Development of *Escherichia coli*-based biocatalysts using CRISPR/Cas9-assisted genome engineering

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Uwe Joost Lülf
aus Wuppertal

Düsseldorf, Mai 2025

aus dem Institut für Biochemie, Lehrstuhl II

der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der

Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

Berichterstattende:

1. Prof. Dr. Vlada B. Urlacher

2. Prof. Dr. Lutz Schmitt

Tag der mündlichen Prüfung: 20.11.2025

Eidesstattliche Erklärung

Ich, Uwe Joost Lülf, 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.

Diese Dissertation wurde weder in der vorliegenden noch in ähnlicher Form bei einer anderen Fakultät eingereicht. Bisher habe ich keine erfolglosen oder erfolgreichen Promotionsversuche unternommen.

Düsseldorf, 20.05.2025

„[…] der unermeßlich reichen, stets sich erneuernden Natur gegenüber wird der Mensch, so weit er auch in der wissenschaftlichen Erkenntnis fortgeschritten sein mag, immer das sich wundernde Kind bleiben und muß sich stets auf neue Überraschungen gefaßt machen.“

Max Planck

Abstract

Since the first steps in biocatalysis, the introduction of molecular biological tools has revolutionized the field many times. The introduction of recombinant DNA has allowed the transfer of foreign genes and thus new functions into fast-growing host organisms like the workhorse *Escherichia coli*. Furthermore, the properties of enzymes including selectivity, activity, and stability can be changed using protein engineering. In 2012, the discovery and subsequent application of CRISPR/Cas as a programmable genome editing tool enabled scar-less genome engineering including chromosomal integration, deletion and substitution of any DNA fragment ranging from single nucleotides to whole metabolic pathways.

In the course of this dissertation, various applications of CRISPR/Cas-assisted genome engineering of *E. coli* in the context of the entire spectrum of biocatalysis were examined. This includes applications in fermentation with growing cells, biotransformation with resting cells and catalysis with isolated enzymes. Part of this work focuses on the oxyfunctionalization of non-activated C-H bonds catalyzed by different oxygenases (cytochrome P450 monooxygenases and peroxygenases as well as unspecific peroxygenases), which is of particular interest for organic synthesis. Here, the knockout of two *E. coli* genes encoding catalases enabled the use of resting cells and cell-free extracts for H_2O_2-driven biocatalysis with peroxygenases, omitting the need for enzyme purification. Further, the effects of chromosomal integration on the catalytic performance of a multi-component P450 system to produce a human drug metabolite with resting *E. coli* cells were evaluated. In contrast to plasmid-based expression systems, the use of selectable markers as antibiotic resistance could be avoided after chromosomal integration and a more stable expression could be achieved. In order to facilitate the transfer from plasmid-based to plasmid-free chromosomally integrated biocatalysts, a versatile toolbox was designed and established which can be applied in various *E. coli* strains.

Apart from oxyfunctionalization reactions catalyzed by a single enzyme, the multi-step synthesis of valuable plant secondary metabolites is a particularly interesting field of

application for biocatalysis. Many active pharmaceutical ingredients are derived from scarce natural resources which can be substituted by biotechnological solutions. For this reason, an artificial biosynthetic pathway for the plasmid-free production of the high value key lignan pinoresinol from the relatively inexpensive phenylpropanoid ferulic acid was reconstituted in growing *E. coli* cells by combining exogenous and endogenous enzymes.

Finally, a proof-of-concept study for combinatorial promoter substitution after chromosomal integration of heterologous genes in *E. coli* is presented. This aims to address the challenges of potential bottlenecks in the construction of multi-step reaction cascades *in vivo*.

In summary, this thesis explores various possible biocatalytic applications of CRISPR/Cas-assisted genome engineering in *E. coli*. These include the prevention of side reactions by gene knockouts, the transfer of plasmid-based to plasmid-free expression systems as well as the reconstitution and optimization of an artificial biosynthetic pathway for biocatalysis both *in vivo* and *in vitro*.

Publications and conference contributions

Publications included in this work:

Luelf, U. J., Böhmer, L. M., Li, S., Urlacher, V. B. (2023). "Effect of chromosomal integration on catalytic performance of a multi-component P450 system in *Escherichia coli*." Biotechnology and Bioengineering **120**: 1762-1772.

Luelf, U. J., Wassing, A., Böhmer, L. M., Urlacher, V. B. (2024). "Plasmid-free production of the plant lignan pinoresinol in growing *Escherichia coli* cells." Microbial Cell Factories **23**, 289: 1-12.

Ebrecht, A.C., **Luelf, U. J.**, Govender, K., Opperman, D. J., Urlacher, V. B., Smit, M. S. (2025). "Use of whole cells and cell-free extracts of catalase-deficient *E. coli* for peroxygenase-catalyzed reactions." Biotechnology and Bioengineering **122**: 1376-1385.

Luelf, U. J., Urlacher, V. B. (2025). "Combinatorial promoter replacement after integration of heterologous genes into the *Escherichia coli* chromosome." *In preparation.*

Publications not included in this work:

Reiss, G. J., Urlacher V. B., **Luelf U. J.** (2020). "Enzyme-mediated synthesis and crystal structure of (2*R*,4*S*)-hydroxyketamine, $C_{13}H_{16}ClNO_2$." Zeitschrift für Kristallographie - New Crystal Structures **235**: 1037-1039.

Luelf, U. J., Reiss, G. J., Bokel A., Urlacher V. B. (2021). "Selective biocatalytic synthesis and crystal structure of (2*R*,6*R*)-hydroxyketaminium chloride, $C_{13}H_{17}Cl_2NO_2$." Zeitschrift für Kristallographie - New Crystal Structures **236**: 827-829.

Lappe, A.*, **Luelf, U. J.***, Keilhammer, M., Bokel A., Urlacher V.B. (2023). „Bacterial cytochrome P450 enzymes: Semi-rational design and screening of mutant libraries in recombinant *Escherichia coli* cells." H. Renata, Academic Press. **693**: 133-170. **both authors contributed equally*

Conference contributions:

Luelf, U. J. & Urlacher, V. B. (2022). Plasmid-free whole-cell biotransformation with recombinant P450. 10th International Congress on Biocatalysis. Hamburg.

Raffaele, A., **Luelf, U. J.**, Bokel, A., Urlacher, V. B. (2022). Screening of promiscuous P450s for forskolin oxidation. 10th International Congress on Biocatalysis. Hamburg.

Knöfel, R., **Luelf, U. J.**, Bechthold, P. A., Urlacher, V. B. (2023). Optimized synthesis of plant lignans via chromosomal integration of a multi-enzyme cascade in *E. coli*. 16th International Symposium on Biocatalysis and Biotransformations series (BIOTRANS). La Rochelle.

Lappe, A., Keilhammer, M., **Luelf, U. J.**, Urlacher, V. B. (2023). Engineering of a bacterial cytochrome P450 monooxygenase for the synthesis of (-)-podophyllotoxin. 16th International Symposium on Biocatalysis and Biotransformations series (BIOTRANS). La Rochelle.

Smit, M. S., Opperman, D. J., Urlacher, V. B., **Luelf, U. J.**, Ebrecht, A. C., Govender, K. (2024). Use of cell free extracts of catalase deficient *E. coli* for hydrogen peroxide driven CYP reactions. 16th International Symposium – Cytochrome P450 Biodiversity & Biotechnology. Torino, Italy.

Table of contents

1 Introduction

1.1 Biocatalysis and biotransformation

1.1.1 A brief history of biocatalysis

The use of enzymes or microbial cells to catalyze chemical reactions has emerged over a century ago. It has become a well-established practice in the chemical industry including food and animal feed production as well as in the pharmaceutical industry. To date, hundreds of industrial-scale biotransformation processes are known. Examples include the production of high fructose corn syrup, which is used as a sweetener in processed food, the production of amino acids for animal feed, or the (semi-)synthesis of antibiotics. Enzymes are also utilized in the paper and textile industry in non-chlorine bleaching processes. Furthermore, the use of enzymes as catalysts has made its way into our everyday lives: lipases, amylases, and proteases are part of many modern laundry detergents and catalyze the degradation of lipid, starch, and protein stains.[1, 2]

The first steps of biocatalysis date back to the 19th century when not only the concept of chemical catalysis was introduced by Berzelius, but also the first enzymes were extracted and applied. For example, in 1833, Payen and Persoz extracted diastase (amylase) from malt and observed the transformation of starch into glucose. Shortly after this, in 1837, baker's yeast (*Saccharomyces cerevisiae*), which has been used unknowingly in fermentation for millenia, was identified as a microorganism.[3] Nevertheless, the term '*enzyme*' was only introduced in 1878, and the chemical nature of enzymes as proteins was not discovered until 1926.[4, 5] From then on, microorganisms and enzymes were increasingly used in the production of bulk chemicals, pharmaceuticals and food ingredients. Apart from fermentation processes like the syntheses of penicillin G and citric acid using fungi (*Penicillium chrysogenum*[6] and *Aspergillus niger*[7]), or the acetone-butanol-ethanol (ABE) fermentation using bacteria of the genus Clostridium[8], the production of high fructose corn syrup using immobilized glucose isomerase might be one of the most prominent examples of the early days of industrial biocatalysis and still remains one of the largest biocatalytic processes today.[9, 10] All the aforementioned

processes relied on naturally occurring microorganisms or enzymes thereof with all their limitations in terms of stability, selectivity, and productivity.

With the introduction of recombinant DNA technology in the 1970s,[11, 12] a new era of biocatalysis dawned. Often, natural compounds of interest are only produced in low quantities or by slow-growing, sometimes endangered species.[13, 14] Moreover, the cultivation of newly identified microbes from environmental samples in the laboratory often fails. Recombinant DNA technology as well as the development of DNA sequencing methods enabled not only the discovery of new genes and enzyme functions as well as the elucidation of biosynthetic pathways, but also enabled their transfer into easy-to-culture microbial hosts.[15]

Furthermore, for the first time, the properties of enzymes could be modified using protein engineering techniques. The extension of their substrate specificity including non-natural substrates broadened their applicability in the synthesis of fine chemicals like active pharmaceutical ingredients.[2] In addition, protein engineering enabled changes in selectivity, including regio- and stereoselectivity. Although enzymes as chiral catalysts are known for their high stereoselectivity, this is not true in all cases, especially not for non-natural substrates. Moreover, protein engineering can help to switch the selectivity if a different stereoisomer is of interest.[16, 17] The improvement of enzyme stability including pH stability, thermostability and stability in organic solvents using protein engineering techniques further increased the applicability of biocatalysts. There are generally two main approaches to protein engineering: rational protein design and directed evolution. The first relies on a fundamental knowledge of the enzyme's structure and mode of action. Using site-directed mutagenesis, individual amino acid residues can be specifically altered to achieve desired enzyme properties. For directed evolution, no structural information is needed. Through random mutation or recombination and subsequent high-throughput screening, gradual improvements in the desired properties can be achieved by mimicking natural evolution. Between these two approaches are semi-rational techniques like saturation mutagenesis, in which a reduced number of positions and amino acids are selected, which reduce the screening effort compared to directed evolution.[18, 19]

With the significant advances made in bioinformatics in recent years, particularly in the field of machine learning and protein folding prediction, it is reasonable to expect that the introduction of *de novo* enzymes will usher in yet another new era of biocatalysis.[20] In 2024, Google DeepMind's AlphaFold developers Demis Hassabis and John Humper as well as David Baker for his pioneering work in the field of *de novo* protein design were awarded the Nobel Prize in Chemistry.[21-23] The open source artificial intelligence software AlphaFold is used for protein structure prediction and demonstrated an accuracy competitive with experimental protein structures.[21] It has already begun to revolutionize the fields of biology and medicine by facilitating the discovery of drugs and their targets, the prediction of protein functions or the design of proteins, to name just a few examples.[24]

Nowadays, biocatalysis is often associated with claims regarding its sustainability and in context of the concept of green chemistry.[25] Catalysis as such belongs to the main principles of green chemistry, as it improves the atom economy and reduces waste. However, biocatalysts show some inherent properties that are considered even more beneficial in this context. For one, they are biodegradable as they are part of all life on earth. Secondly, biocatalytic reactions are usually performed under mild conditions in an aqueous milieu. This does not only include low, biocompatible temperatures which results in lower energy consumption, but also no toxic organic solvents. On the other hand, the use of large amounts of water could pose a challenge in some places facing global warming and its consequences. Finally, due to their specificity as well as chemo-, regio- and stereoselectivity, enzymes can catalyze reactions in a single step that can only be achieved through laborious multi-step syntheses and by using protective groups in organic chemistry.[26]

1.1.2 Fermentation, whole-cell biotransformation or catalysis with isolated enzymes?

The form of the applied biocatalysts and the selected starting materials can be used to differentiate between biocatalytic processes ranging from *de novo* fermentation to biotransformation with whole cells and catalysis with isolated enzymes (Figure 1).

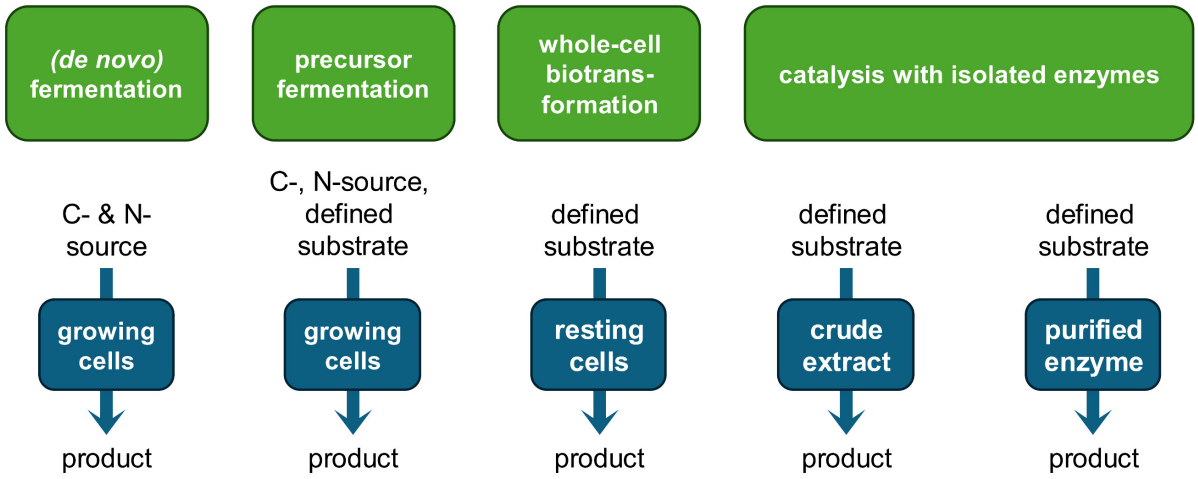


Figure 1: From *de novo* fermentation to whole-cell biotransformation and catalysis with isolated enzymes. Biocatalytic processes and their differences as defined by Bommarius and Riebel.[27] Products can include, for example, small organic molecules, proteins, nucleic acids, and carbohydrates.

In fermentation processes, living microbial cells use raw materials like sugar or starch to build up all metabolites essential for their growth.[27] The aim of fermentation can vary greatly and can involve the cultivation of the organism itself, the production of enzymes/proteins, or the synthesis of a chemical product.[15] For example, the yeast *S. cerevisiae* is produced in large quantities by fermentation and can be bought off the shelf for baking and brewing. In addition, all biocatalytic processes rely on enzymes and therefore depend on their production in living microbial cells.[28] Furthermore, the production of organic compounds of the primary and secondary metabolism of microbes, like amino acids and antibiotics, has increasingly come into focus over the last century.[26] With the help of metabolic engineering, new metabolic pathways can be established so that even complex synthetic compounds like taxadiene can

be produced from simple carbon sources.[29] Sometimes it is advisable to use an easily available more complex precursor as starting material. It is therefore necessary to distinguish between *de novo* fermentations, in which more complex compounds are built up starting from a carbon source, and precursor fermentations, where a defined substrate is added.[27] An example for the latter is the production of vinegar from ethanol by *Acetobacter aceti*.[30] In addition, precursor fermentation must be distinguished from biotransformation with non-growing, so-called resting cells.[27] However, it should be noted that the boundaries between precursor fermentation with growing cells and biotransformation with resting cells are often unclear or crossed, because both are whole-cell approaches.

A whole-cell approach like *de novo* fermentation or biotransformation generally includes two phases: (i) a growing phase in which the biocatalyst is multiplied, and (ii) a conversion phase which is started upon substrate addition.[31] For growing cells, the conversion is usually carried out in the same medium and in the same reactor as cell growth which is particularly attractive due to its simplicity. In contrast, reactions with resting cells are carried out in separate reaction solutions. Although this involves a higher amount of work during cell harvesting and preparation, it can improve the downstream processing as substances contained in the growth medium are removed. In addition, biotransformation with resting cells enables controlled reaction conditions including optimal and constant cell densities, temperatures that might not be compatible with cell growth, and a reduced number of byproducts due to a reduced cell metabolism.[26, 32] Furthermore, for resting cells, the mass transfer over the cell membrane can be improved by treatment with surfactants or freeze-thaw processes, which is particularly interesting for large or polar substrates.[31] It should be noted that the toxicity of substrates, accumulating intermediates or products, and extraction agents can cause an (undesired) stop in cell growth and the transition of a growing cell to a resting cell approach. Both have in common that the cell metabolism can be exploited for catalysis: cost-intensive cofactors such as nicotinamide adenine dinucleotide (phosphate) ($NAD(P)^+$/NAD(P)H) or adenosine triphosphate (ATP) and *S*-adenosyl methionine (SAM) are provided by the cells and do not have to be supplied manually.[31, 33]

Catalysis with isolated enzymes, on the other hand, describes *in vitro* setups for biotransformation with crude cell extracts or purified enzymes.[27] The simple reaction setup, consisting only of a substrate and the catalyst in buffer, can prevent side reactions that might be catalyzed by endogenous enzymes in the whole-cell approach, depending on the level of purity of the isolated enzymes.[34] Other advantages include controlled reaction conditions like pH value, no mass transfer limitations due to the absence of a cell membrane and higher productivity compared to whole-cell approaches.[15, 35-37] Even multi-step reaction cascades can be accomplished *in vitro*, in which the quantities of enzymes added can be perfectly balanced, taking their activity into account.[34] The possibility of adding different enzymes at different times can also be very useful.[36] However, reactions with enzymes that require cost-intensive cofactors are challenging. As these are consumed during a catalyzed reaction in stoichiometric amounts, a cofactor recycling system is usually included for economic reasons, which increases the complexity of this initially very simple approach.[38] Moreover, enzymes are inevitably lost during the purification process, may show a reduced stability or be inhibited or deactivated by reaction mixture components.[14] In industrial production, enzymes are often immobilized in order to increase stability as well as cost efficiency by reusing the enzymes.[2]

With all the different options available, it always depends on the individual case which of the various methods is suitable for the particular biocatalytic application.

1.1.3 Oxyfunctionalization - an important field for biocatalysis

Biocatalysis can show its unique strengths especially in the case of selective oxyfunctionalization of non-activated C-H bonds.[39-41] These reactions are of great interest to synthetic chemists, primarily because they enable late-stage oxyfunctionalization of complex structures and thus improve the step efficiency.[42] Nature's toolbox contains several different enzymes in the class of oxidoreductases that can catalyze C-H oxidation. The most prominent example is the family of heme *b* containing cytochrome P450s (CYPs) which are widely distributed throughout all domains of life with more than 350,000 classified members known

today.[43, 44] The diverse functions of CYPs include, *inter alia*, secondary metabolism and detoxification of xenobiotics. In order to execute these functions, CYPs catalyze over 20 different oxidation reactions including not only hydrocarbon hydroxylation but also alkene and arene epoxidation, *N*-, *S*- and *O*-dealkylation, *N*-hydroxylation or oxidative deamination (Figure 2).[43]

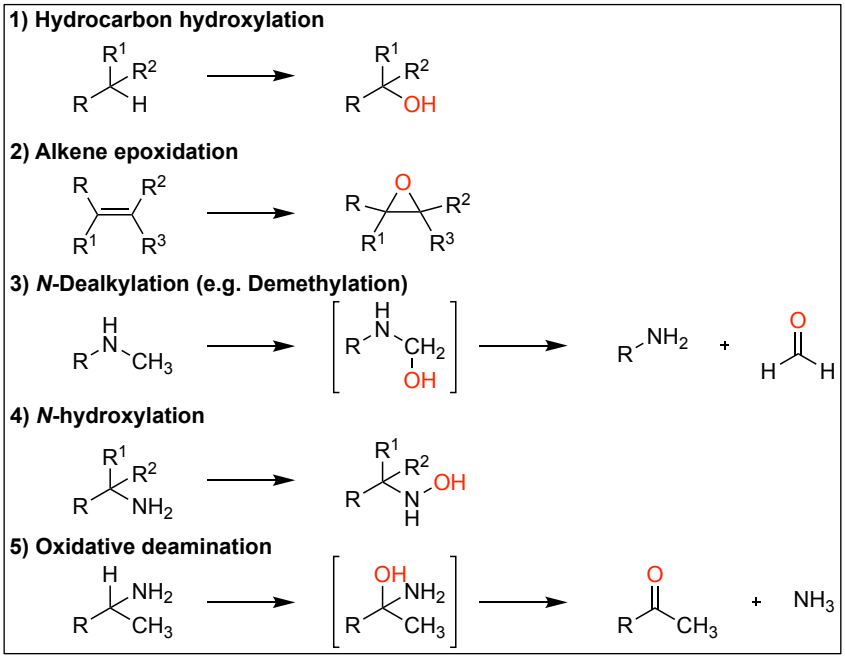


Figure 2: A few examples of reactions known to be catalyzed by CYPs.[45-47]

The vast majority of P450s or CYPs are external oxidases which require specific redox partner proteins for the electron transfer from the cofactor NAD(P)H to the heme iron. However, the respective redox partner proteins differ not only from P450 to P450, but also often remain unidentified.[48, 49] One solution to this problem is to use well-studied redox partners such as the flavodoxin YkuN from *Bacillus subtilis* and the flavodoxin reductase FdR from *Escherichia coli*, as in the three-component P450 system described in Chapter 2.2.[47, 50] Due to the dependence of P450 monooxygenases on the costly redox cofactors NADH or NADPH, a whole-cell approach, in which cell metabolism is used for cofactor supply and regeneration, is generally preferred.[43] Examples include the industrial hydroxylation of steroids using fungi of the genus *Curvularia*[51], the synthesis of pravastatin from compactin with *Streptomyces* sp.[52], or the biosynthesis of artemisinic acid, a precursor of the antimalarial drug artemisinin, in

recombinant *S. cerevisiae*.[53] Furthermore, P450s are used in drug development. Many drug metabolites originate from reactions catalyzed by hepatic P450s since they play a crucial role in the first phase of drug metabolism in humans.[54] As the use of human P450s to synthesize these metabolites faces expression and stability issues as well as low turnover numbers, bacterial or fungal P450s have been used to specifically produce human drug metabolites to further study their effects and evaluate their safety.[55, 56] An example for the synthesis of a human metabolite of the anesthetic (*S*)-ketamine is described as a model case for P450-catalyzed biotransformation with resting *E. coli* cells in Chapter 2.2.

