# Molecular tools for genome engineering of *Corynebacterium glutamicum*

Christiane Katharina Sonntag

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
Forschungszentrum

# Molecular tools for genome engineering

# of *Corynebacterium glutamicum*

**Inaugural Dissertation**

submitted to

the Faculty of Mathematics and Natural Sciences

of the Heinrich Heine University Düsseldorf

presented by

**Christiane Katharina Sonntag**

born in Bad Mergentheim

Jülich, April 2020

# Molecular tools for genome engineering of *Corynebacterium glutamicum*

Christiane Katharina Sonntag

This thesis in hand has been performed at the Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, from November 2016 until April 2020 under the supervision of Prof. Dr. Michael Bott and Prof. Dr. Jan Marienhagen.

Printed with the permission of

the Faculty of Mathematics and Natural Sciences

of the Heinrich Heine University Düsseldorf

Examiner:          **Prof. Dr. Michael Bott**
                   Institute of Bio- and Geosciences, IBG-1: Biotechnology
                   Forschungszentrum Jülich GmbH

Co-examiner:       **Prof. Dr. Matias Zurbriggen**
                   Institute of Synthetic Biology
                   Heinrich Heine University Düsseldorf

Date of oral examination: 2nd November 2020

Results presented in this thesis have been published in two original articles.

**Krumbach K., Sonntag C.K., Eggeling L., Marienhagen J.** (2019) CRISPR/Cas12a Mediated Genome Editing to Introduce Amino Acid Substitutions into the Mechanosensitive Channel MscCG of *Corynebacterium glutamicum*. ACS Synth. Biol. 8:2726–2734. doi: 10.1021/acssynbio.9b00361

Permission to reuse in this dissertation

This article was published in ACS Synthetic Biology, 8, 2726–2734, Copyright American Chemical Society (2019).

Accessible under https://pubs.acs.org/doi/10.1021/acssynbio.9b00361

License: CC-BY-NC-ND pubs.acs.org/page/policy/authorchoice_ccbyncnd_termsofuse.html

**Sonntag C.K., Flachbart L.K., Maass C., Vogt M., Marienhagen J.** (2020) A Unified Design Allows Fine-tuning of Biosensor Parameters and Application Across Bacterial Species. Metab. Eng. Commun. 11 (September), e00150. doi: 10.1016/j.mec.2020.e00150

Permission to reuse in this dissertation

This article was published in Metabolic Engineering Communications, 11, e00150, Copyright Elsevier (2020).

Accessible under https://www.sciencedirect.com/science/article/pii/S221403012030050X

License: Attribution 4.0 International creativecommons.org/licenses/by/4.0/

# Table of contents

## Abstract

Facing the demand for environmental friendly and sustainable production processes, microorganisms are engineered for the industrial biosynthesis of chemicals, fuels, or food and feed additives from renewable resources. However, microbial strain development is still laborious, time-consuming and expensive, which constricts the transition to a more bio-based economy. Therefore, development and consistent improvement of molecular tools for genetic engineering as well as methods for the high-throughput characterization of engineered strain variants are of great importance.

For this purpose, the CRISPR/ Cas12a recombineering method for *Corynebacterium glutamicum*, a well-characterized microorganism employed in the industrial amino acid production, was refined by developing the flexible and easy to assemble crRNA delivery vector pJYScr. Targeting and editing efficiency of this new CRISPR/ Cas12a system was systematically evaluated by inserting genetic mutations proximal and distal to a selected PAM site in a genomic *lacZ* gene encoding for β-galactosidase. Subsequently, this improved method allowing for accelerated genome editing of *C. glutamicum* was applied in a strain engineering campaign aiming for improved ʟ-glutamate efflux. For this purpose single-stranded DNA oligonucleotides targeting critical amino acid residues in the mechanosensitive channel MscCG of *C. glutamicum* were used for CRISPR/ Cas12 recombineering. Several generated strain variants were characterized with regard to their respective ʟ-glutamate efflux identifying new gain-of-function mutations, which improve ʟ-glutamate export in *C. glutamicum*.

To the same extent as fast and reliable genetic engineering, rapid identification of producing strain variants in larger libraries is a crucial step in strain development. In this respect, transcription factor-based, fluorescent biosensors are valuable tools in metabolic engineering allowing for semiquantitative determination of metabolites in single cells. However, transcriptional biosensors are often limited by intrinsic characteristics of the used native regulatory circuit. Moreover, signal saturation at low inducer concentrations typically limits their use in producer strains at advanced engineering stages, and the application of biosensors in heterologous host systems is often not possible.

Therefore, a unified biosensor design was established, which allows fine-tuning of important sensor parameters and ensures a sensor response in a heterologous expression host. As a key feature of the design, the regulator activity can be controlled through modulation of the regulator gene expression level by using different (synthetic) constitutive promoters. Several biosensors based on transcriptional regulators LysG and PhdR and their cognate promoters from *C. glutamicum* were constructed for applications in the native host and in *Escherichia coli*. Detailed characterization of these biosensors in liquid cultures and on the single-cell level using flow cytometry showed that the sensor design enables customization of important biosensor parameters as well as application of these sensors in two different bacterial species.

## Zusammenfassung

Angesichts der Forderungen nach umweltfreundlichen und nachhaltigen Produktionsverfahren werden Mikroorganismen für die biotechnologische Produktion von Chemikalien, Kraftstoffen oder Lebensmittel- und Futtermittelzusätzen aus erneuerbaren Ressourcen entwickelt. Die Stammentwicklung ist jedoch mühsam, zeitaufwendig und teuer. Daher ist die Entwicklung und Verbesserung von molekularen Werkzeugen für die Gentechnik und Methoden zur Hochdurchsatz-Charakterisierung in der Stammentwicklung von großer Bedeutung.

Aus diesem Grund wurde das CRISPR/ Cas12a-Rekombinations-System in *Corynebacterium glutamicum*, durch die Entwicklung des leicht zu modifizierenden crRNA-delivery Vektors pJYScr verfeinert. Die Targeting- und Editiereffizienz des CRISPR/ Cas12a-Systems in Kombination mit dem neuen pJYScr Plasmid und einzelsträngigen DNA-Oligonukleotiden wurde systematisch evaluiert, indem genetische Mutationen proximal und distal zu einer ausgewählten PAM-Stelle in ein genomisches *lacZ*-Gen eingefügt wurden, welches für die β-Galaktosidase kodiert. Durch die Vereinfachung der Anwendung des CRISPR/ Cas12a-Rekombinationssystems wurde dies schließlich eingesetzt, um eine Codon-Sättigungsmutagenese an kritischen Aminosäureresten im mechanosensitiven Kanal MscCG durchzuführen. Die erzeugten Varianten wurden hinsichtlich eines erhöhten L-Glutamat-Effluxes charakterisiert, wodurch neue Mutationsorte in MscCG identifiziert werden konnten, die zu einer verbesserten L-Glutamatsekretion führen.

Ebenso wie der Einsatz von zuverlässigen Methoden zur genetischen Manipulation in der Stammentwicklung sind auch geeignete Werkzeuge für die schnelle Identifizierung produzierender Stammvarianten in größeren Mutanten-Bibliotheken bedeutend. In diesem Zusammenhang ermöglichen auf Transkriptionsregulatoren basierende, fluoreszierende Biosensoren die semiquantitative Bestimmung von Metaboliten auf Einzelzellebene. Die Anwendungsmöglichkeiten solcher transkriptionellen Biosensoren sind jedoch oft durch die intrinsischen Eigenschaften des nativen Regulationskreislaufs limitiert. So kann der Einsatz von Biosensoren durch Signalsättigung bei niedrigen Konzentrationen oder einer verringerten oder gar fehlenden Funktionalität in heterologen Organismen stark eingeschränkt sein.

Aus diesem Grund wurde ein vereinheitlichtes Biosensordesign entwickelt, das nicht nur die Feinabstimmung wichtiger Sensorparameter erlaubt, sondern auch den Einsatz von Sensoren in heterologen Organismen ermöglicht. Die Sensorantwort wird anhand der Expressionsstärke des jeweiligen Gens für den Transkriptionsregulator durch verschiedene (synthetische) konstitutive Promotoren gesteuert. Anhand dieses neuen Designprinzips wurde ein Set von Biosensoren auf Basis der Transkriptionsregulatoren LysG und PhdR und ihrer jeweiligen Zielpromotoren aus *C. glutamicum* für den Einsatz im nativen Wirtssystem und in *Escherichia coli* konstruiert. Die detaillierte Charakterisierung dieser Biosensoren in Flüssigkultur sowie in Durchflusszytometrieexperimenten zeigte, dass das Sensordesign die Modifizierung wichtiger Biosensorparameter erlaubt, und eine Anwendung in verschiedenen Bakterien ermöglicht.

## Abbreviations

| | |
|---|---|
| AEC | L-aminoethylcysteine |
| ATCC | American Type Culture Collection |
| bp | Base pair |
| *C. glutamicum* | *Corynebacterium glutamicum* |
| Cas | CRISPR-associated gene |
| cf | Latin *confer/conferatur* |
| CFP | Cyan fluorescent protein |
| cfu | Colony forming units |
| coA | Coenzyme A |
| cou | *p*-coumaric acid |
| CRISPR | Clustered regularly interspaced short palindromic repeat |
| crRNA | CRISPR RNA |
| DNA | Deoxyribonucleic acid |
| EU | European Union |
| *E. coli* | *Escherichia coli* |
| eYFP | Enhanced yellow fluorescent protein |
| FACS | Fluorescence-activated cell sorting |
| FRET | Förster resonance energy transfer |
| GOF | Gain-of-function |
| GRAS | Generally recognized as safe |
| HDR | Homology-directed repair |
| HPLC | High pressure liquid chromatography |
| HTS | High-throughput screening |
| IPP | Isopentenyl diphosphate |
| IPTG | Isopropyl-β-D-thiogalactopyranosid |
| $K_D$ | Dissociation constant |
| M | Molar |
| M | Moderate |
| MNNG | 1-Methyl-3-nitroguanidine |
| MS | Mass spectrometry |
| NHEJ | Non-homologous end joining |
| nt | Nucleotide |
| NUC | Nuclease |
| OCG | Originally controlled gene |
| PAM | Protospacer adjacent motif |

| | |
|---|---|
| P$_{BAD}$ | Arabinose-inducible *araBAD* promoter |
| RBS | Ribosome binding site |
| REC | Recognition |
| RNA | Ribonucleic acid |
| S | Strong |
| SacB | Levansucrase-activity |
| SELEX | Systematic Evolution of Ligands by Exponential enrichment |
| sgRNA | Single guide RNA |
| TALEN | Transcription Activator-Like Effector Nucleases |
| TF | Transcription factor |
| tracrRNA | Trans-activating crRNA |
| UV | Ultra-violet |
| W | Weak |
| X-Gal | 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside |
| YFP | Yellow fluorescent protein |

## Scientific context and key results of the thesis

## 1    Limiting greenhouse gas emissions– shifting to a bio-based chemical production

Products derived from crude oil, such as fossil fuels and petrochemicals play an essential role in our modern society. The dependence on these products for the maintenance of modern lifestyles is mainly founded on petroleum-based transportation and energy generation. In 2017, both sectors had the highest share of 85.9% in crude oil consumption for the 28 states of the European Union (EU) (Eurostat, 2019a). A third major industrial sector contributing to the overall crude oil consumption is the chemical industry, which accounts for 14.1% of total crude oil consumption in the EU (Eurostat, 2019a). As a result of the high demand in the above mentioned sectors, crude oil production and import reached 565.7 and 66.7 million tons of oil equivalent in the EU in 2017, respectively (Eurostat, 2019a).

Relying on petroleum-based feedstocks as well as on high energy consumption for production processes, the chemical industry contributes to greenhouse gas emissions with 66 million tons of $CO_2$ equivalents per year (Eurostat, 2019b). These emissions are one major driving force in the anthropogenic climate change causing the melting of polar ice caps and extreme weather conditions such as increased precipitation and rising temperatures. Tackling climate change, the EU committed itself to be $CO_2$-neutral until 2050 by enacting the first European Climate Law, the Paris Agreement on climate change (United Nations Framework Convention on Climate Change, 2016).

Following the worldwide efforts to limit the rise of the average global temperature to a maximum of 2°C with respect to pre-industrial levels (United Nations Framework Convention on Climate Change, 2016), amongst others energy consumption and greenhouse gas emissions must be reduced. Aiming for these targets, a sustainable production of fuels, chemicals, and materials is inevitable by using energy-saving processes and replacing petroleum-derived feedstocks with renewable resources (Buijs *et al.*, 2013; Harmsen *et al.*, 2014; Lokko *et al.*, 2018). In this respect, feedstocks such as plant-derived waste products or side products from industry need to be recycled, and hence being reintroduced into industrial production processes (Harmsen *et al.*, 2014). Comparing energy consumption of bio-based processes to the respective chemical synthesis, rather low energy input is requested compared to chemical reactions (Straathof *et al.*, 2019). Therefore, a shift from conventional chemical production towards biotechnological processes can support the ongoing path towards a sustainable bioeconomy (Scarlat *et al.*, 2015; Straathof *et al.*, 2019).

Against this background, industrial biotechnology was identified as key industry and received a great deal of attention (Harmsen *et al.*, 2014; Lokko *et al.*, 2018; Scarlat *et al.*, 2015; Straathof

*et al.*, 2019). In industrial biotechnology, the catalytic potential of enzymes or whole microorganisms is used to produce chemical compounds of interest for virtually all different industries. A multitude of biotechnological processes have already been established and commercialized for the industrial production of a plethora of different products (Figure 1).
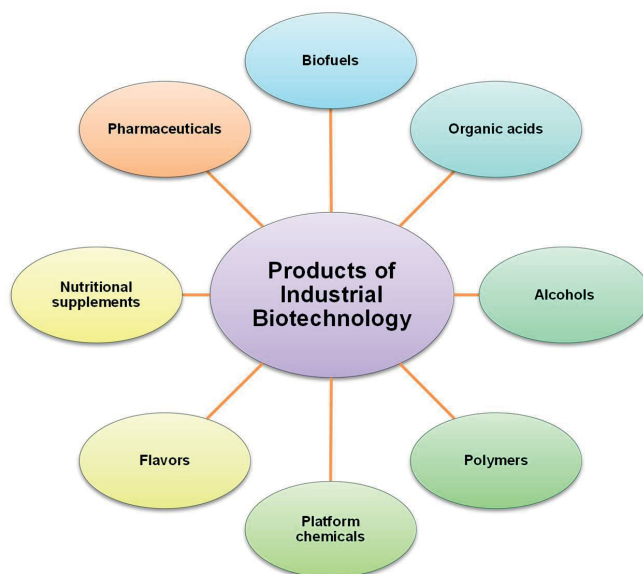


**Figure 1: Typical biotechnological products.**

A broad array of compounds such as biofuels, organic acids, alcohols, polymers, platform chemicals, flavors, nutritional supplements and pharmaceuticals are produced at an industrial scale (Adrio & Demain, 2010; Leuchtenberger *et al.*, 2005; Soetaert & Vandamme, 2006). Typically, the respective enzymes or microorganisms intended for production, need to be modified to enable or increase formation of the desired products.

## 2 Microbial strain development in Industrial Biotechnology

In classical strain development, isolated wild type strains are assessed for their potential as producer strains for a compound of interest (Parekh *et al.*, 2000). In most cases, microbial strains identified in broad screening campaigns only accumulate insufficient amounts of a desired compound in terms of industrial needs. Usually, the metabolism of these wild type strains is tightly regulated to avoid wasteful expenditure of energy or available carbon for the accumulation of metabolites to industrially relevant levels. In this respect, genetic manipulation can unleash or even overcome genetic control as well as rate-limiting enzyme reactions within metabolic pathways. Immediately after isolation of an organism, there is only limited knowledge of physiology or genetic background present. Therefore, the application of undirected strain development techniques following a lab-scale evolutionary approach are straight-forward to isolate strain variants with the desired production phenotype (Parekh 1999). In this context, whole genome mutagenesis methods such as ultra-violet (UV)-irradiation and chemical mutagenesis with *e. g.* 1-Methyl-3-nitroguanidine (MNNG) are used to introduce random mutations into the genetic background. This allows rapid generation of vast mutant libraries, which might encompass clones with beneficial mutations resulting in increased product formation (Parekh *et al.*, 2000; Patnaik, 2008). The identification of the anticipated phenotype involves characterization of every single generated variant by cultivation in liquid cultures followed by the evaluation of product formation.

The production phenotype characterization of mutants involves laborious and time-consuming evaluation of product titers using analytical methods such as high pressure liquid chromatography (HPLC) and mass spectrometry (MS) (Patnaik, 2008; Schallmey *et al.*, 2014). Despite increased product formation, not all randomly introduced mutations in the genetic strain background contribute to the improved production of the target molecule. Hence, poor growth, and thus, low space-time yields of the isolated clones is one of the main and undesired consequences (Ikeda *et al.*, 2009; Warner *et al.*, 2009).

Despite these limitations, this classical strain engineering approach yielded many producing microorganisms, which were subsequently studied in more detail (Parekh *et al.*, 2000). One of these microorganisms, which is nowadays the most important platform organisms for the industrial-scale production of proteinogenic amino acids is *Corynebacterium glutamicum* (Ikeda & Takeno, 2013).

## 2.1 *Corynebacterium glutamicum* - a versatile platform organism in industrial biotechnology

*C. glutamicum* was first isolated in a screening campaign searching for a microbial L-glutamate producer in Japan in 1957 (Kinoshita *et al.*, 1957; Udaka, 1960). This Gram-positive, non-pathogenic and biotin-auxotrophic soil bacterium belongs to the suborder of *Corynebacterianeae* (Figure 2) (Eggeling & Bott, 2005). The determination and annotation of the complete genomic sequence of *C. glutamicum* ATCC 13032 (American Type Culture Collection) was accomplished in 2003 (Kalinowski *et al.*, 2003). Furthermore, decades of studies on cell biology, physiology and metabolic regulation conducted in many laboratories around the world led to detailed understanding of this microorganism, which is comparable to *E. coli*.

Due to its comparatively fast growth and ability to reach high cell densities in defined liquid medium throughout different cultivation scales, *C. glutamicum* is used as biocatalyst in several biotechnological processes (Becker *et al.*, 2018; Grünberger *et al.*, 2013). Since *C. glutamicum* does not produce endotoxins, fermentation products have the "Generally Recognized As Safe" (GRAS) status (Baritugo *et al.*, 2018).



**Figure 2: Electron microscopic image of *C. glutamicum* ATCC 13032** (Sahm *et al.*, 2000)**.**

Out of the various products obtained with *C. glutamicum* up to date, L-amino acids were the first industrially produced compounds and hold the largest market share of these. Strain engineering allowed for the industrial production of L-glutamate and L-lysine at a million ton scale per year (Eggeling & Bott, 2015; Lee & Wendisch, 2017). Due to the increasing demand for amino acids in the food and feed industry, this market still continues to rise (IMARC Group, 2020; Leuchtenberger *et al.*, 2005; Wendisch *et al.*, 2016). The annual amino acid production was 8.9 million tons in 2018 and is expected to further increase to 11.5 million tons by 2024 (IMARC Group, 2020).

### 2.1.1  *C. glutamicum* strain engineering for industrial production of L-amino acids

In the early 1960s, only limited knowledge was present with respect to *C. glutamicum* physiology and genetics associated with amino acid production. Therefore, classical strain development approaches were used to improve product formation. A main focus was set on L-glutamate production from *C. glutamicum,* due to its broad application as flavor enhancer (Kinoshita *et al.*, 1957).

With the expanding feed industry, the interest for microbial production of other amino acids such as L-lysine sparked (Eggeling & Bott, 2015). With the aim to establish L-lysine production using *C. glutamicum*, a random mutagenesis/screening approach was best suited to aim for decoupling of allosteric inhibitory effects of key enzymes in biosynthetic pathways, improving substrate uptake and product export as well as eliminating product degradation (Eggeling, 1994). Due to the generation of large mutant libraries, the isolation of improved producer strains asked for an elaborated selection/screening process, in which L-lysine producers could be identified easily. Isolation of mutants with the desired phenotype was achieved by applying a selection pressure on the generated mutant library using the L-lysine analogue L-aminoethylcysteine (AEC) (Eggeling, 1994). AEC can substitute for L-lysine in feedback inhibition of the aspartate kinase in *C. glutamicum* (Kalinowski *et al.*, 1991), which is a key enzyme in the amino acid metabolism. A typical feature of some L-lysine producers is AEC resistance, due to a mutated aspartate kinase, which is then insensitive to allosteric control by both L-lysine and L-threonine (Tosaka *et al.*, 1983). Therefore, mutants that grew despite the presence of AEC often obtained increased L-lysine production resulting also in L-lysine excretion. Starting from these isolated mutants, further targets in *C. glutamicum*, such as pathway drains and L-lysine efflux were assessed to increase L-lysine titers, in order to finally isolate improved production strains reaching up to 170 g/L L-lysine titers (Eggeling, 1994; Eggeling & Sahm, 1999).

## 2.2  Rational engineering of *C. glutamicum* through metabolic engineering

The application of classical strain development allowed for the isolation of *C. glutamicum* L-glutamate and L-lysine producer strains, which are still used in an industrial setting (Eggeling & Sahm, 1999). With technological progress, new routes were taken in the engineering of *C. glutamicum* to enlarge the product array of the industrial workhorse. The expanding knowledge on the metabolism and physiology of microorganisms as well as the development of genetic tools complemented existing strain engineering strategies and allowed for the development of new powerful methods. In particular targeted DNA manipulation allowed for engineering of genes encoding for key enzymes, which are involved in product formation and metabolic regulation (Kirchner & Tauch, 2003).

### 2.2.1  Development of a molecular toolbox for genetic engineering

For successful genetic manipulation of *C. glutamicum*, a set of established methods and tools for DNA editing and gene expression is compulsory. A straightforward approach, to introduce a gene or gene cluster of interest into a particular organism, is episomal gene expression from plasmids. Therefore, a well-characterized set of compatible cloning vectors and the development of DNA transfer methods are a prerequisite for the introduction of foreign DNA into a new expression host.

For *C. glutamicum,* a multitude of expression plasmids, mostly *Escherichia coli* / *C. glutamicum* shuttle vectors, have been constructed, which allow for the controlled expression of native and heterologous genes. Ensuring gene expression, different well-established recombinant gene expression systems have also been introduced to *C. glutamicum*. The regulation of gene transcription in these systems is controlled by Isopropyl-β-D-thiogalactopyranosid (IPTG)-, heat-, carbon source- and anhydrotetracycline inducible promoters (Eikmanns *et al.*, 1991; Jakoby *et al.*, 1999; Kirchner & Tauch, 2003; Kortmann *et al.*, 2015; Lausberg *et al.*, 2012; Okibe *et al.*, 2010; Park *et al.*, 2008; Tsuchiya & Morinaga, 1988). Furthermore, a broad range of different constitutive promoters are available in *C. glutamicum,* amongst others the well characterized *dapA* promoter library (Pátek & Nešvera, 2011; Vašicová *et al.*, 1999).

Stable genetic manipulation of the *C. glutamicum* genome, such as gene integration, inactivation or deletion is most frequently performed by a two-step homologous recombination method using the "SacB system" (Niebisch & Bott, 2001; Schäfer *et al.*, 1994). The SacB system, based on the conditionally lethal levansucrase-activity (SacB) in the presence of sucrose, allows positive selection of double recombinants involving allelic exchange from a suicide vector (Jäger *et al.*, 1992).

### 2.2.2  System-wide engineering of microorganisms

With the advent of stoichiometric models of the microbial metabolism, a system-wide perspective on metabolic engineering helped to further improve *C. glutamicum*-based production strains. These methods are referred to as omics technologies, such as genome-, transcriptome-, proteome-, metabolome- and fluxome analysis (Kohlstedt *et al.*, 2010). In systems metabolic engineering approaches, data derived from systems biology and computational modelling are combined to assist the construction process of new efficient cell factories (Figure 3) (Dai & Nielsen, 2015; Lee & Kim, 2015).

In recently performed studies, systems metabolic engineering was applied in *C. glutamicum* enabling the production of L-glutarate, L-histidine and the formation of L-lysine from mannitol (Hoffmann *et al.*, 2018; Rohles *et al.*, 2018; Schwentner *et al.*, 2019). Following this systems

metabolic engineering approach, more versatile and universal applicable recombinant DNA methods for genome engineering are needed.
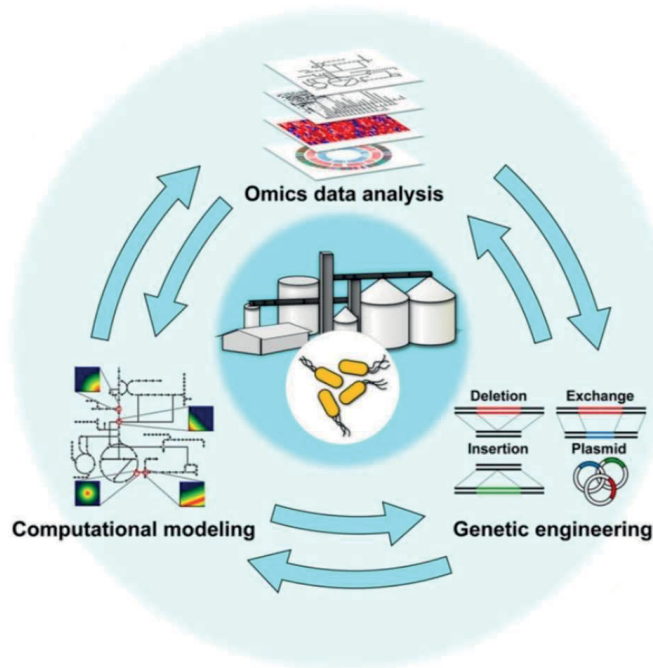


**Figure 3: Systems- and Systemic Metabolic engineering of microbial cell factories. Adapted from Yang et al. 2017**

Recently, the molecular toolbox for genetic manipulation of *C. glutamicum* and other industrially relevant microorganisms was extended by the clustered regularly interspaced short palindromic repeat – CRISPR-associated gene (CRISPR-Cas) system, which allows for precise genome engineering at defined positions in the genome.

# 3    The CRISPR-Cas system – a module of the bacterial immune system

The CRISPR-Cas system is derived from the adaptive bacterial immune response, which is part of the bacterial immune system. Bacteria are constantly threatened by viruses (bacteriophages) and have evolved a wide array of defense mechanisms to cope with phage infections (Rostøl & Marraffini, 2019). These emerged mechanisms of the innate prokaryotic immune system are based on different modes of action to combat phage infection at different stages of the microorganisms life cycle (van Houte *et al.*, 2016). For the defense against phage infections, different mechanisms have been described: hindrance of phage adsorption and/ or gene injection, phage-DNA sensing proteins, which induce cell apoptosis and several DNA degradation mechanisms such as the well-characterized restriction-modification system (Bikard & Marraffini, 2012).

Furthermore, protection is conferred by the CRISPR-Cas system, which is part of the adaptive immune system of prokaryotes to provide immunity against viral or plasmid DNA (Barrangou *et al.*, 2007). The CRISPR-Cas system offers the ability to retain a memory of previous phage infections to elicit a protective immune response upon reinfection. Present in approximately 45% of bacteria and 85% of archaea, CRISPR systems have been classified with regard to the *cas* gene content into two classes, six types and more than 20 subtypes (Figure 4) (Koonin *et al.*, 2017). The first described Cas9 system originating from *Streptococcus thermophilus* is characterized as a class 2 type II system (Garneau *et al.*, 2010; Koonin *et al.*, 2017).
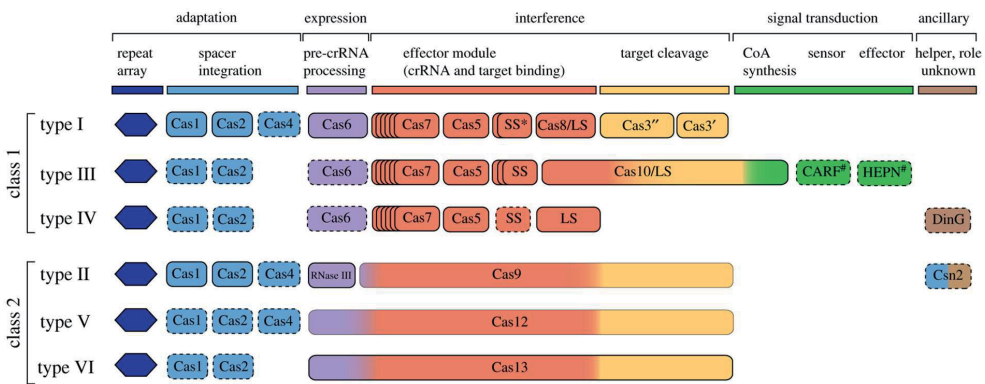


**Figure 4: Modular organization of CRISPR-Cas genes of class 1 and 2 systems from Koonin and Makarova 2019.**

A functional genomic CRISPR-Cas locus consists of a CRISPR array of eponymous identical repeats intercalated with foreign nucleic acid spacer sequences as well as an adjacently located *cas* operon encoding for the respective Cas proteins (Figure 5). Even though different CRISPR-Cas systems have evolved, in general, CRISPR immunity can be categorized in three

main steps: spacer acquisition, CRISPR RNA (crRNA) biogenesis, and interference (Wright *et al.*, 2016).
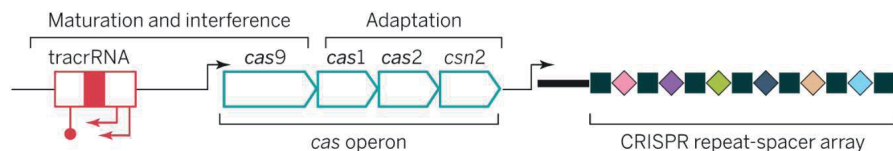


**Figure 5: Schematic illustration of a genomic CRISPR-Cas9 locus.** Adapted from Knott and Doudna 2018.

## 3.1 The process of the CRISPR-mediated, prokaryotic immune response

### 3.1.1 Spacer acquisition and integration

The adaptive immune response in bacterial cells is triggered upon intracellular identification of foreign DNA. In the adaptation or acquisition phase, new spacers derived from the foreign invading DNA, are incorporated into the genomic CRISPR array (Barrangou *et al.*, 2007). The process of spacer acquisition and integration into the CRISPR array is mediated by the Cas1-Cas2 complex, which is highly conserved between species and is present in all CRISPR-Cas types (Figure 6) (Nuñez *et al.*, 2015a; Yosef *et al.*, 2012).

The Cas1-Cas2 complex, which assembles in two Cas1 dimers and a single Cas2 dimer, comprises a catalytic activity similar to viral integrases and transposases (Nuñez *et al.*, 2015b, 2014; Rollie *et al.*, 2015).
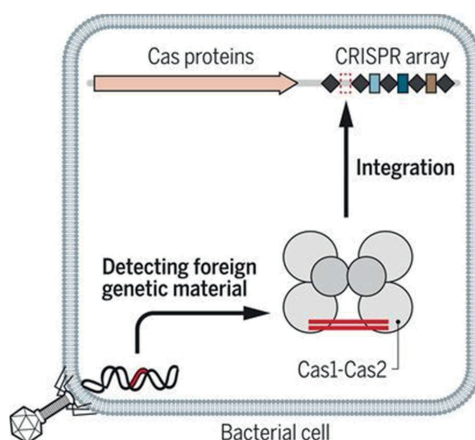


**Figure 6: Adaptation step in CRISPR-Cas immunity.** Foreign template DNA (red) is recognized by the Cas1-Cas2 complex and subsequently is integrated into the CRISPR array with different spacer fragments (colored squares) flanked by repeat sequences (black diamonds). Adapted from Knott and Doudna 2018.

From previous studies performed in *E. coli*, it is known that the Cas1–Cas2 complex requires a 33 nucleotide (nt) sequence for spacer acquisition *in vivo*, suggesting that protospacer length is predetermined before integration by a hitherto unknown mechanism (Nuñez *et al.*, 2015b). Prerequisite for an acquired protospacer sequence is the directly upstream situated 3-5 nt long PAM (protospacer adjacent motif) site (Deveau *et al.*, 2008; Mojica *et al.*, 2009; Shah *et al.*, 2013). In the later interference process, the PAM site initiates binding of the Cas-surveillance complex at the target site. Therefore, autoimmunity against the genomic CRISPR array is prevented by trimming of the PAM sequence before protospacer integration into the CRISPR locus (Kieper *et al.*, 2018; Shiimori *et al.*, 2018).

The processing of the protospacer sequence is followed by integration in the genomic CRISPR locus starting with an A/T rich- and poorly conserved leader sequence. It contains a promoter sequence, which controls the expression of the CRISPR array (Pul *et al.*, 2010). Downstream of the leader sequence, the CRISPR array encompasses a repeat-spacer architecture, in which spacer sequences derived from invading DNA are flanked by repetitive~20–50 base pair (bp) sequences termed repeats (Barrangou *et al.*, 2007; Jansen *et al.*, 2002).

Similar to protospacer acquisition, spacer integration into the CRISPR array is also catalyzed by the Cas1-Cas2 complex, with Cas1 mediating the integration reaction. The integration of the protospacer sequence is initiated by nicking of the DNA double strand at the first leader sequence proximal repeat by nucleophilic attack of the 3'-OH groups of the protospacer. Using this mechanism, the protospacer is integrated as spacer into the CRISPR array while duplicating the repeat sequence to maintain the repeat-spacer architecture (Ivančić-Bace *et al.*, 2015; Nuñez *et al.*, 2015a; Wright & Doudna, 2016).

### 3.1.2 CRISPR RNA biogenesis and CRISPR interference

In all CRISPR-Cas systems, biogenesis of crRNA is essential to confer the specific programming of the activated Cas interference protein-RNA complex resulting in recognition and degradation of invading DNA (Figure 7). The transcription of the CRISPR array is controlled by a promoter embedded in the leader sequence (Pul *et al.*, 2010). Expression of the CRISPR locus often results in a single long transcript, designated as pre-crRNA, which is subsequently processed to mature crRNA by cellular RNases or by affiliated Cas enzymes depending on the respective CRISPR system (Charpentier *et al.*, 2015; Hille *et al.*, 2018).
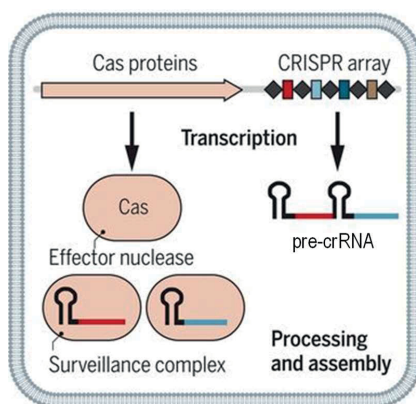


**Figure 7: Expression step in CRISPR-Cas immunity.** The CRISPR array and associated Cas proteins are expressed. The pre-crRNA is processed and Cas effector nucleases, such as Cas9 or Cas12a, associate with a mature sgRNA or crRNA to form the activated Cas surveillance complex. Adapted from Knott and Doudna 2018.

A further established CRISPR-Cas class 2 system is the recently identified type V CRISPR-associated protein Cas12a (Zetsche *et al.*, 2015). An advantageous characteristic of Cas12a over Cas9 is the fact that the Cas12a system does not require a trans-activating crRNA (tracrRNA) to form a single guide RNA (sgRNA) for DNA cooperation of crRNA (Figure 8 A). Cas12a crRNA assembles in a pseudoknot structure, a scaffold which allows for direct incorporation into the Cas protein. (Dong *et al.*, 2016; Yamano *et al.*, 2016). Whereas processing of the 5' end of the pre-crRNA is auto-catalyzed by the Cas12a protein (Fonfara *et al.*, 2016; Zetsche *et al.*, 2016), downstream trimming of the crRNA is mediated via an unknown housekeeping RNAse (Swarts, 2019). The specific target recognition of the activated Cas12a complex is mediated by a spacer-derived 24 nt sequence of the crRNA (Figure 8 B) (Yamano *et al.*, 2016).
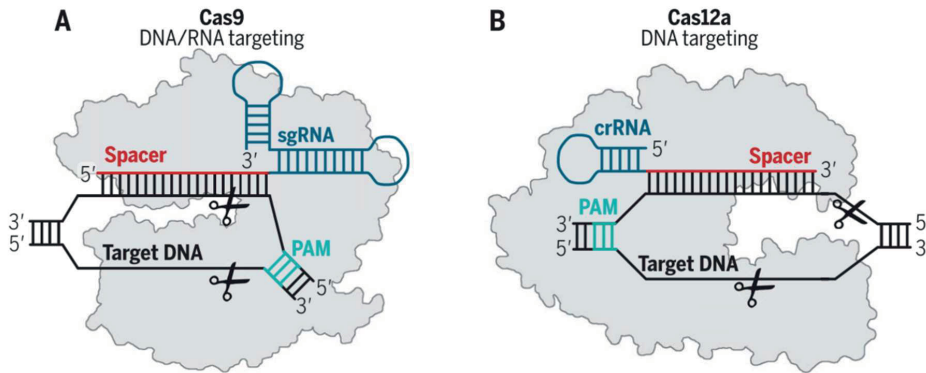
**Figure 8: Schematic illustration of the activated Cas surveillance complex. (A)** Class 2 type II CRISPR-Cas9 complex bound to sgRNA (blue) containing a spacer sequence (red) for specific target recognition. Spacer of sgRNA is hybridized to target dsDNA (black) proximal to a PAM (teal). Correct base-pairing activates nuclease activity to introduce a double strand break (scissors). **(B)** Class 2 type V CRISPR-Cas12a activated by crRNA (blue) encoding a spacer (red). The surveillance complex is bound to its complementary dsDNA target site (black) proximal to a PAM (teal). The nuclease activity is initiated upon correct base-pairing, which introduces cleavage of both strands (scissors). Adapted from Knott and Doudna 2018.

### 3.1.3   Cas interference mediated by activated Cas12a

During the interference state, the activated Cas surveillance complex ensures specific recognition of the target site to induce a conformational change in the Cas protein, which initiates its nuclease activity for DNA cleavage (Figure 9). Due to the recent application of the CRISPR/ Cas12a recombineering system in *C. glutamicum* and its relevance for this thesis*,* this section will focus on Cas12a interference (Jiang *et al.*, 2017).

Structural studies on Cas12a revealed a bilobed protein structure containing a recognition (REC) lobe and a nuclease (NUC) lobe (Dong *et al.*, 2016; Swarts *et al.*, 2017; Yamano *et al.*, 2016). The assembly of the surveillance complex is initiated via the recognition and processing of pre-crRNA (Swarts *et al.*, 2017). In the activated surveillance complex, DNA cleavage is exclusively restricted to the target site. In a DNA unbound state, the catalytic sites of the Cas12a-crRNA complex are not accessible for the DNA strand, and hence, the DNase activity is prohibited (Stella *et al.*, 2018).

Upon PAM site recognition of Cas12a, binding of the Cas12a surveillance complex to the target site is initiated (Gao *et al.*, 2016; Swarts *et al.*, 2017; Yamano *et al.*, 2016). The PAM site of Cas12a, a T-rich conserved sequence with 5'-(T)TTV-3' (V = A, C, or G), is located adjacent to the 5' end of the protospacer sequence (Yamano *et al.*, 2016; Zetsche *et al.*, 2015).
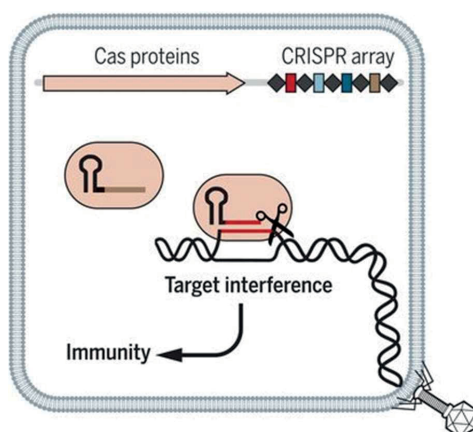
**Figure 9: Interference step in CRISPR-Cas-mediated immune response.** The activated Cas surveillance complex recognizes foreign genetic DNA complementary to its crRNA. Immunity is conferred by targeted interference, through the introduction of a double strand break (scissors). Adapted from Knott and Doudna 2018.

Base-pairing of the spacer-derived sequence in the crRNA to the complementary target DNA strand will propagate full binding of the surveillance complex to the target site resulting in DNA cleavage (Yamano *et al.*, 2016; Zetsche *et al.*, 2015). Cas12a performs *cis*-cleavage of both DNA strands resulting in a staggered double-stranded DNA break with 5' overhangs (Yamano *et al.*, 2016). Eventually, the double-strand cleavage is followed by the release of the PAM distal target DNA from the Cas surveillance complex. The catalytically activated Cas12a complex remains bound to the cleaved DNA target strand and assists in DNA degradation of the ssDNA (Chen *et al.*, 2018).

### 3.1.4   Versatile genome editing strategies using Cas12a and RecET

In genetic engineering, previously developed methods for specific DNA cleavage involved nucleases using protein–DNA recognition such as meganucleases, zink-finger proteins and Transcription Activator-Like Effector Nucleases (TALEN). However, a major draw-back of the afore mentioned nucleases is the impeded programming of DNA recognition to achieve specific sequence targeting, which makes the application laborious and time consuming (Adli, 2018). In contrast, CRISPR-Cas12a associated nucleases circumvent this hurdle by using a spacer-derived sequence in their crRNA, which can be exchanged to any sequence of interest allowing for an easy programming of the target sequence (Zetsche *et al.*, 2016).

CRISPR-Cas methods, using either Cas9 or Cas12a, have already been widely used in genome editing approaches in different bacterial species (Doudna & Charpentier, 2014; Yan *et al.*, 2017). Due to toxic effects of the Cas9 system and therefore prohibited genome editing, a switch to Cas12a-based systems broadened the application range for genome editing also

in different *Cyanobacteria* sp., *E. coli*, *Yersinia pestis*, *Mycobacterium smegmatis* as well as *C. glutamicum* (Ungerer & Pakrasi, 2016; Yan *et al.*, 2017; Zetsche *et al.*, 2016).

Previous to CRISPR-Cas mediated genetic engineering approaches in *C. glutamicum*, targeted genome engineering was also performed via RecET-mediated recombineering. This technique based on homologous recombination of single-stranded DNA oligonucleotides, enabled site-specific genome engineering in *C. glutamicum* (Binder *et al.*, 2013). The application in combination with transcriptional biosensors allowed for direct phenotype characterization of different recombinant variants circumventing the screening bottleneck (Binder *et al.*, 2013).

In contrast to RecET-mediated recombineering, the application of CRISPR/ Cas12a genome engineering enables genetic modification on a defined genomic target using sequence-specific nucleases (Zetsche *et al.*, 2016, 2015). In *C. glutamicum*, the application of an established plasmid-based CRISPR/ Cas12a genome editing system ensures the controlled expression of Cas12a and the crRNA targeting a specific sequence of interest (Jiang *et al.*, 2017). Upon, the assembly of the activated Cas12a-crRNA complex, a staggered double-strand break is introduced at a specific target site into the host genome (Zetsche et al. 2015; Jiang et al. 2017). The double strand break is restored by the DNA repair system using either non-homologous end joining (NHEJ) or homology-directed repair (HDR). Since NHEJ is more susceptible to random insertions, deletions and gene disruption at the target site, for precise gene editing, HDR is favored to insert single-stranded template DNA (Adli, 2018). Using single-stranded oligonucleotides, successful CRISPR/ Cas12a-mediated genome engineering was achieved by generating genomic point mutations, deletions and insertions in *C. glutamicum* (Jiang *et al.*, 2017). Due to CRISPR-mediated inactivation of wild-type cells, recombinant variants can easily be selected (Jiang *et al.*, 2013). Limitations for the application of the CRISPR/ Cas12a-mediated genome editing system are only given by the choice of functional PAM sites proximal to the target sequence. Furthermore, in *C. glutamicum,* CRISPR/ Cas12a-mediated recombineering was applied to perform genomic codon saturation mutagenesis *in situ* with NNK oligonucleotides for the respective target codon (Jiang *et al.*, 2017). Targeting of several different codons of interest in a multiplex approach, would allow for an easy generation of mutant libraries for the respective amino acid residues (Jiang *et al.*, 2017).

## 4 Sensor-based methods to characterize generated phenotypes

In targeted genome engineering using advanced methods such as CRISPR-Cas, genetic alterations can be easily and precisely introduced. At the same time, it has become quite apparent that even the simplest prokaryotic metabolism, such as *C. glutamicum*, is tremendously complex. This implies that even system-based approaches of rational engineering are constrained in the identification of novel genomic targets for strain improvement.

Counteracting this limitation of rational engineering approaches, classical strain engineering approaches, such as undirected whole genome mutagenesis, were recently rediscovered. In particular, genome-wide and random approaches allow for the identification of novel and unintuitive targets. Subsequently, these identified targets can be harnessed in rational engineering approaches. However, this approach is still limited, due to the restricted screening capacity for phenotype characterization, which allows only a certain proportion of the generated library to be analyzed (Eggeling *et al.*, 2015).

