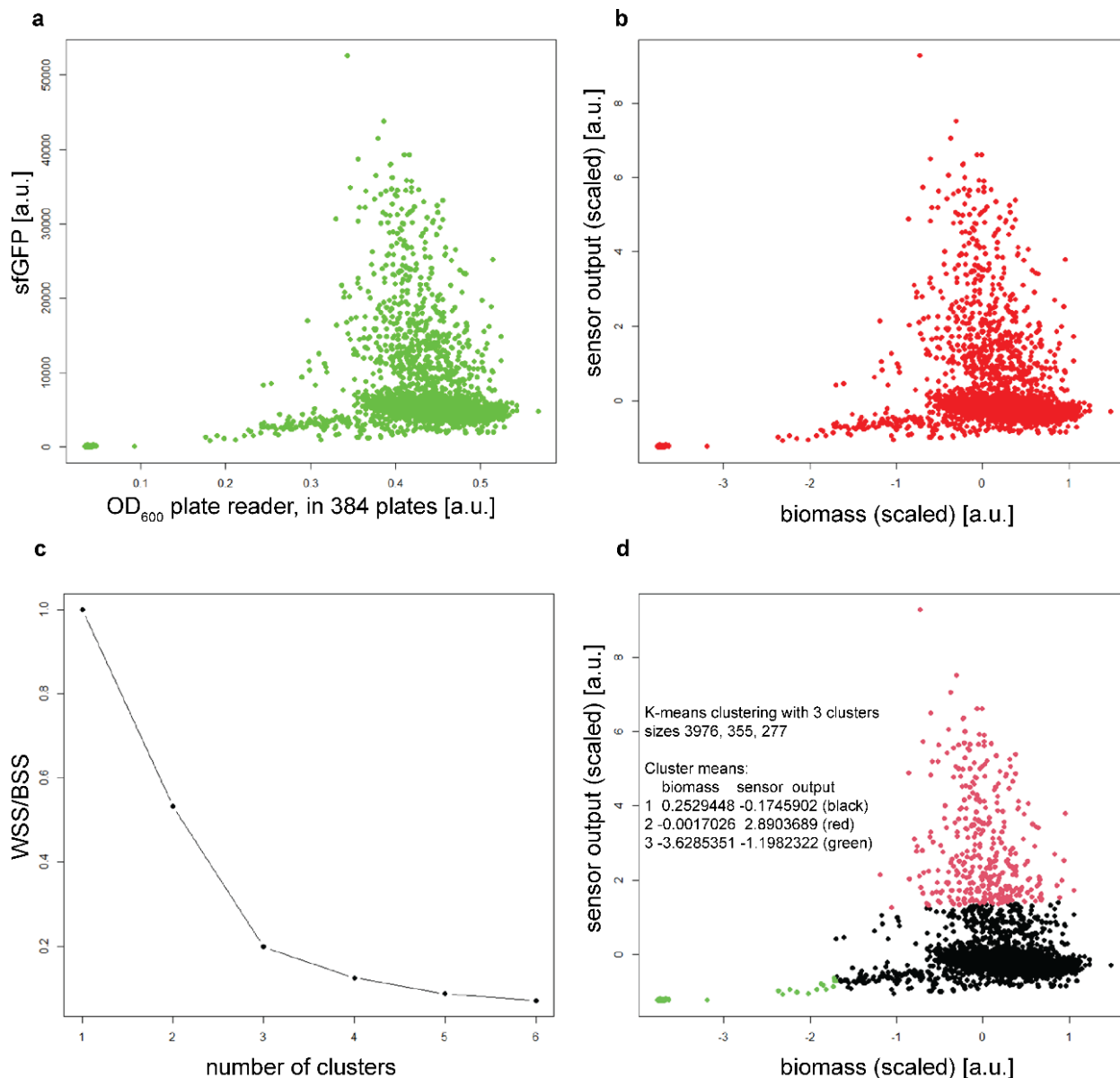
Supplementary Figure 3-15 Biosensor outputs in 384-well plates after automated screening of a combinatorial *leuA* library (in plasmid p221_araC_P_{BAD}_leuABCD) for isopentanol production with *E. coli* BW25113 $\Delta ilvE$ gISO2 [pCK01_P_{alkS}_alkS_L401G pAB_P_{alkB}_sfgfp_AmpR]. Plate reader measurements at 600 nm (OD₆₀₀) and 488 nm with a 530/20 nm filter (sfGFP fluorescence, Tecan M1000 Pro) after 20 h induction, 60 μ L culture volume. The single plate in row five is obtained from a mock library run, initially containing 10% of an isopentanol production strain (*leuABCD* overexpression) and 90% of an isobutanol production strain (no *leuABCD* overexpression). 12% of wells showed increased fluorescence as expected for the isopentanol strain (44/360), and strains selected based on fluorescence output had all the expected genotype (8/8 for each phenotype, by colony PCR). The corresponding image (470 nm excitation and 580 nm amber filter; Kodak EDAS 290 with Biohelix BluPad) was re-colored with Adobe Photoshop software.



Supplementary Figure 3-16 Automation workflow for biosensor-based screening of isopentanol production strain library with a Tecan Evo robotic liquid-handling platform. First, single colonies are picked and cultivated as LB precultures in 384-well plates. Next, the cultures are stored as cryoplates until further usage or directly replicated in production medium for isopentanol synthesis. Finally, the strains are evaluated based on biomass formation (OD₆₀₀) and biosensor output (fluorescence) as a proxy for the isopentanol titer. Here, we automated the most labor and material intensive steps, colony picking and culture replication (green boxes). All plating, incubation and measurement steps (black boxes) could be performed on the same robotic platform as well



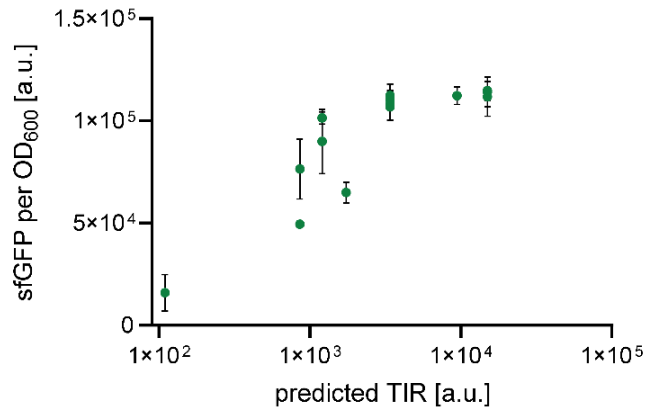
Supplementary Figure 3-17 Images from automation workflow. a) Custom made lid on tip wash station of liquid handling robot during automated colony picking. The lid prevents contamination due to microbe-containing aerosols that emerge during the tip washing procedure. b) Raw image of Omintray agar plate as acquired during colony picking from the Pickolo colony picker module. c) Background corrected image after processing using ImageJ macro. d) Analysis result of background corrected image in Pickolo software. Single colonies to pick are marked with red circles. e) Omintray agar plate on light table during picking of colonies using the automated platform. One of the fixed stainless steel pipetting tips is picking up a colony. f) Washing of 96-pin replicator using the robotic manipulator arm of the liquid handling robot. The pin replicator is agitated for 20 seconds in a bath of 2% hypochlorite provided in an Omintray plate.



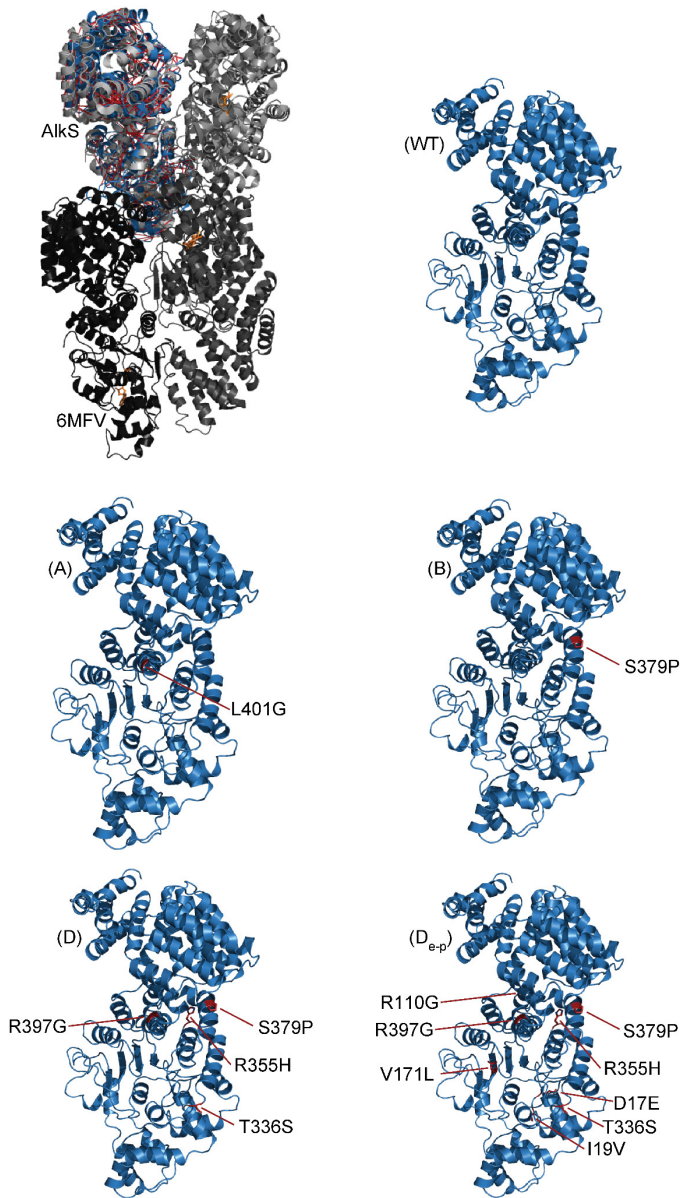
Supplementary Figure 3-18 Screening of combinatorial *leuA* library (variation of RBS and introduction of FBR mutation) for isopentanol production and biosensor output analysis. a) Raw data as obtained from plate reader measurements of *E. coli* BW25113 $\Delta ilvE$ gISO2 [pCK01_P_{alkS}_alkS_L401G pAB_P_{alkB}_sfGFP_AmpR p221_araC_P_{BAD}_leuABCD] strain library after 20 h of alcohol production in 384 plates (60 μ L culture volume), with biomass measured at 600 nm (OD₆₀₀) and green fluorescence biosensor output at 488 nm with a 530/20 nm filter (sfGFP, Tecan M1000 Pro). b) The same data set scaled for variable comparability and cluster analysis. c) Ratio of within sum of squares (WSS) and between sum of squares (BSS) for different numbers of clusters. At a ratio of about 0.2 the slope decreased, indicating three clusters (“Elbow method”²⁴¹). d) k-means clustering of scaled data set, with n=3 clusters and nstart=20 iterations. All data analysis was carried out with R version 4.0.3.

Sequencing of colonyPCRed fragments (not full *leuA(BCD)*)

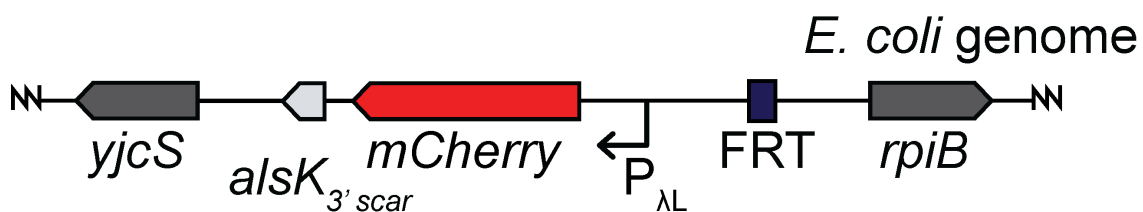
strain ID	RBS	pred. TIR [a.u.]	codon	G462X	comment	
isoC4OH	n/a	n/a	n/a	n/a	empty plasmid control, pSEVA221	
cluster 1 +2	2	n/a	n/a	n/a	no PCR-band obtained, <i>leuA</i> probably missing	
	4	AGGAGG	15023	ACT	T additional G at position 159, frameshift	
	6	AGGAWG	n/a	AGT	S noise only from base 78 onwards , no clear PCR band, error in pathway assembly	
	5	TGGACG	110	ATT	I	
	1	n/a	n/a	n/a	n/a	no PCR-band obtained, <i>leuA</i> probably missing
	3	AGGAGG	15023	n/a	n/a	deletion after about 677 bp
cluster 3	12	TGGAGG	1208	CAT	H wrong gene order or deletion in pathway	
	11	n/a	n/a	n/a	n/a	noise in sequencing data , no clear PCR band, probably mixed variants
	13	GGGAAG	862	CTG	L	
	15	AGGACG	1740	TCT	S	
	22	GGGACG	862	TTT	F	
	23	TGGAGG	1208	TAT	Y	
	16	TGGAGG	1208	CAT	H	
	8	CGGAGG	3402	CGT	R	
	18	CGGAGG	3402	CAT	H	
	19	CGGAGG	3402	CCG	P	
	9	AGGAGG	15023	CGG	R	
	17	AGGAGG	15023	AGT	S	
	7	GGGAGG	9450	TGG	W	
	10	CGGAGG	3402	CCT	P	
24	AGGAGG	15023	GAG	E		
21	AGGAGG	15023	ATG	M		
14	AGGAGG	15023	TTT	F		



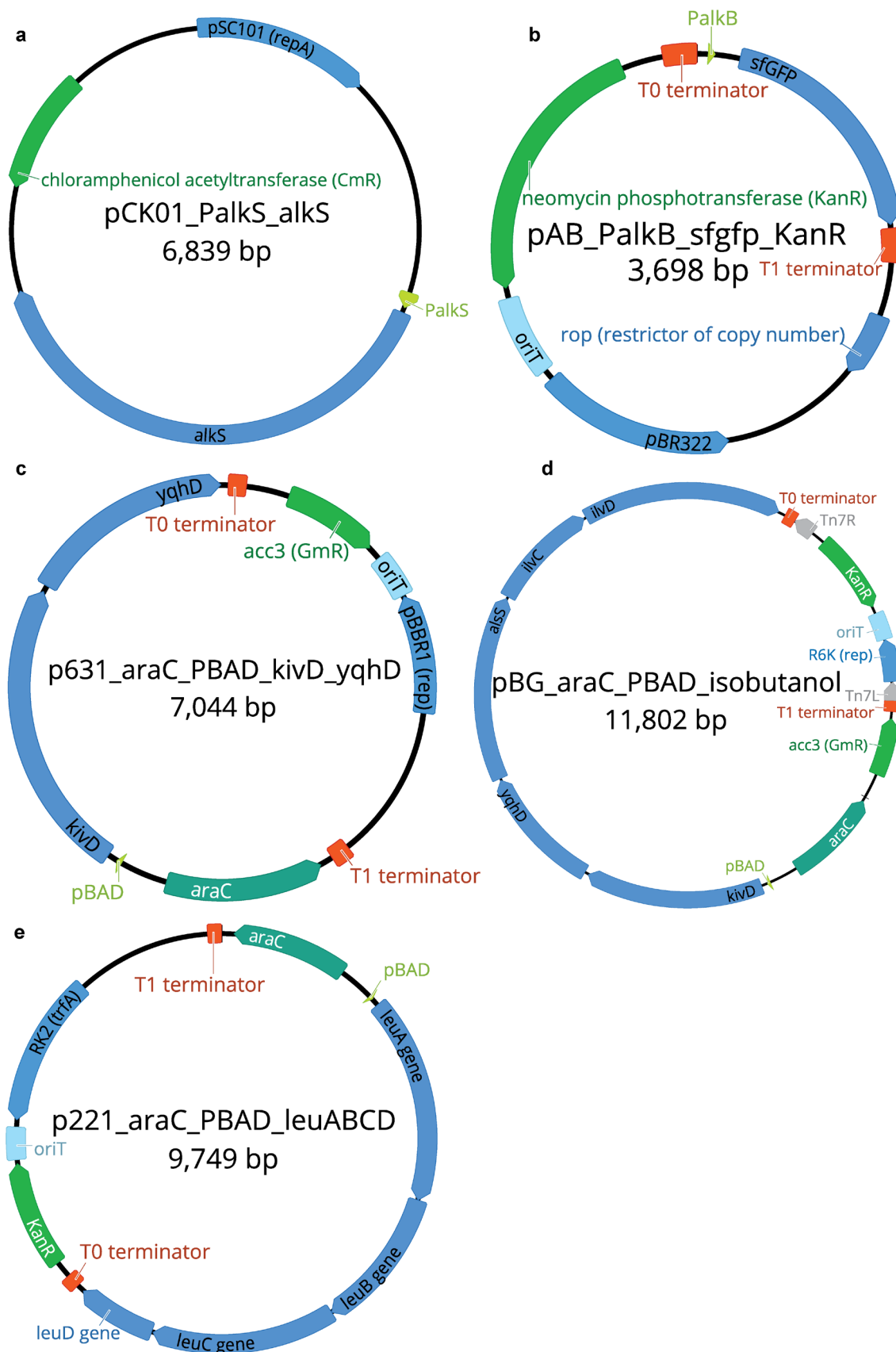
Supplementary Figure 3-19 Overview of *E. coli* BW25113 $\Delta ilvE$ glISO2 [pCK01_*P_{alkS}*_alkS_L401G pAB_*P_{alkB}*_sfgfp_AmpR p221_araC_*P_{BAD}*_leuABCD] *leuA* combinatorial library members selected for re-screening in deep well plates. The table indicates the corresponding cluster as well as the sequences obtained for the degenerate RBS part, the predicted translation ignition rate (TIR), and the amino acid at position G462X. The inlet graph shows the predicted TIR versus the observed biosensor output (750 μ L culture volume, 20 h alcohol production). Data points are means with standard deviation (n=3).



Supplementary Figure 3-21 Homology model of AlkS created with Phyre2²²⁴ (72% sequence coverage, 100% confidence), based on the structure of a signal transduction ATPase (PDB 6MFV). The mutations present in defined AlkS variants used for biosensor characterization were mapped to the predicted structure. Please note that the full tetrameric model structure of AlkS is shown in the first panel only, otherwise the monomer is depicted.



Supplementary Figure 3-22 Integration of a constitutively expressed red-fluorescent gene (*mCherry*) in *E. coli* 10 β (NEB #C3020K) as a biomass marker for fluorescence-assisted cell sorting. The synthetic gene construct $P_{\lambda L}::mCherry$ was integrated by homology-based λ -red recombineering and replaced the allose operon (*alsRBACEK*). The promoter and *mCherry* sequences are available from the Registry of Standard Biological Parts via identifiers BBa_R0051 and BBa_J06504, respectively. The FRT-scar resulted from removal of the Kan^R cassette, while also a 27 bp part of *alsK* is left. For details, see: Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **97**, 6640-6645 (2000).



Supplementary Figure 3-23 Plasmid maps for biosensor construction (a, b), whole-cell catalysis of precursor keto-acid to isopentanol (c), genomic integration of isobutanol pathway (d), and *leuABCD* overexpression for isopentanol biosynthesis (e). Plasmids pCK01_P_{alkS}_alkS (addgene ID 166502, with L401G mutation) and pAB_P_{alkB}_sfGFP_KanR (ID 166503) are available from addgene. Maps were created with Geneious software version 11.1 (Biomatters, San Diego, CA, USA).

Chapter 4: Short-Chain Alcohol and Fatty Acid Biosynthesis in *Pseudomonas putida* with Co-Culture Based *in vivo* Product Detection

Maximilian Ole Bahls, Felix Pfeifer, Sven Panke

4.1 Abstract

Bioprocesses starting from renewable input-resources have the potential to reduce greenhouse gas emissions, as compared to petrochemical processes relying on ultimately limited fossil resources. *Pseudomonas putida* is an up-and-coming microorganism well-suited for industrial process development, due to solvent tolerance and a versatile metabolism. Here, we tested the feasibility of *P. putida* EP2 for isopentanol production, explore varied expression constructs for production via the Ehrlich degradation route and developed a fed-batch process for improved isopentanol titers. Besides, we investigate a co-cultivation strategy for *in situ* product detection with an orthogonal biosensor strain.

P. putida EP2 showed no consumption of isopentanol and tolerance for up to 2 g L⁻¹ isopentanol in cultivations with glucose as C-source, indicating its suitability as a production host. When supplied with plasmid-based and genome-integrated overexpression pathways for isopentanol biosynthesis, we found up to 170 mg L⁻¹ isopentanol for a two-plasmid expression system, while genome integrated pathways showed a reduced expression strength and lower isopentanol titers (66 mg L⁻¹ isopentanol). Besides, we found up to 620 mg L⁻¹ isovaleric acid as a side product. In a fed-batch process with a microaerobic production phase, the isopentanol titer increased to 553 mg L⁻¹, while restricting isovaleric acid formation to 151 mg L⁻¹. Product levels could also be reliably determined with an orthogonal *E. coli* biosensor strain in co-culture with alcohol-producing *P. putida*, with up to seven-fold increased sensor output for co-cultures with the induced production strain.

P. putida EP2 was shown to be a feasible host for isopentanol and isovaleric acid production. In combination with the biosensor-based co-cultivation, this could be the basis for high-throughput screens for further improved strains.

4.2 Introduction

Fermentation-based industrial processes have the potential to critically improve the carbon footprint of products traditionally derived from limited fossil resources¹⁶⁶. For chemical building blocks and transportation fuels, low crude oil prices and the established production lines tailor-made for petrochemistry have limited the introduction of bioprocesses starting from renewable substrates⁶. However, the reduction of greenhouse gas emissions is an urgent task making environmentally benign bioprocess development increasingly favorable⁷.

A major bottleneck for such bioprocesses is the availability of microorganisms producing small molecules, such as alcohols and fatty acids, at high titers, rates and yields^{6, 15}. In order to design improved microbial strains, a vast and growing molecular biology toolset, facilitated by the fields of synthetic and systems biology, is available for metabolic engineering of microbes such as *Escherichia coli* and *Corynebacterium glutamicum*^{5, 163, 245, 246}. Further, a particularly interesting and suitable microbial host chassis for industrial processes is *Pseudomonas putida*^{247, 248}. This strictly aerobic soil bacterium is naturally endowed with features such as solvent resistance²⁴⁹, exceptionally high NAD(P)H regeneration rates²⁵⁰ and a versatile metabolism. *P. putida* strains not only metabolize renewable substrates with a negative carbon footprint such as glucose but also lignin-derived waste compound²⁵¹. Besides, *P. putida* is known to produce little byproducts such as acetate, ethanol or lactate²⁵², which are known to accumulate during cultivations with *E. coli* under unfavourable growth conditions due to glucose overflow metabolism or mixed-acid fermentation²⁵³. Since recently, most synthetic biology methodologies available for *E. coli*, are available for engineering *P. putida* as well²⁴⁷. Besides, *P. putida* KT2440 is safe to use²⁵⁴ and the genome fully sequenced and well annotated^{255, 256}.

Building up on this knowledge base, genome-engineered *P. putida* strains were shown to be unable to consume butanol as a C-source^{184, 257}, in contrast to *P. putida* KT2440, and further to be suitable for the production of isobutanol on the gram per liter scale in pilot-scale fed-batch fermentations²⁵⁸. Also the further oxidized product, isobutyric acid, was previously produced with titers of up to 2.4 g L⁻¹ using a metabolically engineered *Pseudomonas* sp. strain VLB120²⁵².

Nevertheless, many more small molecule products are not yet available from microbial fermentation processes using *P. putida*. For example, isopentanol (3-methylbutan-1-ol) is annually produced on the kilo-ton scale and can directly be used as a chemical building

block, drop-in biofuel or organic solvent¹⁷¹. Simple chemical processing steps allow the production of isopentanol-derived molecules that are suitable for use in the flavor and fragrance industry as well as for use as jet fuel¹⁶⁹. In particular aviation fuel grade compounds are crucial for a more sustainable transportation industry, yet they are relatively difficult to produce at economic titers with microorganisms, as more established alcohol products including ethanol have insufficient physiochemical properties¹⁷². So far, *E. coli*¹⁸³ and *C. glutamicum*¹⁷⁹ as well as *Saccharomyces cerevisiae*²⁵⁹ were the main metabolic engineering targets for improved production of isopentanol, but with the list of favorable properties mentioned above *P. putida* would be an attractive target host. Any attempt to engineer such a strain would benefit from simple and high-throughput compatible product detection. However, the detection of for example isopentanol production still relies on relatively slow and expensive chemical analysis²⁶, and a detection scheme based on a biosensor system, converting the product titer into a fluorescence readout, would be highly desirable for high-throughput screening of production strain libraries^{21, 189}.

Here, we evaluate and engineer *P. putida* EP2 for the production of isopentanol, derived from the commonly used Ehrlich degradation pathway of amino acids in yeast¹⁷³ (Figure 4-1 a). Different Ehrlich pathway overexpression strategies are tested and the best design is used for fed-batch process development for isopentanol production on the bench-top bioreactor scale. Besides, we demonstrate the suitability of co-cultivations with an orthogonal *E. coli* biosensor strain²⁶⁰ for *in situ* alcohol titer detection per sfGFP readout.

4.3 Results

Feasibility of *P. putida* strains for isopentanol production

As a first step in bioprocess development, *P. putida* KT2440 and a genome-engineered derivative for butanol production, *P. putida* EP2¹⁸⁴, were compared for their feasibility for isopentanol production. *P. putida* KT2440 is known to metabolize alcohols such as n-butanol²⁶¹ as a carbon source and accordingly we found consumption of isopentanol as well as biomass formation when growing on 1 g L⁻¹ isopentanol as the sole C-source in defined medium (Figure 4-1b). For the EP2 strain, harboring knockouts of multiple alcohol, aldehyde and 2-oxoisovalerate dehydrogenases ($\Delta pedE$, $\Delta pedH$, $\Delta pedI$, $\Delta aldB$, $\Delta bkdAA$), no isopentanol consumption was found and no significant biomass growth was possible with isopentanol as C-source. The final biomass obtained for growth on glucose was not affected and similar to the one obtained with *P. putida* KT2440. In isopentanol tolerance experiments with defined medium and 5.0 g L⁻¹ glucose as C-source, we found biomass growth of *P. putida* EP2 in the presence of up to 2 g L⁻¹ isopentanol. Further, cells recovered from such an experiment after 24 h of incubation were successfully regrown, indicating their viability after 24 h of incubation in the presence of isopentanol (Figure 4-1c).

Branched-chain alcohol pathway overexpression

The Ehrlich degradation pathway was used for isopentanol production with *P. putida* EP2. The enzymes of the entire pathway from pyruvate to 4-methyl-2-oxopentanoic acid (a precursor in leucine biosynthesis) were overexpressed, as well as two additional heterologous enzymes for decarboxylation (KivD, originating from *Lactococcus lactis*) and reduction (YqhD, *E. coli*) of 4-methyl-2-oxopentanoic acid to 3-methylbutanal and isopentanol, respectively. The upper part of this pathway, i.e. from pyruvate to 2-ketoisovalerate (the 2-ketoacid precursor of valine), is also part of the pathway to isobutanol¹⁸⁴, which can be reached in two steps from 2-ketoisovalerate, catalyzed again by KivD and YqhD (Figure 4-1a). It consists of AlsS, an acetolactate synthase with high pyruvate affinity of *Bacillus subtilis*²⁶², catalyzing the first reaction from pyruvate to 2-acetolactate, and IlvCD from *P. putida* KT2440 catalyzing the reduction to 2,3-dihydroxyisovalerate and the subsequent dehydration to 2-ketoisovalerate. The lower part of the pathway is catalyzed by LeuABCD recruited from *E. coli*, including an acetylation catalyzed by LeuA, a rearrangement by LeuCD, and a reduction to 2-isopropyl-3-oxosuccinate, which spontaneously decarboxylates to 4-methyl-2-oxopentanoic acid. For

the first committed step towards isopentanol production from 2-ketoisovalerate, a feed-back resistant *LeuA* variant was used (G462D)^{218, 262}.