The catalytic cycle of P450s (Figure 3) starts with the binding of a substrate molecule in close proximity to the heme iron which triggers the release of the previously bound sixth ligand water molecule and induces a switch of the spin state. This results in an increased reduction potential of the ferric iron and enables the acceptance of one electron from the redox partner protein. After the binding of molecular oxygen, the second electron transfer and a proton transfer result in the formation of *Compound 0*. Subsequently, the second protonation of the distal oxygen atom leads to a heterolytic cleavage of the dioxygen bond under release of water and the formation of an iron(IV) oxo complex (*Compound I*). The hydroxylation of the substrate follows a rebound mechanism with retention of configuration at the functionalized carbon. Depending on the chemical nature of the hydroxylated carbon atom, different subsequent reactions such as *N*-dealkylation or deamination might occur (Figure 2).[49, 57-60]

However, the P450 catalytic cycle is not 100% efficient. Various reactions can result in decoupling of NAD(P)H consumption and substrate oxidation. In this context, higher degrees of decoupling are observed for non-natural substrates. One of the possible decoupling pathways involves the release of hydrogen peroxide, in which *Compound 0* returns to the substrate-bound ground state via the so-called peroxide shunt. This release of hydrogen peroxide is extremely disadvantageous in enzyme catalysis since it can impair the lifespan of the enzyme. For this reason, catalases are often added to *in vitro* reactions which scavenge hydrogen peroxide and are among the enzymes with the highest known turnover numbers.[43, 61]

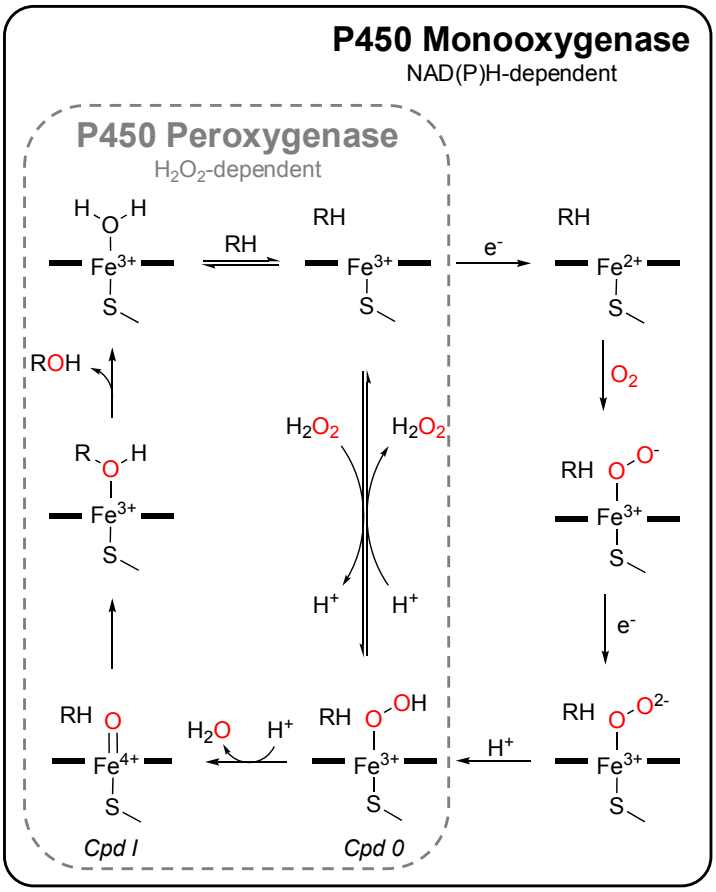


Figure 3: Comparison of catalytic cycles of P450 monooxygenases and peroxygenases. The peroxide shunt of the monooxygenase function can be reversed by the addition of hydrogen peroxide to form *Compound 0 (Cpd 0)* in peroxygenases.

Nevertheless, there is a small subgroup of P450 enzymes capable of using hydrogen peroxide as an oxidant in a “reversed peroxide shunt” rather than NAD(P)H-dependent fixation of molecular oxygen. One group of these P450 peroxygenases is involved in hydroxylation and decarboxylation of fatty acids in bacteria and of particular interest in applied biocatalysis.[62-65] The simple addition of hydrogen peroxide instead of the need for redox partner proteins and the expensive cofactor NAD(P)H is highly preferred from an economical perspective. This is also the reason why repeated attempts have been made to improve the peroxygenase properties of P450 monooxygenases through protein engineering.[66-69] Still, the results are often poor due to the limited enzyme stability and heme inactivation in presence of hydrogen peroxide, which rules out any industrial application.[70, 71]

In this gap, a relatively new group of enzymes have emerged: so-called unspecific peroxygenases (UPOs) which also use hydrogen peroxide for substrate oxidation.[72] Although these are heme-containing enzymes, too, there are some differences to P450s. UPOs are exclusively found in fungi and are suspected to be involved in lignin degradation and other environmental oxidation processes.[73] As extracellular enzymes, UPOs are considered relatively stable, making them particularly interesting for industrial oxidation processes using hydrogen peroxide as a co-substrate.[72] Nevertheless, the hydrogen peroxide concentration cannot be increased *ad libitum* either. In order to not endanger the enzyme's stability or cause heme bleaching, various methods for *in situ* generation of hydrogen peroxide have been established. Even simpler is the slow but steady addition of hydrogen peroxide to mitigate its negative effects.[74] As their name implies, UPOs show a high substrate promiscuity but in turn, wild-type UPOs often lack selectivity and tend to substrate overoxidation. A fact that currently limits their application and has led to attempts to overcome these issues by protein engineering.[73, 75]

Whether it is P450 peroxygenases or UPOs, H_2O_2-driven biocatalytic reactions are almost exclusively limited to catalysis with isolated enzymes *in vitro*.[74] This is due to several factors. First, UPOs are usually secreted both in nature and when using recombinant yeasts for heterologous expression. Thus, the active enzyme is obtained from the fermentation supernatant rather than from within the cell which improves the downstream processing.[76, 77] For this, either the UPO's native or a heterologous signal peptide like the α-factor secretion signal from *S. cerevisiae* can be used.[76, 78] In addition, the use of engineered signal peptides for improved titers is an option.[79] Although the yeasts *S. cerevisiae* and *Pichia pastoris* (i.e. *Komagataella phaffii*) are the preferred expression hosts for UPOs due to their fungal origin, several examples of successful heterologous expression in *E. coli* were described in the literature as well. Compared to yeasts, *E. coli* has the advantage of faster growth and easier accessibility for engineering.[80] However, almost all these examples investigate the catalytic activity using purified enzymes.[81-83] In cells, the degradation of supplied or *in situ* generated H_2O_2 by catalases competes with the peroxygenase-catalyzed reaction. This also applies to

crude extracts of recombinant *E. coli*, so that purification of the peroxygenase is required to remove the catalases. However, as already mentioned, this is associated with a higher workload, which makes the screening of new potential biocatalysts or their optimization by iterative protein engineering a tedious task. For this reason, Chapter 2.1 describes the development of a catalase-deficient *E. coli* strain for heterologous expression of different peroxygenases and subsequent H_2O_2-driven biocatalysis with crude extracts and whole cells.

1.1.4 Heterologous expression of biocatalysts in *Escherichia coli*

The production of a new biocatalyst usually starts with finding a promising enzyme candidate and its DNA sequence in a database. Technical advances in DNA sequencing have resulted in billions of gene sequence entries being accessible online which can be screened using bioinformatic methods.[84] Promising enzyme candidates can be found based on sequence homology to known enzymes or based on conserved motifs, if applicable.[2] Since the original organism is often impossible to cultivate in a laboratory,[15] or does not exhibit sufficient productivity, the gene sequence is then synthesized to introduce it into a microbial host for heterologous expression. A suitable host organism should combine a high growth rate, the availability of gene and genome engineering tools and established culture techniques with high cell density. Over the years, various prokaryotic and eukaryotic host organisms have been established for heterologous expression. These include, for example, the previously mentioned yeasts *S. cerevisiae*[85] and *P. pastoris*[86], or bacteria such as the gram-positive *Bacillus subtilis*[87] and the gram-negative *E. coli.*[88]

The latter is probably the best-studied model organism in microbiology and the host of choice for many biocatalytic processes.[89, 90] *E. coli* is commonly found in the colon of humans and animals and was first described in 1885 by Theodor Escherich as *Bacterium coli*.[91] In the 20th century, the study of *E. coli* and its bacteriophages laid the foundation for our knowledge of genetics. Furthermore, many techniques in modern molecular biology originate from work in *E. coli* or the utilization of its enzymes, and it was among the first organisms whose genome

was fully sequenced.[12, 92-95] Industrial applications include not only biocatalysis, but also the production of many recombinant therapeutics like insulin.[96, 97]

E. coli shows a great genetic diversity.[98] Most *E. coli* strains are non-pathogenic and laboratory strains like K-12 even lost their ability to colonize the intestine over the years.[99, 100] However, strains that are not adapted to laboratory cultivation possess the ability to form biofilms and can be opportunistic pathogens.[99, 101] Some serotypes can even cause severe infections in otherwise healthy people, which has a significant impact on global health.[99] In the laboratory, many non-pathogenic strains and derivatives were optimized for heterologous expression and cloning procedures. For example, the K-12 strain DH5α is optimized for the replication and isolation of DNA vectors. The well-known *E. coli* B strain BL21(DE3) lacks proteases that could degrade heterologous proteins and is equipped with tools derived from the bacteriophage T7 for strong heterologous expression.[102] Its derivatives *E. coli* C41(DE3) and C43(DE3) were optimized for the expression of membrane-associated and toxic proteins,[103, 104] and the Rosetta strains compensate for rare codons in *E. coli*, facilitating the expression of eukaryotic genes.[105] There are also various K-12 strains such as the ancestral *E. coli* MG1655 or W3110 that are widely used in industrial applications.[106, 107] The choice of the right strain and genotype for the specific biocatalytic application is crucial and depends on various factors such as the gene sequence to be expressed, its origin, and the properties of the encoded enzyme as well as the nature of the biocatalytic reaction and the reactants. Some *E. coli* endogenous enzymes may interfere with the planned biocatalytic application. However, using the methods of modern molecular biology, particularly in the field of genome engineering, strains can be modified to suit their intended application (cf. Chapter 1.2.3).

For heterologous expression in *E. coli*, small extrachromosomal circular DNA molecules, so-called plasmids, are often used (Figure 2).

Figure 2: Plasmid-based and chromosomal expression in *E. coli.*

Cloning procedures for the introduction of heterologous genes or manipulation of regulatory elements of plasmids are well established and they can be easily introduced into the host cell by transformation or conjugation.[91, 108] Especially for protein engineering with fast cycles of mutation, transformation, testing, and isolation of mutants, a plasmid-based approach is advantageous. Furthermore, in most cases, plasmids replicate independently of the chromosome and can therefore be present in the cell at different copy numbers. High copy plasmids like the cloning vector pUC can reach up to hundreds of copies per cell,[109] whereas medium copy plasmids like the common pET vector series (***p***lasmids for ***e***xpression by ***T***7 RNA polymerase) reach around 40 copies per cell.[110] In general, a higher copy number is correlated with higher expression and thus a higher enzyme concentration. However, this is not a linear correlation,[111, 112] and the expression might be limited by the cell's resources which can negatively affect the co-expression of other genes.[113, 114] In addition, it is known for example from P450s that the protein folding under high expression conditions can outpace heme incorporation, leading to a significant portion of apoprotein without a prosthetic group instead of the heme-loaded holoenzyme.[115] Apart from varying the gene copy number, the modification of regulatory elements such as promoters or ribosomal binding sites can be used to tune the expression.[116, 117] It should be noted that higher expression levels do not necessarily equal higher enzyme activity and therefore higher product titers in biocatalysis.[112, 118] Another issue of plasmid-based expression is that the plasmid copy number can vary from cell to cell or suddenly go into overflow when expression is induced. The replication of plasmids thus unnecessarily binds the limited resources of the cell. Since plasmids are extrachromosomal, non-essential DNA, some sort of selectable marker such as antibiotic resistance or auxotrophy

has to be used to maintain the plasmids in the host and prevent plasmid loss. On an industrial scale, the extensive use of antibiotics is a cost factor and contributes to the development of multidrug-resistant pathogens.[114, 119, 120]

Instead of plasmid-based expression, it is also possible to insert genes into the host's chromosome, eliminating the need for antibiotics. The rapid advancements, especially since the discovery and application of the CRISPR/Cas systems, have facilitated genome engineering in many organisms, including *E. coli*. The integration of the heterologous gene into the chromosome of *E. coli* not only fixes the copy number of the gene, but also prevents the gene from being lost, as is the case with plasmid instability and segregation. In addition, the metabolic burden associated with plasmid replication and the expression of antibiotic resistance genes is eliminated.[118, 120] Despite lower gene copy numbers, higher productivity of the plasmid-free whole-cell biocatalyst could be observed in many cases.[120-122] The challenges involved in the chromosomal integration of heterologous genes in *E. coli* are described in detail in Chapter 1.2 “Genome Engineering”.

1.1.5 Enzyme cascades and reconstitution of biosynthetic pathways

In addition to the use of enzymes and cells to catalyze single reaction steps, multi-enzymatic cascade reactions have been the focus of research in recent years. Inspired by nature's metabolic pathways, increasingly complex reaction sequences were designed for the synthesis of chemical products, both *in vitro* and *in vivo*.[123, 124] Ideally, a single vessel ("one-pot synthesis") is used for the entire process, which saves time and money by eliminating work-up steps and increases the yield.[125] A prominent example is the *in vitro* production of the investigational HIV drug islatravir reported by Merck & Co and Codexis. This enzyme cascade consists of five engineered enzymes and shows a significantly improved atom economy and reduced number of steps compared to the previous chemical syntheses.[126]

In nature, metabolic pathways and their enzymes have been established and optimized over millions of years.[15] The individual steps are perfectly coordinated, both in terms of activity and selectivity, to suit the demand of the respective organism. However, many natural products, such as plant secondary metabolites, are of pharmaceutical interest and the global demand cannot always be covered by isolation from natural sources. Therefore, the reconstitution and optimization of biosynthetic pathways in heterologous hosts is of great interest.[29, 127, 128] This includes the combination of enzymes from different source organisms in one host, the modulation of expression and protein engineering to improve activity or selectivity.[128] Even new-to-nature pathways like the THETA cycle for CO_2 fixation can be designed. However, this can be associated with issues such as substrate promiscuity leading to by-products or enzymes with low activity causing bottlenecks.[129] While this can be compensated *in vitro* by timing the enzyme addition or increasing its concentration, it is more difficult to deal with in whole cells.

An effective strategy for overcoming these challenges in growing *E. coli* cells is described in Chapter 2.3, illustrated by the example of an artificial cascade for lignan biosynthesis. Lignans are secondary plant metabolites which show various biological activities and are of great pharmaceutical interest.[130, 131] Particularly prominent among these is (-)-podophyllotoxin, which

is used to synthesize the chemotherapeutics etoposide and teniposide. Since the natural source of (-)-podophyllotoxin is scarce and highly endangered, biotechnological production is favored.[132, 133] The biosynthetic pathway of (-)-podophyllotoxin is almost completely elucidated and its direct precursor (-)-deoxypodophyllotoxin was successfully produced from pinoresinol via combination of different enzymes originating from different plants in recombinant *E. coli*.[134, 135] In order to make this reaction more feasible from an economic perspective, in this thesis, the key lignan pinoresinol (259€ / 10 mg)[1] was produced from relatively inexpensive ferulic acid (640 € / kg)[2]. To this end, enzymes from four different organisms, including the host itself, were combined in an artificial cascade reaction.

[1] https://www.sigmaaldrich.com/DE/de/product/sigma/40674 (27.10.2024)

[2] https://www.sigmaaldrich.com/DE/de/product/aldrich/w518301 (27.10.2024)

1.2 Genome engineering

1.2.1 State of the art

Genome engineering, the process of making targeted alterations to an organism's DNA, has become an essential tool in modern molecular biology.[136] By editing specific loci in an organism's genome, deleterious mutations can be eliminated, useful properties introduced, and even completely new biological functions created. This enables, for example, targeted therapies for genetic diseases,[137] the development of crops with improved nutritional profiles,[138] and the synthesis of valuable chemicals by metabolically engineered microorganisms.[29] However, the rapid development of the field of genome engineering has also raised ethical concerns and debates for example about the implications of modifying the human germline.[139]

The beginnings of synthetic biology including genome engineering were mainly driven by the desire to understand biological systems and their functions. As for modern biocatalysis, the foundations of genome engineering were laid in the 1970s and 1980s with the introduction of recombinant DNA technology,[11, 12] and the development of DNA sequencing[140] as well as polymerase chain reaction (PCR).[94] Since then, the field has benefited from exponential growth in oligonucleotide synthesis and sequencing capacities while costs have simultaneously declined dramatically.[141] By studying biological systems, particularly bacteriophages and DNA repair mechanisms, various techniques for targeted genome editing have been developed.[136]

Among the first methods was site-specific recombination using the recombinases Cre and Flp. This enabled integration, deletion, or inversion of genome segments by flanking them with the recognition sequences loxP and FRT, respectively.[142, 143] While site-specific recombination was a breakthrough in genome engineering, it comes with limitations: requiring flanking recognition sites adds complexity, and the resulting scars make broad genome editing challenging.[144]

Alongside recombinases, the exploitation of a natural DNA repair process, homologous recombination, has been developed and applied in various organisms.[145-147] Despite its potential, spontaneous homologous recombination occurs at a low frequency, with a baseline editing efficiency of only 10^{-7}-10^{-6}. Nonetheless, mutants can be obtained by employing positive or negative selectable markers.[148] Furthermore, the introduction of the phage-derived recombination systems ET cloning[149] and λ-Red recombineering[150] greatly enhanced the editing frequency in *E. coli*. It enabled the use of linear PCR fragments for homologous recombination and made the use of flanking recognition sites obsolete. Instead, short homology arms are introduced via PCR. However, antibiotic resistance cassettes should be introduced as a selectable marker and their flippase-assisted removal leaves a scar in the chromosome.[151]

An increase in editing frequency was also observed after introducing chromosomal double-strand breaks (DSBs) either chemically or by meganucleases.[152-154] Chromosomal DSBs can be fixed via two distinct repair mechanisms: non-homologous end joining (NHEJ) or homology-directed repair (HDR). The first results in random short insertion or deletion (indel) mutations, the latter requires the provision of a repair template (donor DNA). Depending on the design of this donor DNA, targeted deletions, insertions or replacements can be carried out (Figure 4).[136, 155]

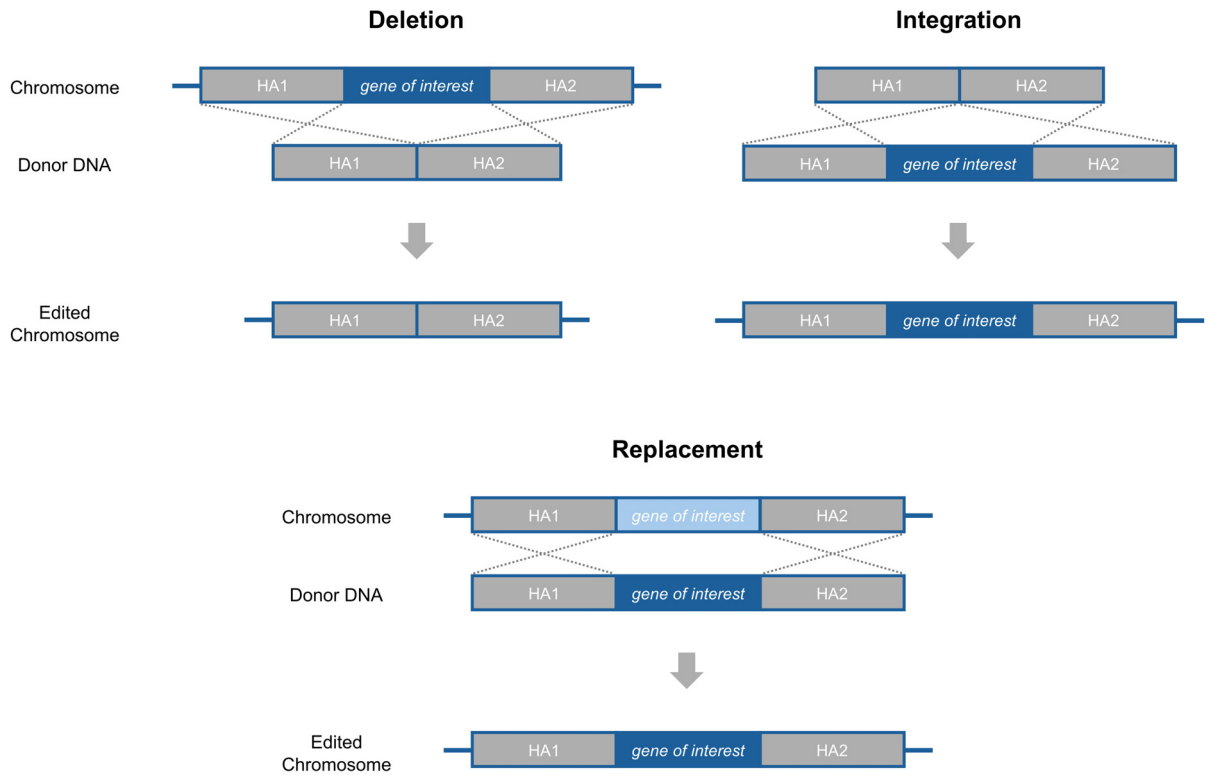


Figure 4: Genome editing via homology-directed repair. Depending on the supplied repair template (donor DNA), deletion, integration, or replacement of any DNA sequence (e.g. gene of interest) is enabled.