Since phenotype characterization of the created library represents the major bottleneck in the overall strain engineering process today, efficient high-throughput approaches are highly aspired to quickly identify strain variants with the desired phenotype. In cases where the product molecules are either chromogenic or fluorogenic, product formation can be directly linked to a machine-readable signal, which in turn allows for a straight-forward isolation of producing strain variants. However, most compounds of biotechnological interest do not confer such an easy distinguishable phenotype to the producing cells (Dietrich et al., 2010; Mahr and Frunzke, 2016; Zhang et al., 2015b). Here, genetically-encoded biosensors, which link the intracellular presence of a product of interest to a measureable optical output signal, represent valuable and universal screening tools (Dietrich *et al.*, 2010; Eggeling *et al.*, 2015; Mahr & Frunzke, 2016; Schallmey *et al.*, 2014). Based on this principle, three different types of biosensors were developed and have found an application in several areas of research.

### 4.1 RNA aptamer-based sensors

Aptamer biosensors are based on riboswitches consisting of RNA elements obtained from the untranslated regions of mRNA. Specific ligand binding of low-molecular metabolites to these aptamer structures induces a change in conformation, thereby regulating transcription or translation of its encoding gene (Figure 10) (Serganov & Nudler, 2013). Using this mechanism, RNA-based small-molecule biosensors have been constructed to sense a series of metabolites by introducing aptamer structures from nature into a synthetic circuit, which controls the expression of a reporter gene (Hallberg *et al.*, 2017). However, aptamer-based biosensors are confined in their chemical diversity. Exploring new aptamer architectures for biosensor design,

molecular evolution or computational design such as the Systematic Evolution of Ligands by Exponential enrichment (SELEX) protocol (Tuerk & Gold, 1990) can be used to design potential new metabolite sensing capabilities *de novo* (Ellington & Szostak, 1990). However, the *in vivo* application of synthetic *in vitro* or *in silico* designed RNA aptamer structures resulted only in a few functional riboswitches until now (Kopniczky *et al.*, 2015; Liang *et al.*, 2011; Schallmey *et al.*, 2014).



**Figure 10: Schematic illustration of an RNA aptamer-based biosensor.** The biosensor is shown in inactive (left) and activated state (right), which is initiated upon ligand binding (light blue). Abbreviation: ribosome binding site (RBS).

Nevertheless, in strain engineering approaches engineered and native aptamer structures were used as sensors responding to a range of different molecules of interest such as folinic acid, theophylline, caffeine and tetracycline (Berens *et al.*, 2001; Hanson *et al.*, 2005; Jenison *et al.*, 1994; Trausch *et al.*, 2011; Wachsmuth *et al.*, 2013). Application of a theophylline sensing synthetic aptamer sensor in *Saccharomyces cerevisiae*, when screening of a diverse caffeine demethylase library, yielded an improved enzyme variant with 33-fold increased enzymatic activity *in vivo* and 22-fold enhanced product selectivity (Michener & Smolke, 2012). In a similar screening approach, application of a flavin mononucleotide-responsive aptamer sensor enabled the identification of a *Bacillus subtilis* strain showing increased vitamin B2 production capabilities (Meyer *et al.*, 2015). Nevertheless, construction of aptamer-based biosensors is quite demanding, since there is only a limited functional sequence space for ligand binding of aptamer structures compared to the broad substrate range of protein activity-based biosensors such as FRET- and transcription factor-based biosensors (Liu *et al.*, 2015).

## 4.2 FRET-based biosensors

Sensors based on Förster resonance energy transfer (FRET) create a fluorescence response due to an energy transfer process amongst a pair of light-sensitive fluorophores. Interacting fluorophores, most commonly cyan- (CFP) and yellow fluorescent protein (YFP), are immobilized in close proximity (<10 nm) on a ligand-binding peptide/ protein and contain an overlapping excitation and emission spectrum between the FRET acceptor and donor, respectively (Hochreiter *et al.*, 2015). In a ligand-bound state the protein alters its conformation

causing a change in distance of donor and acceptor fluorophores, and hence, inducing a FRET change (Figure 11) (Bermejo *et al.*, 2011). In FRET-based metabolite sensors, the changed efficiency of energy transfer between the fluorescent proteins results in a measurable ratiometric index that allows the assessment of the present intracellular metabolite concentration (Constantinou & Polizzi, 2013; Frommer *et al.*, 2009). During the last years, numerous ligand-binding protein scaffolds found in nature were used as basis for the construction of a wide array of FRET-based biosensors being adopted to sense sugars phosphates, amino acids, carboxylic acids, metal ions, redox states and other intracellular reactions which would be challenging to monitor otherwise (Zhang *et al.*, 2015).
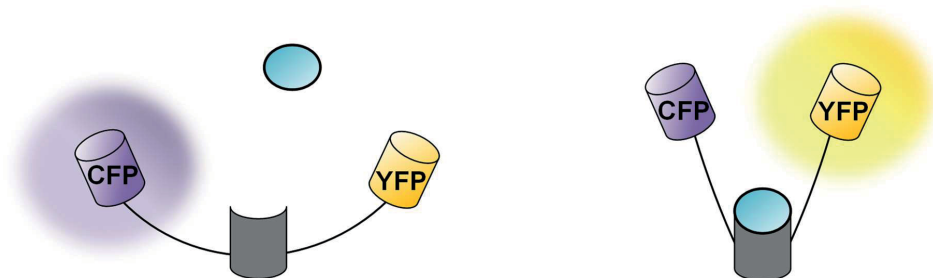


**Figure 11: Schematic illustration of a FRET-based biosensor.** The biosensor is shown in an inactive (left) and activated state (right), which is initiated upon ligand binding (light blue). Abbreviation: cyan- (CFP) and yellow fluorescent protein (YFP).

Due to the relatively low operational range between basal output signal and fully shifted fluorescence state, application of FRET-based biosensors in biotechnological processes is limited to online monitoring of product formation in solution (Zhang *et al.*, 2015). Noteworthy, changes of the fluorescence output signal can only be detected within a small window of concentration, typically in the nM to µM range. In direct comparison, transcriptional biosensors are more useful as they allow for an increased differentiation at higher inducer concentrations (Zhang *et al.*, 2015).

## 4.3 Transcription factor-based biosensors

A transcriptional biosensor consists of a transcription factor (TF) that selectively and specifically recognizes a molecule of interest (ligand) with a ligand-binding domain. Ligand binding causes a conformational change in the structure of the DNA-binding domain of the transcriptional activator or repressor. Inducing either regulator binding or release from its operator site in the cognate promoter sequence, initiation of RNA polymerase recruitment leads to the transcription of a reporter gene such as the yellow fluorescent protein (Figure 12). (De Paepe *et al.*, 2017; Fernandez-López *et al.*, 2015; Mahr & Frunzke, 2016). Based on their regulatory circuit, a broad range of ligand binding transcriptional regulators can be harnessed for the construction of biosensors. A prerequisite for the construction of a transcriptional

biosensor is the specific binding of the molecule of interest to the regulator protein. The identification and the selection of a matching promoter/regulator pair, which activates the native regulatory circuit upon ligand binding, can be simplified by using databases and scientific literature.
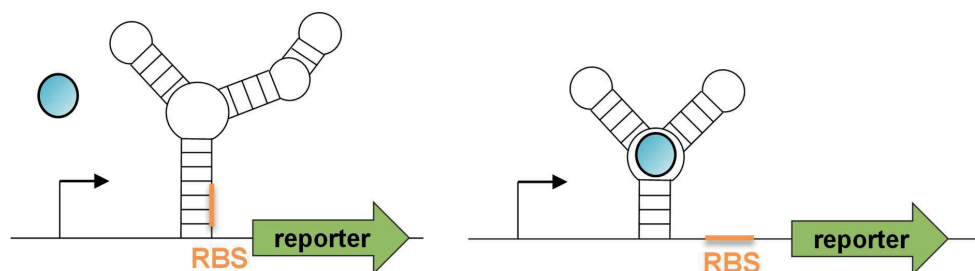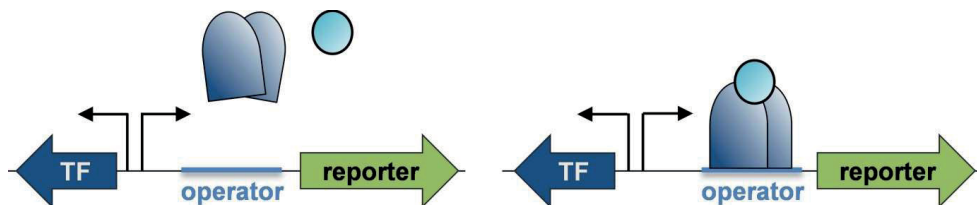


**Figure 12: Schematic illustration of a transcription factor-based biosensor.** The biosensor is shown in inactive (left) and activated state (right) which is initiated upon ligand binding (light blue). Abbreviation: transcription factor (TF).

Various biosensors based on native transcriptional regulators were shown to be functional and used for screening purposes in their native host organisms (Mahr & Frunzke, 2016). Initially, transcriptional biosensors found applications in the detection and monitoring of environmental pollutants, including heavy metal ions (Billinton *et al.*, 1998; de Lorenzo *et al.*, 1993; Ikariyama *et al.*, 1997; Sticher *et al.*, 1997). Furthermore, the regulatory circuit was adapted for dynamic feedback regulation of synthetic pathways in response to the intracellular concentration of pathway intermediates (Chou & Keasling, 2013; Xu *et al.*, 2014; Zhang *et al.*, 2016).

Most commonly, transcriptional biosensors are used in the context of strain engineering where vast and genetically diverse libraries have to be screened (Mahr & Frunzke, 2016). The application of transcriptional biosensors allowed for a rapid assessment of generated mutant libraries on the single-cell level. In combination with flow cytometry, high-throughput screening of diverse libraries by analyzing up to 80,000 single cells per second is possible.

Using elaborated screening campaigns, the identification of improved producer strains is facilitated based on the fluorescence response on the single-cell level (Flachbart *et al.*, 2019). Also for *C. glutamicum,* different transcriptional biosensors were constructed to investigate product formation in the platform organism (Eggeling *et al.*, 2015).

### 4.3.1   Transcriptional biosensors in *C. glutamicum* - industrial and basic research

In the past, *C. glutamicum* derived industrial producer strains for amino acid production have been thoroughly optimized by directed mutagenesis but also rational methods. Thus, it is difficult to further enhance product titers and yields in these evolved strains using current methods. Again, targeting the whole genome by using random mutagenesis is key to identify

unknown targets, which are involved in amino acid accumulation to further increase product formation even in highly engineered *C. glutamicum* strains.

In a chemically mutagenized *C. glutamicum* wild type library, different sensor-based screening approaches were applied to identify mutants with improved product formation characteristics. An Lrp-based (Leucine-responsive protein) biosensor, which detects branched-chain amino acids L-valine, L-leucine, L-isoleucine and L-methionine, enabled the identification of mutants that produced up to 8 mM L-valine, 1 mM of L-leucine and 1 mM L-isoleucine after 48 h cultivation in minimal medium (Mustafi *et al.*, 2012). In comparison to the isolated mutant, the wild type strain could not accumulate any of the respective amino acids (Mustafi *et al.*, 2012). Further increase in L-valine titers, due to the identification of an unintuitive target gene, was achieved with an Lrp biosensor-based screening of strains gained from adaptive laboratory evolution. The highest L-valine titer was obtained upon genetic engineering of the identified novel target gene encoding for urease accessory protein UreD. Cultivation of the constructed variant in defined medium, which carried the loss-of-function mutation *ureD*-E188*, reached 50.5 mM L-valine which is 4-5-fold increased compared to the original strain (Mahr *et al.*, 2015).

In a second biosensor-based screening approach, a LysG-based biosensor, responding to basic amino acids L-histidine, L-lysine or L-arginine, was used to isolate improved L-lysine producers from a chemically mutagenized *C. glutamicum* library (Binder *et al.*, 2012). Based on the isolated mutants which showed increased L-lysine synthesis compared to the wild type, the *murE* gene, encoding for UDP-N-acetylmuramyl-tripeptide synthetase, was identified as an attractive target to engineer L-lysine synthesis. In a subsequent LysG-based screening in combination with RecET-recombineering, the previously identified codon 81 of *murE* was targeted, since it was shown to enhance L-lysine formation (Binder *et al.*, 2012). Fluorescence-activated cell sorting (FACS) screening of the created recombineering library resulted in the isolation of eleven mutants with significantly improved L-lysine production titers (Binder *et al.*, 2013).

### 4.3.2   Limits of traditional sensor design and how to improve sensor parameters

However, the application of native transcriptional biosensors is restricted to the natural ligands of the used transcriptional regulators. In order to enable the application of transcriptional biosensors also for novel substrates, the substrate range of existing transcription factors can be broadened or shifted using the methods of directed protein evolution. The transcriptional regulator AraC with its natural ligand arabinose regulates expression of arabinose utilization genes from the arabinose-inducible *araBAD* promoter ($P_{BAD}$) (Soisson, 1997). In recent studies, the substrate specificity of transcriptional regulator AraC was successfully altered from arabinose to isoprenoid precursor isopentenyl diphosphate (IPP) by replacing the original

ligand-binding domain with a binding site of enzymes which naturally bind IPP (Chou & Keasling, 2013). Further engineering of AraC resulted in the acceptance of a broad range of different ligand molecules such as D-arabinose, mevalonate, ectoine, triacetic acid lactone, vanillin or salicylic acid (Chen *et al.*, 2015; Frei *et al.*, 2018; Tang *et al.*, 2013, 2008; Tang & Cirino, 2011).

When a genetic circuit (regulator and target promoter), either natural or synthetic, is turned into a biosensor, sensor characteristics such as range of inducer concentrations (operational range) and the range of sensor fluorescence response (dynamic range), are predetermined and cannot be adjusted individually for specific sensor applications (Figure 13).
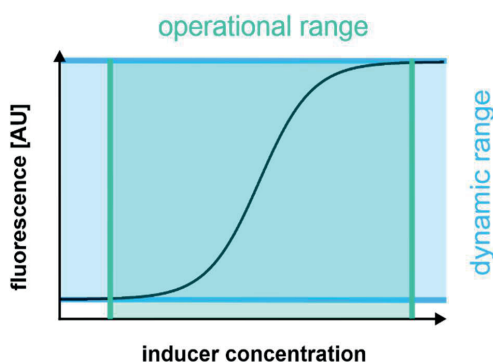


**Figure 13: Schematic illustration of the operational and dynamic range.**

The defined native dynamic and operational range determines the field of possible biosensor applications to either basic research, monitoring of environmental pollutions (low inducer concentrations), or strain engineering in an industrial setting (high inducer concentrations) (Schallmey *et al.*, 2014; Webster *et al.*, 2014). Hence, adaption of sensor parameters would give the opportunity to broaden the application range of these native biosensors.

### 4.3.3  Optimizing sensor response with regard to operational and dynamic range

With the aim to fine-tune different sensor parameters and therefore to overcome the native limitations of transcriptional biosensors, different stages in the genetic circuit can be modified to achieve this goal. With respect to ligand binding to the regulator protein, protein engineering can be applied to either lower or increase the binding affinity of the ligand towards its binding site. Furthermore, transcription factor binding or release from the operator site in the regulated promoter can be modified either by sequence optimization of the operator site or by engineering of the DNA binding domain in the transcriptional regulator (De Paepe *et al.*, 2017).

In a more advanced approach, the intracellular regulator level is fine-tuned to modulate the dynamic- and operational range of the transcriptional biosensor. The regulator concentration

can be modified by influencing the translation efficiency of the regulator *e.g.* by adapting the sequence of the ribosome binding site. Investigating the influence of RBS engineering on the sensor response of a natural LysR-type naringenin biosensor revealed a change in the sensor parameters depending of the RBS sequence and its respective strength (De Paepe *et al.*, 2018). Furthermore, modulation of the response of the transcriptional biosensor can be attained by gradual and constitutive expression of the gene of the transcriptional regulator (Lin *et al.*, 2018; Rogers *et al.*, 2015; Skjoedt *et al.*, 2016; Zhang *et al.*, 2017). In this context, utilization of different constitutive promoters allows for modulation of the expression level of the transcriptional regulator, and in this respect also a change in the sensor response.

Engineering these different targets in the genetic circuit of a transcriptional biosensor, allows the sensor dynamic and operational range to be adjusted for different applications. However, heterologous application of transcriptional biosensors can also be restricted by their exclusive functionality in the native microorganism from which the sensor components are derived (Mannan *et al.*, 2017). For example, the insertion of the regulatory circuit of the 3,4-dihydroxybenzoate biosensor based on the transcriptional regulator PcaU from *Acinetobater* sp. ADP1 in *E. coli* resulted in an absent or weak biosensor responses (Jha *et al.*, 2014).

Here, a unified biosensor design, which allows fine-tuning of important sensor parameters and restores the sensor response in a heterologous expression host would be beneficial to further broaden the application range of transcriptional biosensors in different host organisms.

## 5   Aims of thesis

The main goal of this thesis is the development and characterization of molecular tools to advance strain engineering in *C. glutamicum*. Since the introduction of genetic modification is a major part in strain engineering, the first aim of this thesis should be the establishment of a simplified protocol for CRISPR/ Cas12a-mediated recombineering with ssDNA oligonucleotides in *C. glutamicum*. The CRISPR efficiency of the novel system should be characterized and evaluated by the insertion of point mutations as well as the application of site-saturation mutagenesis in the genomic background of *C. glutamicum*.

Counteracting this refined method, classical strain engineering methods, such as undirected whole genome mutagenesis, are now undergoing a renaissance. Resolving the huge bottleneck of restricted screening capacity for phenotype characterization in this approach, transcriptional biosensors have been widely applied for the characterization of a diverse mutant library. Aiming for facilitated sensor construction, fine-tuning of sensor parameters and transfer of a biosensor to a heterologous expression host, a unified sensor design should be developed. The constructed biosensors should be characterized in liquid cultivations and on single-cell level to assess the applicability of the constructed sensors in future HTS campaigns.

# 6 Key results

Most results presented in the chapters 6.1 and 6.2 are described in detail in chapter 8.1 and 8.2.

## 6.1 Advancing CRISPR/ Cas12a application in *C. glutamicum*

**(Krumbach, K.; Sonntag, C. K. *et al.* 2019, ACS Synthetic Biology*, cf. chapter 8.1)**

### 6.1.1 Construction of a crRNA delivery vector – simplifying plasmid design and assembly

Regarding the CRISPR/ Cas12a system established in *C. glutamicum* (Jiang *et al.*, 2017), the standard protocol for crRNA delivery vector assembly often showed undesired double integration of the crRNA-encoding inserts. Therefore, an improved crRNA delivery vector design was constructed to facilitate and, at the same time, accelerate correct plasmid assembly. Vector pJYS2_crtYf was used as a template to construct the crRNA delivery vector pJYScr, which comprises a *Bam*HI/*BstB*I cloning site allowing for easy insertion of different spacer sequences into the linearized delivery vector. The introduction of the *Bam*HI/*BstB*I cloning site into pJYS2_crtYf was achieved by a synthetic DNA fragment, which contains the parental sequence but replaces the present *Dra*I with a *BstB*I restriction site. The cloning site was made available for insertion of spacer sequences by elimination of a second *BstB*I restriction site in the pJYS2 vector backbone using site-directed mutagenesis. For the construction of individual crRNA delivery vectors, the linearized pJYScr backbone can be ligated with synthetic double-stranded DNA inserts encoding the loop region and spacer of the crRNA.

### 6.1.2 Evaluation of CRISPR/ Cas12a mediated genome editing efficiency using the novel pJYScr crRNA delivery vector in *C. glutamicum*

The efficiency of CRISPR/ Cas12a-mediated ssDNA-recombineering with the novel crRNA delivery vector pJYScr was investigated with respect to the elimination of non-recombinant cells. For evaluating successful genetic engineering in *C. glutamicum* as part of these experiments, the *lacZ* gene encoding for β-galactosidase from *E. coli* was selected, as it enables blue:white screening on 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) agar plates. Therefore, the strain *C. glutamicum* WT::*lacZ* carrying a chromosomal integration of the *lacZ* gene was constructed. In this strain, expression of *lacZ* was under control of the strong synthetic promoter pH36 (Yim *et al.*, 2013). Constitutive Cas12a- and RecT-activity in *C. glutamicum* WT::*lacZ* was ensured by episomal expression from pJYS1-petFU (Jiang *et al.*, 2017).

In *C. glutamicum,* high transformation efficiency of pJYScr is essential for the functionality of CRISPR/ Cas12a recombineering. Therefore, the transformation rate of the empty pJYScr-vector with or without the respective recombineering oligonucleotide was compared to the established cloning vector pEKEx3 as a control. Using the same vector quantity of both plasmids, comparable results of ~$10^6$ colony forming units (cfu)/ transformation for both, the pEKEx3 control and pJYScr with and without oligonucleotides, were obtained. Hence, sufficient uptake of the delivery vector is ensured to perform CRISPR/ Cas12a-based recombineering. However, when using the advanced CRISPR system, it is important to note that in the protocol for electrocompetent cell preparation, cell harvesting must be performed at an optical density $OD_{600}$ below 0.6 in contrast to the standard protocol of Jiang *et al*. preparing competent cells at $OD_{600}$ below 1.0.

Assembly of the activated Cas12a-crRNA complex, introduction of a double-strand DNA break and site specific genome editing in the chromosomally integrated *lacZ* gene was investigated by using the constructed crRNA delivery vector pJYScr-lacZ-1574 and its respective recombineering oligonucleotide. The vector encodes for a crRNA with a spacer sequence of 21 nts adjacent to an intragenic PAM site 5´-(T)TT(C)-3´ in *lacZ*. Transformation of *C. glutamicum* WT::*lacZ* pJYS1-petFU with pJYScr-lacZ-1574 yielded 51±21 cfu, demonstrating the efficiency of Cas12a-mediated killing in *C. glutamicum*. Comparable results were obtained for a second targeted PAM site using pJYScr-lacZ-1654. The assessment of specific genome editing was performed with 59 nt recombineering oligonucleotides, which were designed to achieve PAM site-inactivation by introduction of a stop codon. Successful recombineering of the aforementioned PAM sites in the *lacZ* gene with oligonucleotides o-1574a and o-1654a resulted in an increased cell recovery and with ≥96% of the mutants carrying the desired genomic modification.

### 6.1.3   Site specific nucleotide targeting of PAM site proximal residues allows for reliable genetic engineering with CRISPR/ Cas12a

Further investigation of CRISPR/ Cas12a-mediated ssDNA recombineering efficiency was performed on the accessibility of nucleotides for genetic engineering located proximal or distal from the targeted PAM site in the *lacZ* gene. Nine different recombineering oligonucleotides were designed to introduce point mutations to generate a stop codon over a stretch of 11 nts up- and 16 nts downstream of the targeted PAM site. While inserting point mutations in the target site simultaneously, the respective PAM site was inactivated (Figure 14). Generated mutants using CRISPR/ Cas12a-assisted ssDNA recombineering were assessed by blue:white screening on X-Gal agar plates and subsequent sequencing of the *lacZ* gene (n=4-16) (Figure 14).

Initially, investigation of target sites located up-stream and distal to the targeted PAM site (oligonucleotide o-1574b), was performed. The results showed that the introduction of point mutations could not be achieved. However, PAM site inactivation could be attained in these mutants. Successful introduction of a single point mutation in a target site located proximal and upstream of the PAM site with simultaneous PAM site inactivation could be confirmed for more than 96% of the surviving *C. glutamicum* mutants. Equally efficient was the introduction of a 3-bp mutation proximal to the PAM site in the crRNA-targeting region with a PAM site inactivation in parallel.

| white colonies [%] | 0 | 96 | 98 | 99 | 98 | 92 | 56 | 86 | 5 |
|---|---|---|---|---|---|---|---|---|---|
| mutation type [count] | 0 8 0 | 4 0 0 | 8 0 0 | 16 0 0 | 15 0 0 | 11 0 0 | 10 0 2 | 8 0 0 | 6 1 1 |

Legend: PAM+stop [%], PAM only [%], stop only [%]

```
                                        -3 -2 -1
5`-C C A T C A A A A A A T G G C T T T C G C T A C C T G G A G A G A C G C G G C C G C T -3`
3`-G G T A G T T T T T T A C C G A A A G C G A T G G A C C T C T C T G C G C C G G C G A -5´
```
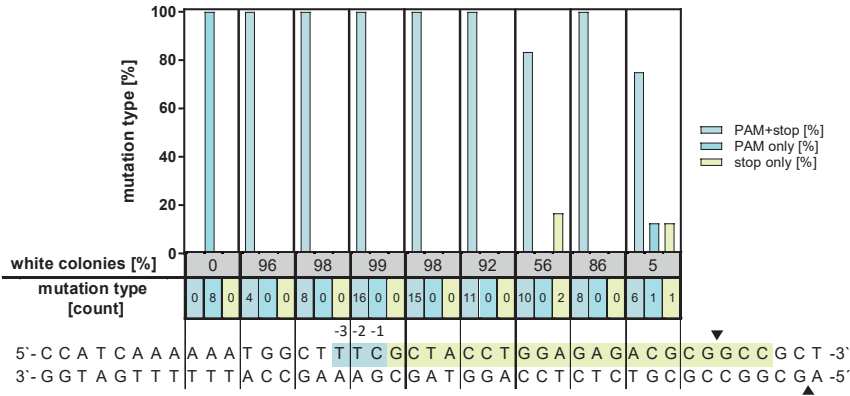
**Figure 14: Genome editing efficiency relative to the PAM site in *C. glutamicum* using the combined recombineering-CRISPR/ Cas12a selection method.** Shown is a sequence fragment of the chromosomally integrated *lacZ* gene with a PAM (blue) and the respective protospacer (yellow) sequence. The Cas12a-cleavage site yielding a staggered cut is indicated by black triangles. The resulting proportion of white colonies is specified as a share in percent. After each assay, the sequence of the *lacZ* gene of randomly selected white clones was assessed (n=4-16). Investigating the genome editing efficiency, results showed either that both the PAM site was mutated and the stop codon was introduced (dark blue bars), only the PAM site was mutated (light blue bars), or solely the stop codon was introduced (yellow bars). Additionally, for each type of mutation (PAM+stop/ stop only/ PAM only), the count of samples assessed via DNA sequencing are specified below the colored bars.

However, introduction of 1 to 3 bp mutations located in a distal and downstream situated sequence of the PAM site resulted in a reduced mutation efficiency. Depending on the required number of substitutions to generate a stop codon, mutation efficiency reached 75% to 90%. Interestingly, in certain cases DNA sequencing revealed that exclusively the stop codon was successfully inserted whereas the PAM site was still intact. In summary, the applied CRISPR/ Cas12a-mediated ssDNA recombineering allows for reliable site-directed genome engineering in a stretch of four nucleotides upstream and ten nucleotides downstream of the respective PAM site.

In previous studies, it has been shown that CRISPR-assisted ssDNA recombineering also allows to perform site-saturation mutagenesis in chromosomal DNA of *C. glutamicum* to create a diverse library for a codon of interest (Jiang *et al.*, 2017). With regard to the successful introduction of a 3-bp mutation in the PAM-site proximal crRNA-targeting sequence, oligonucleotides were designed to perform site-saturation mutagenesis at the respective target codon, while introducing a stop codon in the targeted PAM site. Sequence analysis of the generated mutants (white colonies) revealed that out of the 24 assessed variants 16 and 13, for two distinct target sites respectively, carried mutations in each codon. The resulting amino acid substitutions on protein level were unbiased for a specific chemical property. In a further approach, site saturation in an 18 nt sequence was performed containing the previously targeted PAM (*lacZ*-1574) site plus five nucleotides upstream and ten nucleotides downstream of the respective PAM site. However, only single nucleotides were exchanged. Therefore, this method cannot be used to exchange a longer stretch of DNA and is currently limited to the introduction of three consecutive nucleotide substitutions into the *C. glutamicum* genome.

### 6.1.4   Site-directed mutagenesis targeting *mscCG* in *C. glutamicum*

MscCG, a mechanosensitive channel of *C. glutamicum*, is the major efflux channel for L-glutamate of this bacterium (Hashimoto *et al.*, 2012; Nakamura *et al.*, 2007). Upon increased membrane tension, the export of L-glutamate mediated by MscCG in *C. glutamicum* is induced (Hirasawa & Wachi, 2016). However, also several gain-of-function (GOF) variants of MscCG were identified, which enable constitutive L-glutamate efflux (Nakamura *et al.*, 2007; Yamashita *et al.*, 2013). Amongst others, two GOF mutants were characterized with mutations in the C-terminal extracytoplasmic domain (Nakamura *et al.*, 2007; Yamashita *et al.*, 2013). Especially, the transition of the fourth transmembrane helix to the periplasmic space of the channel structure seems to have a major influence on the L-glutamate efflux mechanism (Nakamura *et al.*, 2007). Presumably, this site determines the overall structure of MscCG and is crucial for the interaction of the channel protein with other components of the cell wall. Therefore, the effect of an amino acid exchange in this particular region at residues V422 and E423 of MscCG was investigated.

Applying site-saturation mutagenesis to generate a diverse library of both target positions in MscCG, *C. glutamicum* pJYS1petFU was transformed using crRNA delivery vector pJYScr_mscCG1269c and the recombineering oligonucleotides o-mscCG1269-V422nnn and o-mscCG1269-E423nnn. From the generated library different variants were randomly selected and sequenced. In both assessed targets, unbiased amino acid substitutions were achieved. Recombinant strains were cured from plasmids and subsequently characterized for L-glutamate accumulation in cultivation experiments.

L-glutamate production in the constructed mutant strains was evaluated in defined medium with biotin excess (50 µg/L) in combination with low glucose concentration (40 g/L). In the used cultivation medium, natural accumulation of L-glutamate does not occur in the *C. glutamicum* WT as for MscCG variants V422A, V422C, V422T, V422L, and V422S (Figure 15A). In contrast, a considerable L-glutamate accumulation of up to 13 mM was detected in the supernatant of the MscCG variants V422K and V422D, in which the native valine residue was replaced with a charged amino acid (Figure 15A). Mutants E423K and E423D did not accumulate L-glutamate, nor did variants E423N, E423T, E423A and E423H (Figure 15B). Contrary, MscCG mutants E423P, E423S as well as E423stop stimulated the L-glutamate export in *C. glutamicum*. Therefore, the substitution by charged amino acid residues at position V422 and the exchange at position E423 with amino acid residues proline and serine yielded an open channel conformation. In cultivation experiments, a decrease in L-glutamate titers over time due to reuptake and utilization of this amino acid as known from previous studies was observed (Becker *et al.*, 2013).
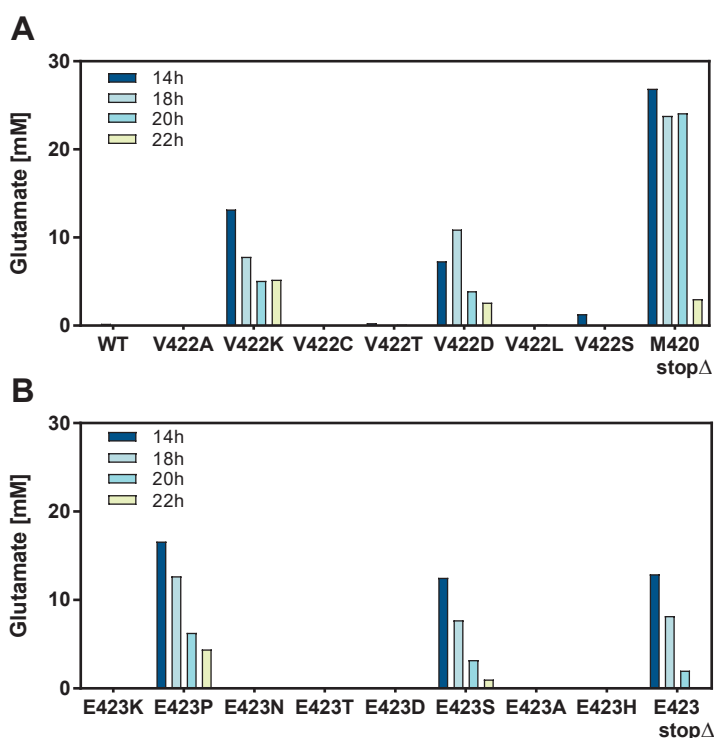


**Figure 15: L-Glutamate titers of *C. glutamicum* variants of the mechanosensitive channel MscCG.** Different strains carrying mutations in the *mscCG* gene at position (A) V422 and (B) E423 as well as variants with stop codons at positions M420 (M420stopΔ) and E423 (E423stop), were characterized to assess the accumulation of L-glutamate. L-glutamate concentrations for each variant were measured after 14, 18, 20 and 22 hours of cultivation.

With regard to results of recent studies, the carboxy-terminal domain of MscCG plays a major role in regulating the L-glutamate channel activity (Yamashita *et al.*, 2013). In order to gain a deeper understanding of the influence of the carboxy-terminal domain on the L-glutamate efflux mechanism, two mutants were generated in which different sequences of the extra-cytoplasmic domain of MscCG were deleted. Construction of these mutants was performed with the well-established SacB system (Niebisch & Bott, 2001; Schäfer *et al.*, 1994). In both generated mutants M420stopΔ and P424stopΔ, the respective codon was replaced by a stop codon while the sequence encoding for the remaining carboxy-terminal end of the channel structure was deleted. Cultivation experiments assessing L-glutamate accumulation in both mutants revealed L-glutamate concentrations up to 27 mM and 130 mM for M420stopΔ and P424stopΔ, respectively. These observations are in line with previous data sets emphasizing the importance of this protein domain for the regulation of the channel activity.

## 6.2   A unified design allows fine-tuning of biosensor parameters and application across bacterial species

**(Sonntag, C. K. *et al.* 2020, Metabolic Engineering Communications, cf. chapter 8.2)**

### 6.2.1   Establishing a unified biosensor design to control important sensor parameters

A unified sensor design was established to enable the gradual and constitutive expression of a transcriptional regulator for influencing the dynamic- and operational range of transcriptional biosensors. The sensor architecture of the sensor design can be separated into a sensing and a reporting module. In the sensing module, the expression of the transcriptional regulator is regulated by a constitutive promoter selected from a synthetic promoter library. In the reporting module, the expression of the reporter gene *eyfp* is controlled by inducer binding to a transcriptional regulator (Figure 16).

Since the unified sensor design was evaluated in *E. coli* and *C. glutamicum*, the well-characterized PLTetO1 and the *dapA* promoter libraries were selected for sensor plasmid construction, respectively (Alper *et al.*, 2006; Vašicová *et al.*, 1999). From both promoter libraries, strong (S), moderate (M) and weak (W) constitutive promoter variants were selected. Assuring full functionality of the regulated promoter in the sensing module, the first 45 nucleotides of the open reading frame of the regulated gene in the original genetic circuit were always included to ensure stability in mRNA folding, which provides efficient translation initiation (Figure 16) (Kudla *et al.*, 2009). Downstream of the extended regulated promoter sequence, a stop codon and an additional RBS (AAGGAGG-N$_{6-7}$) were inserted upstream of the reporter gene start codon (Figure 16). In addition, the often divergently orientated promoter

architecture of a cognate promoter can cause an undesired transcriptional read-through leading to uncontrolled expression of the regulator gene (Maddocks & Oyston, 2008).
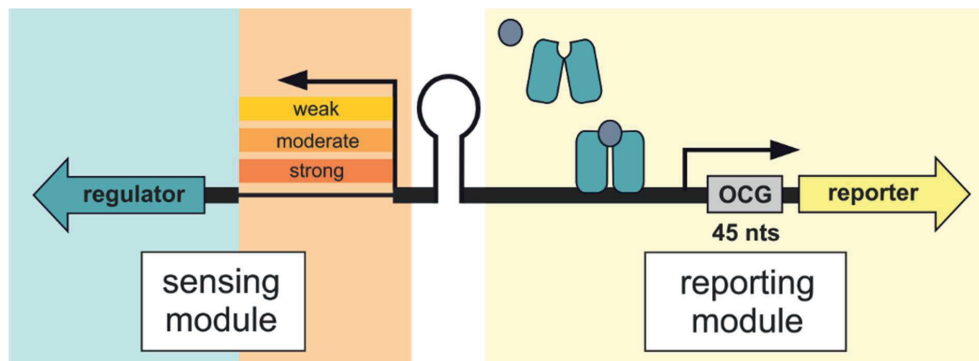


**Figure 16: Schematic illustration of the unified biosensor design.** Expression of the transcriptional regulator (blue arrow) is controlled by a native or synthetic constitutive promoter of the host organism. Upon binding of an inducer molecule, the transcriptional regulator undergoes a conformational change and either binds to or releases the operator site in the regulated promoter including the first 45 nts of the originally controlled gene (OCG) to promote expression of a reporter gene (yellow arrow).

These undesired effects can be reduced or prevented by insertion of a terminator sequence between the sensing and the reporting module (De Paepe *et al.*, 2018; Rogers *et al.*, 2015). Finally, both modules were integrated in divergent orientation in medium-copy vector backbones: pJC1 in case of *C. glutamicum*, and pBR322 for *E. coli*.

### 6.2.2   Construction of PhdR-based transcriptional biosensors for *C. glutamicum*

Initially, the native genetic circuit of repressor PhdR and its target promoter $P_{phdB}$ was used to construct an array of transcriptional biosensors. PhdR originates from *C. glutamicum*, where it naturally regulates the transcription of genes involved in phenylpropanoid degradation (Kallscheuer *et al.*, 2016a). Diffusion of PhdR from its cognate promoter $P_{phdB}$ is specifically induced upon binding of coenzyme A (CoA)-activated and ring-hydroxylated phenylpropanoids such as *p*-coumaroyl-CoA, which finally induces transcription of the *phd* operon (Kallscheuer *et al.*, 2016a).

Serving as a reference for the transcriptional biosensors built according to the unified sensor design, the native regulatory circuit of PhdR was used for the construction of the sensor plasmid pSenPhdR, which included the sequence of *phdR* and the region covering $P_{phdB}$ and 45 nucleotides of the *phdB* coding sequence. Furthermore, PhdR-based biosensors were constructed according to the unified sensor design, expressing the regulator gene *phdR* under control of either a weak (W), moderate (M) or strong (S) constitutive promoter (pSC$_{Cg}$-PhdR-W, pSC$_{Cg}$-PhdR-M, pSC$_{Cg}$-PhdR-S). *C. glutamicum* DelAro$^4$-*4cl*$_{Pc}$ $\Delta phdR$ was selected as host strain, as it is deficient in phenylpropanoid degradation and additionally harbors an IPTG-

inducible plant-derived *4cl* gene encoding a 4-coumarate: CoA-ligase necessary for CoA-activation of phenylpropanoids in this host organism (Kallscheuer *et al.*, 2016b). A stringent control of *phdR* expression is beneficial for the sensor response. Therefore, sole episomal expression of *phdR* was assured by chromosomal deletion of the repressor gene.

### 6.2.3   Fluorescence response of PhdR-based biosensors differs in dependency of inducer concentration and regulator expression level

The influence of inducer concentration and repressor level on the fluorescence response of PhdR-based transcriptional biosensors was evaluated in dose-response experiments in liquid culture using inducer concentrations ranging from 0 - 4,000 µM externally applied *p*-coumaric acid. The evaluation of the results revealed that an increasing expression level of the repressor gene correlates negatively with the dynamic range of the PhdR-based biosensors constructed according to the unified sensor design. However, the highest dynamic range (96-fold) was determined for pSenPhdR, which is based on the native regulatory circuit of the transcriptional regulator (Figure 17A). The comparison of the different sensor variants, with respect to the operational range, revealed distinct differences among the single constructs (Figure 17A). Due to an increased basal fluorescence response of pSC$_{Cg}$-PhdR-S, an overall lower dynamic range could be recorded for this sensor construct. The broadest operational range was covered by pSC$_{Cg}$-PhdR-M with moderate regulator gene expression. However, at low inducer concentrations, the biosensor pSC$_{Cg}$-PhdR-W showed an increased fluorescence response compared to all other PhdR-based sensor variants. This enhanced biosensor sensitivity allows for sensor applications at low ligand concentrations.

Since transcriptional biosensors, in the best case, are applied in FACS for high-throughput screening (HTS) campaigns, the dynamic- and operational range of the constructed sensors was also assessed on the single-cell level. In this context, a homogeneous fluorescence response of a cell population with an identical genetic background to a certain inducer concentration is highly desired. Furthermore, with increasing inducer concentration, a homogeneous shift of the cell population to a higher level of fluorescence intensity is anticipated. Hence, in an actual FACS-based screening campaign of a diverse mutant library, the sensor response allows to identify the desired phenotype based on the increased fluorescence signal correlating with increased product formation.

**A**



**B**



**Figure 17: Sensor response of constructed PhdR-based biosensors in *C. glutamicum.*** (A) Dose-response curves to assess the operational and dynamic range of PhdR-based sensor constructs based on the native regulatory circuit (pSenPhdR) and on the unified sensor design (pSC$_{Cg}$-PhdR-S/M/W). The fluorescence response was assessed in different BioLector cultivations supplemented with seven different p-coumaric acid (cou) concentrations ranging from 4 - 4,000 µM. The fluorescence response of each sensor construct was plotted as fold-change in specific eYFP fluorescence. Error bars represent standard deviations of three biological replicates. (B) Fluorescence response on a single-cell level with *C. glutamicum* strains carrying pSC$_{Cg}$-PhdR-S/M/W or pSenPhdR ranging from 16 - 4,000 µM p-coumaric acid. In each experiment, 95,000 representative single cells were analyzed.

The assessment of the fluorescence response of all PhdR-based biosensors on the single-cell level, except for pSC$_{Cg}$-PhdR-S, revealed a heterogeneous biosensor response for single

ligand concentrations. The respective histogram plots of the cell-count against fluorescence intensity showed two separate fluorescent populations with two global maxima (Figure 17B). However, depending on inducer concentration and expression level of the transcriptional repressor, a full shift of the whole cell population from low to high fluorescence intensity could be achieved.

Detailed analysis of the different constructs revealed that the lower the expression level of the repressor gene, the lower is the *p*-coumaric acid concentration required to homogeneously shift the fluorescence of the whole cell population. However, due to the increased basal fluorescence of sensor construct pSC$_{Cg}$-PhdR-S only a slight but homogeneous shift of the cell population to a higher fluorescence level in the tested inducer concentration range was assessed.

### 6.2.4   Ferulic acid and caffeic acid trigger a weaker biosensor response

In addition to *p*-coumaroyl-CoA, PhdR also accepts other CoA-activated ring-hydroxylated phenylpropanoids, like ferulic acid and caffeic acid, as inducer molecules (Kallscheuer *et al.*, 2016a). Thus, these compounds were also applied to the constructed PhdR-based biosensors to understand whether different inducers trigger a different fluorescence response. Sensor constructs pSC$_{Cg}$-PhdR-W and pSenPhdR were selected for sensor characterization using both phenylpropanoids as inducers to perform 48-well platform cultivations and subsequent flow cytometry analysis.

The fluorescence response of both biosensors showed that induction with ferulic acid and caffeic acid compared to *p*-coumaric acid resulted in a reduced dynamic range by about 30%. However, when comparing both biosensors directly, pSC$_{Cg}$-PhdR-W showed an overall lower operational and dynamic range in response to ferulic acid and caffeic acid in comparison to pSenPhdR. Nevertheless, on the single-cell level the fluorescence response of both PhdR-based biosensors resulted in a shift of the respective population from low to high fluorescence intensity with increasing inducer concentrations similar to inducer *p*-coumaric acid.

### 6.2.5   Construction of LysG-based biosensors according to the unified sensor design and subsequent characterization

In addition to the set of repressor-based biosensors, the unified sensor design was further evaluated by the construction of biosensors using a transcriptional regulator with a different mode of action. In search of possible candidates in *C. glutamicum*, the well-studied transcriptional activator LysG was considered. Naturally, a conformational change of LysG is induced upon binding of any of the three basic amino acids L-lysine, L-arginine and L-histidine. Upon activator binding to the operator site, expression of the transporter gene *lysE* is induced (Bellman *et al.*, 2001). In previous applications of biosensor pSenLysG, based on the native

regulatory circuit of LysG, novel targets to enhance L-lysine production in *C. glutamicum* were identified (Binder *et al.*, 2012). Therefore, the well-characterized genetic circuit of the LTTR-type transcriptional activator LysG was selected to verify the applicability of the unified sensor design (Bellman *et al.*, 2001; Binder *et al.*, 2012). Several LysG-based biosensors according to the unified sensor design were constructed to evaluate the influence of different regulator expression levels on the fluorescence response of transcriptional biosensors (pSC$_{Cg}$-LysG-W, pSC$_{Cg}$-LysG-M, pSC$_{Cg}$-LysG-S). All constructed sensor plasmids were characterized in a *C. glutamicum* Δ*lysEG* strain background to prevent interfering regulator and/or exporter gene expression from the genome.