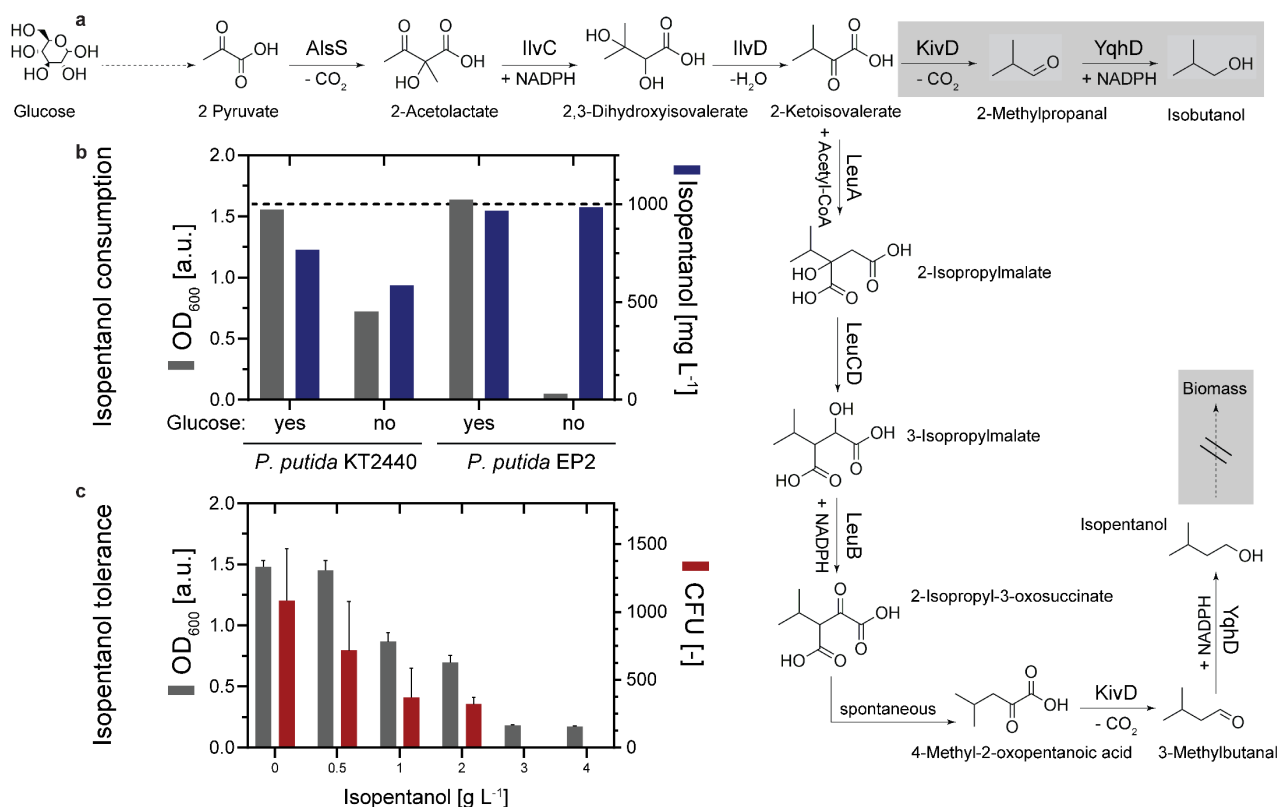


Figure 4-1 Overview of prerequisites for engineering *P. putida* for isopentanol production. a) Overexpression pathway for isopentanol production via Ehrlich degradation. Grey boxes indicate unfavorable reactions leading to side product formation (isobutanol) or product metabolization. b) Isopentanol consumption and biomass formation with *P. putida* KT2440 and *P. putida* EP2 (2 mL culture volume, 24 h, 30°C, n=1). All minimal media cultures were supplemented with 1 g L⁻¹ isopentanol (dotted line), additional 5.0 g L⁻¹ glucose were added as indicated. c) Isopentanol tolerance of *P. putida* EP2 and cell viability after incubation (750 μ L culture volume, minimal medium with 5.4 g L⁻¹ glucose, 24 h, 30°C, n=3). OD₆₀₀ was measured with a plate reader (Tecan M1000 Pro). Abbreviations: CFU, colony forming units.

First, the isobutanol pathway and the *leuABCD* operon were separately implemented under the control of an L-arabinose inducible *araC/P_{BAD}* expression system on two different broad-host range plasmids with pBBR1 and RK2 origins of replication, respectively (overexpression constructs A and B in Figure 4-2). For overexpression of the isobutanol pathway, induced by addition of 0.2% (w/v) L-arabinose, we found titers of 165 mg L⁻¹ isobutanol and 28 mg L⁻¹ of isopentanol as a side product (strain no. 1). The isopentanol side-product was expected, as the strain still contains the natural *leuABCD* genes which are essential for leucine synthesis and encoded by the host genome. Upon additional heterologous *leuABCD* overexpression, the main product shifted to 170 mg L⁻¹ of isopentanol with about 35 mg L⁻¹ isobutanol as side product (strain no. 2). For both strains,

we found a reduced final biomass with an OD₆₀₀ of about 0.7, as compared to 2.2 for a negative control strain that contained 2 plasmids without any of the pathway genes. We also tried to overexpress the full nonacistronic operon for isopentanol production from a single plasmid (construct C). However, here only a comparatively small titer of 23 mg L⁻¹ isopentanol was obtained. Next, we stably integrated the isobutanol pathway downstream of *glmS* into the host genome (construct D). For this strain variant (no. 5), we obtained a mixture of 30 mg L⁻¹ isobutanol and 22 mg L⁻¹ isopentanol. Further, we randomly integrated the *leuABCD* operon with a Tn5-based approach²⁶³. This allowed a shift towards more isopentanol specific production with 66 mg L⁻¹ isopentanol and 13 mg L⁻¹ isobutanol as side product. For plasmid-based *leuABCD* overexpression (construct F), this shift towards isopentanol production was further increased with 95 mg L⁻¹ isopentanol and not significant isobutanol product formation. It should be noted that here the *leuABCD* operon contained additionally a gene encoding a far-red fluorescent protein (*mCardinal*, constructs E and F) in order to characterize differences in expression levels. For the plasmid-based, multi-copy expression strain (no. 6) a 14-fold stronger fluorescence was found as compared to the genome-integrated variant (no. 4, see Supplementary Figure 4-1).

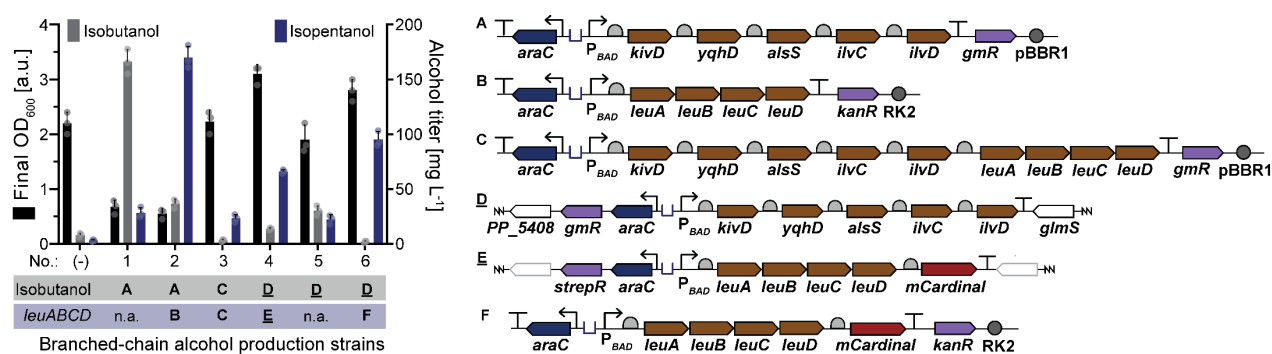
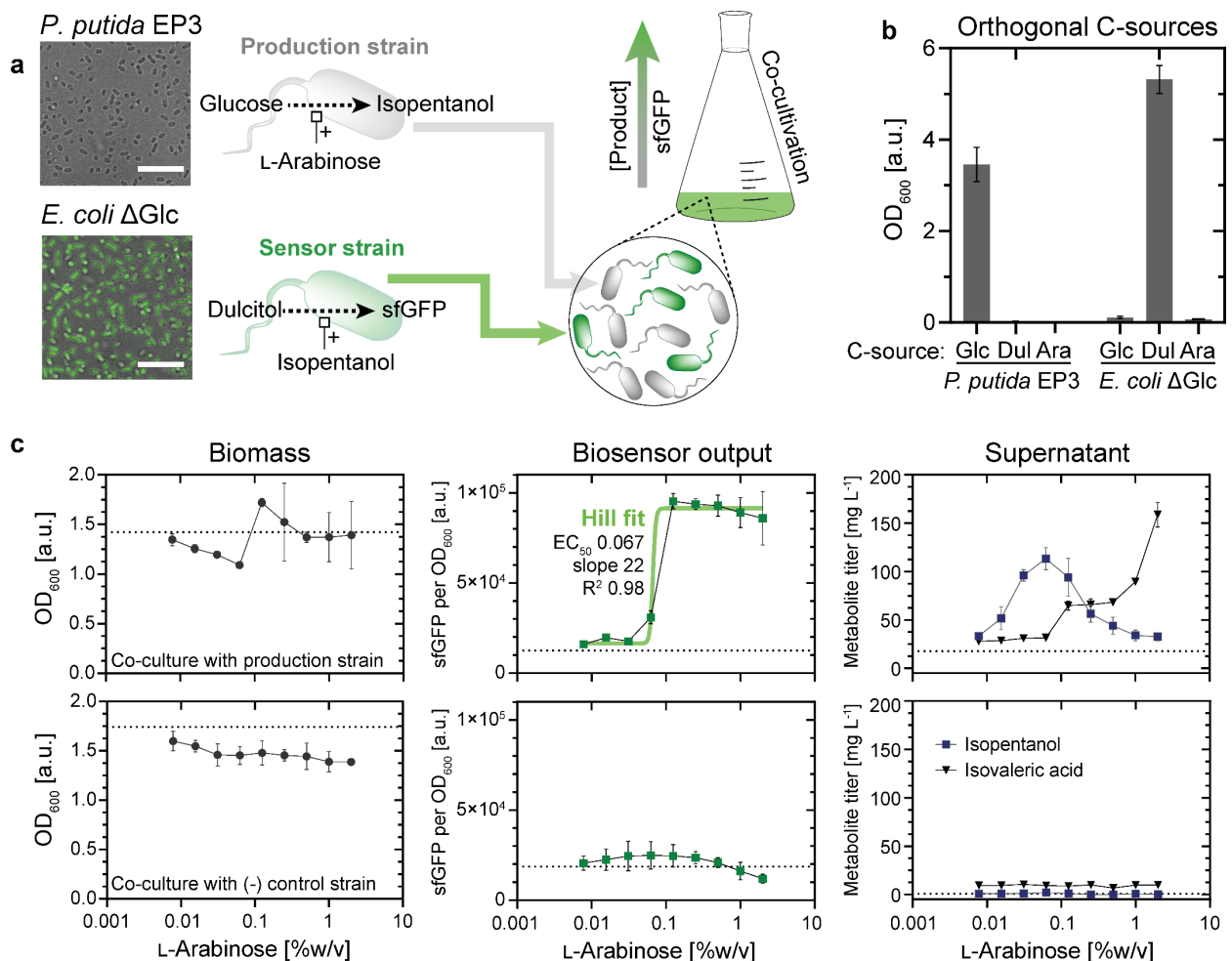


Figure 4-2 Branched-chain higher alcohol production with *P. putida* EP2 and different pathway expression variants. Left: Biomass (OD₆₀₀) and alcohol titers obtained with different production strains in defined medium with 5.4 g L⁻¹ glucose (750 μL culture volume, 36 h, 30°C, n=3). Right: Genetic constructs for enzyme overexpression as indicated for the production strain variants. Please note, underlined constructs D and E were integrated in the host genome. The (-) control strain was transformed with two empty plasmids used for the construction of A and B (pSEVA631 and pSEVA221, respectively).

Isopentanol production monitored *in situ* with a biosensor co-culture

As the evaluation of strain variants for isopentanol production by gas- or liquid-chromatography becomes very laborious and time-consuming for increasing numbers of variants due to the relatively cumbersome protocols²⁶, we developed a co-culture based detection system employing a biosensor. To this end, we constructed an *E. coli* strain with

three knockouts that no longer grew with glucose^{264, 265} (Δglk , $\Delta ptsG$, $\Delta manZ$, abbreviated as ΔGlc), which was supposed to serve as the C-source for alcohol production with *P. putida* (here EP3, a derivative of EP2, but with an unrelated, additional knock-out of the *benABCD* cluster for prevention of 3-methyl benzoate consumption²⁶⁶). In order to find an orthogonal C-source used only by *E. coli*, both sensor and production strain were tested with a commercial kit for utilization of 49 different C-sources (API 50 CH, BioMérieux, see Supplementary Figure 4-2). First, this experiment confirmed that only *P. putida* could metabolize glucose under the test conditions, but not the *E. coli* ΔGlc strain. Second, we found that orthogonal C-sources for *E. coli* ΔGlc included three sugar alcohols (dulcitol, mannitol, sorbitol) as well as N-acetylglucosamine, fucose and ribose, none of which served as a C-source for *P. putida* EP3. We continued with dulcitol, a reduction product of galactose (Figure 4-3b). The *E. coli* ΔGlc strain was transformed with two plasmids encoding an AlkS L401G-based biosensor system²⁶⁰. Upon induction with isopentanol, the activated AlkS regulator binds to its cognate promoter sequence thus activating sfGFP production as biosensor output. Consequently, isopentanol production could be assessed as fluorescence output instead of chemical analysis, saving time and materials in strain evaluation²⁶. As isopentanol readily diffuses across microbial membranes, we hypothesized that a co-culture of the *P. putida* EP3 isopentanol producer (harboring plasmid-based overexpression constructs A and B, see Figure 4-2) with the *E. coli* ΔGlc biosensor strain for product-detection should be feasible. The co-cultures were implemented as follows (also see Figure 4-3a): The defined cultivation medium included two different C-sources, glucose for the production strain and dulcitol for the sensor strain. Additionally, for varied induction of isopentanol production, varied concentrations of L-arabinose were added to the culture medium (up to 2% (w/v), Figure 4-3c). After 24 h of incubation at 30°C for isopentanol production by *P. putida*, the co-cultures were incubated for an additional 12 h at 37°C in order to allow further growth of the *E. coli* biosensor strain. The fluorescence per biomass of the induced co-cultures increased up to seven-fold, as compared to the non-induced ones. In chemical analyses of the culture supernatants, we found the minimal isopentanol titer for biosensor activation to be approximately 100 mg L⁻¹. However, we also found that isopentanol titers decreased for increasing L-arabinose concentrations from 0.13% (w/v) onwards, with a concomitant increase of isovaleric acid titers. For the same co-cultivation experiments but with a *P. putida* EP3 control strain lacking the overexpression pathway, we found neither biosensor activation nor isopentanol or isovaleric acid production, further validating the co-culture approach for product detection.



Branched-chain alcohol production in shake flasks and bioreactors

For the upscaling of isopentanol production from deep well plate format (750 μL) to bench-scale bioreactors ($V_0 = 1 \text{ L}$), we used a two plasmid-based expression system (Figure 4-4a). First, we validated the switch to isopentanol from isobutanol formation upon expression of the *leuABCD* cluster and checked for basal alcohol production without induction of pathway overexpression. As before, we found isopentanol as the main product (125 mg L^{-1}) with small amounts of isobutanol as side product (10 mg L^{-1}). For the uninduced cultures we found about 17% of the isopentanol titer remaining (22 mg L^{-1}), indicating little but not absent basal enzyme expression (Figure 4-4b). Next, we evaluated biomass formation and alcohol production over time in shake flasks (Figure 4-4c). Upon reaching stationary phase after approximately 15 h, 167 mg L^{-1} isopentanol had been produced with an OD_{600} of 2.4 and a maximum specific growth rate μ_{max} of 0.31 h^{-1} . This corresponds to a volumetric isopentanol production rate of $11 \text{ mg L}^{-1} \text{ h}^{-1}$. Assuming full glucose consumption and $0.32 \text{ g}_{\text{CDW}} \text{ OD}_{600}^{-1}$ (see Supplementary Figure 4-3), this corresponds to biomass and product yields of $Y_{X/S} = 0.14 \text{ g}_{\text{CDW}} \text{ g}_{\text{Glc}}^{-1}$ and $Y_{P/S} = 31 \text{ mg}_{\text{isoC5OH}} \text{ g}_{\text{Glc}}^{-1}$, respectively. After 15 h of cultivation, the OD_{600} increased slightly to 2.9 while the isopentanol titer decreased to 125 mg L^{-1} . Over the whole production time isobutanol titers were below 6 mg L^{-1} . For a negative control strain, we found an increased maximum OD_{600} of 4.2 after 15 h (μ_{max} of 0.37 h^{-1} , $Y_{X/S} = 0.25 \text{ g}_{\text{CDW}} \text{ g}_{\text{Glc}}^{-1}$), with no alcohols produced.

In order to investigate the limitations in isopentanol production and the formation of isovaleric acid as observed for co-culture experiments, we also conducted shake flask experiments for isopentanol production in culture media with varied product (isopentanol) or precursor (4-methyl-2-oxopentanoic acid) concentrations added prior to incubation (Supplementary Figure 4-5 and 4-6, respectively). In general, we found isopentanol titers of up to 150 mg L^{-1} isopentanol, with minimal isobutanol (5 mg L^{-1}) and approximately 550 mg L^{-1} isovaleric acid as side products when none of the two compounds had been supplemented to the cultures. Interestingly, for isopentanol added to the culture medium at 100 or 400 mg L^{-1} (well below the toxicity level, see Figure 4-1), we found no improvement in isopentanol titers. On top of 100 mg L^{-1} isopentanol added, only 25 mg L^{-1} were produced, resulting in a very similar titer as obtained for cultures without pre-added product (125 and 129 mg L^{-1} , respectively). For 400 mg L^{-1} isopentanol added, we measured a reduction to 225 mg L^{-1} over time, a net loss of 175 mg L^{-1} . However, for isovaleric acid we found titers increasing with pre-added isopentanol concentrations, from 515 mg L^{-1} (no isopentanol added) to 604 mg L^{-1} (100 mg L^{-1} isopentanol added) and 726 mg L^{-1} (400 mg L^{-1} isopentanol added). Further, the

isopentanol precursor 4-methyl-2-oxopentanoic acid was added to 150, 300, or 600 mg L⁻¹ in the culture medium. For the two highest concentrations added, the maximum OD₆₀₀ was reduced to 3.1 and 1.8, respectively, as compared to an OD₆₀₀ of 3.7 otherwise. For these two highest precursor concentrations added, we also found improved isopentanol titers of 159 and 192 mg L⁻¹, respectively, as compared to 140 mg L⁻¹ otherwise. For isovaleric acid, we again found constantly increasing titers of 702 mg L⁻¹, 923 mg L⁻¹, and 1.4 g L⁻¹, respectively, as compared to 620 mg L⁻¹ without precursor added. Overall, the experiments showed that non-toxic isopentanol concentrations added to the culture medium did not improve overall product titers, while supplementation of an isopentanol pathway precursor allowed increased isopentanol titers (up to 1.4-fold). However, this effect was more pronounced for the titers of the by-product isovaleric acid (up to 2.4-fold increase).

In order to increase biomass and isopentanol production as well as improving process control, we developed two fed-batch protocols for bench-top bioreactors ($V_0 = 1$ L, see Figure 4-4d). First, a batch phase with a defined high cell density medium²⁶⁷ and an initial glucose concentration of 10 g L⁻¹ was implemented. Until glucose depletion after approximately 12 h, as indicated by a sudden increase of the dissolved oxygen tension (DOT), the cells grew with a maximum specific growth rate of $\mu_{\max} = 0.31$ h⁻¹ to a cell dry weight (CDW) of 2.7 g L⁻¹. This corresponds to a $Y_{X/S} = 0.27$ g_{CDW} g_{Glc}⁻¹, a value similar to the biomass yield calculated for the control strain in shake flasks. Next, glucose was fed to the culture at a rate of 1.5 g h⁻¹. After 1 h of feeding, isopentanol production was induced by adding L-arabinose to 0.2% (w/v) in the culture. In the next 4 h, 271 mg L⁻¹ isopentanol were produced, as well as 185 mg L⁻¹ isovaleric acid. In the following 12 h the isopentanol concentration increased slowly to 312 mg L⁻¹, however the isovaleric acid concentration increased further to 650 mg L⁻¹. Over the full induction time of 15 h no significant isobutanol production (concentrations remained below 20 mg L⁻¹) and no accumulation of glucose was found (below 200 mg L⁻¹), while the biomass reached a maximum CDW of 5.9 g L⁻¹.

In the second protocol, we changed the aeration conditions during the production process. While air was supplied with a submerged sparger as before (DOT > 40%) during batch-phase, for production of isopentanol the aeration was restricted to the bioreactor headspace ("microaerobic conditions"). Previously, a similar strategy had been shown to be beneficial for NADPH-dependent isobutanol production²⁵⁸. As the pathway to isopentanol requires even an additional NADPH per product molecule, oxygen limitation could potentially improve production here as well. Besides, conditions with an increased NAD(P)H/NAD(P)⁺ ratio could be beneficial for decreasing side product formation, as aldehyde dehydrogenases,

potentially catalyzing the reaction from isovaleraldehyde (3-methylbutanal) to isovaleric acid, require NAD(P)⁺ ²⁶⁸. For the first hour of isopentanol pathway induction, the DOT was still maintained above 40% in order to allow for unrestricted enzyme overexpression. During this time, 50 mg L⁻¹ isopentanol were produced, while isobutanol and isovaleric acid concentrations were below 10 mg L⁻¹. Next, the conditions were switched to microaerobic conditions, which led to an immediate DOT reduction to 0% and no further increase in CDW, which remained at a maximum of about 4.0 g L⁻¹. Under these oxygen-limited conditions, 553 mg L⁻¹ isopentanol were produced in a total process time of 27 h (including 9 h of induction). Isobutanol titers increased as well, to 210 mg L⁻¹. At the same time, isovaleric acid production was limited to 151 mg L⁻¹. However, glucose started to accumulate linearly at a rate of approximately 0.9 g L⁻¹ h⁻¹ to a final concentration of 9.2 g L⁻¹. In comparison to aerobic production, this microaerobic process-strategy allowed a 1.8-fold improvement in isopentanol titer, while isovaleric acid production was reduced 4.3-fold. However, additional studies are required to investigate how the strictly aerobic *P. putida* metabolism reacts to microaerobic conditions and how NAD(P)H availability could be improved further.

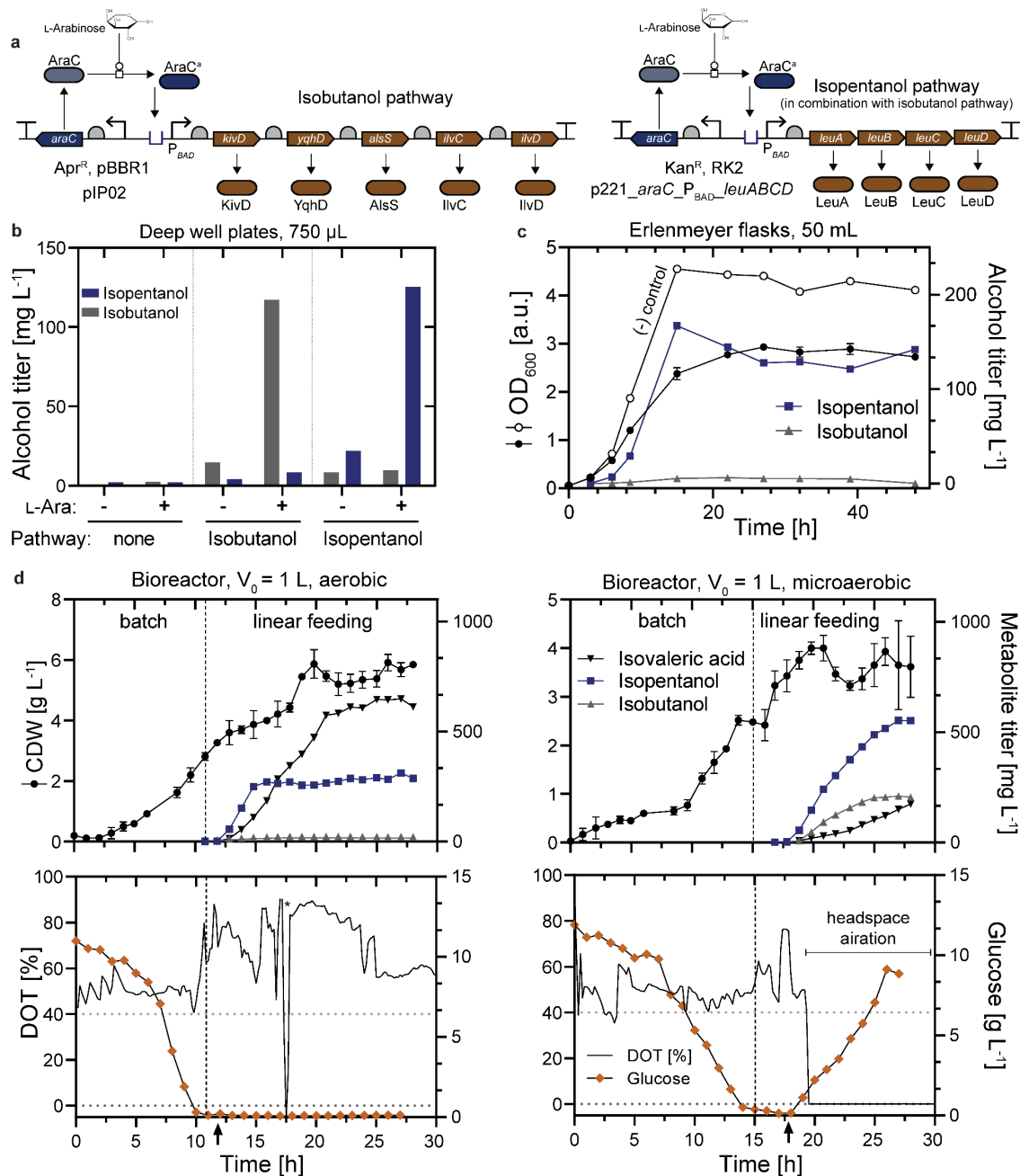


Figure 4-4 Branched-chain alcohol production with *P. putida* EP2. a) Implementation of Ehrlich degradation pathway for isopentanol production. First, overexpression of the full isobutanol pathway is required (*ori* pBBR1, apramycin resistance). Second, for isopentanol production *leuABCD* is overexpressed in addition to the isobutanol pathway (*ori* RK2, kanamycin resistance). b) Verification of isobutanol and isopentanol production with the two overexpression systems upon induction with L-arabinose (L-ara). Triplicate cultures, bars represent pooled supernatant analysis after 48 h of cultivation in deep well plates. c) Biomass formation (OD_{600}) and isopentanol production over time in shake flasks. $n=3$ for isopentanol production strain with pooled supernatant analysis, $n=1$ for (-) control. For clarity only the biomass of the (-) control strain is shown, but no alcohol titers (no alcohol production found). d) Two fed-batch protocols with either normal aeration (“aerobic”, $\text{DOT} > 40\%$) or headspace aeration (“microaerobic”) during isopentanol production. The batch phase started with an initial glucose concentration of 10 g L^{-1} , followed by a linear feed with a rate of $1.5 \text{ g}_{\text{Glucose}} \text{ L}^{-1} \text{ h}^{-1}$. Black arrows indicate time points of induction of isopentanol production. Abbreviations: CDW, cell dry weight; DOT, dissolved oxygen tension; *, air-supply filter replaced. The (-) control strain was transformed with two empty plasmids (pNG413.1 and pSEVA221), the isobutanol production strain with pIP02 and pSEVA221.

4.4 Discussion

Initial experiments with *P. putida* EP2, a strain with multiple alcohol and aldehyde dehydrogenases deleted from the genome¹⁸⁴, indicated the feasibility of the strain for isopentanol production, with a tolerance of up to 2 g L⁻¹ isopentanol while also not metabolizing the product. Further metabolic engineering of the strain allowed the production of about 170 mg L⁻¹ isopentanol in shake flask experiments and this titer was further increased to more than 0.5 g L⁻¹ in a two-stage production process with aerobic biomass production followed by microaerobic conditions for isopentanol synthesis. Similar observations were previously made for the production of isobutanol with the same strain¹⁸⁴. The biomass and product yields of $Y_{X/S} = 0.14 \text{ g}_{\text{CDW}} \text{ g}_{\text{Glc}}^{-1}$ and $Y_{P/S} = 31 \text{ mg}_{\text{isoC5OH}} \text{ g}_{\text{Glc}}^{-1}$, respectively, found here for isopentanol production in shake flasks fit well to the ones previously published for isobutanol production, with $Y_{X/S} = 0.13 \text{ g}_{\text{CDW}} \text{ g}_{\text{Glc}}^{-1}$ and $22 \text{ mg}_{\text{isoC4OH}} \text{ g}_{\text{Glc}}^{-1}$. Further, in a recently published bioprocess study with a 30 L bioreactor, more than 3 g L⁻¹ isobutanol were successfully produced with an optimized microaerobic production phase²⁵⁸, suggesting that g per liter product titers might be feasible for isopentanol production. To the best of our knowledge, this study is also the first metabolic engineering effort targeting isopentanol production with *P. putida*. Previously, isopentanol was only reported as a side product (about 5 mg L⁻¹ in defined medium with 20 g L⁻¹ glucose²⁵²). While higher isopentanol titers with up to approximately 3 g L⁻¹ were previously reported for comparable cultivations in shake flasks with more established host organisms such as *E. coli*¹⁸³ and *C. glutamicum*¹⁷⁹, further engineering of *Pseudomonas* spp. could yield additional benefits. For example, processes harnessing more solvent tolerant strains²⁴⁹ and lignin-derived C-sources would be highly desirable²⁵¹.