The rise of sequence-specific and programmable endonucleases like zinc finger nucleases (ZFNs)[156] and transcription-activator-like effector nucleases (TALENs)[157] facilitated the introduction of DSBs and therefore revolutionized targeted genome engineering. In addition, parallel progress in DNA sequencing made the first whole genome sequences available,[95, 158] which enabled the design of genome-wide applications of these methods. In the process, more and more new genes were discovered and their biological function could be deciphered through targeted knockouts.[159, 160] However, the process of designing and validating ZFNs is difficult and labor-intensive.[136] Frequently, cell toxicity is described, which is probably due to significant off-target effects.[148] In comparison, TALENs show a significantly lower cytotoxicity and are easier to produce. However, both require protein engineering to change their sequence specificity.

Not long after the introduction of TALENs, the first description of the ***c**lustered **r**egularly **i**nterspaced **s**hort **p**alindromic **r**epeats* (CRISPR) / ***C**RISPR-**as**sociated system* (CRISPR/Cas

system) and its potential for genome editing pushed all other methods into the background.[157, 161] Since then, various CRISPR/Cas systems have been employed for scarless genome engineering in a multitude of organisms, including not only microorganisms like *E. coli*, but also plants and mammals, including human.[161-163] In 2020, Emmanuelle Charpentier and Jennifer A. Doudna were awarded the Nobel Prize in Chemistry for the discovery of CRISPR/Cas and its application as a method for genome editing.[164]

In nature, CRISPR/Cas provides adaptive immunity against viruses in prokaryotes. For this purpose, short sequences of viral origin (spacers) are stored in between repetitive palindromic sequences in the prokaryote's genome (CRISPR array). For defense against phages, the CRISPR array is transcribed and processed. So-called guide RNAs (gRNA), which contain only one spacer sequence, are used to direct the cleavage of the phage genome by Cas endonucleases.[165] The application of CRISPR/Cas systems in targeted genome editing is usually based on two components: (i) the adaptation of the guided DNA endonuclease activity of a Cas protein as used in the defense mechanism described above, and (ii) a repair method to fix the resulting double-strand break in the genome of the targeted organism.[166-168] In addition, the guide function of Cas proteins have been utilized in gene regulating approaches omitting the introduction of DSBs. For this, a catalytically inactive variant of a Cas protein (e.g. dCas9) can be used to regulate the transcription of a target gene either by blocking the transcription or fusing dCas9 with an effector (CRISPR interference, CRISPRi and CRISPR activation, CRISPRa).[169]

Various CRISPR/Cas systems from different prokaryotes with different modes of action have been investigated and adapted for genome engineering. The most prominent one, CRISPR/Cas9 originating from *Streptococcus pyogenes*, is used in this work for the development of *E. coli*-derived biocatalysts. Its function as well as challenges and limits are described in detail in the next chapter.

1.2.2 CRISPR/Cas9-assisted recombineering in *E. coli*

First introduced by Jiang et al. in 2013,[170] the combination of CRISPR/Cas9-induced chromosomal DSBs and λ-Red recombineering in *E. coli* has since been described repeatedly in the literature.[171] In contrast to eukaryotes, Cas9-induced chromosomal DSBs are lethal to many bacteria including *E. coli.* Like most microbes, *E. coli* does not contain a NHEJ pathway and relies on homologous recombination with sister chromosomes to repair DSBs. However, the simultaneous cleavage of all copies of the targeted sequence cannot be repaired.[166, 167] Thus, the combination with phage-derived recombineering methods (e.g. λ-Red) is necessary for efficient CRISPR/Cas9-assisted genome engineering in *E. coli*.

The approaches described in the literature differ in (i) the choice of the target *E. coli* strain, (ii) the details in the experimental design, and (iii) the aim, e.g. chromosomal integration or deletion of genes.

Most studies rely on *E. coli* K-12 derivatives (MG1655, BW25113, W3110)[117, 163, 172-177], but successful genome editing was also demonstrated in *E. coli* B strains (BL21 and BL21(DE3))[178-181] and a W strain.[181] In general, K-12 strains are preferred over B strains, partly because the latter is considered *hard-to-edit* and shows a reduced editing efficiency.[177, 178, 181]

In the following, different aspects of the experimental design are discussed. The very first application of CRISPR/Cas9 in *E. coli* utilized a gRNA complex as found in nature which is formed by two RNAs: a ***CR****ISPR-RNA* (crRNA) and a ***tr****ans-****a****ctivating* ***CR****ISPR-RNA* (tracrRNA). The crRNA contains a 20-nucleotide-long sequence complementary to the target DNA sequence (sometimes still referred to as *spacer* based on their biological role). The tracrRNA contains the scaffold for binding of the RNA to Cas9, forming the Cas9 ribonucleoprotein. However, today, a less complex chimeric single gRNA design (*single-guide RNA*, sgRNA) instead of the two components is usually preferred.[161] Thus, the Cas9 ribonucleoprotein is an easily programmable endonuclease which does not require protein engineering like ZFNs for changing the target. Instead, the target is defined by the

gRNA which is transcribed independently of Cas9. The simultaneous transcription of multiple gRNAs can be used for multiplexed CRISPR technologies.[182]

The binding of the Cas9 ribonucleoprotein to its target sequence (sometimes also referred to as protospacer) is initiated at the so-called ***p****rotospacer* ***a****djacent* ***m****otif* (PAM). The PAM is located downstream of the target sequence and consists of only a few nucleotides which differ between different CRISPR/Cas systems. In case of Cas9 from *S. pyogenes*, the PAM is the three-nucleotide long sequence 5'-NGG-3', in which N stands for any nucleotide and G stands for guanine. In theory, any 5'-NGG-3' motif could be used to design a specific gRNA. More than 400,000 potential PAMs can be found in *E. coli* which enables targeting of almost every part of its 4.6 million base pair long genome.[151] However, not all 20-nucleotide-long sequences are suitable as target sequences and the targeting efficiency as well as specificity (off-target effects) of potential gRNAs vary dramatically. As it is relatively well understood which sequence properties have which influence on the editing efficiency, various bioinformatic tools are available to facilitate experimental design.[183] The Cas9 protein contains two independent nuclease domains, HNH and RuvC, each targeting one strand of the DNA. The single mutations H840A in the HNH domain and D10A in the RuvC domain can independently render each domain inactive. Apart from the catalytically dead variant (dCas9) which can be used for CRISPRa or CRISPRi as mentioned before, CRISPR-Prime editing using a Cas9-nickase (Cas9n) has been developed for DSB-free base editing.[184]

Most studies that applied CRISPR/Cas9-assisted recombineering in *E. coli* used plasmids for the expression of the *cas9* and λ-Red genes and the transcription of the gRNA. Especially when CRISPR/Cas9 is only a means to an end and not the actual subject of the study, this route is more suitable due to its easy application and removability compared to chromosomally integrated *cas9*. Since the *cas9* gene is prone to inactivation,[185] the chromosomal integration might be wasted effort. It should also be noted that the Cas9 protein can lead to double-strand breaks even without gRNA present and thus has a cytotoxic effect.[186, 187] Therefore, a two-plasmid approach is widely used, which encodes *cas9* and λ-Red genes on one plasmid and

the target-specific gRNA on the second plasmid. The latter can then be quickly exchanged for iterative genome engineering.[163]

If heterologous genes or expression cassettes are introduced into *E. coli*'s chromosome via CRISPR/Cas9-assisted recombineering, there are few aspects to consider. First and foremost, a suitable integration locus has to be selected. Integration can be carried out either in between chromosomal genes (intergenic) or into a host gene (intragenic), inactivating it. Although this can be exploited for screening purposes, e.g. integration into *lacZ* for blue-white selection using X-gal,[181] it should be carefully considered whether the lost gene function might negatively affect the host. Furthermore, the integration locus can influence the expression of the integrated expression cassette. Depending on the locus, the expression strength varies 2-4-fold with individual outliers of up to ~300-fold. The search for suitable integration loci is therefore of great relevance.[188-190] In addition, the design of the repair template which contains the construct to be integrated has a major influence on the success of the integration procedure. It should be noted that although CRISPR/Cas9-induced DSBs are considered lethal to *E. coli*, so-called "escaper" colonies evade this counter-selection mainly because of inactivating mutations in the *cas9* gene.[185] It is therefore necessary to screen for successful integration clones, e.g. by colony PCR. The integration efficiency decreases with the length of the construct[172, 177, 181] and increases with the length of the homology arms.[174] Successful integration of up to 12-kb-constructs has been demonstrated before.[172, 191] However, the screening effort increases with increasing length as it becomes less likely to find a clone with successful integration. Furthermore, for the integration of a smaller gene short homology arms of only 50 bp can be sufficient. For longer constructs (>1 kb), longer homology arms are preferred.[174]

1.2.3 Genome engineering in context of biocatalysis

Genome engineering can be applied in biocatalysis for different reasons. In general, three different purposes can be distinguished:

(i) introduction of new functions,
(ii) deletion of unwanted functions (prevention of side reactions), and
(iii) regulation of activity.

This can be most easily illustrated using the example of the development of a biocatalyst in *E. coli* (Figure 5). As outlined in Chapter 1.1.3, once a suitable sequence has been found, heterologous expression can be carried out either plasmid-based or chromosomal. Plasmid-based expression is preferable for the development and optimization phase, including protein engineering. Once optimization of the enzyme's properties has been completed, it is then possible to switch to plasmid-free chromosomal expression to generate a genetically stable recombinant *E. coli* strain and to achieve antibiotic-free production of the enzyme. Chromosomal integration also enables the expression of biosynthetic pathways without the problems associated with the simultaneous propagation of several plasmids, such as slow cell growth.

However, the application of genome engineering in context of biocatalysis is not limited to the transfer of plasmid-based to plasmid-free expression systems. For example, by introducing a new RNA polymerase, a different promoter can be used for optimized expression both from plasmids and the chromosome.

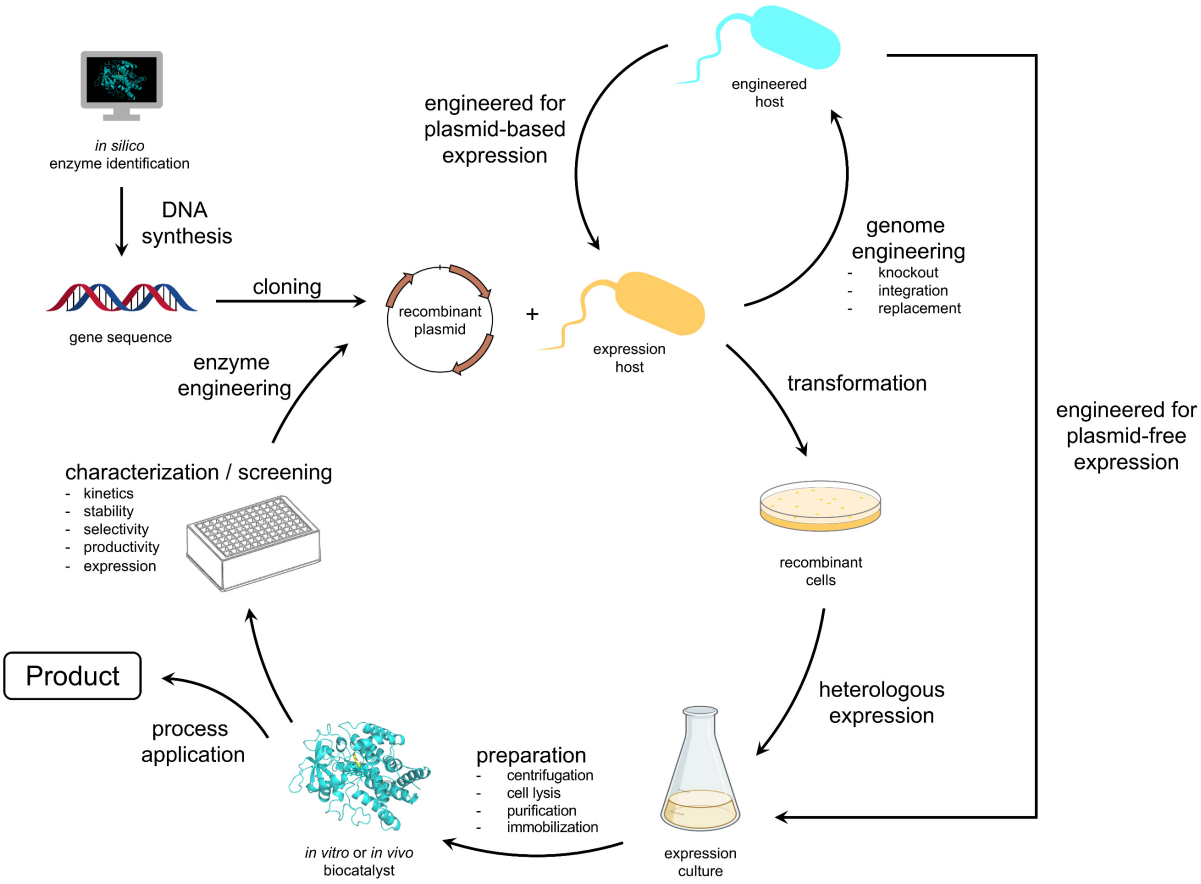


Figure 5: Genome engineering in context of the development of a new biocatalyst.

Sometimes, in biocatalysis, there are also problems with side reactions catalyzed by enzymes of the host. Although these can be avoided by laborious purification of the target enzyme, a targeted knockout of the competing enzyme solves the problems permanently. This not only eliminates the need for purification, but also opens up new possibilities for catalysis *in vivo*.

The third reason for genome engineering, regulation of activity, is particularly interesting in context of enzyme cascades and pathways. By selectively adjusting the copy number of individual genes or modifying regulatory elements such as promoters or ribosomal binding sites, bottlenecks in pathways can be eliminated.

1.3 Objectives

A wide variety of challenges can arise during biocatalysis and biotransformation. In addition to intrinsic problems of the catalyst such as stability, selectivity and productivity, which can be addressed e.g. by protein engineering, there are also challenges in the production and application of the catalyst. Enzyme catalysis can be performed *in vitro* either using crude extracts or purified enzymes. The latter requires a significant amount of work and thus is cost-intensive. However, the use of the less expensive crude extracts can be impaired by side reactions catalyzed by the host's enzymes, just as in whole cells. Furthermore, plasmid-based heterologous expression is accompanied by issues with plasmid stability, cell-to-cell variability and antibiotic resistance. All these issues can be addressed with the application of CRISPR/Cas9-assisted recombineering which allows marker-free modifications to the chromosome of *E. coli* including gene deletion and integration. Apart from optimizations of the host which affect plasmid-based expression, the transition from plasmid-based to plasmid-free expression systems was the main focus of this work. Moreover, broad applicability of the developed tools in context of biocatalysis and the comparison of various *E. coli* laboratory strains were in the center of attention.

In particular, the following applications and further developments of CRISPR/Cas9-assisted recombineering were realized in this work:

i) Gene knockouts to prevent side reactions catalyzed by host enzymes, using the example of catalases that degrade hydrogen peroxide added for peroxygenase-catalyzed oxidation reactions with *E. coli* resting cells or cell-free extracts

ii) Extension of *E. coli* K-12 strains capable of T7 expression by chromosomal integration of the gene encoding T7 RNA polymerase

iii) Evaluation of the effects of chromosomal integration on the catalytic performance of a multi-component P450 system for the production of a human metabolite of the anesthetic (*S*)-ketamine with *E. coli* resting cells

iv) Development and evaluation of a versatile toolbox for the simple and reliable transfer of plasmid-based to plasmid-free expression systems in *E. coli* K-12 and B strains

v) Reconstitution of an artificial biosynthetic pathway in growing *E. coli* cells for the plasmid-free production of the high value furofuran lignan pinoresinol from the relatively inexpensive phenylpropanoid ferulic acid

vi) Development of a system for combinatorial substitution of promoters after chromosomal integration of heterologous genes

2 Results

The results are divided into four subchapters and are presented in the form of three peer-reviewed publications and a manuscript containing so far unpublished results. These four subchapters describe multiple applications of CRISPR/Cas-assisted genome engineering in different *Escherichia coli* strains.

In Chapter 2.1, targeted in-frame gene knockouts of the two catalase genes *katE* (encoding hydroperoxidase II) and *katG* (encoding hydroperoxidase I) in *E. coli* BL21-Gold(DE3) were applied to prevent the degradation of hydrogen peroxide, which is used as co-substrate by the actual biocatalysts of interest - different peroxygenases. This catalase-deficient *E. coli* strain was used for the expression of peroxygenases, and then applied either in form of whole-cell catalysts or after lysis in form of crude extracts without the need for time-consuming purification of the enzymes.

In Chapter 2.2, a P450 system consisting of a CYP154E1 variant from *Thermobifida fusca* YX as well as the redox partner proteins YkuN from *B. subtilis* and FdR from *E. coli*, was used for the enantioselective two-step oxidation of the anesthetic (*S*)-ketamine to its human metabolite (2*S*,6*S*)-hydroxynorketamine in resting *E. coli* cells. Plasmid-based and chromosomal expression in different *E. coli* strains were compared to examine the effect of chromosomal integration on the catalytic performance of this three-component P450 system. In the course, the host's ability for T7 expression was extended to different *E. coli* strains.

Chapter 2.3 describes the reconstitution of a plasmid-free four-step synthesis of the valuable lignan pinoresinol from the relatively cheap phenylpropanoid ferulic acid. The endogenous reducing capabilities of the host were combined with heterologous enzymes from the plants *Petroselinum crispum* and *Zea mays* to yield the monolignol coniferyl alcohol. The final single electron oxidation was catalyzed by a laccase from *Corynebactrium glutamicum* and led to subsequent oxidative coupling of coniferyl alcohol to the target compound pinoresinol.

In Chapter 2.4, a proof-of-concept study for combinatorial promoter substitution after chromosomal integration of heterologous genes in *E. coli*, is presented. To this end, the two

genes *cyp154e1m* and *ykuN* were first integrated into two different loci under the control of the T7 promoter and then the promoters were simultaneously replaced at both loci to obtain combinations of T7, T5 and tac promoters.

The individual contribution of the author of this dissertation to each manuscript is outlined at the start of each chapter.

2.1 Gene knockouts for the prevention of side reactions in crude extracts and whole cells

Title: Use of Whole Cells and Cell-Free Extracts of Catalase-Deficient *E. coli* for Peroxygenase-Catalyzed Reactions

Authors: A. C. Ebrecht, **U. J. Luelf**, K. Govender, D. J. Opperman, V. B. Urlacher and M. S. Smit

Published in: Biotechnology and Bioengineering, 122, 1376-1385

DOI: 10.1002/bit.28959

Contribution: Design and conduction of CRISPR/Cas-assisted genome engineering experiments for targeted in-frame deletion of the catalase genes *katE* and *katG* to construct a catalase-deficient *E. coli* strain. Writing of the corresponding section of the manuscript and supporting information.

Biotechnology and Bioengineering

WILEY

BIOTECHNOLOGY BIOENGINEERING

ARTICLE OPEN ACCESS

Use of Whole Cells and Cell-Free Extracts of Catalase-Deficient *E. coli* for Peroxygenase-Catalyzed Reactions

Ana C. Ebrecht[1] | U. Joost Luelf[2] | Kamini Govender[1] | Diederik J. Opperman[1] | Vlada B. Urlacher[2] | Martha S. Smit[1]

[1]Department of Microbiology and Biochemistry, University of the Free State, Bloemfontein, South Africa | [2]Institute of Biochemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

Correspondence: Martha S. Smit (smitms@ufs.ac.za)

Received: 15 October 2024 | **Revised:** 4 February 2025 | **Accepted:** 15 February 2025

Funding: This work was supported by the South African Council for Scientific and Industrial Research—Industrial Biocatalysis Hub (CSIR-IBH) initiative funded by the South African Department of Science and Innovation (DSI) and the Technology Innovation Agency (TIA).

Keywords: catalase-deficient *E. coli* | epoxidation | hydrogen peroxide | hydroxylation | peroxygenase | sulfoxidation

ABSTRACT

Unspecific peroxygenases (UPOs) and cytochrome P450 monooxygenases (CYPs) with peroxygenase activity are becoming the preferred biocatalysts for oxyfunctionalization reactions. While whole cells (WCs) or cell-free extracts (CFEs) of *Escherichia coli* are often preferred for cofactor-dependent monooxygenase reactions, hydrogen peroxide (H_2O_2) driven peroxygenase reactions are generally performed with purified enzymes, because the catalases produced by *E. coli* are expected to quickly degrade H_2O_2. We used the CRISPR/Cas system to delete the catalase encoding chromosomal genes, *katG*, and *katE*, from *E. coli* BL21-Gold (DE3) to obtain a catalase-deficient strain. A short UPO, *Dca*UPO, and two CYP peroxygenases, *Ssca*CYP_E284A and CYP102A1_21B3, were used to compare the strains for peroxygenase expression and subsequent sulfoxidation, epoxidation, and benzylic hydroxylation activity. While 10 mM H_2O_2 was depleted within 10 min after addition to WCs and CFEs of the wild-type strain, at least 60% remained after 24 h in WCs and CFEs of the catalase-deficient strain. CYP peroxygenase reactions, with generally lower turnover frequencies, benefited the most from the use of the catalase-deficient strain. Comparison of purified peroxygenases in buffer versus CFEs of the catalase-deficient strain revealed that the peroxygenases in CFEs generally performed as well as the purified proteins. We also used WCs from catalase-deficient *E. coli* to screen three CYP peroxygenases, wild-type SscaCYP, *Ssca*CYP_E284A, and *Ssca*CYP_E284I for activity against 10 substrates comparing H_2O_2 consumption with substrate consumption and product formation. Finally, the enzyme-substrate pair with highest activity, *Ssca*CYP_E284I, and *trans*-β-methylstyrene, were used in a preparative scale reaction with catalase-deficient WCs. Use of WCs or CFEs from catalase-deficient *E. coli* instead of purified enzymes can greatly benefit the high-throughput screening of enzyme or substrate libraries for peroxygenase activity, while they can also be used for preparative scale reactions.