First, the fluorescence response of the constructed biosensors was assessed in 48-well plate cultivations using amino acid dipeptides His-Ala, Lys-Ala and Arg-Ala at different concentrations ranging from 0 - 10,000 μM. In order to evaluate the fluorescence response of the different biosensor constructs compared to the native regulatory circuit, pSenLysG was used as a reference. For all constructed biosensors, the fold-induction in specific eYFP fluorescence was plotted against the respective inducer concentration (Figure 18A).

When comparing the generated fluorescence response of the different sensor constructs in liquid cultures, results revealed that the dynamic range increases with an increasing expression level of the regulator gene *lysG* (Figure 18A). Regarding the operational range of the respective sensor constructs, with increasing regulator gene expression an enhanced sensor response at low inducer concentrations was assessed. Proceeding from these findings, a shift of the operational range to higher concentrations was found for constructs with a low regulator gene expression level (pSC$_{Cg}$-LysG-W). Corresponding to the determined operational range from liquid cultivations (Figure 18B), an increase and gradual shift in fluorescence intensity from a lower to a higher state could be followed in flow cytometry measurements for all sensor variants with increasing inducer concentration.
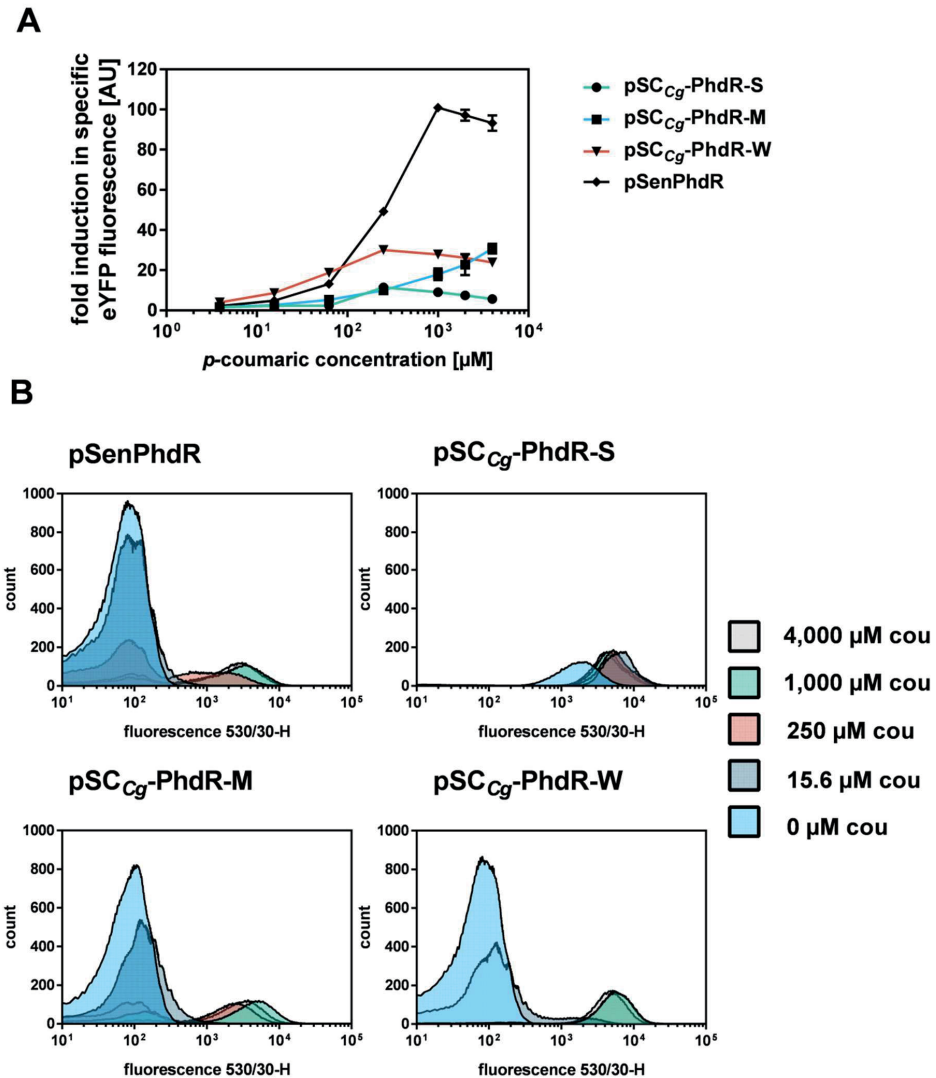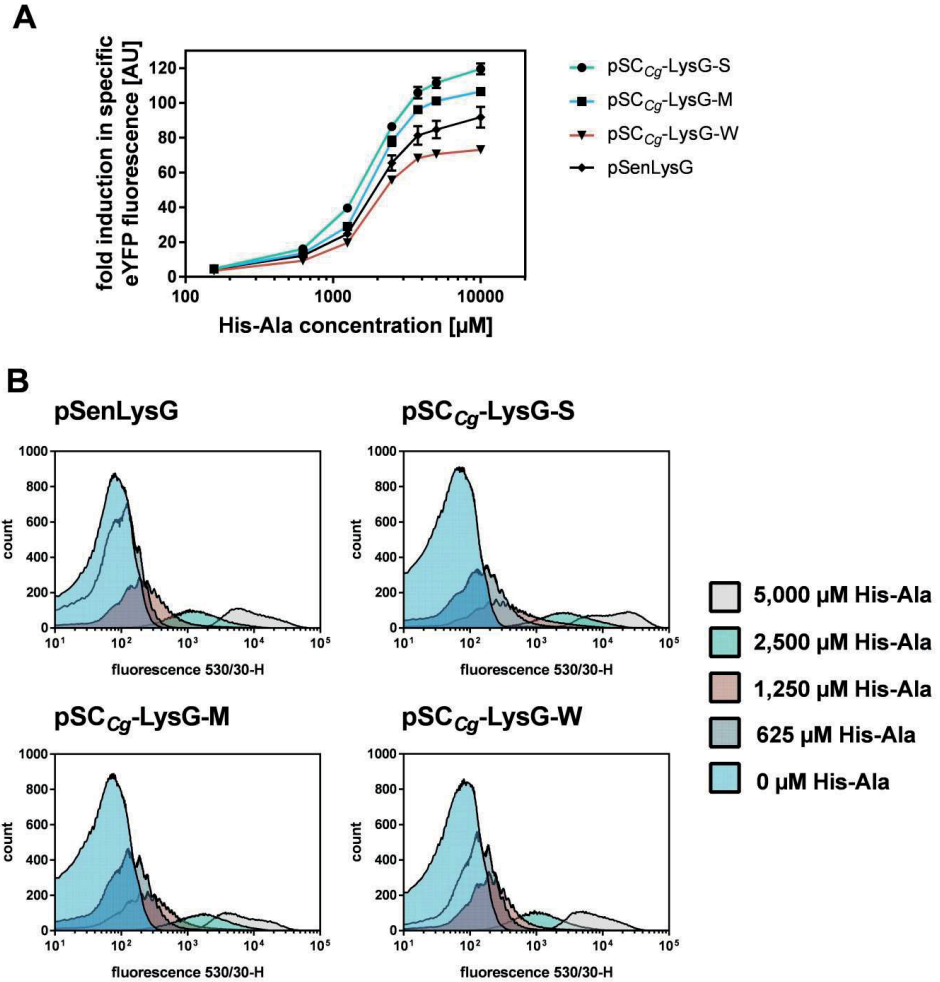
**Figure 18: Sensor response of constructed LysG-based biosensors in *C. glutamicum.*** (A) Dose-response curves to assess the operational and dynamic range of LysG-based sensor constructs based on the native regulatory circuit (pSenLysG) and on the unified sensor design (pSC*Cg*-LysG-S/M/W). The fluorescence response was investigated in different BioLector cultivations supplemented with seven different inducer concentrations ranging from 125 - 10,000 µM dipeptide His-Ala. The fluorescence response of each sensor construct was plotted as fold-change in specific eYFP fluorescence. Error bars represent standard deviations of three biological replicates. (B) Fluorescence response on single-cell level with *C. glutamicum* strains carrying pSC*Cg*-LysG-S/M/W or pSenLysG was performed with externally supplemented His-Ala dipeptides ranging from 0 - 5,000 µM. In each experiment, 95,000 representative single cells were analyzed.

### 6.2.6   Ligand spectrum of LysG-based biosensors

Since LysG also accepts basic amino acids L-arginine and L-lysine as inducers, the respective dipeptides were used to evaluate the ligand specificity of the transcriptional activator. Previously performed binding studies of LysG showed that the transcriptional regulator

interacts with all basic amino acids L-histidine-, L-lysine- and L-arginine with different affinities corresponding to the dissociation constants ($K_D$) (L-His $16 \pm 1.1$ $10^{-6}$ M/ L-Lys $3.29 \pm 0.62$ $10^{-3}$ M/ L-Arg $1.15 \pm 0.06$ $10^{-3}$ M) (personal communication, Hugo van Beek). With regard to the highest dynamic- and operational range from dose-response experiments of pSC$_{Cg}$-LysG-S with inducer His-Ala, this biosensor and pSenLysG were used to characterize the biosensor performance in presence of Lys-Ala and Arg-Ala in 48-well plate cultivations and in flow cytometry experiments. Contradictory to previous results, for both dipeptides, the dynamic range of pSC$_{Cg}$-LysG-S was decreased compared to His-Ala. Comparing these results to pSenLysG, a broader dynamic range for the respective dipeptides was assessed. Analyzing the operational range of both pSC$_{Cg}$-LysG-S and pSenLysG, an increasing $K_D$ of the tested inducer molecules, hence, a decreased affinity of the ligand to LysG, shifted the operational range to higher concentrations. In addition, results showed that the lower the affinity of the respective amino acid to LysG, the higher is the required amino acid concentration to obtain the same biosensor response for both LysG-based sensor variants.

The results of the previously performed 48-well cultivations were confirmed by the fluorescence response at the single-cell level. The fluorescence response for the whole cell population was assessed for pSenLysG and pSC$_{Cg}$-LysG-S with both dipeptide inducers, since both sensors showed a homogeneous fluorescence response on the single-cell level at different inducer concentrations for inducer His-Ala. For both sensor constructs, a shift of the cell population from lower to higher fluorescence intensity was achieved with increasing inducer concentration. A homogeneous fluorescence response was recorded for all cells carrying pSenLysG within a culture, whereas pSC$_{Cg}$-LysG-S only yielded a heterogeneous fluorescence response for the whole cell population.

### 6.2.7 The unified biosensor design enables the functional transfer of biosensors from *C. glutamicum* to *E. coli*.

Unable to generate a fully functional LysG-based biosensor in *E. coli* using the native promoters from *C. glutamicum* (Wang *et al.*, 2016), the unified sensor design was applied to construct LysG-based biosensors for *E. coli.* Demonstrating the broad applicability of the unified sensor design for transcriptional regulators with different modes of action, the transcriptional repressor PhdR was also used to construct biosensors to functionally implement in *E. coli.* In this context, biosensors carrying either *lysG* or *phdR* under control of four different constitutive promoters of the PLTetO1 promoter library were constructed. Since CoA-activation of the ring-hydroxylated phenylpropanoids is essential for ligand-binding to PhdR, co-expression of a 4CL-encoding gene (pCDF-BAD-4*cl*) from *Streptomyces coelicolor* enabled regulator activity in *E. coli* (van Summeren-Wesenhagen & Marienhagen, 2015). Biosensor characterization was performed in *E. coli* DH10B pCDF-BAD-4*cl* with plasmid-based *4cl* expression for all PhdR-based sensor variants. All PhdR- and LysG-based biosensors were

first characterized in 48-well plate cultivations with regard to their fluorescence response at different inducer concentrations ranging from 0 - 200 µM *p*-coumaric acid and 0 - 90,000 µM L-histidine for PhdR- and LysG based biosensors, respectively (Figure 19A and 20A).

The obtained results from liquid cultivation for PhdR-based biosensors confirmed the suitability of the unified sensor design to construct functional biosensors in *E. coli*, since a fluorescence response could be detected for all biosensor constructs (Figure 19A). However, the antibiotic concentration for plasmid maintenance and arabinose concentration controlling *4cl* gene expression had to be adjusted for every single variant to lower metabolic burden for the host organism. The results for PhdR-based biosensors in *E. coli* showed that the highest dynamic- and operational range could be determined for the biosensor variants pSC_{Ec}-PhdR-W with the weakest constitutive promoter (Figure 19A). Assessing the fluorescence response on the single-cell level, a rise in fluorescence intensity with increasing *p*-coumaric acid concentrations could be confirmed (Figure 19B).
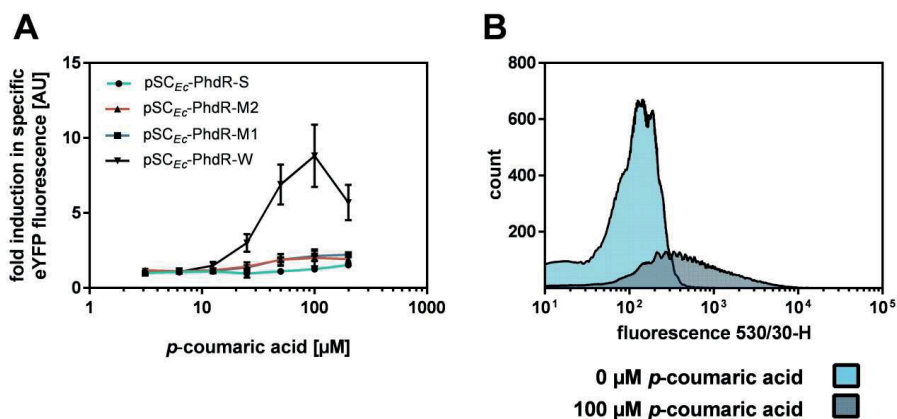


**Figure 19: Biosensor response of PhdR-based biosensors in *E. coli*.** (A) Dose-response plot to assess the operational and dynamic range of PhdR-based biosensors in *E. coli* with *p*-coumaric acid. BioLector cultivations were externally supplemented with seven different inducer concentrations ranging from 4 - 200 µM *p*-coumaric acid. The fluorescence response was plotted as fold change in specific eYFP fluorescence. Error bars represent standard deviations calculated from three biological replicates. (B) The fluorescence response on single-cell level with *E. coli* strains carrying pSC_{Ec}-PhdR-W were determined in the presence of 0 µM *p*-coumaric acid (blue) and 100 µM *p*-coumaric acid (dark blue). In each experiment 95,000 representative single cells were analyzed.

Similar to PhdR-based biosensors, the sensor response of all constructed LysG-based biosensors could be restored in *E. coli*. The maximal fold-induction of all variants was reached for pSC_{Ec}-LysG-M at 90,000 µM L-histidine with 15-fold induction also generating the widest operational range (Figure 20A). Nevertheless, a considerable difference in dynamic- and operational range between different sensor variants comprising a weak and moderate promoter was not apparent. Illustrative for LysG-based biosensors in *E. coli*, the fluorescence

response of pSC$_{Ec}$-LysG-M was investigated on the single-cell level showing a clear shift from lower to higher fluorescence intensity with increasing inducer concentrations (Figure 20B).

Independent from the mode of action of the used transcriptional regulator, the fluorescence response of the respective transcriptional biosensor from *C. glutamicum* could be restored in *E. coli* and additionally showed a rise in fluorescence intensity with increasing inducer concentrations on the single-cell level.
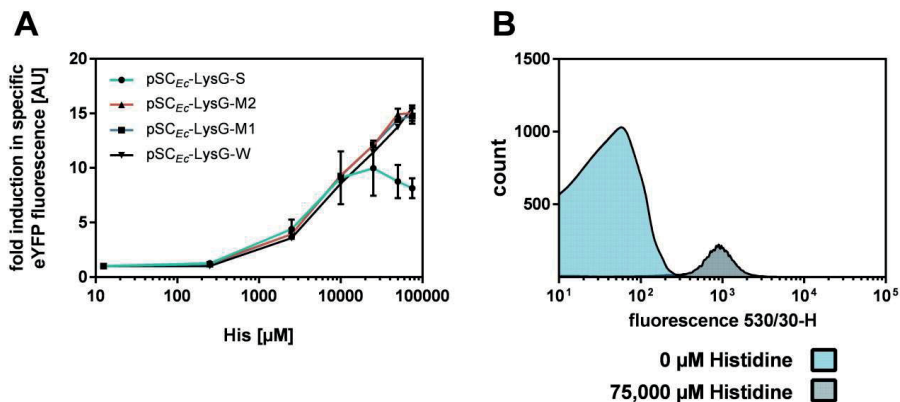


**Figure 20: Sensor response of LysG-based biosensors in *E. coli*.** Dose-response plot to assess the operational and dynamic range of LysG-based biosensors in *E. coli* with inducer L-histidine. BioLector cultivations were externally supplemented with seven different inducer concentrations ranging from 125 - 75,000 µM L-His. The fluorescence response was plotted as fold change in specific eYFP fluorescence. Error bars represent standard deviations calculated from three biological replicates. (B) The fluorescence response on a single-cell level with *E. coli* strains carrying pSC$_{Ec}$-LysG-W were performed in the presence of 0 µM His (blue) and 75,000 µM His-Ala (dark blue) (all externally supplemented). In each experiment 95,000 representative single cells were analyzed.

# 7    Conclusion and Outlook

## 7.1    Advancing CRISPR/ Cas12a applications in *C. glutamicum*

The CRISPR/ Cas genome editing system for in *C. glutamicum*, which is based on the Cas12a endonuclease from *Francisella novicida* was transformed into a more usable method by the design, construction and validation of the vector pJYScr, which now enables fast and convenient vector assembly for an efficient crRNA delivery. CRISPR/ Cas12a-assisted recombineering with ssDNA oligonucleotides using this novel vector enables the efficient and precise introduction of nucleotide substitutions into the genome of *C. glutamicum*. Reliable site-directed genome engineering and successful site-directed mutagenesis of three subsequent base pairs can be performed within a stretch of four nucleotides upstream and ten nucleotides downstream of the respective PAM site.

Furthermore, the L-glutamate efflux in *C. glutamicum* mediated by mechanosensitive MscCG channel was studied in detail when performing site-directed mutagenesis using the developed CRISPR/ Cas12a system. Amino acid substitutions targeting positions V422 and E423 as well as complete deletion of the periplasmic domain of MscCG cause deregulated L-glutamate efflux in *C. glutamicum,* probably due to drastic structural alterations of this domain.

Taken together, the constructed crRNA delivery vector pJYScr simplifies plasmid constructions and is a first step to promote CRISPR/ Cas12a applications in *C. glutamicum*.

## 7.2    A unified design allows fine-tuning of biosensor parameters and application across bacterial species

The unified biosensor design introduced here enables an easy and time-saving construction of transcriptional biosensors for applications in *C. glutamicum* and *E. coli*. Furthermore, this design allows for fine-tuning of important sensor parameters including the dynamic range and the operational range.

Independent from the transcriptional regulator's mode of action, the adaption of the regulator expression level for a specific application was sufficient to ensure a biosensor response usable for sensor applications either in liquid cultures or at the single-cell level. In addition, it could be demonstrated that an individual characterization of each biosensor in the presence of different ligands is absolutely essential, as different binding affinities of the respective inducer molecules strongly influence the overall biosensor response. Likewise, it should also be noted that if a substrate still needs to be metabolized in order to act as an inducer of a transcriptional regulator, the substrate specificity as well as the turn-over rate of the enzyme(s) involved must be considered. Further consideration should be taken with regard to substrate uptake to ensure

that the externally applied inducer concentration is congruent to the intracellular inducer concentration.

The set of constructed PhdR-based sensors can be directly used in both organisms, *C. glutamicum* and *E. coli*, to screen genetically diverse libraries of enzymes involved in natural product synthesis to isolate improved variants. A potential target might be enzyme 4CL which mediates the CoA-activation of *p*-coumaric acid, ferulic acid and caffeic acid as precursors for biotechnologically interesting flavonoids or stilbenes.

With regard to the above criteria, the unified biosensor design described can be applied to construct tailor-made biosensors from regulatory circuits. Subsequent characterization allows for the selection of distinct sensor variants suitable for application in screening campaigns at the single-cell level. With a given constitutive promoter library, the sensor design may even be applied for other prokaryotic species as well as novel uncharacterized regulatory circuits. Thus, the unified sensor design enables semiquantitative intracellular detection of small molecules in a broad variety of organisms and represents a powerful tool for future strain engineering campaigns.

## 8 Peer-reviewed publications

### 8.1 CRISPR/ Cas12a mediated genome editing to introduce amino acid substitutions into the mechanosensitive channel MscCG of *Corynebacterium glutamicum*

# CRISPR/Cas12a Mediated Genome Editing To Introduce Amino Acid Substitutions into the Mechanosensitive Channel MscCG of Corynebacterium glutamicum

Karin Krumbach,[†] Christiane Katharina Sonntag,[†] Lothar Eggeling,[†] and Jan Marienhagen*,[†,‡]

[†] Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich, D-52425 Jülich, Germany
[‡] Institute of Biotechnology, RWTH Aachen University, Worringer Weg 3, D-52074 Aachen, Germany

S Supporting Information

**ABSTRACT:** Against the background of a growing demand for the implementation of environmentally friendly production processes, microorganisms are engineered for the large-scale biosynthesis of chemicals, fuels, or food and feed additives from sustainable resources. Since strain development is expensive and time-consuming, continuous improvement of molecular tools for the genetic modification of the microbial production hosts is absolutely vital. Recently, the CRISPR/ Cas12a technology for the engineering of *Corynebacterium glutamicum* as an important platform organism for industrial amino acid production has been introduced. Here, this system was advanced by designing an easy-to-construct crRNA delivery vector using simple oligonucleotides. In combination with a *C. glutamicum* strain engineered for the chromosomal expression of the β-galactosidase-encoding *lacZ* gene, this new plasmid was used to investigate CRISPR/Cas12a targeting and editing at various positions relative to the PAM site. Finally, we used this system to perform codon saturation mutagenesis at critical positions in the mechanosensitive channel MscCG to identify new gain-of-function mutations for increased L-glutamate export. The mutations obtained can be explained by particular demands of the channel on its immediate lipid environment to allow L-glutamate efflux.

**KEYWORDS:** CRISPR/Cas, recombineering, L-glutamate, Corynebacterium glutamicum, metabolic engineering, genome editing

*Corynebacterium glutamicum*
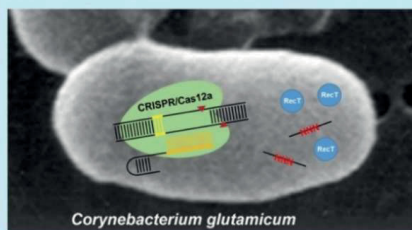
C lustered regularly interspaced short palindromic repeats (CRISPR) together with CRISPR-associated (Cas) proteins constitute an immune system in bacteria and archaea that cut foreign DNA entering the cell.[1] DNA degradation is achieved by DNA cleavage at specific positions *via* the Cas-encoded nuclease activity, which is guided to the target site by a short CRISPR-RNA (crRNA) complementary to one of the strands of the target DNA. The simplicity and programmability of the CRISPR/Cas system, established for a broad range of different organisms, has enabled its straightforward application to specifically target any genomic location. Its applicability for engineering bacteria was first demonstrated in the context of recombineering in *Escherichia coli* where 65% of recombinant cells could be recovered by specifically killing of nonmutated cells.[2] Since then, this approach has been successfully used for manipulating numerous bacterial species, including *Streptococcus pneumoniae*,[2] *Lactobacillus reuteri*,[3] *Clostridium beijerinckii*,[4] and *Streptomyces coelicolor*,[5] and more recently also *Corynebacterium glutamicum*.[6,7]

While Cas9 is currently the best-characterized and most widely used endonuclease among available CRISPR editing methods, Cas12a (previously named Cpf1) has recently emerged as an alternative for Cas9. This can be reasoned by

the cytotoxicity of Cas9 for a number of bacteria;[6,8−10] possible reasons could be the generation of blunt ends by Cas9 and the lack of or reduced activity of nonhomologous end-joining (NHEJ) in bacteria.[11] In contrast, Cas12a generates staggered cut DNA with 5′-overhangs, which can be easier repaired by prokaryotes. Recently, Cas12a found additional applications in CRISPR genome editing methods for bacteria, such as *E. coli*,[12] *Mycobacterium smegmatis*,[12] Cyanobacteria,[13] *Pseudomonas putida*,[14] and *Clostridium difficile*.[15]

The first developed CRISPR/Cas method for bacteria taking advantage of Cas12a was developed for *C. glutamicum*, directly combined with single-stranded DNA (ssDNA) recombineering to relieve allosteric feedback inhibition of the γ-glutamyl kinase by L-proline.[16] *C. glutamicum*, generally recognized as safe (GRAS), has been used for decades in industry for the million ton-scale production of monosodium L-glutamate.[17,18] Despite intensive efforts to understand L-glutamate synthesis in this organism in detail, it was only recently possible to solve the mystery of L-glutamate export by identifying the responsible mechanosensitive channel MscCG.[19] In addition, *C. glutamicum*

has also been engineered for the production of organic acids,[20] alcohols,[21,22] polyphenols,[23,24] and specialty chemicals.[25]

Methodical milestones for the engineering of *C. glutamicum* include the availability of DNA transfer methods,[26] a two-step homologous recombination method (the "sacB-system"),[27] and, more recently, the introduction of RecT-mediated recombineering.[28] The sacB-system, based on the conditionally lethal levansucrase-activity (SacB), is currently the method of choice for the site-specific introduction of genomic mutations. In principle, RecT-mediated recombineering with oligonucleotides (ssDNA) represents a true and less laborious alternative to the sacB-system, but applicability of this method requires a selection step since the frequency of obtaining recombinants in *C. glutamicum* is about $10^{-3}$.[28] Here, biosensors, enabling direct screening for mutated cells with increased metabolite formation *via* fluorescence-activated cell sorting (FACS), were successfully used to isolate recombineered variants.[29-32] However, depending on the application in question, development of a biosensor-based screen is not an option. In such cases, CRISPR/Cas12a mediated killing of nonrecombineered *C. glutamicum* cells is another option for quickly identifying variants with the desired genotype by selection, which holds the potential to decrease the experimental effort even further.

In this manuscript, we characterized the CRISPR/Cas12a system with regard to the introduction of point mutations into the genome of *C. glutamicum* and developed a simplified CRISPR/Cas12a protocol. Furthermore, we used this system to identify new gain-of-function mutations in the mechanosensitive channel MscCG leading to ʟ-glutamate efflux.

## ■ MATERIALS AND METHODS

**Plasmid Constructions.** In the used CRISPR/Cas12a system,[7] vector pJYS2_crtYf is applied for crRNA delivery. Using this vector as a template, plasmid pJYScr was constructed comprising a *Bam*HI/*Bst*BI cloning site which facilitates the insertion of protospacer sequences into the delivery vector. Removal of the original crtYf protospacer sequence was performed by DNA restriction using *Bam*HI and *Dra*I. The respective restriction sites for these enzymes are located at position 3196 and 3244. Since the available vector pJYS2_crtYf contains three additional *Dra*I sites,[16] the selected *Dra*I restriction site was not suitable for cloning. However, to make this site available for cloning, the restriction site was replaced by a *Bst*BI site using a synthesized 313 bp *Xho*I-*Afl*II fragment (Thermo Fisher Scientific, Waltham, MA, USA), comprising the parental sequence but replacing two nucleotides in the *Dra*I site to create a *Bst*BI restriction site. For plasmid construction, the delivery vector pJYS2_crtYf was cut with *Xho*I/*Afl*II. Subsequently, the resulting plasmid backbone fragment was purified and combined with the synthesized fragment. Since the vector backbone contains an additional *Bst*BI restriction site, site-directed mutagenesis was performed to introduce a C1165T mutation thereby eliminating this restriction site. The QuikChange II XL Site-Directed Mutagenesis Kit (Agilent, Santa Clara, CA, USA) was used for this purpose. However, we experienced that even after gel purification of vector DNA treated with *Bam*HI/*Bst*BI still intact vector was obtained after ligation with oligonucleotides. Therefore, the 49 bp *Bam*HI/*Bst*BI fragment was replaced by a 931 bp dummy fragment (nt 379847−383111 of BA000036) yielding vector pJYScr. This vector makes it easy to separate individual DNA fragments after *Bam*HI/*Bst*BI restriction for further processing. The DNA sequence of pJYScr and the

vector itself are deposited at Addgene (#132719). As control experiment, pJYScr:crtYf bearing the same crRNA as pJYS2_crtYf was constructed to evaluate, whether the replacement of restriction sites has an effect on the editing efficiency in comparison to pJYS2_crtYf. In these experiments, $21 \pm 7$ cfu were obtained with the pJYScr:crtYf delivery vector and with the original vector pJYS2_crtYf $34 \pm 10$ cfu as escaper cells could be counted. This indicated that the vector backbone modifications did not affect the formation of an active Cas12a/crRNA complex in *C. glutamicum* cells.

For construction of the individual crRNA delivery vectors, annealed oligonucleotides encoding loop region and spacer were used. For annealing, 10 μL of both oligonucleotides (200 nmol/mL $H_2O$) were mixed, incubated for 10 min at 98 °C, at room temperature for 1 h, and subsequently stored on ice. The annealed oligonucleotides were then ligated with the 4332 bp *Bam*HI/*Bst*BI vector fragment of pJYScr prepared *via* gel purification. Then *E. coli* DH5α was transformed with the ligation mix. Spectinomycin resistant clones indicating a successful transformation event were identified on LB plates containing the antibiotic at a concentration of 100 mg/L. Oligonucleotides pJYS2_fw2 and pJYS2_rv2 to amplify a 995 bp PCR fragment and DNA sequencing were used to verify the presence the crRNA coding region in the vector. If desired, recombinant strains were cured from the pJYScr plasmids by cultivation at 34 °C without supplementation of antibiotics as described earlier.[7]

**Strain Constructions.** *C. glutamicum* ATCC13032 was used throughout this study. Integration of the *lacZ* gene under control of the synthetic PH36 promoter was achieved by isolating a 3190 bp *Sma*I-fragment from the plasmid pMK-RQ_PH36_LacZ synthesized by Invitrogen (Thermo Fisher Scientific, Waltham, MA, USA). This fragment was inserted into the pK18mobsacB vector carrying the two flanking regions of the chromosomal noncoding region at nucleotide 1.404.651 (acc. no. NC_006958) as previously described.[33] Using two homologous recombination events strain *C. glutamicum* WT::*lacZ* was obtained. DNA sequencing confirmed a correct chromosomal integration. *C. glutamicum* WT::*lacZ* colonies plated onto BHIS-X-gal (5-bromo-4-chloro-3-indolyl-β-ᴅ-gal-actopyranoside) turned blue after 3 days, whereas the *C. glutamicum* wild type cells remained yellowish. The two constructed *mscCG* mutants devoid of the carboxyterminal end were constructed by homologous recombination using pK19mobsacB-420stopΔ and pK19mobsacB-424stopΔ (Supporting Table S1), each carrying two flanking fragments of ~500 bp.[34]

**Transformation and Recombination.** Preparation of electrocompetent *C. glutamicum* pJYScr cells providing inherent expression of RecT and Cas12a was performed as described previously with only minor modifications.[16] In short, a single colony from a BHI agar plate was used to inoculate a preculture of 4 mL BHISG-Kan and incubated overnight at 30 °C and 220 rpm. Subsequently, 500 μL of the preculture was transferred into 50 mL of BHISG-Kan supplemented with 1 mL/L Tween 80 and 4 g/L glycine. The culture was harvested when the optical density at 600 nm ($OD_{600}$) reached approximately 1. Cells were chilled on ice for 20 min and centrifuged at 2,600 g and 4 °C for 10 min. Subsequently, two washing steps in 50 mL of 10% glycerol were performed. Competent cells were resuspended in 500 μL 10% glycerol, and aliquots of 90 μL were stored at −80 °C. For transformation, an aliquot was thawed on ice, mixed with 0.5

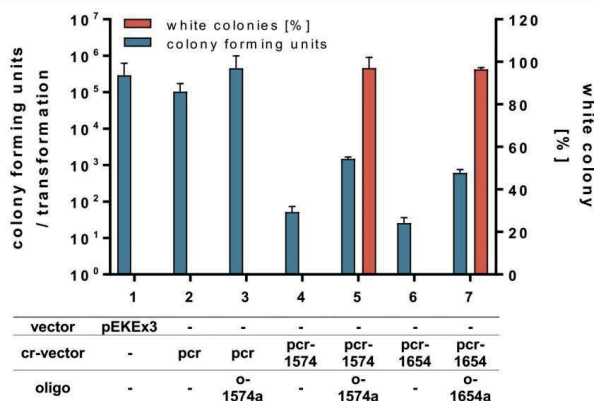| vector | pEKEx3 | - | - | - | - | - | - |
|--------|--------|---|---|--------------|--------------|--------------|--------------|
| cr-vector | - | pcr | pcr | pcr-1574 | pcr-1574 | pcr-1654 | pcr-1654 |
| oligo | - | - | o-1574a | - | o-1574a | - | o-1654a |

**Figure 1.** Efficiency of genome targeting and editing in *C. glutamicum*. The transformation assays to transform *C. glutamicum* WT::lacZ pJYS1-petFU contained 500 ng plasmid and 10 μg oligonucleotide where indicated. The transformants were plated on X-Gal for blue-white screening. The results are averages from at least two independent experiments, and error bars depict standard deviations. In control experiments, *C. glutamicum* was either transformed with the cloning vector pEKEx3 (column 1), the empty vector pJYScr (column 2) inscribed as pcr, or empty vector plus oligo o-1574a targeting a PAM site in the chromosomal *lacZ* gene (column 3). When the crRNA was provided by vector pJYScr-lacZ-1574 (pcr-1574, column 4) or vector pJYScr_lacZ-1654 (pcr-1654, column 6), only few cells survived, whereas addition of the corresponding oligonucleotides (o-1574a, column 5 or o-1654a, column 7) for PAM site-inactivation and introduction of a stop codon resulted in a drastic increase of surviving colonies with >96% of having the desired mutation.

μg of the pJYScr derivative and 10 μg of PAGE purified oligonucleotide (Eurofins MWG Operon (Ebersberg, Germany), and transferred into 4 °C precooled electroporation cuvettes (width 2 mm). The oligonucleotides used were 59 bp in size and encoded a mutated PAM site, spacer region, and, if applicable, the mutation to be introduced. After electroporation, cells were immediately transferred to 900 μL of prewarmed BHISG medium and heat-shocked for 6 min at 46 °C. The cells were incubated to recover for 2 h at 30 °C with shaking at 170 rpm. Cells were plated on BHISG containing kanamycin and spectinomycin (BHISG-Kan-Spc-Xgal), and incubated for 3−4 days to determine the number of colony forming units and their respective color.

**Media and L-Glutamate Formation.** *E. coli* was grown on LB and *C. glutamicum* on BHI (Difco Laboratories, Detroit, USA). BHIS contained in addition 90 g/L sorbitol, BHISG moreover 10 g/L glucose, BHISG-Kan 15 mg/L kanamycin, BHISG-Kan-Spec 100 mg/L spectinomycin, and BHISG-Kan-Spc-Xgal 50 mg/L. For glutamate formation, precultures were grown overnight in 20 mL BHI. Cells were harvested by centrifugation, washed with 20 mL 0.9% NaCl and inoculated to an $OD_{600}$ of 1 in defined CGXII medium (Eggeling and Bott 2005) containing 50 μg/L biotin. L-Glutamate in supernatants was determined as its *o*-phthaldialdehyde derivative *via* HPLC and fluorescence detection as previously described.[35]

## ■ RESULTS

**Simple Construction of crRNA Delivery Vectors Using Oligonucleotides.** According to the protocol currently used, amplification of the entire pJYS2_crtYf vector is necessary for inserting the desired crRNA-encoding spacers.[7] Despite different PCR conditions and the use of megaprimers, this procedure resulted in a low yield and often a doubling of the oligonucleotides. With the original protocol we were able to construct only 1 of 5 desired vectors, which motivated us to

develop an improved crRNA delivery vector expediting necessary constructions. The starting vector pJYS2_crtYf encodes the loop region of the crRNA in combination with the specific crtYf-targeting sequence on a 49 bp *Bam*HI-*Dra*I-fragment. The single *Bam*HI site is located between the strong promoter pJ23119 and the specific spacer, whereas one of the four *Dra*I sites of the vector is located between the spacer and the T1rrnB terminator. Using a synthesized fragment and applying site-directed mutagenesis, we replaced the *Dra*I site by a *Bst*BI site and removed an existing *Bst*BI site in the vector backbone. The resulting vector in principle allows replacing the crtYf-specific crRNA-encoding sequence on the 49 bp *Bam*HI-*Bst*BI fragment by any other sequence.

However, we substituted the crtYf-fragment by a larger fragment of 931 bp resulting in the vector pJYScr of 5.252 kb, which enables easy preparation of the 4.321 kb *Bam*HI-*Bst*BI vector fragment for the insertion of oligonucleotides encoding the desired crRNA (Figure S1). This new vector for genome editing has been made available through Addgene (#132719).

To validate the functionality of pJYScr in a CRISPR/Cas12a selection experiment, a crRNA targeting *murE* of *C. glutamicum* was cloned into the vector. The *murE* gene encodes the essential UDP-*N*-acetylmuramoylalanyl-D-glutamate 2,6-diaminopimelate ligase.[36] The corresponding oligonucleotides comprising the PAM site 5'-GTTG-3' (nt 255−258) within *murE*, were annealed and ligated with the 4.321 kb *Bam*HI-*Bst*BI fragment of pJYScr, resulting in pJYScr:murE. After transformation of *C. glutamicum* pJYS1-petFU (providing constitutive Cas12a-activity) with 1 μg of the pJYScr:murE plasmid, only 81 ± 26 cfu as escaper cells were obtained demonstrating the formation of an active Cas12a/crRNA complex in the cells. A control experiment, in which *C. glutamicum* wild type cells without pJYS1-petFU were transformed with the same vector yielded $3.2 \times 10^5$ cfu.
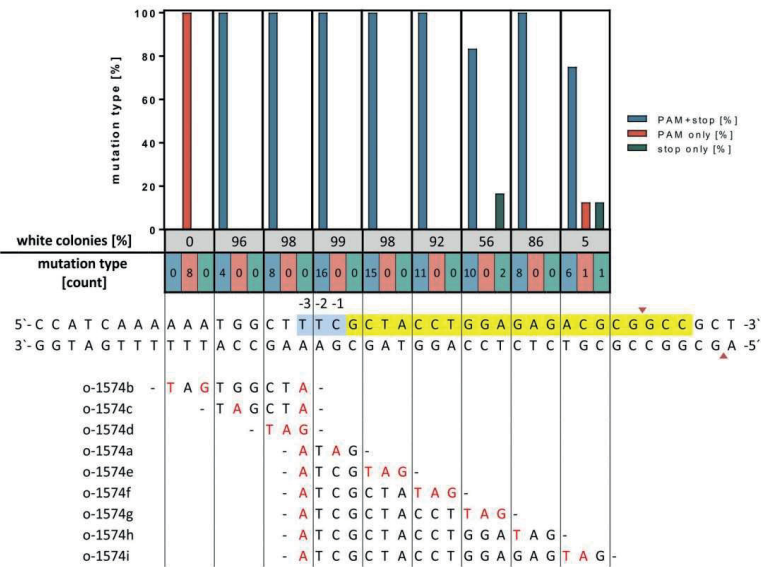
**Figure 2.** Genome editing efficiency relative to the PAM site in *C. glutamicum* using the combined recombineering-CRISPR/Cas12a selection method. Shown is a portion of the chromosome-integrated *lacZ* gene with the PAM sequence highlighted in blue and the protospacer highlighted in yellow. The Cas12a-cleavage site yielding staggered ends is indicated by red triangles. The relevant DNA sequence of the oligonucleotides for site-directed mutagenesis is shown at the bottom with all mutations highlighted in red. The fraction of white colonies obtained is given as percentage share. From each assay, the *lacZ* gene of randomly selected white clones was sequenced ($n = 4-16$). Either the PAM site was mutated and the stop codon was introduced (blue bars), only the PAM site was mutated (red bars), or only the stop codon was introduced (green bars). The respective numbers for each type of mutation (PAM+stop/stop only/PAM only) assessed *via* DNA sequencing are also given as numbers below the colored bars. Detailed DNA sequencing results are given in Supporting Table S1.

**Efficiency of CRISPR/Cas12a-Assisted ssDNA-Recombineering in *C. glutamicum*.** With the aim to explore the capabilities of combining ssDNA-recombineering with CRISPR/Cas12a-mediated elimination of nonrecombinant cells, we first constructed a *C. glutamicum* test strain carrying a chromosomal integration of the *lacZ* gene encoding for the β-galactosidase from *E. coli* (*C. glutamicum* WT::lacZ). In this particular strain, *lacZ*-expression is under control of the synthetic promoter pH36,[37] enabling easy identification of generated *lacZ* mutants on agar plates supplemented with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) by blue-white screening. Subsequently, this strain was transformed with pJYS1-petFU providing constitutive Cas12a- as well as RecT-activity to this strain for the following recombineering-CRISPR/Cas12a experiments.[16] When transformed with the empty pJYScr-vector with or without the recombineering oligonucleotide o-1574a, high transformation rates comparable to that with the standard pEKEx3 vector as a control were obtained (Figure 1). A high transformation rate is essential to obtain recombinants in the combined recombineering-CRISPR/Cas12a-killing approach. Noteworthy, the transformation rates obtained also confirmed the nontoxicity of Cas12a opposed to Cas9 as it has been already described for various organisms[13] including *C. glutamicum*.[16] As crRNA delivery vector targeting the chromosomal *lacZ* gene, we have built pJYScr-lacZ-1574 using the intragenic PAM site 5′-(T)TT(C)-3′ with a guide sequence of 21 nts. When *C. glutamicum* WT::lacZ pJYS1-petFU was transformed with this vector, only $51 \pm 21$ cells survived underlining the efficiency of Cas12a targeting in *C. glutamicum* (Figure 1). To assay for DNA-editing plus targeting of unmutated cells, *C. glutamicum* WT::lacZ pJYS1-petFU was transformed with pJYScr-lacZ-1574 in combination with the editing oligonucleotide o-1574a (all recombinogenic oligos are listed in Supporting Table S2). This oligonucleotide was designed to replace the PAM site and the adjacent codon TCG (5′-TTTCG-3′) by 5′-TATAG-3′, which results in an inactivated PAM site and the introduction of a stop codon within the open reading frame of *lacZ*. This experiment yielded $1500 \pm 172$ colonies of which more than 97% were white (Figure 1). Comparable numbers could be obtained when targeting a second PAM-site (5′-(T)TT(A)-3′) within *lacZ* using the oligonucleotide o-1654a and the simultaneous introduction of a stop codon by substituting a single nucleotide (Figure 1).

These results raised the question at which frequency a reliable introduction of mutations at different positions relative to the PAM site is possible with this system. To answer this, nine different oligonucleotides with a length of 59 nt were designed for introducing a stop codon over a stretch of 18 nt along with a mutation inactivating the PAM site. Oligonucleotides used for recombineering and introduction of the desired mutations in the *lacZ* gene and the PAM site are given in Table S2. Experiments conducted with these oligonucleotides yielded no white colonies when oligonucleotide o-1574b was used,

which would introduce two nucleotides substitutions upstream of the PAM-site (Figure 2). In this case, sequencing of eight randomly picked blue clones revealed indeed that only the PAM site was mutated. When performing recombineering-CRISPR/Cas12a with the editing oligonucleotides o-1574c, -d, -a, e, and -f, more than 96% of the surviving *C. glutamicum* variants were white in all cases indicating a successful introduction of the stop codon into the *lacZ* gene. This observation was supported by DNA sequencing of randomly selected clones, which confirmed correct incorporation of both mutations in all clones analyzed. However, when using the recombinogenic oligonucleotides o-1574g, o-1574h, or o-1574i designed for introducing a stop codon further downstream of the PAM site, introduction of the desired mutations was still possible, but only at a reduced frequency ranging from 75% to 90% (blue bar), depending on the number of substitutions necessary to generate a stop codon. Interestingly, DNA-sequencing revealed that in a few cases only the stop codon was successfully introduced whereas the PAM site was still intact (given in Figure 2 as "stop only"). We obtained a similar result of such partially successful editing events in experiments targeting a second PAM site in *lacZ* (Table S2). In these cases, possibly only a short sequence of the genomic DNA was exchanged for the respective oligonucleotide during recombineering. These clones escape Cas12a cleavage when only mismatches in the PAM-distal region of the DNA target are present.[38,39] Collectively, these data suggest that site-directed mutagenesis using the combined recombineering/Cas12a selection strategy is possible over a stretch of four nucleotides upstream and ten nucleotides downstream of the PAM sequence.