Considerable formation of isovaleric acid as a side-product was observed, in particular when supplementing pathway precursors to the culture medium or during extended incubation times under unfavorable conditions such as increased cultivation temperature. In addition to process based solutions, as shown here for a microaerobic production phase, side-product formation could likely be reduced by deleting further aldehyde dehydrogenases (ALDH) from the host genome²⁵⁷, assuming that isovaleric acid accumulation is due to aldehyde dehydrogenase activity. A crucial first step of subsequent studies of alcohol metabolism would be to test which of the about 30 known ALDHs are the most detrimental ones left for pure alcohol biosynthesis²⁶⁸. In a recently published random transposon knock-out library, additional targets for the prevention of alcohol consumption were found²⁶⁹. As before²⁵⁷, multiple knock-outs were required for preventing growth on n-butanol, however, partly

different target genes were identified. A single *pedF* (PP_2675, a cytochrome c oxidase) knock-out prevented growth on isopentanol as sole C-source, but isopentanol was still consumed by this $\Delta pedF$ strain when the alcohol was supplemented in complex media cultures. Overall, these results support that the alcohol as well as aldehyde dehydrogenase functionalities in *P. putida* KT2440 are widely redundant, but *pedF* might be a good target for engineering further reduced isopentanol consumption. It needs to be tested which influence the overexpression of YqhD, an alcohol dehydrogenase from *E. coli* catalyzing the final reaction of the isopentanol pathway, had on the metabolism of *P. putida* EP2 in this study. Initial isopentanol supplementation experiments indicated no isopentanol consumption for the EP2 strain. However, in isopentanol production experiments using strain EP2 transformed with the overexpression pathway, pre-added isopentanol was consumed again, indicating that the final metabolic pathway reactions are not orthogonal to the host metabolism. Further, the isopentanol titers in batch cultures were relatively insensitive to the addition of isopentanol or a pathway intermediate (4-methyl-2-oxopentanoic acid) to the culture medium and limited to approximately 200 mg L⁻¹, while isovaleric acid titers significantly increased to up to 1.4 g L⁻¹ with increasing intermediate concentrations added. It remains to be seen why the isopentanol titers achievable were limited, but both enzyme properties, e.g. different substrate affinity for 3-methylbutanal, or regulatory mechanisms, e.g. induction of ALDH expression at increased isopentanol concentrations, seem to be valid reasons.

In co-cultures for *in vivo* product detection, for L-arabinose inducer concentrations of at least 0.063% (w/v) the biosensor output of the co-cultures was increased, while reaching a plateau of maximum sensor output from 0.13% (w/v) onwards. This in good agreement with L-arabinose concentration used previously for induction of isobutanol production of about 0.2% (w/v)^{184, 258}. However, we found that isopentanol titers decreased for increasing L-arabinose concentrations from 0.13% (w/v) onwards, with a concomitant increase of isovaleric acid titers. For *P. putida* it is known that higher L-arabinose concentration can lead to further increased expression strength^{270, 271}. Increased induction levels might reduce the cellular NAD(P)H/NAD(P)⁺ ratio further, thus favoring the ALDH reaction. Theoretically, a potential imbalance of KivD and YqhD activity could become observable at increased induction levels, with upregulation of ALDHs allowing to prevent the accumulation of the toxic intermediate 3-methylbutanal²⁷². In any case, future analysis of the kinetics of isopentanol and isovaleric acid production would be helpful for understanding side product accumulation in more detailed.

However, isovaleric acid is a relevant chemical product as well²⁷³ and most likely its targeted production could be significantly improved in terms of titer and purity by replacing the heterologous alcohol dehydrogenase (YqhD) with an aldehyde dehydrogenase in the overexpression pathway. Previously, a similar strategy based on KivD expression allowed the production of up to 2.4 g L⁻¹ isobutyric acid with *Pseudomonas* sp. VLB120²⁵².

Overall, the co-culture scheme with an orthogonal isopentanol biosensor could greatly improve the scope of future high-throughput screens for alcohol production strains. A previous co-culture screen for isobutanol production relied on two glucose-consuming, auxotrophic *E. coli* strains cross-feeding a pathway intermediate (2-ketoisovalerate)²⁷⁴. However, schemes directly detecting the product in miniaturized, microfluidic co-culture assays could be highly beneficial, for instance as shown previously for screening of a *Bacillus subtilis* strain library for improved vitamin B2 production⁵⁶. An evident example application of such co-culture based screens for further improved isopentanol production strains would be the screening of random, transposon based libraries²⁷⁵. For example, this could yield highly expressed pathway variants integrated at suitable positions on the host-genome, or in the case of random knock-out libraries, strain variants with improved isopentanol production due to reduced side-product formation.

In conclusion, the feasibility of *P. putida* EP2 for isopentanol biosynthesis was demonstrated. We expect the developed production strains to be a well-suited starting point for further strain engineering and process development, in particular in combination with co-culture based biosensor screens.

4.5 Methods

Materials and chemicals

If not stated otherwise, all chemicals including short DNA oligomers were purchased from Sigma Aldrich (Buchs, Switzerland), DNA isolation and purification kits from Macherey-Nagel (Dueren, Germany) or Zymo Research (Irvine, CA, USA), enzymes from New England Biolabs (Ipswich, MA, USA) and DNA fragments from IDT (Coralville, IA, USA). PCRs were carried out using high-fidelity Phusion or Q5 polymerases and DNA parts were verified by Sanger sequencing at Microsynth (Balgach, Switzerland).

Bacterial strains, media, and cultivation conditions

In general, all cultivation and cloning procedures followed standard molecular biology protocols²³³ and supplier's recommendations. The strains and plasmids used in this study are summarized in Supplementary Table 4-1 and relevant plasmid maps are available in Supplementary Figure 4-7. For cloning and plasmid maintenance *E. coli* 10 β (*mCherry*)²⁶⁰ was used, except for plasmids with R6K origin of replication, for which *E. coli* PIR2 was used. *E. coli* JKE201²⁷⁶ was used for conjugation. *E. coli* Δ Glc is a BW25113 derivative created by iterative phage transduction²⁷⁷ with Δ *glk*, Δ *ptsG*, and Δ *manZ* strains²¹⁴, followed by removal of the kanamycin resistance gene as described elsewhere²³⁴. *P. putida* KT2440 was used as a control strain in alcohol consumption experiments, derivatives EP2¹⁸⁴ and EP3 were used for alcohol production. EP3 is derivative of strain EP2, but with the *benABCD* cluster deleted from the genome as described previously²⁶⁶. Briefly, the EP2 strain was transformed first with pSNW2 Δ benABCD for homologous recombination and subsequently with pQURE6H for I-SceI based resolution of co-integration. *E. coli* strains were transformed with DNA by electroporation with a micropulser (Bio-Rad, Hercules, CA, USA) at 1.8 kV and cuvettes with a 1 mm gap (Cell Projects, Harrietsham, UK). *P. putida* strains were prepared with 300 mM sucrose as described previously²⁷⁸ and transformed by electroporation at 2.5 kV and cuvettes with a 2 mm gap.

cloning procedures and precultures Lysogeny broth²³³ (LB) was used (Becton Dickinson, Franklin Lakes, NJ, USA). Unless stated otherwise, experiments were conducted in defined minimal medium as described by de Bont et al.²⁷⁹, containing 3.88 g L⁻¹ K₂HPO₄, 1.63 g L⁻¹ NaH₂PO₄, 2.0 g L⁻¹ (NH₄)₂SO₄, 10 mg L⁻¹ EDTA, 100 mg L⁻¹ MgCl₂·6 H₂O, 2.0 mg L⁻¹ ZnSO₄·7 H₂O, 1.0 mg L⁻¹ CaCl₂·2 H₂O, 5.0 mg L⁻¹ FeSO₄·7 H₂O, 0.20 mg L⁻¹ Na₂MoO₄·2 H₂O, 0.20 mg L⁻¹ CuSO₄·5 H₂O, 0.22 mg L⁻¹ CoCl₂, 1.0 mg L⁻¹ MnCl₂·2 H₂O and 5.4 g L⁻¹ glucose. For solid media plates 15 g L⁻¹ agar was added (Applichem, Darmstadt,

Germany). Antibiotics were used, where appropriate, at the following concentrations: 50 µg mL⁻¹ apramycin, 50 µg mL⁻¹ kanamycin, 10 µg mL⁻¹ gentamycin, and 50 µg mL⁻¹ streptomycin. Cultures were routinely cultivated at 37°C (*E. coli*) or 30°C (*P. putida*). Culture volumes occupied 20% of the nominal volume for cultivation tubes (incubated at 1200 rpm, Thermomixer C, Eppendorf, Hamburg, Germany) and 10% (v/v) for Erlenmeyer flasks (120 rpm, Minitron, Infors, Bottmingen, Switzerland) unless stated otherwise. Cultivations in deep well plates were carried out in a miniaturized fermentation system²²⁷ (EnzyScreen, Hemsteede, The Netherlands) with a shaking amplitude of 50 mm at 300 rpm (LT-X, Kuhner, Birsfelden, Sitzerland). For experiments including alcohol supplementation or production, plates were sealed with aluminum foil or screw-cap tubes and flasks were used in order to prevent alcohol evaporation. Cells were kept at 4°C for short-term storage or as cryostocks with 20% (v/v) glycerol at -80°C otherwise.

Cloning of alcohol expression pathways and biosensor strain construction

For insertion of mutation G1385A¹⁸³ (encoding G462D²¹⁸) into *leuA* of plasmid p221_araC_P_{BAD}_leuABCD, site-directed mutagenic PCR with primers pQC_leuA_fwd and pQC_leuA_rev and template p221_araC_P_{BAD}_leuABCD was used. All primers used in this study are summarized in Supplementary Table 4-2. For generating a nonacistronic isopentanol operon, the *leuABCD* operon was amplified from p221_araC_P_{BAD}_leuABCD with primers pleu_operon_fwd and pleu_operon_rev and combined with pIP02 (amplified with pfwd_ISO2 and prev_ISO2) by two-fragment isothermal assembly (ITA), resulting in plasmid pIP02*leuABCD*. The full *araC_P_{BAD}_kivD_yqhD_alsS_ilvC_ilvD_leuABCD* part of pIP02*leuABCD* was amplified with pAvrII_araC_isoalcohol_forward and pNotI_araC_isoalcohol_reverse, digested with AvrII and NotI, and inserted within the corresponding sites of the MCS of pSEVA631, yielding plasmid p631_araC_P_{BAD}_isoC5OH. Similarly, the isobutanol operon (*araC_P_{BAD}_kivD_yqhD_alsS_ilvC_ilvD*) was amplified from pIP02 with the same primer pair, digested and inserted within the corresponding sites of the MCS of pSEVA631, yielding p631_araC_P_{BAD}_isobutanol.

The *mCardinal* coding sequence (BBa_K2348002, obtained from Registry of Standard Biological Parts) with an upstream RBS (predicted TIR of 100043 a.u.²²¹) was synthesized as a DNA fragment ("RBS_mCardinal", see Supplementary Table 4-3) and directly inserted into p221_araC_P_{BAD}_leuABCD by ITA. The corresponding plasmid-based part was amplified by PCR with primers leuD_fwd_for_FXP and leuD_rev_for_FXP using p221_araC_P_{BAD}_leuABCD as template. For cloning of pBAMD1-4_araC_P_{BAD}_isobutanol, an *araC_P_{BAD}_leuABCDmCardinal* fragment was PCR-amplified with primers

araC_leuABCD_for_pBAMD_Tn5_KpnI_fwd and araC_leuABCD_for_pBAMD_Tn5_Sall_rev. This fragment was inserted into pBAMD1-4²⁶³ by restriction (KpnI and Sall) and ligation.

For cloning of sensor input plasmid p671_P_{alkS}_alkS (harboring AlkS L401G), the MCS of pSEVA671 was cut with BamHI and KpnI. A 2.9 kb fragment encoding P_{alkS}_alkS_L401G was PCR-amplified with primers BamHI_alkS_regulator_fwd and KpnI_alkS_regulator_rev using plasmid pCK01_P_{alkS}_alkS_L401G²⁶⁰ as template, subsequently cut with the same restriction enzymes and ligated into the digested MCS of pSEVA671. Subsequently *E. coli* ΔGlc was co-transformed with p671_P_{alkS}_alkS and pAB_P_{alkB}_sfgfp_KanR for construction of the biosensor strain.

Genomic pathway integration

For genomic integration of the isobutanol pathway downstream of *glmS* of *P. putida* EP2, a Tn7 transposase containing plasmid (pTNS2²⁴³) was co-transformed with pBG_araC_P_{BAD}_isobutanol²⁶⁰ and recovered for 8 h (120 rpm, 30°C). Pathway integration at the correct genomic site was verified by PCR with primers Pput_glmSUP and kivD_rev (approx. 2.5 kb fragment). For additional Tn5-based genomic (random) integration of the *leuABCDmCardinal* operon, *E. coli* JKE201 was transformed with pBAMD1-4_araC_P_{BAD}_leuABCDmCardinal and the plasmid subsequently conjugated into a *P. putida* EP2 strain already equipped with the genome-integrated isobutanol pathway. Briefly, about 5 µL of cell culture of each strain were scraped from an agar plate, mixed with an inoculation loop, and restreaked on LB agar. After approximately 12 h, cells were scraped and resuspended in PBS, washed twice and plated on LB. All cultivation steps including *E. coli* JKE201, except the final plating for selection, were carried out with 1 mM meso-2,6-diaminopimelic acid²⁷⁶. 48 cfu were picked into 750 µL LB in deep well plates, incubated for 16 h and subsequently inoculated in minimal medium with 5.4 g L⁻¹ glucose and 0.2% (w/v) L-arabinose as described above. After 20 h, mCardinal fluorescence was measured with a plate reader (Tecan M1000 Pro, Maennedorf, Switzerland) with an excitation wavelength of 604(9) nm and an emission wavelength of 659(20) nm. Ten variants with varied mCardinal expression strength were sampled for metabolite analysis. The strain variant with the highest isopentanol titer produced was further used for pathway expression variant comparisons.

Isopentanol consumption and tolerance

Isopentanol consumption was tested for *P. putida* KT2440 and EP2 in culture tubes with minimal medium and 1 g L⁻¹ isopentanol as the sole C-source (initial OD₆₀₀ of 0.1). For control cultures, additional 5 g L⁻¹ glucose were added. After 24 h, the optical density at 600 nm (OD₆₀₀) was measured in a plate reader (Tecan M1000 Pro) and samples for metabolite analysis were taken. The tolerance of *P. putida* EP2 against isopentanol was tested in minimal medium with 5.4 g L⁻¹ glucose in deep well plates (750 µL culture volume, inoculated with 2.5% (v/v) preculture). Varying concentrations of isopentanol were added to the medium prior to inoculation (up to 4 g L⁻¹). After 24 h of incubation, the OD₆₀₀ was measured as described above. Additionally, the viability of the cells was determined by cultivating 50 µL of samples (diluted by a factor of 10⁶ in LB) on LB-agar plates. After 24 h at 30°C, the plates were photographed and the colony forming units were counted using the image analysis software ImageJ²⁴².

Orthogonal C-source tests

For qualitative C-source experiments API 50 CH test kits (BioMérieux, Marcy-l'Étoile, France) were used. Briefly, precultures of *P. putida* EP3 and *E. coli* ΔGlc were grown in LB medium for 16 h at 30°C, spun down (4000 rcf, 3 min), resuspended in PBS, and subsequently diluted to an OD₆₀₀ of 0.1 in M9 minimal medium²³³ (i.e., without a C-source). Prior to inoculation of the test wells with the cell suspension, either phenol red (pH indicator, approximately 480 µM) or resazurin (redox indicator, approximately 110 µM) was added. C-source utilization was read out via a red to yellow color shift (phenol red) or a more sensitive shift from pink to clear (resazurin), also see Supplementary Figure 4-2.

Alcohol production and co-cultures in deep well plates

Production experiments were conducted in minimal medium with 5.4 g L⁻¹ glucose in deep well plates (750 µL culture volume, inoculated with 2.5% (v/v) preculture). Pathway overexpression was induced directly at the time point of inoculation by addition of 0.2% (w/v) L-arabinose. The plates were sealed with an adhesive aluminum foil and cultivated at 30°C and 300 rpm. For OD₆₀₀ measurements 50 µL samples were taken and diluted in 150 µL PBS and measured with a plate reader (Tecan M1000 Pro). For metabolite analysis, 500 µL (pooled measurements) or 700 µL culture samples were taken. Co-culture experiments were conducted similarly, but with the following differences. The precultures of the production strain *P. putida* EP3 [p631_araC_P_{BAD}_isobutanol; p221_araC_P_{BAD}_leuABCD] and the biosensor strain *E. coli* ΔGlc [p671_P_{alkS}_alkS; pAB_P_{alkB}_sfgfp_KanR] were grown in M9

medium²³³ with 5 g L⁻¹ glucose or dulcitol, respectively. Co-culture experiments were conducted in M9 medium with 5 g L⁻¹ glucose and 2 g L⁻¹ dulcitol as C-sources, respectively, and varied concentrations of L-arabinose for induction of isopentanol production. Strains were inoculated to an initial OD₆₀₀ of 0.2 (production strain) and 0.1 (sensor strain). After 24 h of incubation at 30°C for isopentanol production, the co-cultures were incubated for an additional 12 h at 37°C. OD₆₀₀ and sfGFP fluorescence (excitation wavelength 488(9) nm, emission wavelength 530(20) nm) were measured with a plate reader as described above.

Alcohol production in shake flasks

Shake flask experiments for alcohol production were carried out with minimal medium containing 5.4 g L⁻¹ glucose and with 50 mL culture volume in 500 mL Erlenmeyer flasks. Production cultures were inoculated to an OD₆₀₀ of 0.2 and isopentanol production was induced at the same time by addition of 0.2% (w/v) L-arabinose. Biomass production was regularly measured as OD₆₀₀ with a bench-top photometer (D30, Eppendorf) and 1.4 mL samples were taken for metabolite analysis at the same time.

Alcohol production in bioreactors

Benchtop bioreactors (Minifors 1, Infors) were prepared according to the manual and sterilized by autoclaving at 121°C for 1 h. The sensor for dissolved oxygen (FDA 2, Hamilton, Reno, USA) was calibrated with a two-point calibration program using compressed air and nitrogen (100% and 0% dissolved oxygen tension, respectively). For bioreactor inoculation, first 10 mL of precultures in minimal medium (5.4 g L⁻¹ glucose) were grown for 16 h and then re-inoculated into 90 mL of batch medium (5.0 g L⁻¹ glucose) in 1 L shake flasks and cultivated for another 16 h (seed culture). The batch medium was adapted from Sun et al.²⁶⁷ and contained: 24 g L⁻¹ Na₂HPO₄·12 H₂O, 4.05 g L⁻¹ KH₂PO₄ and 4.7 g L⁻¹ (NH₄)₂SO₄, 0.8 g L⁻¹ MgSO₄·7 H₂O, 10 g L⁻¹ FeSO₄·7 H₂O, 3 g L⁻¹ CaCl₂·2 H₂O, 2.2 mg L⁻¹ ZnSO₄·7 H₂O, 1 mg L⁻¹ CuSO₄·5 H₂O, 0.36 mg L⁻¹ MnCl₂·2 H₂O, 0.3 mg L⁻¹ H₃BO₃, 0.15 mg L⁻¹ Na₂MoO₄·2 H₂O, 0.11 mg L⁻¹ CoCl₂ and 0.02 mg L⁻¹ NiCl₂·6 H₂O. Next, fermentations were inoculated with 100 mL seed culture in 900 mL batch medium (10 g L⁻¹ glucose), resulting in an initial OD₆₀₀ of approximately 0.3. Throughout the fermentation, the dissolved oxygen tension (DOT) was controlled via the rotational speed of the stirrer and maintained above 40%, except for microaerobic production phases. pH was maintained at 7.0 and regulated by the addition of diluted sulfuric acid (10% (v/v) H₂SO₄) or ammonium hydroxide solution (30% (w/v) NH₄OH). During fermentations, rotational speed of the stirrer, DOT, temperature, pH and the rates of all pumps were controlled and recorded via IRIS V5

software (Infors). Samples of 10 mL were taken regularly for analysis of OD₆₀₀, cell dry weight, as well as metabolite and glucose concentrations. After completion of the batch phase, as indicated by a sudden increase in DOT, the culture was fed at a rate of 7.5 mL h⁻¹ with feed solution containing: 200 g L⁻¹ glucose, 3.33 g L⁻¹ MgSO₄·7 H₂O, 10 mg L⁻¹ FeSO₄·7 H₂O, 3 mg L⁻¹ CaCl₂·2 H₂O, 2.2 mg L⁻¹ ZnSO₄·7 H₂O, 1 mg L⁻¹ CuSO₄·5 H₂O, 0.36 mg L⁻¹ MnCl₂·2 H₂O, 0.3 mg L⁻¹ H₃BO₃, 0.15 mg L⁻¹ Na₂MoO₄·2 H₂O, 0.11 mg L⁻¹ CoCl₂ and 0.02 mg L⁻¹ NiCl₂·6 H₂O. Approximately two hours after the start of the feeding process, L-arabinose was added to 0.2% (w/v) in the culture for induction of isopentanol production. For microaerobic production phases, after 1 h of induction the compressed air flow was switched from entering through the submerged sparger inlet to entering via the vessel headspace only.

Metabolite, glucose and biomass analysis

For quantitative alcohol determination, samples were spun down (21130 rcf, 4°C, 10 min; Eppendorf 5424R) and 500 µL of the culture supernatant were diluted in 500 µL water (deep well plate experiments), otherwise 1 mL undiluted supernatant was used. Next, 100 µL of internal standard solution (1% (v/v) n-pentanol in water) were added and the sample briefly vortexed. Subsequently 500 µL methyl tert-butyl ether (SupraSolv, Merck, Darmstadt, Germany) were added and the mixture thoroughly vortexed. If necessary, samples were briefly centrifuged for complete phase separation. For gas chromatography (GC) analysis, 400 µL of the organic phase were transferred into glass vials closed with PTFE snap-caps (Wicom, Maienfeld, Switzerland). The GC system consisted of a 6890N GC-FID with a 7683 auto-sampler and a DB-WAX UI capillary column (30 m with 0.25 mm diameter; all Agilent) with helium as the carrier gas (2 mL min⁻¹; PanGas, Dagmersellen, Switzerland). Samples were injected in splitless mode with a volume of 1 µL. The temperature of the injector and the FID were maintained constant at 225°C and 250°C, respectively. The oven temperature was initially held at 40°C for five minutes. Next, the temperature was increased to 120°C at a rate of 15°C min⁻¹, followed by heating to 230°C at a rate of 50°C min⁻¹ and held constant for four minutes. Concentrations were determined by extrapolation from standard curves based on samples extracted from diluted pure compounds. The glucose concentrations in culture supernatants were determined by an enzymatic assay (D-Glucose UV-Test, R-Biopharm, Pfungstadt, Germany) based on the generation of NAD(P)H upon enzymatic phosphorylation of glucose. Briefly, for each measurement 5 µL of sample supernatant were mixed with 145 µL of master solution containing the enzyme solution, ATP and NADP⁺ in 96-well plates. The absorption at a wavelength of 340 nm was measured in a plate reader

(Tecan M1000 Pro). The measured absorption values were translated into glucose concentrations via a glucose calibration curve (0, 50, 100, 250, and 500 mg L⁻¹). Samples were diluted in PBS whenever the absorption exceeded the calibration range. For cell dry weight analysis, 2 mL culture samples were spun-down in pre-weighted 2 mL micro centrifuge tubes. Supernatants were removed and the remaining cell pellet was dried at 42°C for at least 48 h until the measured weight remained constant.

4.6 Acknowledgements

We thank A. Ankenbauer, R. Nitschel, R. Takors and Ingenza Inc. for providing plasmids pIP02, pNG413.1 and *P. putida* EP2; V. de Lorenzo and his lab members for providing pSEVA and pBAMD plasmids; S. Koebbing and L. Blank for providing pBG14g; N. Wirth for pSNWΔbenABCD and pQURE6H; M-P. Fischer, B. Nebel and B. Hauer for GC-MS support; A. Schulz and FGen GmbH for *E. coli* ΔGlc; J. Koerner for *E. coli* JKE201; T. Lummen and the Single Cell Facility of D-BSSE for microscopic images. We acknowledge financial support from The European Union Project EmPowerPutida (grant ID 635536) for MOB and are indebted to the whole consortium for most valuable feedback and discussions.

4.7 Author contributions

M.O.B. conceived and carried out experimental work as well as data analysis. F. P. developed fermentation protocols and contributed to alcohol production experiments and data analysis (supervised by M.O.B and S.P, summarized in a written MSc Thesis: Pfeifer, Felix (2019) Bioprocess development for branched-chain alcohol production in *Pseudomonas putida*, D-MAVT, ETH Zurich). M.O.B. and S.P. wrote the manuscript. S.P. provided the project outline, supervision and resources.

4.8 Supplementary information

Supplementary Table 4-1 Overview of strains and plasmids used in this study. AmpR, ampicillin resistance; AprR, apramycin resistance, CmR, chloramphenicol resistance; GmR, gentamycin resistance; KanR, kanamycin resistance; StrepR, streptomycin resistance.

Strain or plasmid name	Relevant characteristics	Source/Reference
Strains		
<i>Escherichia coli</i> 10β(<i>mCherry</i>)	<i>E. coli</i> NEB 10β derivative (New England Biolabs, #C3019I), StrepR, allose operon replaced with constitutive <i>mCherry</i> expression cassette, $\Delta alsRBACEK::\Delta PL-mCherry$	260
<i>E. coli</i> PIR2	F ⁻ $\Delta(argF-lac)169$ <i>rpoS</i> (Am) <i>robA1</i> <i>creC510</i> <i>hsdR514</i> <i>endA</i> <i>recA1</i> <i>uidA</i> ($\Delta Mlul$):: <i>pir</i> ⁺	Invitrogen
<i>E. coli</i> JKE201	MG1655 derivative, RP4-2 Tc::[$\Delta Mu1::\Delta aac(3)IV::lacIq-\Delta aphA-\Delta nic35-\Delta Mu2::zeo$] $\Delta dapA::(erm-pir)$ $\Delta recA$ $\Delta mcrA$ $\Delta(mrr-hsdRMS-mcrBC)$	276
<i>E. coli</i> BW25113	KEIO collection, <i>rrnB3</i> $\Delta lacZ4787$ <i>hsdR514</i> $\Delta(araBAD)567$ $\Delta(rhaBAD)568$ <i>rph-1</i>	214
<i>E. coli</i> BW25113 ΔGlc	As above, but Δglk , $\Delta ptsG$, $\Delta manZ$	214, this study
<i>Pseudomonas putida</i> KT2440	Plasmid-free derivative of <i>P. putida</i> mt-2	280, 281
<i>P. putida</i> EP2	As KT2440, but Δupp , $\Delta pedE$, $\Delta pedI$, $\Delta pedH$, $\Delta aldB-I$, $\Delta bkdAA$, $\Delta sthA$	184
<i>P. putida</i> EP3	As EP2, but $\Delta benABCD$	This study

Continues on next page

Strain or plasmid name	Relevant characteristics	Source/Reference
Plasmids		
pSNW2ΔbenABCD	Homology regions for deletion of benABCD, P _{EM7_gfp} , R6K, KanR	266
pQUIRE6H	Conditionally replicating plasmid, XylS, P _{m_trfA} , P _{m_I-SceI} , P _{14g_mrfp} , RK2, GmR	266
pCK01_P _{alkS_alkS}	P _{alkS_alkS} , pSC101, CmR	260
p671_P _{alkS_alkS}	P _{alkS_alkS} L401G, pSC101, GmR	This study
pAB_P _{alkB_sfgfp} _KanR	P _{alkB_sfgfp} , pBR322/rop, KanR	260
pNG413.1	araC_P _{BAD_lacZ} , pBBR1, AprR	184
pip02	araC_P _{BAD_kivD_yqhD_alsS_ilvC_ilvD} , pBBR1, AprR	184
pSEVA221	MCS, RK2, KanR	195
p221_araC_P _{BAD_leuABCD}	araC_P _{BAD_leuABCD} , RK2, KanR (Figure 4-2, construct B)	260
p221_araC_P _{BAD_leuABCDmCardinal}	As above, but additional mCardinal (Figure 4-2, construct F)	This study
pip02leuABCD	araC_P _{BAD_kivD_yqhD_alsS_ilvC_ilvD_I} euABCD, pBBR1, AprR	This study
pSEVA631	MCS, pBBR1, GmR	195
p631_araC_P _{BAD_isobutanol}	araC_P _{BAD_kivD_yqhD_alsS_ilvC_ilvD} , pBBR1, GmR (Figure 4-2, construct A)	This study
p631_araC_P _{BAD_isoC5OH}	araC_P _{BAD_kivD_yqhD_alsS_ilvC_ilvD_I} euABCD, pBBR1, GmR (Figure 4-2, construct C)	This study

Continues on next page

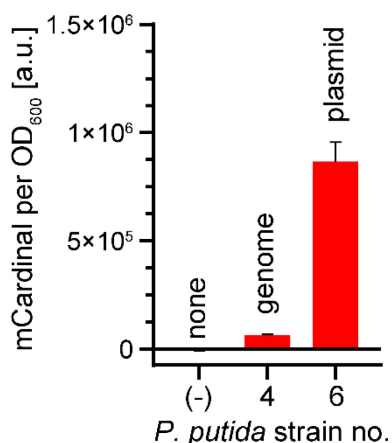
Strain or plasmid name	Relevant characteristics	Source/Reference
pBAMD1-4	Mini-Tn5, R6K, StrepR	263
pBAMD1-4_araC_P _{BAD} _leuABCDmCardinal	As above, with <i>araC_P_{BAD}_leuABCDmCardinal</i> as cargo (Figure 4-2, construct E)	This study
pTNS2	TnsABC+D transposition pathway, R6K, AmpR	243
pBG_araC_P _{BAD} _isobutanol	Tn7L/R, <i>araC_P_{BAD}_kivD_yqhD_alsS_ilvC_ilvD</i> , R6K, KanR, GmR (Figure 4-2, construct D)	260

Supplementary Table 4-2 DNA oligomers used in this study.