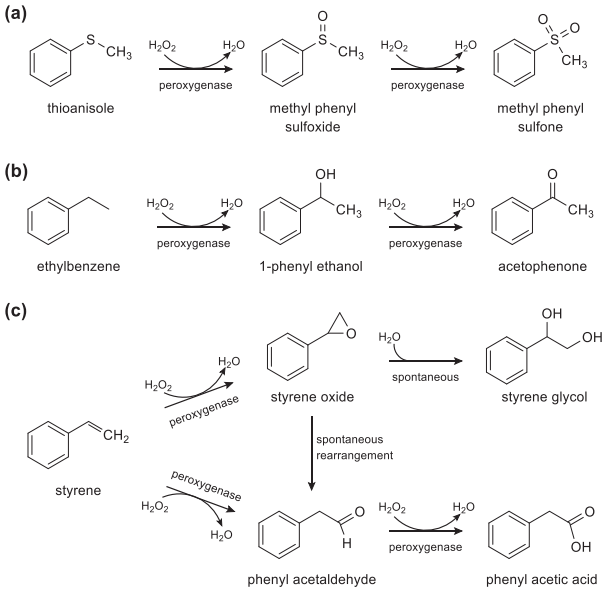
1 | Introduction

Unspecific peroxygenases (UPOs) and cytochrome P450 monooxygenases (CYPs) with peroxygenase activity are becoming the preferred catalysts for reactions traditionally performed with monooxygenases (Monterrey et al. 2023; Xu et al. 2023). While whole cells (WCs) or cell-free extracts (CFEs) of *Escherichia coli* are often preferred for monooxygenase-catalyzed reactions which require reduced cofactors and their regeneration, hydrogen peroxide (H_2O_2) driven peroxygenase reactions are generally performed with purified enzymes, since it is assumed that the catalases produced by *E. coli* will quickly catalyze the disproportionation of H_2O_2. Two

catalases (katG and katE), together with an alkyl hydroperoxide reductase (AhP) are responsible for scavenging H_2O_2 in *E. coli* with the catalases dominating when H_2O_2 levels exceed 20 μM (Xu et al. 2020; Liu et al. 2021). In the case of UPOs, *Pichia pastoris* is the preferred host for heterologous expression because these are extracellular glycosylated fungal enzymes that generally do not express well in *E. coli* (Kinner et al. 2021, Monterrey et al. 2023). However, *E. coli* is gradually more often used for the expression of short UPOs (Linde et al. 2020, Kinner et al. 2021), and recently a superfolder-green-fluorescent-protein (sfGFP) mediated secretion system was developed that facilitated the expression of four different UPOs, including the long UPO, *Aae*UPO, in *E. coli* (Yan et al. 2024). The activity of these UPOs displayed on the cell surface of *E. coli* was detected using WCs and CFEs of *E. coli* BL21(DE3), despite possible disproportionation of H_2O_2 by the well-described catalases of *E. coli*.

Catalase-deficient strains of *E. coli* have been constructed and characterized (Nakagawa et al. 1996; Hui et al. 2014; Xu et al. 2020; Liu et al. 2021). Nakagawa et al. (1996) constructed a catalase-deficient strain of *E. coli* for the large-scale production of catalase-free uricase preparations. They found that deletion of both chromosomal genes, *katG,* and *katE*, from a strain derived from *E. coli* K-12 did not affect growth or uricase production. The same catalase-deficient *E. coli* strain was subsequently used by the Arnold group to develop variants of CYP102A1 with peroxygenase activity (Cirino and Arnold 2002, 2003; Salazar et al. 2003). More recently Xu et al. (2020) deleted *katG* and *katE* from *E. coli* BL21 (DE3), the commercially available B strain commonly used for heterologous protein expression. They used this catalase-deficient *E. coli* strain to develop a high-throughput screening method relying on H_2O_2 consumption detected by the colorimetric Amplex Red assay. This screening method was used to screen large DNA shuffling and random mutagenesis libraries of the fatty acid decarboxylases OleTJE (CYP152L1) and CYP-Sm46Δ29 (CYP152L2) for improved variants (Xu et al. 2020).

We deleted *katG* and *katE* from *E. coli* BL21-Gold(DE3), to investigate the use of WCs and CFEs from catalase-deficient strains expressing different known peroxygenases for different hydrogen peroxide driven reactions. Three heme-thiolate-based peroxygenases were selected, namely *Dca*UPO, a short UPO from *Daldinia caldariorum* (Linde et al. 2020), *Ssca*CYP_E284A, a variant of *Ssca*CYP a CYP peroxygenase from *Streptomyces scabiei* which contains an Asp instead of the usual Thr in the I helix, both previously described by our group (Ebrecht et al. 2023), and the CYP102A1_21B3 CYP peroxygenase developed by the Arnold group from the N-terminal heme-domain of CYP102A1 (Cirino and Arnold 2003). Three reactions which all three enzymes can catalyze to various degrees were selected to evaluate peroxygenase activity using WCs and CFEs. These were sulfoxidation of thioanisole, benzylic hydroxylation of ethylbenzene, and conversion, mainly epoxidation, of styrene (Scheme 1). Next, we evaluated the robustness of the colorimetric H_2O_2 consumption assay by screening the wild-type *Ssca*CYP, together with two mutants, *Ssca*CYP_E284A used above and *Ssca*CYP_E284I also previously described by us (Ebrecht et al. 2023), for activity against 10 different substrates. From this screening, we finally selected the enzyme-substrate pair with highest activity, *Ssca*CYP_E284I, and *trans*-β-



SCHEME 1 | Reactions used for comparing activities were (a) sulfoxidation of thioanisole to methyl phenyl sulfoxide, (b) hydroxylation of ethylbenzene to 1-phenyl ethanol and further to acetophenone, and (c) epoxidation or anti-Markovnikov type oxidation of styrene to styrene oxide and phenylacetaldehyde, respectively. Methyl phenyl sulfone, 1-phenyl ethanol, and phenylacetaldehyde can be further oxidized in a second round of peroxygenase reactions. Phenylacetaldehyde might also be formed by rearrangement of styrene oxide (Aschenbrenner et al. 2024).

methylstyrene, for a preparative scale reaction using catalase-deficient WCs.

2 | Results and Discussion

CRISPR/Cas-assisted λ-red recombineering was used for successive in-frame deletions of the *katE* and *katG* genes in *E. coli* BL21-Gold(DE3). Successful knockout mutants were identified using colony PCR and verified by Sanger sequencing. Deletion of the catalase-encoding genes was further confirmed by monitoring the fate of H_2O_2 (10 mM) added to WCs and CFEs of wild-type *E. coli* BL21-Gold(DE3) (WT) and its catalase-deficient derivative (CD), both transformed with empty pET-28a (+). As expected, H_2O_2 was quickly depleted by catalase-containing WCs and CFEs from the catalase-containing *E. coli*, with no H_2O_2 left after 20 min (Figure 1a). However, H_2O_2 levels dropped slowly in WCs and CFEs from the catalase-deficient strain, with more than 5 mM left after 8 h which remained up to 24 h. Xu et al. (2020) also observed H_2O_2 consumption (ca. 20%) in their catalase-deficient strain with *katE* and *katG* genes deleted and ascribed it to the activity of the alkyl hydroperoxide reductase (AhP) which functions at low H_2O_2 concentrations (< 10 μM) or other unknown scavenging enzymes. In their experience a triple deletion strain with the *ahp* gene also deleted displayed severe growth defects. After 8 h H_2O_2 stability was similar to what we observed in buffer without WCs or CFEs, where 7 mM H_2O_2 remained after 24 h.

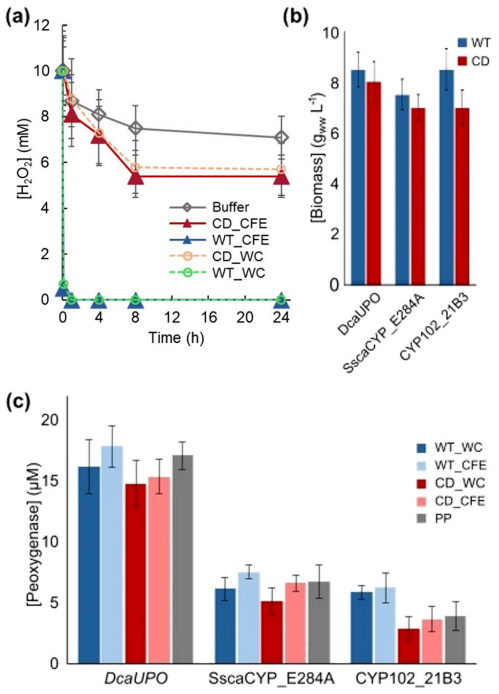


FIGURE 1 | (a) Stability of H_2O_2 (10 mM) in buffer containing WCs (0.1 g wet weight mL^{-1}) (open circles) or CFEs (produced after cell disruption at 0.1 g wet weight mL^{-1}) (closed triangles) from wild-type (WT) and catalase-deficient (CD) *E. coli* and in buffer without WCs or CFEs (gray open diamonds) at 25°C, pH7. (b) Biomass (wet weight) recovered from cultures of WT and CD *E. coli* when cells were harvested for biotransformations. (c) Concentrations of different peroxygenases (based on CO difference spectra) in biotransformation reactions containing WCs (dark colors) or CFEs (light colors) from WT (blue) and CD (red) *E. coli* (0.1 g wet weight mL^{-1}) as well as purified protein (grey). Averages and standard deviations were calculated from at least three independent repeats.

The catalase-deficient strain displayed satisfactory growth and enzyme production. Biomass harvested from its cultures expressing the different peroxygenases was 6%–18% less than from wild-type cultures (Figure 1b). *Dca*UPO and *Ssca*CYP_E284A expressed equally well in both catalase-deficient and wild-type strains, while expression of CYP102A1_21B3 was lower in the catalase-deficient strain (Figure 1c).

Given how quickly H_2O_2 is depleted by catalase-containing WCs and CFEs, it was surprising that in 24 h reactions, sulfoxidation of thioanisole by all three peroxygenases could be detected with WCs and CFEs of the wild-type strain even when H_2O_2 (20 mM) was added in a single dose to start reactions (Figure 2a). The initial sulfoxidation of thioanisole by *Dca*UPO is evidently a very fast reaction since there was essentially no difference between reactions with catalase-containing and catalase-deficient CFEs both giving approximately 80% conversion of the substrate. WC sulfoxidation by *Dca*UPO benefited from the use of catalase-deficient cells, most likely because H_2O_2 diffusion into the cells is quicker than thioanisole diffusion. When H_2O_2 was produced in situ by glucose oxidation with glucose oxidase (GOx) there was a significant difference in the product distribution from catalase-containing and catalase-deficient CFEs, with further oxidation of the sulfoxide to sulfone by *Dca*UPO more prevalent in catalase-deficient CFEs (Figure 2a). In the case of *Ssca*CYP_E284A, which was expressed at lower concentrations, there was a clear benefit to using the catalase-deficient strain, whether H_2O_2 was added in a single dose or produced in situ by GOx. With catalase-deficient WCs and CFEs, however, significant sulfoxidation to the sulfone was observed. Chiral analysis of extracts from reactions with catalase-deficient WCs and CFEs containing *Ssca*CYP_E284A revealed that use of WCs or CFEs did not reduce enantioselectivity when compared with previous results obtained with purified protein when the (*S*)-enantiomer of the sulfoxide was produced with 81% *ee* (Ebrecht et al. 2023) (Figure S7). Activity of CYP102A1_21B3 towards thioanisole was much lower than that of the other two enzymes, with only traces of sulfoxide observed with wild-type WCs and CFEs. Highest conversions were achieved when H_2O_2 was supplied in situ by GOx. However, no further oxidation to the sulfone was observed with CYP102A1_21B3 (Figure 2a).

All three peroxygenases displayed lower activity toward ethylbenzene than to thioanisole. *Ssca*CYP_E284A displayed barely detectable activity with wild-type and catalase-deficient CFEs only when H_2O_2 was supplied in situ by GOx (Figure 2b). In *Dca*UPO-catalyzed reactions activity was low with WCs in both strains. In similar CFE reactions with H_2O_2 also added as a single dose,

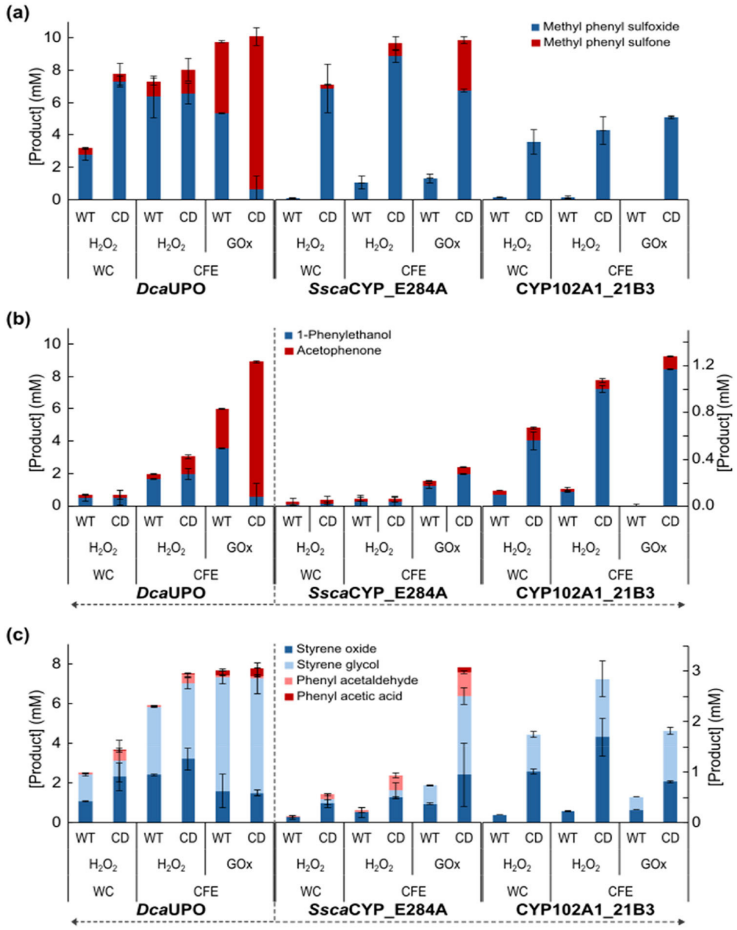


FIGURE 2 | Concentrations of different products formed by *Dca*UPO, *Ssca*CYP_E284A, and CYP102A1_21B3 from (a) thioanisole, (b) ethylbenzene and (c) styrene in 24 h reactions using WCs or CFEs of wild-type (WT) and catalase-deficient (CD) *E. coli*. H_2O_2 was added in a single dose to start reactions or H_2O_2 was produced in situ by oxidation of glucose with GOx. Concentrations of products from thioanisole are displayed by the axis on the left. Arrows point to products formed by *Dca*UPO from ethylbenzene and styrene displayed by the axis on the left and those from ethylbenzene and styrene produced by *Ssca*CYP_E284A and CYP102A1_21B3 by the axis on the right. Reactions contained WCs or CFEs of WT and CD *E. coli* (0.1 g wet weight mL^{-1}), substrate (10 mM), acetone (5% (v/v)), H_2O_2 (20 mM) or GOx (0.2 U mL^{-1}) with glucose (100 mM). Averages and standard deviations were calculated from at least three independent reactions.

activity in catalase-deficient CFEs was higher with more of the initially formed 1-phenyl ethanol further oxidized to acetophenone. When H_2O_2 was supplied in situ by GOx the advantage to using catalase-deficient CFEs became even more apparent with ca. 80% of ethylbenzene hydroxylated and subsequently oxidized to acetophenone while with wild-type CFEs ca. 60% was hydroxylated with only ca. 24% further oxidized to acetophenone. Activity of CYP102A1_21B3 toward ethylbenzene was much lower than that of *Dca*UPO with in the best reactions only ca. 10% conversion and very little oxidation to acetophenone. However, activities with both WCs and CFEs were significantly improved when the catalase-deficient strain was used (Figure 2b).

Oxidation of styrene by these peroxygenases yielded mainly styrene oxide which was in 24 h reactions spontaneously hydrolyzed to styrene glycol. In the case of *Dca*UPO, which again yielded the most product given its high concentration, small amounts of phenylacetaldehyde were also detected, some of which were oxidized to phenylacetic acid (Figure 1c). In these reactions, activity was only slightly improved when catalase-deficient WCs and CFEs were used with H_2O_2 added in a single dose. With in situ generation of H_2O_2 by GOx there was in 24 h reactions, no benefit to using catalase-deficient CFEs. CYP102A1_21B3 and *Ssca*CYP_E284A activities in all cases benefited significantly from the use of the catalase-deficient

strain. *Ssca*CYP_E284A produced relatively more phenylacetaldehyde than the other two peroxygenases.

Next, we compared the performance of purified peroxygenases in buffer and peroxygenases in CFEs of catalase-deficient *E. coli* in a series of time-course experiments. In these experiments, H_2O_2 was added as a single dose or produced in situ using either GOx or formate oxidase (FOx). A total number of oxygenations was calculated and plotted assuming that the formation of sulfone from thioanisole, acetophenone from ethylbenzene, and phenylacetic acid from phenylacetaldehyde required in each case two sequential peroxygenase reactions rather than peroxidase activity (Figure S1). These values were used to calculate turnover numbers (TONs) at the times when reactions leveled off (Figure 3, Table S3). It is evident from these results that the peroxygenases in CFEs of the catalase-deficient strain generally performed at least as well as the purified peroxygenases in buffer. The exception was the purified *Dca*UPO which, when H_2O_2 was produced in situ, achieved maximum product concentrations already after 8 h in thioanisole and ethylbenzene conversions, while these reactions with CFE took 24 h (Figure S1, Table S3). This might be due to the GOx and FOx being unable to supply sufficient H_2O_2 quickly enough for the high concentration of *Dca*UPO when enzymes in the CFEs consumed some of the H_2O_2, glucose and/or formate. On the other hand, in the case of CYP102A1_21B3, which displayed low activity on all substrates, the CFEs generally yielded better results than the purified peroxygenases, possibly because the CFEs improved stability of the CYP102A1_21B3. Notable from these time course experiments are the reactions with styrene which in all cases with all three enzymes leveled off within 2 h, indicating that styrene most likely inhibited or denatured these peroxygenases. Activity of *Ssca*CYP_E284A toward ethylbenzene, in all cases, ceased within 1 h.

Xu et al. (2020) used CFEs of catalase-deficient *E. coli* in the H_2O_2-dependent high throughput screening method they had used to screen libraries derived from $OleT_{JE}$ (CYP152L1) and CYP-Sm46Δ29 (CYP152L2) for decarboxylase activity against lauric acid. We explored whether WCs of catalase-deficient *E. coli* can similarly be used by screening *Ssca*CYP_E284A as well as WT *Ssca*CYP and a second mutant *Ssca*CYP_E284I for activity against 10 different substrates (Figure S8). Although H_2O_2 consumption could reliably be detected in WC assays, the correlation between H_2O_2 consumption and activity detected with GC analysis (substrate consumption and product formation) was evidently influenced by the enzymes and substrates (Figure 4). H_2O_2 consumption accurately indicated the activity of *Ssca*CYP_E284I against *trans*-β-methylstyrene, α-methylstyrene, 4-methoxybenzyl alcohol, and vanillyl alcohol and the activity of *Ssca*CYP_E284A against *trans*-β-methylstyrene. However, it did not indicate activity of *Ssca*CYP_E284I against 1,5-cyclooctadiene and of *Ssca*CYP_E284A against 4-methoxybenzyl alcohol, both for which products were detected with GC, or any activity against styrene, which all three enzymes converted. The presence of the CYPs in the absence of substrate increased H_2O_2 consumption in an enzyme-dependent manner, with *Ssca*CYP_E284I consuming at least 60% more H_2O_2 than the no-CYP control. In some instances, the presence of substrate reduced H_2O_2 consumption, particularly in the case of *Ssca*CYP_E284I where propylbenzene, 3-methoxybenzyl

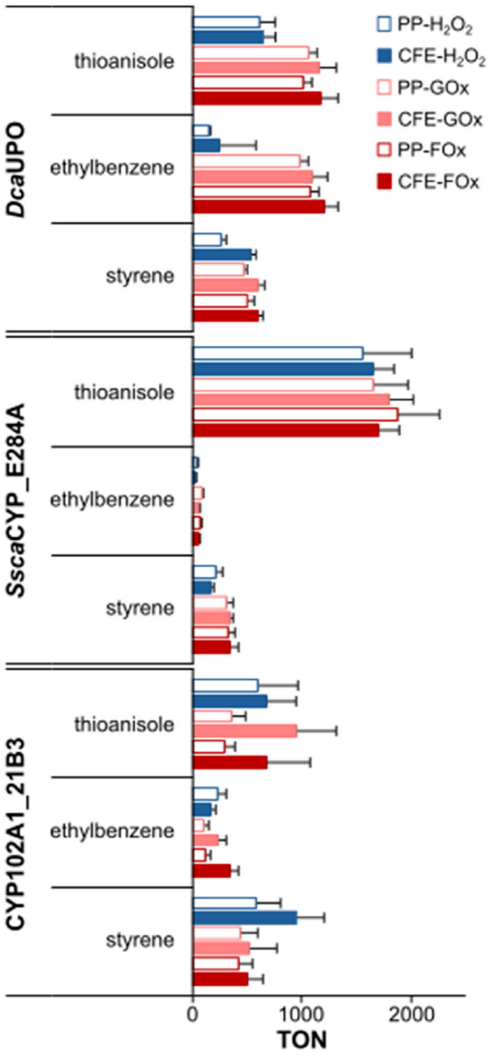


FIGURE 3 | Comparison of turnover numbers (TON) of *Dca*UPO, *Ssca*CYP_E284A, and CYP102A1_21B3 in CFEs of catalase-deficient *E. coli* and as purified protein. TONs were calculated from time course experiments at times when activities leveled off (Table S3, Figure S1). H_2O_2 was added in a single dose to start reactions or H_2O_2 was produced in situ by oxidation of glucose by GOx or by oxidation of formate by FOx. Reactions contained CFEs of catalase-deficient *E. coli* (0.1 g wet weight mL^{-1}) or a corresponding concentration (based on CO difference spectra) of purified protein, substrate (10 mM), acetone (5% [v/v]), H_2O_2 (20 mM) or GOx (0.2 U mL^{-1}) with glucose (100 mM) or FOx (0.2 U mL^{-1}) with formate (100 mM). Standard deviations of TONs were calculated using standard deviations of the product concentrations and the catalyst concentrations.