Inspired by these results, the mutagenic oligonucleotide o-1574fe-random was designed, which carries a continuous stretch of 18 random nucleotides spanning the previously targeted PAM sequence plus five nucleotides upstream and ten nucleotides downstream of that PAM site (Supporting Table S3). After performing the recombineering-CRISPR/Cas12a selection, only eight white colonies were obtained. DNA sequencing of the *lacZ* gene of all clones revealed a mutated and thus nonfunctional PAM sequence in all cases with four clones carrying additional mutation(s) downstream of the targeted position (nucleotides +11 to +16, Supporting Table S3). Apparently, when considering the low number of obtained colonies one can assume, that with the given oligonucleotide bearing such a long random sequence, recombineering does not work. Hence, this method cannot be used to exchange a longer stretch of DNA and is limited to the introduction of three consecutive nucleotide substitutions into the genome of *C. glutamicum*.

This conclusion is supported by the results of two additionally performed experiments using oligonucleotides o-1574e-mix and o-1574g-mix, which were both designed for the saturation mutagenesis of a selected codon (Supporting Table S3). Twenty-four white colonies obtained by recombineering-CRISPR/Cas12a selection using o-1574e-mix were randomly picked and sequenced. All variants were characterized by a mutated PAM site and 16 of them carried additional mutation(s) in the desired codon leading to 12 amino acid substitutions on the protein level without a discernible preference for specific substitutions (Supporting Table S3). Similarly, targeted DNA sequencing of 24 white clones resulting from the genome editing experiment using oligonucleotide o-1574g-mix yielded 13 variants bearing

mutations in the desired codon, leading to eight different amino acids at this particular position (Supporting Table S3).

**Site-Directed Mutagenesis Targeting *mscCG* in *C. glutamicum*.** An efficient L-glutamate export is an important prerequisite for achieving high L-glutamate titers with *C. glutamicum*. In case of this bacterium, the mechanosensitive channel MscCG (533 aa) is mainly responsible for L-glutamate efflux.[19] However, L-glutamate efflux only occurs in response to changes in the structure of this channel, induced either by growth under biotin limitation or addition of penicillin, ethambutol, or fatty acid ester surfactants to the medium.[40] In addition, an altered lipid composition may cause efflux.[41,42] According to the current model, these cultivation/production conditions mentioned increase membrane tension and thus activate the MscCG channel.[43] Interestingly, several gain-of-function (GOF) variants of MscCG are known causing constitutive L-glutamate efflux.[19,44] These mutational hot spots are located in the pore-lining helix of the N-terminal domain (1−286 aa), which has been also shown to be responsible for L-glutamate excretion in response to the above-mentioned treatments.[44] Furthermore, two additional GOF mutants, characterized by mutations in the C-terminal extracytoplasmic domain are known. In one case, insertion of an IS-element at position 419 was described, resulting in the introduction of a stop codon at position 424.[19] The other GOF mutant bears an P424L-substitution.[19] Finally, a third mutant devoid of the entire cytoplasmic domain (amino acids 420−533) has been described.[44] We were interested in studying consequences of mutations in this particular region at positions 422 and 423 of MscCG and used the recombineering-CRISPR/Cas12a method to assay additional amino acid substitutions at these particular positions. Both targeted amino acid positions of MscCG are predicted to be located at the transition of the fourth transmembrane helix to the periplasmic space.[19] Presumably, this site determines the overall structure of MscCG and is crucial for the interaction of the channel protein with other components of the cell wall.

*C. glutamicum* pJYS1petFU was transformed with pJYScr_mscCG1269c providing the crRNA and with 10 μg of the mutagenic oligo o-mscCG1269-V422nnn for site saturation mutagenesis (Table S4). The assay yielded 43 clones of which 16 were randomly chosen and sequenced. In these clones, altogether seven different amino acid substitutions at amino acid position V422 of MscCG could be identified. Similarly, position E423 was also targeted, eventually yielding 27 clones surviving CRISPR/Cas12a selection after recombineering. In 19 clones, DNA sequencing revealed introduction of eight different amino acids and one stop codon. In addition, one variant with an undesired deletion of two nucleotides in the PAM sequence was found. This frameshift mutation inevitably resulted in a stop codon (E423AQstop). Recombinant strains were cured from the respective plasmids and subsequently assayed for L-glutamate accumulation in the culture supernatant.

As previous studies have shown, L-glutamate formation, even in described MscCG mutants, is very much dependent on biotin concentration, medium composition and sugar concentration.[19,44,45] For instance, the known MscCG mutant devoid of the extracytoplasmic domain[44] excretes L-glutamate only at low biotin concentrations, whereas the mutant carrying the IS-element insertion at position 419 accumulates L-glutamate even in the presence of excess biotin.[19] For assaying the L-glutamate accumulation capabilities of the newly constructed
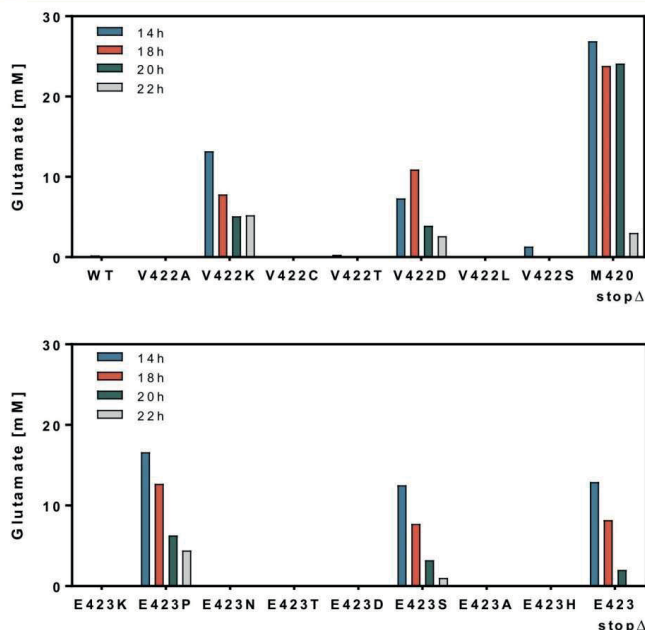
**Figure 3.** L-Glutamate accumulation in the supernatant of *C. glutamicum* variants with different mutations in the mechanosensitive channel MscCG gene at position V422 (top) and E423 (bottom). In addition, MscCG variants with stop codons at positions M420 (M420stopΔ) and E423 (E423stop), respectively, were also characterized with regard to the accumulation of L-glutamate. L-Glutamate concentrations for each variant were determined after 14, 18, 20, and 22 h.

MscCG-variants obtained by recombineering-CRISPR/Cas12a selection, we chose stringent conditions in defined medium with biotin excess (50 μg/L) in combination with a low glucose concentration (40 g/L). Under these conditions, the *C. glutamicum* wild type does not excrete L-glutamate (Figure 3a). The MscCG variants V422A, V422C, V422T, V422L, and V422S did not accumulate L-glutamate. In contrast, in the supernatant of the MscCG-variants V422 K and V422D, in which the valine residue was substituted by charged amino acids, a substantial L-glutamate accumulation of up to 13 mM could be detected (Figure 3a). However, variants E423 K and E423D did not show L-glutamate accumulation, as did mutants E423N, E423T, E423A, and E423H (Figure 3b). In contrast, mutant E423P, E423S as well as the strain with the MscCG variant E423stop promoted L-glutamate export. Noteworthy, L-glutamate concentrations decreased with time due to reuptake and utilization of this amino acid as known from previous studies.[45]

Interestingly, the MscCG variant E423AQstop did not cause L-glutamate accumulation as expected (data not shown). We also used the conventional sacB-system to delete two different portions of the extracytoplasmic domain of MscCG. In mutant M420stopΔ methionine in position 420 was replaced by a stop codon and the carboxyterminal end of 113 amino acids deleted. Similarly, the *C. glutamicum* strain carrying the shortened MscCG variant P424stopΔ missing the carboxyterminal 109 amino acids was constructed. With variant M420stopΔ, L-glutamate concentrations of up to 27 mM could

be detected, and with mutant P424stopΔ concentrations of up to 130 mM could be determined in the supernatant (Supporting Table S5). These observations are in line with previous data also implying that presence of the carboxyterminal domain has a negative impact on L-glutamate channeling activity.[44]

### ■ DISCUSSION

CRISPR/Cas selection is a powerful technology to perform genome editing in living cells. On the basis of this system, methods for inhibiting transcription (CRISPRi), activation of gene expression, or modification of nucleobases have been developed.[46] In bacteria, for the most part, methods introducing double-strand breaks using the crRNA/Cas-complex to eliminate unaltered strain variants find an application.[47] The main challenge during combined recombineering-selection experiments using RecT-recombineering and CRISPR/Cas represents balancing activities of all individual components including the recombinase, the crRNA and the endonuclease involved. The need to ensure the availability of the respective components in appropriate quantities at the right time initially led to the development of CRISPR/Cas methods for *E. coli*, which require up to three vectors or three independent induction systems.[48,49] However, these established Cas9-based systems were not suitable for applications in *C. glutamicum*, for the most part due to the cytotoxicity of the Cas9 endonuclease.

However, a CRISPR/Cas genome editing system based on the Cas12a endonuclease originating from *Francisella novicida* was introduced for *C. glutamicum*.[16] This genome editing method is more suitable for this bacterium as (i) Cas12a is not toxic for *C. glutamicum*; (ii) induction of gene expression (to establish endonuclease or recombinase activity and crRNA formation) is not necessary due to a fortunate choice of constitutive promoters;[16] and (iii) the crRNA, 43 nt in length, is about half the size of a typical Cas9 crDNA.[39] However, the construction of the crRNA delivery vector for a distinct protospacer sequence remained laborious and time-consuming, lowering the overall applicability of this method. The reconstructed vector pJYScr presented here, overcomes this bottleneck as it allows for fast and simple vector assembly to enable crRNA delivery. With CRISPathBrick, a similar strategy was developed by the Koffas group, which allows for the modular assembly of CRISPR-bricks to be used for CRISPR-Cas9 applications in *E. coli*.[50] In this study, CRISPR/Cas12a-assisted recombineering with ssDNA oligonucleotides using this new vector was assessed by performing experiments to introduce point mutations up- and downstream of PAM-sites in a chromosomally integrated *lacZ* gene. Interestingly, the results are consistent with previous experiments in *Mycobacterium smegmatis*, showing that any targeting distal from the PAM site results in a drastically decreased recombineering efficiency.[12] Nonetheless, the results obtained in this study confirm the potential of CRISPR/Cas12a-assisted recombineering using ssDNA oligonucleotides for the efficient and precise introduction of nucleotide substitutions into the genome of *C. glutamicum*. This was demonstrated when performing site-directed mutagenesis of the gene encoding the MscCG channel of *C. glutamicum*. We focused on positions 422 and 423 of the MscCG protein located at the transition of the fourth transmembrane helix to the periplasmic space.[19] It is striking that only substitution by charged amino acids residues resulted in an open-channel conformation at position 422. Given that the hydrophobic head groups of cardiolipin and phosphatidylglycerol are present in *C. glutamicum* in this region,[41] the direct local membrane environment may be altered, which in turn creates the necessary conditions for the amino-terminal part of MscCG to mediate the L-glutamate efflux.[44] In contrast, amino acid substitutions E423P and E423S, presumably causing drastic structural alterations, support L-glutamate efflux at position 422. Against the background that complete deletion of the periplasmic domain also fosters L-glutamate efflux,[44] there are two obvious possible explanations. Again, the immediate channel environment could be affected, which could result in a deformed channel. Alternatively, the extracytoplasmic domain could be forced away from the channel opening, thus enabling the efflux of L-glutamate. Such models where a periplasmic domain functions as an efflux-controlling lid and where this lid further interacts with the extracellular matrix have been described for eukaryotes.[51]

Taken together, the crRNA delivery vector pJYScr introduced here simplifies plasmid constructions and expands the number of CRISPR/Cas12a applications in *C. glutamicum*. In particular in combination with recombineering, CRISPR/Cas12a is a powerful tool when performing site-directed mutagenesis in the genome of *C. glutamicum* by selectively eliminating unedited cells as presented in this study.

## ◼ ASSOCIATED CONTENT

**ⓈSupporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.9b00361.

> Figure S1: Plasmid map of the pJYScr vector; Table S1: Plasmids and oligonucleotides used for plasmid constructions; Table S2: Recombinogenic oligos and obtained mutations at two *lacZ* locations; Table S3: Recombinogenic oligos for the insertion of a random stretch of nucleotides and codon saturation mutagenesis; Table S4: Recombinogenic oligonucleotides used for introducing mutations in *mscCG* and the mutations obtained; Table S5: L-Glutamate accumulation in the culture supernatant of constructed *mscCG* deletion mutants (PDF)

## ◼ AUTHOR INFORMATION

**Corresponding Author**

*Tel: +49 2461 61 2843. Fax: +49 2461 61 2710. E-mail: j.marienhagen@fz-juelich.de.

**ORCID** ◉

Jan Marienhagen: 0000-0001-5513-3730

**Author Contributions**

L.E. designed the experiments, K.K. and C.K.S. conducted the experiments. L.E. and J.M. wrote the manuscript.

**Notes**

The authors declare no competing financial interest.

## ◼ ACKNOWLEDGMENTS

## ◼ REFERENCES

(1) Park, H. M., Liu, H., Wu, J., Chong, A., Mackley, V., Fellmann, C., Rao, A., Jiang, F., Chu, H., Murthy, N., and Lee, K. (2018) Extension of the crRNA enhances Cpf1 gene editing in vitro and in vivo. *Nat. Commun.* 9, 1–12.

(2) Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L. A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31, 233–239.

(3) Oh, J. H., and Van Pijkeren, J. P. (2014) CRISPR-Cas9-assisted recombineering in Lactobacillus reuteri. *Nucleic Acids Res.* 42, 1–11.

(4) Diallo, M., Hocq, R., Collas, F., Chartier, G., Wasels, F., Wijaya, H. S., Werten, M. W. T., Wolbert, E. J. H., Kengen, S. W. M., van der Oost, J., Ferreira, N. L., and López-Contreras, A. M. (2019) Adaptation and application of a two-plasmid inducible CRISPR-Cas9 system in Clostridium beijerinckii. *Methods*, DOI: 10.1016/j.ymeth.2019.07.022.

(5) Huang, H., Zheng, G., Jiang, W., Hu, H., and Lu, Y. (2015) One-step high-efficiency CRISPR/Cas9-mediated genome editing in Streptomyces. *Acta Biochim. Biophys. Sin.* 47, 231–243.

(6) Cho, J. S., Choi, K. R., Prabowo, C. P. S., Shin, J. H., Yang, D., Jang, J., and Lee, S. Y. (2017) CRISPR/Cas9-coupled recombineering for metabolic engineering of Corynebacterium glutamicum. *Metab. Eng.* 42, 157–167.

(7) Jiang, Y., Qian, F., Yang, J., Liu, Y., Dong, F., Xu, C., Sun, B., Chen, B., Xu, X., Li, Y., Wang, R., and Yang, S. (2017) CRISPR-Cpf1 assisted genome editing of Corynebacterium glutamicum. *Nat. Commun.* 8, 15179.

(8) Wendt, K. E., Ungerer, J., Cobb, R. E., Zhao, H., and Pakrasi, H. B. (2016) CRISPR/Cas9 mediated targeted mutagenesis of the fast

growing cyanobacterium Synechococcus elongatus UTEX 2973. *Microb. Cell Fact.* 15, 1−8.

(9) Xu, T., Li, Y., Shi, Z., Hemme, C. L., Li, Y., Zhu, Y., Van Nostrand, J. D., He, Z., and Zhou, J. (2015) Efficient genome editing in clostridium cellulolyticum *via* CRISPR-Cas9 nickase. *Appl. Environ. Microbiol.* 81, 4423−4431.

(10) Sternberg, S. H., and Doudna, J. A. (2015) Expanding the Biologist's Toolkit with CRISPR-Cas9. *Mol. Cell* 58, 568−574.

(11) Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L., and Corn, J. E. (2016) Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* 34, 339−344.

(12) Yan, M.-Y., Yan, H., Ren, G., Zhao, J.-P., Guo, X., and Sun, Y.-C. (2017) CRISPR-Cas12a-Assisted Recombineering in Bacteria. *Appl. Environ. Microbiol.* 83, 1−13.

(13) Ungerer, J., and Pakrasi, H. B. (2016) Cpf1 Is A Versatile Tool for CRISPR Genome Editing Across Diverse Species of Cyanobacteria. *Sci. Rep.* 6, 39681.

(14) Sun, J., Wang, Q., Jiang, Y., Wen, Z., Yang, L., Wu, J., and Yang, S. (2018) Genome editing and transcriptional repression in Pseudomonas putida KT2440 *via* the type II CRISPR system. *Microb. Cell Fact.* 17, 41.

(15) Hong, W., Zhang, J., Cui, G., Wang, L., and Wang, Y. (2018) Multiplexed CRISPR-Cpf1-Mediated Genome Editing in Clostridium difficile toward the Understanding of Pathogenesis of C. difficile Infection. *ACS Synth. Biol.* 7, 1588−1600.

(16) Jiang, Y., Qian, F., Yang, J., Liu, Y., Dong, F., Xu, C., Sun, B., Chen, B., Xu, X., Li, Y., Wang, R., and Yang, S. (2017) CRISPR-Cpf1 assisted genome editing of Corynebacterium glutamicum. *Nat. Commun.*, DOI: 10.1038/ncomms15179.

(17) Kinoshita, S. (1972) Amino-acid production by the fermentation process. *Nature* 240, 211.

(18) Eggeling, L., and Bott, M. (2015) A giant market and a powerful metabolism: L-lysine provided by Corynebacterium glutamicum. *Appl. Microbiol. Biotechnol.* 99, 3387−94.

(19) Nakamura, J., Hirano, S., Ito, H., and Wachi, M. (2007) Mutations of the Corynebacterium glutamicum NCgl1221 gene, encoding a mechanosensitive channel homolog, induce L-glutamic acid production. *Appl. Environ. Microbiol.* 73, 4491−8.

(20) Wieschalka, S., Blombach, B., Bott, M., and Eikmanns, B. J. (2013) Bio-based production of organic acids with Corynebacterium glutamicum. *Microb. Biotechnol.* 6, 87−102.

(21) Blombach, B., Riester, T., Wieschalka, S., Ziert, C., Youn, J. W., Wendisch, V. F., and Eikmanns, B. J. (2011) Corynebacterium glutamicum tailored for efficient isobutanol production. *Appl. Environ. Microbiol.* 77, 3300−3310.

(22) Jojima, T., Noburyu, R., Sasaki, M., Tajima, T., Suda, M., Yukawa, H., and Inui, M. (2015) Metabolic engineering for improved production of ethanol by Corynebacterium glutamicum. *Appl. Microbiol. Biotechnol.* 99, 1165−1172.

(23) Kallscheuer, N., Vogt, M., Stenzel, A., Gätgens, J., Bott, M., and Marienhagen, J. (2016) Construction of a Corynebacterium glutamicum platform strain for the production of stilbenes and (2S)-flavanones. *Metab. Eng.* 38, 47−55.

(24) Milke, L., Ferreira, P., Kallscheuer, N., Braga, A., Vogt, M., Kappelmann, J., Oliveira, J., Silva, A. R., Rocha, I., Bott, M., Noack, S., Faria, N., and Marienhagen, J. (2019) Modulation of the central carbon metabolism of Corynebacterium glutamicum improves malonyl-CoA availability and increases plant polyphenol synthesis. *Biotechnol. Bioeng.* 116, 1380−1391.

(25) Kogure, T., and Inui, M. (2018) Recent advances in metabolic engineering of Corynebacterium glutamicum for bioproduction of value-added aromatic chemicals and natural products. *Appl. Microbiol. Biotechnol.* 102, 8685−8705.

(26) Schwarzer, A., and Pühler, A. (1991) Manipulation of Corynebacterium glutamicum by gene disruption and replacement. *Nat. Biotechnol.* 9, 84−7.

(27) Jäger, W., Schäfer, A., Pühler, A., Labes, G., and Wohlleben, W. (1992) Expression of the Bacillus subtilis sacB gene leads to sucrose

sensitivity in the gram-positive bacterium Corynebacterium glutamicum but not in Streptomyces lividans. *J. Bacteriol.* 174, 5462−5.

(28) Binder, S., Siedler, S., Marienhagen, J., Bott, M., and Eggeling, L. (2013) Recombineering in Corynebacterium glutamicum combined with optical nanosensors: A general strategy for fast producer strain generation. *Nucleic Acids Res.* 41, 6360−6369.

(29) Binder, S., Schendzielorz, G., Stäbler, N., Krumbach, K., Hoffmann, K., Bott, M., and Eggeling, L. (2012) A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. *Genome Biol.* 13, R40.

(30) Eggeling, L., Bott, M., and Marienhagen, J. (2015) Novel screening methods-biosensors. *Curr. Opin. Biotechnol.* 35, 30.

(31) Siedler, S., Khatri, N. K., Zsohár, A., Kjærbølling, I., Vogt, M., Hammar, P., Nielsen, C. F., Marienhagen, J., Sommer, M. O. A., and Joensson, H. N. (2017) Development of a Bacterial Biosensor for Rapid Screening of Yeast p-Coumaric Acid Production. *ACS Synth. Biol.* 6, 1860−1869.

(32) Jang, S., Jang, S., Xiu, Y., Kang, T. J., Lee, S.-H., Koffas, M. A. G., and Jung, G. Y. (2017) Development of Artificial Riboswitches for Monitoring of Naringenin In Vivo. *ACS Synth. Biol.* 6, 2077−2085.

(33) Blombach, B., Hans, S., Bathe, B., and Eikmanns, B. J. (2009) Acetohydroxyacid synthase, a novel target for improvement of L-lysine production by Corynebacterium glutamicum. *Appl. Environ. Microbiol.* 75, 419−27.

(34) Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., and Pühler, A. (1994) Small mobilizable multi-purpose cloning vectors derived from the Escherichia coli plasmids pK18 and pK19: selection of defined deletions in the chromosome of Corynebacterium glutamicum. *Gene* 145, 69−73.

(35) Marienhagen, J., Kennerknecht, N., Sahm, H., and Eggeling, L. (2005) Functional analysis of all aminotransferase proteins inferred from the genome sequence of Corynebacterium glutamicum. *J. Bacteriol.* 187, 7639−7646.

(36) Hochheim, J., Kranz, A., Krumbach, K., Sokolowsky, S., Eggeling, L., Noack, S., Bocola, M., Bott, M., and Marienhagen, J. (2017) Mutations in MurE, the essential UDP-N-acetylmuramoyla-lanyl-d-glutamate 2,6-diaminopimelate ligase of Corynebacterium glutamicum: effect on l-lysine formation and analysis of systemic consequences. *Biotechnol. Lett.* 39, 283−288.

(37) Yim, S. S., An, S. J., Kang, M., Lee, J., and Jeong, K. J. (2013) Isolation of fully synthetic promoters for high-level gene expression in corynebacterium glutamicum. *Biotechnol. Bioeng.* 110, 2959−2969.

(38) Swarts, D. C., van der Oost, J., and Jinek, M. (2017) Structural Basis for Guide RNA Processing and Seed-Dependent DNA Targeting by CRISPR-Cas12a. *Mol. Cell* 66, 221−233.

(39) Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A., and Charpentier, E. (2016) The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* 532, 517−21.

(40) Hirasawa, T., and Wachi, M. (2016) Glutamate Fermentation-2: Mechanism of L-Glutamate Overproduction in Corynebacterium glutamicum. *Adv. Biochem. Eng./Biotechnol.* 159, 57−72.

(41) Hoischen, C., and Krämer, R. (1990) Membrane alteration is necessary but not sufficient for effective glutamate secretion in Corynebacterium glutamicum. *J. Bacteriol.* 172, 3409−16.

(42) Nampoothiri, K. M., Hoischen, C., Bathe, B., Möckel, B., Pfefferle, W., Krumbach, K., Sahm, H., and Eggeling, L. (2002) Expression of genes of lipid synthesis and altered lipid composition modulates L-glutamate efflux of Corynebacterium glutamicum. *Appl. Microbiol. Biotechnol.* 58, 89−96.

(43) Nakayama, Y., Hashimoto, K.-I., Sawada, Y., Sokabe, M., Kawasaki, H., and Martinac, B. (2018) Corynebacterium glutamicum mechanosensitive channels: towards unpuzzling "glutamate efflux" for amino acid production. *Biophys. Rev.* 10, 1359−1369.

(44) Yamashita, C., Hashimoto, K.-I., Kumagai, K., Maeda, T., Takada, A., Yabe, I., Kawasaki, H., and Wachi, M. (2013) L-Glutamate secretion by the N-terminal domain of the Corynebacterium glutamicum NCgl1221 mechanosensitive channel. *Biosci., Biotechnol., Biochem.* 77, 1008−13.

### 8.1.1 CRISPR/ Cas12a mediated genome editing to introduce amino acid substitutions into the mechanosensitive channel MscCG of *Corynebacterium glutamicum* – Supplementary

# CRISPR/Cas12a mediated genome editing to introduce amino acid substitutions into the mechanosensitive channel MscCG of *Corynebacterium glutamicum*

Karin Krumbach[1], Christiane Katharina Sonntag[1], Lothar Eggeling[1], and Jan Marienhagen[1,2,*]

[1]Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich, D-52425 Jülich, Germany

[2]Institute of Biotechnology, RWTH Aachen University, Worringer Weg 3, D-52074 Aachen, Germany

\* To whom correspondence should be addressed.

Tel: +49 2461 61 2843; Fax: +49 2461 61 2710; Email: j.marienhagen@fz-juelich.de
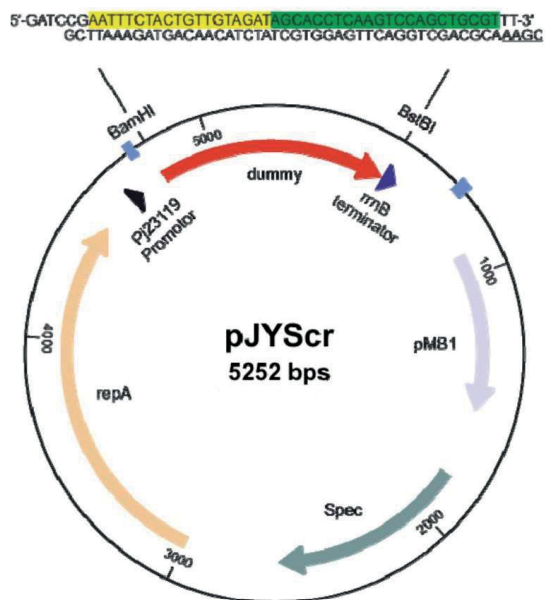
**Supplementary Figure S1:** Plasmid map of the pJYScr vector constructed in this study for easy incorporation of crRNA encoding oligonucleotides. The 931 *Bam*HI-*Bst*BI-"dummy" fragment with the relevant restrictions sites for easy subcloning is highlighted.

**Supplementary Table S1.** Plasmids and oligonucleotides used for plasmid constructions.

| Plasmids | Characteristics | Source |
|---|---|---|
| pK19mobsacB | Integrative vector, Km$^r$ *mob sacB* | (1) |
| pEKEx3 | Shuttle vector, Spc$^r$ P*tac lacI*$^q$ | (2) |
| pJYS2_crtYf | Shuttle vector based on pMB1, Spc$^r$, Pj23119-crRNAtargeting *crtYf* | (3) |
| pJYS1 | Shuttle vector based on pBL1, Km$^r$, RecT, *Fn*Cpf1, identical with pJYS1-petFu | (3) |
| pMK-RQ_PH36_LacZ | Vector providing *lacZ* under constant high expression of PH36 | invitrogen Thermo Fisher Scientific |
| pMA-RQ_crRNA_crtyf | Vector providing a 313 bp fragment in which a *Dra*I site of pJYS2_crtYf was replaced by a *Bst*BI site | invitrogen Thermo Fisher Scientific |

| Oligonucleotides | Sequence (5´-3´) | Purpose |
|---|---|---|
| cr:lacZ 1574f_ fw | CTGTTGTAGATGCTACCTGGAGAGACGC GCCCATTCGAAATAAAACGAAAGGCTC | pJYScr-lacZ-1574 |
| cr:lacZ 1574f _rv | CGTTTTATTTCGAATGGGCGCGTCTCTC CAGGTAGCATCTACAACAGTAGAAATTC G | pJYScr-lacZ-1574f |
| cr:lacZ1656c_fw | CTGTTGTAGATGCGAAACCGCCAAGACT GTTAATTCGAAATAAAACGAAAGGCTC | pJYScr-lacZ-1654c |
| cr:lacZ1656c _rev | CGTTTTATTTCGAATTAACAGTCTTGGCG GTTTCGCATCTACAACAGTAGAAATTCG | pJYScr-lacZ-1654c |
| cr:murE-2 _BamHI | GATCCGAATTTCTACTGTTGTAGATAGC ACCTCAAGTCCAGCTGCGTTT | pJYScr:murE |
| cr:murE-2_BstBI | CGAA ACGCAGCTGGACTTGAGGTGCTATCTAC AACAGTAGAAATTCG | pJYScr:murE |
| cr:mscCG_1269c_fw | GATCCGAATTTCTACTGTTGTAGATCAC AGTCATGACCTTAAATAGTTT | pJYScr_mscCG1269c |
| cr:mscCG_1269c_rv | CGAAACTATTTAAGGTCATGACTGTGAT CTACAACAGTAGAAATTCG | pJYScr_mscCG1269c |
| m420stopΔ-1fwd | tgcatgcctgcaggtcgactGGCGGATCGACCAC GGCT | pK19mobsacB-420stopΔ |
| m420stopΔ-2rev | cagcgtcctaGACCTTAAATAGTGACAGCAA | pK19mobsacB- |

| | GAGCAGC | 420stopΔ |
|---|---|---|
| m420stopΔ-3fwd | atttaaggtcTAGGACGCTGATTACAGAC | pK19mobsacB-420stopΔ |
| m420stopΔ-4rev | ttgtaaaacgacggccagtgGTAGCCGTCTTCTTGAAC | pK19mobsacB-420stopΔ |
| p424stopΔ-1fwd | cctgcaggtcgactctagagCCGGCAATCAATGGCTGG | pK19mobsacB-424stopΔ |
| p424stopΔ-2rev | AATCAGCGTCTTATTATTCCACAGTCATGACCTTAAATAGTG | pK19mobsacB-424stopΔ |
| p424stopΔ-3fwd | GGAATAATAAGACGCTGATTACAGACGTG | pK19mobsacB-424stopΔ |
| p424stopΔ-4rev | ttgtaaaacgacggccagtgGTAGCCGTCTTCTTGAAC | pK19mobsacB-424stopΔ |
| pJYS2_fw2 | GTGTCAGTGAAAGGCGCATCC | Sequencing primer |
| pJYS2_rv2 | TCGCCACCTCTGACTTGAGCG | Sequencing primer |

**Supplementary Table S2.** Recombinogenic oligos and obtained mutations at two *lacZ* locations. Shown on top is the *lacZ* sequence targeted for mutations using recombineering in combination with CRISPR/Cas12. The PAM sequence is underlined. Used recombinogenic oligonucleotides (o-1574x or o-1654x) have a grey background with the nucleotides deviating from the *lacZ* sequence in red. The DNA sequencing results for individual clones are given below the oligonucleotide sequences.

| *lacZ* | CC ATC AAA AAA TGG CT<u>T TC</u>G CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| --- | --- |
| o-1574a | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 1 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 2 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 3 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 4 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 5 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 6 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 7 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 8 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 8 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 9 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 10 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 11 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 12 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 13 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 14 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 15 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 16 | CC ATC AAA AAA TGG CTA TAG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| o-1574b | CC ATC AAA TAG TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 1 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 2 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 3 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 4 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 5 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 6 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 7 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 8 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| o-1574c | CC ATC AAA AAA TAG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 1 | CC ATC AAA AAA TAG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 2 | CC ATC AAA AAA TAG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 3 | CC ATC AAA AAA TAG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 4 | CC ATC AAA AAA TAG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 5 | CC ATC AAA AAA TAG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 6 | CC ATC AAA AAA TAG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 7 | CC ATC AAA AAA TAG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |

| o-1574fd | CC ATC AAA AAA TGG TAG TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
|----------|---|
| clone 1 | CC ATC AAA AAA TGG TAG TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 2 | CC ATC AAA AAA TGG TAG TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 3 | CC ATC AAA AAA TGG TAG TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 4 | CC ATC AAA AAA TGG TAG TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 5 | CC ATC AAA AAA TGG TAG TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 6 | CC ATC AAA AAA TGG TAG TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 7 | CC ATC AAA AAA TGG TAG TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 8 | CC ATC AAA AAA TGG TAG TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |

| o-1574fe | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
|----------|---|
| clone 1 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 2 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 3 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 4 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 5 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 6 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 7 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 8 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 9 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 10 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 11 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 12 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 13 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 14 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 15 | CC ATC AAA AAA TGG CTA TCG TAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |

| o-1574f | CC ATC AAA AAA TGG CTA TCG CTA TAG GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
|----------|---|
| clone 1 | CC ATC AAA AAA TGG CTA TCG CTA TAG GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 2 | CC ATC AAA AAA TGG CTA TCG CTA TAG GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 3 | CC ATC AAA AAA TGG CTA TCG CTA TAG GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 4 | CC ATC AAA AAA TGG CTA TCG CTA TAG GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 5 | CC ATC AAA AAA TGG CTA TCG CTA TAG GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 6 | CC ATC AAA AAA TGG CTA TCG CTA TAG GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 7 | CC ATC AAA AAA TGG CTA TCG CTA TAG GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 8 | CC ATC AAA AAA TGG CTA TCG CTA TAG GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 9 | CC ATC AAA AAA TGG CTA TCG CTA TAG GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 10 | CC ATC AAA AAA TGG CTA TCG CTA TAG GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 11 | CC ATC AAA AAA TGG CTA TCG CTA TAG GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |

| o-1574g | CC ATC AAA AAA TGG CTA TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
|----------|---|
| clone 1 | CC ATC AAA AAA TGG CTA TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 2 | CC ATC AAA AAA TGG CTA TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 3 | CC ATC AAA AAA TGG CTA TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 4 | CC ATC AAA AAA TGG CTA TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 5 | CC ATC AAA AAA TGG CTA TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |
| clone 6 | CC ATC AAA AAA TGG CTA TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC |

```
clone 7     CC ATC AAA AAA TGG CTA TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 8     CC ATC AAA AAA TGG CTA TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 9     CC ATC AAA AAA TGG CTA TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 10    CC ATC AAA AAA TGG CTT TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 11    CC ATC AAA AAA TGG CTT TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 12    CC ATC AAA AAA TGG CTA TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC


o-1574h     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA TAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 1     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA TAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 2     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA TAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 3     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA TAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 4     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA TAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 5     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA TAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 6     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA TAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 7     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA TAG ACG CGG CCG CTG ATC CTT TGC GAA TAC
clone 8     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA TAG ACG CGG CCG CTG ATC CTT TGC GAA TAC


o-1574i     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG TAG CGG CCG CTG ATC CTT TGC GAA TAC
clone 1     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG TAG CGG CCG CTG ATC CTT TGC GAA TAC
clone 2     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG TAG CGG CCG CTG ATC CTT TGC GAA TAC
clone 3     CC ATC AAA AAA TGG CTT TCG CTA CCT GGA GAG TAG CGG CCG CTG ATC CTT TGC GAA TAC
clone 4     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG TAG CGG CCG CTG ATC CTT TGC GAA TAC
clone 5     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG TAG CGG CCG CTG ATC CTT TGC GAA TAC
clone 6     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG TAG CGG CCG CTG ATC CTT TGC GAA TAC
clone 7     CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG TAG CGG CCG CTG ATC CTT TGC GAA TAC


lacZ        TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC GCT AAA TAC TGG CAG GCG TTT CG


o-1654a     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 1     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 2     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 3     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 4     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 5     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 6     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 7     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 8     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC GCT TAA TAC TGG CAG GCG TTT CG


o-1654ca    TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC TGA TAA TAC TGG CAG GCG TTT CG
clone 1     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC TGA TAA TAC TGG CAG GCG TTT CG
clone 2     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC TGA TAA TAC TGG CAG GCG TTT CG
clone 3     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC TGA TAA TAC TGG CAG GCG TTT CG
clone 4     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC TGA TAA TAC TGG CAG GCG TTT CG
clone 5     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC TGA TAA TAC TGG CAG GCG TTT CG
clone 6     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC TGA TAA TAC TGG CAG GCG TTT CG
clone 7     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC TGA TAA TAC TGG CAG GCG TTT CG
```

```
clone 8      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC TGA TAA TAC TGG CAG GCG TTT CG

o-1654cb     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TGA GCT TAA TAC TGG CAG GCG TTT CG
clone 1      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TGA GCT TAA TAC TGG CAG GCG TTT CG
clone 2      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TGA GCT TAA TAC TGG CAG GCG TTT CG
clone 3      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TGA GCT TAA TAC TGG CAG GCG TTT CG
clone 4      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TGA GCT TAA TAC TGG CAG GCG TTT CG
clone 5      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TGA GCT TAA TAC TGG CAG GCG TTT CG
clone 6      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TGA GCT TAA TAC TGG CAG GCG TTT CG
clone 7      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TGA GCT TAA TAC TGG CAG GCG TTT CG
clone 8      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TGA GCT TAA TAC TGG CAG GCG TTT CG

o-1654cc     TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC TGA TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 1      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC TGA TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 2      TAC GCC CAC GCG ATG GGT AAC AG- CTT GGC GGT TTC GCT AAA TAC TGG CAG GCG TTT CG
clone 3      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC TGA TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 4      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC TGA TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 5      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC TGA TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 6      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC TGA TTC GCCT TAA TAC TGG CAG GCG TTT CG
clone 7      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC TGA TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 8      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC TGA TTC GCT TAA TAC TGG CAG GCG TTT CG

o-1654cd     TAC GCC CAC GCG ATG GGT AAC AGT CTT TGA GGT TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 1      TAC GCC CAC GCG ATG GGT AAC AGT CTT TGA GGT TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 2      TAC GCC CAC GCG ATG GGT AAC AGT CTT TGA GGT TTC GCT AAA TAC TGG CAG GCG TTT CG
clone 3      TAC GCC CAC GCG ATG GGT AAC AGT CTT TGA GGT TTC GCT AAA TAC TGG CAG GCG TTT CG
clone 4      TAC GCC CAC GCG ATG GGT AAC AGT CTT TGA GGT TTC GCT AAA TAC TGG CAG GCG TTT CG
clone 5      TAC GCC CAC GCG ATG GGT AAC AGT CTT TGA GGT TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 6      TAC GCC CAC GCG ATG GGT AAC AGT CTT TGA GGT TTC GCT AAA TAC TGG CAG GCG TTT CG
clone 7      TAC GCC CAC GCG ATG GGT AAC AGT CTT GGC GGT TTC GCT TAA TAC TGG CAG GCG TTT CG
clone 8      TAC GCC CAC GCG ATG GGT AAC AGT CTT TGA GGT TTC GCT AAA TAC TGG CAG GCG TTT CG
```

8

**Supplementary Table S3.** Recombinogenic oligos for the insertion of a random stretch of nucleotides and codon saturation mutagenesis. On top always the lacZ sequence is shown, targeted by recombineering in combination with CRISPR/Cas12a. The PAM sequence is underlined. Used recombinogenic oligonucleotides have a grey background with the nucleotides deviating from the *lacZ* sequence in red. The DNA sequencing results for individual clones are given below the oligonucleotide sequences. With the oligonucleotide o-1574fe-random carrying a stretch of 18 random nucleotides, only few white colonies were obtained (see results section in the main text). This was also the case with an additional, independently synthesized oligonucleotide with identical sequence. Noteworthy, clones 7 and 8 obtained during experiments with this oligonucelotide appeared white, but do not carry any additional mutation near the PAM site. It is possible that these two clones carry additional unwanted mutations in the *lacZ* gene. However, the nucleotide sequence nt 1188 to nt 1902 of these two clones was not at variance from the wild type sequence. Results obtained during codon saturation mutagenesis (oligonucleotides o-1574e-mix and o-1574g-mix), the amino acid substitution obtained is indicated in addition to the nucleotide sequence.

| | Sequence | amino acid |
|---|---|---|
| *lacZ* | CC ATC AAA AAA TGG CTT TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| o-lacZ_1574fe-random | CC ATC AAA AAA TCC GAA AGC GAT GGA TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 1 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG TAA CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 2 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA TAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 3 | CC ATC AAA AAA TGG TAG TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 4 | CC ATC AAA AAA TGG TAA TCG CTA CCT GGA AAG ACA CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 5 | CC ATC AAA AAA TGG TAG TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 6 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA TCG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 7 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 8 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| | | amino acid |
| o-1574e-mix | CC ATC AAA AAA TGG CTA TCG NNN CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 1 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 2 | CC ATC AAA AAA TGG CTA TCG GCC CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527A |
| clone 3 | CC ATC AAA AAA TGG CTA TCG AGA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527R |
| clone 4 | CC ATC AAA AAA TGG CTA TCG AAC CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527N |
| clone 5 | CC ATC AAA AAA TGG CTA TCG TAC CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527Y |
| clone 6 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 7 | CC ATC AAA AAA TGG CTA TCG TCT CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527S |
| clone 8 | CC ATC AAA AAA TGG CTA TCG ATT CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527I |
| clone 9 | CC ATC AAA AAA TGG CTA TCG TAA CCT GGA --- ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527stop |
| clone 10 | CC ATC AAA AAA TGG CTA TCG ACC CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527T |
| clone 11 | CC ATC AAA AAA TGG CTA TCG TCT CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527S |
| clone 12 | CC ATC AAA AAA TGG CTA TCG GGG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527G |
| clone 13 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 14 | CC ATC AAA AAA TGG CTA TCG TGT CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527C |
| clone 15 | CC ATC AAA AAA TGG CTA TCG CAG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527Q |
| clone 16 | CC ATC AAA AAA TGG C-A TCG GCG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 17 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 18 | CC ATC AAA AAA TGG CTA TCG GCG CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527A |

| clone 19 | CC ATC AAA AAA TGG CTA TCG ATC CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527I |
|---|---|---|
| clone 20 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 21 | CC ATC AAA AAA TGG CTT TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 22 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 23 | CC ATC AAA AAA TGG CTA TCG TAT CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | L527Y |
| clone 24 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTGG ATC CTT TGC GAA TAC | |

| o-1574g-mix | CC ATC AAA AAA TGG CTA TCG CTA CCT NNN GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
|---|---|---|
| clone 1 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGC GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529G |
| clone 2 | CC ATC AAA AAA TGG CTA TCG CTA CCT AAA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529K |
| clone 3 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 4 | CC ATC AAA AAA TGG CTA TCG CTA CCT AGC GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529S |
| clone 5 | CC ATC AAA AAA TGG CTA TCG CTA CCT GAA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529E |
| clone 6 | CC ATC AAA AAA TGG CTA TCG CTA CCT TAC GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529Y |
| clone 7 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 8 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 9 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 10 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 11 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 12 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 13 | CC ATC AAA AAA TGG CTA TCG CTA CCT TTA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529L |
| clone 14 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 15 | CC ATC AAA AAA TGG CTA TCG CTA CCT AGC GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529S |
| clone 16 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 17 | CC ATC AAA AAA TGG CTA TCG CTA CCT TAG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529stop |
| clone 18 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 19 | CC ATC AAA AAA TGG CTA TCG CTA CCT GTA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529V |
| clone 20 | CC ATC AAA AAA TGG CTA TCG CTA CCT GGA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | |
| clone 21 | CC ATC AAA AAA TGG CTA TCG CTA CCT TTG GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529L |
| clone 22 | CC ATC AAA AAA TGG CTA TCG CTA CCT TAA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529stop |
| clone 23 | CC ATC AAA AAA TGG CTA TCG CTA CCT GAA GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529E |
| clone 24 | CC ATC AAA AAA TGG CTA TCG CTA CCT TGC GAG ACG CGG CCG CTG ATC CTT TGC GAA TAC | G529C |

10

**Supplementary Table S4.** Recombinogenic oligonucleotides used for introducing mutations in *mscCG* and the mutations obtained. Shown on top is the targeted *mscCG* sequence with the PAM sequence underlined (PAM sequence is on the opposite strand). The recombinogenic oligonucleotides o-mscCG1269-V422nnn and o-mscCG1xxx-E423nnn and o-mscCG1269-V422nnn and o-mscCG1xxx-E423 used are highlighted by a grey background. The DNA sequencing results for all individual clones with nucleotides deviating from the wild type sequence are presented in red. In the right column the relevant amino acid sequences (amino acids 420 - 426) are shown.

| Name | DNA sequence | Amino acids |
|---|---|---|
| *mscCG* | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG GAA CCA AGT GAG AAT TGG CAA | – Met Thr Val Glu Pro Ser Glu – |
| o-mscCG1269-V422nnn | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT nnn GAG CCA AGT GAG AAT TGG CAA | |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT tTa GAG CCA AGT GAG AAT TGG CAA | – Met Thr Leu Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GCT GAG CCA AGT GAG AAT TGG CAA | – Met Thr Ala Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT AAA GAG CCA AGT GAG AAT TGG CAA | – Met Thr Lys Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT TGC GAG CCA AGT GAG AAT TGG CAA | – Met Thr Cys Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT ACT GAG CCA AGT GAG AAT TGG CAA | – Met Thr Thr Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT CTC GAG CCA AGT GAG AAT TGG CAA | – Met Thr Leu Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT TTA GAG CCA AGT GAG AAT TGG CAA | – Met Thr Leu Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GAT GAG CCA AGT GAG AAT TGG CAA | – Met Thr Asp Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT TCG GAG CCA AGT GAG AAT TGG CAA | – Met Thr Ser Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT ACT GAG CCA AGT GAG AAT TGG CAA | – Met Thr Thr Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT ACT GAG CCA AGT GAG AAT TGG CAA | – Met Thr Thr Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT CTC GAG CCA AGT GAG AAT TGG CAA | – Met Thr Leu Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT ACA GAG CCA AGT GAG AAT TGG CAA | – Met Thr Thr Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT TGC GAG CCA AGT GAG AAT TGG CAA | – Met Thr Cys Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT ACT GAG CCA AGT GAG AAT TGG CAA | – Met Thr Thr Glu Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT CTG GAG CCA AGT GAG AAT TGG CAA | – Met Thr Leu Glu Pro Ser Glu – |
| *mscCG* | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG GAA CCA AGT GAG AAT TGG CAA | – Met Thr Val Glu Pro Ser Glu – |
| o-mscCG1269-E423nnn | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG nnn CCA AGT GAG AAT TGG CAA | |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG CCG CCA AGT GAG AAT TGG CAA | – Met Thr Val Pro Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG AAC CCA AGT GAG AAT TGG CAA | – Met Thr Val Asn Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG ACC CCA AGT GAG AAT TGG CAA | – Met Thr Val Thr Pro Ser Glu – |
| | CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG TCC CCA AGT GAG AAT TGG CAA | – Met Thr Val Ser Pro Ser Glu – |

11

```
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG GAC CCA AGT GAG AAT TGG CAA - Met Thr Val Asp Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG TAA CCA AGT GAG AAT TGG CAA - Met Thr Val STOP
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG GAA CCA AGT GAG AAT TGG CAA - Met Thr Val Glu Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG TCA CCA AGT GAG AAT TGG CAA - Met Thr Val Ser Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GT- -AG CTC AGT GAG AAT TGG TGA - Met Thr Val Ala Gln STOP
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG ACA CCA AGT GAG AAT TGG CAA - Met Thr Val Thr Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG GCA CCA AGT GAG AAT TGG CAA - Met Thr Val Ala Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG GAC CCA AGT GAG AAT TGG CAA - Met Thr Val Asp Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG AAC CCA AGT GAG AAT TGG CAA - Met Thr Val Asn Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG AAA CCA AGT GAG AAT TGG CAA - Met Thr Val Lys Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG GCC CCA AGT GAG AAT TGG CAA - Met Thr Val Ala Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG GCC CCA AGT GAG AAT TGG CAA - Met Thr Val Ala Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG CAC CCA AGT GAG AAT TGG CAA - Met Thr Val His Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG CAC CCA AGT GAG AAT TGG CAA - Met Thr Val His Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG GAC CCA AGT GAG AAT TGG CAA - Met Thr Val Asp Pro Ser Glu -
CG CTG CTC TTG CTG TCA CTA TTT AAG GTC ATG ACT GTG GCA CCA AGT GAG AAT TGG CAA - Met Thr Val Ala Pro Ser Glu -
```

**Supplementary Table S5.** L-Glutamate accumulation in the culture supernatant of constructed *mscCG* deletion mutants.

| | Sampling time [h] and L-glutamate concentration [mM] | | | | | |
|---|---|---|---|---|---|---|
| *C. glutamicum* variant | 14 h | 16 h | 18 h | 20 h | 22 h | 24 h |
| wild type | 0 | 0 | 0 | 0 | 0 | 0 |
| M420stopΔ  Clone 1 | 27.1 | | 24 | 24.6 | 2.7 | |
| M420stopΔ  Clone 2 | 35.2 | 45.3 | 32.2 | 15.1 | 5.7 | 0.3 |
| P424stopΔ  Clone 1 | 42.4 | 89.1 | 82.6 | 61.4 | 53.6 | 59.5 |
| P424stopΔ  Clone 2 | 57.5 | 135.1 | 103.8 | 67.8 | 74.3 | 85.2 |
| P424stopΔ  Clone 3 | 62.1 | 97 | 91.1 | 47.7 | 57.4 | 33 |

# References

(1)     Schäfer, A.; Tauch, A.; Jäger, W.; Kalinowski, J.; Thierbach, G.; Pühler, A. Small Mobilizable Multi-Purpose Cloning Vectors Derived from the Escherichia Coli Plasmids PK18 and PK19: Selection of Defined Deletions in the Chromosome of Corynebacterium Glutamicum. *Gene* **1994**, *145* (1), 69–73. https://doi.org/10.1016/0378-1119(94)90324-7.