Primer	Sequence (5' to 3')
pQC_leuA_fwd	TGCGCTGGATCAGGTGGATATCGTC
pQC_leuA_rev	ACCTGATCCAGCGCATCTTTACCG
pleu_operon_fwd	GGAAGGCCTCTGAGTCATTCGAGCTCTAAGG AGGTTATAAAAACATATG
pleu_operon_rev	CTTCGAACTGCAGGTCTTAATTCATAAACGC AGGTTGTTTTGC
pfwd_ISO2	GCGTTTATGAATTAAGACCTGCAGTTCGAAG TTCCTATTCTC
prev_ISO2	CTCCTTAGAGCTCGAATGACTCAGAGGCCTT CCAGCATTG
pAvrII_araC_isoalcohol_forward	ATACCTAGGGCTTCCGGTAGTCAATAAACCG GTGGATC
pNotI_araC_isoalcohol_reverse	ATAGCGGCCGCGAGAATAGGAACTTCGAACT GCAGGTC
leuD_fwd_for_FXP	CGACTAGTCTTGGACTCCTGTTG
leuD_rev_for_FXP	GCTGCAAGGTAAGCCCAATAC
araC_leuABCD_for_pBAMD_Tn5_KpnI_fwd	AAAATAGGTACCGGATCTCAAGAAGATCCTT TGATCTTTTCTACG
araC_leuABCD_for_pBAMD_Tn5_Sall_rev	AAAATAGTCGACGTAACATCAGAGATTTTGA GACACAAATTTAAATCG
BamHI_alkS_regulator_fwd	ATAGGATCCGTCTTCAAATCGGCGAAGGCC GAAG
KpnI_alkS_regulator_rev	TAGGTACCGCCATAATTCTCTCTCCGGTATA CTTTTCAC
Pput_glmSUP	AGTCAGAGTTACGGAATTGTAGG
kivD_rev	CTAAAAATTGTAAGTTATAGTC

Supplementary Table 4-3 DNA parts synthesized for this study.

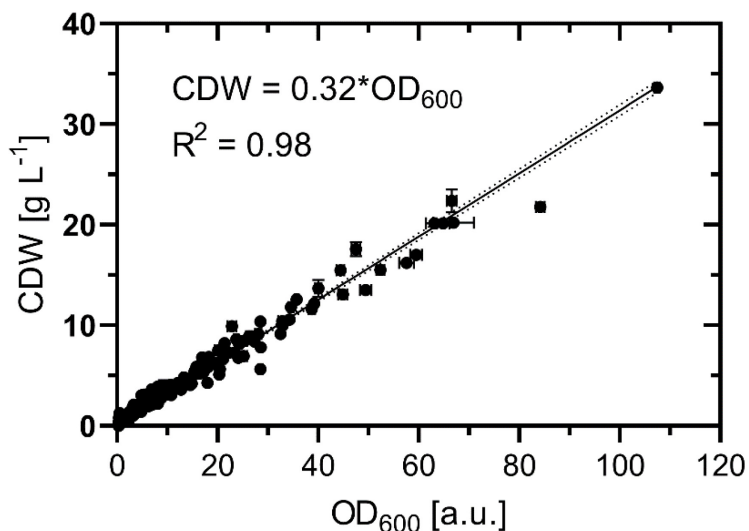
DNA part	Sequence (5' to 3')
RBS_mCardinal (with homology overlaps for isothermal assembly)	CATGATGAACGGTCTGGACAGTATTGGGCTTACCTTGCAGCACGACG ACGCCATTGCCGCTTATGAAGCAAACAACCTGCGTTTATGAATTAA CAACCAGAACTTTAAGGAGGTATAGATGGTGAGCAAGGGCGAGGAGC TGATCAAGGAGAACATGCACATGAAGCTGTACATGGAAGGCACCGTG AACAAACCACCACTTCAAGTGCACCACCGAAGGGGAGGGCAAGCCCTA CGAGGGCAGCCAGACCCAGAGGATTAAGGTGGTGGAGGGAGGCCCCC TGCCGTTTCGATTCGACATCCTGGCCACCTGCTTTTATGTACGGGAGC AAGACCTTCATCAACCACACCCAGGGCATCCCCGATTTCTTTAAGCA GTCCTTCCCTGAGGGCTTCACATGGGAGAGAGTCACCACATACGAAG ACGGGGGGCGTGCTTACCGTTACCCAGGACACCAGCCTCCAGGACGGC TGCTTGATCTACAACGTCAAGCTCAGAGGGGTGAACTTCCCATCCAA CGGCCCTGTGATGCAGAAGAAAACACTCGGCTGGGAGGCCACCACCG AGACCCTGTACCCCGCTGACGGCGGCCCTGGAAGGCAGATGCGACATG GCCCTGAAGCTCGTGGGCGGGGGCCACCTGCACTGCAACCTGAAGAC CACATACAGATCCAAGAAACCCGCTAAGAACCTCAAGATGCCCGGCG TCTACTTTGTGGACCGCAGACTGGAAAGAATCAAGGAGGCCGACAAT GAGACCTACGTCGAGCAGCACGAGGTGGCTGTGGCCAGATACTGCGA CCTCCCTAGCAAACCTGGGGCACAACTTAATGGCATGGACGAGCTGT ACAAGTAACCTGCAGGCATGCAAGCTTGCGGCCGCGTCGTGACTGGG AAAACCCTGGCGACTAGTCTTGGACTCCTGTTGATAGATCCAGTAAT GACCTC



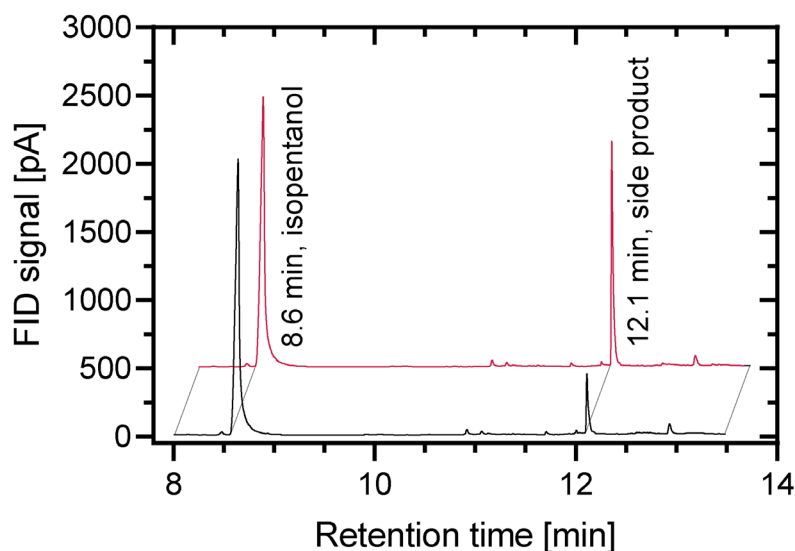
Supplementary Figure 4-1 mCardinal fluorescence of induced *P. putida* branched chain alcohol production strains. (-) is the mean fluorescence of strains not containing mCardinal (n=15), for strains no. 4 and no. 6 mCardinal is expressed as part of a *leuABCD_mCardinal* operon which is either genome-integrated or located on a plasmid (*ori* RK2), respectively (see Figure 4-2). mCardinal fluorescence was measured with a Tecan M1000 Pro plate reader at an excitation wavelength of 604(9) nm and emission wavelength of 659(20) nm. Bars represent means and error-bars are standard deviations (n=3).

C-source	Growth		C-source	Growth	
	<i>P. putida</i>	<i>E. coli</i> ΔGlc		<i>P. putida</i>	<i>E. coli</i> ΔGlc
0 none	-/-	-/-	25 Esculin ferric citrate	-/-	-/-
1 Glycerol	-/w	w/+	26 Salicin	-/-	-/-
2 Erythritol	-/-	-/-	27 D-cellobiose	-/-	-/-
3 D-arabinose	-/-	-/-	28 D-maltose	-/-	-/-
4 L-arabinose	-/w	-/-	29 D-lactose (bovine)	-/-	-/-
5 D-ribose	-/-	w/+	30 D-melibiose	-/-	-/-
6 D-xylose	w/+	+/+	31 D-saccharose	-/-	-/-
7 L-xylose	-/-	-/-	32 D-trehalose	-/-	-/-
8 D-adonitol	-/-	-/-	33 Inulin	-/-	-/-
9 Methyl-β-D-xylopyranoside	-/-	-/-	34 D-melezitose	-/-	-/-
10 D-galactose	+/+	+/+	35 D-raffinose	-/-	-/-
11 D-glucose	+/+	-/-	36 Starch	-/-	-/-
12 D-fructose	-/w	+/+	37 Glycogen	-/-	-/-
13 D-mannose	w/+	-/w	38 Xylitol	-/-	-/-
14 L-sorbose	-/-	-/-	39 Gentiobiose	-/-	-/-
15 L-rhamnose	-/-	-/-	40 D-turanose	-/-	-/-
16 Dulcitol	-/-	w/+	41 D-lyxose	-/-	-/-
17 Inositol	-/-	-/-	42 D-tagatose	-/-	-/-
18 D-mannitol	-/-	+/+	43 D-fucose	+/+	-/-
19 D-sorbitol	-/-	w/+	44 L-fucose	-/-	w/+
20 Methyl-α-D-mannopyranoside	-/-	-/-	45 D-arabitol	-/-	-/-
21 Methyl-α-D-glucopyranoside	-/-	-/-	46 L-arabitol	-/-	-/-
22 N-acetylglucosamine	-/-	+/+	47 Potassium gluconate	-/w	+/+
23 Amygdalin	-/-	-/-	48 Potassium 2-ketogluconate	-/-	-/w
24 Arbutin	-/-	-/-	49 Potassium 5-ketogluconate	-/-	-/-

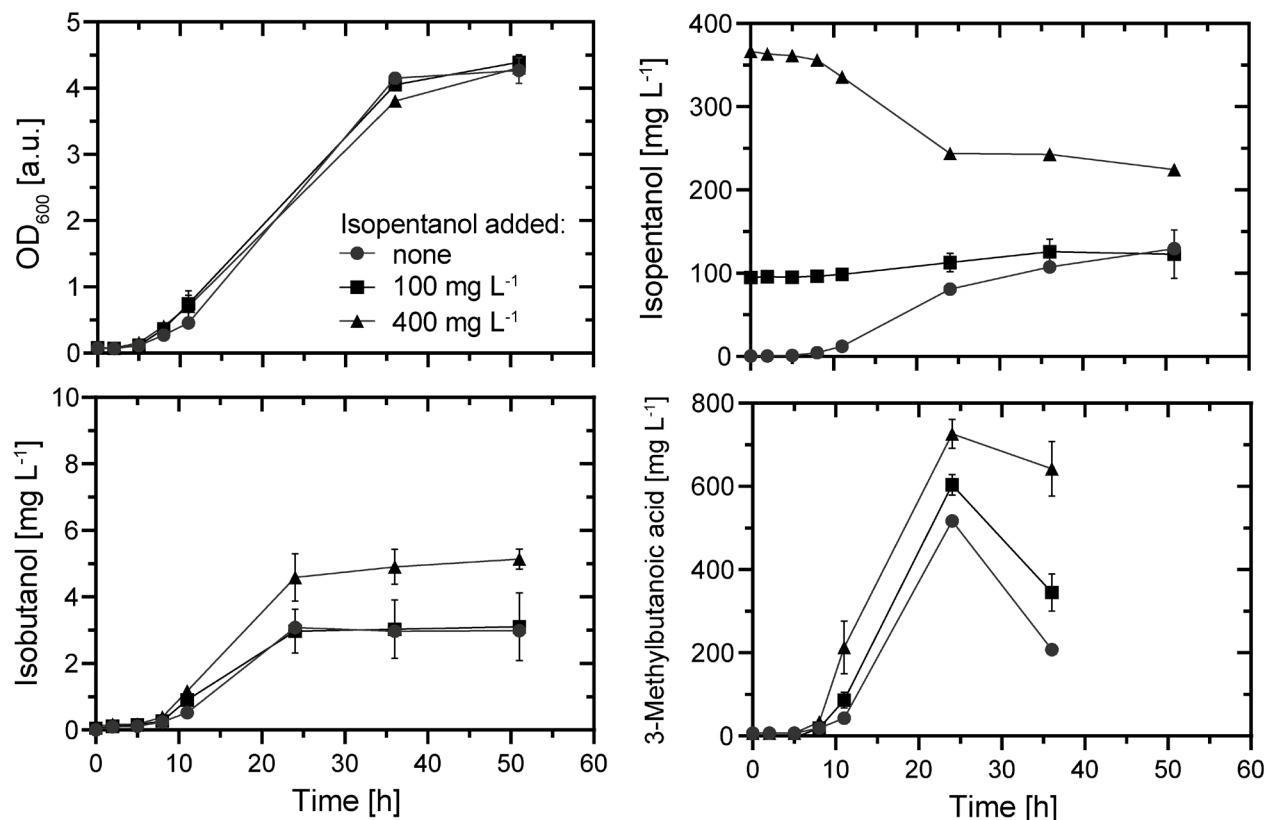
Supplementary Figure 4-2 Qualitative C-source test with M9 medium and *Pseudomonas putida* EP3 or *E. coli* ΔGlc with API 50 CH test kits (BioMérieux, Marcy-l'Étoile, France). Precultures were grown in LB medium, washed twice with PBS and diluted to an OD₆₀₀ of 0.1 in minimal medium without a C-course prior to inoculation of the test wells. C-source utilization was read out via a pH indicator (phenol red, red to yellow color shift) or a more sensitive redox indicator (resazurin, pink to clear) and denoted as phenol red signal/resazurin signal after 72 h at 30°C: -, no color change; +, full color change; w, weak color change.



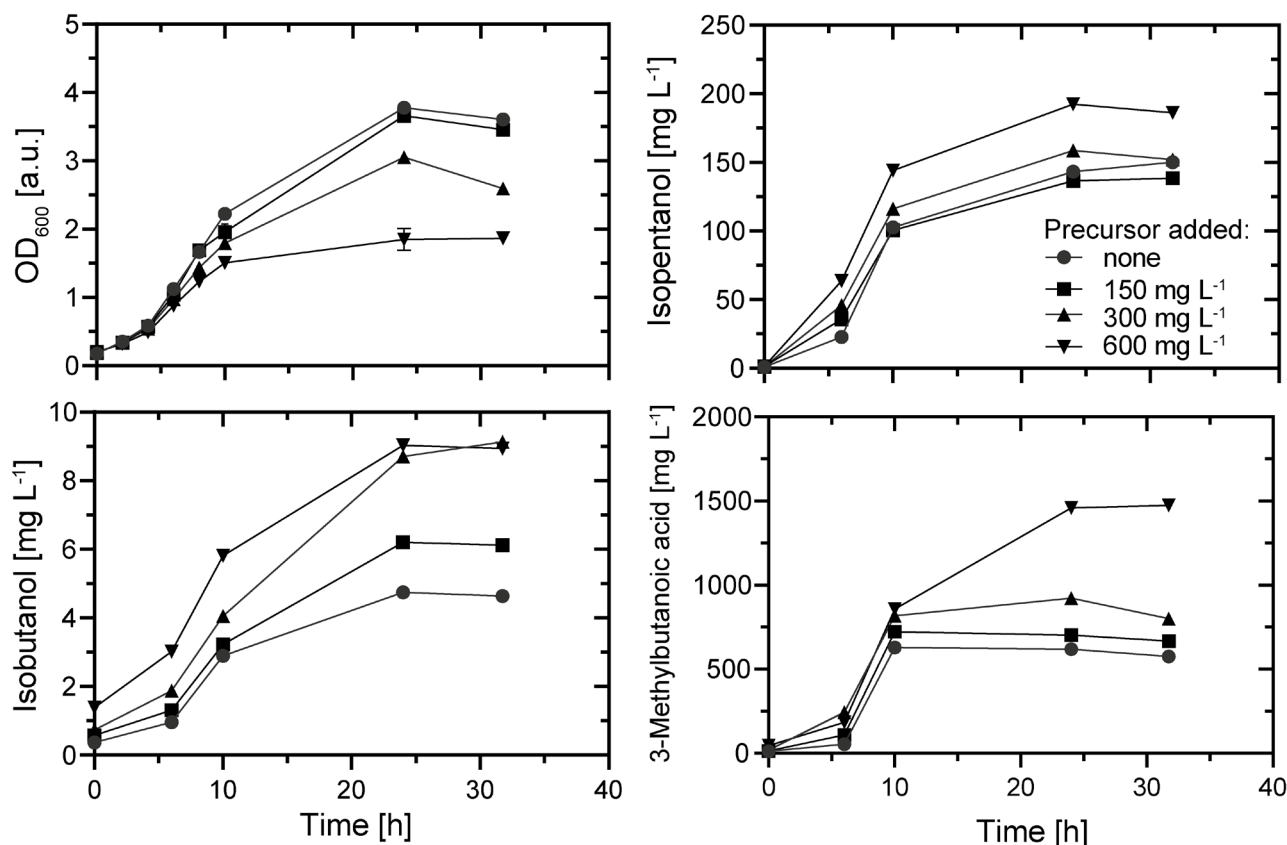
Supplementary Figure 4-3
 Linear fit of *P. putida* EP2 biomass measurements of OD₆₀₀ and cell dry weight (CDW) obtained from independent bioreactor cultivations. Samples with an OD₆₀₀ above 0.8 were diluted in PBS to the linear range of the spectrophotometer. Single data points with error-bars representing standard deviation of triplicate measurements (n=198), dotted lines indicate 95% confidence interval. Fitting with Prism 8.4.2, GraphPad, San Diego, CA, USA.



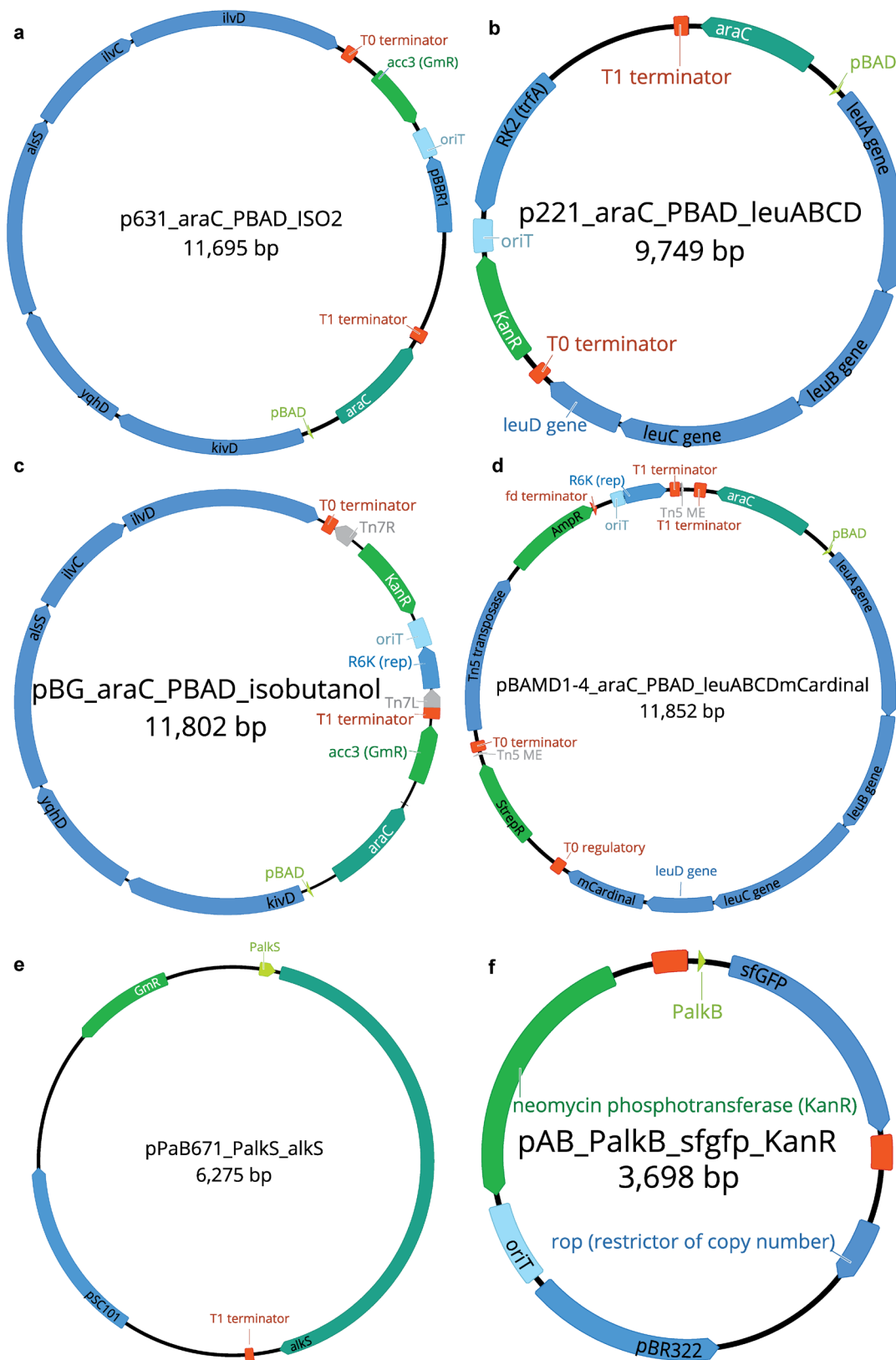
Supplementary Figure 4-4
 Chromatogram of culture supernatant from bioreactor cultivation for isopentanol production (front, black) and the same sample spiked with additional isovaleric acid (back, red). The peak at a retention time of 12.1 min was found as a side product and initial GC-MS analysis suggested isovaleric acid by comparison of the fractionation pattern with a NIST data base (GC-MS is courtesy of M-P. Fischer and B. Nebel, IBT, Stuttgart).



Supplementary Figure 4-5 Kinetics of biomass (OD₆₀₀), isopentanol, 3-methylbutanoic acid (isovaleric acid), and isobutanol production with *P. putida* EP2 [pIS02 p221_araC_P_{BAD}_leuABCD] in shake flask cultures with different concentrations of pre-added isopentanol (50 mL culture volume, defined medium with 5.4 g L⁻¹ glucose, 0.2% (w/v) L-arabinose, initial OD₆₀₀ of 0.2, 30°C). Data points are means of duplicate cultivations with error bars representing standard deviations.



Supplementary Figure 4-6 Kinetics of biomass (OD₆₀₀), isopentanol, 3-methylbutanoic acid (isovaleric acid), and isobutanol production with *P. putida* EP2 [pIS02 p221_araC_P_{BAD}_leuABCD] in shake flask cultures with different concentrations of pre-added 4-methyl-2-oxopentanoic acid, a precursor in isopentanol production (50 mL culture volume, defined medium with 5.4 g L⁻¹ glucose, 0.2% (w/v) L-arabinose, initial OD₆₀₀ of 0.2, 30°C). OD₆₀₀ data points are means of duplicate cultivations with error bars representing standard deviations, for metabolite analyses n=1.



Supplementary Figure 4-7 Plasmid maps for isobutanol and isopentanol production (a, b), genomic integration of isobutanol pathways via Tn7 transposase (c, with pTNS2, not shown), genomic integration of *leuABCDmCardinal* pathway via Tn5 transposase (d), and biosensor plasmids for isopentanol detection in co-cultures (e, f). Maps were created with Geneious software version 11.1 (Biomatters, San Diego, CA, USA).

Chapter 5: Towards Genome Wide sRNA-Based Gene Knock Down for Improved *P. putida* Production Strains

5.1 Abstract

In order to find novel metabolic engineering targets, genome wide and exhaustive gene knock out libraries are a particularly helpful tool. However, most established methods require a one-by-one implementation of deletions or extensive sequencing for random library characterization. In contrast, sRNAs for knock down of gene expression can readily be constructed for a virtually unlimited number of target genes due to their modular and relatively unconstrained architecture. If conditionally expressed from a plasmid, they are also portable between strains and can even target essential genes that cannot easily be deleted. Here, we constructed an inducible expression system for sRNAs, based on the XylS/ P_m system and a Spot 42-derived sRNA scaffold. We demonstrated a reduction in specific fluorescence to about 60% for an anti-GFP sRNA, as compared to an unspecific control sRNA. Further, a general workflow for the construction of genome wide sRNA libraries was established and a corresponding library synthesized and cloned for *P. putida* KT2440. We expect sRNA libraries to be a helpful tool for finding new target genes in metabolic engineering efforts and ultimately for the development of improved production strains.

5.2 Introduction

A particularly promising novel synthetic biology tool is the application of synthetic small RNAs (sRNAs) for metabolic engineering in order to regulate gene translation, as initially shown for *E. coli*²⁸² and more recently also for *P. putida*²⁸³. While classical metabolic engineering approaches usually involve relatively static changes in plasmid copy number, promoter strength, or gene knock out of non-essential genes in order to achieve the intended changes in metabolic fluxes, translational regulation at the RNA level is inducible, modular, multiplex-able, highly dynamic²⁸⁴, and can be applied even to essential genes. Conceptually, small RNAs are designed in order to target the ribosomal binding region²⁸⁵ or the first bases of the coding region^{282, 286} of an mRNA of interest. As a consequence, the formation of the sRNA-mRNA duplex, based on complementary base pairing and facilitated by a chaperone, reduces the mRNA translation rate (see Figure 5-1). The detailed mode-of-action is not fully understood yet, and for example increased mRNA degradation might play an important role in regulation as well^{287, 288}. The modular blueprint of the sRNAs allows to have a common scaffold for interaction with the Hfq chaperone, a general facilitator and stabilizer of the bimolecular binding of sRNA and mRNA, which is preceded by the variable anti-sense target or “seed” region²⁸⁶. Therefore, the modular sRNA structure allows multiplexing gene

regulation with multiple easily programmable sRNAs and without the requirement to, for example, laboriously re-engineer multiple genomic sites²⁸⁹. Furthermore, also essential genes can be regulated in this fashion, which could not be knocked-out at the genome level with methods such as random Tn5-based deletions²⁹⁰. Further, sRNAs targeting essential genes could be exploited for the development of economic zero-growth bioprocesses²⁹¹. Generally, sRNAs are constructed in the same way for any target with a known coding sequence and are thus a molecular biology tool well suited for naïve, genome-wide screens for novel targets in metabolic engineering.