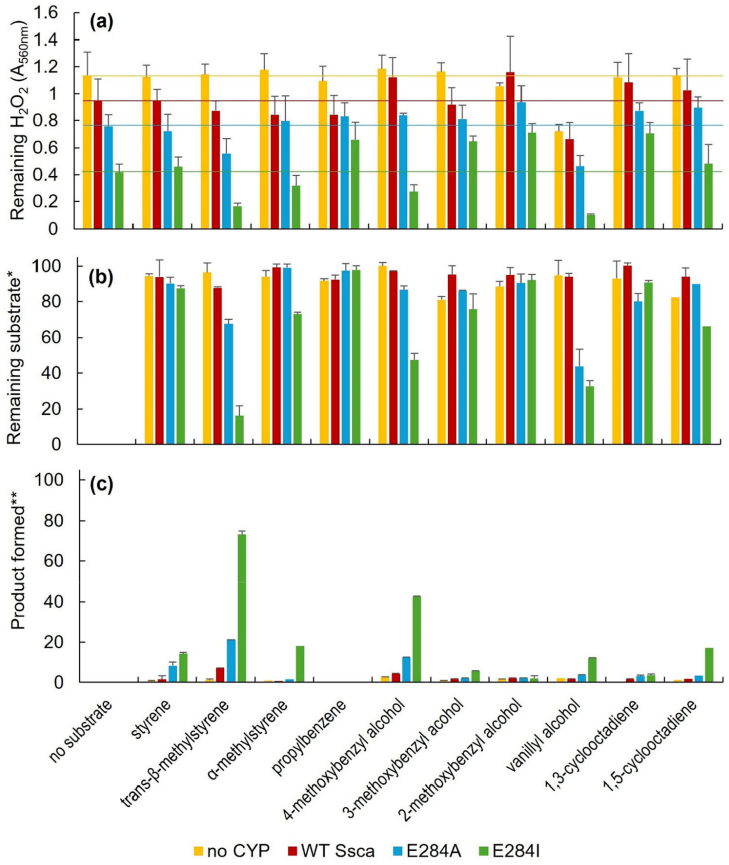


FIGURE 4 | (a) Remaining H_2O_2 (measured as absorbance at A_{560nm} in Ampliflu assay), (b) percentage remaining substrate, and (c) percentage product formed in reactions of WT *Ssca*CYP and its E284A and E284I mutants with 10 different substrates. Reactions (1 mL) contained WCs of catalase-deficient *E. coli* (0.1 g wet weight mL^{-1}), substrate (10 mM), acetone (5% [v/v]) and H_2O_2 (20 mM). H_2O_2 was added in a single dose to start reactions and assays and extractions were done after 4 h. The percentage remaining substrate and percentage product formed were calculated as percentage of maximum ratio of given substrate to internal standard in GC-FID assays. The percentage substrate and product are for reactions done in duplicate with Ampliflu assay for each reaction also done in duplicate.

alcohol, 2-methoxybenzyl alcohol, and 1,3-cyclooctadiene reduced H_2O_2 consumption by between 34% and 40%. On the other hand, with vanillyl alcohol residual H_2O_2 was markedly reduced even in the absence of enzyme. Although it is possible that H_2O_2 might directly react with a phenolic compound such as vanillyl alcohol, no product was detected in the GC analyses of the no-CYP control samples. It is more likely that this is an apparent reduction in H_2O_2 since interference of phenolic compounds with the Ampliflu Red assay has been described (Tama et al. 2023). The effects of enzymes and substrates on H_2O_2 consumption might explain why H_2O_2 consumption is not always suitable for detecting low activity levels. It might also explain why 20 variants selected by Xu et al. (2020) out of 8000 clones in their high-throughput screening based on H_2O_2 consumption, did not as purified enzymes display dramatically improved activity over the parental $OleT_{JE}$ (CYP152L1) and CYP-Sm46Δ29 (CYP152L2). High-throughput screening of peroxygenase libraries need not depend solely on H_2O_2 consumption, and the use of catalase-deficient WCs or CFEs can facilitate the development of complementary high-throughput screening methods for discovering improved or novel peroxygenases.

Finally, we used the conversion of *trans*-β-methylstyrene to cinnamaldehyde via cinnamyl alcohol by *Ssca*CYP_E284I to explore the use of catalase-deficient WCs in a preparative scale reaction. In this reaction, the biomass concentration was increased to 20 g wet weight $mL^{-1,}$ and H_2O_2 was added in 10 mM aliquots every 4 h up to 12 h and then at 24 h. H_2O_2 accumulation, CYP stability (CO-difference spectra) and product formation was followed over 28 h at which time the total reaction mixture was extracted (Figure 5). Although the

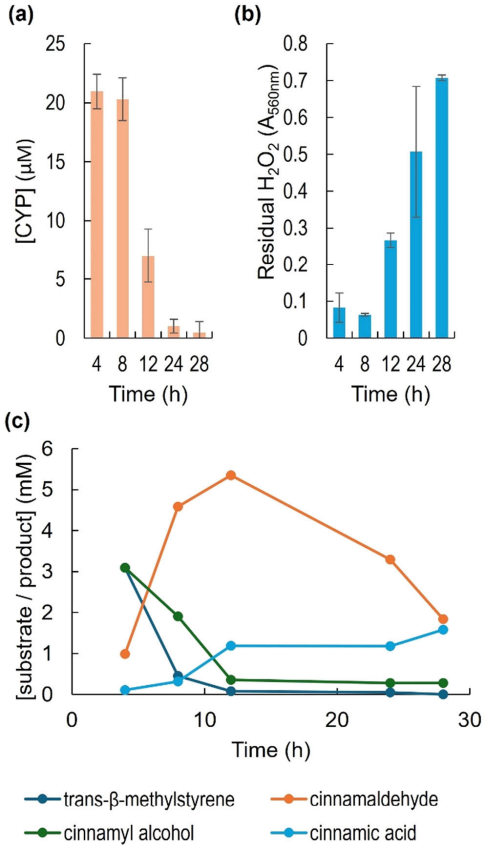


FIGURE 5 | Results from preparative scale conversion of *trans*-β-methylstyrene by the *Ssca*CYP_E284I mutant. The reaction mixture (50 mL) contained WCs of catalase-deficient *E. coli* (0.2 g wet weight mL^{-1}), initial (CYP) 25 μM, substrate (10 mM) and acetone (5% [v/v]). The reaction was started by the addition of H_2O_2 (10 mM) with subsequent additions every 4 h (up to 12 h, and then at 24 h). Samples were taken every 4 h (up to 12 h, then at 24 and 28 h) to analyze enzyme stability (CO-difference spectra) (a), H_2O_2 accumulation (Ampliflu assay) (b) and *trans*-β-methylstyrene conversion (GC-FID) (c) with the total reaction mixture extracted after 28 h.

trans-β-methylstyrene was essentially completely (99%) converted after 12 h, we allowed the reaction to proceed for 28 h to obtain complete conversion of cinnamyl alcohol to cinnamaldehyde. This unfortunately led to significant oxidation of the cinnamaldehyde to cinnamic acid. Additionally, cinnamaldehyde could have been epoxidized and then oxidized to benzaldehyde by the H_2O_2 (Chen et al. 2012). Traces of benzaldehyde were detected in the final extracts (Figure S11). The eventual poor recovery of cinnamaldehyde (only 18%) might also be ascribed to Schiff base adduct formation of cinnamaldehyde in the amino acid-rich environment of the WCs (Wei et al. 2011). Further experiments will be required to determine the role residual H_2O_2 and WCs played in the loss of cinnamaldehyde and to optimize reaction conditions and product recovery. However, our experiment demonstrates that catalase-deficient WCs are suitable for a preparative scale reaction in which over 5 mM cinnamaldehyde was formed after 12 h of reaction (Figure 5c).

The Arnold group employed a spectrophotometric assay in 96-well microtiter plates using clarified CFEs from mutant libraries expressed in catalase-deficient *E. coli* K-12 to screen for H_2O_2-driven hydroxylase activity (Cirino and Arnold 2003). It is surprising that there were no further reports on the use of such catalase-deficient *E. coli* strains in peroxygenase research until 2020 when Xu et al. for the first time reported using CFEs of a catalase deficient derivative of the commercially available, commonly used B strains in a H_2O_2 dependent high throughput screening for improved variants of CYP152 decarboxylases. Although it is possible, as demonstrated by Yan et al. (2024), to use WCs or CFEs of catalase-containing *E. coli* to screen large numbers of UPO variants for an improved characteristic, our results show that when activities are low, and specifically in the case of CYP peroxygenases, the use of catalase-deficient *E. coli* will be advantageous. We have demonstrated that peroxygenase containing WCs or CFEs from catalase-deficient *E. coli* generally perform as well as purified peroxygenases and that not only CFEs but also WCs of catalase-deficient *E. coli* can be used when screening for improved peroxygenases. However, it should be kept in mind that H_2O_2 is a very potent oxidizing agent which might react with the enzyme of interest destroying it or react directly with the substrate or products. Additionally, substrates or even products might interfere with H_2O_2 assays such as the peroxidase-dependent Ampliflu or Amplex Red assays. Thus, for both screening and preparative scale reactions extensive controls and optimization will be required for every enzyme-substrate combination whether purified peroxygenase, CFEs, or WCs are used.

3 | Materials and Methods

3.1 | Deletion of *katE* and *katG* in *E. coli* BL21-Gold(DE3)

CRISPR/Cas-assisted λ-red recombineering was used to create in-frame gene deletions. The start codon and the codons encoding the six residues of the C-terminus as well as the stop codon were left as described by Baba et al. (2006) for the construction of the Keio Collection.

The genome sequence of *E. coli* BL21(DE3) was used as a reference genome (GenBank accession number CP001509.3). Plasmids pEcCas (Addgene plasmid #73227) (Li et al. 2021) and pgRNA-bacteria (Addgene plasmid #44251) (Qi et al. 2013) were obtained from Addgene. The CRISPR/Cas-assisted λ-red recombineering was carried out as described before (Luelf et al. 2023). Briefly, gRNA targeting sequences were designed using the CHOPCHOP web toolbox (Labun et al. 2019) and cloned into the pgRNA-bacteria plasmid. For construction of the repair template, homology arms of approx. 500 bp length were amplified from boiled cells and combined by fusion PCR. Electrocompetent cells of *E. coli* BL21-Gold(DE3) harboring pEcCas were transformed with modified pgRNA and the repair

template as described in Supporting Information S1. The knockout was verified by colony PCR and sequencing (Eurofins Genomics, Germany). Curing of the plasmids pEcCas and pgRNA-bacteria was performed as described before (Luelf et al. 2023). Primers for amplification of the homology arms and exchange of the targeting sequence in the plasmid pgRNA are listed in Tables S1 and S2.

3.2 | Plasmids and Enzymes

For expression of the (CYP-)peroxygenases, the constructs pET-22b(+):CYP102A1_21B3, pET-28a(+):*Dca*UPO, pET-28a(+):*Ssca*CYP_E284A, pET28a(+):*Ssca*CYP and pET-28a(+):*Ssca*CYP_E284I were obtained as described elsewhere (Aschenbrenner et al. 2024; Ebrecht et al. 2023; Ebrecht et al. 2023). The construct pET-21c(+):AoFOx for expression of the formate oxidase (FOx) from *Aspergillus oryzae* was kindly provided by Prof. Frank Hollmann (Delft University of Technology, the Netherlands) with permission of Prof. Andreas Bommarius (Georgia Institute of Technology, USA) (Tieves et al. 2019; Willot et al. 2020). All constructs were introduced into *E. coli* BL21-Gold(DE3) (WT) and the catalase-deficient strain *E. coli* BL21-Gold(DE3) Δ*katE* Δ*katG* (CD). Heterologous expression and protein purification are described in the Supporting Information S1. GOx from *Aspergillus niger* was purchased from Sigma-Aldrich.

Concentrations of CYP peroxygenases were calculated using the usual extinction coefficient at 450 nm of 91 $mM^{-1}cm^{-1}$ from the respective CO-difference spectra recorded using a Spectra-max M2 Microtiter Plate Reader (Molecular Devices Corporation) (Omura and Sato 1964, Guengerich et al. 2009). For *Dca*UPO, concentrations of purified protein were determined using the Pierce BCA assay kit (ThermoFisher Scientific), with bovine serum albumin as a standard. Purified *Dca*UPO was then used to calculate a concentration factor to determine the final concentration of enzyme in the CFEs and WCs from CO-difference spectra (at 450 nm).

3.3 | H_2O_2 Measurement

H_2O_2 stability was tested in WCs and CFEs from wild-type and catalase-deficient *E. coli* (0.1 g wet weight mL^{-1}) in 200 mM potassium phosphate buffer pH 7.0. Reactions containing 800 μL of WCs or CFEs and 10 mM of H_2O_2 were incubated at 25°C. Samples were taken at different time points and H_2O_2 concentrations were quantified with the Ampliflu Red assay (Sigma-Aldrich) (560 nm $\varepsilon = 71\,000\ M^{-1}\ cm^{-1}$) using the Spectramax M2 spectrophotometer.

3.4 | Biotransformations

Biotransformations were performed in 4 mL glass vials, in a final reaction volume of 1 mL. Reactions were incubated at 25°C with shaking at 200 rpm.

Reaction mixtures consisted of 800 μL of WCs or CFEs from wild-type and catalase-deficient *E. coli* (0.1 g wet weight mL^{-1}) in 200 mM potassium phosphate buffer pH 7.0, 10 mM substrate, 5% (v/v) acetone, 20 mM H_2O_2 or 0.2 U GOx or FOx, 100 mM glucose or sodium formate. In reactions with purified proteins WCs or CFEs were replaced with 800 μL 200 mM potassium phosphate buffer pH 7.0 containing purified peroxygenases at similar concentrations.

Reactions were stopped and extracted by addition of 1 mL ethyl acetate containing 2 mM internal standard (1-undecanol). Samples were analyzed by GC-FID (Shimadzu GC-2010) and GC-MS (Thermo Scientific TraceGC ultra—Trace DSQ) (Figures S2–S6) using a FactorFour VF-5ms column (60 m × 0.32 mm × 0.25 μm, Varian) column. Temperature program: 100°C hold 1 min, then 8°C min^{-1} up to 200°C. Concentrations of products and remaining substrates were calculated from standard curves of commercial standards and corrected for low levels of H_2O_2 oxidation in the absence of enzyme detected with cells transformed with empty plasmid. Chiral analysis of extracts from thioanisole reactions was performed by GC-FID (Thermo Scientific TraceGC ultra) using a CHIRALDEX B-TA column (30 m × 0.25 mm × 0.12 μm). Temperature program: 100°C hold 1 min then 1.5°C min^{-1} up to 136°C.

Averages and standard deviations were calculated from at least three independent reactions. Standard deviations of TONs were calculated using standard deviations of the product concentrations and the catalyst concentrations.

3.5 | Screening Experiment

Reactions for the screening experiment were performed as described above using WCs of catalase-deficient *E. coli* (0.1 g wet weight mL^{-1}) containing no CYP, *Ssca*CYP (4 μM), *Ssca*CYP_E284A (8 μM) and *Ssca*CYP_E284I (8 μM). After 4 h 10 μL aliquots from each reaction mixture were transferred in duplicate to a 96-well microtiter plate and diluted 200 times in 200 mM potassium phosphate buffer pH 7.0. Remaining H_2O_2 levels were measured at 560 nm using the Ampliflu Red assay. The remaining reaction mixtures were extracted as described above and analyzed with GC-MS/FID (Thermo Scientific TraceGC ultra—Trace DSQ) (Figures S9 and S10) using the FactorFour VF-5ms column (60 m × 0.32 mm × 0.25 μm, Varian) column.

3.6 | Preparative Scale Reaction

A WC suspension (50 mL) of catalase-deficient *E. coli* (0.2 g wet weight mL^{-1} in 200 mM potassium phosphate buffer pH 7.0) containing *Ssca*CYP E284I (ca. 20 μM) was supplemented with *trans*-β-methyl styrene (10 mM, 5% v/v acetone). The reaction was started by the addition of H_2O_2 (10 mM) and then incubated at 25°C with shaking at 200 rpm. Samples were taken every 4 h up to 12 h and then at 24 h. Enzyme stability (CO-difference spectra), H_2O_2 accumulation (quantification by Ampliflu Red assay), and substrate conversion (GC-MS/FID) (Figure S10) were evaluated, followed by the addition of H_2O_2 (10 mM). After 28 h enzyme, stability and H_2O_2 accumulation were evaluated and then the total reaction mixture was extracted with an equal volume of ethyl acetate.

Author Contributions

Martha Smit, Diederik J. Opperman, and Vlada B. Urlacher conceptualized the study and contributed to the design of experiments. Joost Luelf created the catalase-deficient *E. coli* strain. Ana C. Ebrecht designed and conducted most of the biotransformation experiments and analyzed the data. Kamini Govender performed the first biotransformation experiment with the catalase-deficient strain. Martha S. Smit prepared the first draft, and all authors then edited the manuscript. All authors approved the manuscript.

Acknowledgments

This work was supported by the South African Council for Scientific and Industrial Research—Industrial Biocatalysis Hub (CSIR-IBH) initiative funded by the South African Department of Science and Innovation (DSI) and the Technology Innovation Agency (TIA). We also thank Sarel Marais for technical assistance with GC analyses.

Ethics Statement

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Aschenbrenner, J. C., A. C. Ebrecht, M. S. Smit, and D. J. Opperman. 2024. "Revisiting Strategies and Their Combinatorial Effect for Introducing Peroxygenase Activity in CYP102A1 (P450BM3)." *Molecular Catalysis* 557, no. February: 113953. https://doi.org/10.1016/j.mcat.2024.113953.

Baba, T., T. Ara, M. Hasegawa, et al. 2006. "Construction of *Escherichia coli* K-12 in-Frame, Single-Gene Knockout Mutants: The Keio Collection." *Molecular Systems Biology* 2: 2006.0008. https://doi.org/10.1038/msb4100050.

Chen, H. Y., Z. J. Yang, X. T. Zhou, and H. B. Ji. 2012. "Oxidative Cleavage of C=C Bond of Cinnamaldehyde to Benzaldehyde in the Presence of β-Cyclodextrin Under Mild Conditions." *Supramolecular Chemistry* 24, no. 4: 247–254. https://doi.org/10.1080/10610278.2012.655277.

Cirino, P. C., and F. H. Arnold. 2002. "Regioselectivity and Activity of Cytochrome P450 BM-3 and Mutant F87A in Reactions Driven by Hydrogen Peroxide." *Advanced Synthesis & Catalysis* 344, no. 9: 932–937. https://doi.org/10.1002/1615-4169(200210)344:9<932::AID-ADSC932>3.0.CO;2-M.

Cirino, P. C., and F. H. Arnold. 2003. "A Self-Sufficient Peroxide-Driven Hydroxylation Biocatalyst." *Angewandte Chemie International Edition* 42, no. 28: 3299–3301. https://doi.org/10.1002/anie.200351434.

Ebrecht, A. C., T. M. Mofokeng, F. Hollmann, M. S. Smit, and D. J. Opperman. 2023. "Lactones From Unspecific Peroxygenase-Catalyzed in-Chain Hydroxylation of Saturated Fatty Acids." *Organic Letters* 25, no. 27: 4990–4995. https://doi.org/10.1021/acs.orglett.3c01601.

Ebrecht, A. C., M. S. Smit, and D. J. Opperman. 2023. "Natural Alternative Heme-Environments Allow Efficient Peroxygenase Activity by Cytochrome P450 Monooxygenases." *Catalysis Science & Technology* 13, no. 21: 6264–6273. https://doi.org/10.1039/d3cy01207g.

Guengerich, F. P., M. V. Martin, C. D. Sohl, and Q. Cheng. 2009. "Measurement of Cytochrome P450 and NADPH-Cytochrome P450 Reductase." *Nature Protocols* 4, no. 9: 1245–1251. https://doi.org/10.1038/nprot.2009.121.

Hui, D., S. Khaiat, T. Uy, and H. Xu. 2014. "Partial Confirmation of Single *katG* and *katE* Knockouts and Double *katG*/*katE* Knockouts Created From Isogenic Background of *Escherichia coli* K-12 Strains." *Journal of Experimental Microbiology and Immunology (JEMI)* 18, no. April: 139–145.

Kinner, A., K. Rosenthal, and S. Lütz. 2021. "Identification and Expression of New Unspecific Peroxygenases—Recent Advances, Challenges and Opportunities." *Frontiers in Bioengineering and Biotechnology* 9: 705630. https://doi.org/10.3389/fbioe.2021.705630.

Labun, K., T. G. Montague, M. Krause, Y. N. Torres Cleuren, H. Tjeldnes, and E. Valen. 2019. "CHOPCHOP v3: Expanding the CRISPR Web Toolbox Beyond Genome Editing." *Nucleic Acids Research* 47, no. W1: W171–W174. https://doi.org/10.1093/nar/gkz365.

Li, Q., B. Sun, J. Chen, Y. Zhang, Y. Jiang, and S. Yang. 2021. "A Modified pCas/pTargetF System for CRISPR-Cas9-Assisted Genome Editing in *Escherichia coli*." *Acta Biochimica et Biophysica Sinica* 53, no. 5: 620–627. https://doi.org/10.1093/abbs/gmab036.

Linde, D., A. Olmedo, A. González-Benjumea, et al. 2020. "Two New Unspecific Peroxygenases From Heterologous Expression of Fungal Genes in *Escherichia coli*." *Applied and Environmental Microbiology* 86, no. 7: e02899-19. https://doi.org/10.1128/AEM.02899-19.

Liu, F., R. Min, J. Hong, G. Cheng, Y. Zhang, and Y. Deng. 2021. "Quantitative Proteomic Analysis of ahpC/F and *katE* and *katG* Knockout *Escherichia coli*—A Useful Model to Study Endogenous Oxidative Stress." *Applied Microbiology and Biotechnology* 105, no. 6: 2399–2410. https://doi.org/10.1007/s00253-021-11169-2.

Luelf, U. J., L. M. Böhmer, S. Li, and V. B. Urlacher. 2023. "Effect of Chromosomal Integration on Catalytic Performance of a Multi-Component P450 System in *Escherichia coli*." *Biotechnology and Bioengineering* 120, no. 7: 1762–1772. https://doi.org/10.1002/bit.28404.