(2)     Hoffelder, M.; Raasch, K.; van Ooyen, J.; Eggeling, L. The E2 Domain of OdhA of Corynebacterium Glutamicum Has Succinyltransferase Activity Dependent on Lipoyl Residues of the Acetyltransferase AceF. *J. Bacteriol.* **2010**, *192* (19), 5203–5211. https://doi.org/10.1128/JB.00597-10.

(3)     Jiang, Y.; Qian, F.; Yang, J.; Liu, Y.; Dong, F.; Xu, C.; Sun, B.; Chen, B.; Xu, X.; Li, Y.; et al. CRISPR-Cpf1 Assisted Genome Editing of Corynebacterium Glutamicum. *Nat. Commun.* **2017**, *8*, 15179. https://doi.org/10.1038/ncomms15179.

## 8.2   A unified design allows fine-tuning of biosensor parameters and application across bacterial species

# A unified design allows fine-tuning of biosensor parameters and application across bacterial species

Check for updates

Christiane Katharina Sonntag [a], Lion Konstantin Flachbart [a], Celine Maass [a], Michael Vogt [a], Jan Marienhagen [a,b,*]

[a] Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich, D-52425, Jülich, Germany
[b] Institute of Biotechnology, RWTH Aachen University, Worringer Weg 3, D-52074, Aachen, Germany

ABSTRACT

In recent years, transcriptional biosensors have become valuable tools in metabolic engineering as they allow semiquantitative determination of metabolites in single cells. Although being perfectly suitable tools for high-throughput screenings, application of transcriptional biosensors is often limited by the intrinsic characteristics of the individual sensor components and their interplay. In addition, biosensors often fail to work properly in heterologous host systems due to signal saturation at low intracellular metabolite concentrations, which typically limits their use in high-level producer strains at advanced engineering stages.

We here introduce a biosensor design, which allows fine-tuning of important sensor parameters and restores the sensor response in a heterologous expression host. As a key feature of our design, the regulator activity is controlled through the expression level of the respective gene by different (synthetic) constitutive promoters selected for the used expression host. In this context, we constructed biosensors responding to basic amino acids or ring-hydroxylated phenylpropanoids for applications in *Corynebacterium glutamicum* and *Escherichia coli*. Detailed characterization of these biosensors in liquid cultures and during single-cell analysis using flow cytometry showed that the presented sensor design enables customization of important biosensor parameters as well as application of these sensors in relevant heterologous hosts.

## 1. Introduction

Metabolic engineering of microorganisms is still a very laborious task and despite detailed knowledge of the microbial metabolism, sophisticated rational engineering strategies are often flanked by traditional random mutagenesis and screening approaches. Methods for random mutagenesis of genes and genomes are well-established and very large strain libraries with a diverse genetic background can be easily generated. However, the production phenotype of the generated strain variants has to be investigated manually, which usually requires individual cultivation of each variant and costly analytical techniques (i.e. HPLC or GC) (Eggeling et al., 2015). A few years ago, transcriptional biosensors emerged as powerful high-throughput screening tools as they allow for the conversion of an intracellular metabolite concentration to a machine-readable fluorescence output signal at the single-cell level (Schallmey et al., 2014). In combination with fluorescence activated cell sorting (FACS), producing cells can be directly isolated from a strain library eliminating the need for individual cultivation of every variant

prior to screening (Binder et al., 2012; Siedler et al., 2014). Considering these advantages, it is not surprising that transcriptional biosensors are widely applied in several areas of research in academia and industry (Schallmey et al., 2014; Zhang et al., 2015).

Transcriptional biosensors are based on simple regulatory genetic circuits comprised of a transcriptional regulator, its cognate target promoter, and a reporter gene, *e.g.* coding for a fluorescent protein or a positive selection marker, which is under control of the cognate promoter (Binder et al., 2012). The transcriptional regulator specifically recognizes the inducer molecule of interest (ligand) as it binds to a dedicated ligand binding domain. This induces a conformational change of the transcriptional regulator. Depending on the regulator type, either regulator-binding to or regulator-release from its target operator site is induced. Both modes of action ultimately lead to initiation of transcription of the reporter gene (Mahr and Frunzke, 2016). A large variety of transcriptional regulator/promoter pairs have already been used for the construction of many different biosensors ranging from amino acids to more complex molecules such as antibiotics (Binder et al., 2012; Rebets

et al., 2018). Recently, constructed transcriptional biosensors illustrate the broad field of different sensor applications e. g. a BenM-based transcriptional biosensor, which allowed for selection of high *cis,cis*-muconic acid producer strains in yeast or a set of protocatechuate (3,4-dihydroxybenzoate) sensors based on the transcriptional regulator PcaU, which was applied to optimize screening applications in *Pseudomonas putida* (Jha et al., 2018; Snoek et al., 2018).

However, when a genetic circuit (regulator and target promoter) is directly turned into a biosensor, sensor characteristics such as the operational- and dynamic range (Fig. 1B), are predetermined and cannot be adjusted individually for specific sensor applications. The defined native dynamic and operational range assigns the field of possible biosensor applications to either basic research, monitoring of environmental pollution (low inducer concentrations), or strain engineering in an industrial setting (high inducer concentrations) (Schallmey et al., 2014; Webster et al., 2014). Furthermore, application of transcriptional

biosensors can also be limited by their exclusive functionality in the native microorganism from which the sensor components are derived from (Mannan et al., 2017). For example, insertion of the regulatory circuit as part of *cis,cis*-muconic acid biosensor from *Acinetobacter* sp. ADP1 in yeast required insertion and adaption of a constitutive yeast promoter to establish a functional phenylpropanoic acid sensor in this organism (Skjoedt et al., 2016). Similarly, a 3,4-dihydroxybenzoate biosensor, which was originally also constructed in *Acinetobater* sp ADP1 was not functional in *E. coli* until the promoter controlling the expression of the transcriptional regulator PcaU of the biosensor was adapted to the new expression host (Jha et al., 2014).

With the aim to address these biosensor limitations and to improve the applicability of biosensors across different species, we developed a unified biosensor design, which allows for fine-tuning of important sensor parameters. As key feature, this design allows for exchanging the original promoter of the transcriptional regulator gene to different
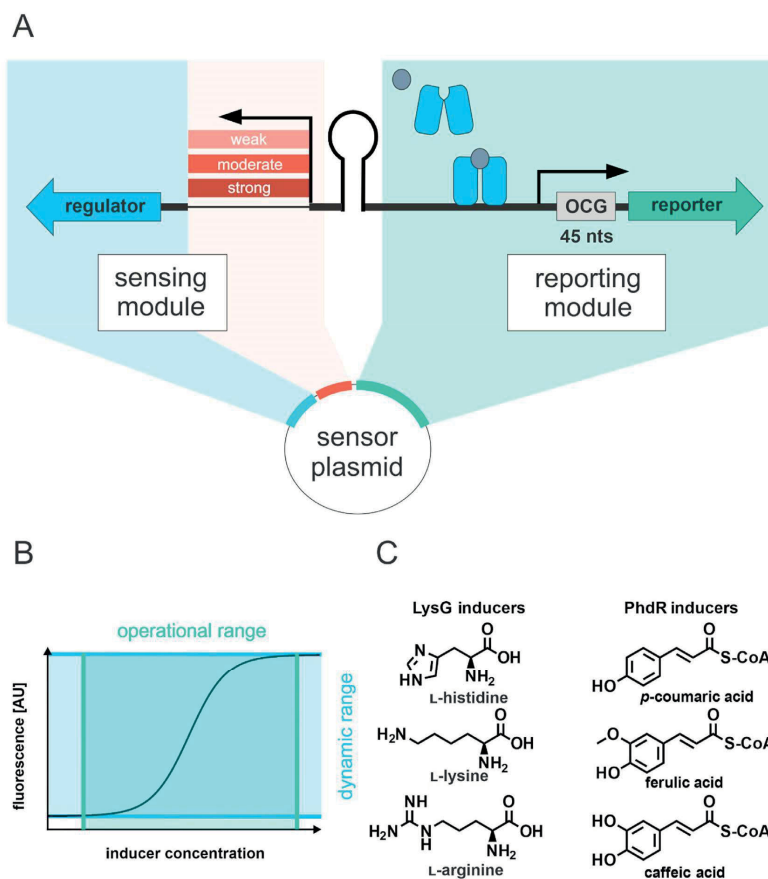


**Fig. 1. Unified biosensor design.** (A) Schematic modular sensor architecture. Expression of the transcriptional regulator (blue arrow) is under control of a native or synthetic constitutive promoter. Upon binding of an inducer molecule, the transcriptional regulator undergoes a conformational and binds to the operator site in the cognate promoter including the first 45 nucleotides of the originally controlled gene (OCG) to promote expression of a reporter gene (green arrow). (B) Definition of operational range (green) and dynamic range (blue) based on the sensor output signal. (C) Ligands of PhdR- and LysG-based biosensors. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(synthetic) constitutive promoters preselected for different host systems. With this design, we successfully constructed different transcription factor-based biosensors for amino acids and phenylpropanoids, which can be directly used for applications in the industrially important microbial workhorses *C. glutamicum* and *E. coli*. Demonstrating the versatility of this concept, we included both, transcriptional activator- and repressor-based biosensors, in our study.

## 2. Materials and methods

### 2.1. Plasmid and strain construction

Recombinant DNA work was performed according to standard protocols of molecular cloning such as polymerase chain reaction (PCR), DNA restriction and ligation (Sambrook and Russel, 2001). Electroporation for transformation of *C. glutamicum* strains was carried out as described previously (Eggeling and Bott, 2005). Restriction enzymes were obtained from Thermo Fisher Scientific (Waltham, MA, USA). Genes, promoter- and terminator sequences were amplified by PCR using Novagen KOD Polymerase (Merck KGaA, Darmstadt, Germany). Cloning of amplified PCR products was performed using classical DNA restriction ligation or Gibson assembly (Gibson et al., 2009). For plasmids containing more than one insert or small inserts (<500 bp), fragments were joined by overlap-extension PCR. In-frame deletion of genes in the *C. glutamicum* genome was carried out with the pK19mobsacB system by a two-step homologous recombination method (Niebisch and Bott, 2001; Schäfer et al., 1994). Both, synthesis of oligonucleotides and DNA sequencing using Sanger sequencing, was performed by Eurofins MWG Operon (Ebersberg, Germany). The sequence of all oligonucleotides used in this work are listed in Table S3.

### 2.2. Bacterial strains, media and growth conditions

All bacterial strains and plasmids used in this study and their relevant characteristics are listed in Tables S1 and S2, respectively. *E. coli* DH5α was used for plasmid constructions and was cultivated in Luria-Bertani (LB) medium at 37 °C (Bertani, 1951). Where appropriate, kanamycin (25 µg/mL) or spectinomycin (100 µg/mL) was added to the medium. Bacterial growth was followed by measuring the optical density at 600 nm ($OD_{600}$). *E. coli* DH10B-based strains were cultivated at 37 °C in LB medium or yeast nitrogen base (YNB) medium (Merck KGaA, Darmstadt, Germany) containing kanamycin (15–25 µg/mL) and spectinomycin (100 µg/mL). For the preparation of 1 L of YNB medium, 100 mL of 10xYNB, containing 5.1% (v/v) glycerol was added to 900 mL YNB base solution. YNB base solution contained 6 g/L $K_2HPO_4$, 3 g/L $KH_2PO_4$ and 10 g/L 3-(*N*-morpholino)propanesulfonic acid (MOPS), pH 7. Due to L-leucine auxotrophy of *E. coli* DH10B, L-leucine was supplemented to a final concentration of 2 mM to YNB medium (Durfee et al., 2008). The *C. glutamicum* ATCC 13032 wild type (American Type Culture Collection) and derived strains were cultivated at 30 °C in brain heart infusion (BHI) medium (Difco Laboratories, Detroit, MI, USA) or defined CGXII medium with 2% (w/v) glucose as sole carbon and energy source (Abe et al., 1967; Keilhauer et al., 1993). For plasmid maintenance in the respective strains, kanamycin (25 µg/mL) was supplemented.

Constructed strains harboring LysG- or PhdR-based biosensors were characterized with respect to their biomass and fluorescence response using the BioLector cultivation system (m2p-labs GmbH, Baesweiler, Germany) (Funke et al., 2009; Kensy et al., 2009). For this purpose, *E. coli*- and *C. glutamicum* sensor strains, were grown on a rotary shaker at 170 rpm for 6–8 h in reaction tubes with 5 mL rich medium (either LB (*E. coli*) or BHI (*C. glutamicum*)), which was supplemented with antibiotics where appropriate. Precultures were used to inoculate 15 mL defined medium (YNB (*E. coli*) or CGXII (*C. glutamicum*)) in 50 mL baffled flasks, which were subsequently cultivated overnight on a rotary shaker at 120 rpm. Main cultures, cultivated in 48-well microtiter FlowerPlates (MFPs) using the BioLector cultivation system in YNB or CGXII medium

were inoculated to an $OD_{600}$ of 0.1 and 1.0 for *E. coli* and *C. glutamicum*, respectively. In PhdR-based sensor strains, heterologous expression of the 4-coumarate: CoA ligase (4CL)-encoding gene was induced using 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) in *C. glutamicum* DelAro[4]-4$cl_{Pc}$ ΔphdR or 0.002%–0.02% L-arabinose in *E. coli* DH10B immediately after inoculation (Kallscheuer et al., 2016b). The operational range of all biosensors was investigated with various inducers, basic amino acids for LysG-based biosensors (either L-histidine, L-lysine or L-arginine for *E. coli* or the respective alanine dipeptides in *C. glutamicum*) and ring-hydroxylated phenylpropanoids for PhdR-based biosensors (either *p*-coumaric acid, caffeic acid or ferulic acid dissolved in dimethyl sulfoxide (DMSO)). These molecules of interest were supplemented at different final concentrations in the cultivation medium in the flower plates before inoculation of the main cultures. Cell growth was measured by backscatter light intensity (wavelength 620 nm; signal gain of 25 (*E. coli*) or 10 (*C. glutamicum*)). The fluorescence signal of the enhanced yellow fluorescence protein (eYFP) was measured as fluorescence emission at a wavelength of 532 nm (signal gain factor of 30) after excitation at 510 nm. Specific fluorescence was calculated as ratio of 532 nm fluorescence to 620 nm backscatter using BioLection software version 2.2.0.6 (m2p-labs GmbH, Baesweiler, Germany). Engineered biosensors were assessed by the sensor response given as fold-induction of the specific eYFP fluorescence. The basal specific fluorescence signal in cultures, corresponding to the sensor output signal at an uninduced state, was used to normalize the specific fluorescence response at different inducer concentrations.

Based on the resulting fold-induction at different inducer concentrations, important sensor parameters such as operational- and dynamic range were assessed. The operational range is defined as the range of inducer concentrations for which a change in fluorescent output signal is given until a plateau or maximum is reached. The dynamic range refers to the maximal fluorescence response signal of the biosensor in relation to the basal fluorescence response in an uninduced state. Dynamic range and the operational range of all tested sensors were determined on the basis of sensor response functions obtained from dose-response experiments. The dynamic range is given by maximal fold-induction reached. The operational range was determined in accordance with the fold-induction signal. As bottom value the concentration was used where 2 fold-induction was exceeded. The concentration at which the highest fold-induction was reached was used as top value.

### 2.3. Flow cytometry and data analysis

The specific fluorescence was followed during cultivations in 48-well plates using the BioLector cultivation system. After a global maximum or plateau was reached for the highest inducer concentration, the cultivation was paused to collect samples from all cultures containing the respective inducer to analyze the biosensor fluorescence response on a single-cell level by flow cytometry. Such experiments were performed using a BD FACSAria II cell sorter (BD Biosciences, Franklin Lakes, NJ, USA) equipped with a 70 µm nozzle run with a sheath pressure of 70 psi. A 488 nm blue solid laser was used for excitation. Front scatter (FSC) was recorded as small-angle (axial) scatter and side-scatter (SSC) was recorded as perpendicular scatter of the 488-nm laser beam. Detection of emitted eYFP fluorescence from the SSC signal was achieved by combining a 502 nm long-pass and 530/30 nm band-pass filter. Acquisition of fluorescence data was always performed following a two-step gating strategy: By using a FSC-H versus SSC-H plot, a first population was gated to exclude signals from cell debris and electronic noise. From the resulting population, the FSC-H signal was plotted against FSC-W to perform singlet discrimination. The gated singlet population was used for fluorescence acquisition in all experiments. The afore mentioned gating strategy for fluorescence measurement of single cells was used for both, *E. coli* and *C. glutamicum*. For each biosensor variant bearing strain, the fluorescence signal of 100,000 events at different inducer concentrations was routinely recorded. The total event rate during measurements never

exceeded 15,000–25,000 events per second for *E. coli* and *C. glutamicum* strains, respectively. FACSDiva 7.0.1 software (BD Biosciences, San Jose, CA, USA) was used for FACS device control and data analysis. Gated events ($n = 95{,}000$) were used for data analysis in FlowJo for Windows 10.4.2 (FlowJo, LLC, Ashland, OR, USA) and Prism 7.04 (GraphPad Software, San Diego, CA, USA) to visualize FACS data.

## 3. Results & discussion

### 3.1. A unified biosensor design enables control over important sensor parameters

Key to modulating the response of a transcriptional biosensor is the gradual and constitutive expression of the gene of the transcriptional regulator (Lin et al., 2018; Rogers et al., 2015; Skjoedt et al., 2016; Zhang et al., 2017). Hence, a unified sensor design was established to control the expression level of the transcriptional regulator in a direct manner to influence the dynamic- and operational range of the transcriptional biosensor (Fig. 1B and S1). The overall sensor architecture can be separated into two modules: The regulatory module which encompasses the expression of the transcriptional regulator, and the sensing module in which the expression of the reporter gene *eyfp* is under control of a regulated promoter (Fig. 1A). In the regulatory module, a small selection of constitutive promoters from a previously characterized promoter library was used to compare effects of different transcriptional regulator levels on the sensor response. Since the biosensor design is evaluated in both, *E. coli* and *C. glutamicum*, two constitutive promoter libraries were chosen for the two host organisms. The well-characterized PLTetO1 promoter library was selected for *E. coli*, as this library was previously used for the successful construction of an arsenite-responsive biosensor (Alper et al., 2006; Merulla et al., 2013). In *C. glutamicum*, no such experience could be drawn on, as no studies have been carried out on changes in sensor architecture and the resulting sensor response, yet. Therefore, a well-characterized library of different constitutive *dapA* promoter variants was used to fine-tune the expression level of the regulator gene. Originally, these promoter variants were constructed during characterization of the $-10$ region of the promoter of the *dapA* gene, which encodes for dihydrodipicolinate synthase in *C. glutamicum* (Bonnassie et al., 1990; Vašicová et al., 1999). In this context, chloramphenicol acetyltransferase (CAT) assays were performed to quantify and compare the promoter activity of more than 20 mutated promoter variants. The results showed that the promoter activity covers a range from 5 to 500% when compared to the wild-type *dapA* promoter. Noteworthy, variants of this *dapA* promoter library were successfully used to down-regulate the expression of the citrate synthase gene *gltA* in the context of ʟ-lysine production with *C. glutamicum* (van Ooyen et al., 2012).

From both libraries, strong (S), moderate (M) and weak (W) constitutive promoters were selected to examine the effects of promoter strength on the fluorescence response (Table S4). In the sensing module, the regulated promoter and the first 45 nucleotides of the open reading frame of the regulated gene in the original genetic circuit were always included in order to not impair the stability of any mRNA folding, which could have a negative effect on translation initiation (Fig. 1A) (Kudla et al., 2009). Hence, this sequence was always inserted upstream of the reporter gene followed by a stop codon and an additional ribosome binding site (RBS) (AAGGAG-N$_{6.7}$) in front of the reporter gene start codon (Fig. 1A). By adding this RBS sequence, insufficient promoter activity can be circumvented as it has been described in the context of previous biosensor studies in *C. glutamicum* (Binder et al., 2012). In addition, the often divergently orientated promoter architecture of a cognate promoter might cause an undesired transcriptional read-through leading to uncontrolled expression of the regulator gene (Maddocks and Oyston, 2008). These effects can be minimized or averted by insertion of a terminator sequence between regulated and constitutive promoter (De Paepe et al., 2018; Rogers et al., 2015). Eventually, the sensing and

regulatory module was integrated in divergent orientation in medium-copy vector backbones: pJC1 in case of *C. glutamicum*, and pBR322 in case of *E. coli*. Both vector backbones were selected as they represent a low metabolic burden for the respective host organism (Wu et al., 2016).

### 3.2. Construction of PhdR-based transcriptional biosensors in C. glutamicum

As first example, a set of transcriptional biosensors using the repressor PhdR and its regulated promoter P$_{phdB}$ was constructed. In *C. glutamicum*, PhdR naturally represses the expression of genes involved in phenylpropanoid degradation in absence of any phenylpropanoids, which can be readily taken up and utilized as sole carbon and energy source by this bacterium (Kallscheuer et al., 2016a). PhdR derepression is specifically initiated by CoA-activated ring-hydroxylated phenylpropanoids such as *p*-coumaroyl-CoA. The proposed mode of action of PhdR is typical for MarR-type regulators (Otani et al., 2016). Upon binding of a CoA-activated phenylpropanoid, PhdR undergoes a conformational change and is unable to bind to its operator site, which in turn promotes binding of the RNA polymerase to P$_{phdB}$, the promoter of the regulated *phd* operon (Kallscheuer et al., 2016a).

For the construction of the native sensor plasmid pSenPhdR, the sequence of *phdR* and the region covering P$_{phdB}$ and 45 nucleotides of the *phdB* coding sequence were amplified from the genome of *C. glutamicum*. As the exact position of promoter elements in the intergenic region is not known, the information was deduced from the homologous region coding for the MarR-type regulator CouR and the CouR-controlled catabolic genes in *Rhodopseudomonas palustris* and *Rhodococcus jostii* RHA1 (Hirakawa et al., 2012; Otani et al., 2016). The resulting plasmid pSenPhdR was used to compare the sensor response of the native regulatory circuit to the biosensors constructed according to the unified biosensor design. These included derivatives of pSenPhdR, in which the expression of *phdR* is under control of either a weak (W), moderate (M) or strong (S) constitutive promoter (pSC$_{Cg}$-PhdR-W, pSC$_{Cg}$-PhdR-M, pSC$_{Cg}$-PhdR-S). *C. glutamicum* DelAro$^4$-4cl$_{Pc}$ was selected as host strain, as it is deficient in phenylpropanoid degradation and additionally harbors an IPTG-inducible plant-derived 4cl gene encoding a 4-coumarate: CoA-ligase necessary for CoA-activation of phenylpropanoids (Kallscheuer et al., 2016b). Both features are a prerequisite to ensure sufficient intracellular levels of the PhdR inducer *p*-coumaroyl-CoA. This is particularly important since the intracellular *p*-coumaroyl-CoA level should reflect the different extracellular concentrations of supplemented *p*-coumaric acid in biosensor characterization experiments. Since the chromosomal copy of *phdR* may interfere with the sensor response, sole episomal expression of *phdR* would be beneficial for a stringent control of the sensor response. Therefore, the chromosomal *phdR* gene was deleted yielding *C. glutamicum* DelAro$^4$-4cl$_{Pc}$ Δ*phdR* as strain background for all subsequent experiments.

### 3.3. The expression level of phdR determines the biosensor response

The influence of inducer concentration and repressor level on the fluorescence response of PhdR-based transcriptional biosensors was first evaluated in dose-response experiments with *p*-coumaric acid (Fig. S2). Interestingly, performed experiments for all PhdR-based biosensor variants revealed that the expression strength of the repressor gene correlates negatively with the dynamic range. The constructed biosensors pSC$_{Cg}$-PhdR-W and pSC$_{Cg}$-PhdR-M, displayed a dynamic range of 30-fold induction (Fig. 2A) whereas the highest dynamic range (101-fold) could be determined for pSenPhdR employing the native regulatory circuit of the transcriptional regulator (Table 1). Due to an increased basal fluorescence response of pSC$_{Cg}$-PhdR-S, an overall lower dynamic range could be observed for this sensor construct (Fig. S2). This may be reasoned by the leakiness of the regulated promoter at high intracellular levels of transcriptional regulator possibly caused by protein aggregation of the
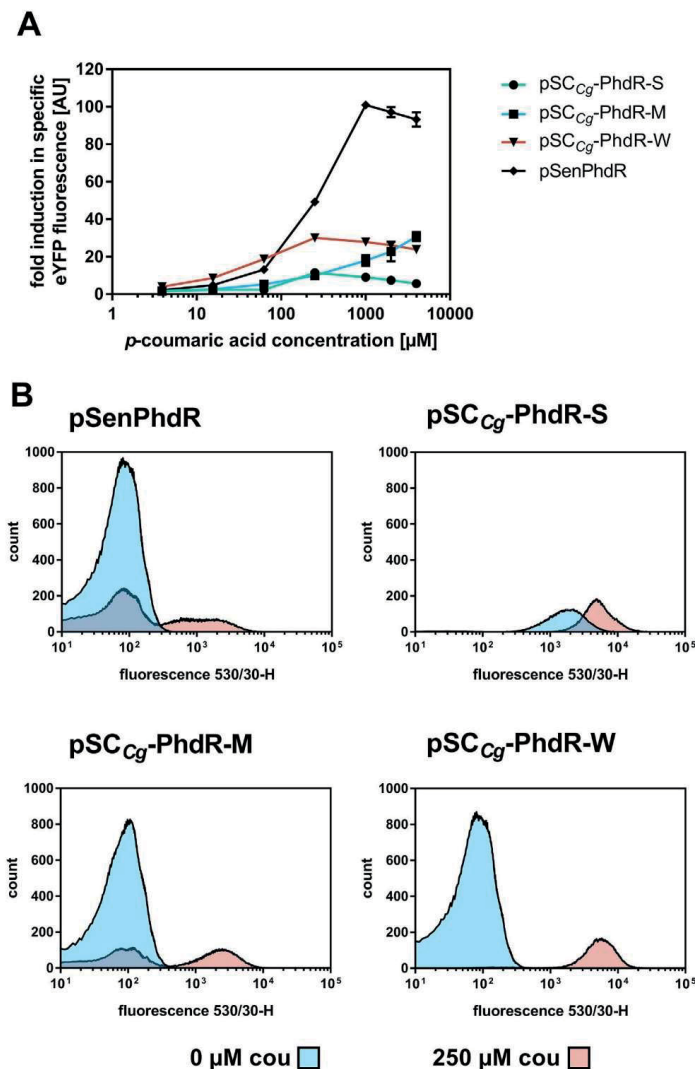
Fig. 2. Biosensor response of constructed PhdR-based biosensors in *C. glutamicum*. (A) Dose-response plot for PhdR-based sensor constructs based on the native regulatory circuit (pSenPhdR) and on the unified sensor design (pSC$_{Cg}$-PhdR-S/M/W). All different cultivations were supplemented with eight different *p*-coumaric acid concentrations ranging from 4 to 4000 μM (externally supplemented). The respective fluorescence response was plotted as fold-change in specific eYFP fluorescence. Error bars represent standard deviations calculated from three biological replicates. (B) FACS experiments with *C. glutamicum* DelAro[4]-4cl$_{Pc}$Δ*phdR* strains carrying pSC$_{Cg}$-PhdR-S/M/W or pSenPhdR in the presence of externally supplemented 0 μM *p*-coumaric acid (blue) or 250 μM *p*-coumaric acid (pink). In each case, 95,000 representative single cells were analyzed. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

regulator protein, which then loses its functionality and cannot tightly repress the expression of the reporter gene (Ellis and Minton, 2006). The reduced dynamic range of the other biosensor variants compared to pSenPhdR (Fig. 2A) can be explained by an increased basal specific fluorescence (Fig. S2), which is in turn due to the leakiness of the regulated promoter at high intracellular transcriptional regulator concentrations. Based on these results, the native genetic circuit of PhdR provides an optimal intracellular regulator activity to ensure high signal amplification. Neither high nor low constitutive expression of *phdR* could restore this sensor response, indicating that an autoregulatory circuit might

control the expression of *phdR* similar to other MarR-type transcriptional regulators (Wilkinson and Grove, 2006).

When comparing the operational range of the sensor variants constructed according to the unified sensor design, distinct differences became apparent (Table 1). The broadest operational range was covered by pSC$_{Cg}$-PhdR-M with moderate regulator gene expression. In contrast, compared to all other constructed PhdR-based sensors pSC$_{Cg}$-PhdR-W showed an increased fluorescence response at lower inducer concentrations. This enhanced biosensor sensitivity allows for sensor applications at low ligand concentrations.

**Table 1**
**Characterization of PhdR and LysG-based biosensors in *C. glutamicum* and *E. coli.*** Operational- and dynamic range were determined for all sensor used in this study. Obtained values are based on biosensor response functions obtained from dose-response experiments. The dynamic range is given as maximal fold-induction. For the operational range, the lowest and the highest inducer concentration is shown for which a change in fluorescent output signal is given until a plateau or maximum is reached. In all experiments, biological triplicates were analyzed.

| Sensor | Inducer | Host | Dynamic range [fold-induction] | Operational range [µM] | |
|---|---|---|---|---|---|
| | | | | low | high |
| pSC$_{Cg}$-PhdR-S | *p*-coumaric acid | *C. glutamicum* DelAro$^4$-4cl$_{Pc}$ $\Delta phdR$ | 11 | 63 | 250 |
| pSC$_{Cg}$-PhdR-M | *p*-coumaric acid | | 30 | 16 | 4000 |
| pSC$_{Cg}$-PhdR-W | *p*-coumaric acid | | 30 | 4 | 250 |
| pSenPhdR | *p*-coumaric acid | | 101 | 4 | 1000 |
| pSC$_{Cg}$-PhdR-W | ferulic acid | | 17 | 4 | 63 |
| pSC$_{Cg}$-PhdR-W | caffeic acid | | 15 | 4 | 250 |
| pSenPhdR | ferulic acid | | 60 | 4 | 250 |
| pSenPhdR | caffeic acid | | 55 | 4 | 1000 |
| pSC$_{Cg}$-LysG-S | L-His | *C. glutamicum* $\Delta lysEG$ | 120 | 125 | 10,000 |
| pSC$_{Cg}$-LysG-S | L-Arg | | 10 | 1250 | 3750 |
| pSC$_{Cg}$-LysG-S | L-Lys | | 85 | 1250 | 10,000 |
| pSC$_{Cg}$-LysG-M | L-His | | 107 | 125 | 10,000 |
| pSC$_{Cg}$-LysG-W | L-His | | 74 | 125 | 5000 |
| pSenLysG | L-His | | 92 | 125 | 10,000 |
| pSenLysG | L-Arg | | 68 | 1250 | 10,000 |
| pSenLysG | L-Lys | | 102 | 625 | 10,000 |
| pSC$_{Ec}$-PhdR-S | *p*-coumaric acid | *E. coli* DH10B pCDF-BAD-4cl$_{Sc}$ | 1 | – | – |
| pSC$_{Ec}$-PhdR-W | *p*-coumaric acid | | 9 | 25 | 100 |
| pSC$_{Ec}$-PhdR-M2 | *p*-coumaric acid | | 2 | 50 | 100 |
| pSC$_{Ec}$-PhdR-M1 | *p*-coumaric acid | | 2 | 50 | 100 |
| pSC$_{Ec}$-LysG-S | L-His | *E. coli* DH10B | 10 | 2500 | 25,000 |
| pSC$_{Ec}$-LysG-W | L-His | | 16 | 2500 | 75,000 |
| pSC$_{Ec}$-LysG-M2 | L-His | | 15 | 2500 | 50,000 |
| pSC$_{Ec}$-LysG-M1 | L-His | | 15 | 2500 | 50,000 |

### 3.4. Homogeneous fluorescence response of PhdR-based biosensors can be restored by decreased regulator expression levels

Transcriptional biosensors show their true potential in combination with FACS in high-throughput screening (HTS) campaigns in which every cell is characterized with respect to fluorescence intensity individually. However, the biosensor response of a whole culture can be very different from the response of individual cells. Hence, an in-depth characterization of the different PhdR-based biosensors at the single-cell level using flow cytometry was carried out to judge their suitability for such applications.

In this context, a low basal fluorescence signal in an uninduced state and a homogeneous fluorescence response of a cell population with an identical genetic background to a certain inducer concentration is highly desired. Thus, individual cells with the desired phenotype (showing strong fluorescence) can be reliably distinguished in real FACS-based screening campaigns. First experiments comparing the basal fluorescence of *C. glutamicum* DelAro$^4$-4cl$_{Pc}$$\Delta phdR$ to the same strain carrying the empty sensor plasmid pSC$_{Cg}$ without any biosensor components as well as the strain carrying the pSenPhdR sensor, confirmed a similarly low fluorescence response (Fig. S3A). However, in the first experiments conducted, the histograms of all PhdR-based biosensors showed a heterogeneous biosensor response at selected ligand concentrations as the fluorescent populations always separated into two global maxima (Fig. 2B and S4). Since, a full shift of the population from un-induced to induced state defines the field of application for repressor-based biosensors, the fluorescence response of different PhdR-based biosensors was investigated for different inducer concentrations (Fig. S4). Interestingly, the lower the expression level of the repressor gene, the lower the required *p*-coumaric acid concentration to homogeneously shift the fluorescence of the whole cell population. Explicitly, this could be shown by comparing the fluorescence response of all assessed PhdR-based biosensors at 250 µM *p*-coumaric acid. Here, both, pSenPhdR and pSC$_{Cg}$-PhdR-M, gave a heterogeneous fluorescence response resulting in two distinct populations during FACS analysis, whereas the fluorescence response of all cells bearing pSC$_{Cg}$-PhdR-W showed a single global maximum upon supplementation of *p*-coumaric acid. (Fig. 2B). For all constructed sensors the gradual shift of the fluorescent population induced with increasing ligand concentration is sufficient to apply PhdR-based biosensors in screening applications in the determined operational range (Fig. S4).

### 3.5. Ferulic acid and caffeic acid trigger a weaker biosensor response

In addition to *p*-coumaroyl-CoA, PhdR can also detect other ring-hydroxylated phenylpropanoids such as ferulic and caffeic acid (Kallscheuer et al., 2016a). Therefore, it was of interest to know whether the sensor response also depends on the supplemented ligands. For this purpose, the fluorescence response of pSC$_{Cg}$-PhdR-W and pSenPhdR was investigated in 48-well platform cultivations and subsequent FACS experiments in the presence of both phenylpropanoids. pSC$_{Cg}$-PhdR-W was selected as this biosensor already offered a strong and homogeneous fluorescence response at low *p*-coumaric acid concentrations in previous experiments.

In general, the dynamic range of both biosensors determined for ferulic acid and caffeic acid was reduced by approximately 30% in comparison to *p*-coumaric acid (Table 1). However, when comparing both biosensors directly, pSC$_{Cg}$-PhdR-W showed an overall decreased performance with regard to operational and dynamic range in response to ferulic acid and caffeic acid. Additional characterization of the fluorescence response of both PhdR-based sensors using FACS showed a shift of the respective population from low to a high fluorescence intensity with increasing inducer concentrations (Fig. S5). When comparing the histogram plots at 1000 µM inducer concentration of both biosensors (Fig. 3D and Fig. S5 B,D), a clear shift from low to high fluorescence could be observed for pSC$_{Cg}$-PhdR-W in the presence of *p*-coumaric acid and caffeic acid, but not in response to supplementation of ferulic acid. In contrast, pSenPhdR gave a heterogeneous fluorescence response in presence of all three phenylpropanoids (Fig. 3 C and Fig. S5 A,C). This might indicate that the overall intracellular repressor activity of pSenPhdR was higher compared to pSC$_{Cg}$-PhdR-W, which resulted in a stronger repression of the reporter gene at identical inducer concentrations ultimately leading to a more heterogeneous biosensor response. Noteworthy, not only regulator gene expression and inducer concentration, but also affinity of the respective inducer to the 4CL and the affinity of the CoA-activated inducer on the regulator protein have an impact on the biosensor response in these experiments.
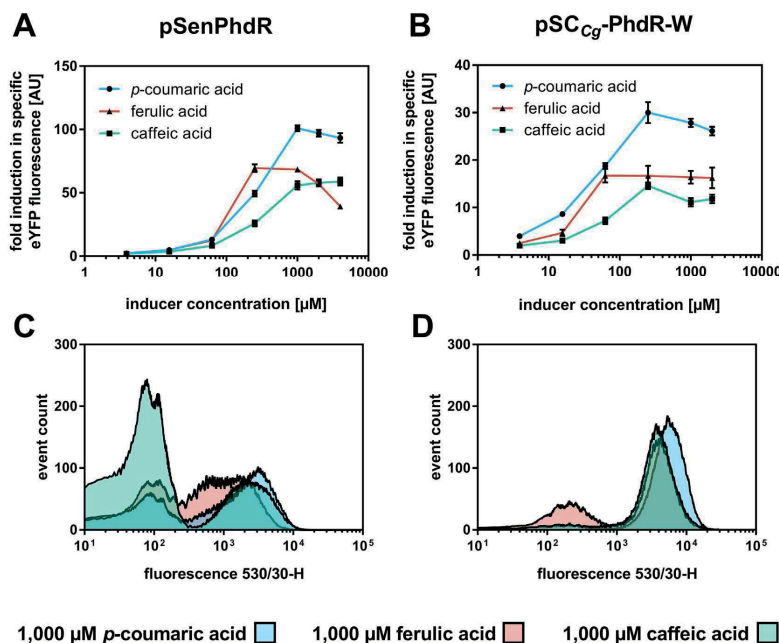
**Fig. 3. Ligand spectrum of PhdR-based biosensors in *C. glutamicum*.** Dose-response of (A) pSenPhdR and (B) pSC$_{Cg}$-PhdR-W in the presence of externally supplemented *p*-coumaric acid (blue), ferulic acid (pink) and caffeic acid (green). Cultivations were supplemented with eight different inducer concentrations ranging from 4 to 4000 μM. The respective fluorescence response was plotted as fold-change in specific eYFP fluorescence. Error bars represent standard deviations calculated from three biological replicates. FACS experiments with *C. glutamicum* DelAro$^4$-4cl$_{Pc}$Δ*phdR* strains carrying (C) pSenPhdR or (D) pSC$_{Cg}$-PhdR-W in the presence of externally supplemented 1000 μM *p*-coumaric acid (blue), ferulic acid (pink) and caffeic acid (green). In each case, 95,000 representative single cells were analyzed. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

However, based on the results from analysis of liquid cultures and FACS analysis, *phdR* expression requires tight control since high repressor concentrations result in permanent repression of the reporter gene, whereas low repressor levels result in an incomplete or partial repression of the reporter gene. Taken together, pSC$_{Cg}$-PhdR-W appears to be most suitable biosensor for FACS applications since a homogeneous shift of the entire population with respect to the fluorescence intensity was observable for all tested phenylpropanoids, even at low inducer concentrations.

### 3.6. Construction and modulation of transcriptional activator-based biosensors

In addition to the constructed repressor-based PhdR-biosensors, we set out to also use the unified biosensor design for constructing activator-based biosensors. In this context, the genetic circuit controlling basic amino acid export in *C. glutamicum* was selected, which involves the LTTR-type transcriptional activator LysG (Bellman et al., 2001). In *C. glutamicum*, LysG regulates the expression of the transporter gene *lysE*, which exports excess amounts of basic proteinogenic amino acids L-lysine and L-arginine as well as of L-citrulline and L-ornithine (Bellman et al., 2001). Based on this native regulatory circuit, the transcriptional biosensor pSenLysG was constructed previously (Binder et al., 2012). Subsequently, this biosensor was successfully used to screen chemically mutagenized *C. glutamicum* wild type cells for identifying novel genetic hot spots contributing to L-lysine overproduction using FACS (Binder et al., 2012). Here, several LysG-based biosensors using the same

constitutive promoters for regulator gene expression were constructed according to the unified sensor design for *C. glutamicum* (pSC$_{Cg}$-LysG-W, pSC$_{Cg}$-LysG-M, pSC$_{Cg}$-LysG-S). All constructed sensor plasmids were characterized in a *C. glutamicum* Δ*lysEG* strain background to prevent interfering regulator and/or exporter gene expression from the genome.