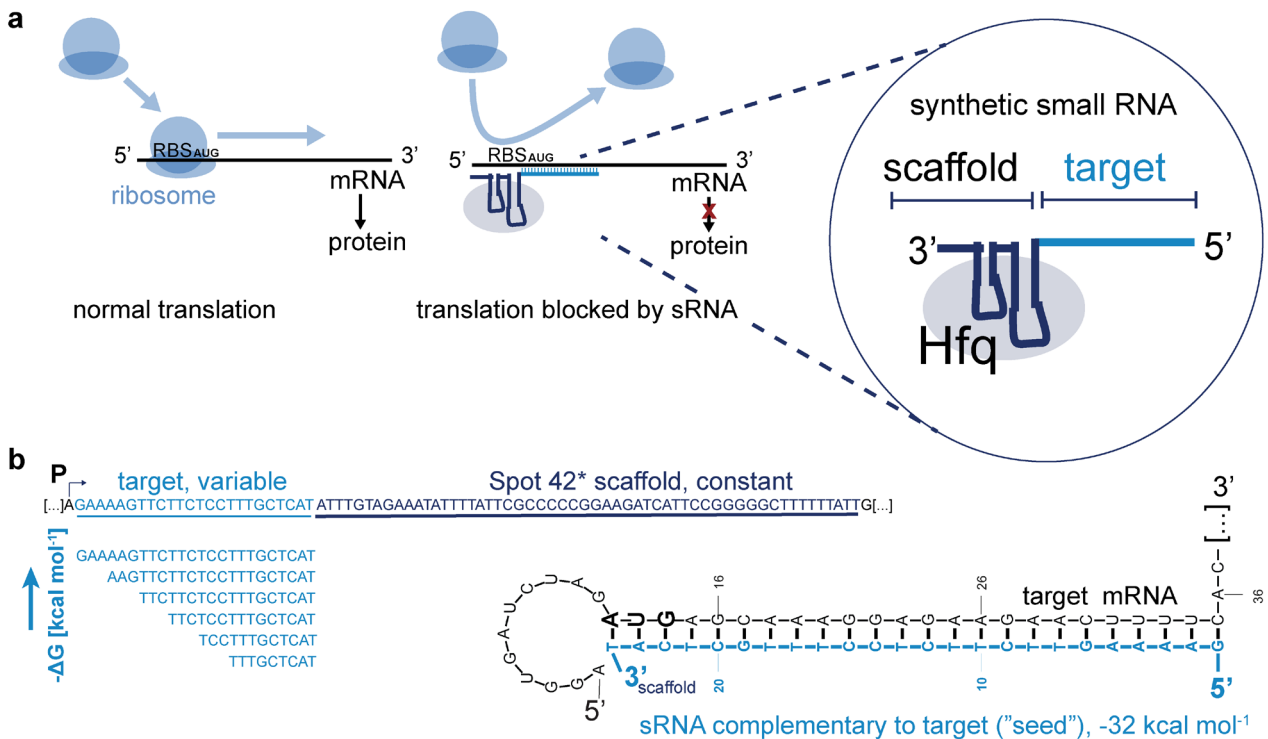


Figure 5-1 Modular design of synthetic small RNA (sRNA) for targeted gene knock down. a) sRNA specifically binds to 5' region of a target mRNA and thus blocks translation initiation and protein production. The sRNA consists of a modular target region, which allows binding to its target mRNA due to complementary base pairing. At the 3' end of the sRNA, a constant scaffold binds the chaperon Hfq (here adapted from Spot 42 sRNA), facilitating sRNA efficiency. b) The length of the target region (also called seed region, i.e. the reverse complement of the target mRNA region) determines the binding strength. Exemplarily, a 24 bp target region against the corresponding mRNA bases is shown (3' scaffold sequence omitted for clarity).

Further, the sRNA seed sequences encode the information of the target gene, thus the library hits are readily verifiable. For example, transposon-based (barcoded) libraries usually require extensive sequencing in order to establish the barcode-knock out connection²⁹². The feasibility of extensive synthetic sRNA libraries was previously shown by screening a library covering 1858 target genes (about 45% of total) in *E. coli* for improved malonyl-CoA

production²⁹³. For the construction of new sRNAs multiple design rules are known^{282, 294}. For reliable repression, the sRNA target region should be around the start codon of mRNAs and about 20 bp to 30 bp in length^{282, 286}, with a binding energy in the range of -30 to 40 kcal mol⁻¹ (see Figure 5-1b). For the constant Hfq binding region²⁸⁷, multiple natural sRNA scaffolds were tested, e.g. of MicC and Spot 42²⁸², as well as sequence-optimized scaffolds developed for improved repression²⁸⁵. More recently, also (inactivated) dCas9, guided by sgRNA to its genomic target²⁹⁵, was successfully used for conditional knock down at the transcriptional level in both *E. coli*²⁹⁵ and *P. putida*^{296, 297}. However, this strategy requires the additional expression of dCas9 and the sgRNA design is more intricate as it requires an adjacent NGG PAM sequence. Further, in a genome-wide library screen sequence-specific toxicity of dCas9 overexpression as well as yet unexplainable off-target effects were found²⁹⁸.

Here we set out to demonstrate the comparable straightforward construction of a genome-wide sRNA library for *P. putida* KT2440, based on available genome annotations²⁵⁶ and economic on-chip, pooled DNA synthesis¹¹.

5.3 Results

For inducible expression of sRNAs we chose the well-characterized XylS/*P_m* system²⁹⁹, which is activated by the addition of 3-methyl benzoate (3MBz). Further, the corresponding pSEVA258 plasmid contained an origin of transfer for conjugation (*oriT*, for efficient transfer of large plasmid libraries), an antibiotic resistance gene (*neo*, kanamycin resistance) and a medium to high copy number origin of replication (RSF1010²⁷⁰). For the constant scaffold region for Hfq-binding, an optimized Spot 42 sequence²⁸⁵ was used (indicated as Spot 42*). In order to test the efficiency of knock-down with this construct, we designed a sRNA target seed against GFP (anti-*gfp*), with 25 bp length and a calculated³⁰⁰ binding efficiency of -46 kcal mol⁻¹. In general, here seed sequence were designed to target the first 25 coding bases²⁸⁶, which facilitates the automated design because coding sequences are precisely annotated while the corresponding 5'UTR regions are usually not part of the genome annotation²⁵⁶. As a control sRNA, we used the reverse-complement sequence of this seed (-17 kcal mol⁻¹), which still can partially bind the target region but with a binding energy well above the recommended -30 kcal mol⁻¹ or lower for strong and specific binding.

As a host strain we constructed *P. putida* EP3_{GFP}, a strain derived from EP2¹⁸⁴, optimized for alcohol production, but further constitutively expressing GFP from the genome²¹⁵ and not metabolizing 3MBz²⁶⁶. This strain was transformed with the two sRNA plasmids, as well as

an empty plasmid control (Figure 5-2). For anti-*gfp* sRNA expression found a significant reduction in GFP fluorescence, to about 60% of the one measured for the strain transformed with the control sRNA. However, the reduction as compared to the no sRNA control was smaller and the overall reduction still substantially different from a control strain expressing no GFP, indicating that a high intracellular GFP concentration was still present. Overall, the results were similar to previous sRNA experiments in *P. putida*²⁸³ and thus we continued with the circuit design for sRNA library construction.

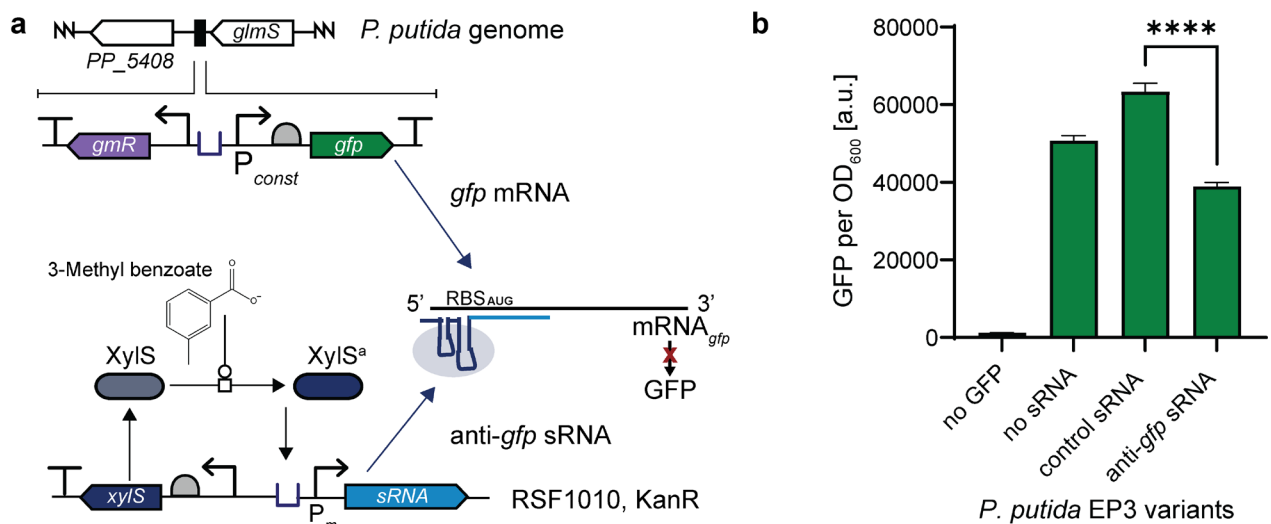


Figure 5-2 Genetic circuits for sRNA-based knock down of GFP production in *P. putida* EP3_{GFP}. a) sRNA is conditionally expressed from a XylS/*P_m* system upon induction with 3-methyl benzoate (3MBz) from a plasmid (RSF1010 origin of replication, kanamycin resistance). GFP is under the control of a strong constitutive promoter, expressed from the genome downstream of *glmS*. b) GFP fluorescence per OD₆₀₀ for *P. putida* EP3_{GFP} variants induced with 5 mM 3MBz after 6 h of induction. The anti-*gfp* sRNA variant has a 25 bp target region for the GFP mRNA (-46 kcal mol⁻¹), the control sRNA is the reverse complement (-17 kcal mol⁻¹). The “no GFP” (*P. putida* EP3) and “no sRNA” control strains were both transformed with pSEVA258. Cultures of 25 mL in shake flasks, measurements with a plate reader (Tecan M1000 Pro, excitation wavelength 488 nm, emission wavelength 530 nm), n=3. For a t-test comparing the two sRNAs, we obtained a P value < 0.0001 (****, with GraphPad Prism software).

To this end, we introduced two additional restriction sites for library cloning in the expression plasmid. An EcoRI site was introduced upstream of the promoter *P_m* and a SacI site downstream of the Spot 42* scaffold (Figure 5-3). Next, we designed three length variants (15 bp, 20 bp, and 25 bp) of seed regions for all annotated coding regions of *P. putida*²⁵⁶, a total of 5564 genes as well as 165 RNAs. At the 5' site, a constant sequence up to and including the EcoRI restriction site was added. At the 3' site, a constant sequence including the Spot 42* scaffold as well as the SacI site was added. This library pool included all

annotated genome sequences as targets, with three seed length (a total of 17187 variants). Further, four additional sub pools with only one seed length (25 bp) were added: all targets (5729 variants), all targets annotated as regulators (436 variants), all targets annotated as RNAs (165 variants), as well as all targets annotated as ncRNAs (67 variants). All sub pools included specific primers³⁰¹ encoded at the 3' end in order to allow for dial out PCRs³⁰² from the full DNA oligomer pool. Here, we continued with PCR amplification of the full pool containing about $1.7 \cdot 10^4$ sRNA variants. These variant were ligated into the expression plasmid backbone and used for transformation of *E. coli*, yielding $3.2 \cdot 10^4$ strain variants. For Sanger sequencing of the corresponding plasmid library pool, we found an error-free P_m sequence but an abrupt signal stop at the variable seed sequence site (NNN), indicating correct assembly. Future NGS experiments will be required to reveal the actual library diversity.

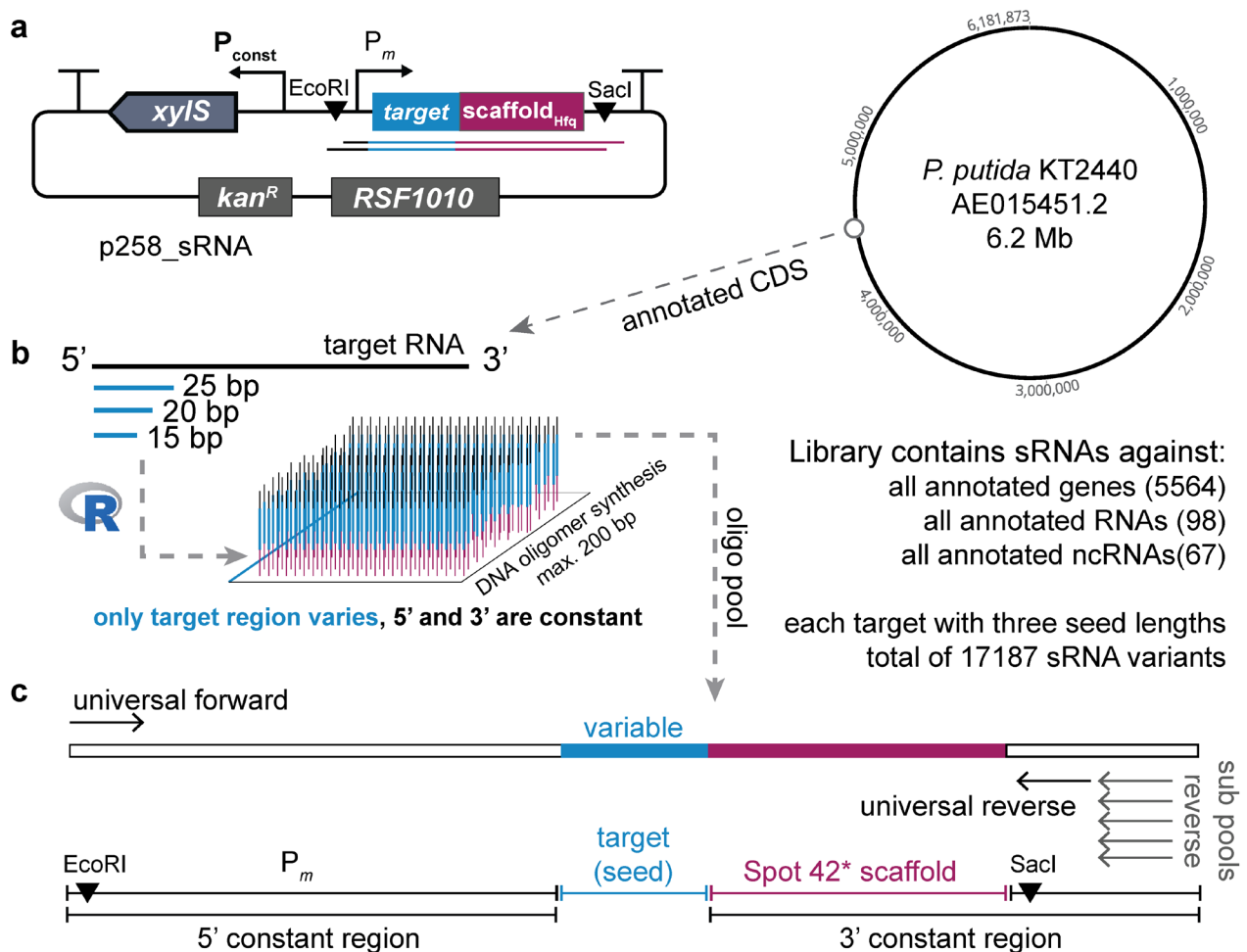


Figure 5-3 Design and synthesis of a genome wide sRNA library for *P. putida* KT2440. a) Genetic circuit for sRNA expression with a 3-methyl benzoate inducible XylS/ P_m system, based on plasmid pSEVA258. The sRNA part consists of a variable target region (up to 25 bp, light blue) and a constant scaffold region for binding chaperone Hfq (adapted from sRNA Spot42²⁸⁵, 55 bp, magenta). The full sRNA (incl. promoter P_m) is located between *EcoRI* and *SacI* restriction sites (up to 167 bp fragment length). b) Genome annotation AE015451.2²⁵⁶ was used in order to generate sRNA target regions of three different lengths against all annotated genes and RNAs, starting at the first base of the predicted RNA sequence. Constant 5' and 3' regions were added to the target (or "seed") regions and the full library synthesized as single-stranded, pooled DNA oligomers of slightly different length (no equalization of DNA molecule length required). c) The full library could be amplified with universal primers and cloned via restriction-ligation into the expression plasmid. Alternatively, sub pools of sRNAs could be selectively amplified using the universal forward primer and pool-specific reverse primers ("dial out"). Annotations²⁵⁶: "RNAs", rRNAs and tRNAs; "ncRNAs", mainly regulatory RNAs.

5.4 Discussion

As expected from the modular nature of sRNA blueprints, it was demonstrated that designing sRNAs for genome wide screens is comparatively straightforward. First, we constructed an underlying RNA expression circuit using a Spot 42-derived scaffold sequence for Hfq binding and an inducible XylS/ P_m system on a plasmid with a RSF1010 origin of replication. For a dummy knock down of constitutively expressed GFP, we found a reduction to 60% of specific fluorescence. Similar results were reported previously for a comparable system²⁸³, with about 75% of specific fluorescence remaining, but using a different origin of replication with a lower copy number²⁷⁰ (pBBR1) as well as a different Hfq scaffold (MicC). As GFP is highly stable, reductions in fluorescence in the timeframe of hours are expected mainly from cell division and the resulting dilution of GFP³⁰³. Nevertheless, previous sRNA studies with *E. coli* found reduction in fluorescent protein expression down to about 1%²⁸², however often with constitutively expressed sRNA²⁸⁵. Interestingly, an optimized sRNA construct allowing more than 150-fold repression in *E. coli*, was found to be less efficient in *Salmonella typhimurium* and *Pseudomonas protegens* (30- and 15-fold, respectively)²⁸⁵, indicating that further sRNA construct improvement for *P. putida* could be possible. For example, improvements might be achieved by optimizing the Hfq-scaffold as well as expression strength³⁰⁴. Besides, actual targets might have a shorter half-life than GFP. Once targets are identified, their downregulation might be optimized or they could be deleted if non-essential^{305, 306}. Genomic targets could also be recoded for inducible protease-based removal³⁰⁷ at the protein level, which might be a powerful approach in combination with sRNA-based knock down at the translational level.

The construction and synthesis of a genome wide library targeting all annotated sequences of *P. putida* KT2440 currently known was efficient and economic due to the availability of software packages³⁰⁸ for the manipulation of DNA sequences and increasingly cheap DNA synthesis, in particular for pooled oligomer synthesis¹¹. Further, this DNA oligomer pool was used for simple one-pot cloning methods, as compared to inserting the single target seed sequences by individual PCR reactions as done previously^{286, 293}.

A potential metabolic engineering application of the sRNA library would be the optimization of isopentanol production with *P. putida*³⁰⁹. Knock down of yet to be identified target enzymes could improve NAD(P)H availability¹⁸⁴ by removing unfavourable reduction equivalent sinks. Further, reactions leading to side product formation, for instance isovaleric acid³⁰⁹, could be reduced and the corresponding enzymes be identified. However, screening of all approximately $2 \cdot 10^4$ strain variants, a number further increased six- to ten-fold by

oversampling²⁹³, is challenging. For example, the previous genome-wide sRNA library in *E. coli* was screened by measuring flaviolin, a red compound, enzymatically derived from (and thus indicative of) increased malonyl-CoA concentrations. Subsequently, only the top fourteen performing sRNAs were applied for exploring actual production of polyketides and phenylpropanoids²⁹³. For the detection of isopentanol, a previously developed co-culture system with an *E. coli* biosensor could be used³⁰⁹. Here, at least for the sub pool targeting every coding sequences with a single sRNA (about $5 \cdot 10^3$ variants), library oversampling in the range of three-fold should be feasible. Still, with the current liquid handling set ups, this requires about forty 384-well plates for each culturing step and an initial cfu picking time of about 75 h²⁶⁰, which might be improved by using more rapid cell arraying methods, e.g. with a flow cytometer. Crucially, before running such a resource-demanding screening campaign and in order to prevent multiple rounds of trial-and-error, it would be most helpful to gain more insight into when to induce sRNAs. An (early) time point of sRNA induction is crucial in order to be able to see an effect on enzyme and product concentrations, while in the optimal case, not leading to the loss of essential gene targets (due to too early induction). Overall, the optimal time point for induction might depend to a large extent on protein half-life and mRNA transcript abundance, as well as how many division the cells are able undergo under screening conditions after induction. Unfortunately, in high-throughput screens culture volumes and biomass formation are restricted and data on enzyme half-lives and transcript concentrations, especially under production (stress) conditions, is only slowly emerging^{310, 311}.

All in all, genome wide sRNA libraries have the clear potential of being a great asset for synthetic biology. sRNAs have a comparatively simple and well-understood blueprint and can be automatically designed, based on widely available genome annotations. As long as there is a suitable experimental readout for testing the comprehensive libraries, they will be of particular help in finding novel gene functions and targets relevant for metabolic engineering.

5.5 Methods

Materials and chemicals

If not stated otherwise, all chemicals including short DNA oligomers were purchased from Sigma Aldrich (Buchs, Switzerland), DNA isolation and purification kits from Macherey-Nagel (Dueren, Germany) or Zymo Research (Irvine, CA, USA), enzymes from New England Biolabs (Ipswich, MA, USA) and DNA fragments from IDT (Coralville, IA, USA). PCRs were carried out using high-fidelity Phusion or Q5 polymerases and DNA parts were verified by Sanger sequencing at Microsynth (Balgach, Switzerland).

Bacterial strains, media, and cultivation conditions

In general, all cultivation and cloning procedures followed standard molecular biology protocols²³³ and supplier's recommendations. The strains and plasmids used in this study are summarized in Supplementary Table 5-1. For cloning and plasmid maintenance *E. coli* 10 β (*mCherry*)²⁶⁰ was used, except for plasmids with R6K origin of replication, for which *E. coli* PIR2 was used. *P. putida* KT2440 EP3_{GFP} was constructed for constitutive GFP production. Briefly, a EP3 strain³⁰⁹ was co-transformed with plasmids pBG14g²¹⁵ and pTNS2²⁴³ for genomic integration of P_{14g}*BCD2_msf_{gfp}* downstream of *glmS*. The correct integration site was verified with primers Pput_glmSUP and Pput_glmSDN²⁴³. *E. coli* strains were transformed with DNA by electroporation with a micropulser (Bio-Rad, Hercules, CA, USA) at 1.8 kV and cuvettes with a 1 mm gap (Cell Projects, Harrietsham, UK). *P. putida* strains were prepared with 300 mM sucrose as described previously²⁷⁸ and transformed by electroporation at 2.5 kV and cuvettes with a 2 mm gap. For cloning procedures and experiments Lysogeny broth²³³ (LB) was used (Becton Dickinson, Franklin Lakes, NJ, USA). For solid media plates 15 g L⁻¹ agar was added (Applichem, Darmstadt, Germany). Antibiotics were used, where appropriate, at the following concentrations: 50 μ g mL⁻¹ kanamycin, 10 μ g mL⁻¹ gentamycin. Cultures were routinely cultivated at 37°C (*E. coli*) or 30°C (*P. putida*). Culture volumes were 20% (v/v) for cultivation tubes (incubated at 1200 rpm, Thermomixer C, Eppendorf, Hamburg, Germany) and 10% (v/v) for Erlenmeyer flasks (120 rpm, Minitron, Infors, Bottmingen, Switzerland) unless stated otherwise. Cells were kept at 4°C for short-term storage and as cryostocks with 20% (v/v) glycerol at -80°C otherwise.

Construction of sRNA expression plasmids

First, pSEVA258¹⁹⁵ was digested with *Age*I and *Spe*I. Subsequently a synthetic DNA part (sRNA_backbone_empty_for_lib), cut with the same two enzymes, was inserted by ligation

to construct plasmid p258_sRNA_scaffold. For primers and synthetic DNA parts used in this study, see Supplementary Table 5-2 and Supplementary Table 5-3, respectively. This DNA part contained a P_m promoter with a 5' EcoRI site and an adapted Spot 42* Hfq-binding scaffold²⁸⁵ with a 3' SacI restriction site. These two restriction sites were used for inserting sRNA library oligomers (see below). In order to improve the purity of gel-purified plasmid backbone DNA via an increased size difference in cut and non-cut p258_sRNA_scaffold, a 1.9 kb DNA dummy part was cut (EcoRI and SacI) and ligated from pCK01_ P_{alkS} _alkS²⁶⁰ (encoding P_{alkS} and a 5' fragment of *alkS*) into p258_sRNA_scaffold.

For insertion of sRNA target sequences against GFP, plasmid p258_sRNA_scaffold was PCR-amplified with phosphorylated primers sRNA_Tn7gfp_25_fwd and sRNA_Tn7gfp_25_rev, containing the sRNA target sequence (25 bp, $-46.1 \text{ kcal mol}^{-1}$, calculated by two-state hybridization³⁰⁰, RNA at 30°C with default settings), and ligated to yield plasmid p258_anti-gfp_sRNA as described previously²⁸⁶. The same strategy was used for constructing plasmid p258_control_sRNA_gfp (25 bp seed length, but reverse complement, $-17.4 \text{ kcal mol}^{-1}$) with primers sRNA_Tn7gfp_25_cont_fwd and sRNA_Tn7gfp_25_cont_rev.

Construction of genome-wide sRNA expression library

All annotated coding regions (incl. RNAs) of *P. putida* KT2440 (GenBank AE015451.2²⁵⁶, accessed via ncbi.nlm.nih.com) were used as targets and exported as csv spreadsheets from Geneious 11.1 cloning software (Biomatters, San Diego, CA, USA). The first 15 bp, 20 bp, and 25 bp of the corresponding DNA sequences were selected and changed to the reverse complements. For library DNA oligo construction, a constant 5' region was added (including a binding region for primer oligo_pool_uni_fwd and promoter P_m), as well as a constant 3' region (including a binding region for primer oligo_pool_uni_rev and the Spot 42* scaffold). Depending on the library sub pools constructed (see below), an additional unique “dial out” reverse primer binding region was incorporated in the 3' constant region (see Supplementary Table 5-3). The required DNA string manipulations were carried out with R software²⁴⁰ and the Bioconductor package³⁰⁸ (see supplementary Script 5-1). The five resultant library sub pools contained: a) all targets with sRNAs of all three binding length (17187 variants, reverse primer 144), b) all targets with sRNAs of 25 bp binding length (5729 variants, reverse primer 112), c) all targets annotated as regulators with sRNAs of 25 bp binding length (436 variants, primer 91), d) all targets annotated as RNAs with sRNAs of 25 bp binding length (165 variants, primer 88), and e) all targets annotated as ncRNAs with sRNAs of 25 bp binding length (67 variants, primer 67). The total of 23584 DNA oligomers

was synthesized using microchip technology and received as a single, pooled library (Twist Bioscience, San Francisco, CA, USA).

The DNA pool was diluted in ultra-pure water to a concentration of 20 ng μL^{-1} . The single-stranded oligomers were amplified as recommended by the DNA supplier. Briefly, for 12 cycles of PCR-based amplification with primers oligo_pool_uni_fwd and 144 of sub-pool a) consisting of 1.7×10^4 variants, Kapa HiFi HotStart polymerase was used (Kapa Biosystems, Cape Town, South Africa). Subsequently, the PCR product was column purified, digested with EcoRI and SacI, and column purified again. The library fragments were ligated into the cut and gel-purified pSEVA258_sRNA_scaffold plasmid. *E. coli* 10 β (*mCherry*) was transformed with the pooled sRNA library, yielding 3.2×10^4 variants. The cfu were scraped, the plasmids isolated and stored at -20°C as a library pool. The sRNA region of the plasmid pool was Sanger sequenced with primer g_block_sRNA_fwd.

sRNA expression to knock down GFP production

P. putida strain EP3_{GFP} was transformed with either pSEVA258 (no sRNA), p258_control_sRNA_gfp (control sRNA) or p258_anti-gfp_sRNA (anti-gfp sRNA). A negative control strain EP3 was transformed with pSEVA258 (no GFP). Precultures were grown in culture tubes for 16 h and diluted in fresh medium supplemented with 5 mM 3-methyl benzoate to an initial OD₆₀₀ of 0.05 (25 mL culture volume in Erlenmeyer flasks). After 6 h of induction, OD₆₀₀ and GFP fluorescence (excitation wavelength 488(9) nm, emission wavelength 530(20) nm) were measured with a plate reader (Tecan M1000 Pro, Maennedorf, Switzerland). Data was visualized and analyzed with Prism 8.4.2 (GraphPad, San Diego, CA, USA).