Monterrey, D. T., A. Menés-Rubio, M. Keser, D. Gonzalez-Perez, and M. Alcalde. 2023. "Unspecific Peroxygenases: The Pot of Gold at the End of the Oxyfunctionalization Rainbow?" *Current Opinion in Green and Sustainable Chemistry* 41: 100786. https://doi.org/10.1016/j.cogsc.2023.100786.

Nakagawa, S., S. Ishino, and S. Teshiba. 1996. "Construction of Catalase Deficient *Escherichia coli* Strains for the Production of Uricase." *Bioscience, Biotechnology, and Biochemistry* 60, no. 3: 415–420. https://doi.org/10.1271/bbb.60.415.

Omura, T., and R. Sato. 1964. "The Carbon Monoxide-Binding Pigment of Liver Microsomes." *Journal of Biological Chemistry* 239, no. 7: 2370–2378.

Qi, L. S., M. H. Larson, L. A. Gilbert, et al. 2013. "Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression." *Cell* 152, no. 5: 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022.

Salazar, O., P. C. Cirino, and F. H. Arnold. 2003. "Thermostabilization of a Cytochrome P450 Peroxygenase." *ChemBioChem* 4, no. 9: 891–893. https://doi.org/10.1002/cbic.200300660.

Tama, A., G. Bartosz, and I. Sadowska-Bartosz. 2023. "Phenolic Compounds Interfere in the Ampliflu Red/Peroxidase Assay for Hydrogen Peroxide." *Food Chemistry* 422, no. April: 136222. https://doi.org/10.1016/j.foodchem.2023.136222.

Tieves, F., S. J. P. Willot, M. M. C. H. van Schie, et al. 2019. "Formate Oxidase (FOx) From *Aspergillus oryzae*: One Catalyst Enables Diverse H_2O_2-Dependent Biocatalytic Oxidation Reactions." *Angewandte Chemie International Edition* 58, no. 23: 7873–7877. https://doi.org/10.1002/anie.201902380.

Wei, Q. Y., J. J. Xiong, H. Jiang, C. Zhang, and Ye Wen Ye. 2011. "The Antimicrobial Activities of the Cinnamaldehyde Adducts With Amino

Acids." *International Journal of Food Microbiology* 150, no. 2–3: 164–170. https://doi.org/10.1016/j.ijfoodmicro.2011.07.034.

Willot, S. J. P., M. D. Hoang, C. E. Paul, et al. 2020. "FOx News: Towards Methanol-Driven Biocatalytic Oxyfunctionalisation Reactions." *ChemCatChem* 12, no. 10: 2713–2716. https://doi.org/10.1002/cctc.202000197.

Xu, H., W. Liang, L. Ning, et al. 2020. "Directed Evolution of P450 Fatty Acid Decarboxylases Via High-Throughput Screening Towards Improved Catalytic Activity." *ChemCatChem* 12, no. 1: 80–84. https://doi.org/10.1002/cctc.201901347.

Xu, X., T. Hilberath, and F. Hollmann. 2023. "Selective Oxyfunctionalisation Reactions Catalysed by P450 Monooxygenases and Peroxygenases—A Bright Future for Sustainable Chemical Synthesis." *Current Opinion in Green and Sustainable Chemistry* 39: 100745. https://doi.org/10.1016/j.cogsc.2022.100745.

Yan, X., X. Zhang, H. Li, et al. 2024. "Engineering of Unspecific Peroxygenases Using a Superfolder-Green-Fluorescent-Protein-Mediated Secretion System in *Escherichia coli*." *JACS Au* 4, no. 4: 1654–1663. https://doi.org/10.1021/jacsau.4c00129.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

2.1.1 Supporting information

Supporting Information

Use of whole cells and cell free extracts of catalase deficient *E. coli* for peroxygenase catalyzed reactions

Ana C. Ebrecht[1], U. Joost Luelf[2], Kamini Govender[1], Diederik J. Opperman[1], Vlada B. Urlacher[2] and Martha S. Smit[1*]

[1] Department of Microbiology and Biochemistry, University of the Free State, PO Box 339, Bloemfontein 9300, South Africa
[2] Institute of Biochemistry, Heinrich Heine University Düsseldorf, 40225, Düsseldorf, Germany

[*]Corresponding Author

Email: smitms@ufs.ac.za

ORCID: Martha Smit 0000-0002-1295-6507

CRISPR/Cas-assisted λ-Red Recombineering

To prepare electrocompetent cells for CRISPR/Cas-assisted λ-Red Recombineering, 5 ml 2xYT medium were inoculated with 100 µl of an overnight culture of *E. coli* BL21-Gold(DE3) harboring the pEcCas plasmid. For induction of the λ-Red genes, 100 µL 1M L-arabinose was added and the cultures were incubated at 37°C, 180 revolutions per minute (rpm) for 2 h. The cells were harvested by centrifugation and washed twice with 1 ml ice-cold 10%(vol/vol) glycerol. After a final centrifugation step, the cells were resuspended in 100 µl ice-cold 10%(vol/vol) glycerol. The pgRNA-bacteria plasmid (100 ng) and the repair template (500 ng) were added, and the suspension was transferred to a 2 mm gap electroporation cuvette. Electroporation was carried out by applying 2.5 kV (MicroPulser electroporator from Bio-Rad). Immediately after electroporation, 1 ml of SOC medium was added and the cells recovered at 37 °C, 180 rpm for 1 h. Finally, the cells were plated on LB agar plates (30 µg/ml kanamycin, 100 µg/ml ampicillin) and were grown overnight at 37 °C.

Heterologous expression and protein purification

Heterologous expression was performed in auto-induction media (ZYP-5052) (Studier, 2005) supplemented with 0.5 mM δ-aminolevulinic acid hydrochloride and 50 µM $FeCl_3 \cdot 6H_2O$ (20 °C for 48 h, 200 rpm, for CYP102A1_21B3 and SscaCYP_E284A, and 16 °C, for 5 days, 200 rpm, for *Dca*UPO). Cells were harvested by centrifugation (5000 x *g*, 10 min, 4 °C) and resuspended (0.1 g wet weight mL^{-1}) in 200 mM potassium phosphate buffer pH 7.0. To obtain CFEs, cells were disrupted by single passage through a One-Shot Cell disrupter System (Constant Systems Ltd) at 30 kPsi, followed by centrifugation (20000 x *g,* 20 min, 4°C). For purification, cells were harvested by centrifugation (5000 x *g*, 10 min, 4°C) and resuspended (0.2 g wet weight mL^{-1}) in buffer A (25 mM Tris-HCl pH 8.0). Disruption of the cells was carried out by single passage through a One-Shot Cell disrupter System (Constant Systems Ltd) at 30 kPsi, followed by centrifugation (30 000 x *g*, 45 min, 4°C). The resulting soluble fraction was loaded onto a 1 mL His GraviTrap™ column (Cytiva), previously equilibrated with buffer B (25 mM Tris-HCl, 300 mM NaCl, 40 mM imidazole, pH 8.0). The loaded column was washed with 5 column volumes of buffer B. Protein was eluted in buffer E (25 mM Tris-HCl, 300 mM NaCl, 300 mM imidazole, pH 8.0), and desalted using PD-10 columns (GE Healthcare) equilibrated with buffer D (10 mM potassium phosphate buffer, 150 mM NaCl, pH 7.4). For *Dca*UPO, the disrupted cells were incubated with 1% (w/v) Triton-X100 for 1 h before clarification step (30 000 x g, 45 min, 4 °C), followed by purification with GraviTrap™ columns.

column. Temperature program: 100 ºC hold 1 min then 8 °C min^{-1} up to 200 °C.

Table S1. List of primers for in-frame deletion of *katE* in *E. coli* BL21-Gold(DE3)

Description	Sequence
forward primer for amplification of homology arm 1	CAGAATTTCAGGTCATGTAAC
reverse primer for amplification of homology arm 1 (5' extension for fusion with homology arm 2 underlined)	CAGGCAGGAATTTTGTCAATCATTGAACTCGTCTCCTTA
forward primer for amplification of homology arm 2 (5' extension for fusion with homology arm 2 underlined)	TAAGGAGACGAGTTCAATGATTGACAAAATTCCTGCCTG
reverse primer for amplification of homology arm 2	CGGTGAAGAGATCAGTGAG
forward primer for exchange of targeting sequence (5' extension for digestion underlined, targeting sequence in bold)	ATATATACTAGTCCGAATGGGTGAAGTGACTG **GTTTTAGAGCTAGAAATAGC**
reverse primer for exchange of targeting sequence (5' extension for digestion underlined)	ATATATACTAGTATTATACCTAGGACTGAGCTAG
sequencing primer 1	GAGCGACGCCACAGGATG
sequencing primer 2	TTCCCGATTGTTGCCAGGTTTG

Table S2. List of primers for in-frame deletion of *katG* in *E. coli* BL21-Gold(DE3)

Description	Sequence
forward primer for amplification of homology arm 1	TTCGACGGTCTGGATGATG
reverse primer for amplification of homology arm 1 (5' extension for fusion with homology arm 2 underlined)	CAGCAGGTCGAAACGGTCCATCAATGTGCTCCCCTC
forward primer for amplification of homology arm 2 (5' extension for fusion with homology arm 2 underlined)	GAGGGGAGCACATTGATGGACCGTTTCGACCTGCTG
reverse primer for amplification of homology arm 2	CGAAATATTGCCCACGTCG
forward primer for exchange of targeting sequence designed as described by [6] (5' extension for digestion underlined, targeting sequence in bold)	ATATATACTAGTCATTGCTCAATACGCCAACG **GTTTTAGAGCTAGAAATAGC**

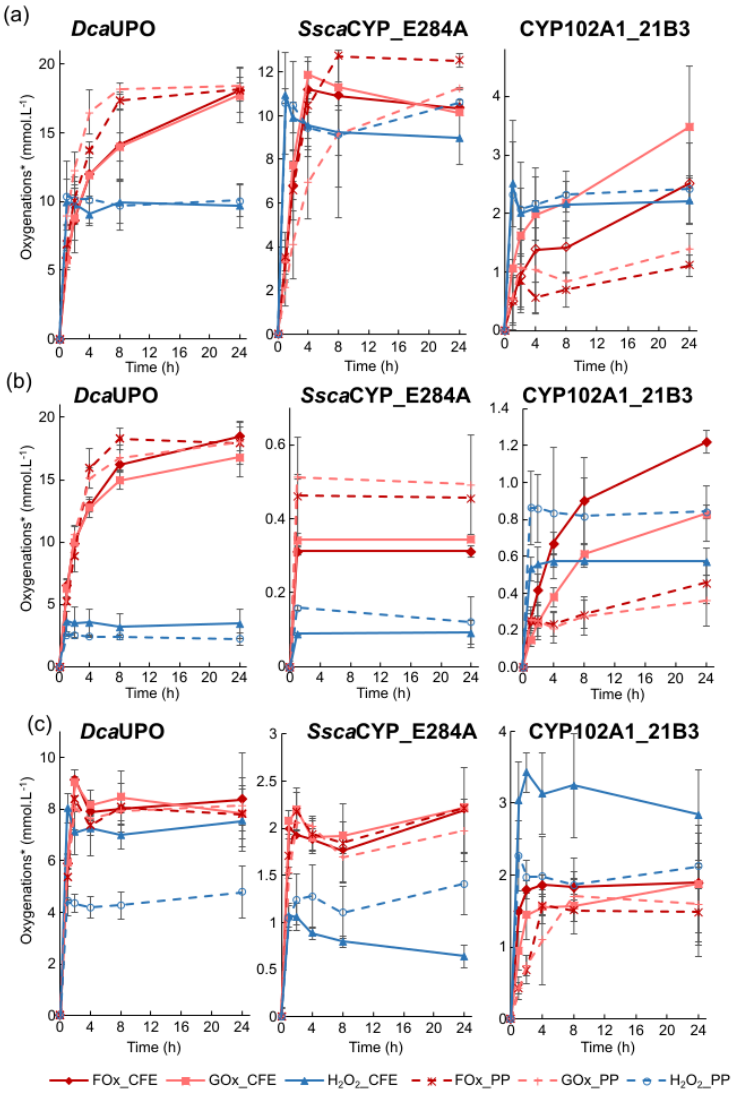


Figure S1. Comparison of activities towards (a) thioanisole, (b) ethylbenzene and (c) styrene of *Dca*UPO, *Ssca*CYP_E284A and CYP102A1_21B3 in CFEs of catalase-deficient *E. coli* BL21-Gold(DE3) and as the purified enzyme in buffer (PP), in time course experiments over 24 h. Oxygenations* represent the total number of peroxygenase reactions assuming that methyl phenyl sulfone, 1-phenyl ethanol and phenyl acetaldehyde are formed by two rounds of peroxygenase-catalyzed reactions. H_2O_2 was added in a single dose to start reactions or H_2O_2 was produced *in situ* by oxidation of glucose by GOx or by oxidation of formate by FOx. Reactions contained CFEs of catalase-deficient *E. coli* BL21-Gold(DE3) (0.1 g wet weight mL^{-1}) or a corresponding concentration of the purified enzyme (PP), substrate (10 mM), acetone (5% (v/v)), H_2O_2 (20 mM) or GOx (0.2 U mL^{-1}) with glucose (100 mM) or FOx (0.2 U mL^{-1}) with formate (100 mM).

Table S3. Times when reactions with the different substrates levelled off and total turnover numbers (TONs) at these times for *Dca*UPO, *Ssca*CYP_E284A, and CYP102A1_21B3 using purified proteins (PP) (17, 7, 4 μM peroxygenase) and cell free extracts (CFE) (18, 7, 4 μM peroxygenase). H_2O_2 (20 mM) was added in a single dose or produced *in situ* by oxidation of glucose (100 mM) with GOx or formate (100mM) with FOx.

	H_2O_2 (20 mM)				**GOx-glucose (100 mM)**				**FOx-formate (100 mM)**			
	PP		**CFE**		**PP**		**CFE**		**PP**		**CFE**	
	Time (h)	**TON ($x10^2$)**	**Time (h)**	**TON ($x10^2$)**	**Time (h)**	**TON ($x10^2$)**	**Time (h)**	**TON ($x10^2$)**	**Time (h)**	**TON ($x10^2$)**	**Time (h)**	**TON ($x10^2$)**
Thioanisole												
*Dca*UPO	1	6.1±1.5	1	6.5±1.2	8	10.6±0.7	24	11.5±1.7	8	10.2±0.7	24	11.8±1.5
*Ssca*CYP_E284A	1	15.5±4.5	1	16.5±1.8	24	16.5±3.2	4	18.0±2.2	8	18.7±4.0	4	17.0±2.0
CYP102A1_21B3	1	6.0±3.7	1	6.8±2.7	24	3.6±1.3	24	9.4±3.8	24	2.9±1.0	24	6.8±3.9
Ethylbenzene												
*Dca*UPO	1	1.5±0.1	1	2.4±3.5	8	9.8±0.9	24	10.9±1.4	8	10.7±0.8	24	12.0±1.3
*Ssca*CYP_E284A	1	0.2±0.1	1	0.1±0.1	1	0.8±0.1	1	0.5±0.1	1	0.7±0.1	1	0.6±0.1
CYP102A1_21B3	1	2.2±0.9	8	1.6±0.4	24	0.9±0.6	24	2.2±0.7	24	1.2±0.5	24	3.3±0.9
Styrene												
*Dca*UPO	1	2.6±0.4	1	5.2±0.6	2	4.7±0.3	2	5.9±0.6	2	4.9±0.8	2	5.9±0.5
*Ssca*CYP_E284A	1	2.1±0.6	1	1.6±0.3	1	3.0±0.6	1	3.2±0.4	1	3.2±0.7	1	3.3±0.8
CYP102A1_21B3	1	5.8±2.2	2	9.4±2.7	8	4.4±1.6	24	5.1±2.3	4	4.2±1.3	4	5.0±1.4

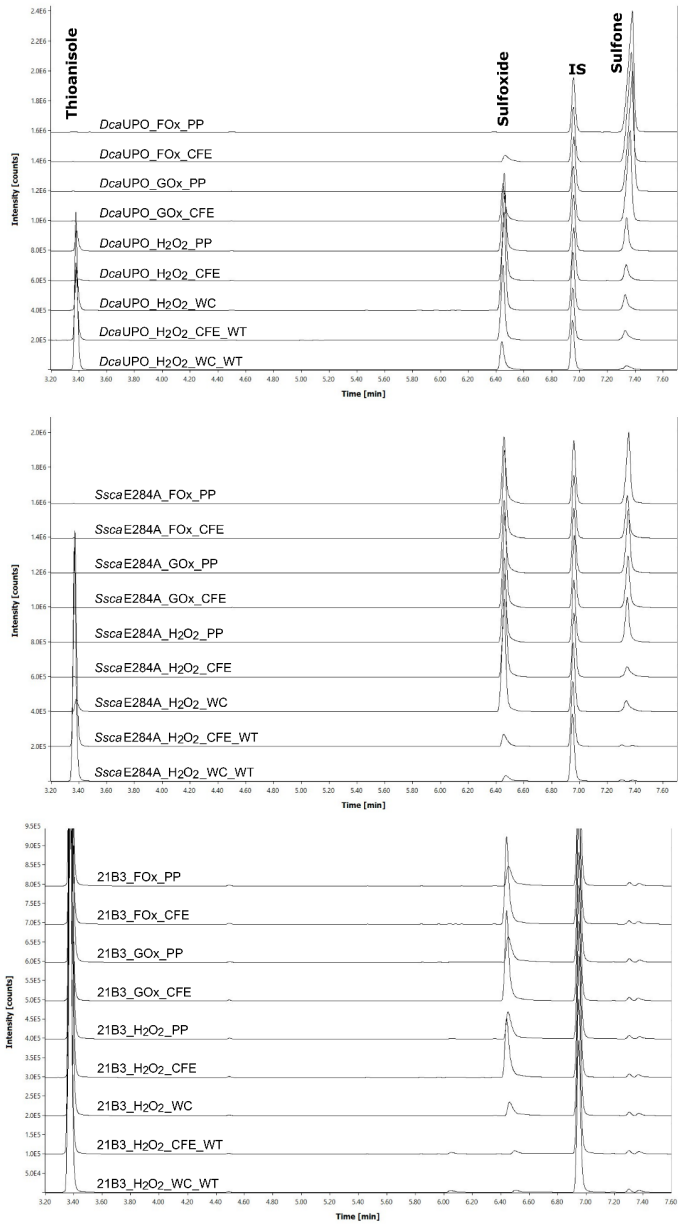


Figure S2. Analysis of products formed in biotransformations with thioanisole. Chromatograms for 24 h reactions with *Dca*UPO, *Ssca*CYP_E284A, and CYP102A1_21B3. PP: purified protein; CFE: catalase deficient cell-free extract, WC: catalase deficient whole cells, CFE_WT: cell-free extract from wild-type strain; WC_WT: whole cells from wild-type strain.

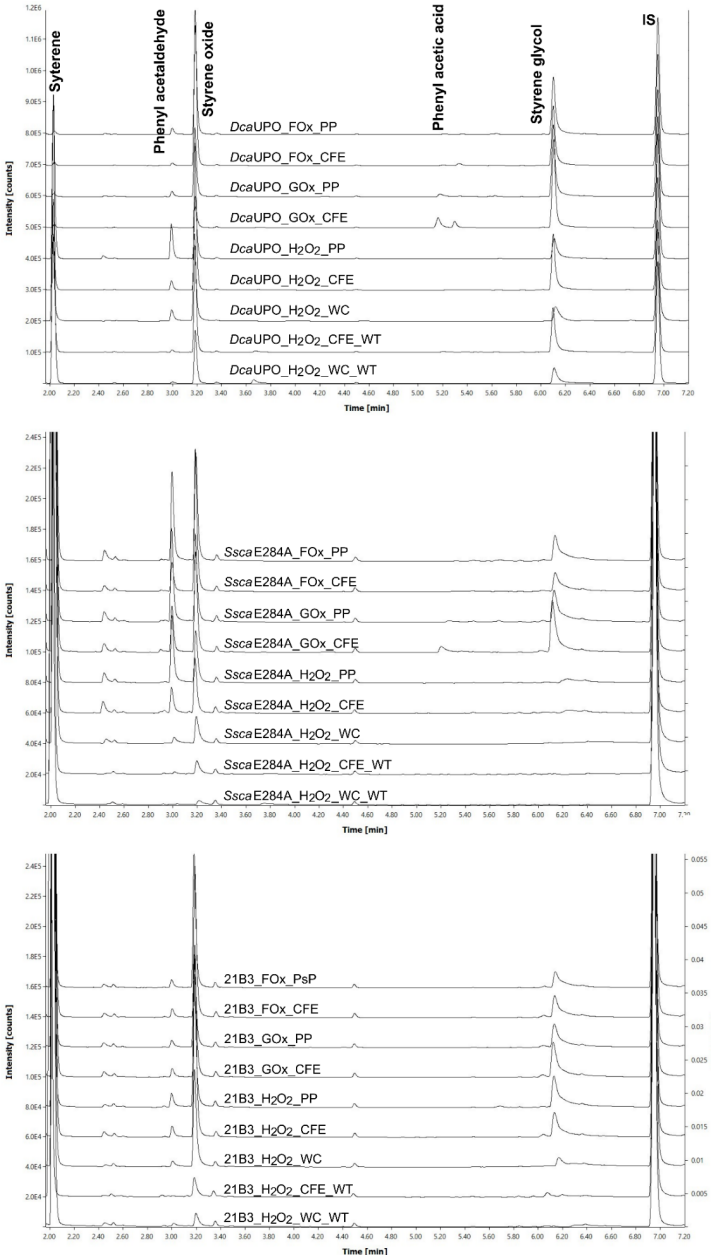


Figure S3. Analysis of products formed in biotransformations with styrene. Chromatograms for 24 h reactions with *Dca*UPO, *Ssca*CYP_E284A, and CYP102A1_21B3. PP: purified protein; CFE: catalase deficient cell-free extract, WC: catalase deficient whole cells, CFE_WT: cell-free extract from wild-type strain; WC_WT: whole cells from wild-type strain.