Initially, the fluorescence response of the constructed biosensors was assessed in 48-well plate cultivations using amino acid dipeptides His-Ala, Lys-Ala and Arg-Ala at different concentrations ranging from 0 to 10,000 μM (Fig. S6). Dipeptide inducers were used instead of free amino acids, as dipeptides are more easily taken up by *C. glutamicum*, independent from their composition and sequence (Erdmann et al., 1993). In these experiments, the original biosensor pSenLysG served as benchmark. For all constructed biosensors the fold-induction in specific eYFP fluorescence was plotted against the respective inducer concentration (Fig. 4A). These dose-response experiments revealed, that the dynamic range increases with an increasing expression of the regulator gene *lysG* (Table 1 and Fig. 4A). Based on the "activating" nature of LysG, an increase of regulator gene expression resulted in an enhanced sensor response at low inducer concentrations. In line with this sensor response at low regulator gene expression levels, a slight shift of the operational range to higher concentrations could be observed. Therefore, the design principles developed can be also used to modulate the transcriptional activator activity, which ensures a broad applicability of LysG-based biosensors.

The potential of the different LysG-based biosensors in HT-FACS applications was assessed by detailed FACS analysis using all constructed biosensor variants (Fig. 4B). A low basal fluorescence level is crucial for
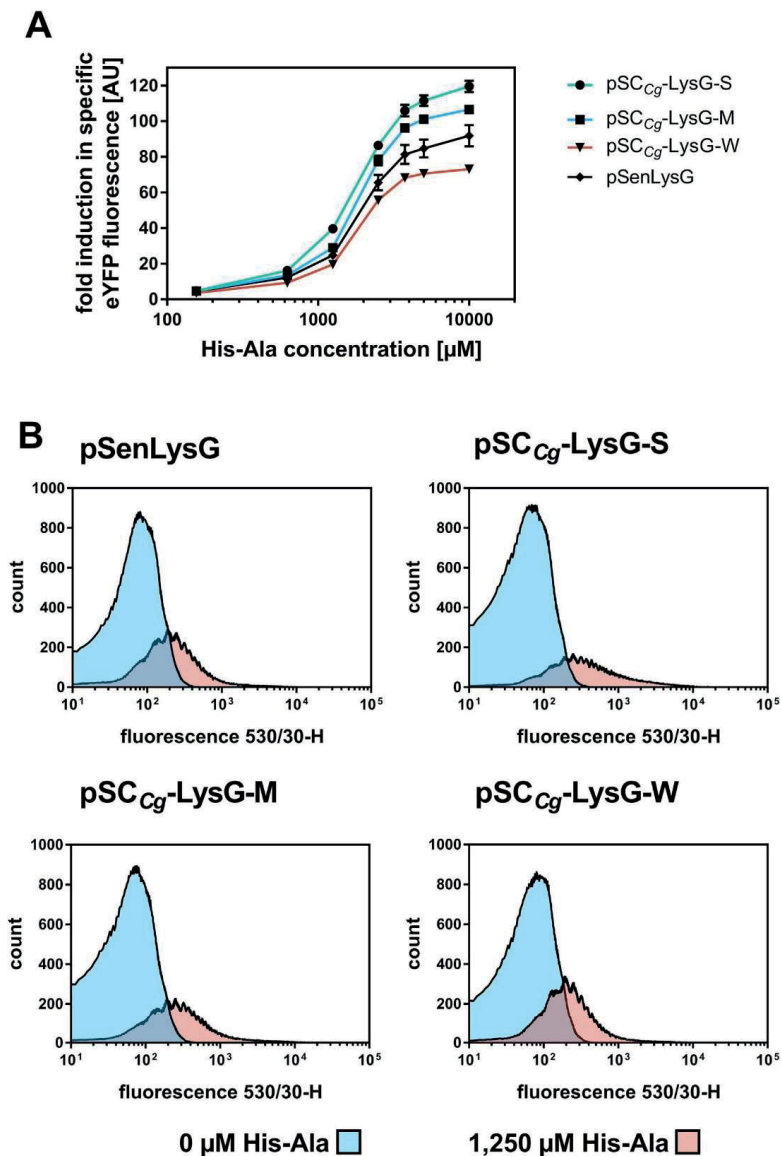
Fig. 4. **Biosensor response of constructed LysG-based biosensors in *C. glutamicum*.** (A) Dose-response plot for LysG-based sensor constructs based on the native regulatory circuit (pSenLysG) and on the unified sensor design (pSC$_{Cg}$-LysG-S/M/W). All different cultivations were supplemented with eight different inducer concentrations ranging from 125 to 10,000 μM dipeptide His-Ala (externally supplemented). The respective fluorescence response was plotted as fold-change in specific eYFP fluorescence. Error bars represent standard deviations calculated from three biological replicates. (B) FACS experiments with *C. glutamicum* Δ*lysEG* strains carrying pSenLysG or pSC$_{Cg}$-LysG-S/M/W in the presence of externally supplemented 0 μM His-Ala (blue), 1250 μM His-Ala (pink) and 2500 μM His-Ala (green). In each case 95,000 representative single cells were analyzed. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

any future sensor application in high-throughput screening campaigns. Therefore, a first characterization comparing the basal fluorescence level of *C. glutamicum* Δ*lysEG* to the same strain carrying the empty sensor plasmid pSC$_{Cg}$ without any biosensor components as well as the same strain carrying the pSenLysG sensor was performed. These experiments confirmed a very similar fluorescence response under the same conditions rendering strain and plasmid suitable for further experiments (Fig. S3B).

Corresponding to the determined operational range from performed BioLector cultivations (Fig. 4A), an increase and gradual shift in fluorescence intensity from lower to higher state could be followed for all sensor variants with increasing inducer concentration (Fig. S7). An increased fluorescence response for pSC$_{Cg}$-LysG-S and pSC$_{Cg}$-LysG-M could be observed at low inducer concentrations (625 μM His-Ala) whereas the same shift could be detected for pSenLysG and pSC$_{Cg}$-LysG-W at higher inducer concentrations (>625 μM His-Ala) (Fig. S7). This result was confirmed by comparing the histogram plots at extracellular supplemented inducer concentration of 1250 μM His-Ala of all biosensor constructs in which the different stages of induction of the sensor constructs were clearly visible (Fig. 4B). A distinct shift of the fluorescence response from an uninduced to an induced state allows for the application of LysG-based biosensors in FACS screening approaches to achieve a sufficient separation between different cell populations in a genetically diverse mutant library. Since the operational range of the

constructed sensor variants differs as a result of the different promoter strengths, a suitable biosensor can be selected based on the expected target molecule concentration in the respective screening application.

### 3.7. Ligand-spectrum of LysG-based biosensors

Further sensor characterization with regard to the biosensor's ligand spectrum was performed with different dipeptide inducers containing the basic amino acids L-arginine and L-lysine. Previous binding studies of LysG showed that the transcriptional regulator responds to all basic amino acids L-histidine-, L-lysine- and L-arginine with different affinities corresponding to the dissociation constants (K$_D$) (L-His $16 \pm 1.1$ $10^{-6}$ M/ L-Lys $3.29 \pm 0.62$ $10^{-3}$ M/L-Arg $1.15 \pm 0.06$ $10^{-3}$ M) (Della Corte et al., 2020). Due to the observed broader dynamic- and operational range of pSC$_{Cg}$-LysG-S with inducer His-Ala, this biosensor and pSenLysG were used to characterize the biosensor performance in the presence of Lys-Ala and Arg-Ala in 48-well plate cultivations and in FACS experiments (Table 1). The dynamic range of pSC$_{Cg}$-LysG-S in response to the presence both dipeptides was decreased compared to His-Ala. Here, pSenLysG showed a broader dynamic range compared to pSC$_{Cg}$-LysG-S in case of the same dipeptides. However, in case of both biosensors, with decreasing binding affinity of the respective ligand to LysG, the operational range of the different biosensor variants was more shifted to higher concentrations (Table 1, Fig. 5A and B).
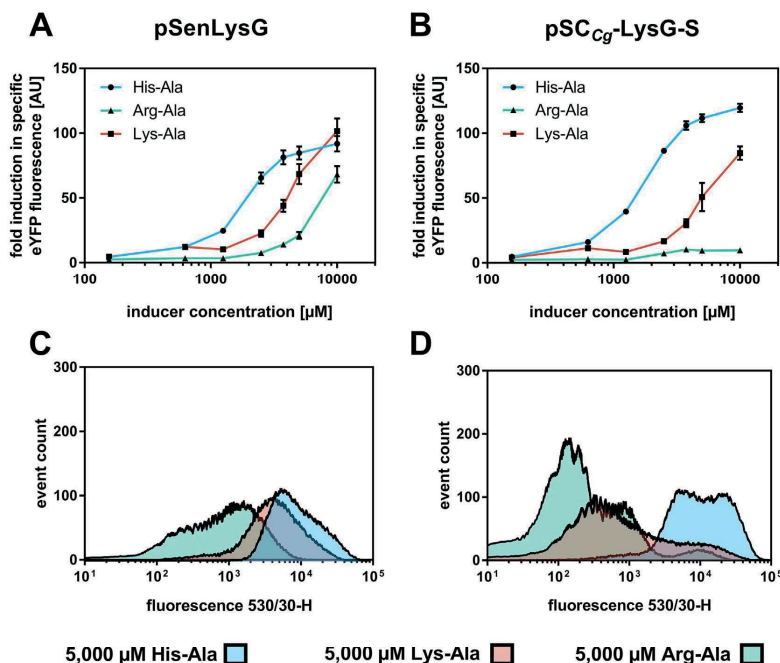


**Fig. 5. Ligand spectrum of LysG-based biosensors in *C. glutamicum*.** (A) Dose-response plot for reference pSenLysG and (B) sensor construct pSC$_{Cg}$-LysG-S in the presence of externally supplemented inducers His-Ala (blue), Lys-Ala (pink) and Arg-Ala (green). Cultivations were supplemented with eight different inducer concentrations ranging from 125 μM −10,000 μM (externally supplemented). The respective fluorescence response was plotted as fold-change in specific eYFP fluorescence. Error bars represent standard deviations calculated from three biological replicates. FACS experiments with *C. glutamicum* Δ*lysEG* strains carrying (C) pSenLysG or (D) pSC$_{Cg}$-LysG-S in the presence of 5000 μM His-Ala (blue), Lys-Ala (pink) and Arg-Ala (green) (externally supplemented). In each case 95,000 representative single cells were analyzed. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

The fluorescence response at the single-cell level confirmed the results obtained in liquid cultures and showed that an increased binding affinity to the inducer allowed for a shift from lower to a higher fluorescence levels at lower inducer concentrations. With respect to a homogeneous sensor response on a single-cell level for the whole cell population, the respective sensor response of pSenLysG and pSC$_{Cg}$-LysG-S in the presence of different dipeptide inducers at the same concentration was investigated. A homogeneous fluorescence response was recorded for all cells carrying pSenLysG within a culture (Fig. 5C and S8 A,C), whereas pSC$_{Cg}$-LysG-S only yielded a heterogeneous fluorescence response for the whole cell population (Fig. 5D and S8 B,D). In case of pSenLysG, the native genetic circuit controlling *lysG* expression always provides sufficient regulator activity in response to the respective inducer concentration, whereas in pSC$_{Cg}$-LysG-S a strong constitutive expression of *lysG* always ensures an excess of LysG in the cell. Intracellularly, LysG is present either in an activated ligand-bound state or a ligand-free inactive state (Bellman et al., 2001). This equilibrium is influenced by many parameters, namely the regulator concentration, the number of ligand binding sites per regulator as well as by the inducer concentration and binding affinity of the inducer to its designated regulator binding site. Hence, a shift of this equilibrium results in a change in fluorescence response of the sensor. Therefore, a weaker or incoherent activation of the transcriptional regulator depending on inducer concentration, may result in a heterogeneous fluorescence response of the entire cell population on the single-cell level.

Hence, fine-tuning of the expression level of a transcriptional repressor or activator in different sensor constructs enables the adaption of the sensor's operational and dynamic range depending on the desired application in *C. glutamicum*. Successful sensor characterization on a single-cell level allowed for the identification of the concentration range in which a gradual shift of the fluorescent population from low to higher fluorescence levels can be assessed, even for different inducer molecules. Based on the results of the sensor characterization of both LysG and PhdR-based biosensors, the repertoire of biosensor variants constructed allows for selection of a specific sensor construct, which meets the requirements for FACS screenings in a specific field of application.

### 3.8. The unified biosensor design allows for the functional transfer of biosensors to E. coli

Previously, a LysG-based biosensor using the native promoters from *C. glutamicum* was constructed for applications in *E. coli* (Wang et al., 2016). However, no fluorescence response could be detected in performed dose-response experiments (Wang et al., 2016). With the aim to construct LysG-based biosensors for *E. coli* and to show that also PhdR-based sensors can be functionally implemented, several biosensor variants were built following the unified biosensor design. In this context, biosensors carrying either *lysG* or *phdR* under control of four different constitutive promoters of the PLTetO1 promoter library were constructed. For the application of the transcriptional regulator PhdR in
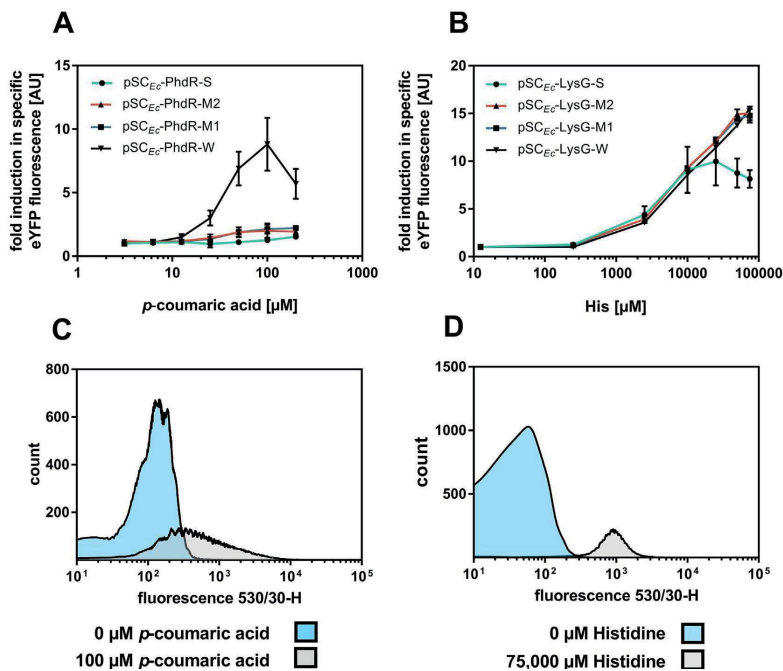


**Fig. 6. Biosensor response of PhdR and LysG-based biosensors in *E. coli*.** Dose-response plot for (A) PhdR-based and (B) LysG-based biosensors in *E. coli* with inducers and L-histidine and *p*-coumaric acid, respectively. Cultivations were supplemented with eight different externally supplemented inducer concentrations ranging from 125– to 75,000 μM L-His and 4–200 μM *p*-coumaric acid for LysG- and PhdR-based biosensors respectively. The biosensor response was plotted as fold change in specific eYFP fluorescence. Error bars represent standard deviations calculated from three biological replicates. FACS experiments with *E. coli* strains carrying (C) pSC$_{Ec}$-PhdR-W in the presence of 0 μM *p*-coumaric acid (blue) and 100 μM *p*-coumaric acid (grey) and (D) pSC$_{Ec}$-LysG-W were performed in the presence of 0 μM His (blue) and 75,000 μM His-Ala (grey) (all externally supplemented). In each case 95,000 representative single cells were analyzed. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

*E. coli*, CoA-activation of the ring-hydroxylated phenylpropanoids is essential. Hence, the functionality of the transcriptional biosensor could only be achieved by co-expression of a 4CL-encoding gene in *E. coli*. For this purpose, the open reading frame of the 4cl-gene originating from *Streptomyces coelicolor* was integrated in a pCDF-Duet-1 vector under control of an arabinose-inducible promoter yielding pCDF-BAD-4cl (van Summeren-Wesenhagen and Marienhagen, 2015). For all experiments concerning PhdR-based biosensors, the respective biosensor plasmid and the 4cl expression plasmid were co-expressed in *E. coli* DH10B.

All PhdR- and LysG-based biosensors were first characterized in 48-well plate cultivations with regard to their fluorescence response at different inducer concentrations ranging from 0 to 200 μM *p*-coumaric acid and 0–90,000 μM L-histidine for PhdR- and LysG based biosensors, respectively (Fig. 6 A,B and S 9,10). Obtained results showed that PhdR-based biosensors are fully functional in *E. coli*, however, growth of sensor constructs pSC$_{Ec}$-PhdR-S, pSC$_{Ec}$-PhdR-M1 and pSC$_{Ec}$-PhdR-M2 was impaired. For an optimal fluorescence response of PhdR-based biosensors, optimization of the kanamycin concentration for plasmid maintenance and the arabinose concentration controlling 4cl gene expression was crucial to reduce the metabolic burden. With 0.02% (w/v) arabinose corresponding to low expression levels of the 4cl gene, the biosensor pSC$_{Ec}$-PhdR-W reached the highest dynamic range with up to 9 fold-induction at 100 μM *p*-coumaric acid inducer concentration also offering the widest operational range (Fig. 6 B). Due to the employed 4CL not accepting any other substrate than *p*-coumaric acid, only this inducer could be tested. For PhdR-based biosensors in *E. coli*, the highest dynamic- and operational range could be determined for the biosensor variant with the weakest constitutive promoter. These results showed that low levels of the transcriptional repressor lead to the highest fluorescence response, which might be due to a low metabolic burden for *E. coli* at low expression levels. In accordance to the performed dose-response experiments in 48-well plate cultivations, the conducted FACS experiments showed an increase in fluorescence over increasing *p*-coumaric acid concentrations (Fig. 6C and S11).

Similar to PhdR-based biosensors, a sensor response of all constructed LysG-based biosensors could be obtained in *E. coli*. Key was the expression the transcriptional regulator gene under control of the constitutive promoter originating from *E. coli*. The maximal fold-induction of all variants was reached for pSC$_{Ec}$-LysG-W at 90,000 μM L-histidine with 16-fold induction (Fig. 6B) and the widest operational range (Table 1). However, a substantial difference in dynamic- and operational range between different sensor variants comprising a weak and moderate promoter was not observable. Exemplarily, the fluorescence response of pSC$_{Ec}$-LysG-W was investigated on the single cell-level using FACS. Increasing inducer concentrations also resulted in a shift of fluorescence intensity. Based on the homogeneous fluorescence response over the entire cell population at different inducer concentrations and a well differentiable fluorescence response, the sensor can be applied in HTS in combination with FACS to detect basic amino acids in *E. coli* (Fig. 6D and S12).

Independent from the mode of action of the used transcriptional regulator, the fluorescence response of the respective transcriptional biosensor from *C. glutamicum* could always be restored in *E. coli*. Furthermore, the constitutive promoter library controlling the regulator expression always allowed for the selection of a suitable transcriptional biosensor variant for FACS applications.

## 4. Conclusions

In this study, we introduced a unified biosensor design, which allows not only for the construction of transcriptional biosensors in *c. glutamicum* and *e. coli*, but also for the fine-tuning of important sensor parameters, such as dynamic range and operational range. Depending on the mode of action of the transcriptional regulator, regulator expression strength was an important parameter when tailoring the biosensor activity to a specific application, either in liquid cultures or at the single-cell

level using FACS. Furthermore, an individual characterization of the biosensor in the presence of different ligands is essential, as different binding affinities of the respective inducer molecules strongly influence the overall biosensor response. In the future, the biosensor variability can be expanded by modulation of the target gene promoter, which would allow additional fine-tuning of the biosensor performance. We believe that the biosensor design described here, can also be used to construct tailor-made biosensors for other prokaryotic species.

## Author contributions

C.K.S. and L.K.F designed the experiments; C.K.S., C.M. and M.V. performed all experiments, and C.K.S. and J.M. wrote the manuscript. All authors read and approved the final version of the manuscript.

## CRediT authorship contribution statement

**Christiane Katharina Sonntag:** Conceptualization, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Lion Konstantin Flachbart:** Conceptualization, Verification. **Celine Maass:** Investigation. **Michael Vogt:** Investigation. **Jan Marienhagen:** Conceptualization, Funding acquisition, Supervision, Resources, Writing - original draft, Writing - review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mec.2020.e00150.

## References

Abe, S., Takayama, K.-I., Kinoshita, S., 1967. Taxonomical studies on glumatic acid-producing bacteria. J. Gen. Appl. Microbiol. 13, 279–301. https://doi.org/10.2323/jgam.13.279.

Alper, H., Fischer, C., Nevoigt, E., Stephanopoulos, G., Demasi, J., Huh, K., Nakatani, Y., Mu, K., 2006. Tuning genetic control through promoter engineering. Pnas 103, 3006–3007. https://doi.org/10.1073/pnas.0507062103.

Bellman, A., Vrljić, M., Pátek, M., Sahm, H., Krämer, R., Eggeling, L., 2001. Expression control and specificity of the basic amino acid exporter *lysE* of *Corynebacterium glutamicum*. Microbiol. 147, 1765–1774. https://doi.org/10.1099/00221287-147-7-1765.

Bertani, G., 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. J. Bacteriol. 62, 293 doi:citeulike-article-id:149214.

Binder, S., Schendzielorz, G., Stäbler, N., Krumbach, K., Hoffmann, K., Bott, M., Eggeling, L., 2012. A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. Genome Biol. 13, R40. https://doi.org/10.1186/gb-2012-13-5-r40.

Bonnassie, S., Oreglia, J., Sicard, A.M., 1990. Nucleotide sequence of the *dapA* gene from *Corynebacterium glutamicum*. Nucleic Acids Res. 18, 6421.

De Paepe, B., Maertens, J., Vanholme, B., De Mey, M., 2018. Modularisation and response curve engineering of a naringenin-responsive transcriptional biosensor. ACS Synth. Biol. 7, 1303–1314. https://doi.org/10.1021/acssynbio.7b00419.

Della Corte, D., van Beek, H.L., Syberg, F., Schallmey, M., Tobola, F., Cormann, K.U., Schlicker, C., Baumann, P.T., Krumbach, K., Sokolowsky, S., Morris, C.J., Grünberger, A., Hofmann, E., Schröder, G.F., Marienhagen, J., 2020. Engineering and application of a biosensor with focused ligand specificity. Nat. Commun. 11, 4851. https://doi.org/10.1038/s41467-020-18400-0.

Durfee, T., Nelson, R., Baldwin, S., Plunkett, G., Burland, V., Mau, B., Petrosino, J.F., Qin, X., Muzny, D.M., Ayele, M., Gibbs, R.A., Csörgo, B., Pósfai, G., Weinstock, G.M., Blattner, F.R., 2008. The complete genome sequence of *Escherichia coli* DH10B:

insights into the biology of a laboratory workhorse. J. Bacteriol. 190, 2597–2606. https://doi.org/10.1128/JB.01695-07.

Eggeling, L., Bott, M., 2005. Handbook of Corynebacterium Glutamicum. CRC Press. https://doi.org/10.1201/9781420039696.

Eggeling, L., Bott, M., Marienhagen, J., 2015. Novel screening methods-biosensors. Curr. Opin. Biotechnol. 35, 30–36. https://doi.org/10.1016/j.copbio.2014.12.021.

Ellis, R.J., Minton, A.P., 2006. Protein aggregation in crowded environments. Biol. Chem. 387, 485–497. https://doi.org/10.1515/BC, 064.

Erdmann, A., Weil, B., Kramer, R., 1993. Lysine secretion by wild-type *Corynebacterium glutamicum* triggered by dipeptide uptake. J. Gen. Microbiol. 139, 3115–3122. https://doi.org/10.1099/00221287-139-12-3115.

Funke, M., Diederichs, S., Kensy, F., Müller, C., Büchs, J., 2009. The baffled microtiter plate: increased oxygen transfer and improved online monitoring in small scale fermentations. Biotechnol. Bioengineering 103, 1118–1128. https://doi.org/10.1002/bit.22341.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–345. https://doi.org/10.1038/nmeth.1318.

Hirakawa, H., Schaefer, A.L., Greenberg, E.P., Harwood, C.S., 2012. Anaerobic *p*-coumarate degradation by *Rhodopseudomonas palustris* and identification of *couR*, a MarR repressor protein that binds p-coumaroyl coenzyme A. J. Bacteriol. 194, 1960–1967. https://doi.org/10.1128/JB.06817-11.

Jha, R.K., Bingen, J.M., Johnson, C.W., Kern, T.L., Khanna, P., Trettel, D.S., Strauss, C.E.M., Beckham, G.T., Dale, T., 2018. A protocatechuate biosensor for *Pseudomonas putida* KT2440 via promoter and protein evolution. Metab. Eng. Commun. Now. 6, 33–38. https://doi.org/10.1016/j.meteno.2018.03.001.

Jha, R.K., Kern, T.L., Fox, D.T., M Strauss, C.E., 2014. Engineering an *Acinetobacter* regulon for biosensing and high-throughput enzyme screening in *E. coli* via flow cytometry. Nucleic Acids Res. 42, 8150–8160. https://doi.org/10.1093/nar/gku444.

Kallscheuer, N., Vogt, M., Kappelmann, J., Krumbach, K., Noack, S., Bott, M., Marienhagen, J., 2016a. Identification of the *phd* gene cluster responsible for phenylpropanoid utilization in *Corynebacterium glutamicum*. Appl. Microbiol. Biotechnol. 100, 1871–1881. https://doi.org/10.1007/s00253-015-7165-1.

Kallscheuer, N., Vogt, M., Stenzel, A., Gätgens, J., Bott, M., Marienhagen, J., 2016b. Construction of a *Corynebacterium glutamicum* platform strain for the production of stilbenes and (2S)-flavanones. Metab. Eng. 38, 47–55. https://doi.org/10.1016/j.ymben.2016.06.003.

Keilhauer, C., Eggeling, L., Sahm, H., 1993. Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon. J. Bacteriol. 175, 5595–5603. https://doi.org/10.1128/jb.175.17.5595-5603.1993.

Kensy, F., Zang, E., Faulhammer, C., Tan, R.-K.K., Büchs, J., 2009. Validation of a high-throughput fermentation system based on online monitoring of biomass and fluorescence in continuously shaken microtiter plates. Microb. Cell Factories 8, 31. https://doi.org/10.1186/1475-2859-8-31.

Kudla, G., Murray, A.W., Tollervey, D., Plotkin, J.B., 2009. Coding-sequence determinants of gene expression in *Escherichia coli*. Sci. 324, 255–258. https://doi.org/10.1126/science.1170160.

Lin, C., Jair, Y.C., Chou, Y.C., Chen, P.S., Yeh, Y.C., 2018. Transcription factor-based biosensor for detection of phenylalanine and tyrosine in urine for diagnosis of phenylketonuria. Anal. Chim. Acta 1041, 108–113. https://doi.org/10.1016/j.aca.2018.08.053.

Maddocks, S.E., Oyston, P.C.F., 2008. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiol. 154, 3609–3623. https://doi.org/10.1099/mic.0.2008/022772-0.

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

Mannan, A.A., Liu, D., Zhang, F., Oyarzún, D.A., 2017. Fundamental design principles for transcription-factor-based metabolite biosensors. ACS Synth. Biol. 6, 1851–1859. https://doi.org/10.1021/acssynbio.7b00172.

Merulla, D., Hatzimanikatis, V., Van der Meer, J.R., 2013. Tunable reporter signal production in feedback-uncoupled arsenic bioreporters. Microb. Biotechnol 6, 503–514. https://doi.org/10.1111/1751-7915.12031.

Niebisch, A., Bott, M., 2001. Molecular analysis of the cytochrome bc 1-aa 3 branch of the *Corynebacterium glutamicum* respiratory chain containing an unusual diheme

cytochrome c 1. Arch. Microbiol. 175, 282–294. https://doi.org/10.1007/s002030100262.

Otani, H., Stogios, P.J., Xu, X., Nocek, B., Li, S.-N., Savchenko, A., Eltis, L.D., 2016. The activity of CouR, a MarR family transcriptional regulator, is modulated through a novel molecular mechanism. Nucleic Acids Res. 44, 595–607. https://doi.org/10.1093/nar/gkv955.

Rebets, Y., Schmelz, S., Gromyko, O., Tistechok, S., Petzke, L., Scrima, A., Luzhetskyy, A., 2018. Design, development and application of whole-cell based antibiotic-specific biosensor. Metab. Eng. 47, 263–270. https://doi.org/10.1016/j.ymben.2018.03.019.

Rogers, J.K., Guzman, C.D., Taylor, N.D., Raman, S., Anderson, K., Church, G.M., 2015. Synthetic biosensors for precise gene control and real-time monitoring of metabolites. Nucleic Acids Res. 43, 7648–7660. https://doi.org/10.1093/nar/gkv616.

Sambrook, J., Russel, D.W., 2001. Molecular Cloning: A Lab Manual, third ed. Cold Spring Harbor Laboratory Press, New York.

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

Schallmey, M., Frunzke, J., Eggeling, L., Marienhagen, J., 2014. Looking for the pick of the bunch: high-throughput screening of producing microorganisms with biosensors. Curr. Opin. Biotechnol. 26, 148–154. https://doi.org/10.1016/j.copbio.2014.01.005.

Siedler, S., Schendzielorz, G., Binder, S., Eggeling, L., Bringer, S., Bott, M., 2014. SoxR as a single-cell biosensor for NADPH-consuming enzymes in *Escherichia coli*. ACS Synth. Biol. 3, 41–47. https://doi.org/10.1021/sb400110j.

Skjoedt, M.L.M., Snoek, T., Kildegaard, K.K.R., Arsovska, D., Eichenberger, M., Goedecke, T.T.J., Rajkumar, A.S., Zhang, J., Kristensen, M., Lehka, B.J., Siedler, S., Borodina, I., Jensen, M.K., Keasling, J.J.D., 2016. Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast. Nat. Chem. Biol. 12, 951–958. https://doi.org/10.1038/nchembio.2177.

Snoek, T., Romero-Suarez, D., Zhang, J., Ambri, F., Skjoedt, M.L., Sudarsan, S., Jensen, M.K., Keasling, J.D., 2018. An orthogonal and pH-tunable sensor-selector for muconic acid biosynthesis in yeast. ACS Synth. Biol. 7, 995–1003. https://doi.org/10.1021/acssynbio.7b00439.

van Ooyen, J., Noack, S., Bott, M., Reth, A., Eggeling, L., 2012. Improved L-lysine production with *Corynebacterium glutamicum* and systemic insight into citrate synthase flux and activity. Biotechnol. Bioeng. 109, 2070–2081. https://doi.org/10.1002/bit.24486.

van Summeren-Wesenhagen, P.V., Marienhagen, J., 2015. Metabolic engineering of *Escherichia coli* for the synthesis of the plant polyphenol pinosylvin. Appl. Environ. Microbiol. 81, 840–849. https://doi.org/10.1128/AEM.02966-14.

Vašicová, P., Pátek, M., Nešvera, J., Sahm, H., Eikmanns, B., 1999. Analysis of the *Corynebacterium glutamicum dapA* promoter. J. Bacteriol. 181, 6188–6191. https://doi.org/10.1128/JB.181.19.6188-6191.1999.

Wang, Y., Li, Q., Zheng, P., Guo, Y., Wang, L., Zhang, T., Sun, J., Ma, Y., 2016. Evolving the L-lysine high-producing strain of *Escherichia coli* using a newly developed high-throughput screening method. J. Ind. Microbiol. Biotechnol. 43, 1227–1235. https://doi.org/10.1007/s10295-016-1803-1.

Webster, D.P., TerAvest, M.A., Doud, D.F.R., Chakravorty, A., Holmes, E.C., Radens, C.M., Sureka, S., Gralnick, J.A., Angenent, L.T., 2014. An arsenic-specific biosensor with genetically engineered *Shewanella oneidensis* in a bioelectrochemical system. Biosens. Bioelectron. 62, 320–324. https://doi.org/10.1016/j.bios.2014.07.003.

Wilkinson P., Steven, Grove, Anne, 2006. Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins. Curr. Iss. Mol. Biol. 8 (1), 51–62.

Wu, G., Yan, Q., Jones, J.A., Tang, Y.J., Fong, S.S., Koffas, M.A.G.G., 2016. Metabolic burden: cornerstones in synthetic biology and metabolic engineering applications. Trends Biotechnol. 34, 652–664. https://doi.org/10.1016/j.tibtech.2016.02.010.

Zhang, J., Barajas, J.F., Burdu, M., Ruegg, T.L., Dias, B., Keasling, J.D., 2017. Development of a transcription factor-based lactam biosensor. ACS Synth. Biol. 6, 439–445. https://doi.org/10.1021/acssynbio.6b00136.

Zhang, J., Jensen, M.K., Keasling, J.D., 2015. Development of biosensors and their application in metabolic engineering. Curr. Opin. Chem. Biol. 28, 1–8. https://doi.org/10.1016/j.cbpa.2015.05.013.

### 8.2.1 A unified design allows fine-tuning of biosensor parameters and application across bacterial species - Supplementary

# A unified design allows fine-tuning of biosensor parameters

# and application across bacterial species

Christiane Katharina Sonntag[1], Lion Konstantin Flachbart[1], Celine Maass[1], Michael Vogt[1]

and Jan Marienhagen[1,2,*]

[1]Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich, D-52425 Jülich, Germany

[2]Institute of Biotechnology, RWTH Aachen University, Worringer Weg 3, D-52074 Aachen, Germany

* To whom correspondence should be addressed.

Tel: +49 2461 61 2843; Fax: +49 2461 61 2710; Email: j.marienhagen@fz-juelich.de

**Supplementary Table S1. Strains constructed and used in this study**

| Strains | Relevant characteristics | Reference or source |
|---|---|---|
| *C. glutamicum* DelAro$^4$-4cl$_{Pc}$ | DelAro$^4$ derivative with chromosomally encoded 4cl$_{Pc}$ gene under control of the T7 promoter (Δcg0344-47:PT7-4clP) | (Kallscheuer et al., 2016) |
| *C. glutamicum* DelAro$^4$-4cl$_{Pc}$ Δ*phdR* | MB001(DE3) derivative with in-frame deletions of cg0344-47, cg2625-40, cg1226, cg0502 and chromosomally encoded 4cl$_{Pc}$ gene under control of the T7 promoter (Δcg0343-47:PT7-4clP) | This work |
| *C. glutamicum* Δ*lysEG* | *C. glutamicum* ATCC 13032 wild type with an in-frame deletion of cg1424-5 | (Vrljic et al., 1996) |
| *E. coli* DH10B | F– *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15 Δ*lacX*74 *recA*1 *endA*1 *ara*D139 Δ(*ara, leu*)7697 *gal*U *gal*K λ– *rps*L *nup*G | Invitrogen (Karlsruhe, Germany) |
| *E. coli* DH5α | F– Φ80lacZΔM15 Δ(lacZYA-argF)U169 recA1 endA1 hsdR17 (rK–,mKþ) phoA supE44 λ– thi-1 gyrA96 relA1 | Invitrogen (Karlsruhe, Germany) |

**Supplementary Table S2. Plasmids constructed and used in this study.**

| Plasmids | Relevant characteristics | Reference or source |
|---|---|---|
| pBAD/myc-His A | Cloning vector for the synthesis of His6-tagged (c-terminal) fusion protein, AmpR | Invitrogen |
| pBR322 | pMB1 (oriV), Tet$^R$, Amp$^R$ | M10785 |
| pCDF-BAD | pCDFDuet1 derived arabinose inducible cloning vector, Spt$^r$, $Nco$I linearization | This work |
| pCDF-BAD-4cl$_{Sc}$ | pCDF-Duet-1 vector variant with an arabinose-inducible promoter controlling the expression of a $4cl$-gene originating from $Streptomyces\ coelicolor$ | This work |
| pCDFDuet1 | Spt$^r$;2 T7 lac promoters, CDF replicon, $lacI$, His$_6$ tag, S tag | Merck Millipore (Billerica, MA, USA) |
| pC-$P$dsts3-$Sc4cl$ | pCDFDuet1 derivative containing sts3 from P. densiflora and 4cl A294G from $S.\ coelicolor$ | (van Summeren-Wesenhagen and Marienhagen, 2015) |
| pEC-XT99A | pGA1 minireplicon pGA1 $per$ gene; Tet$^r$ | AY219684 |
| pJC1 | Km$^r$, pHM1519 ori | AJ012294 |
| pK19mobsacB | Km$^r$, Suc$^r$, mobilizable (oriT), oriV | (Schäfer et al., 1994) |
| pK19mobsacB_cg0343-del | vector for in-frame deletion of cg0343 | This work |
| pSC$_{Cg}$-LysG-M | Transcriptional biosensor which encodes $C.\ glutamicum$ LysG, and its target promoter of $lysE$ with a transcriptional fusion to $eyfp$, Km$^r$ Expression of transcriptional regulator is under control of synthetic promoter variant $dapA$ A14. | This work |
| pSC$_{Cg}$-LysG-S | Transcriptional biosensor which encodes $C.\ glutamicum$ LysG, and its target promoter of $lysE$ with a transcriptional fusion to $eyfp$, Km$^r$ Expression of transcriptional regulator is under control of synthetic promoter variant $dapA$ A16 | This work |
| pSC$_{Cg}$-LysG-W | Transcriptional biosensor which encodes $C.\ glutamicum$ LysG, and its target promoter of $lysE$ with a transcriptional fusion to $eyfp$, Km$^r$ Expression of transcriptional regulator is under control of synthetic promoter variant $dapA$ B27. | This work |
| pSC$_{Cg}$-PhdR-M | Transcriptional biosensor which encodes $C.\ glutamicum$ PhdR, and its target promoter of $phdB$ with a transcriptional fusion to $eyfp$, Km$^r$ Expression of transcriptional regulator is under control of synthetic promoter variant $dapA$ A14. | This work |
| pSC$_{Cg}$-PhdR-S | Transcriptional biosensor which encodes $C.\ glutamicum$ PhdR, and its target promoter of $phdB$ with a transcriptional fusion to $eyfp$, Km$^r$ | This work |

| | | |
|---|---|---|
| | Expression of transcriptional regulator is under control of synthetic promoter variant *dapA* A16. | |
| pSC*Cg*-PhdR-W | Transcriptional biosensor which encodes *C. glutamicum* PhdR, and its target promoter of *phdB* with a transcriptional fusion to *eyfp*, Km^r Expression of transcriptional regulator is under control of synthetic promoter variant *dapA* B27. | This work |
| pSC*Ec*-LysG-M1 | Transcriptional biosensor which encodes *C. glutamicum* LysG, and its target promoter of *lysE* with a transcriptional fusion to *eyfp*, Km^r Expression of transcriptional regulator is under control of synthetic promoter variant PLTetO1 K. | This work |
| pSC*Ec*-LysG-M2 | Transcriptional biosensor which encodes *C. glutamicum* LysG, and its target promoter of *lysE* with a transcriptional fusion to *eyfp*, Km^r Expression of transcriptional regulator is under control of synthetic promoter variant PLTetO1 S. | This work |
| pSC*Ec*-LysG-S | Transcriptional biosensor which encodes *C. glutamicum* LysG, and its target promoter of *lysE* with a transcriptional fusion to *eyfp*, Km^r Expression of transcriptional regulator is under control of synthetic promoter variant PLTetO1. | This work |
| pSC*Ec*-LysG-W | Transcriptional biosensor which encodes *C. glutamicum* LysG, and its target promoter of *lysE* with a transcriptional fusion to *eyfp*, Km^r Expression of transcriptional regulator is under control of synthetic promoter variant PLTetO1 JJ. | This work |
| pSC*Ec*-PhdR-M1 | Transcriptional biosensor which encodes *C. glutamicum* PhdR, and its target promoter of *phdB* with a transcriptional fusion to *eyfp*, Km^r Expression of transcriptional regulator is under control of synthetic promoter variant PLTetO1 K. | This work |
| pSC*Ec*-PhdR-M2 | Transcriptional biosensor which encodes *C. glutamicum* PhdR, and its target promoter of *phdB* with a transcriptional fusion to *eyfp*, Km^r Expression of transcriptional regulator is under control of synthetic promoter variant PLTetO1 S. | This work |
| pSC*Ec*-PhdR-S | Transcriptional biosensor which encodes *C. glutamicum* PhdR, and its target promoter of *phdB* with a transcriptional fusion to *eyfp*, Km^r Expression of transcriptional regulator is under control of synthetic promoter variant PLTetO1. | This work |
| pSC*Ec*-PhdR-W | Transcriptional biosensor which encodes *C. glutamicum* PhdR, and its target promoter of *phdB* with a transcriptional fusion to *eyfp*, Km^r Expression of transcriptional regulator is under control of synthetic promoter variant PLTetO1 JJ. | This work |

| pSC$_{Ec}$-term_PLTetO1 | Plasmid for the amplification of the PLTetO1 promoter library fused to the terminators T7$_{term}$ and tonb$_{term}$<br>pMA-RQ vector backbone, col E1 ori, amp$^R$ | This work, Invitrogen (Karlsruhe, Germany) |
|---|---|---|
| pSenCA | Transcriptional biosensor construct inducing *eyfp* expression in response to the presence of *trans*-cinnamic acid or phenylpropionic acid, Km$^r$ | (Flachbart et al., 2019) |
| pSC$_{Cg}$ | pJC1 based biosensor chassis plasmid for *C. glutamicum* with an *eyfp* reporter gene, Km$^r$. Linearization via *Bam*HI and *Sal*I digest. | This work |
| pSC$_{Ec}$ | pBR223 based biosensor chassis plasmid for *E. coli* with an *eyfp* reporter gene, Km$^r$. Linearization via *Hind*III digest. | This work |
| pSenLysG | pJC1 based transcriptional biosensor which encodes *C. glutamicum* LysG, and its target promoter of *lysE* with a transcriptional fusion to *eyfp*, Km$^r$ | (Binder et al., 2012) |
| pSenPhdR | pJC1 based transcriptional biosensor which encodes *C. glutamicum* PhdR, and its target promoter of *phdB* with a transcriptional fusion to *eyfp*, Km$^r$ | This work |