5.6 Acknowledgements

We thank P. Apura, S. Domingues, S. C. Viegas and C. Arraiano for support with the sRNA library; V. de Lorenzo and his lab members for providing pSEVA plasmids; S. Koebbing and L. Blank for providing pBG14g. We acknowledge financial support from The European Union Project EmPowerPutida (grant ID 635536) for MOB and are indebted to the whole consortium for most valuable feedback and discussions.

5.7 Supplementary information

Script 5-1 R-code for DNA oligo construction.

```
#load the libraries needed
library(Biostrings)
library(stringr)
library(tidyverse)
library(writexl)

sRNA <- read.csv("input.csv", header = T) # read spreadsheet containing the coding
sequences to be converted into sRNA target sequences (here all genome annotations
available for for P. putida via AE015451.2 exported from Geneious cloning software)

sRNA_tibble <- as_tibble(sRNA) #creates a tibble (tidyverse)
sRNA_tibble #or print(sRNA_tibble, n=Inf) #shows full tibble, change n to any
amount of rows to be shown

sRNA_tibble$Target_25 <- str_sub(sRNA_tibble$Sequence, 1,25) #define the sRNA
target length, e.g. 25 bp, 20 bp, and 15 bp, starting here at first coding base
sRNA_tibble$Target_20 <- str_sub(sRNA_tibble$Sequence, 1,20)
sRNA_tibble$Target_15 <- str_sub(sRNA_tibble$Sequence, 1,15)

DNA_Target_25 <- DNASTringSet(sRNA_tibble$Target_25) #converts column Target_25
into a DNASTringSet
sRNA_25 <- reverseComplement(DNA_Target_25) #produce the rev. comp of the target
region, i.e. the actual sRNA to be expressed
sRNA_25 <- as_vector(as.character(sRNA_25)) #get rid of DNASTringSet format, back
to normal vector

sRNA_tibble$sRNA25 <- paste(sRNA_25) #add tibble to table

DNA_Target_20 <- DNASTringSet(sRNA_tibble$Target_20) #converts column Target_20
into a DNASTringSet
sRNA_20 <- reverseComplement(DNA_Target_20) #produce the rev. comp of the target
region, i.e. the actual sRNA to be expressed
sRNA_20 <- as_vector(as.character(sRNA_20)) #get rid of DNASTringSet format, back
to normal vector

sRNA_tibble$sRNA20 <- paste(sRNA_20) #add tibble to table

DNA_Target_15 <- DNASTringSet(sRNA_tibble$Target_15) #converts column Target_20
into a DNASTringSet
sRNA_15 <- reverseComplement(DNA_Target_15) #produce the rev. comp of the target
region, i.e. the actual sRNA to be expressed
sRNA_15 <- as_vector(as.character(sRNA_15)) #get rid of DNASTringSet format, back
to normal vector

sRNA_tibble$sRNA15 <- paste(sRNA_15) #add tibble to table

five_prime_const <-
"CGATGAATTCAGCCTTGCAAGAAGCGGATACAGGAGTGCAAAAAATGGCTATCTCTAGTAAGGCCTACCCCTTAGGCTT
TATGCA" #adds constant region
three_prime_const <-
"ATTTGTAGAAATATTTTATTCGCCCCCGGAAGATCATTCCGGGGGCTTTTTTATTGAGCTCGTTCGTGACTCTCCTGTGC
```

```
ATTTCGTGGTA" #adds constant region, varied ones for sub-pools to be dialled out  
by PCR (i.e. includes dial-out specific primer sequence)
```

```
as_tibble(five_prime_const)  
as_tibble(three_prime_const)
```

```
sRNA_tibble$oligo_25 <- paste(five_prime_const, sRNA_tibble$sRNA25, sep = "|")  
#adds five prime const sequence, separator is for QC, replace/delete in final  
sheet if necessary
```

```
sRNA_tibble$oligo_25 <- paste(sRNA_tibble$oligo_25, three_prime_const, sep = "|")  
#same but three prime
```

```
sRNA_tibble$oligo_20 <- paste(five_prime_const, sRNA_tibble$sRNA20, sep = "|")  
#adds five prime const sequence
```

```
sRNA_tibble$oligo_20 <- paste(sRNA_tibble$oligo_20, three_prime_const, sep = "|")  
#same but three prime
```

```
sRNA_tibble$oligo_15 <- paste(five_prime_const, sRNA_tibble$sRNA15, sep = "|")  
#adds five prime const sequence
```

```
sRNA_tibble$oligo_15 <- paste(sRNA_tibble$oligo_15, three_prime_const, sep = "|")  
#same but three prime
```

```
write_xlsx(x = sRNA_tibble, path = "output.csv", col_names = TRUE)#data output  
sheet, including columns with the oligos to be ordered (copy-paste to web shop,  
e.g. here Twist Bioscience). To anyone who understands R well, apologies for this  
improvised hack.
```

Supplementary Table 5-1 Overview of strains and plasmids used in this study. AmpR, ampicillin resistance; GmR, gentamycin resistance; KanR, kanamycin resistance.

Strain or plasmid name	Relevant characteristics	Source/Reference
Strains		
<i>Escherichia coli</i> 10β(<i>mCherry</i>)	<i>E. coli</i> NEB 10β derivative (New England Biolabs, #C3019I), StrepR, allose operon replaced with constitutive <i>mCherry</i> expression cassette, $\Delta alsRBACEK::\lambda P_L-mCherry$	260
<i>E. coli</i> PIR2	F ⁻ $\Delta(argF-lac)169$ <i>rpoS</i> (Am) <i>robA1</i> <i>creC510</i> <i>hsdR514</i> <i>endArecA1</i> <i>uidA</i> ($\Delta MluI$):: <i>pir</i> ⁺	Invitrogen
<i>P. putida</i> EP3	KT2440 derivative, Δupp , $\Delta pedE$, $\Delta pedI$, $\Delta pedH$, $\Delta aldB-I$, $\Delta bkdAA$, $\Delta sthA$ $\Delta benABCD$	309
<i>P. putida</i> EP3 _{GFP}	<i>P</i> _{14g} <i>_BCD2_msfGFP</i> integrated downstream of <i>glmS</i> in strain EP3, GmR	This study.
Plasmids		
pBG14g	Tn7L/R, <i>P</i> _{14g} <i>_BCD2_msfGFP</i> , R6K, KanR, GmR	215
pTNS2	Tn7 transposase, R6K, AmpR	243
pSEVA258	XylS/P _m , RSF1010, KanR	195
p258_sRNA_scaffold	XylS/P _m _spot42*, RSF1010, KanR	This study.
p258_control_sRNA_gfp	XylS/P _m _controlsRNA_spot42*, RSF1010, KanR	This study.
p258_anti-gfp_sRNA	XylS/P _m _anti_gfp_sRNA_spot42*, RSF1010, KanR	This study.

Supplementary Table 5-2 DNA oligomers used in this study. Orthogonal primers 86, 88, 91, 112, and 144 from³⁰¹.

Primer	Sequence (5' to 3')
Pput_glmSUP	AGTCAGAGTTACGGAATTGTAGG
Pput_glmSDN	TTACGTGGCCGTGCTAAAGGG
oligo_pool_uni_fwd	CGATGAATTCAGCCTTGC
oligo_pool_uni_rev	CGACGAGCTCAATAAAAAAGC
144	TACCACGAAATGCACAGGAG
112	CGGGAGGAAGTCTTTAGACC
91	ATCCTAGAAAAGGCGAAGGC
88	ATCACAACAAAGGACGGGTC
86	ATAGATGGTGCCTACATGCG
sRNA_Tn7gfp_25_fwd	TCCCATGATCATATTTGTAGAAATATTTTATTCGCCCCCG AAG
sRNA_Tn7gfp_25_rev	ATTCATAAAGGTGTGCATAAAGCCTAAGGGGTAGGCCTTAC
sRNA_Tn7gfp_25_cont_fwd	ATTCATAAAGGTGATTTGTAGAAATATTTTATTCGCCCCCG GAAG
sRNA_Tn7gfp_25_cont_rev	TCCCATGATCATTGCATAAAGCCTAAGGGGTAGGCCTTAC
g_block_sRNA_fwd	CAACGACCGGTAGCGGAGCTATCCAAC

Supplementary Table 5-3 DNA parts synthesized in this study.

DNA part	Sequence (5' to 3')
sRNA_backbone_empty_for_lib (IDT gBlock)	CAACGACCGGTAGCGGAGCTATCCAACGGCGGTATACCAGGAA AACACACAGCAGGTACATCAGAACAGTACCATGACTGAAGAAC AAATAGTTTTTTTCCTGATCCATAAAGCAGAACGGCCTGCTCCA TGACAAATCTGGCTCCCCAACTAATGCCCCATGCAGCCAGCAT AACCAGCATAAAGTGCAGTGTCCGGTTTGATAGCGATGAATTC AGCCTTGCAAGAAGCGGATACAGGAGTGCAAAAAATGGCTATC TCTAGTAAGGCCTACCCCTTAGGCTTTATGCAATTTGTAGAAA TATTTTATTCGCCCCCGGAAGATCATTCCGGGGGCTTTTTTAT TGAGCTCGTCGTGACTGGGAAAACCCTGGCGACTAGTCTTGGA CTCCCTCCTGTGCATTTCTGTGGTAACCTCAGAACGCCTTCGCC TTTTCTAGGATACGCTCGGTTGGTCTAAAGACTTCCTCCCGGG TG
5' constant region, as part of oligo library synthesis (all Twist Bioscience)	CGATGAATTCAGCCTTGCAAGAAGCGGATACAGGAGTGCAAAA AATGGCTATCTCTAGTAAGGCCTACCCCTTAGGCTTTATGCA
3' region incl. primer 144 binding region	ATTTGTAGAAATATTTTATTCGCCCCCGGAAGATCATTCCGGG GGCTTTTTTATTGAGCTCGTCGTGACTCTCCTGTGCATTTCTG GGTA
3' region incl. primer 112 binding region	ATTTGTAGAAATATTTTATTCGCCCCCGGAAGATCATTCCGGG GGCTTTTTTATTGAGCTCGTCGTGACTGGTCTAAAGACTTCCT CCCG
3' region incl. primer 91 binding region	ATTTGTAGAAATATTTTATTCGCCCCCGGAAGATCATTCCGGG GGCTTTTTTATTGAGCTCGTCGTGACTGCCTTCGCCTTTTCTA GGAT
3' region incl. primer 88 binding region	ATTTGTAGAAATATTTTATTCGCCCCCGGAAGATCATTCCGGG GGCTTTTTTATTGAGCTCGTCGTGACTGACCCGTCCTTTGTTG TGAT
3' region incl. primer 86 binding region	ATTTGTAGAAATATTTTATTCGCCCCCGGAAGATCATTCCGGG GGCTTTTTTATTGAGCTCGTCGTGACTCGCATGTAGGCACCAT CTAT

Chapter 6: Conclusion and Outlook

In this thesis work, two industrially relevant microorganisms, *E. coli* and *P. putida* KT2440, were investigated for branched chain higher alcohol production as well as for *in vivo* alcohol detection enabled by biosensor circuits based on evolved AlkS transcription factor variants.

For the production of isopentanol, the commonly used Ehrlich degradation pathway was used¹⁷⁴. To this end, pyruvate, the final product of glycolysis, was channelled through the metabolic pathways for valine and leucine production, ultimately yielding 4-methyl-2-oxopentanoic acid, the α -keto acid precursor of leucine. The α -keto acid was subsequently decarboxylated and the resultant aldehyde reduced to the end product isopentanol. In total, nine enzymes encoding the metabolic pathway from pyruvate to isopentanol were overexpressed from an L-arabinose inducible system. For both microbes significant isopentanol production was achieved with this strategy. As expected, the iterative design of pathway variants allowed for the improvement of product titers as compared to initial designs. Also not surprisingly, achieving high product titers was more straightforward for *E. coli*, due to extensive prior knowledge. For isopentanol production with *E. coli*, first an aminotransferase (*ilvE*) was knocked-out in order to prevent α -keto acid flux towards valine and leucine. Next, depending on the overexpression of the leucine pathway module (encoding the *leuABCD* gene cluster), alcohol production was fine-tuned from almost exclusive isobutanol formation towards increasingly specific isopentanol production. By stable genomic integration of the isobutanol pathway downstream of *glmS*, consisting of the valine pathway as well as the decarboxylase and an alcohol dehydrogenase, about 0.53 g L⁻¹ isobutanol were made, with 20 mg L⁻¹ isopentanol as a side product. Upon additional, optimized expression of the leucine pathway from a combinatorial *leuA* library, varying expression strength as well as introducing a negative feedback release mutation, 0.73 g L⁻¹ isopentanol were produced, with 31 mg L⁻¹ isobutanol as a side product. This isopentanol titer corresponds to a product yield on glucose ($Y_{P/S}$) of about 0.14 g g⁻¹ or 29% of the theoretical maximum yield. The overall results are on par with previous metabolic engineering efforts for isopentanol production with *E. coli*¹⁸³. However, the absolute titers could likely be further improved by process optimization. For example, up to 9.5 g L⁻¹ isopentanol were produced previously with an extractive two-phase fermentation set-up³¹² and 56.5 g L⁻¹ isobutanol were produced with a fed-batch process using gas-stripping and cooling-traps for alcohol recovery¹²⁴.

For isopentanol production with *P. putida*, a strictly aerobic microorganism, multiple expression modules encoding the Ehrlich degradation pathway were tested in a multiple-knock out strain previously optimized for isobutanol production¹⁸⁴. Strains with genome-integrated alcohol pathways produced up to about 60 mg L⁻¹ isopentanol, while plasmid-based pathway overexpression allowed titers of up to 0.17 g L⁻¹ in batch cultures. Besides, about 0.6 g L⁻¹ of a side product, identified as isovaleric acid, were produced. Similar titers were found for an aerobic fed-batch process, with isopentanol titers close to 0.3 g L⁻¹ and about 0.65 g L⁻¹ of isovaleric acid as side product. The development of a fed-batch process with a microaerobic production phase, distinguished by switching the air supply to the head space of the bioreactor and thus limiting the dissolved oxygen, allowed to increase the isopentanol titer to 0.55 g L⁻¹ while the isovaleric acid production was reduced about four-fold. Besides, no further increase in biomass was found during the microaerobic production phase and glucose started to accumulated (which was not observed for the same glucose feeding-rate under aerobic conditions), pointing to the potential of optimizing the glucose-feeding during the microaerobic production phase for improved yields. To the best of my knowledge, no previous metabolic engineering efforts were made for targeted isopentanol production with *P. putida* and the highest titers published are about 125-fold lower (about 4.5 mg L⁻¹ as a side product)²⁵².

However, compared to other model microorganisms, these titers are still comparatively low, but as *P. putida* strains are endowed with abilities that for example *E. coli* lacks, e.g. the usage of lignin-derived carbon sources²⁵¹ and high organic solvent resistance²⁴⁹, it is encouraging to see that this microorganism is rapidly evolving as a more established host. The showcased production of isopentanol is certainly a valuable contribution towards the creation of additional production strains. Also the observed side product isovaleric acid is an attractive, industrially relevant target molecule and future research could develop targeted isovaleric acid biosynthesis for its sustainable production. Similar projects were previously conducted with *E. coli*²⁷³, allowing isovaleric acid titers of up to 32 g L⁻¹. Nevertheless, the results also highlight the current limitations in using DNA circuits as plug-and-play devices. While for *E. coli* a relatively pure isopentanol product was found (with titers close to the maximum achievable toxicity level), only for *P. putida* the product was found to be partly over-oxidized to the corresponding isovaleric acid. Overall, for *P. putida* the limitation of oxygen during the production phase improved the formation of the reduction product isopentanol from the aldehyde precursor (3-methylbutanal, reduced by a NADPH dependent alcohol dehydrogenase) over the oxygenation product isovaleric acid (potentially oxidized

by a NAD(P)⁺ dependent by aldehyde dehydrogenase). This observation suggests that the optimization of the NAD(P)H/NAD(P)⁺ ratio could further improve the achievable isopentanol titers with *P. putida*. A similar microaerobic fed-batch production process in a 30 L bioreactor was found to be beneficial for isobutanol production derived from the Ehrlich degradation pathway with *P. putida*²⁵⁸, with isobutanol titers of up to 3.4 g L⁻¹. Further, here the maximum achievable isopentanol titers were most likely limited by yet unknown regulatory mechanisms and not product toxicity, as indicated by experiments with a pathway intermediate or the end product externally added to the production cultures. Those different results were obtained despite using the same underlying genetic circuitry as in *E. coli*. Future research has to show if it is possible to domesticate²⁴⁸, or artificially built, a single minimal or reduced genome organism that is understood to a degree that such unintentional interferences can be avoided already during the design phase. In any case, minimal genome platform strains will most likely lack the benefit of, for example, improved solvent tolerance or mixed carbon source usage as known for *P. putida* strains. Consequently, it might be more feasible to select a particular strain from a set of reasonably well understood host cells, for example harnessing *P. putida* for the production of highly oxidized compounds from lignin-derived C-sources. Ultimately, the solution to this chassis issue will depend on how cheaply artificial bacterial genomes can be synthesized, potentially enabling projects in the (distant) future to start from a minimal genome sequence adding on top tailor-made DNA parts in order to embed only metabolic pathways beneficial for the product of interest.

The development of AlkS-based biosensor systems demonstrated how powerful and readily useable the method of directed evolution is for transcription factor engineering. AlkS variants were optimized through three rounds of evolution and the defined AlkS mutants obtained were characterized on the single-cell level for a panel of high production volume C₄- and C₅-alcohols. For all alcohols tested, maximum fold-changes in biosensor output of at least 19-fold were found, while the dynamic range included the low to mid mM range making the developed sensor systems highly suitable for the detection of further improved alcohol production strains (up to the toxicity limit of *E. coli*). In particular, an AlkS variant with a single L401G mutation allowed the recognition of isopentanol or n-butanol, without being responsive to isobutanol. Such alcohol isomer-specific recognition profiles were previously not found, despite engineering efforts made with the regulator BmoR¹⁹². These results also confirmed that protein engineering leads to useful experimental outcomes without insights in the underlying protein structure or functional mechanism. Nevertheless, there is a clear gap in the learning part of the corresponding DBTL cycle. In particular, a current limitation

is how to choose a wild-type regulator for evolving it to a sensor system for detection of a given target compound. After all this choice is an educated guess and it is hard to predict how much re-shaping a regulatory protein requires to change its inducer spectrum or how (un)specific the mutant variants will be. The development of screening and selection methods that allow the collection of large data sets including information on rare as well as detrimental and neutral mutations by long-read NGS in combination with machine learning could reveal improved protein engineering strategies and insides into rational regulator design in the near future. The interest in biosensor systems will most likely grow further, as they fit well into the scope of automation-based screening platforms as championed by central biofoundry facilities working on full integration of DBTL cycles with robotic systems.

In this work, the biosensor system was applied for two screens concerning key steps of alcohol production by Ehrlich degradation. First, a KivD site-saturation mutagenesis library for isopentanol production via whole-cell catalysis from supplied α -keto acid was screened, and second, a combinatorial library for *leuA* expression optimizing isopentanol production from glucose. For the latter library screen, more than 4000 strain variants were successfully screened in 384-well plates with an automated protocol using liquid handling robots and utilizing a biosensor system with the isopentanol specific AlkS variant, thus not being activated by isobutanol side product formation. Previously, biosensor-based screens for n-butanol¹²³ or isobutanol¹²⁴ relied and on manual pipetting and 96-well plates, increasing material usage while limiting scalability. Further, these two set-ups either required a more laborious two-step screening protocol, adding the sensor strain only after n-butanol was produced from the corresponding supplemented α -keto acid, or showed limited fold-changes in biosensor output of less than 1.5-fold for isobutanol production from glucose. In general, the AlkS-based biosensor system enabled significantly faster and more economic screening runs than would have been possible with traditional chemical analyses, while also creating no additional plastic and solvent waste. Still, bioprocess upscale requires these more exact traditional methods, but biosensor systems will be a handy tool for providing screening hits at the beginning of the process development pipeline.

A yet unexplored approach for optimizing *P. putida* production strains is individual, but genome wide, sRNA-based knock down of RNA encoding sequences. This strategy could lead to the discovery of unexpected engineering targets, including essential genes which cannot easily be deleted from the host genome. Here, it was demonstrated that genome wide annotations available for *P. putida* KT2440 can be readily used for designing synthetic sRNAs against all annotated coding frames, while their relatively cost-efficient synthesis was

also feasible due to increasingly powerful but cheap on-chip oligo synthesis. For example, the resultant plasmid-based sRNA library could be used for finding knock down targets for improved alcohol production in *P. putida* in co-culture screens employing the orthogonal *E. coli* biosensor strain established in this work. This co-culture scheme and the orthogonal C-sources used could further be fine-tuned in order to iteratively adapt product titers to fit to the dynamic range of the biosensor strain. The current limitation is to carry out all necessary (genetic) engineering steps in concert. First, the alcohol product titer has to be high enough to allow biosensor activation and the corresponding strain has to maintain an additional sRNA expression plasmid. A reasonable way for implementation would start with an alcohol pathway stably integrated in the host genome, thus not requiring a selection marker or origin of replication for pathway maintenance. Second, it needs to be explored when sRNA expression should be induced. Most likely, the optimal time point varies for different target genes, for example depending on the amount of the mRNA transcript present, as well as the half-life and essentiality of the corresponding enzyme.

Chapter 7: References

1. Walker, G.M. & Stewart, G.G. *Saccharomyces cerevisiae* in the Production of Fermented Beverages. *Beverages* **2**, 1-12 (2016).
2. Buchner, E. Alkoholische Gahrung ohne Hefezellen. *Berichte der deutschen chemischen Gesellschaft* **30**, 117-124 (1897).
3. Platt, D. Memoire sur la fermentation appelee lactique (Louis Pasteur). *Del Med J* **60**, 23-24 (1988).
4. Awang, G.M., Jones, G.A. & Ingledew, W.M. The acetone-butanol-ethanol fermentation. *Crit. Rev. Microbiol.* **15**, S33-67 (1988).
5. Becker, J. & Wittmann, C. Advanced biotechnology: metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products. *Angew. Chem. Int. Ed. Engl.* **54**, 3328-3350 (2015).
6. Hoff, B. et al. Unlocking Nature's Biosynthetic Power-Metabolic Engineering for the Fermentative Production of Chemicals. *Angew. Chem. Int. Ed. Engl.* **60**, 2258-2278 (2021).
7. Takors, R. Biochemical engineering provides mindset, tools and solutions for the driving questions of a sustainable future. *Eng. Life Sci.* **20**, 5-6 (2020).
8. Straathof, A.J.J. et al. Grand Research Challenges for Sustainable Industrial Biotechnology. *Trends Biotechnol.* **37**, 1042-1050 (2019).
9. Kuhn, D., Blank, L.M., Schmid, A. & Buhler, B. Systems biotechnology–Rational whole-cell biocatalyst and bioprocess design. *Eng. Life Sci.* **10**, 384-397 (2010).
10. Lee, S.Y. & Kim, H.U. Systems strategies for developing industrial microbial strains. *Nat. Biotechnol.* **33**, 1061-1072 (2015).
11. Kosuri, S. & Church, G.M. Large-scale de novo DNA synthesis: technologies and applications. *Nat. Methods* **11**, 499-507 (2014).
12. Heinemann, M. & Panke, S. Synthetic biology - Putting engineering into biology. *Bioinformatics* **22**, 2790-2799 (2006).
13. Nielsen, J. & Keasling, J.D. Synergies between synthetic biology and metabolic engineering. *Nat. Biotechnol.* **29**, 693-695 (2011).
14. Bailey, J.E. Toward a science of metabolic engineering. *Science* **252**, 1668-1675 (1991).
15. Liu, Y. & Nielsen, J. Recent trends in metabolic engineering of microbial chemical factories. *Curr. Opin. Biotechnol.* **60**, 188-197 (2019).
16. Rogers, J.K. & Church, G.M. Multiplexed Engineering in Biology. *Trends Biotechnol.* **34**, 198-206 (2016).

17. Arnold, F.H. Innovation by Evolution: Bringing New Chemistry to Life (Nobel Lecture). *Angew. Chem. Int. Ed. Engl.* **58**, 14420-14426 (2019).
18. de Lorenzo, V., Krasnogor, N. & Schmidt, M. For the sake of the Bioeconomy: define what a Synthetic Biology Chassis is! *N Biotechnol* **60**, 44-51 (2021).
19. Kwok, R. Five hard truths for synthetic biology. *Nature* **463**, 288-290 (2010).
20. Zhang, J. et al. Accelerating strain engineering in biofuel research via build and test automation of synthetic biology. *Curr. Opin. Biotechnol.* **67**, 88-98 (2021).
21. Dietrich, J.A., McKee, A.E. & Keasling, J.D. High-throughput metabolic engineering: advances in small-molecule screening and selection. *Annu. Rev. Biochem.* **79**, 563-590 (2010).
22. Kim, B., Kim, W.J., Kim, D.I. & Lee, S.Y. Applications of genome-scale metabolic network model in metabolic engineering. *J. Ind. Microbiol. Biotechnol.* **42**, 339-348 (2015).
23. Sandberg, T.E., Salazar, M.J., Weng, L.L., Palsson, B.O. & Feist, A.M. The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. *Metab. Eng.* **56**, 1-16 (2019).
24. Agresti, J.J. et al. Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proc Natl Acad Sci U S A* **107**, 4004-4009 (2010).
25. Charmantray, F., Legeret, B., Helaine, V. & Hecquet, L. Fluorogenic substrates for the screening assay of transketolase through beta-elimination of umbelliferone - Development, scope and limitations. *J. Biotechnol.* **145**, 359-366 (2010).
26. Morgan, S.A., Nadler, D.C., Yokoo, R. & Savage, D.F. Biofuel metabolic engineering with biosensors. *Curr. Opin. Chem. Biol.* **35**, 150-158 (2016).
27. Mehrer, C.R. et al. Growth-coupled bioconversion of levulinic acid to butanone. *Metab. Eng.* **55**, 92-101 (2019).
28. Fernandez-Cabezon, L., Cros, A. & Nikel, P.I. Evolutionary Approaches for Engineering Industrially Relevant Phenotypes in Bacterial Cell Factories. *Biotechnol J* **14**, e1800439 (2019).
29. Pontrelli, S. et al. Metabolic repair through emergence of new pathways in Escherichia coli. *Nat. Chem. Biol.* **14**, 1005-1009 (2018).
30. Rogers, J.K., Taylor, N.D. & Church, G.M. Biosensor-based engineering of biosynthetic pathways. *Curr. Opin. Biotechnol.* **42**, 84-91 (2016).
31. Mahr, R. & Frunzke, J. Transcription factor-based biosensors in biotechnology: current state and future prospects. *Appl. Microbiol. Biotechnol.* **100**, 79-90 (2016).
32. Nielsen, J. & Keasling, J.D. Engineering Cellular Metabolism. *Cell* **164**, 1185-1197 (2016).
33. Lawson, C.E. et al. Machine learning for metabolic engineering: A review. *Metab. Eng.* **63**, 34-60 (2021).

34. Presnell, K.V. & Alper, H.S. Systems Metabolic Engineering Meets Machine Learning: A New Era for Data-Driven Metabolic Engineering. *Biotechnol J* **14**, e1800416 (2019).
35. Siegel, J. & Krishnan, S. Cultivating Invisible Impact with Deep Technology and Creative Destruction. *Journal of Innovation Management* **8**, 6-19 (2020).
36. Camacho, D.M., Collins, K.M., Powers, R.K., Costello, J.C. & Collins, J.J. Next-Generation Machine Learning for Biological Networks. *Cell* **173**, 1581-1592 (2018).
37. Hillson, N. et al. Building a global alliance of biofoundries. *Nat Commun* **10**, 2040 (2019).
38. Carbonell, P., Le Feuvre, R., Takano, E. & Scrutton, N.S. In silico design and automated learning to boost next-generation smart biomanufacturing. *Synth Biol (Oxf)* **5**, ysaa020 (2020).
39. Tas, H. et al. Are synthetic biology standards applicable in everyday research practice? *Microb Biotechnol* **13**, 1304-1308 (2020).
40. Abraham, M.A. & Nguyen, N. Green engineering: Defining the principles. *Environ. Prog.* **22**, 233-236 (2003).
41. Ribeiro, B.D., Coelho, M.A.Z. & Machado de Castro, A. Principles of Green Chemistry and White Biotechnology (in White Biotechnology for Sustainable Chemistry). *Royal Society of Chemistry*, 1-8 (2015).
42. Anastas, P.T. & Warner, J.C. Green chemistry: theory and practice. (Oxford University Press, Oxford; 1998).
43. Dai, Z. & Nielsen, J. Advancing metabolic engineering through systems biology of industrial microorganisms. *Curr. Opin. Biotechnol.* **36**, 8-15 (2015).
44. Jeschek, M., Gerngross, D. & Panke, S. Combinatorial pathway optimization for streamlined metabolic engineering. *Curr. Opin. Biotechnol.* **47**, 142-151 (2017).
45. Tee, K.L. & Wong, T.S. Polishing the craft of genetic diversity creation in directed evolution. *Biotechnol. Adv.* **31**, 1707-1721 (2013).
46. Sundberg, S.A. High-throughput and ultra-high-throughput screening: solution- and cell-based approaches. *Curr. Opin. Biotechnol.* **11**, 47-53 (2000).
47. Xiao, H., Bao, Z. & Zhao, H. High Throughput Screening and Selection Methods for Directed Enzyme Evolution. *Ind. Eng. Chem. Res.* **54**, 4011-4020 (2015).
48. Wong, T.S., Zhurina, D. & Schwaneberg, U. The diversity challenge in directed protein evolution. *Comb Chem High Throughput Screen* **9**, 271-288 (2006).
49. Ruff, A.J., Dennig, A. & Schwaneberg, U. To get what we aim for - Progress in diversity generation methods. *FEBS J.* **280**, 2961-2978 (2013).
50. Dower, W.J., Miller, J.F. & Ragsdale, C.W. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* **16**, 6127-6145 (1988).