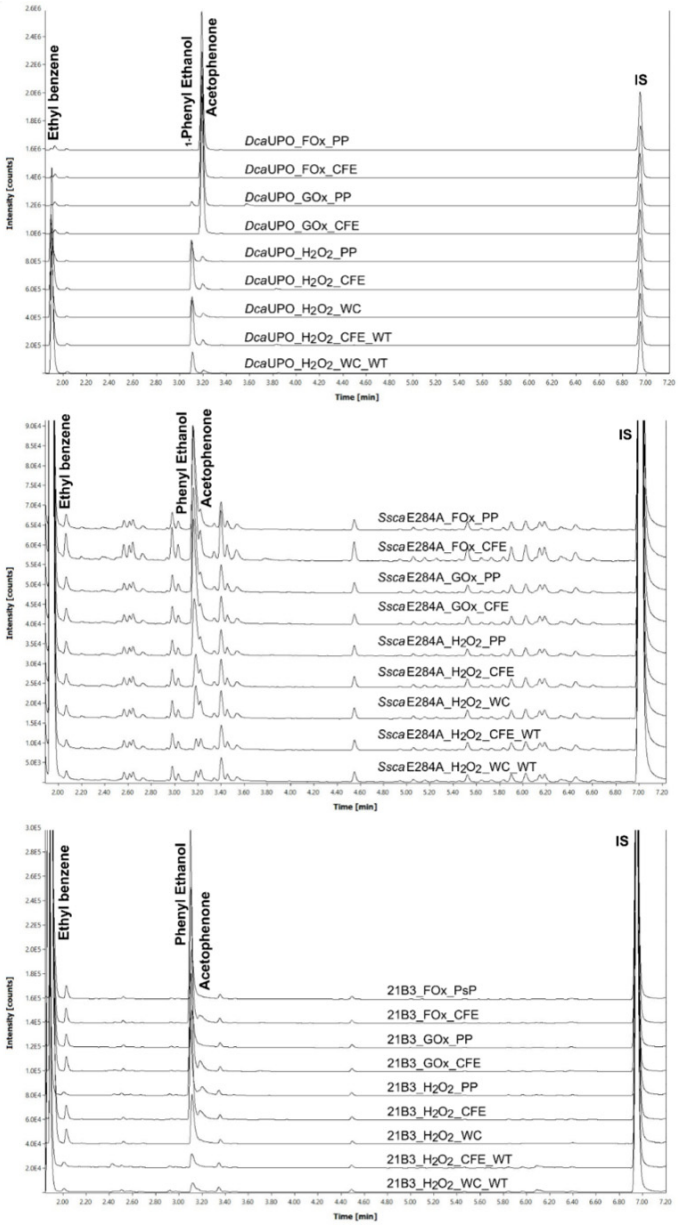


Figure S4. Analysis of products formed in biotransformations with ethyl benzene. Chromatograms for 24 h reactions with *Dca*UPO, *Ssca*CYP_E284A, and CYP102A1_21B3. PP: purified protein; CFE: catalase deficient cell-free extract, WC: catalase deficient whole cells, CFE_WT: cell-free extract from wild-type strain; WC_WT: whole cells from wild-type strain.

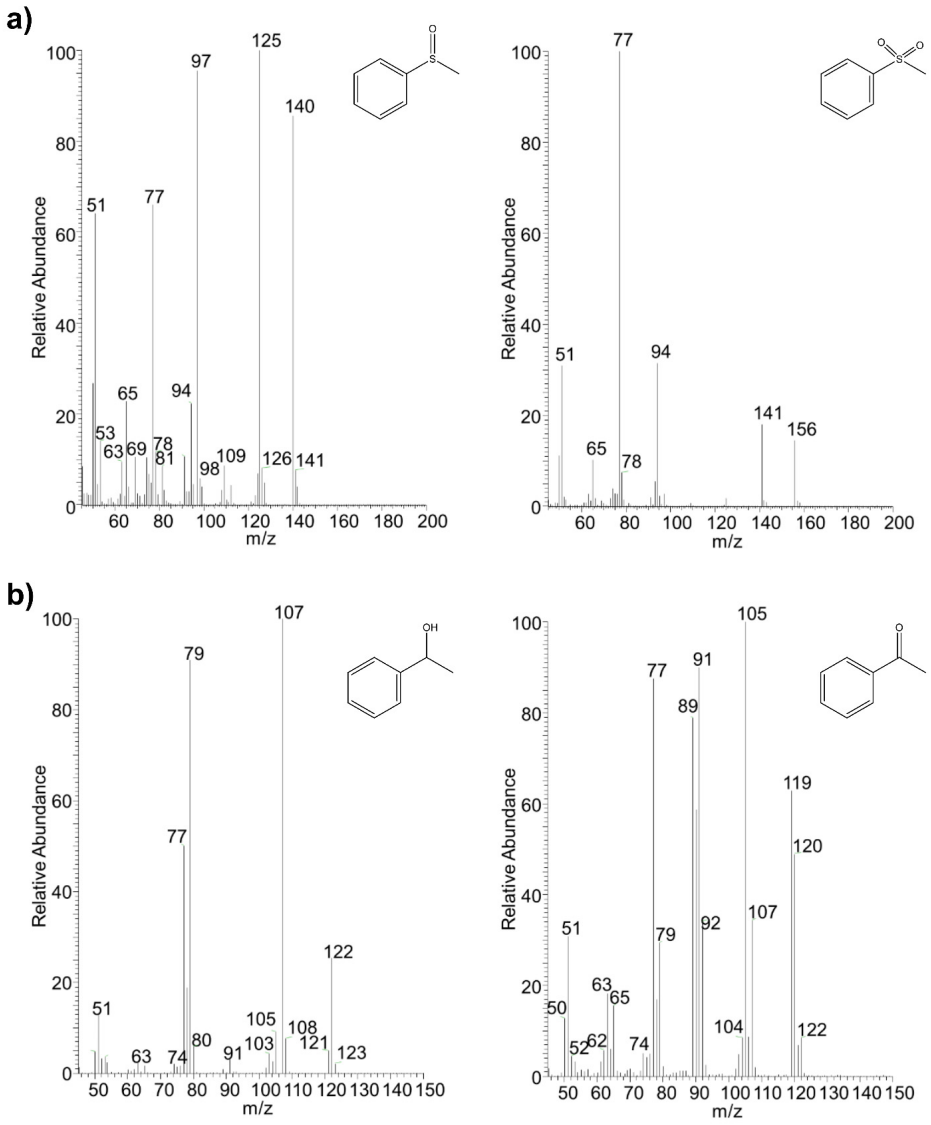


Figure S5. Mass spectra from the analysis of products from biotransformations with **a)** thioanisole, and **b)** ethyl benzene.

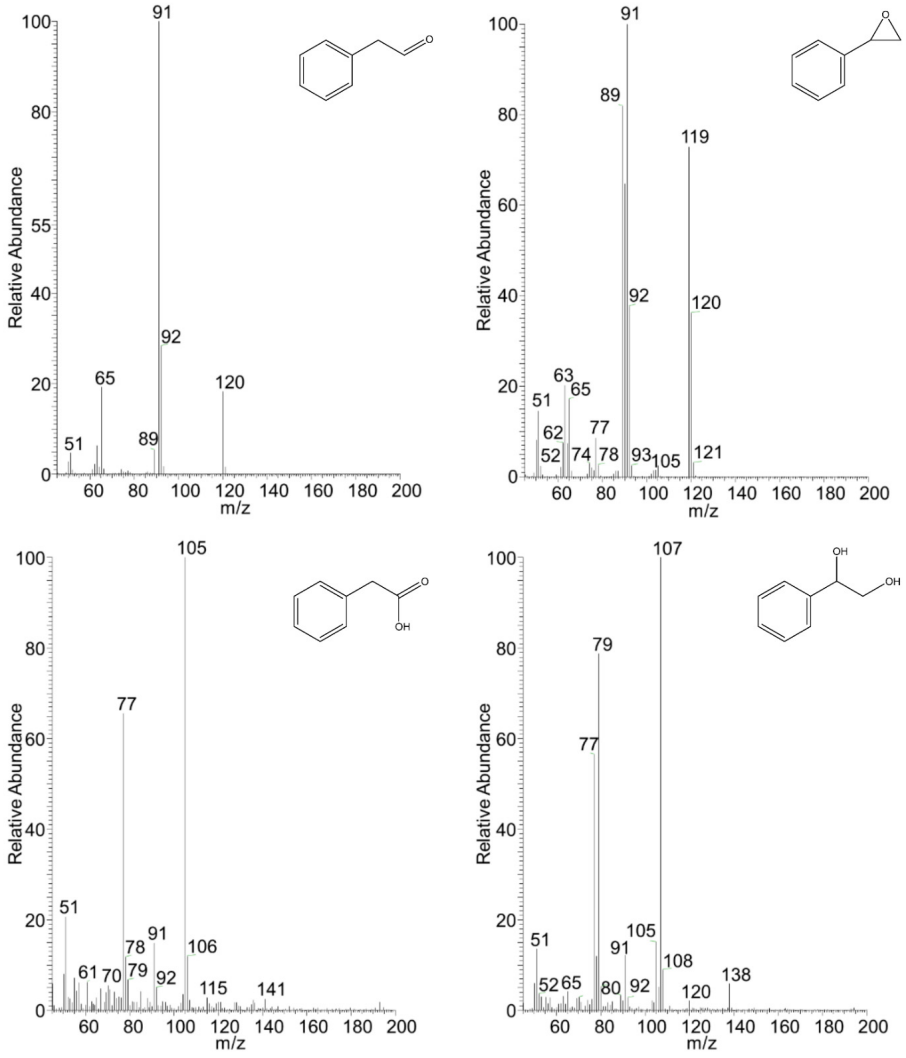


Figure S6. Mass spectra from the analysis of products from biotransformations with styrene.

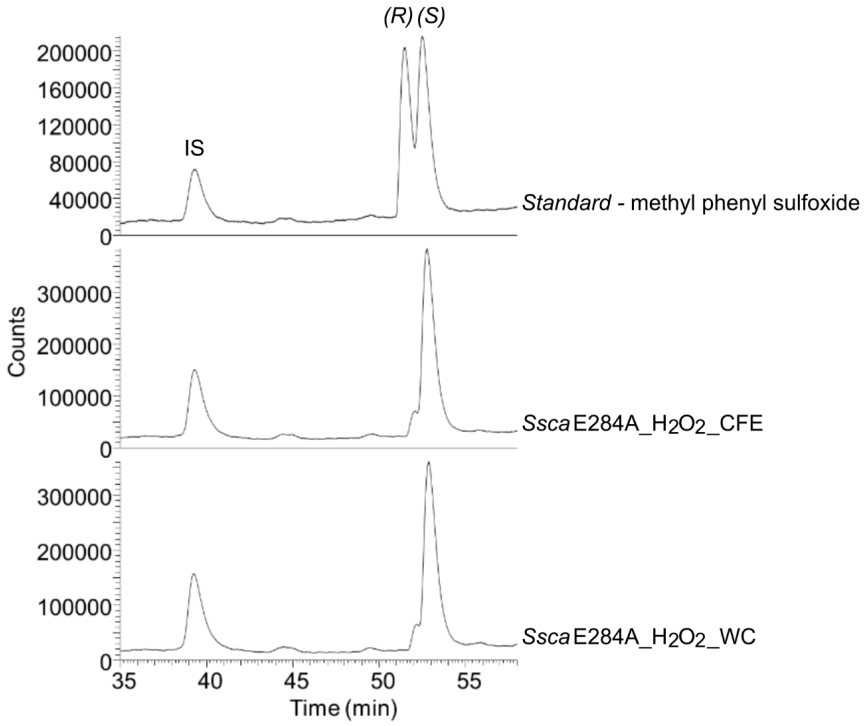


Figure S7. Chromatograms of chiral analysis of the sulfoxidation of thioanisole using catalase-deficient WCs and CFEs containing *Ssca*CYP_E284A. Enantiomeric excess with CFE and WCs is ca. 92 %.

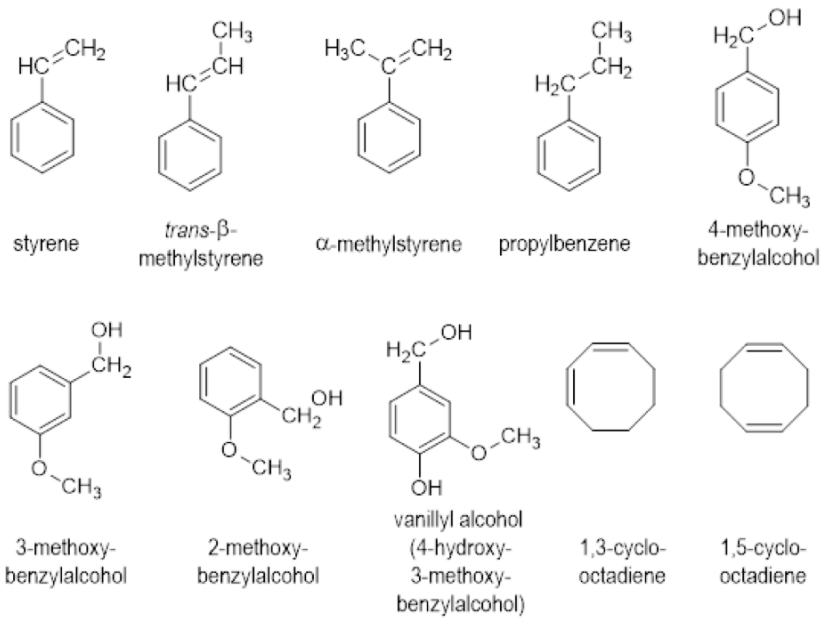


Figure S8. Ten substrates used to investigate whether H_2O_2 consumption could be used with WCs of catalase-deficient *E. coli* expressing *Ssca*CYP and its E284A and E284I mutants to screen for substrate specificity.

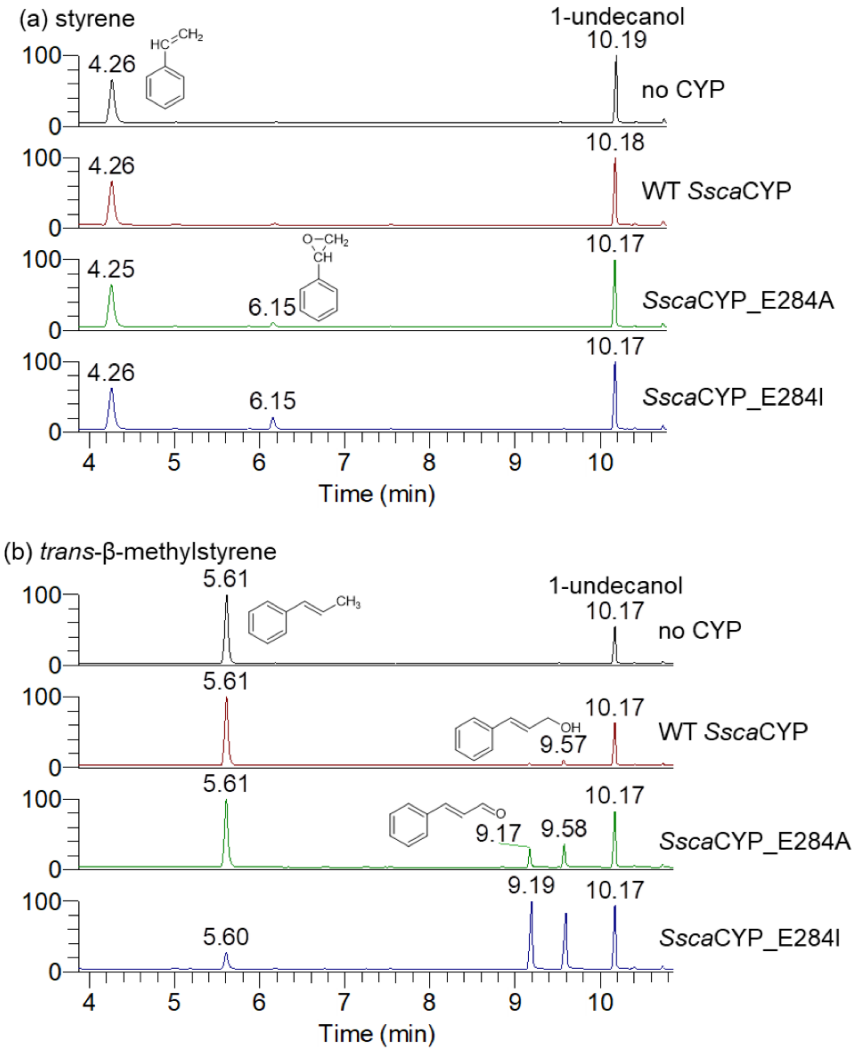


Figure S9. (a-j) Chromatograms of GC-FID analysis of extracts of substrate screen using WCs of catalase-deficient *E. coli* expressing no CYP, WT *SSca*CYP, *Ssca*CYP_E284A and *Ssca*CYP_E284I. Analysis of extracts obtained after 4 h biotransformations for (a) styrene, (b) *trans*-β-methylstyrene (c) α-methylstyrene, (d) propylbenzene, (e) 4-methoxybenzyl alcohol, (f) 3-methoxybenzyl alcohol , (g) 2-methoxybenzyl alcohol (h) vanillyl alcohol (i) 1,3-cyclooctadiene, (j) 1,5-cyclooctadiene.