5

**Supplementary Table S3. Oligonucleotides used in this study.**

| Primer name | Sequence | Resulting plasmid |
|---|---|---|
| fw_pMB1 ori | AACGGATTCACCACTCCAAG | pSC$_{Ec}$ |
| rev_pMB1 ori | AAGCTTTCATGACCAAAATCCCTTAACGTG | pSC$_{Ec}$ |
| fw_spacer | TTAAGGCATTTTGGTCATGAAAGCTTTTATGC CGTTACGCTTGCC | pSC$_{Ec}$ |
| rev_spacer | CCTTAAGCTTCCATTGGCTTTGTGCCATC | pSC$_{Ec}$ |
| Fw_RBS_*eyfp*_term | AAGCCAATGGAAGCTTAAGGAGGTTAATTATG GTGAGC | pSC$_{Ec}$ |
| Rev_RBS_*eyfp*_term | GTTCTTCTGAGAGCTCAGTCAAAAGCCTCCGG TC | pSC$_{Ec}$ |
| Fw_*ntp* | GCTTTTGACTGAGCTCTCAGAAGAACTCGTCA AGAAGGC | pSC$_{Ec}$ |
| Rev_*ntp* | CTTGGAGTGGTGAATCCGTTCGAACCGGAATT GCCAGC | pSC$_{Ec}$ |
| Fw_*eyfp* | CGGGATCCCGTTATTACTTGTACAGCTCGTCC ATGC | pSC$_{Cg}$ |
| Rev_*eyfp* | GCTCTAGAGCATGGTGAGCAAGGGCGAGGAG CTGTTC | pSC$_{Cg}$ |
| Fw-SenPhdR-*eyfp* (*Xba*I) | GGGTTCTAGATTATTACTTGTACAGCTCGTCC ATGCC | pSenPhdR |
| Rev-SenPhdR-*eyfp* | TCGTTCCTGAACTTTAAGAAGGAGATATCATA TGGTGAGCAAGGGC | pSenPhdR |
| Fwd-SenPhdR | CTCCTTCTTAAAGTTCAGGAACGACCAAGTCC TGCACCAGATCCG | pSenPhdR |
| Rev-SenPhdR | GGGGTCTAGACAATAAGTTTGCCCCCGATCTT CACAATTGTGC | pSenPhdR |
| Rev_AraC_P$_{araC}$_MCS_ T*rrB* | TTATGTCTATTGCTGGTTTACCGGTAGGGAAT AAGGGCGACACGGAAATGTTGAATACTC | pCDF-BAD |
| Fwd_AraC_P$_{araC}$_MCS_ T*rrB* | ATCTTTTCTACTGAACCGCTTCTAGATACTCC GTCAAGCCGTCAATTGTCTGATTCGTTAC | pCDF-BAD |
| Fwd_4*cl*$_{Sc}$ | GGCTAACAGGAGGAATTAACATGTTTCGTAGC GAATATGCAGATGTTCCGCCTGTTG | pCDF-BAD-4*cl*$_{Sc}$ |
| Rev_4*cl*$_{Sc}$ | GCAGATCTCGAGCTCGGATCTTATTAACGCG GTTCACGCAGCTGACGACG | pCDF-BAD-4*cl*$_{Sc}$ |

| Fwd_up_PhdR | CAAGCACGGGTGTTGCCCCAATGAGGTTCG | pK19mobsacB_cg0343-del |
| Rev_up_PhdR | ACAAATTCAAATCACCTACCCGAACAACTAAGGGAACTGCTGCTCAAAATCGCCTTTTAAGTCTCTTAACCAC | pK19mobsacB_cg0343-del |
| Fwd_down_PhdR | GGGTAGGTGATTTGAATTTGTGGTGGGGTTGCTGGTGGTCATAGTGGCTCCATGTGAACTG | pK19mobsacB_cg0343-del |
| Rev_down_PhdR | CCCGGATAACCCAAGAATTAGTGCCTTCCAGCGAGCTGTAACC | pK19mobsacB_cg0343-del |
| Fwd_Cg_P*lysE* | CACACTACCATCGGCGCTACCGAAGCTGCCTTCATCAATGATTGAGAGCAAAGTGTC | pSC*Cg*-LysG-S/M/W |
| Rev_Cg_ P*lysE* | TGAACAGCTCCTCGCCCTTGCTCACCATGCTATGATATCTCCTTCTTAAAGTTCATCTAGGTCCGATGGACAGTAAAAGACTGGCCCCCAAAAG | pSC*Cg*-LysG-S/M/W |
| Fwd_*lysG*_term | CATTGATGAAGGCAGCTTCGGTAGCGCCGATGGTAGTG | pSC*Cg*-LysG-S/M/W |
| Rev_*lysG*_term | ACAACCCCGCAAAAAACCTACATGAGCGGATACATATTTGAATGTATTTAG | pSC*Cg*-LysG-S/M/W |
| Fwd_P constitutive S *lysG* | TAGGTTTTTTGCGGGGTTGTTTAACCCCCAAATGAGGGAAGAAGGTATAATTGAACTCTGAACTTTAAGAAGGAGATATCATATGAACCCCATTCAACTGGACACTTTGCTCTCAATC | pSC*Cg*-LysG-S |
| Rev_P constitutive S/M/W *lysG* | TTGTTGCCATTGCTGCAGGTCGACTCTAAGGCCGCAATCCCTCGATTGCTG | pSC*Cg*-LysG-S/M/W |
| Fwd_P constitutive S/MW *lysG* overhang | ATTCAAATATGTATCCGCTCATGTAGGTTTTTTGCGGGGTTGTTTAACCCCCAAATG | pSC*Cg*-LysG-S/M/W |
| Fwd_P constitutive M *lysG* | TAGGTTTTTTGCGGGGTTGTTTAACCCCCAAATGAGGGAAGAAGGTATCCTTGAACTCTGAACTTTAAGAAGGAGATATCATATGAACCCCATTCAACTGGACACTTTGCTCTCAATC | pSC*Cg*-LysG-M |
| Fwd_P constitutive W *lysG* | TAGGTTTTTTGCGGGGTTGTTTAACCCCCAAATGAGGGAAGAAGGAAACCATGAACTCTGAACTTTAAGAAGGAGATATCATATGAACCCCATTCAACTGGACACTTTGCTCTCAATC | pSC*Cg*-LysG-W |
| Fwd_Cg_P*phdB* | CACACTACCATCGGCGCTACAGTGGCTCCATGTGAACTGGCTGAAAAATAGTTTCG | pSC*Cg*-PhdR-S/M/W |
| Rev_Cg_ P*phdB* | CTCGCCCTTGCTCACCATGCTATATCTCCTTCTTAAAGTTCAGGAACGACCAAGTCCTGCACCAGATCC | pSC*Cg*-PhdR-S/M/W |
| Fwd_Cg_*phdR*_term | ACAACCCCGCAAAAAACCTACATGAGCGGATACATATTTGAATGTATTTAG | pSC*Cg*-PhdR-S/M/W |

7

| Rev_Cg_*phdR*_term | CCAGTTCACATGGAGCCACTGTAGCGCCGAT GGTAGTG | pSC$_{Cg}$-PhdR-S/M/W |
| Fwd_Cg_P$_{constitutive}$ S *phdR* | TAGGTTTTTTGCGGGGTTGTTTAACCCCCAAA TGAGGGAAGAAGGTATAATTGAACTCTGAACT TTAAGAAGGAGATATCATATGACCACCAGCAA CCCCACCGCCGAGATCATTG | pSC$_{Cg}$-PhdR-S |
| Rev_Cg_P$_{constitutive}$ S/M/W *phdR* | GCCATTGCTGCAGGTCGACTCATGTGAACAT GGCCGGCGTGGTTAAGAGAC | pSC$_{Cg}$-PhdR-S/M/W |
| Fwd_Cg_P$_{constitutive}$ S/MW *phdR* overhang | CAAATATGTATCCGCTCATGTAGGTTTTTTGC GGGGTTGTTTAACCCCCAAATG | pSC$_{Cg}$-PhdR-S/M/W |
| Fwd_Cg_P$_{constitutive}$ M *phdR* | TAGGTTTTTTGCGGGGTTGTTTAACCCCCAAA TGAGGGAAGAAGGTATCCTTGAACTCTGAACT TTAAGAAGGAGATATCATATGACCACCAGCAA CCCCACCGCCGAGATCATTG | pSC$_{Cg}$-PhdR-M |
| Fwd_Cg_P$_{constitutive}$ W *phdR* | TAGGTTTTTTGCGGGGTTGTTTAACCCCCAAA TGAGGGAAGAAGGAAACCATGAACTCTGAAC TTTAAGAAGGAGATATCATATGACCACCAGCA ACCCCACCGCCGAGATCATTG | pSC$_{Cg}$-PhdR-W |
| Fwd_Ec_P$_{phdB}$ | CACCATAATTAACCTCCTTAAGCTTGGCATGC AGGAACGACCAAGTCCTGCACCAGATCC | pSC$_{Ec}$-PhdR-S/M/W |
| Rev_Ec_P$_{phdB}$ | GATTACAGCGTAAATGCCGTAGTGGCTCCATG TGAACTGGCTGAAAAAATAGTTTCG | pSC$_{Ec}$-PhdR-S/M/W |
| Fwd_Ec_*phdR*_ P$_{constitutive}$_S/M1/M2/W_t erm | CCAGTTCACATGGAGCCACTACGGCATTTACG CTGTAATCACACTGGCTCAC | pSC$_{Ec}$-PhdR-S/M/W |
| Rev_Ec_*phdR*_ P$_{constitutive}$_S_term | GTGGGGTTGCTGGTGGTCATCATGCTATTCCT CCTTAGGTCAGTGCGTCCTGCTGATG | pSC$_{Ec}$-PhdR-S |
| Rev_Ec_*phdR*_ P$_{constitutive}$_M1/_term | ATGTGCCCAGTGTCTCTATCACTGATAGGGAT GTCAATCCCTATCACTGATAGGGACTCGAGGT GAAGACGAGAGGGCCTCGTGATACGC | pSC$_{Ec}$-PhdR-M1 |
| Rev_Ec_*phdR*_ P$_{constitutive}$_M1_term | GTGGGGTTGCTGGTGGTCATCATGCTATTCCT CCTTAGGTCAGTGCGTCCTGCTGATGTGCCC AGTGTCTCTATCACTGATAGGGATGTCAATCC CTATCACTGATAGGGACTC | pSC$_{Ec}$-PhdR-M1 |
| Rev_Ec_*phdR*_ P$_{constitutive}$_M2_term | GTGGGGTTGCTGGTGGTCATCATGCTATTCCT CCTTAGGTCAGTGCGTCCTGCTGATGTGCTCA GTGTCTCTATCACTGATAGGGATGTCAATCTC TATCACTGATAGGGACTC | pSC$_{Ec}$-PhdR-M2 |
| Rev_Ec_*phdR*_ P$_{constitutive}$_W_term | GTGGGGTTGCTGGTGGTCATCATGCTATTCCT CCTTAGGTCAGTGCGTCCTGCTGATGTGCTCA GTATCTCTATCACTGATAGGGAGGTCAATCTC | pSC$_{Ec}$-PhdR-W |
| Fwd_Ec_*phdR* | ACCTAAGGAGGAATAGCATGATGACCACCAG CAACCCCACCGCCGAGATCATTG | pSC$_{Ec}$-PhdR-S/M/W |

| | | |
|---|---|---|
| Rev_Ec_*phdR* | TAAGGGATTTTGGTCATGAACATGTGAACATG GCCGGCGTGGTTAAGAGAC | pSC$_{Ec}$-PhdR-S/M/W |
| Fwd_Ec_P$_{lysE}$ | GATTACAGCGTAAATGCCGTCGAAGCTGCCTT CATCAATGATTGAGAGCAAAGTGTC | pSC$_{Ec}$-LysG-S/M/W |
| Rev_Ec_P$_{lysE}$ | caccATAATTAACCTCCTTAAGCTTGGCATGCT CTAGGTCCGATGGACAGTAAAAGACTGGCCC CCAAAAG | pSC$_{Ec}$-LysG-S/M/W |
| Fwd_Ec_*lysG*_P$_{constitutive}$_S_term | TCCAGTTGAATGGGGTTCATGCTATTCCTCCT TAGGTCAGTGCGTCCTGCTGATG | pSC$_{Ec}$-LysG-S |
| Rev_Ec_*lysG*_P$_{constitutive}$_S_term | CATTGATGAAGGCAGCTTCGACGGCATTTACG CTGTAATCACACTGGCTCAC | pSC$_{Ec}$-LysG-S |
| Fwd_Ec_*lysG*_P$_{constitutive}$_M1_term_P$_{lysE}$_1 | ATGTGCCCAGTGTCTCTATCACTGATAGGGAT GTCAATCCCTATCACTGATAGGGACTCGAGGT GAAGACGAGAGGGCCTCGTGATACGC | pSC$_{Ec}$-LysG-M1 |
| Fwd_Ec_*lysG*_P$_{constitutive}$_M1_term_P$_{lysE}$_2 | TGTCCAGTTGAATGGGGTTCATGCTATTCCTC CTTAGGTCAGTGCGTCCTGCTGATGTGCCCA GTGTCTCTATCACTGATAGGGATGTCAATCCC TATCACTGATAGGGACTC | pSC$_{Ec}$-LysG-M2 |
| Fwd_Ec_*lysG*_P$_{constitutive}$_M2_term_P$_{lysE}$ | TGTCCAGTTGAATGGGGTTCATGCTATTCCTC CTTAGGTCAGTGCGTCCTGCTGATGTGCTCA GTGTCTCTATCACTGATAGGGATGTCAATCTC | pSC$_{Ec}$-LysG-M2 |
| Fwd_Ec_*lysG*_P$_{constitutive}$_W_term_P$_{lysE}$ | AGCGTAAATGCCGTAGTTCCATGCTATTCCTC CTTAGGTCAGTGCGTCCTGCTGATGTGCTCA GTATCTCTATCACTGATAGGGAGGTCAATCTC TATCACTGATAGGGACTC | pSC$_{Ec}$-LysG-W |

9

**Supplementary Table S4. Constitutive Promoters of the unified sensor design.**

| Promoter | Promoter strength | Organism | Sequence |
|---|---|---|---|
| PLTetO1 | Strong | *E. coli* | TCTTCACCTCGAGTCCCTATCAGTGATAGAGAT TGACATCCCTATCAGTGATAGAGATACTGAGCA CATCAGCAGGACGCACTGACC |
| PLTetO1JJ | Weak | *E. coli* | CAATTCCGACGTCTAAGAAACCATTATTATCATG ACATTAACCTATAAAAATAGGCGTATCACGAGG CCCTTTCGTCTTCACCTCGAGTCCCTATCAGTG ATAGAGATTGACCTCCCTATCAGTGATAGAGAT ACTGAGCACATCAGCAGGACGCACTGACC |
| PLTetO1S | Moderate 1 | *E. coli* | CAATTCCGACGTCTAAGAAACCATTATTATCATG ACATTAACCTATAAAAATAGGCGTATCACGAGG CCCTTTCGTCTTCACCTCGAGTCCCTATCAGTG ATAGAGATTGACATCCCTATCAGTGATAGAGAC ACTGAGCACATCAGCAGGACGCACTGACC |
| PLTetO1K | Moderate 2 | *E. coli* | CAATTCCGACGTCTAAGAAACCATTATTATCATG ACATTAACCTATAAAAATAGGCGTATCACGAGG CCCTCTCGTCTTCACCTCGAGTCCCTATCAGTG ATAGGGATTGACATCCCTATCAGTGATAGAGAC ACTGGGCACATCAGCAGGACGCACTGACC |
| *dapA* A16 | Strong | *C. glutamicum* | TAGGTTTTTTGCGGGGTTGTTTAACCCCCAAAT GAGGGAAGAAGGTATAATTGAACTCT |
| *dapA* A14 | Moderate | *C. glutamicum* | TAGGTTTTTTGCGGGGTTGTTTAACCCCCAAAT GAGGGAAGAAGGTATCCTTGAACTCT |
| *dapA* B27 | Weak | *C. glutamicum* | TAGGTTTTTTGCGGGGTTGTTTAACCCCCAAAT GAGGGAAGAAGGAAACCATGAACTCT |

**Figure S1: Exemplary annotated DNA sequences of the important biosensor components of pSC$_{Cg}$-PhdR-S and pSC$_{Ec}$-PhdR-S.**
All genes and the corresponding amino acid sequence (green) as well as all elements of the unified sensor design (red) are indicated
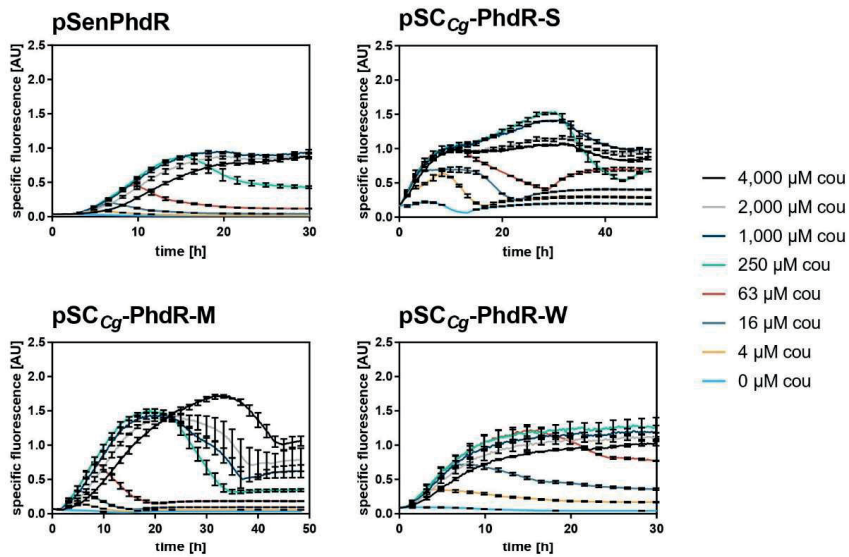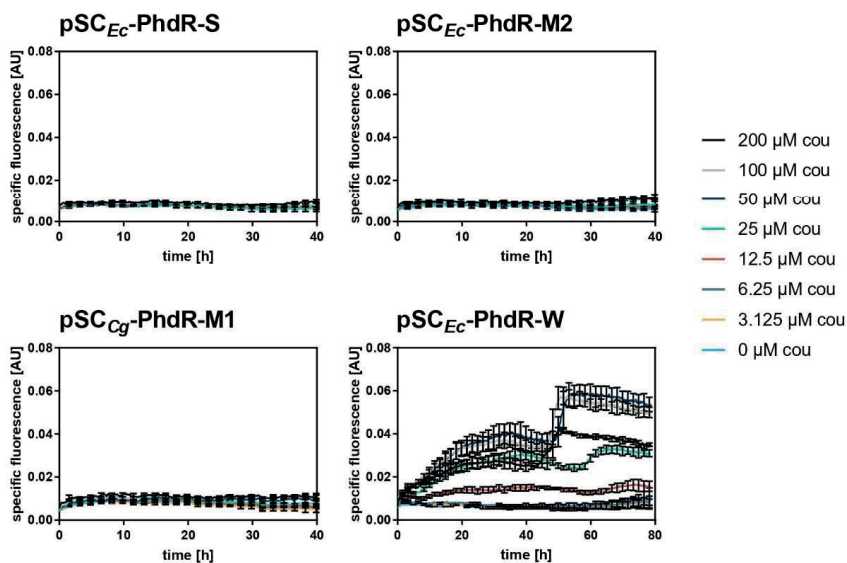
**Figure S2: Specific fluorescence response of PhdR-based biosensors in *C. glutamicum*.**
BioLector cultivations with *C. glutamicum* strains carrying pSC*Cg*-PhdR-S/M/W or pSenPhdR
with eight different inducer concentrations ranging from 4 μM -4,000 μM coumaric acid
(externally supplemented) n=3.
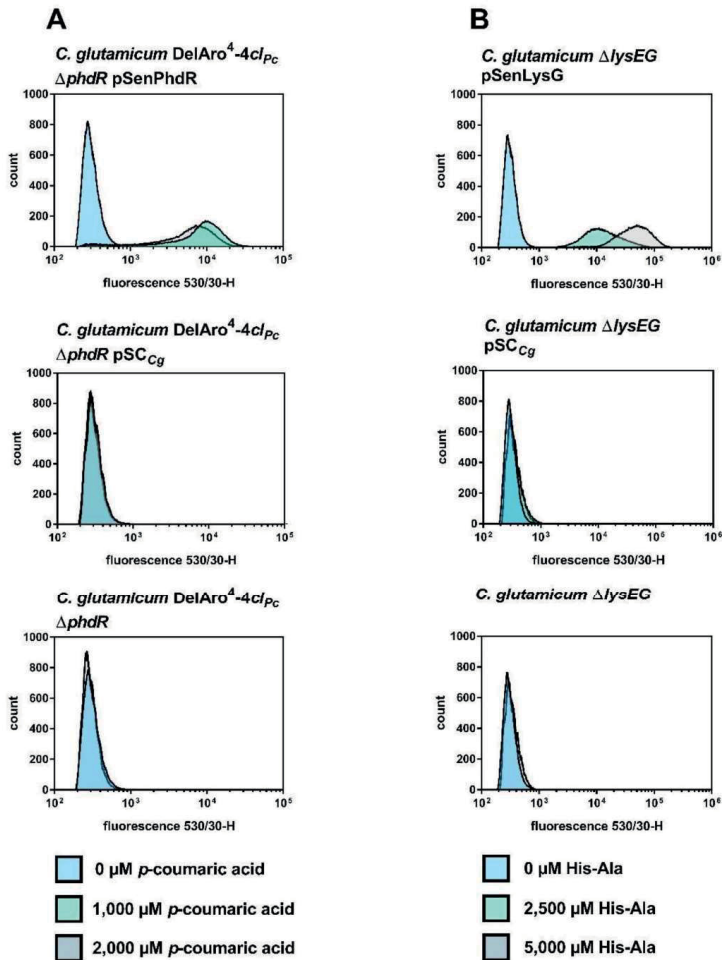
**Figure S3: Basal fluorescence response of LysG- and PhdR-based biosensors in *C. glutamicum* using FACS. (A)** Characterization of fluorescence response of *C. glutamicum* DelAro⁴-4*cl* Δ*phdR*, *C. glutamicum* DelAro⁴-4*cl* Δ*phdR* pSC$_{Cg}$ carrying the sensor backbone pSC$_{Cg}$ and *C. glutamicum* DelAro⁴-4*cl* Δ*phdR* carrying the pSenPhdR sensor with and without inducer *p*-coumaric acid. **(B)** Characterization of fluorescence response of strain background *C. glutamicum* Δ*lysEG*, *C. glutamicum* Δ*lysEG* pSC$_{Cg}$ carrying the sensor backbone pSC$_{Cg}$ and *C. glutamicum* Δ*lysEG* carrying the pSenLysG sensor with and without inducer His-Ala. In all experiments, 95,000 representative single cells were analyzed by FACS.

13
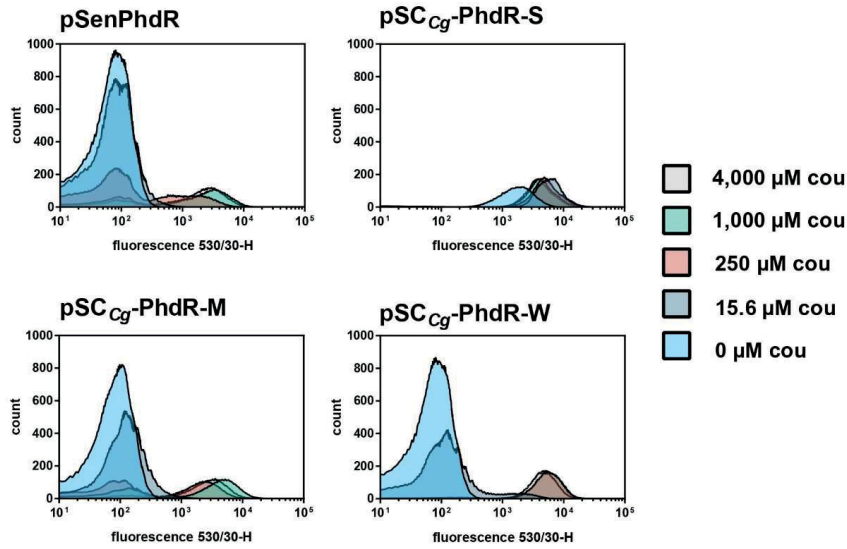
**Figure S4: Biosensor response of PhdR-based biosensors in *C. glutamicum*.** FACS experiments with *C. glutamicum* strains carrying pSC$_{Cg}$-PhdR-S/M/W or pSenPhdR in the presence of externally supplemented 0 - 4,000 μM *p*-coumaric acid (cou). In each case, 95,000 representative single cells were analyzed.

14
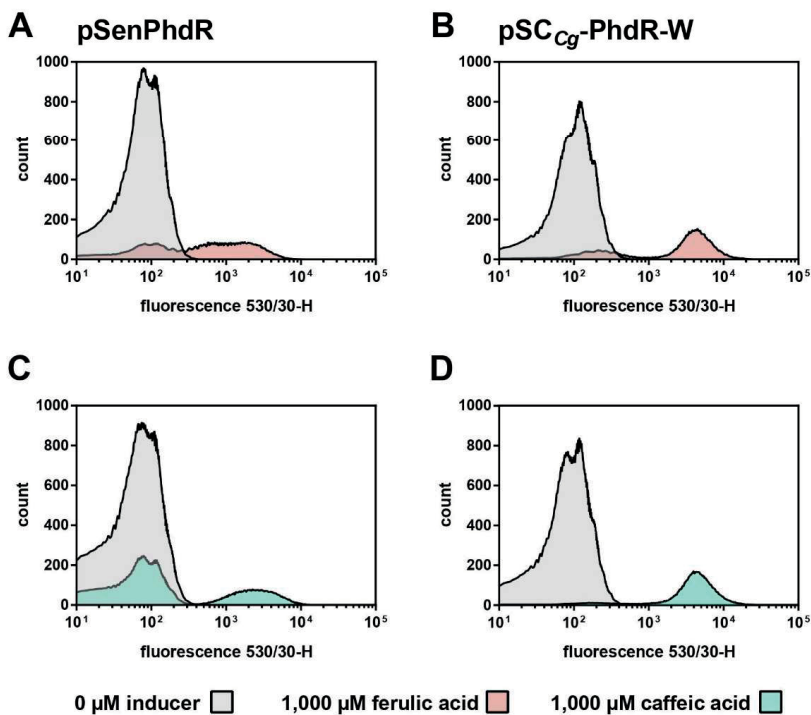
**Figure S5: Substrate specificity of PhdR-based biosensors in *C. glutamicum*.** FACS experiments with *C. glutamicum* strains carrying **(A/C)** pSenPhdR or **(B/D)** pSC$_{Cg}$-PhdR-W in the presence of externally supplemented 0 (grey) and 1,000 μM inducers ferulic acid (pink) and caffeic acid (green). In each case, 95,000 representative single cells were analyzed.

15
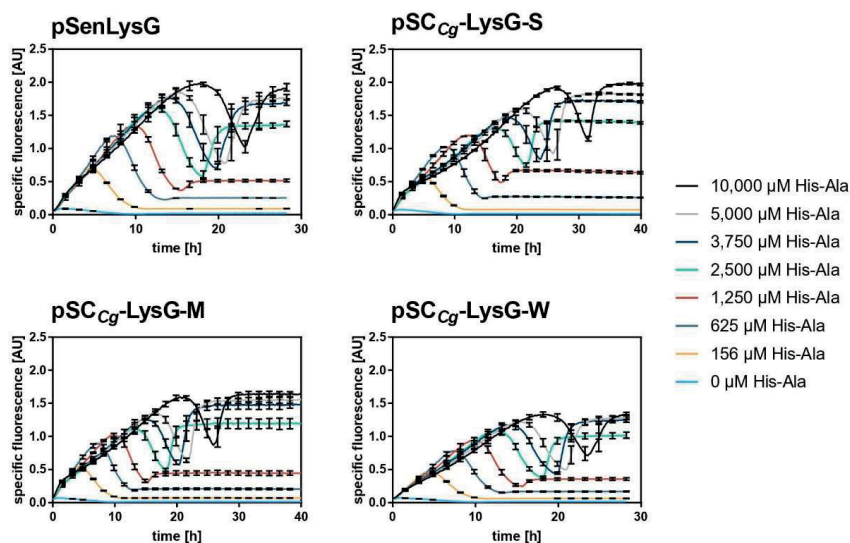
**Figure S6: Specific fluorescence response of LysG-based biosensors in *C. glutamicum.***
BioLector cultivations with *C. glutamicum* strains carrying pSC$_{Cg}$-LysG-S/M/W or pSenLysG with eight different inducer concentrations ranging from 125 µM -10,000 µM His-Ala dipeptides (externally supplemented) n=3.

16
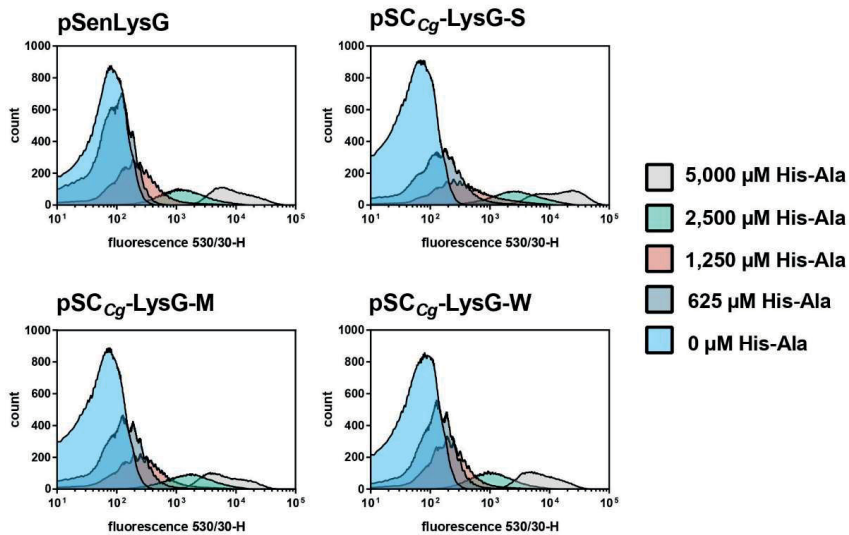
**Figure S7: Biosensor response of LysG-based biosensors in *C. glutamicum*.** FACS experiments with *C. glutamicum* strains carrying pSC$_{Cg}$-LysG-S/M/W or pSenLysG in the presence of externally supplemented 0 - 5,000 µM His-Ala dipeptides. In each case, 95,000 representative single cells were analyzed.
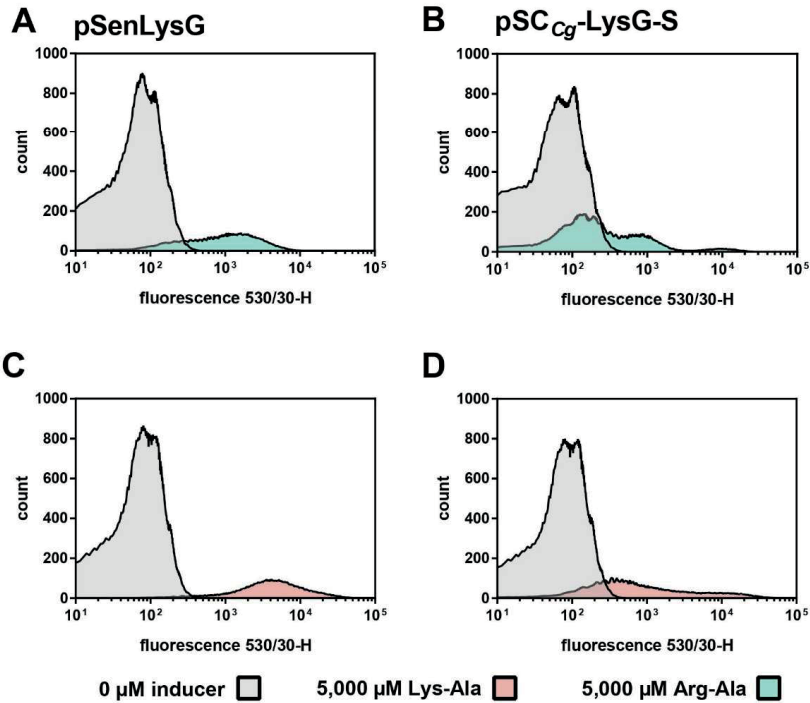
**Figure S8: Substrate specificity of LysG-based biosensors in *C. glutamicum*.** FACS experiments with *C. glutamicum* strains carrying **(A/C)** pSenLysG or **(B/D)** pSC$_{Cg}$-LysG-S in the presence of externally supplemented 0 (grey) and 5,000 μM dipeptide inducers Lys-Ala (pink) and Arg-Ala (green). In each case, 95,000 representative single cells were analyzed.

18

**Figure S9: Specific fluorescence response of PhdR-based biosensors in *E. coli*.** BioLector cultivation with *E. coli* strains carrying pSC*Ec*-PhdR-S/M1/M2/W in the presence of externally supplemented 0 - 200 µM *p*-coumaric acid (cou).

19
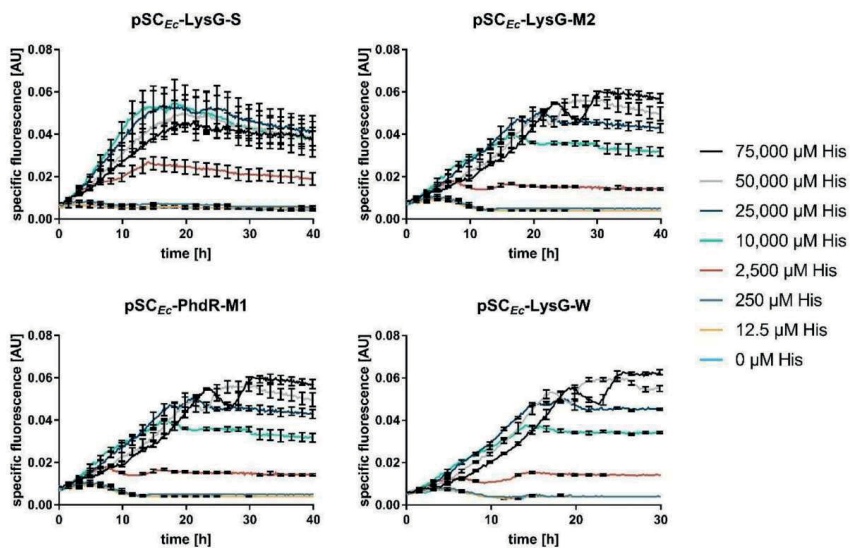
**Figure S10: Specific fluorescence response of LysG-based biosensors in *E. coli*.** BioLector cultivation with *E. coli* strains carrying pSC*Ec*-LysG-S/M1/M2/W in the presence of externally supplemented 0 - 75,000 µM L-histidine (His).
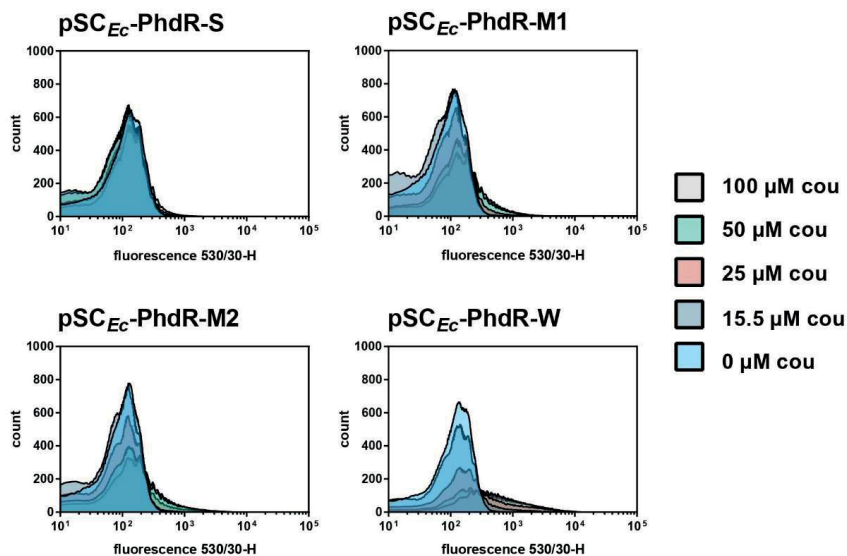
**Figure S11: Biosensor response of PhdR-based biosensors in *E. coli*.** FACS experiments with *E. coli* strains carrying pSC$_{Ec}$-PhdR-S/M1/M2/W in the presence of externally supplemented 0 - 100 µM *p*-coumaric acid (cou). In each case, 95,000 representative single cells were analyzed.
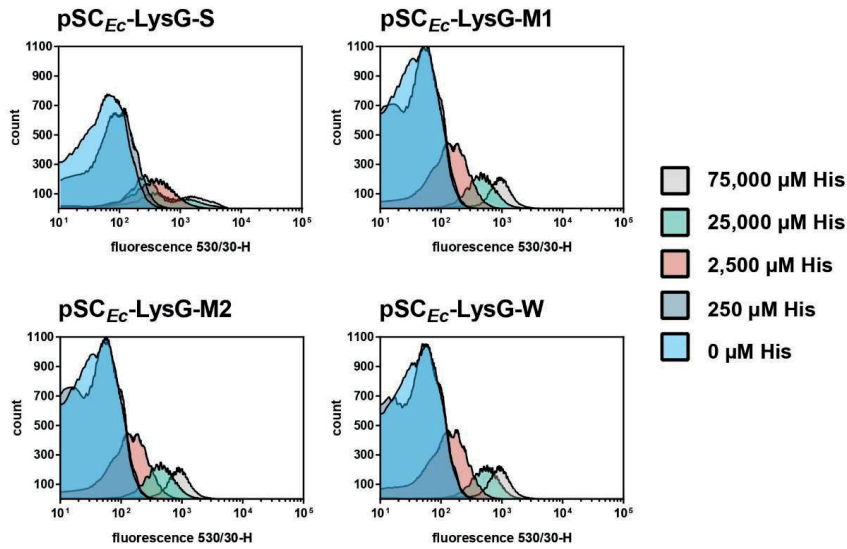
**Figure S12: Biosensor response of LysG-based biosensors in *E. coli*.** FACS experiments with *E. coli* strains carrying pSC*Ec*-LysG-S/M1/M2/W in the presence of externally supplemented 0 - 75,000 µM L-histidine (His). In each case, 95,000 representative single cells were analyzed.

22

# References

Binder, S., Schendzielorz, G., Stäbler, N., Krumbach, K., Hoffmann, K., Bott, M., Eggeling, L., 2012. A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. Genome Biol. 13, R40. https://doi.org/10.1186/gb-2012-13-5-r40

Flachbart, L.K., Sokolowsky, S., Marienhagen, J., 2019. Displaced by deceivers: Prevention of biosensor cross-talk is pivotal for successful biosensor-based high-throughput screening campaigns. ACS Synth. Biol. 8, 1847–1857. https://doi.org/10.1021/acssynbio.9b00149

Kallscheuer, N., Vogt, M., Stenzel, A., Gätgens, J., Bott, M., Marienhagen, J., 2016. Construction of a *Corynebacterium glutamicum* platform strain for the production of stilbenes and (2S)-flavanones. Metab. Eng. 38, 47–55. https://doi.org/10.1016/j.ymben.2016.06.003

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

van Summeren-Wesenhagen, P.V., Marienhagen, J., 2015. Metabolic engineering of *Escherichia coli* for the synthesis of the plant polyphenol pinosylvin. Appl. Environ. Microbiol. 81, 840–849. https://doi.org/10.1128/AEM.02966-14

Vrljic, M., Sahm, H., Eggeling, L., 1996. A new type of transporter with a new type of cellular function: L-lysine export from *Corynebacterium glutamicum*. Mol. Microbiol. 22, 815–826. https://doi.org/10.1046/j.1365-2958.1996.01527.x

## 9    Literature

**Adli, M.**, 2018. "The CRISPR tool kit for genome editing and beyond." *Nature Communications*, 9 (1). https://doi.org/10.1038/s41467-018-04252-2

**Adrio, J.-L. ; Demain, A.L.**, 2010. "Recombinant organisms for production of industrial products." *Bioengineered Bugs*, 1 (2), 116–131. https://doi.org/10.4161/bbug.1.2.10484

**Alper, H. ; Fischer, C. ; Nevoigt, E. ; Stephanopoulos, G. ; Demasi, J. ; Huh, K. ; Nakatani, Y. ; Mu, K.**, 2006. "Tuning genetic control through promoter engineering." *Pnas*, 103 (36), 3006–3007. https://doi.org/10.1073/pnas.0507062103

**Baritugo, K.-A. ; Kim, H.T. ; David, Y. ; Choi, J. ; Hong, S.H. ; Jeong, K.J. ; Choi, J.H. ; Joo, J.C. ; Park, S.J.**, 2018. "Metabolic engineering of *Corynebacterium glutamicum* for fermentative production of chemicals in biorefinery." *Applied Microbiology and Biotechnology*, 102 (9), 3915–3937. https://doi.org/10.1007/s00253-018-8896-6

**Barrangou, R. ; Fremaux, C. ; Deveau, H. ; Richards, M. ; Boyaval, P. ; Moineau, S. ; Romero, D. a ; Horvath, P.**, 2007. "CRISPR provides acquired resistance against viruses in prokaryotes." *Science*, 315 (5819), 1709–1712. https://doi.org/10.1126/science.1138140

**Becker, J. ; Rohles, C.M. ; Wittmann, C.**, 2018. "Metabolically engineered *Corynebacterium glutamicum* for bio-based production of chemicals, fuels, materials, and healthcare products." *Metabolic Engineering*, 50 (April), 122–141. https://doi.org/10.1016/j.ymben.2018.07.008

**Becker, M. ; Börngen, K. ; Nomura, T. ; Battle, A.R. ; Marin, K. ; Martinac, B. ; Krämer, R.**, 2013. "Glutamate efflux mediated by *Corynebacterium glutamicum* MscCG, *Escherichia coli* MscS, and their derivatives." *Biochimica et biophysica acta*, 1828 (4), 1230–40. https://doi.org/10.1016/j.bbamem.2013.01.001

**Bellman, A. ; Vrljić, M. ; Pátek, M. ; Sahm, H. ; Krämer, R. ; Eggeling, L.**, 2001. "Expression control and specificity of the basic amino acid exporter *lysE* of *Corynebacterium glutamicum*." *Microbiology*, 147 (7), 1765–1774. https://doi.org/10.1099/00221287-147-7-1765

**Berens, C. ; Thain, A. ; Schroeder, R.**, 2001. "A tetracycline-binding RNA aptamer." *Bioorganic and Medicinal Chemistry*, 9 (10), 2549–2556. https://doi.org/10.1016/S0968-0896(01)00063-3

**Bermejo, C. ; Haerizadeh, F. ; Takanaga, H. ; Chermak, D. ; Frommer, W.B.**, 2011. "Optical sensors for measuring dynamic changes of cytosolic metabolite levels in yeast." *Nature Protocols*, 6 (11), 1806–1817. https://doi.org/10.1038/nprot.2011.391

**Bikard, D. ; Marraffini, L.A.**, 2012. "Innate and adaptive immunity in bacteria: Mechanisms of programmed genetic variation to fight bacteriophages." *Current Opinion in Immunology*, 24 (1), 15–20. https://doi.org/10.1016/j.coi.2011.10.005

**Billinton, N. ; Barker, M.G. ; Michel, C.E. ; Knight, A.W. ; Heyer, W.D. ; Goddard, N.J. ; Fielden, P.R. ; Walmsley, R.M.**, 1998. "Development of a green fluorescent protein reporter for a yeast genotoxicity biosensor." *Biosensors and Bioelectronics*, 13 (7–8), 831–838. https://doi.org/10.1016/S0956-5663(98)00049-9

**Binder, S. ; Schendzielorz, G. ; Stäbler, N. ; Krumbach, K. ; Hoffmann, K. ; Bott, M. ; Eggeling, L.**, 2012. "A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level." *Genome Biology*, 13 (5), R40. https://doi.org/10.1186/gb-2012-13-5-r40

**Binder, S. ; Siedler, S. ; Marienhagen, J. ; Bott, M. ; Eggeling, L.**, 2013. "Recombineering in *Corynebacterium glutamicum* combined with optical nanosensors: A general strategy for fast producer strain generation." *Nucleic acids research*, 41 (12), 6360–9. https://doi.org/10.1093/nar/gkt312

**Buijs, N.A. ; Siewers, V. ; Nielsen, J.**, 2013. "Advanced biofuel production by the yeast *Saccharomyces cerevisiae*." *Current Opinion in Chemical Biology*, 17 (3), 480–488. https://doi.org/10.1016/j.cbpa.2013.03.036

**Charpentier, E. ; Richter, H. ; van der Oost, J. ; White, M.F.**, 2015. "Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity." *FEMS Microbiology Reviews*, 39 (3), 428–441. https://doi.org/10.1093/femsre/fuv023

**Chen, J.S. ; Ma, E. ; Harrington, L.B. ; Da Costa, M. ; Tian, X. ; Palefsky, J.M. ; Doudna, J.A.**, 2018. "CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity." *Science*, 360 (6387), 436–439. https://doi.org/10.1126/science.aar6245

**Chen, W. ; Zhang, S. ; Jiang, P. ; Yao, J. ; He, Y. ; Chen, L. ; Gui, X. ; Dong, Z. ; Tang, S.Y.**, 2015. "Design of an ectoine-responsive AraC mutant and its application in metabolic engineering of ectoine biosynthesis." *Metabolic Engineering*, 30 , 149–155. https://doi.org/10.1016/j.ymben.2015.05.004

**Chou, H.H. ; Keasling, J.D.**, 2013. "Programming adaptive control to evolve increased metabolite production." *Nature Communications*, 4 (1), 2595. https://doi.org/10.1038/ncomms3595

**Constantinou, A. ; Polizzi, K.M.**, 2013. "Opportunities for bioprocess monitoring using FRET biosensors." *Biochemical Society Transactions*, 41 (5), 1146–1151. https://doi.org/10.1042/BST20130103

**Dai, Z. ; Nielsen, J.**, 2015. "Advancing metabolic engineering through systems biology of industrial microorganisms." *Current Opinion in Biotechnology*, 36 , 8–15. https://doi.org/10.1016/j.copbio.2015.08.006

**de Lorenzo, V. ; Fernández, S. ; Herrero, M. ; Jakubzik, U. ; Timmis, K.N.**, 1993. "Engineering of alkyl- and haloaromatic-responsive gene expression with mini-transposons containing regulated promoters of biodegradative pathways of *Pseudomonas*." *Gene*, 130 (1), 41–46. https://doi.org/10.1016/0378-1119(93)90344-3

**De Paepe, B. ; Maertens, J. ; Vanholme, B. ; De Mey, M.**, 2018. "Modularisation and response curve engineering of a naringenin-responsive transcriptional biosensor." *ACS Synthetic Biology – Revised version submitted*, 7 , 1303–1314. https://doi.org/10.1021/acssynbio.7b00419

**De Paepe, B. ; Peters, G. ; Coussement, P. ; Maertens, J. ; De Mey, M.**, 2017. "Tailor-made transcriptional biosensors for optimizing microbial cell factories." *Journal of Industrial Microbiology & Biotechnology*, 44 (4–5), 623–645. https://doi.org/10.1007/s10295-016-1862-3

**Deveau, H. ; Barrangou, R. ; Garneau, J.E. ; Labonté, J. ; Fremaux, C. ; Boyaval, P. ; Romero, D.A. ; Horvath, P. ; Moineau, S.**, 2008. "Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*." *Journal of Bacteriology*, 190 (4), 1390–1400. https://doi.org/10.1128/JB.01412-07

**Dietrich, J.A. ; McKee, A.E. ; Keasling, J.D.**, 2010. "High-throughput metabolic engineering: Advances in small-molecule screening and selection." *Annual Review of Biochemistry*, 79 (1), 563–590. https://doi.org/10.1146/annurev-biochem-062608-095938

**Dong, D. ; Ren, K. ; Qiu, X. ; Zheng, J. ; Guo, M. ; Guan, X. ; Liu, H. ; Li, N. ; Zhang, B. ; Yang, D. ; Ma, C. ; Wang, S. ; Wu, D. ; Ma, Y. ; Fan, S. ; Wang, J. ; Gao, N. ; Huang, Z.**, 2016. "The crystal structure of Cpf1 in complex with CRISPR RNA." *Nature*, 532 (7600), 522–526. https://doi.org/10.1038/nature17944

**Doudna, J.A. ; Charpentier, E.**, 2014. "The new frontier of genome engineering with CRISPR-Cas9." *Science*, 346 (6213). https://doi.org/10.1126/science.1258096

**Eggeling, L.**, 1994. "Biology of L-lysine overproduction by *Corynebacterium glutamicum*." *Amino Acids*, 6 (3), 261–272. https://doi.org/10.1007/BF00813746

**Eggeling, L. ; Bott, M.**, 2015. "A giant market and a powerful metabolism: L-lysine provided by *Corynebacterium glutamicum*." *Applied microbiology and biotechnology*, 99 (8), 3387–94. https://doi.org/10.1007/s00253-015-6508-2

**Eggeling, L. ; Bott, M.**, 2005. Handbook of Corynebacterium glutamicum. CRC Press. https://doi.org/10.1201/9781420039696

**Eggeling, L. ; Bott, M. ; Marienhagen, J.**, 2015. "Novel screening methods-biosensors." *Current Opinion in Biotechnology*, 35 (ii), 30–36. https://doi.org/10.1016/j.copbio.2014.12.021

**Eggeling, L. ; Sahm, H.**, 1999. "L-glutamate and l-lysine: traditional products with impetuous developments." *Applied Microbiology and Biotechnology*, 52 (2), 146–153. https://doi.org/10.1007/s002530051501

**Eikmanns, B.J. ; Kleinertz, E. ; Liebl, W. ; Sahm, H.**, 1991. "A family of *Corynebacterium glutamicum*/*Escherichia coli* shuttle vectors for cloning, controlled gene expression, and promoter probing." *Gene*, 102 (1), 93–98. https://doi.org/10.1016/0378-1119(91)90545-M

**Ellington, A.D. ; Szostak, J.W.**, 1990. "In vitro selection of RNA molecules that bind specific ligands." *Nature*, 346 (6287), 818–822. https://doi.org/10.1038/346818a0

**Eurostat**, 2019a. Oil and petroleum products - a statistical overview [WWW Document]. *Oil and petroleum products - a statistical overview*,. URL https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Oil_and_petroleum_products_-_a_statistical_overview (accessed 2.21.20).