51. Liu, D., Evans, T. & Zhang, F. Applications and advances of metabolite biosensors for metabolic engineering. *Metab. Eng.* **31**, 35-43 (2015).
52. Zhang, J., Jensen, M.K. & Keasling, J.D. Development of biosensors and their application in metabolic engineering. *Curr. Opin. Chem. Biol.* **28**, 1-8 (2015).
53. Williams, T.C., Pretorius, I.S. & Paulsen, I.T. Synthetic Evolution of Metabolic Productivity Using Biosensors. *Trends Biotechnol.* **34**, 371-381 (2016).
54. Monosik, R., Stred'ansky, M. & Sturdik, E. Biosensors - classification, characterization and new trends. *Acta Chimica Slovaca* **5**, 109-120 (2012).
55. He, W., Yuan, S., Zhong, W.H., Siddikee, M.A. & Dai, C.C. Application of genetically engineered microbial whole-cell biosensors for combined chemosensing. *Appl. Microbiol. Biotechnol.* **100**, 1109-1119 (2016).
56. Meyer, A. et al. Optimization of a whole-cell biocatalyst by employing genetically encoded product sensors inside nanolitre reactors. *Nat Chem* **7**, 673-678 (2015).
57. You, M. & Jaffrey, S.R. Structure and Mechanism of RNA Mimics of Green Fluorescent Protein. *Annu Rev Biophys* **44**, 187-206 (2015).
58. Morris, K.V. & Mattick, J.S. The rise of regulatory RNA. *Nat. Rev. Genet.* **15**, 423-437 (2014).
59. Hallberg, Z.F., Su, Y., Kitto, R.Z. & Hammond, M.C. Engineering and In Vivo Applications of Riboswitches. *Annu. Rev. Biochem.* **86**, 515-539 (2017).
60. Berens, C., Groher, F. & Suess, B. RNA aptamers as genetic control devices: the potential of riboswitches as synthetic elements for regulating gene expression. *Biotechnol J* **10**, 246-257 (2015).
61. Krishnamurthy, M. et al. Tunable Riboregulator Switches for Post-transcriptional Control of Gene Expression. *ACS Synth Biol* **4**, 1326-1334 (2015).
62. Winkler, W., Nahvi, A. & Breaker, R.R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952-956 (2002).
63. Hollands, K. et al. Riboswitch control of Rho-dependent transcription termination. *Proc Natl Acad Sci U S A* **109**, 5376-5381 (2012).
64. Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A. & Breaker, R.R. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**, 281-286 (2004).
65. Davidson, M.E., Harbaugh, S.V., Chushak, Y.G., Stone, M.O. & Kelley-Loughnane, N. Development of a 2,4-dinitrotoluene-responsive synthetic riboswitch in *E. coli* cells. *ACS Chem Biol* **8**, 234-241 (2013).
66. Wieland, M. & Hartig, J.S. Artificial riboswitches: synthetic mRNA-based regulators of gene expression. *ChemBioChem* **9**, 1873-1878 (2008).
67. Berens, C. & Suess, B. Riboswitch engineering - making the all-important second and third steps. *Curr. Opin. Biotechnol.* **31**, 10-15 (2015).

68. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505-510 (1990).
69. Robertson, D.L. & Joyce, G.F. Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature* **344**, 467-468 (1990).
70. Ellington, A.D. & Szostak, J.W. Selection in vitro of single-stranded DNA molecules that fold into specific ligand-binding structures. *Nature* **355**, 850-852 (1992).
71. Penchovsky, R. Computational design and biosensor applications of small molecule-sensing allosteric ribozymes. *Biomacromolecules* **14**, 1240-1249 (2013).
72. Clote, P. Computational Prediction of Riboswitches (Methods in Enzymology, Computational Methods for Understanding Riboswitches, Academic Press). (2015).
73. Weigand, J.E. et al. Screening for engineered neomycin riboswitches that control translation initiation. *RNA* **14**, 89-97 (2008).
74. Michener, J.K. & Smolke, C.D. High-throughput enzyme evolution in *Saccharomyces cerevisiae* using a synthetic RNA switch. *Metab. Eng.* **14**, 306-316 (2012).
75. Jang, S. et al. On-chip analysis, indexing and screening for chemical producing bacteria in a microfluidic static droplet array. *Lab Chip* **16**, 1909-1916 (2016).
76. Wang, J., Gao, D., Yu, X., Li, W. & Qi, Q. Evolution of a chimeric aspartate kinase for L-lysine production using a synthetic RNA device. *Appl. Microbiol. Biotechnol.* **99**, 8527-8536 (2015).
77. Zhou, L.B. & Zeng, A.P. Engineering a Lysine-ON Riboswitch for Metabolic Control of Lysine Production in *Corynebacterium glutamicum*. *ACS Synth Biol* **4**, 1335-1340 (2015).
78. Zhou, L.B. & Zeng, A.P. Exploring lysine riboswitch for metabolic flux control and improvement of L-lysine synthesis in *Corynebacterium glutamicum*. *ACS Synth Biol* **4**, 729-734 (2015).
79. Lünse, C. & Mayer, G. Screening Assays to Identify Artificial glmS Ribozyme Activators (in Therapeutic Applications of Ribozymes and Riboswitches). 199-209 (2014).
80. Paige, J.S., Wu, K.Y. & Jaffrey, S.R. RNA mimics of green fluorescent protein. *Science* **333**, 642-646 (2011).
81. Juskowiak, B. Nucleic acid-based fluorescent probes and their analytical potential. *Anal. Bioanal. Chem.* **399**, 3157-3176 (2011).
82. Dolgosheina, E.V. & Unrau, P.J. Fluorophore-binding RNA aptamers and their applications. *Wiley Interdiscip Rev RNA* **7**, 843-851 (2016).
83. Strack, R.L. & Jaffrey, S.R. New approaches for sensing metabolites and proteins in live cells using RNA. *Curr. Opin. Chem. Biol.* **17**, 651-655 (2013).
84. You, M., Litke, J.L. & Jaffrey, S.R. Imaging metabolite dynamics in living cells using a Spinach-based riboswitch. *Proc Natl Acad Sci U S A* **112**, 2756-2765 (2015).

85. Strack, R.L., Song, W. & Jaffrey, S.R. Using Spinach-based sensors for fluorescence imaging of intracellular metabolites and proteins in living bacteria. *Nat Protoc* **9**, 146-155 (2014).
86. Song, W., Strack, R.L. & Jaffrey, S.R. Imaging bacterial protein expression using genetically encoded RNA sensors. *Nat. Methods* **10**, 873-875 (2013).
87. Su, Y., Hickey, S.F., Keyser, S.G. & Hammond, M.C. In Vitro and In Vivo Enzyme Activity Screening via RNA-Based Fluorescent Biosensors for S-Adenosyl-l-homocysteine (SAH). *J. Am. Chem. Soc.* **138**, 7040-7047 (2016).
88. Hochreiter, B., Garcia, A.P. & Schmid, J.A. Fluorescent proteins as genetically encoded FRET biosensors in life sciences. *Sensors* **15**, 26281-26314 (2015).
89. Baird, G.S., Zacharias, D.A. & Tsien, R.Y. Circular permutation and receptor insertion within green fluorescent proteins. *Proc. Natl. Acad. Sci. USA* **96**, 11241-11246 (1999).
90. Schallmeyer, M., Frunzke, J., Eggeling, L. & Marienhagen, J. Looking for the pick of the bunch: high-throughput screening of producing microorganisms with biosensors. *Curr. Opin. Biotechnol.* **26**, 148-154 (2014).
91. Lim, J.B. & Sikes, H.D. Use of a genetically encoded hydrogen peroxide sensor for whole cell screening of enzyme activity. *Protein Eng. Des. Sel.* **28**, 79-83 (2015).
92. Polisky, B., Bishop, R.J. & Gelfand, D.H. A plasmid cloning vehicle allowing regulated expression of eukaryotic DNA in bacteria. *Proc Natl Acad Sci U S A* **73**, 3900-3904 (1976).
93. Hatfield, D., Hofnung, M. & Schwartz, M. Genetic Analysis of Maltose a Region in Escherichia coli. *J. Bacteriol.* **98**, 559 (1969).
94. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318-356 (1961).
95. Shoeman, R. et al. Regulation of methionine synthesis in Escherichia coli: Effect of metJ gene product and S-adenosylmethionine on the expression of the metF gene. *Proc Natl Acad Sci U S A* **82**, 3601-3605 (1985).
96. de Lorenzo, V., Silva-Rocha, R., Carbajosa, G., Galvao, T.C. & Cases, I. in Sensory Mechanisms in Bacteria: Molecular Aspects of Signal Recognition: Sensing Xenobiotic Compounds: Lessons from Bacteria that Face Pollutants in the Environment 81-105 (2010).
97. Libis, V., Delepine, B. & Faulon, J.L. Sensing new chemicals with bacterial transcription factors. *Curr. Opin. Microbiol.* **33**, 105-112 (2016).
98. Galvao, T.C. & de Lorenzo, V. Transcriptional regulators a la carte: engineering new effector specificities in bacterial regulatory proteins. *Curr. Opin. Biotechnol.* **17**, 34-42 (2006).

99. Jha, R.K., Chakraborti, S., Kern, T.L., Fox, D.T. & Strauss, C.E. Rosetta comparative modeling for library design: Engineering alternative inducer specificity in a transcription factor. *Proteins* **83**, 1327-1340 (2015).
100. Tang, S.Y. & Cirino, P.C. Design and application of a mevalonate-responsive regulatory protein. *Angew. Chem. Int. Ed. Engl.* **50**, 1084-1086 (2011).
101. Cebolla, A., Sousa, C. & de Lorenzo, V. Effector specificity mutants of the transcriptional activator NahR of naphthalene degrading *Pseudomonas* define protein sites involved in binding of aromatic inducers. *J. Biol. Chem.* **272**, 3986-3992 (1997).
102. Santos, C.N. & Stephanopoulos, G. Melanin-based high-throughput screen for L-tyrosine production in *Escherichia coli*. *Appl. Environ. Microbiol.* **74**, 1190-1197 (2008).
103. Kohlmeier, S. et al. Bioreporters: gfp versus lux revisited and single-cell response. *Biosens. Bioelectron.* **22**, 1578-1585 (2007).
104. Rogers, J.K. & Church, G.M. Genetically encoded sensors enable real-time observation of metabolite production. *Proc Natl Acad Sci U S A* **113**, 2388-2393 (2016).
105. Daunert, S. et al. Genetically engineered whole-cell sensing systems: coupling biological recognition with reporter genes. *Chem. Rev.* **100**, 2705-2738 (2000).
106. van der Meer, J.R. & Belkin, S. Where microbiology meets microengineering: design and applications of reporter bacteria. *Nat. Rev. Microbiol.* **8**, 511-522 (2010).
107. Gredell, J.A., Frei, C.S. & Cirino, P.C. Protein and RNA engineering to customize microbial molecular reporting. *Biotechnol J* **7**, 477-499 (2012).
108. Eggeling, L., Bott, M. & Marienhagen, J. Novel screening methods -Biosensors. *Curr. Opin. Biotechnol.* **35**, 30-36 (2015).
109. Li, J.W., Zhang, X.Y., Wu, H. & Bai, Y.P. Transcription Factor Engineering for High-Throughput Strain Evolution and Organic Acid Bioproduction: A Review. *Front Bioeng Biotechnol* **8**, 98 (2020).
110. de Lorenzo, V., Fernández, S., Herrero, M., Jakubzik, U. & Timmis, K.N. Engineering of alkyl- and haloaromatic-responsive gene expression with mini-transposons containing regulated promoters of biodegradative pathways of *Pseudomonas*. *Gene* **130**, 41-46 (1993).
111. Werlen, C., Jaspers, M.C. & van der Meer, J.R. Measurement of biologically available naphthalene in gas and aqueous phases by use of a *Pseudomonas putida* biosensor. *Appl. Environ. Microbiol.* **70**, 43-51 (2004).
112. Sevilla, E., Yuste, L. & Rojo, F. Marine hydrocarbonoclastic bacteria as whole-cell biosensors for n-alkanes. *Microb Biotechnol* **8**, 693-706 (2015).

113. Jaspers, M.C., Meier, C., Zehnder, A.J., Harms, H. & van der Meer, J.R. Measuring mass transfer processes of octane with the help of an alkSalkB::gfp-tagged *Escherichia coli*. *Environ. Microbiol.* **3**, 512-524 (2001).
114. Sticher, P. et al. Development and characterization of a whole-cell bioluminescent sensor for bioavailable middle-chain alkanes in contaminated groundwater samples. *Appl. Environ. Microbiol.* **63**, 4053-4060 (1997).
115. Espinosa-Urgel, M., Serrano, L., Ramos, J.L. & Fernandez-Escamilla, A.M. Engineering Biological Approaches for Detection of Toxic Compounds: A New Microbial Biosensor Based on the *Pseudomonas putida* TtgR Repressor. *Mol. Biotechnol.* **57**, 558-564 (2015).
116. Lacal, J., Busch, A., Guazzaroni, M.E., Krell, T. & Ramos, J.L. The TodS-TodT two-component regulatory system recognizes a wide range of effectors and works with DNA-bending proteins. *Proc Natl Acad Sci U S A* **103**, 8191-8196 (2006).
117. Call, T.P., Akhtar, M.K., Baganz, F. & Grant, C. Modulating the import of medium-chain alkanes in *E. coli* through tuned expression of FadL. *J Biol Eng* **10**, 5 (2016).
118. Wu, W. et al. Genetically assembled fluorescent biosensor for in situ detection of biosynthesized alkanes. *Sci Rep* **5**, 10907 (2015).
119. Berepiki, A., Kent, R., Machado, L.F.M. & Dixon, N. Development of High-Performance Whole Cell Biosensors Aided by Statistical Modeling. *ACS Synth Biol* **9**, 576-589 (2020).
120. Binder, S. et al. A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. *Genome Biol* **13**, R40 (2012).
121. Binder, S., Siedler, S., Marienhagen, J., Bott, M. & Eggeling, L. Recombineering in *Corynebacterium glutamicum* combined with optical nanosensors: a general strategy for fast producer strain generation. *Nucleic Acids Res.* **41**, 6360-6369 (2013).
122. Schendzielorz, G. et al. Taking control over control: use of product sensing in single cells to remove flux control at key enzymes in biosynthesis pathways. *ACS Synth Biol* **3**, 21-29 (2014).
123. Dietrich, J.A., Shis, D.L., Alikhani, A. & Keasling, J.D. Transcription factor-based screens and synthetic selections for microbial small-molecule biosynthesis. *ACS Synth Biol* **2**, 47-58 (2013).
124. Yu, H. et al. Establishment of BmoR-based biosensor to screen isobutanol overproducer. *Microb Cell Fact* **18**, 30 (2019).
125. Cheng, F. et al. A Competitive Flow Cytometry Screening System for Directed Evolution of Therapeutic Enzyme. *ACS synthetic biology* **4**, 768-775 (2015).
126. Siedler, S. et al. SoxR as a single-cell biosensor for NADPH-consuming enzymes in *Escherichia coli*. *ACS Synth Biol* **3**, 41-47 (2014).
127. Novichkov, P.S. et al. RegPrecise 3.0 - A resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC Genomics* **14**, 745 (2013).

128. Carbajosa, G., Trigo, A., Valencia, A. & Cases, I. Bionemo: molecular information on biodegradation metabolism. *Nucleic Acids Res.* **37**, D598-602 (2009).
129. Gao, J., Ellis, L.B. & Wackett, L.P. The University of Minnesota Biocatalysis/Biodegradation Database: improving public access. *Nucleic Acids Res.* **38**, D488-491 (2010).
130. Zaslaver, A. et al. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nat. Methods* **3**, 623-628 (2006).
131. Mahr, R., von Boeselager, R.F., Wiechert, J. & Frunzke, J. Screening of an *Escherichia coli* promoter library for a phenylalanine biosensor. *Appl. Microbiol. Biotechnol.* **100**, 6739-6753 (2016).
132. Uchiyama, T., Abe, T., Ikemura, T. & Watanabe, K. Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes. *Nat. Biotechnol.* **23**, 88-93 (2005).
133. Uchiyama, T. & Watanabe, K. Substrate-induced gene expression (SIGEX) screening of metagenome libraries. *Nat Protoc* **3**, 1202-1212 (2008).
134. Delepine, B., Libis, V., Carbonell, P. & Faulon, J.L. SensiPath: computer-aided design of sensing-enabling metabolic pathways. *Nucleic Acids Res.* **44**, W226-231 (2016).
135. Libis, V., Delepine, B. & Faulon, J.L. Expanding Biosensing Abilities through Computer-Aided Design of Metabolic Pathways. *ACS Synth Biol* **5**, 1076-1085 (2016).
136. Ramos, J.L., Stolz, A., Reineke, W. & Timmis, K.N. Altered effector specificities in regulators of gene expression: TOL plasmid xylS mutants and their use to engineer expansion of the range of aromatics degraded by bacteria. *Proc Natl Acad Sci U S A* **83**, 8467-8471 (1986).
137. Ramos, J.L., Michan, C., Rojo, F., Dwyer, D. & Timmis, K. Signal-regulator interactions. Genetic analysis of the effector binding site of xylS, the benzoate-activated positive regulator of Pseudomonas TOL plasmid meta-cleavage pathway operon. *J. Mol. Biol.* **211**, 373-382 (1990).
138. Monteiro, L.M.O. et al. Reverse Engineering of an Aspirin-Responsive Transcriptional Regulator in *Escherichia coli*. *ACS Synth Biol* **8**, 1890-1900 (2019).
139. Jensen, R.A. Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* **30**, 409-425 (1976).
140. Galvao, T.C., Mencia, M. & de Lorenzo, V. Emergence of novel functions in transcriptional regulators by regression to stem protein types. *Mol. Microbiol.* **65**, 907-919 (2007).
141. Garmendia, J., Devos, D., Valencia, A. & de Lorenzo, V. A la carte transcriptional regulators: unlocking responses of the prokaryotic enhancer-binding protein XylR to non-natural effectors. *Mol. Microbiol.* **42**, 47-59 (2001).

142. Hawkins, A.C., Arnold, F.H., Stuermer, R., Hauer, B. & Leadbetter, J.R. Directed evolution of *Vibrio fischeri* LuxR for improved response to butanoyl-homoserine lactone. *Appl. Environ. Microbiol.* **73**, 5775-5781 (2007).
143. Tang, S.Y., Fazelinia, H. & Cirino, P.C. AraC regulatory protein mutants with altered effector specificity. *J. Am. Chem. Soc.* **130**, 5267-5271 (2008).
144. Tang, S.Y. et al. Screening for enhanced triacetic acid lactone production by recombinant *Escherichia coli* expressing a designed triacetic acid lactone reporter. *J. Am. Chem. Soc.* **135**, 10099-10103 (2013).
145. Looger, L.L., Dwyer, M.A., Smith, J.J. & Hellinga, H.W. Computational design of receptor and sensor proteins with novel functions. *Nature* **423**, 185-190 (2003).
146. Reimer, A., Yagur-Kroll, S., Belkin, S., Roy, S. & van der Meer, J.R. *Escherichia coli* ribose binding protein based bioreporters revisited. *Sci Rep* **4**, 5626 (2014).
147. Tinberg, C.E. et al. Computational design of ligand-binding proteins with high affinity and selectivity. *Nature* **501**, 212-216 (2013).
148. Das, R. & Baker, D. Macromolecular modeling with rosetta. *Annu. Rev. Biochem.* **77**, 363-382 (2008).
149. Taylor, N.D. et al. Engineering an allosteric transcription factor to respond to new ligands. *Nat. Methods* **13**, 177-183 (2016).
150. Zhang, J.H., Chung, T.D. & Oldenburg, K.R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* **4**, 67-73 (1999).
151. Tabor, J.J., Groban, E.S. & Voigt, C.A. Performance Characteristics for Sensors and Circuits Used to Program *E. coli* (in *Systems Biology and Biotechnology of Escherichia coli*). 401-439 (2009).
152. Ang, J., Harris, E., Hussey, B.J., Kil, R. & McMillen, D.R. Tuning response curves for synthetic biology. *ACS Synth Biol* **2**, 547-567 (2013).
153. Mannan, A.A., Liu, D., Zhang, F. & Oyarzun, D.A. Fundamental Design Principles for Transcription-Factor-Based Metabolite Biosensors. *ACS Synth Biol* **6**, 1851-1859 (2017).
154. Merulla, D. & van der Meer, J.R. Regulatable and Modulable Background Expression Control in Prokaryotic Synthetic Circuits by Auxiliary Repressor Binding Sites. *ACS Synth Biol* **5**, 36-45 (2016).
155. de Las Heras, A., Fraile, S. & de Lorenzo, V. Increasing signal specificity of the TOL network of *Pseudomonas putida* mt-2 by rewiring the connectivity of the master regulator XylR. *PLoS Genet.* **8**, e1002963 (2012).
156. Goni-Moreno, A., Benedetti, I., Kim, J. & de Lorenzo, V. Deconvolution of Gene Expression Noise into Spatial Dynamics of Transcription Factor-Promoter Interplay. *ACS Synth Biol* **6**, 1359-1369 (2017).

157. Choi, S.L. et al. Toward a generalized and high-throughput enzyme screening system based on artificial genetic circuits. *ACS Synth Biol* **3**, 163-171 (2014).
158. Xue, H. et al. Design, construction, and characterization of a set of biosensors for aromatic compounds. *ACS Synth Biol* **3**, 1011-1014 (2014).
159. Siedler, S. et al. Development of a Bacterial Biosensor for Rapid Screening of Yeast p-Coumaric Acid Production. *ACS Synth Biol* **6**, 1860-1869 (2017).
160. Ding, N. et al. Programmable cross-ribosome-binding sites to fine-tune the dynamic range of transcription factor-based biosensor. *Nucleic Acids Res.* **48**, 10602-10613 (2020).
161. Saeki, K., Tominaga, M., Kawai-Noma, S., Saito, K. & Umeno, D. Rapid Diversification of BetI-Based Transcriptional Switches for the Control of Biosynthetic Pathways and Genetic Circuits. *ACS Synth Biol* **5**, 1201-1210 (2016).
162. Della Corte, D. et al. Engineering and application of a biosensor with focused ligand specificity. *Nat Commun* **11**, 4851 (2020).
163. Gustavsson, M. & Lee, S.Y. Prospects of microbial cell factories developed through systems metabolic engineering. *Microb Biotechnol* **9**, 610-617 (2016).
164. Wang, C., Pflieger, B.F. & Kim, S.W. Reassessing Escherichia coli as a cell factory for biofuel production. *Curr. Opin. Biotechnol.* **45**, 92-103 (2017).
165. Das, M., Patra, P. & Ghosh, A. Metabolic engineering for enhancing microbial biosynthesis of advanced biofuels. *Renew Sust Energ Rev* **119** (2020).
166. Choi, K.R., Jiao, S. & Lee, S.Y. Metabolic engineering strategies toward production of biofuels. *Curr. Opin. Chem. Biol.* **59**, 1-14 (2020).
167. Nissen, T.L., Kielland-Brandt, M.C., Nielsen, J. & Villadsen, J. Optimization of ethanol production in *Saccharomyces cerevisiae* by metabolic engineering of the ammonium assimilation. *Metab. Eng.* **2**, 69-77 (2000).
168. Nguyen, N.P., Raynaud, C., Meynial-Salles, I. & Soucaille, P. Reviving the Weizmann process for commercial n-butanol production. *Nat Commun* **9**, 3682 (2018).
169. Cheon, S., Kim, H.M., Gustavsson, M. & Lee, S.Y. Recent trends in metabolic engineering of microorganisms for the production of advanced biofuels. *Curr. Opin. Chem. Biol.* **35**, 10-21 (2016).
170. Peralta-Yahya, P.P., Zhang, F., del Cardayre, S.B. & Keasling, J.D. Microbial engineering for the production of advanced biofuels. *Nature* **488**, 320-328 (2012).
171. Cann, A.F. & Liao, J.C. Pentanol isomer synthesis in engineered microorganisms. *Appl. Microbiol. Biotechnol.* **85**, 893-899 (2010).
172. Donnelly, J., Horton, R., Gopalan, K., Bannister, C.D. & Chuck, C.J. Branched Ketone Biofuels as Blending Agents for Jet-A1 Aviation Kerosene. *Energy & Fuels* **30**, 294-301 (2016).

173. Hazelwood, L.A., Daran, J.M., van Maris, A.J.A., Pronk, J.T. & Dickinson, J.R. The Ehrlich pathway for fusel alcohol production: A century of research on *Saccharomyces cerevisiae* metabolism *Appl. Environ. Microbiol.* **74**, 3920-3920 (2008).
174. Lamsen, E.N. & Atsumi, S. Recent progress in synthetic biology for microbial production of C3-C10 alcohols. *Front Microbiol* **3**, 196 (2012).
175. Blombach, B. & Eikmanns, B.J. Current knowledge on isobutanol production with *Escherichia coli*, *Bacillus subtilis* and *Corynebacterium glutamicum*. *Bioeng Bugs* **2**, 346-350 (2011).
176. Nielsen, D.R. et al. Engineering alternative butanol production platforms in heterologous bacteria. *Metab. Eng.* **11**, 262-273 (2009).
177. Li, S., Wen, J. & Jia, X. Engineering *Bacillus subtilis* for isobutanol production by heterologous Ehrlich pathway construction and the biosynthetic 2-ketoisovalerate precursor pathway overexpression. *Appl. Microbiol. Biotechnol.* **91**, 577-589 (2011).
178. Boock, J.T. et al. Engineered microbial biofuel production and recovery under supercritical carbon dioxide. *Nat Commun* **10**, 587 (2019).
179. Vogt, M., Brusseler, C., Ooyen, J.V., Bott, M. & Marienhagen, J. Production of 2-methyl-1-butanol and 3-methyl-1-butanol in engineered *Corynebacterium glutamicum*. *Metab. Eng.* **38**, 436-445 (2016).
180. Blombach, B. et al. *Corynebacterium glutamicum* tailored for efficient isobutanol production. *Appl. Environ. Microbiol.* **77**, 3300-3310 (2011).
181. George, K.W. et al. Metabolic engineering for the high-yield production of isoprenoid-based C(5) alcohols in *E. coli*. *Sci Rep* **5**, 11128 (2015).
182. Zhang, K., Sawaya, M.R., Eisenberg, D.S. & Liao, J.C. Expanding metabolism for biosynthesis of nonnatural alcohols. *Proc Natl Acad Sci U S A* **105**, 20653-20658 (2008).
183. Connor, M.R. & Liao, J.C. Engineering of an *Escherichia coli* strain for the production of 3-methyl-1-butanol. *Appl. Environ. Microbiol.* **74**, 5769-5775 (2008).
184. Nitschel, R. et al. Engineering *Pseudomonas putida* KT2440 for the production of isobutanol. *Eng. Life Sci.* **20**, 148-159 (2020).
185. Miao, R., Xie, H., F, M.H. & Lindblad, P. Protein engineering of alpha-ketoisovalerate decarboxylase for improved isobutanol production in *Synechocystis* PCC 6803. *Metab. Eng.* **47**, 42-48 (2018).
186. Hammer, S.K., Zhang, Y. & Avalos, J.L. Mitochondrial Compartmentalization Confers Specificity to the 2-Ketoacid Recursive Pathway: Increasing Isopentanol Production in *Saccharomyces cerevisiae*. *ACS Synth Biol* **9**, 546-555 (2020).
187. Avalos, J.L., Fink, G.R. & Stephanopoulos, G. Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. *Nat. Biotechnol.* **31**, 335-341 (2013).

188. Jessop, P.G. The use of auxiliary substances (e.g. solvents, separation agents) should be made unnecessary wherever possible and innocuous when used. *Green Chemistry* **18**, 2577-2578 (2016).
189. Bahls, M.O., Kardashliev, T. & Panke, S. Novel Sensors for Engineering Microbiology (in Consequences of Microbial Interactions with Hydrocarbons, Oils, and Lipids: Production of Fuels and Chemicals). 331-357 (2017).
190. Kurth, E.G., Doughty, D.M., Bottomley, P.J., Arp, D.J. & Sayavedra-Soto, L.A. Involvement of BmoR and BmoG in n-alkane metabolism in 'Pseudomonas butanovora'. *Microbiology* **154**, 139-147 (2008).
191. Sayavedra-Soto, L.A., Doughty, D.M., Kurth, E.G., Bottomley, P.J. & Arp, D.J. Product and product-independent induction of butane oxidation in Pseudomonas butanovora. *FEMS Microbiol. Lett.* **250**, 111-116 (2005).
192. Yu, H. et al. Engineering transcription factor BmoR for screening butanol overproducers. *Metab. Eng.* **56**, 28-38 (2019).
193. Makart, S., Heinemann, M. & Panke, S. Characterization of the AlkS/P(alkB)-expression system as an efficient tool for the production of recombinant proteins in Escherichia coli fed-batch fermentations. *Biotechnol. Bioeng.* **96**, 326-336 (2007).
194. Panke, S., Meyer, A., Huber, C.M., Witholt, B. & Wubbolts, M.G. An alkane-responsive expression system for the production of fine chemicals. *Appl. Environ. Microbiol.* **65**, 2324-2332 (1999).
195. Martinez-Garcia, E. et al. SEVA 3.0: an update of the Standard European Vector Architecture for enabling portability of genetic constructs among diverse bacterial hosts. *Nucleic Acids Res.* **48**, D1164-D1170 (2020).
196. van Beilen, J.B. et al. Characterization of two alkane hydroxylase genes from the marine hydrocarbonoclastic bacterium Alcanivorax borkumensis. *Environ. Microbiol.* **6**, 264-273 (2004).
197. Eggink, G. et al. Alkane utilization in Pseudomonas oleovorans. Structure and function of the regulatory locus alkR. *J. Biol. Chem.* **263**, 13400-13405 (1988).
198. Rojo, F. Degradation of alkanes by bacteria. *Environ. Microbiol.* **11**, 2477-2490 (2009).
199. Fennewald, M., Benson, S., Oppici, M. & Shapiro, J. Insertion element analysis and mapping of the Pseudomonas plasmid alk regulon. *J. Bacteriol.* **139**, 940-952 (1979).
200. Kok, M. et al. The Pseudomonas oleovorans alkane hydroxylase gene. Sequence and expression. *J. Biol. Chem.* **264**, 5435-5441 (1989).
201. Moreno, R., Marzi, S., Romby, P. & Rojo, F. The Crc global regulator binds to an unpaired A-rich motif at the Pseudomonas putida alkS mRNA coding sequence and inhibits translation initiation. *Nucleic Acids Res.* **37**, 7678-7690 (2009).
202. Yuste, L. & Rojo, F. Role of the crc gene in catabolic repression of the Pseudomonas putida GPo1 alkane degradation pathway. *J. Bacteriol.* **183**, 6197-6206 (2001).

203. Canosa, I., Sanchez-Romero, J.M., Yuste, L. & Rojo, F. A positive feedback mechanism controls expression of AlkS, the transcriptional regulator of the *Pseudomonas oleovorans* alkane degradation pathway. *Mol. Microbiol.* **35**, 791-799 (2000).
204. van Beilen, J.B. et al. Analysis of *Pseudomonas putida* alkane-degradation gene clusters and flanking insertion sequences: evolution and regulation of the alk genes. *Microbiology* **147**, 1621-1630 (2001).
205. Reed, B., Blazeck, J. & Alper, H. Evolution of an alkane-inducible biosensor for increased responsiveness to short-chain alkanes. *J. Biotechnol.* **158**, 75-79 (2012).
206. Calles, B., Goni-Moreno, A. & de Lorenzo, V. Digitalizing heterologous gene expression in Gram-negative bacteria with a portable ON/OFF module. *Mol. Syst. Biol.* **15**, e8777 (2019).
207. Grund, A. et al. Regulation of alkane oxidation in *Pseudomonas putida*. *J. Bacteriol.* **123**, 546-556 (1975).
208. Beggah, S., Vogne, C., Zenaro, E. & Van Der Meer, J.R. Mutant HbpR transcription activator isolation for 2-chlorobiphenyl via green fluorescent protein-based flow cytometry and cell sorting. *Microb Biotechnol* **1**, 68-78 (2008).
209. <http://parts.igem.org/Promoters/Catalog/Anderson>. (accessed 01.02.2021).
210. Drummond, D.A., Iverson, B.L., Georgiou, G. & Arnold, F.H. Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins. *J. Mol. Biol.* **350**, 806-816 (2005).
211. Zhao, J., Kardashliev, T., Joelle Ruff, A., Bocola, M. & Schwaneberg, U. Lessons from diversity of directed evolution experiments by an analysis of 3,000 mutations. *Biotechnol. Bioeng.* **111**, 2380-2389 (2014).
212. de la Plaza, M., Fernandez de Palencia, P., Pelaez, C. & Requena, T. Biochemical and molecular characterization of alpha-ketoisovalerate decarboxylase, an enzyme involved in the formation of aldehydes from amino acids by *Lactococcus lactis*. *FEMS Microbiol. Lett.* **238**, 367-374 (2004).
213. Chen, G.S., Siao, S.W. & Shen, C.R. Saturated mutagenesis of ketoisovalerate decarboxylase V461 enabled specific synthesis of 1-pentanol via the ketoacid elongation cycle. *Sci Rep* **7**, 11284 (2017).
214. Baba, T. et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006 0008 (2006).
215. Zobel, S. et al. Tn7-Based Device for Calibrated Heterologous Gene Expression in *Pseudomonas putida*. *ACS Synth Biol* **4**, 1341-1351 (2015).
216. Levin-Karp, A. et al. Quantifying translational coupling in *E. coli* synthetic operons using RBS modulation and fluorescent reporters. *ACS Synth Biol* **2**, 327-336 (2013).
217. Jeschek, M., Gerngross, D. & Panke, S. Rationally reduced libraries for combinatorial pathway optimization minimizing experimental effort. *Nat Commun* **7**, 11163 (2016).

218. Gusyatiner, M.M., Lunts, M.G., Kozlov, Y.I., Ivanovskaya, L.V. & Voroshilova, E.B. DNA coding for mutant isopropylmalate synthase L-leucine-producing microorganism and method for producing L-leucine. U.S. Patent No. 6,403,342 (2002).
219. Hartigan, J.A. & Wong, M.A. Algorithm AS 136: A k-means clustering algorithm. *J R Stat Soc* **28**, 100-108 (1979).
220. Duetz, W.A. et al. Methods for intense aeration, growth, storage, and replication of bacterial strains in microtiter plates. *Appl. Environ. Microbiol.* **66**, 2641-2646 (2000).
221. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946-950 (2009).
222. OECD The 2007 OECD List of High Production Volume Chemicals. (2009).
223. Lisa, M.N. et al. Double autoinhibition mechanism of signal transduction ATPases with numerous domains (STAND) with a tetratricopeptide repeat sensor. *Nucleic Acids Res.* **47**, 3795-3810 (2019).
224. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. & Sternberg, M.J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858 (2015).
225. Berthold, C.L. et al. Structure of the branched-chain keto acid decarboxylase (KdcA) from *Lactococcus lactis* provides insights into the structural basis for the chemoselective and enantioselective carbonylation reaction. *Acta Crystallogr D Biol Crystallogr* **63**, 1217-1224 (2007).
226. Liu, F. et al. Bioconversion of distillers' grains hydrolysates to advanced biofuels by an *Escherichia coli* co-culture. *Microb Cell Fact* **16**, 192 (2017).
227. Duetz, W.A. & Witholt, B. Oxygen transfer by orbital shaking of square vessels and deepwell microtiter plates of various dimensions. *Biochem. Eng. J.* **17**, 181-185 (2004).
228. Lin, X. et al. Optimization and validation of a GC-FID method for the determination of acetone-butanol-ethanol fermentation products. *J. Chromatogr. Sci.* **52**, 264-270 (2014).
229. Rugbjerg, P., Dyerberg, A.S.B., Quainoo, S., Munck, C. & Sommer, M.O.A. Short and long-read ultra-deep sequencing profiles emerging heterogeneity across five platform *Escherichia coli* strains. *Metab. Eng.* (2020).
230. Dunlop, M.J., Keasling, J.D. & Mukhopadhyay, A. A model for improving microbial biofuel production using a synthetic feedback loop. *Syst Synth Biol* **4**, 95-104 (2010).
231. Sheppard, M.J., Kunjapur, A.M. & Prather, K.L.J. Modular and selective biosynthesis of gasoline-range alkanes. *Metab. Eng.* **33**, 28-40 (2016).
232. Machas, M.S., McKenna, R. & Nielsen, D.R. Expanding Upon Styrene Biosynthesis to Engineer a Novel Route to 2-Phenylethanol. *Biotechnol J* **12** (2017).
233. Sambrook, J.F. & Russel, D.W. *Molecular Cloning: A Laboratory Manual*, Edn. 3rd. (Cold Spring Harbor Laboratory, 2001).

234. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-6645 (2000).
235. Fernandez, S., de Lorenzo, V. & Perez-Martin, J. Activation of the transcriptional regulator XylR of *Pseudomonas putida* by release of repression between functional domains. *Mol. Microbiol.* **16**, 205-213 (1995).
236. Kapust, R.B. et al. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* **14**, 993-1000 (2001).
237. Fromant, M., Blanquet, S. & Plateau, P. Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal. Biochem.* **224**, 347-353 (1995).
238. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
239. Cherepanov, P.P. & Wackernagel, W. Gene Disruption in *Escherichia coli* - TcR and KmR Cassettes with the Option of Flp-Catalyzed Excision of the Antibiotic-Resistance Determinant. *Gene* **158**, 9-14 (1995).
240. Team, R.C. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing, Vienna, Austria* (2014).
241. Thorndike, R.L. Who Belongs in the Family. *Psychometrika* **18**, 267-276 (1953).
242. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676-682 (2012).
243. Choi, K.H. et al. A Tn7-based broad-range bacterial cloning and expression system. *Nat. Methods* **2**, 443-448 (2005).
244. Steegborn, C., Danot, O., Huber, R. & Clausen, T. Crystal structure of transcription factor MalT domain III: a novel helix repeat fold implicated in regulated oligomerization. *Structure* **9**, 1051-1060 (2001).
245. Choi, K.R. et al. Systems Metabolic Engineering Strategies: Integrating Systems and Synthetic Biology with Metabolic Engineering. *Trends Biotechnol.* **37**, 817-837 (2019).
246. Lee, S.Y. et al. A comprehensive metabolic map for production of bio-based chemicals. *Nat. Catal.* **2**, 18-33 (2019).
247. Martinez-Garcia, E. & de Lorenzo, V. *Pseudomonas putida* in the quest of programmable chemistry. *Curr. Opin. Biotechnol.* **59**, 111-121 (2019).
248. Nikel, P.I., Martinez-Garcia, E. & de Lorenzo, V. Biotechnological domestication of pseudomonads using synthetic biology. *Nat. Rev. Microbiol.* **12**, 368-379 (2014).
249. Ruhl, J., Schmid, A. & Blank, L.M. Selected *Pseudomonas putida* strains able to grow in the presence of high butanol concentrations. *Appl. Environ. Microbiol.* **75**, 4653-4656 (2009).

250. Blank, L.M., Ionidis, G., Ebert, B.E., Buhler, B. & Schmid, A. Metabolic response of *Pseudomonas putida* during redox biocatalysis in the presence of a second octanol phase. *FEBS J.* **275**, 5173-5190 (2008).
251. Wada, A. et al. Characterization of aromatic acid/proton symporters in *Pseudomonas putida* KT2440 toward efficient microbial conversion of lignin-related aromatics. *Metab. Eng.* **64**, 167-179 (2021).
252. 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* **13**, 2 (2014).
253. Xu, B., Jahic, M., Blomsten, G. & Enfors, S.O. Glucose overflow metabolism and mixed-acid fermentation in aerobic large-scale fed-batch processes with *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **51**, 564-571 (1999).
254. Kampers, L.F.C., Volkers, R.J.M. & Martins Dos Santos, V.A.P. *Pseudomonas putida* KT2440 is HV1 certified, not GRAS. *Microb Biotechnol* **12**, 845-848 (2019).
255. Nogales, J. et al. High-quality genome-scale metabolic modelling of *Pseudomonas putida* highlights its broad metabolic capabilities. *Environ. Microbiol.* **22**, 255-269 (2020).
256. Belda, E. et al. The revisited genome of *Pseudomonas putida* KT2440 enlightens its value as a robust metabolic chassis. *Environ. Microbiol.* **18**, 3403-3424 (2016).
257. Simon, O. et al. Analysis of the molecular response of *Pseudomonas putida* KT2440 to the next-generation biofuel n-butanol. *J Proteomics* **122**, 11-25 (2015).
258. Ankenbauer, A. et al. Micro-aerobic production of isobutanol with engineered *Pseudomonas putida*. *Eng. Life Sci.*, (in print) (2021).
259. Park, S.H., Kim, S. & Hahn, J.S. Metabolic engineering of *Saccharomyces cerevisiae* for the production of isobutanol and 3-methyl-1-butanol. *Appl. Microbiol. Biotechnol.* **98**, 9139-9147 (2014).
260. Bahls, M.O., Platz, L., Morgado, G., Schmidt, G.W. & Panke, S. Directed Evolution of Biofuel-Responsive Biosensors for Automated Optimization of Branched-Chain Alcohol Biosynthesis (in preparation, see chapter 3).
261. Vallon, T. et al. Applying systems biology tools to study n-butanol degradation in *Pseudomonas putida* KT2440. *Eng. Life Sci.* **15**, 760-771 (2015).
262. Atsumi, S., Hanai, T. & Liao, J.C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-89 (2008).
263. Martinez-Garcia, E., Aparicio, T., de Lorenzo, V. & Nikel, P.I. New transposon tools tailored for metabolic engineering of gram-negative microbial cell factories. *Front Bioeng Biotechnol* **2**, 46 (2014).
264. Xia, T. et al. Glucose consumption in carbohydrate mixtures by phosphotransferase-system mutants of *Escherichia coli*. *Microbiology* **163**, 866-877 (2017).

265. Curtis, S.J. & Epstein, W. Phosphorylation of D-glucose in *Escherichia coli* mutants defective in glucosephosphotransferase, mannosephosphotransferase, and glucokinase. *J. Bacteriol.* **122**, 1189-1199 (1975).
266. Volke, D.C., Friis, L., Wirth, N.T., Turlin, J. & Nikel, P.I. Synthetic control of plasmid replication enables target- and self-curing of vectors and expedites genome engineering of *Pseudomonas putida*. *Metab Eng Commun* **10**, e00126 (2020).
267. Sun, Z., Ramsay, J.A., Guay, M. & Ramsay, B.A. Automated feeding strategies for high-cell-density fed-batch cultivation of *Pseudomonas putida* KT2440. *Appl. Microbiol. Biotechnol.* **71**, 423-431 (2006).
268. Riveros-Rosas, H., Julian-Sanchez, A., Moreno-Hagelsieb, G. & Munoz-Clares, R.A. Aldehyde dehydrogenase diversity in bacteria of the *Pseudomonas* genus. *Chem. Biol. Interact.* **304**, 83-87 (2019).
269. Thompson, M.G. et al. Fatty Acid and Alcohol Metabolism in *Pseudomonas putida*: Functional Analysis Using Random Barcode Transposon Sequencing. *Appl. Environ. Microbiol.* **86** (2020).
270. Cook, T.B. et al. Genetic tools for reliable gene expression and recombineering in *Pseudomonas putida*. *J. Ind. Microbiol. Biotechnol.* **45**, 517-527 (2018).
271. Basler, G., Thompson, M., Tullman-Ercek, D. & Keasling, J. A *Pseudomonas putida* efflux pump acts on short-chain alcohols. *Biotechnol Biofuels* **11**, 136 (2018).
272. Kunjapur, A.M. & Prather, K.L. Microbial engineering for aldehyde synthesis. *Appl. Environ. Microbiol.* **81**, 1892-1901 (2015).
273. Xiong, M. et al. A bio-catalytic approach to aliphatic ketones. *Sci Rep* **2**, 311 (2012).
274. Saleski, T.E. et al. Syntrophic co-culture amplification of production phenotype for high-throughput screening of microbial strain libraries. *Metab. Eng.* **54**, 232-243 (2019).
275. Jacobs, M.A. et al. Comprehensive transposon mutant library of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* **100**, 14339-14344 (2003).
276. Harms, A. et al. A bacterial toxin-antitoxin module is the origin of inter-bacterial and inter-kingdom effectors of *Bartonella*. *PLoS Genet.* **13**, e1007077 (2017).
277. Thomason, L.C., Costantino, N. & Court, D.L. *E. coli* genome manipulation by P1 transduction. *Curr Protoc Mol Biol* **1**, 1-17 (2007).
278. Choi, K.H. & Schweizer, H.P. An improved method for rapid generation of unmarked *Pseudomonas aeruginosa* deletion mutants. *BMC Microbiol.* **5**, 30 (2005).
279. Hartmans, S., Smits, J.P., van der Werf, M.J., Volkering, F. & de Bont, J.A. Metabolism of Styrene Oxide and 2-Phenylethanol in the Styrene-Degrading *Xanthobacter* Strain 124X. *Appl. Environ. Microbiol.* **55**, 2850-2855 (1989).
280. Bagdasarian, M. et al. 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* **16**, 237-247 (1981).

281. Regenhardt, D. et al. Pedigree and taxonomic credentials of *Pseudomonas putida* strain KT2440. *Environ. Microbiol.* **4**, 912-915 (2002).
282. Na, D. et al. Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nat. Biotechnol.* **31**, 170-174 (2013).
283. Apura, P. et al. Tailor-made sRNAs: a plasmid tool to control the expression of target mRNAs in *Pseudomonas putida*. *Plasmid* **109**, 102503 (2020).
284. Xie, W.H., Deng, H.K., Hou, J. & Wang, L.J. Synthetic small regulatory RNAs in microbial metabolic engineering. *Appl. Microbiol. Biotechnol.* **105**, 1-12 (2021).
285. Ghodasara, A. & Voigt, C.A. Balancing gene expression without library construction via a reusable sRNA pool. *Nucleic Acids Res.* **45**, 8116-8127 (2017).
286. Yoo, S.M., Na, D. & Lee, S.Y. Design and use of synthetic regulatory small RNAs to control gene expression in *Escherichia coli*. *Nat Protoc* **8**, 1694-1707 (2013).
287. Wagner, E.G.H. & Romby, P. Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv. Genet.* **90**, 133-208 (2015).
288. Holmqvist, E. & Vogel, J. RNA-binding proteins in bacteria. *Nat. Rev. Microbiol.* **16**, 601-615 (2018).
289. Noh, M., Yoo, S.M., Kim, W.J. & Lee, S.Y. Gene Expression Knockdown by Modulating Synthetic Small RNA Expression in *Escherichia coli*. *Cell Syst* **5**, 418-426 e414 (2017).
290. Martinez-Garcia, E., Calles, B., Arevalo-Rodriguez, M. & de Lorenzo, V. pBAM1: an all-synthetic genetic tool for analysis and construction of complex bacterial phenotypes. *BMC Microbiol.* **11**, 38 (2011).
291. Lange, J., Takors, R. & Blombach, B. Zero-growth bioprocesses: A challenge for microbial production strains and bioprocess engineering. *Eng. Life Sci.* **17**, 27-35 (2017).
292. Rand, J.M. et al. A metabolic pathway for catabolizing levulinic acid in bacteria. *Nat Microbiol* **2**, 1624-1634 (2017).
293. Yang, D. et al. Repurposing type III polyketide synthase as a malonyl-CoA biosensor for metabolic engineering in bacteria. *Proc Natl Acad Sci U S A* **115**, 9835-9844 (2018).
294. Hoynes-O'Connor, A. & Moon, T.S. Development of Design Rules for Reliable Antisense RNA Behavior in *E. coli*. *ACS Synth Biol* **5**, 1441-1454 (2016).
295. Qi, L.S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173-1183 (2013).
296. Kim, S.K. et al. CRISPR interference-mediated gene regulation in *Pseudomonas putida* KT2440. *Microb Biotechnol* **13**, 210-221 (2020).
297. Batiannis, C. et al. An expanded CRISPRi toolbox for tunable control of gene expression in *Pseudomonas putida*. *Microb Biotechnol* **13**, 368-385 (2020).

298. Cui, L. et al. A CRISPRi screen in *E. coli* reveals sequence-specific toxicity of dCas9. *Nat Commun* **9**, 1912 (2018).
299. Dominguez-Cuevas, P., Ramos, J.L. & Marques, S. Sequential XylS-CTD binding to the Pm promoter induces DNA bending prior to activation. *J. Bacteriol.* **192**, 2682-2690 (2010).
300. Markham, N.R. & Zuker, M. DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res.* **33**, W577-581 (2005).
301. Subramanian, S.K., Russ, W.P. & Ranganathan, R. A set of experimentally validated, mutually orthogonal primers for combinatorially specifying genetic components. *Synth Biol (Oxf)* **3**, ysx008 (2018).
302. Schwartz, J.J., Lee, C. & Shendure, J. Accurate gene synthesis with tag-directed retrieval of sequence-verified DNA molecules. *Nat. Methods* **9**, 913-915 (2012).
303. Leveau, J.H. & Lindow, S.E. Predictive and interpretive simulation of green fluorescent protein expression in reporter bacteria. *J. Bacteriol.* **183**, 6752-6762 (2001).
304. Yang, D. et al. Expanded synthetic small regulatory RNA expression platforms for rapid and multiplex gene expression knockdown. *Metab. Eng.* **54**, 180-190 (2019).
305. Aparicio, T., Nyerges, A., Martinez-Garcia, E. & de Lorenzo, V. High-Efficiency Multi-site Genomic Editing of *Pseudomonas putida* through Thermoinducible ssDNA Recombineering. *IScience* **23**, 100946 (2020).
306. Aparicio, T., de Lorenzo, V. & Martinez-Garcia, E. CRISPR/Cas9-Based Counterselection Boosts Recombineering Efficiency in *Pseudomonas putida*. *Biotechnol J* **13**, e1700161 (2018).
307. Durante-Rodriguez, G., de Lorenzo, V. & Nikel, P.I. A Post-translational Metabolic Switch Enables Complete Decoupling of Bacterial Growth from Biopolymer Production in Engineered *Escherichia coli*. *ACS Synth Biol* **7**, 2686-2697 (2018).
308. Gentleman, R.C. et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**, R80 (2004).
309. Bahls, M.O., Pfeifer, F. & Panke, S. Short-Chain Alcohol and Fatty Acid Biosynthesis in *Pseudomonas putida* with co-culture based in vivo product detection. (in preparation, see chapter 4).
310. Maier, T. et al. Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Mol. Syst. Biol.* **7**, 511 (2011).
311. Lahtvee, P.J., Seiman, A., Arike, L., Adamberg, K. & Vilu, R. Protein turnover forms one of the highest maintenance costs in *Lactococcus lactis*. *Microbiology* **160**, 1501-1512 (2014).
312. Connor, M.R., Cann, A.F. & Liao, J.C. 3-Methyl-1-butanol production in *Escherichia coli*: random mutagenesis and two-phase fermentation. *Appl. Microbiol. Biotechnol.* **86**, 1155-1164 (2010).

Chapter 8: Acknowledgements

It is a great pleasure to acknowledge all the people that made this scientific work and personal adventure possible and fruitful. Apologies to everyone not mentioned personally, but still I want to express my gratitude to all the unnamed and uncountable contributors.

First and foremost, I like to thank Prof. Sven Panke for almost a decade of scientific guidance through master and PhD studies, while allowing enough freedom to explore new ideas and learn from failures. Special thanks to Prof. Victor de Lorenzo and Prof. Ralf Takors for co-examining this thesis and in particular for contributing to this project through feedback during the consortium meetings. Thank you to Prof. Martin Fussenegger for taking the time to co-examine the thesis.

I am grateful to my thesis students, Felix Pfeifer and Lukas Platz, for enthusiasm, hard work and valuable contributions to my research.

Special thanks to the whole EmPowerPutida consortium, led by Prof. Vitor dos Santos, for scientific input as well as outstanding assembly meetings. Additional thanks to Prof. Bernhard Hauer and Max for help with GC-MS, Robert and Andreas of Prof. Takors group for exchange regarding alcohol production, Prof. Cecilia Arraiano, Patricia, Sandra, and Susana for input regarding sRNA expression in *P. putida*, the anonymous SEVA distributors from Prof. de Lorenzo's lab, as well as Belen and Esteban for general help with *P. putida* related questions.

Of course, thanks to all lab members of the BPL. As the years changed, so did the people around. I representatively start with thanking the first colleague I remember graduating, Sonja, to the most recent one joining, Aurellia. I would like to highlight the impact of Daniel, Markus, Martin H, Tania and Tsvetan on my work and thank them for their helpful suggestions. Many thanks to all the additional lab members that became not only colleagues but great friends that will remain in my life way beyond graduation day. Also, thanks for weekend trips to such lovely places as Cambridge, Munich, Paris, Sofia, and Vienna as well as the soccer world cup in Russia (with best regards to Sai's group).

Finally yet importantly, thanks to everyone who contributed to the Thursday afternoon aperos, an integral part of studies preventing psychological meltdowns after unsuccessful experiments. Besides, thanks to everyone who was part of the fun and sports programs, in particular thanks to everyone who was part of the VMB board during my time serving. Last but not least, I have to say thanks to BPL's "Secret Santa" for the Swiss army knife, it was quite a help in a bunch of otherwise (almost) hopeless situations.

Thank you FIS and facility members, for making the D-BSSE a great workplace while also forgiving us the occasional traces of malt after brewing in the science lounge. Thanks to the automation and single cell facilities, namely Gregor, Telma, Tom and Verena.

Finally special thanks to my parents and my brother, as well as a few unnamed but close friends, for giving me the freedom to work many yet happy extra hours - apologies for the little time I was around. However, without you I would not have had the stable environment to make this endeavor happen, thank you for your unconditional love and patience.

Chapter 9: Curriculum Vitae

Maximilian Ole Bahls

Born: 30th June 1989 in Wipperfürth, Germany

Education

- Since 2015 PhD student, Bioprocess Laboratory, Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland
- 2012-2015 Master in Biotechnology (major in Synthetic Biology), ETH Zurich, Switzerland, Thesis: "Development of a fed-batch process for streptavidin production in biotin-independent strains of *Escherichia coli*."
- 2012 IARU Exchange program, NUS, Singapore
- 2009-2012 Bachelor in Biology (chemical direction), ETH Zurich, Switzerland
- 2008-2009 Wadden Sea National Park (civilian service), Norderney, Germany
- 2008 Abitur, Anne-Frank Gymnasium, Halver, Germany

Summer schools and workshops

- 2018 *Pseudomonas putida* Grassroots meeting, Frankfurt, Germany
- 2017 Advanced Course on Metabolic Engineering and Systems Biology, Chalmers University of Technology Gothenburg, Sweden
- 2014 iGEM jamboree Boston, MA, USA, gold medal, "Best Modelling" award
- 2014 ESBES Bioprocess Engineering course, Brač, Croatia

Activities within the scientific community

- 2020 Judge at the International Genetically Engineered Machine Competition (iGEM) jamboree
- Since 2018 Jury member at the St. Gallen Symposium Global Essay Competition

Publications

Bahls, M.O., Kardashliev, T. and Panke, S., 2017. Novel Sensors for Engineering Microbiology. In S. Y. Lee (Ed.) *Production of Fuels and Chemicals: Springer International Publishing*.

Jeschek, M., **Bahls, M.O.**, Schneider, V., Marlière, P., Ward, T.R. and Panke, S., 2017. Biotin-independent Strains of *Escherichia coli* for Enhanced Streptavidin Production. *Metabolic Engineering*.

Conference participations

2019 Synthetic Biology: Engineering, Evolution & Design (SEED), New York City, NY, USA

2018 ChemSpec Europe, Cologne, Germany

2016 European Symposium on Biochemical Engineering Sciences 2016, Dublin, Ireland

2014 International Genetically Engineered Machine (iGEM) competition, Boston, MA, USA