**Eurostat**, 2019b. Climate change - driving forces [WWW Document]. *Climate change - driving forces*,. URL https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Climate_change_-_driving_forces#CO2_emission_intensities_of_economic_activities (accessed 2.21.20).

**Fernandez-López, R. ; Ruiz, R. ; de la Cruz, F. ; Moncalián, G.**, 2015. "Transcription factor-based biosensors enlightened by the analyte." *Frontiers in Microbiology*, 6 (JUL), 1–21. https://doi.org/10.3389/fmicb.2015.00648

**Flachbart, L.K. ; Sokolowsky, S. ; Marienhagen, J.**, 2019. "Displaced by deceivers: Prevention of biosensor cross-talk is pivotal for successful biosensor-based high-throughput screening campaigns." *ACS Synthetic Biology*, 8 (8), 1847–1857. https://doi.org/10.1021/acssynbio.9b00149

**Fonfara, I. ; Richter, H. ; BratoviÄ, M. ; Le Rhun, A. ; Charpentier, E.**, 2016. "The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA." *Nature*, 532 (7600), 517–521. https://doi.org/10.1038/nature17945

**Frei, C.S. ; Qian, S. ; Cirino, P.C.**, 2018. "New engineered phenolic biosensors based on the AraC regulatory protein." *Protein Engineering, Design and Selection*, 31 (6), 213–220. https://doi.org/10.1093/protein/gzy024

**Frommer, W.B. ; Davidson, M.W. ; Campbell, R.E.**, 2009. "Genetically encoded biosensors based on engineered fluorescent proteins." *Chemical Society Reviews*, 38 (10), 2833–2841. https://doi.org/10.1039/b907749a

**Gao, P. ; Yang, H. ; Rajashankar, K.R. ; Huang, Z. ; Patel, D.J.**, 2016. "Type V CRISPR-Cas Cpf1 endonuclease employs a unique mechanism for crRNA-mediated target DNA recognition." *Cell Research*, 26 (8), 901–913. https://doi.org/10.1038/cr.2016.88

**Garneau, J.E. ; Dupuis, M.-È. ; Villion, M. ; Romero, D.A. ; Barrangou, R. ; Boyaval, P. ; Fremaux, C. ; Horvath, P. ; Magadán, A.H. ; Moineau, S.**, 2010. "The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA." *Nature*, 468 (7320), 67–71. https://doi.org/10.1038/nature09523

**Grünberger, A. ; van Ooyen, J. ; Paczia, N. ; Rohe, P. ; Schiendzielorz, G. ; Eggeling, L. ; Wiechert, W. ; Kohlheyer, D. ; Noack, S.**, 2013. "Beyond growth rate 0.6: *Corynebacterium glutamicum* cultivated in highly diluted environments." *Biotechnology and Bioengineering*, 110 (1), 220–228. https://doi.org/10.1002/bit.24616

**Hallberg, Z.F. ; Su, Y. ; Kitto, R.Z. ; Hammond, M.C.**, 2017. "Engineering and *in vivo* applications of riboswitches." *Annual Review of Biochemistry*, 86 (1), 515–539. https://doi.org/10.1146/annurev-biochem-060815-014628

**Hanson, S. ; Bauer, G. ; Fink, B. ; Suess, B.**, 2005. "Molecular analysis of a synthetic tetracycline-binding riboswitch." *RNA*, 11 (4), 503–511. https://doi.org/10.1261/rna.7251305

**Harmsen, P.F.H. ; Hackmann, M.M. ; Bos, H.L.**, 2014. "Green building blocks for bio-based plastics." *Biofuels, Bioproducts and Biorefining*, 8 (3), 306–324. https://doi.org/10.1002/bbb.1468

**Hashimoto, K.I. ; Murata, J. ; Konishi, T. ; Yabe, I. ; Nakamatsu, T. ; Kawasaki, H.**, 2012. "Glutamate is excreted across the cytoplasmic membrane through the NCgl1221 channel of *Corynebacterium glutamicum* by passive diffusion." *Bioscience, Biotechnology and Biochemistry*, 76 (7), 1422–1424. https://doi.org/10.1271/bbb.120366

**Hille, F. ; Richter, H. ; Wong, S.P. ; Bratovič, M. ; Ressel, S. ; Charpentier, E.**, 2018. "The biology of CRISPR-Cas: backward and forward." *Cell*, 172 (6), 1239–1259. https://doi.org/10.1016/j.cell.2017.11.032

**Hirasawa, T. ; Wachi, M.**, 2016. "Glutamate fermentation-2: Mechanism of L-glutamate overproduction in *Corynebacterium glutamicum*," in: Advances in Biochemical Engineering/Biotechnology. pp. 57–72. https://doi.org/10.1007/10_2016_26

**Hochreiter, B. ; Garcia, A.P. ; Schmid, J.A.**, 2015. "Fluorescent proteins as genetically encoded FRET biosensors in life sciences." *Sensors (Switzerland)*, 15 (10), 26281–26314. https://doi.org/10.3390/s151026281

**Hoffmann, S.L. ; Jungmann, L. ; Schiefelbein, S. ; Peyriga, L. ; Cahoreau, E. ; Portais, J.-C. ; Becker, J. ; Wittmann, C.**, 2018. "Lysine production from the sugar alcohol mannitol: Design of the cell factory *Corynebacterium glutamicum* SEA-3 through integrated analysis and engineering of metabolic pathway fluxes." *Metabolic Engineering*, 47 , 475–487. https://doi.org/10.1016/j.ymben.2018.04.019

**Ikariyama, Y. ; Nishiguchi, S. ; Koyama, T. ; Kobatake, E. ; Aizawa, M. ; Tsuda, M. ; Nakazawa, T.**, 1997. "Fiber-optic-based biomonitoring of benzene derivatives by recombinant *E. coli* bearing luciferase gene-fused TOL-plasmid immobilized on the fiber-optic end." *Analytical Chemistry*, 69 (13), 2600–2605. https://doi.org/10.1021/ac961311o

**Ikeda, M. ; Mitsuhashi, S. ; Tanaka, K. ; Hayashi, M.**, 2009. "Reengineering of a *Corynebacterium glutamicum* L-arginine and L-citrulline producer." *Applied and Environmental Microbiology*, 75 (6), 1635–1641. https://doi.org/10.1128/AEM.02027-08

**Ikeda, M. ; Takeno, S.**, 2013. "Amino acid production by *Corynebacterium glutamicum*," in: Corynebacterium Glutamicum. Springer, pp. 107–147.

**IMARC Group**, 2020. Amino Acids Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2019-2024 [WWW Document]. *Amino Acids Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2019-2024*,. URL https://www.imarcgroup.com/amino-acid-technical-material-market-report (accessed 2.14.20).

**Ivančić-Bace, I. ; Cass, S.D. ; Wearne, S.J. ; Bolt, E.L.**, 2015. "Different genome stability proteins underpin primed and naïve adaptation in *E. coli* CRISPR-Cas immunity." *Nucleic Acids Research*, 43 (22), 10821–10830. https://doi.org/10.1093/nar/gkv1213

**Jäger, W. ; Schäfer, A. ; Pühler, A. ; Labes, G. ; Wohlleben, W.**, 1992. "Expression of the *Bacillus subtilis sacB* gene leads to sucrose sensitivity in the gram-positive bacterium *Corynebacterium glutamicum* but not in *Streptomyces lividans*." *Journal of bacteriology*, 174 (16), 5462–5. https://doi.org/10.1128/jb.174.16.5462-5465.1992

**Jakoby, M. ; Carole-Estelle ; Ngouoto-Nkili ; Burkovski, A.**, 1999. "Construction and application of new *Corynebacterium glutamicum* vectors." *Biotechnology Techniques*, 13 (6), 437–441. https://doi.org/10.1023/A:1008968419217

**Jansen, R. ; Embden, J.D.A. van ; Gaastra, W. ; Schouls, L.M.**, 2002. "Identification of genes that are associated with DNA repeats in prokaryotes." *Molecular Microbiology*, 43 (6), 1565–1575. https://doi.org/10.1046/j.1365-2958.2002.02839.x

**Jenison, R.D. ; Gill, S.C. ; Pardi, A. ; Polisky, B.**, 1994. "High-resolution molecular discrimination by RNA." *Science*, 263 (5152), 1425–1429. https://doi.org/10.1126/science.7510417

**Jha, R.K. ; Kern, T.L. ; Fox, D.T. ; M Strauss, C.E.**, 2014. "Engineering an *Acinetobacter* regulon for biosensing and high-throughput enzyme screening in *E. coli* via flow cytometry." *Nucleic acids research*, 42 (12), 8150–60. https://doi.org/10.1093/nar/gku444

**Jiang, W. ; Bikard, D. ; Cox, D. ; Zhang, F. ; Marraffini, L.A.**, 2013. "RNA-guided editing of bacterial genomes using CRISPR-Cas systems." *Nature Biotechnology*, 31 (3), 233–239. https://doi.org/10.1038/nbt.2508

**Jiang, Y. ; Qian, F. ; Yang, J. ; Liu, Y. ; Dong, F. ; Xu, C. ; Sun, B. ; Chen, B. ; Xu, X. ; Li, Y. ; Wang, R. ; Yang, S.**, 2017. "CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum*." *Nature communications*, 8 (1), 15179. https://doi.org/10.1038/ncomms15179

**Kalinowski, J. ; Bathe, B. ; Bartels, D. ; Bischoff, N. ; Bott, M. ; Burkovski, A. ; Dusch, N. ; Eggeling, L. ; Eikmanns, B.J. ; Gaigalat, L. ; Goesmann, A. ; Hartmann, M. ; Huthmacher, K. ; Krämer, R. ; Linke, B. ; McHardy, A.C. ; Meyer, F. ; Möckel, B. ; Pfefferle, W. ; Pühler, A. ; Rey, D.A. ; Rückert, C. ; Rupp, O. ; Sahm, H. ; Wendisch, V.F. ; Wiegräbe, I. ; Tauch, A.**, 2003. "The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins." *Journal of Biotechnology*, 104 (1–3), 5–25. https://doi.org/10.1016/S0168-1656(03)00154-8

**Kalinowski, J. ; Cremer., J. ; Bachmann, B. ; Eggeling, L. ; Sahm, H. ; Pühler, A.**, 1991. "Genetic and biochemical analysis of the aspartokinase from *Corynebacterium glutamicum*." *Molecular Microbiology*, 5 (5), 1197–1204. https://doi.org/10.1111/j.1365-2958.1991.tb01893.x

**Kallscheuer, N. ; Vogt, M. ; Kappelmann, J. ; Krumbach, K. ; Noack, S. ; Bott, M. ; Marienhagen, J.**, 2016a. "Identification of the *phd* gene cluster responsible for phenylpropanoid utilization in *Corynebacterium glutamicum*." *Applied Microbiology and Biotechnology*, 100 (4), 1871–1881. https://doi.org/10.1007/s00253-015-7165-1

**Kallscheuer, N. ; Vogt, M. ; Stenzel, A. ; Gätgens, J. ; Bott, M. ; Marienhagen, J.**, 2016b. "Construction of a *Corynebacterium glutamicum* platform strain for the production of stilbenes and (2S)-flavanones." *Metabolic Engineering*, 38 (1), 47–55. https://doi.org/10.1016/j.ymben.2016.06.003

**Kieper, S.N. ; Almendros, C. ; Behler, J. ; McKenzie, R.E. ; Nobrega, F.L. ; Haagsma, A.C. ; Vink, J.N.A. ; Hess, W.R. ; Brouns, S.J.J.**, 2018. "Cas4 facilitates PAM-compatible spacer selection during CRISPR adaptation." *Cell Reports*, 22 (13), 3377–3384. https://doi.org/10.1016/j.celrep.2018.02.103

**Kinoshita, S. ; Udaka, S. ; Shimono, M.**, 1957. "Studies on the amino acid fermentation part I. Production of L-glutamic acid by various microorganisms." *The Journal of General and Applied Microbiology*, 3 (3), 193–205. https://doi.org/10.2323/jgam.3.193

**Kirchner, O. ; Tauch, A.**, 2003. "Tools for genetic engineering in the amino acid-producing bacterium *Corynebacterium glutamicum*." *Journal of Biotechnology*, 104 (1–3), 287–299. https://doi.org/10.1016/S0168-1656(03)00148-2

**Knott, G.J. ; Doudna, J.A.**, 2018. "CRISPR-Cas guides the future of genetic engineering." *Science*, 361 (6405), 866–869. https://doi.org/10.1126/science.aat5011

**Kohlstedt, M. ; Becker, J. ; Wittmann, C.**, 2010. "Metabolic fluxes and beyond—systems biology understanding and engineering of microbial metabolism." *Applied Microbiology and Biotechnology*, 88 (5), 1065–1075. https://doi.org/10.1007/s00253-010-2854-2

**Koonin, E. V. ; Makarova, K.S. ; Zhang, F.**, 2017. "Diversity, classification and evolution of CRISPR-Cas systems." *Current Opinion in Microbiology*, 37 , 67–78. https://doi.org/10.1016/j.mib.2017.05.008

**Koonin, E. V ; Makarova, K.S.**, 2019. "Origins and evolution of CRISPR-Cas systems." *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 374 (1772), 20180087. https://doi.org/10.1098/rstb.2018.0087

**Kopniczky, M.B. ; Moore, S.J. ; Freemont, P.S.**, 2015. "Multilevel regulation and translational switches in synthetic biology." *IEEE Transactions on Biomedical Circuits and Systems*, 9 (4), 485–496. https://doi.org/10.1109/TBCAS.2015.2451707

**Kortmann, M. ; Kuhl, V. ; Klaffl, S. ; Bott, M.**, 2015. "A chromosomally encoded T7 RNA polymerase-dependent gene expression system for *Corynebacterium glutamicum*: Construction and comparative evaluation at the single-cell level." *Microbial Biotechnology*, 8 (2), 253–265. https://doi.org/10.1111/1751-7915.12236

**Kudla, G. ; Murray, A.W. ; Tollervey, D. ; Plotkin, J.B.**, 2009. "Coding-Sequence Determinants of Gene Expression in Escherichia coli." *Science*, 324 (5924), 255–258. https://doi.org/10.1126/science.1170160

**Lausberg, F. ; Chattopadhyay, A.R. ; Heyer, A. ; Eggeling, L. ; Freudl, R.**, 2012. "A tetracycline inducible expression vector for *Corynebacterium glutamicum* allowing tightly regulable gene expression." *Plasmid*, 68 (2), 142–147. https://doi.org/10.1016/j.plasmid.2012.05.001

**Lee, J.H. ; Wendisch, V.F.**, 2017. "Production of amino acids – Genetic and metabolic engineering approaches." *Bioresource Technology*, 245 , 1575–1587. https://doi.org/10.1016/j.biortech.2017.05.065

**Lee, S.Y. ; Kim, H.U.**, 2015. "Systems strategies for developing industrial microbial strains." *Nature Biotechnology*, 33 (10), 1061–1072. https://doi.org/10.1038/nbt.3365

**Leuchtenberger, W. ; Huthmacher, K. ; Drauz, K.**, 2005. "Biotechnological production of amino acids and derivatives: Current status and prospects." *Applied Microbiology and Biotechnology*, 69 (1), 1–8. https://doi.org/10.1007/s00253-005-0155-y

**Liang, J.C. ; Bloom, R.J. ; Smolke, C.D.**, 2011. "Engineering biological systems with synthetic RNA molecules." *Molecular Cell*, 43 (6), 915–926. https://doi.org/10.1016/j.molcel.2011.08.023

**Lin, C. ; Jair, Y.C. ; Chou, Y.C. ; Chen, P.S. ; Yeh, Y.C.**, 2018. "Transcription factor-based biosensor for detection of phenylalanine and tyrosine in urine for diagnosis of phenylketonuria." *Analytica Chimica Acta*, 1041 , 108–113. https://doi.org/10.1016/j.aca.2018.08.053

**Liu, D. ; Evans, T. ; Zhang, F.**, 2015. "Applications and advances of metabolite biosensors for metabolic engineering." *Metabolic Engineering*, 31 , 35–43. https://doi.org/10.1016/j.ymben.2015.06.008

**Lokko, Y. ; Heijde, M. ; Schebesta, K. ; Scholtès, P. ; Montagu, M. Van**, 2018. "Biotechnology and the bioeconomy — Towards inclusive and sustainable industrial development." *New BIOTECHNOLOGY*, 40 (Part A), 5–10. https://doi.org/10.1016/j.nbt.2017.06.005

**Maddocks, S.E. ; Oyston, P.C.F.**, 2008. "Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins." *Microbiology*, 154 (12), 3609–3623. https://doi.org/10.1099/mic.0.2008/022772-0

**Mahr, R. ; Frunzke, J.**, 2016. "Transcription factor-based biosensors in biotechnology: current state and future prospects." *Applied Microbiology and Biotechnology*, 100 (1), 79–90. https://doi.org/10.1007/s00253-015-7090-3

**Mahr, R. ; Gätgens, C. ; Gätgens, J. ; Polen, T. ; Kalinowski, J. ; Frunzke, J.**, 2015. "Biosensor-driven adaptive laboratory evolution of L-valine production in *Corynebacterium glutamicum*." *Metabolic Engineering*, 32 , 184–194. https://doi.org/10.1016/j.ymben.2015.09.017

**Mannan, A.A. ; Liu, D. ; Zhang, F. ; Oyarzún, D.A.**, 2017. "Fundamental design principles for transcription-factor-based metabolite biosensors." *ACS Synthetic Biology*, 6 (10), 1851–1859. https://doi.org/10.1021/acssynbio.7b00172

**Meyer, A. ; Pellaux, R. ; Potot, S. ; Becker, K. ; Hohmann, H.P. ; Panke, S. ; Held, M.**, 2015. "Optimization of a whole-cell biocatalyst by employing genetically encoded product sensors inside nanolitre reactors." *Nature Chemistry*, 7 (8), 673–678. https://doi.org/10.1038/nchem.2301

**Michener, J.K. ; Smolke, C.D.**, 2012. "High-throughput enzyme evolution in *Saccharomyces cerevisiae* using a synthetic RNA switch." *Metabolic Engineering*, 14 (4), 306–316. https://doi.org/10.1016/j.ymben.2012.04.004

**Mojica, F.J.M. ; Díez-Villaseñor, C. ; García-Martínez, J. ; Almendros, C.**, 2009. "Short motif sequences determine the targets of the prokaryotic CRISPR defence system." *Microbiology*, 155 (3), 733–740. https://doi.org/10.1099/mic.0.023960-0

**Mustafi, N. ; Grünberger, A. ; Kohlheyer, D. ; Bott, M. ; Frunzke, J.**, 2012. "The development and application of a single-cell biosensor for the detection of L-methionine and branched-chain amino acids." *Metabolic Engineering*, 14 (4), 449–457. https://doi.org/10.1016/j.ymben.2012.02.002

**Nakamura, J. ; Hirano, S. ; Ito, H. ; Wachi, M.**, 2007. "Mutations of the *Corynebacterium glutamicum* NCgl1221 gene, encoding a mechanosensitive channel homolog, induce L-glutamic acid production." *Applied and environmental microbiology*, 73 (14), 4491–8. https://doi.org/10.1128/AEM.02446-06

**Niebisch, A. ; Bott, M.**, 2001. "Molecular analysis of the cytochrome bc 1 - aa 3 branch of the *Corynebacterium glutamicum* respiratory chain containing an unusual diheme cytochrome c 1." *Archives of Microbiology*, 175 (4), 282–294. https://doi.org/10.1007/s002030100262

**Nuñez, J.K. ; Harrington, L.B. ; Kranzusch, P.J. ; Engelman, A.N. ; Doudna, J.A.**, 2015a. "Foreign DNA capture during CRISPR-Cas adaptive immunity." *Nature*, 527 (7579), 535–538. https://doi.org/10.1038/nature15760

**Nuñez, J.K. ; Kranzusch, P.J. ; Noeske, J. ; Wright, A. V. ; Davies, C.W. ; Doudna, J.A.**, 2014. "Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity." *Nature Structural and Molecular Biology*, 21 (6), 528–534. https://doi.org/10.1038/nsmb.2820

**Nuñez, J.K. ; Lee, A.S.Y. ; Engelman, A. ; Doudna, J.A.**, 2015b. "Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity." *Nature*, 519 (7542), 193–198. https://doi.org/10.1038/nature14237

**Okibe, N. ; Suzuki, N. ; Inui, M. ; Yukawa, H.**, 2010. "Isolation, evaluation and use of two strong, carbon source-inducible promoters from *Corynebacterium glutamicum*." *Letters in Applied Microbiology*, 50 (2), 173–180. https://doi.org/10.1111/j.1472-765X.2009.02776.x

**Parekh, S. ; Vinci, V.A. ; Strobel, R.J.**, 2000. "Improvement of microbial strains and fermentation processes." *Applied Microbiology and Biotechnology*, 54 (3), 287–301. https://doi.org/10.1007/s002530000403

**Park, J.-U. ; Jo, J.-H. ; Kim, Y.-J. ; Chung, S.-S. ; Lee, J.-H. ; Lee, H.H.**, 2008. "Construction of heat-inducible expression vector of *Corynebacterium glutamicum* and *C. ammoniagenes*: fusion of lambda operator with promoters isolated from C. ammoniagenes." *Journal of microbiology and biotechnology*, 18 (4), 639–47.

**Pátek, M. ; Nešvera, J.**, 2011. "Sigma factors and promoters in *Corynebacterium glutamicum*." *Journal of Biotechnology*, 154 (2–3), 101–113. https://doi.org/10.1016/j.jbiotec.2011.01.017

**Patnaik, R.**, 2008. "Engineering complex phenotypes in industrial strains." *Biotechnol. Prog.*, 24 , 38–47. https://doi.org/10.1002/9781118433034

**Pul, Ü. ; Wurm, R. ; Arslan, Z. ; Geißen, R. ; Hofmann, N. ; Wagner, R.**, 2010. "Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS." *Molecular Microbiology*, 75 (6), 1495–1512. https://doi.org/10.1111/j.1365-2958.2010.07073.x

**Rogers, J.K. ; Guzman, C.D. ; Taylor, N.D. ; Raman, S. ; Anderson, K. ; Church, G.M.**, 2015. "Synthetic biosensors for precise gene control and real-time monitoring of metabolites." *Nucleic Acids Research*, 43 (15), 7648–7660. https://doi.org/10.1093/nar/gkv616

**Rohles, C.M. ; Gläser, L. ; Kohlstedt, M. ; Gießelmann, G. ; Pearson, S. ; del Campo, A. ; Becker, J. ; Wittmann, C.**, 2018. "A bio-based route to the carbon-5 chemical glutaric acid and to bionylon-6,5 using metabolically engineered *Corynebacterium glutamicum*." *Green Chemistry*, 20 (20), 4662–4674. https://doi.org/10.1039/C8GC01901K

**Rollie, C. ; Schneider, S. ; Brinkmann, A.S. ; Bolt, E.L. ; White, M.F.**, 2015. "Intrinsic sequence specificity of the Cas1 integrase directs new spacer acquisition." *eLife*, 4 (AUGUST2015), 1–19. https://doi.org/10.7554/eLife.08716

**Rostøl, J.T. ; Marraffini, L.**, 2019. "(Ph)ighting phages: How bacteria resist their parasites." *Cell Host and Microbe*, 25 (2), 184–194. https://doi.org/10.1016/j.chom.2019.01.009

**Sahm, H. ; Eggeling, L. ; de Graaf, A.A.**, 2000. "Pathway analysis and metabolic engineering in *Corynebacterium glutamicum*." *Biological Chemistry*, 381 (9–10), 899–910. https://doi.org/10.1515/BC.2000.111

**Scarlat, N. ; Dallemand, J.F. ; Monforti-Ferrario, F. ; Nita, V.**, 2015. "The role of biomass and bioenergy in a future bioeconomy: Policies and facts." *Environmental Development*, 15 (2015), 3–34. https://doi.org/10.1016/j.envdev.2015.03.006

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

**Schallmey, M. ; Frunzke, J. ; Eggeling, L. ; Marienhagen, J.**, 2014. "Looking for the pick of the bunch: High-throughput screening of producing microorganisms with biosensors." *Current Opinion in Biotechnology*, 26 , 148–154. https://doi.org/10.1016/j.copbio.2014.01.005

**Schwentner, A. ; Feith, A. ; Münch, E. ; Stiefelmaier, J. ; Lauer, I. ; Favilli, L. ; Massner, C. ; Öhrlein, J. ; Grund, B. ; Hüser, A. ; Takors, R. ; Blombach, B.**, 2019. "Modular systems metabolic engineering enables balancing of relevant pathways for L-histidine production with *Corynebacterium glutamicum*." *Biotechnology for Biofuels*, 12 (1), 65. https://doi.org/10.1186/s13068-019-1410-2

**Serganov, A. ; Nudler, E.**, 2013. "A decade of riboswitches." *Cell*, 152 (1–2), 17–24. https://doi.org/10.1016/j.cell.2012.12.024

**Shah, S.A. ; Erdmann, S. ; Mojica, F.J.M. ; Garrett, R.A.**, 2013. "Protospacer recognition motifs: Mixed identities and functional diversity." *RNA Biology*, 10 (5), 891–899. https://doi.org/10.4161/rna.23764

**Shiimori, M. ; Garrett, S.C. ; Graveley, B.R. ; Terns, M.P.**, 2018. "Cas4 nucleases define the PAM, length, and orientation of DNA fragments integrated at CRISPR loci." *Molecular Cell*, 70 (5), 814-824.e6. https://doi.org/10.1016/j.molcel.2018.05.002

**Skjoedt, M.L.M. ; Snoek, T. ; Kildegaard, K.K.R. ; Arsovska, D. ; Eichenberger, M. ; Goedecke, T.T.J. ; Rajkumar, A.S. ; Zhang, J. ; Kristensen, M. ; Lehka, B.J. ; Siedler, S. ; Borodina, I. ; Jensen, M.K. ; Keasling, J.J.D.**, 2016. "Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast." *Nature Chemical Biology*, 12 (11), 951–958. https://doi.org/10.1038/nchembio.2177

**Soetaert, W. ; Vandamme, E.**, 2006. "The impact of industrial biotechnology." *Biotechnology Journal*, 1 (7–8), 756–769. https://doi.org/10.1002/biot.200600066

**Soisson, S.M.**, 1997. "Structural Basis for Ligand-Regulated Oligomerization of AraC." *Science*, 276 (5311), 421–425. https://doi.org/10.1126/science.276.5311.421

**Stella, S. ; Mesa, P. ; Thomsen, J. ; Paul, B. ; Alcón, P. ; Jensen, S.B. ; Saligram, B. ; Moses, M.E. ; Hatzakis, N.S. ; Montoya, G.**, 2018. "Conformational activation promotes CRISPR-Cas12a catalysis and resetting of the endonuclease activity." *Cell*, 175 (7), 1856-1871.e21. https://doi.org/10.1016/j.cell.2018.10.045

**Sticher, P. ; Jaspers, M.C.M. ; Stemmler, K. ; Harms, H. ; Zehnder, A.J.B. ; Van Der Meer, J.R.**, 1997. "Development and characterization of a whole-cell bioluminescent sensor for bioavailable middle-chain alkanes in contaminated groundwater samples." *Applied and Environmental Microbiology*, 63 (10), 4053–4060. https://doi.org/10.1128/aem.63.10.4053-4060.1997

**Straathof, A.J.J. ; Wahl, S.A. ; Benjamin, K.R. ; Takors, R. ; Wierckx, N. ; Noorman, H.J.**, 2019. "Grand research challenges for sustainable industrial biotechnology." *Trends in Biotechnology*, 37 (10), 1042–1050. https://doi.org/10.1016/j.tibtech.2019.04.002

**Swarts, D.C.**, 2019. "Making the cut(s): how Cas12a cleaves target and non-target DNA." *Biochemical Society Transactions*, 47 (5), 1499–1510. https://doi.org/10.1042/BST20190564

**Swarts, D.C. ; van der Oost, J. ; Jinek, M.**, 2017. "Structural basis for guide RNA processing and seed-dependent DNA targeting by CRISPR-Cas12a." *Molecular Cell*, 66 (2), 221-233.e4. https://doi.org/10.1016/j.molcel.2017.03.016

**Tang, S.Y. ; Cirino, P.C.**, 2011. "Design and application of a mevalonate-responsive regulatory protein." *Angewandte Chemie - International Edition*, 50 (5), 1084–1086. https://doi.org/10.1002/anie.201006083

**Tang, S.Y. ; Fazelinia, H. ; Cirino, P.C.**, 2008. "AraC regulatory protein mutants with altered effector specificity." *Journal of the American Chemical Society*, 130 (15), 5267–5271. https://doi.org/10.1021/ja7109053

**Tang, S.Y. ; Qian, S. ; Akinterinwa, O. ; Frei, C.S. ; Gredell, J.A. ; Cirino, P.C.**, 2013. "Screening for enhanced triacetic acid lactone production by recombinant Escherichia coli expressing a designed triacetic acid lactone reporter." *Journal of the American Chemical Society*, 135 (27), 10099–10103. https://doi.org/10.1021/ja402654z

**Tosaka, O. ; Enei, H. ; Hirose, Y.**, 1983. "The production of L-lysine by fermentation." *Trends in Biotechnology*, 1 (3), 70–74. https://doi.org/10.1016/0167-7799(83)90055-0

**Trausch, J.J. ; Ceres, P. ; Reyes, F.E. ; Batey, R.T.**, 2011. "The structure of a tetrahydrofolate-sensing riboswitch reveals two ligand binding sites in a single aptamer." *Structure*, 19 (10), 1413–1423. https://doi.org/10.1016/j.str.2011.06.019

**Tsuchiya, M. ; Morinaga, Y.**, 1988. "Genetic control systems of *Escherichia coli* can confer inducible expression of cloned genes in Coryneform bacteria." *Nature Biotechnology*, 6 (4), 428–430. https://doi.org/10.1038/nbt0488-428

**Tuerk, C. ; Gold, L.**, 1990. "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase." *Science*, 249 (4968), 505–510. https://doi.org/10.1126/science.2200121

**Udaka, S.**, 1960. "Screening method for microorganisms accumulating metabolites and its use in the isolation of *Micrococcus glutamicus*." *Journal of Bacteriology*, 79 (5), 754–755. https://doi.org/10.1128/JB.79.5.754-755.1960

**Ungerer, J. ; Pakrasi, H.B.**, 2016. "Cpf1 is a versatile tool for CRISPR genome editing across diverse species of Cyanobacteria." *Scientific reports*, 6 , 39681. https://doi.org/10.1038/srep39681

**United Nations Framework Convention on Climate Change**, 2016. Paris Agreement. Paris.

**van Houte, S. ; Buckling, A. ; Westra, E.R.**, 2016. "Evolutionary ecology of prokaryotic immune mechanisms." *Microbiology and Molecular Biology Reviews*, 80 (3), 745–763. https://doi.org/10.1128/mmbr.00011-16

**van Summeren-Wesenhagen, P.V. ; Marienhagen, J.**, 2015. "Metabolic engineering of *Escherichia coli* for the synthesis of the plant polyphenol pinosylvin." *Applied and Environmental Microbiology*, 81 (3), 840–849. https://doi.org/10.1128/AEM.02966-14

**Vašicová, P. ; Pátek, M. ; Nešvera, J. ; Sahm, H. ; Eikmanns, B.**, 1999. "Analysis of the *Corynebacterium glutamicum dapA* promoter." *Journal of Bacteriology*, 181 (19), 6188–6191. https://doi.org/10.1128/JB.181.19.6188-6191.1999

**Wachsmuth, M. ; Findeiß, S. ; Weissheimer, N. ; Stadler, P.F. ; Mörl, M.**, 2013. "*De novo* design of a synthetic riboswitch that regulates transcription termination." *Nucleic Acids Research*, 41 (4), 2541–2551. https://doi.org/10.1093/nar/gks1330

**Wang, Y. ; Li, Q. ; Zheng, P. ; Guo, Y. ; Wang, L. ; Zhang, T. ; Sun, J. ; Ma, Y.**, 2016. "Evolving the L-lysine high-producing strain of *Escherichia coli* using a newly developed high-throughput screening method." *Journal of Industrial Microbiology & Biotechnology*, 43 (9), 1227–1235. https://doi.org/10.1007/s10295-016-1803-1

**Warner, J.R. ; Patnaik, R. ; Gill, R.T.**, 2009. "Genomics enabled approaches in strain engineering." *Current Opinion in Microbiology*, 12 (3), 223–230. https://doi.org/10.1016/j.mib.2009.04.005

**Webster, D.P. ; TerAvest, M.A. ; Doud, D.F.R. ; Chakravorty, A. ; Holmes, E.C. ; Radens, C.M. ; Sureka, S. ; Gralnick, J.A. ; Angenent, L.T.**, 2014. "An arsenic-specific biosensor with genetically engineered *Shewanella oneidensis* in a bioelectrochemical system." *Biosensors and Bioelectronics*, 62 , 320–324. https://doi.org/10.1016/j.bios.2014.07.003

**Wendisch, V.F. ; Jorge, J.M.P. ; Pérez-García, F. ; Sgobba, E.**, 2016. "Updates on industrial production of amino acids using *Corynebacterium glutamicum*." *World Journal of Microbiology and Biotechnology*, 32 (6). https://doi.org/10.1007/s11274-016-2060-1

**Wright, A. V. ; Doudna, J.A.**, 2016. "Protecting genome integrity during CRISPR immune adaptation." *Nature Structural and Molecular Biology*, 23 (10), 876–883. https://doi.org/10.1038/nsmb.3289

**Wright, A. V. ; Nuñez, J.K. ; Doudna, J.A.**, 2016. "Biology and applications of CRISPR systems: Harnessing nature's toolbox for genome engineering." *Cell*, 164 (1–2), 29–44. https://doi.org/10.1016/j.cell.2015.12.035

**Xu, P. ; Li, L. ; Zhang, F. ; Stephanopoulos, G. ; Koffas, M.**, 2014. "Improving fatty acids production by engineering dynamic pathway regulation and metabolic control." *Proceedings of the National Academy of Sciences of the United States of America*, 111 (31), 11299–304. https://doi.org/10.1073/pnas.1406401111

**Yamano, T. ; Nishimasu, H. ; Zetsche, B. ; Hirano, H. ; Slaymaker, I.M. ; Li, Y. ; Fedorova, I. ; Nakane, T. ; Makarova, K.S. ; Koonin, E. V. ; Ishitani, R. ; Zhang, F. ; Nureki, O.**, 2016. "Crystal structure of Cpf1 in complex with guide RNA and target DNA." *Cell*, 165 (4), 949–962. https://doi.org/10.1016/j.cell.2016.04.003

**Yamashita, C. ; Hashimoto, K.-I. ; Kumagai, K. ; Maeda, T. ; Takada, A. ; Yabe, I. ; Kawasaki, H. ; Wachi, M.**, 2013. "L-Glutamate secretion by the N-terminal domain of the *Corynebacterium glutamicum* NCgl1221 mechanosensitive channel." *Bioscience, biotechnology, and biochemistry*, 77 (5), 1008–13. https://doi.org/10.1271/bbb.120988

**Yan, M.-Y. ; Yan, H. ; Ren, G. ; Zhao, J.-P. ; Guo, X. ; Sun, Y.-C.**, 2017. "CRISPR-Cas12a-assisted recombineering in bacteria." *Applied and environmental microbiology*, 83 (17), 1–13. https://doi.org/10.1128/AEM.00947-17

**Yang, D. ; Cho, J.S. ; Choi, K.R. ; Kim, H.U. ; Lee, S.Y.**, 2017. "Systems metabolic engineering as an enabling technology in accomplishing sustainable development goals." *Microbial Biotechnology*, 10 (5), 1254–1258. https://doi.org/10.1111/1751-7915.12766

**Yim, S.S. ; An, S.J. ; Kang, M. ; Lee, J. ; Jeong, K.J.**, 2013. "Isolation of fully synthetic promoters for high-level gene expression in *Corynebacterium glutamicum*." *Biotechnology and Bioengineering*, 110 (11), 2959–2969. https://doi.org/10.1002/bit.24954

**Yosef, I. ; Goren, M.G. ; Qimron, U.**, 2012. "Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*." *Nucleic Acids Research*, 40 (12), 5569–5576. https://doi.org/10.1093/nar/gks216

**Zetsche, B. ; Gootenberg, J.S.S. ; Abudayyeh, O.O.O. ; Slaymaker, I.M.M. ; Makarova, K.S.S. ; Essletzbichler, P. ; Volz, S.E.E. ; Joung, J. ; van der Oost, J. ; Regev, A. ; Koonin, E.V. V. ; Zhang, F. ; Abudayyeh, O.O.O. ; Regev, A. ; Makarova, K.S.S. ; Slaymaker, I.M.M. ; Essletzbichler, P. ; Gootenberg, J.S.S. ; Volz, S.E.E. ; van der Oost, J. ; Zhang, F. ; Koonin, E.V. V. ; Zetsche, B.**, 2015. "Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system." *Cell*, 163 (3), 759–71. https://doi.org/10.1016/j.cell.2015.09.038

**Zetsche, B. ; Heidenreich, M. ; Mohanraju, P. ; Fedorova, I. ; Kneppers, J. ; Scott, D.A. ; van der Oost, J. ; Heidenreich, M. ; Severinov, K. ; DeGennaro, E.M. ; Gootenberg, J.S. ; Mohanraju, P. ; Zhang, F. ; Winblad, N. ; Zetsche, B. ; Abudayyeh, O.O. ; Choudhury, S.R. ; Wu, W.Y. ; Fedorova, I.**, 2016. "Multiplex gene editing by CRISPR–Cpf1 using a single crRNA array." *Nature Biotechnology*, 35 (1), 31–34. https://doi.org/10.1038/nbt.3737

**Zhang, J. ; Barajas, J.F. ; Burdu, M. ; Ruegg, T.L. ; Dias, B. ; Keasling, J.D.**, 2017. "Development of a transcription factor-based lactam biosensor." *ACS Synthetic Biology*, 6 (3), 439–445. https://doi.org/10.1021/acssynbio.6b00136

**Zhang, J. ; Jensen, M.K. ; Keasling, J.D.**, 2015. "Development of biosensors and their application in metabolic engineering." *Current Opinion in Chemical Biology*, 28 , 1–8.

https://doi.org/10.1016/j.cbpa.2015.05.013

**Zhang, J. ; Sonnenschein, N. ; Pihl, T.P.B. ; Pedersen, K.R. ; Jensen, M.K. ; Keasling, J.D.**, 2016. "Engineering an NADPH/NADP+ redox biosensor in yeast." *ACS synthetic biology*, 5 (12), 1546–1556. https://doi.org/10.1021/acssynbio.6b00135

## Appendix

## Author Contributions

LE conceived the design of the study. CKS and KK constructed plasmids and strains and performed the cultivation experiments to assess L-glutamate formation and performed HPLC measurements to monitor product formation. JM and LE wrote the manuscript. CKS and JM revised the manuscript.

Overall contribution CKS: 40%

CKS conceived the design of the study. LKF supported the design of the study. CKS constructed plasmids according to the unified sensor design for LysG- and PhdR-based biosensors for the application in *C. glutamicum* and *E. coli*. MV constructed sensor plasmid pSenPhdR. CKS constructed an arabinose-inducible expression plasmid for TAL expression in *E. coli*. CM determined inducer concentration range for LysG-based biosensors in *C. glutamicum* using cultivation experiments in a 48-well cultivation platform. CKS determined inducer concentration range for PhdR-based biosensors in *C. glutamicum* as well as LysG- and PhdR-based biosensors in *E. coli* using cultivation experiments in a 48-well cultivation platform. CSK assessed episomal 4CL expression (pCDF-BAD-*4cl$_{Pc}$*) in *E.coli*. CKS performed cultivation experiments in a 48-well cultivation platform for LysG- and PhdR-based biosensors in *C. glutamicum* and *E. coli* and subsequent flow cytometry measurements. CKS and JM wrote and revised the manuscript.

Overall contribution CKS: 90%

## Danksagung

Bei **Herrn Prof. Dr. Michael Bott** bedanke ich mich für die Möglichkeit meine Doktorarbeit am IBG-1 anzufertigen sowie für die Übernahme des Erstgutachtens und für das Interesse am Fortgang dieser Arbeit.

**Herrn Prof. Dr. Matias Zurbriggen** danke ich für die Übernahme des Koreferats und die Bereitschaft, meine Doktorandenzeit als Mentor zu begleiten.

Ein besonderer Dank gilt meinem Betreuer **Herrn Prof. Dr. Jan Marienhagen** für die Unterstützung und Motivation während meiner Promotion. Danke für die vielen fordernden und fördernden Diskussionen, Besprechungen und den Meinungs-Austausch à la Jan.

Genauso bedanke ich mich bei der Arbeitsgruppe **„Synthetische Zellfabriken"** für die wunderbare Arbeitsatmosphäre, die große Hilfsbereitschaft und viele Diskussionen zum Thema meiner Dissertation, aber auch für unterhaltsame Kaffeepausen, schöne Gruppenausflüge und lustige Labor-Schwätzle.

Im Besonderen danke ich der **Sensoren-Untergruppe** (Lion, Hugo, Michael und Philipp) mit denen ich meine Sensor-Sorgen austauschen konnte. Genauso bedanke ich mich bei meiner CRISPR-Sister **Karin**, die ebenso wie ich, das ein oder andere Mal an CRISPR/ Cas12a gezweifelt hat.

Außerdem bedanke ich mich bei **Sascha** und meiner **Masterstudentin D-Celine**, die mich während meiner Arbeit immer fleißig und mit großem Interesse unterstützt haben.
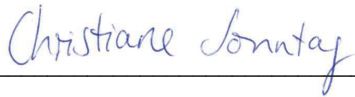
Dem **IBG-1** danke ich für die tolle Arbeitsatmosphäre und die Hilfsbereitschaft in den letzten Jahren.

Zuletzt geht ein besonderer Dank an meine **Familie**, die immer an mich geglaubt hat und die mich während des gesamten Studiums und der Doktorarbeit zu jeder Zeit unterstützt hat. Egal wie weit ich von zu Hause weg war!

## Eidesstattliche Erklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf" erstellt worden ist. Die Dissertation wurde in der vorgelegten oder ähnlichen Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Jülich, den 05.03.2021

**Christiane Katharina Sonntag**

Schriften des Forschungszentrums Jülich
Reihe Schlüsseltechnologien / Key Technologies

Schriften des Forschungszentrums Jülich
Reihe Schlüsseltechnologien / Key Technologies

Weitere _Schriften des Verlags im Forschungszentrum Jülich_ unter
http://wwwzb1.fz-juelich.de/verlagextern1/index.asp

JÜLICH
Forschungszentrum