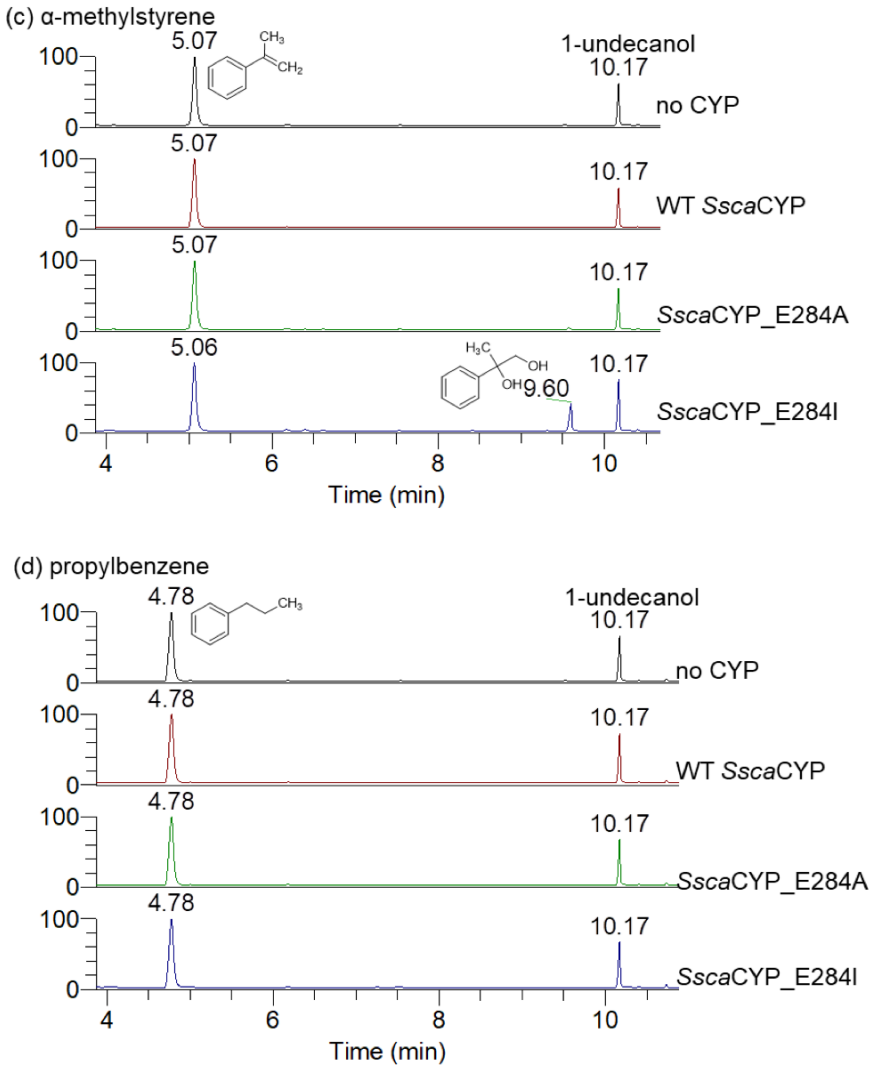


Figure S9 continued

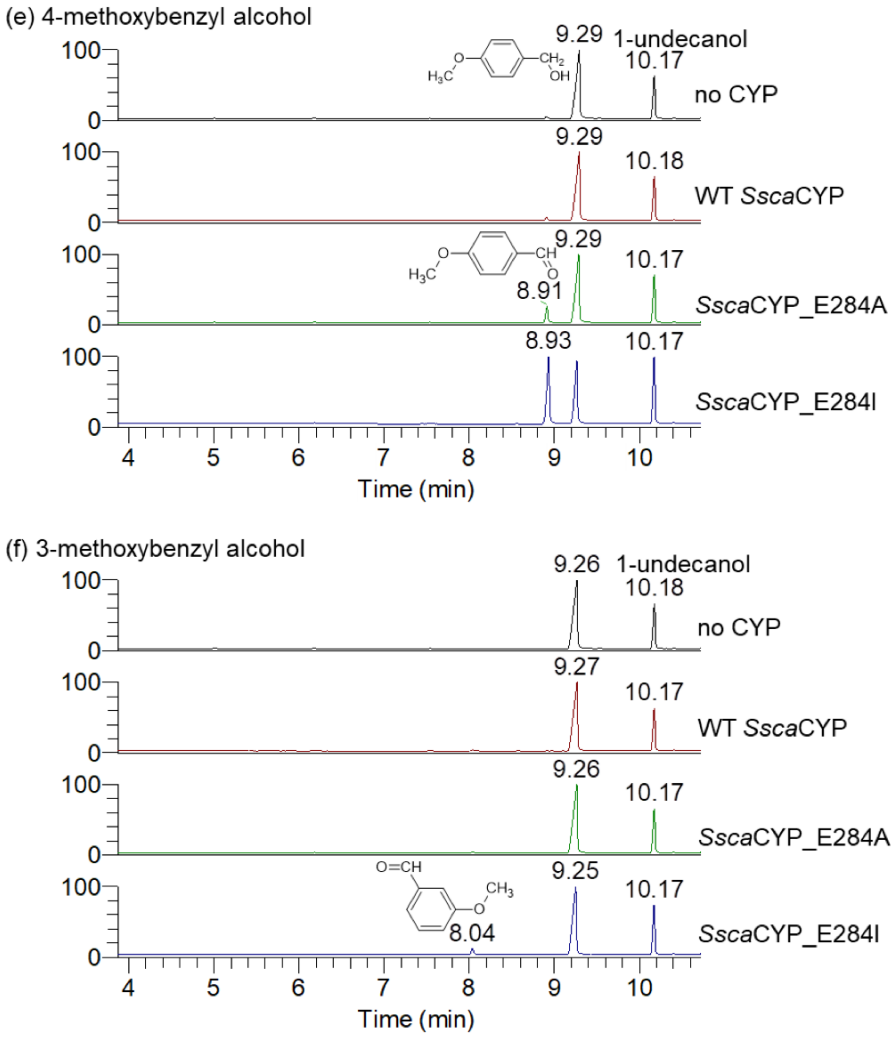


Figure S9 continued

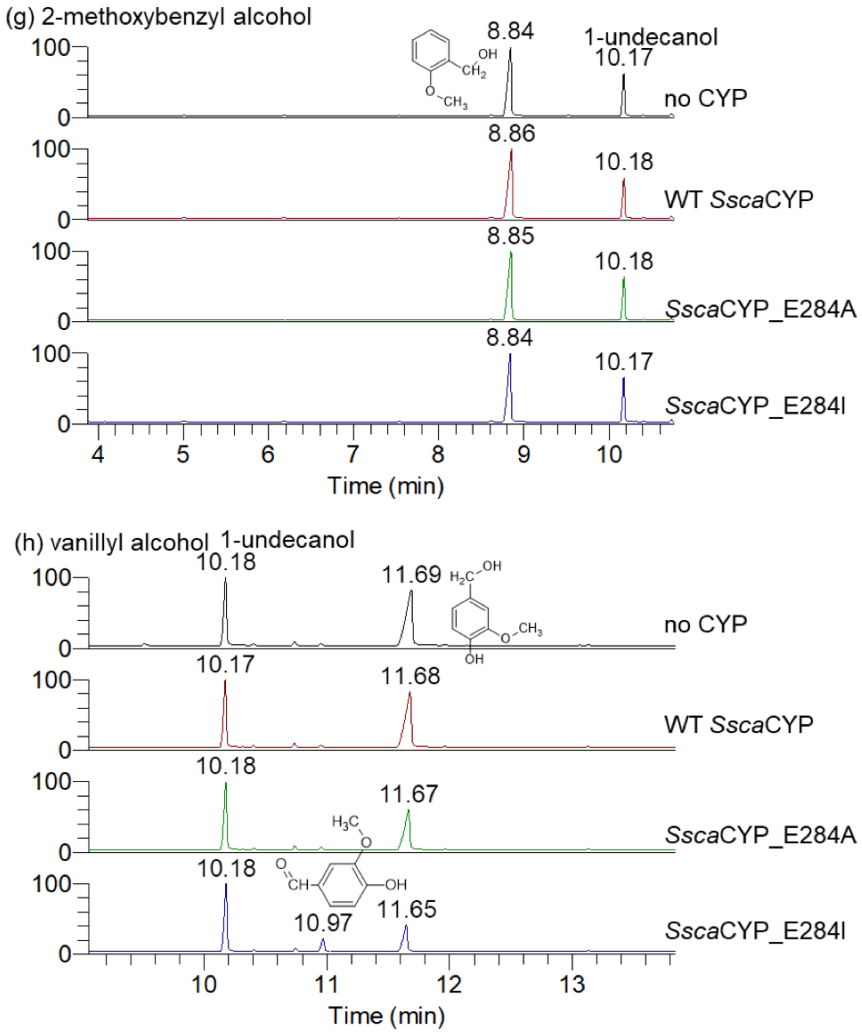


Figure S9 continued

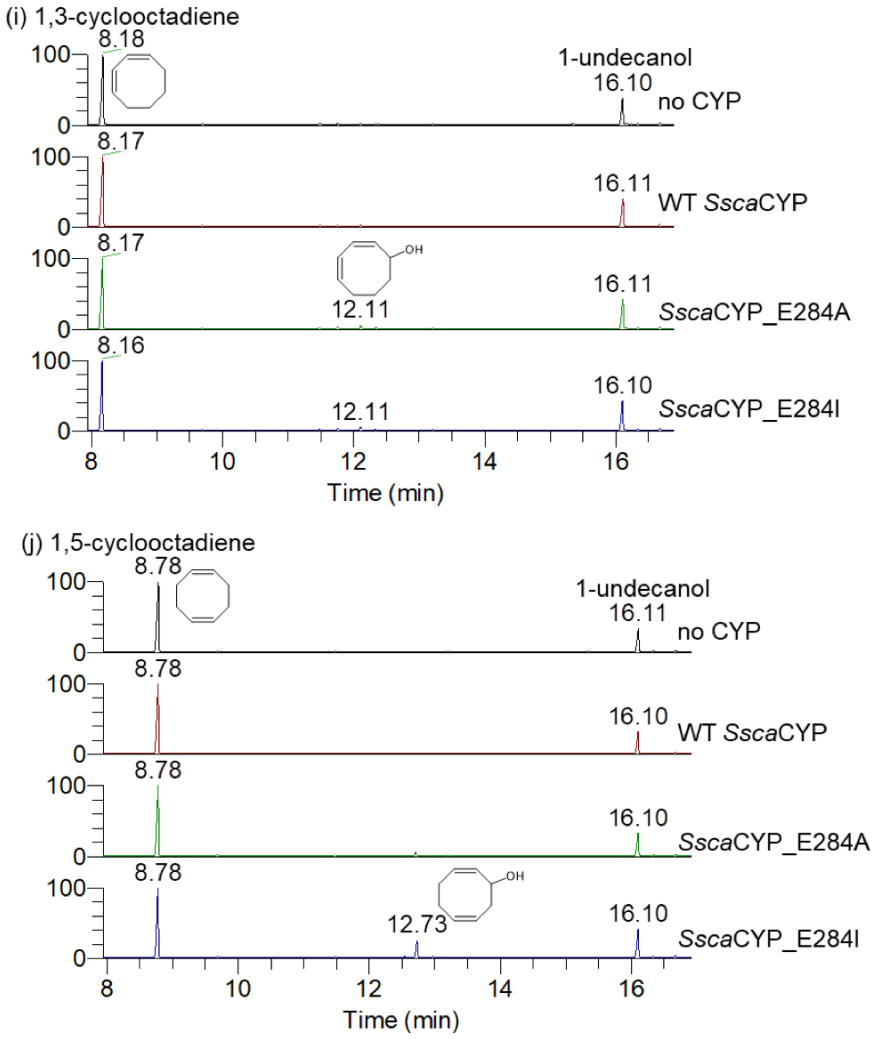


Figure S9 continued

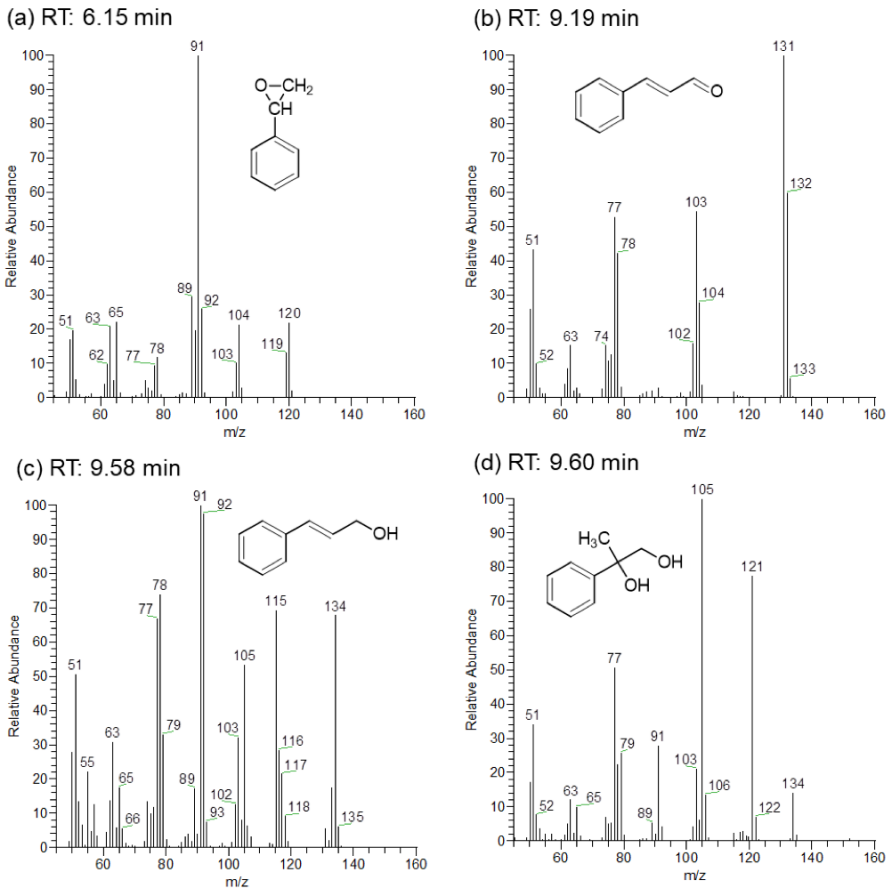


Figure S10 (a-j) Mass spectra of products identified in GC-MS analysis of extracts of substrate screen using WCs of catalase-deficient *E. coli* expressing no CYP, WT *SSca*CYP, *Ssca*CYP_E284A and *Ssca*CYP_E284I. Products are from (a) styrene, (b & c) *trans*-β-methylstyrene (d) α-methylstyrene, (e) 4-methoxybenzyl alcohol, (f) 3-methoxybenzyl alcohol , (g) 2-methoxybenzyl alcohol (h) vanillyl alcohol (i) 1,3-cyclooctadiene, (j) 1,5-cyclooctadiene.

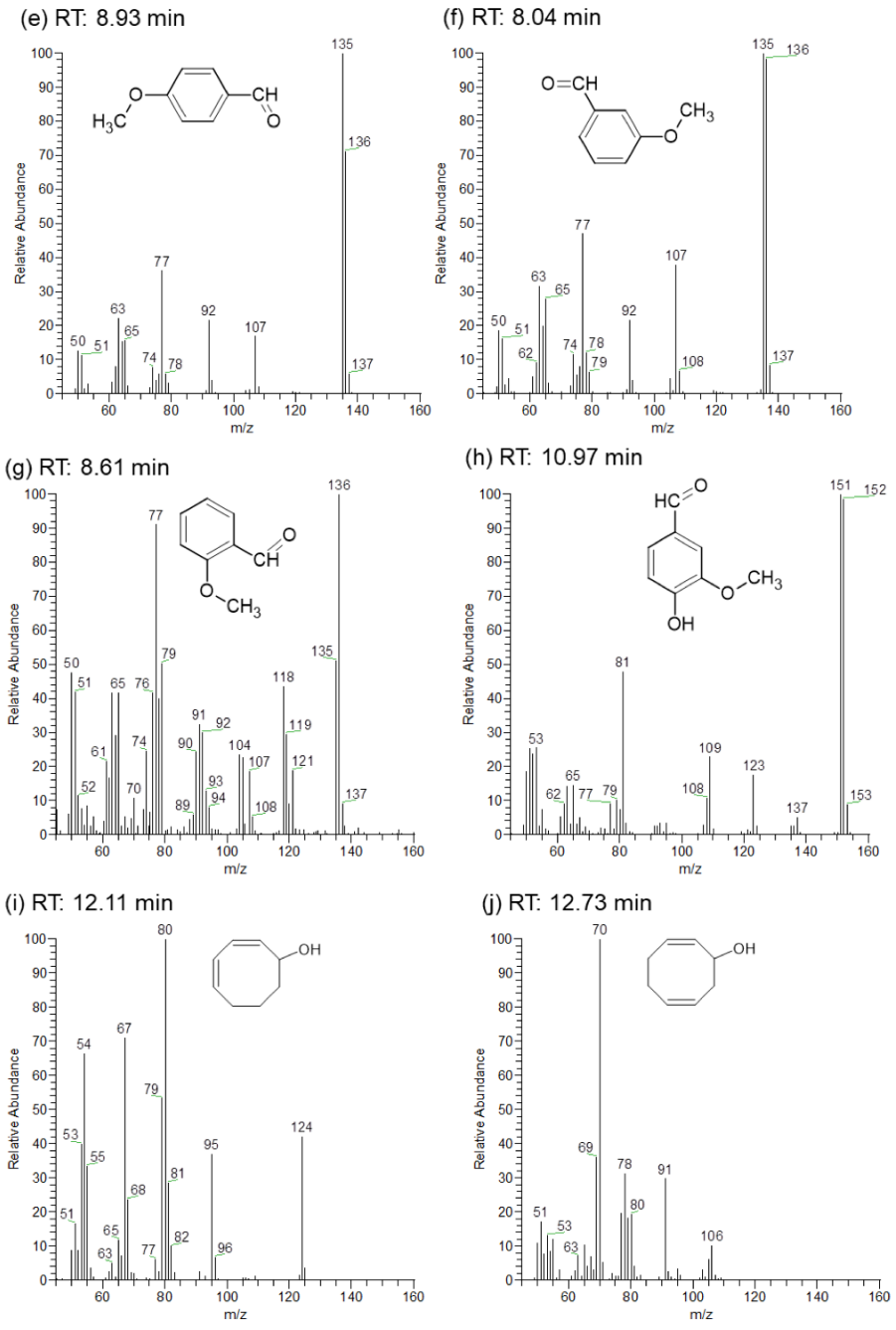


Figure S10 continued

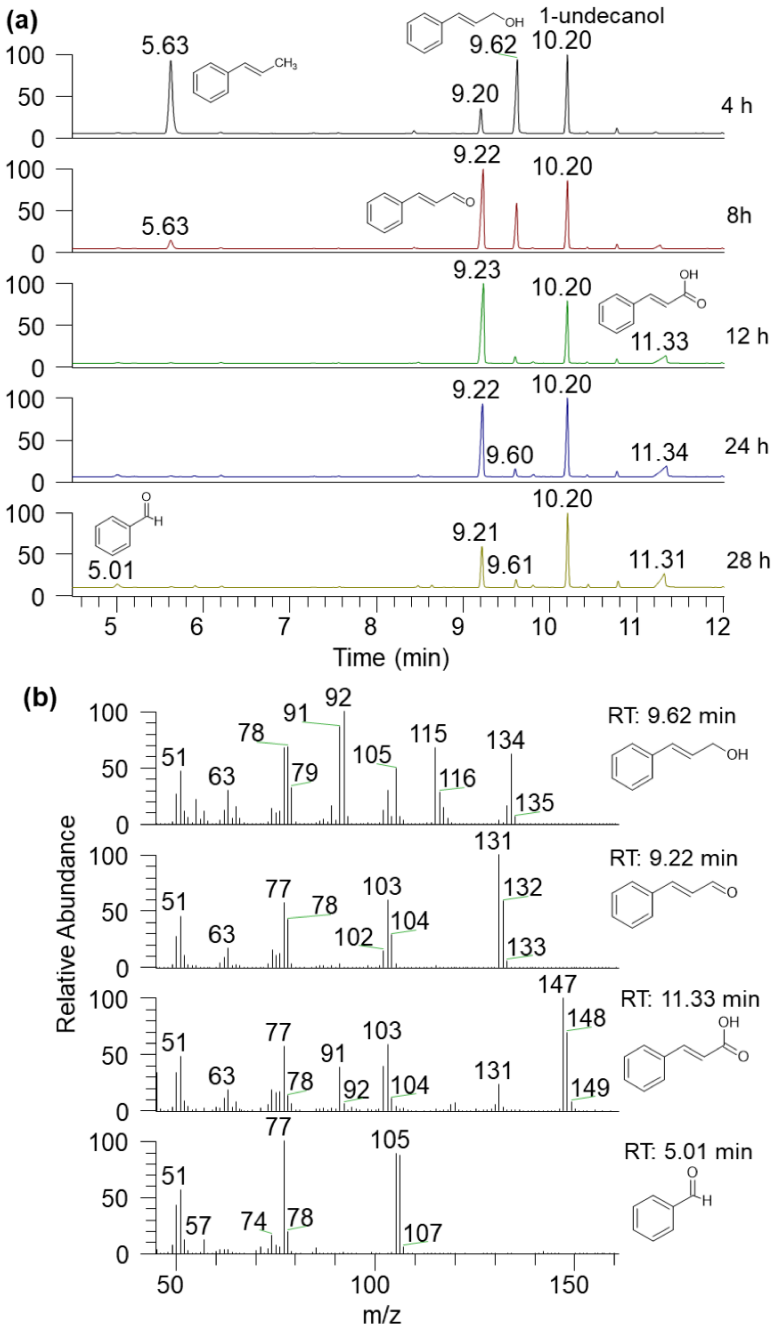


Figure S11 (a)Chromatograms and (b) mass spectra of GC-MS/FID analyses of extracts from preparative scale reaction of SscaCYP_E284I with *trans*-β-methylstyrene.

References

Studier, F. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, *41*(1), 207–234. https://doi.org/10.1016/j.pep.2005.01.016

2.2 Chromosomal integration of three recombinant genes for the synthesis of (2*S*,6*S*)-hydroxynorketamine with resting *E. coli* cells

Title: Effect of chromosomal integration on catalytic performance of a multi-component P450 system in *Escherichia coli*

Authors: **U. J. Luelf**, L. M. Böhmer, S. Li, V. B. Urlacher

Published in: Biotechnology and Bioengineering, 120, 1762–1772

DOI: 10.1002/bit.28404

Contribution: Conceptualized the study, designed and conducted most of the experiments, analyzed the data, prepared the first draft and edited the manuscript.

Received: 7 February 2023 | Revised: 4 April 2023 | Accepted: 9 April 2023

DOI: 10.1002/bit.28404

ARTICLE

BIOTECHNOLOGY & BIOENGINEERING WILEY

Effect of chromosomal integration on catalytic performance of a multi-component P450 system in *Escherichia coli*

U. Joost Luelf[1] | Lisa M. Böhmer[1] | Shengying Li[2] | Vlada B. Urlacher[1]

[1]Institute of Biochemistry, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany

[2]State Key Laboratory of Microbial Technology, Shandong University, Qingdao, Shandong, China

Correspondence
Vlada B. Urlacher, Institute of Biochemistry, Heinrich-Heine University Düsseldorf, 40225 Düsseldorf, Germany.
Email: vlada.urlacher@uni-duesseldorf.de

Funding information
European Regional Development Fund (EFRE), Project "ClusterIndustrial Biotechnology (CLIB) Kompetenzzentrum Biotechnologie (CKB)

Abstract

Cytochromes P450 are useful biocatalysts in synthetic chemistry and important biobricks in synthetic biology. Almost all bacterial P450s require separate redox partners for their activity, which are often expressed in recombinant *Escherichia coli* using multiple plasmids. However, the application of CRISPR/Cas recombineering facilitated chromosomal integration of heterologous genes which enables more stable and tunable expression of multi-component P450 systems for whole-cell biotransformations. Herein, we compared three *E. coli* strains W3110, JM109, and BL21(DE3) harboring three heterologous genes encoding a P450 and two redox partners either on plasmids or after chromosomal integration in two genomic loci. Both loci proved to be reliable and comparable for the model regio- and stereoselective two-step oxidation of (*S*)-ketamine. Furthermore, the CRISPR/Cas-assisted integration of the T7 RNA polymerase gene enabled an easy extension of T7 expression strains. Higher titers of soluble active P450 were achieved in *E. coli* harboring a single chromosomal copy of the P450 gene compared to *E. coli* carrying a medium copy pET plasmid. In addition, improved expression of both redox partners after chromosomal integration resulted in up to 80% higher (*S*)-ketamine conversion and more than fourfold increase in total turnover numbers.

KEYWORDS
chromosomal integration, CRISPR/Cas9, cytochrome P450, episomal expression, redox partners, whole-cell catalyst

1 | INTRODUCTION

Selective oxidation of complex compounds is an important step towards their functionalization in synthetic chemistry and structural diversification during biosynthesis. Several enzymes enable C–H bond oxidation, including α-ketoglutarate-dependent hydroxylases, dioxygenases and unspecific peroxygenases (Aranda et al., 2021). However, cytochrome P450 monooxygenases (P450s or CYPs) currently remain unsurpassed regarding their broad substrate and reaction spectra as well as their targeted specificity. In this context, P450-catalyzed reactions often represent key steps in biosynthetic pathways as well as artificial multienzyme cascades (Fessner, 2019). Two well-known examples are the synthesis of a precursor of the chemotherapeutic drug taxol in *Escherichia coli* (Ajikumar et al., 2010) and the production of a precursor of the antimalarial drug artemisinin in *Saccharomyces cerevisiae* (Ro et al., 2006). Due to these properties,

Dedicated to Karl-Erich Jaeger for his outstanding contributions in the field of molecular enzyme technology.

P450s have been recognized not only as useful biocatalysts for organic chemistry but also as important bio-bricks in synthetic biology (Urlacher & Girhard, 2019).

P450s are normally dependent on the cofactor NAD(P)H and on redox partner proteins which transfer electrons from NAD(P)H to the heme group and thus almost always function as multi-component systems (Hannemann et al., 2007; Li et al., 2020). After substrate binding, ferric (Fe^{3+}) heme iron is reduced by redox partners via one electron reduction to ferrous (Fe^{2+}) form, which binds molecular oxygen. After the second one-electron reduction by the redox partners and two protonation steps, the dioxygen bond is cleaved, yielding the reactive ferryl-oxo species Compound I and water as byproduct. The Compound I is considered the main active species in P450 reactions (Belcher et al., 2014). Traditionally, microorganisms harboring P450s and their cognate redox partner proteins have been used since the 1960s to produce oxidized compounds like hydrocortisone or pravastatin at industrial scale (Park et al., 2003; Petzoldt et al., 1982). After introducing recombinant DNA technology, heterologous microbial hosts have become more attractive as whole-cell biocatalysts or microbial cell factories with better control over enzyme activities and costs. For P450-catalyzed reactions, *E. coli* seems to be a particularly appropriate host because it is easy to manipulate and does not express intrinsic P450s (Brixius-Anderko et al., 2015). Metabolic pathways and gene expression in *E. coli* are extensively studied, and a number of commercial expression strains as well as a variety of genetic engineering methods are available (Pontrelli et al., 2018).

In *E. coli*, heterologous gene expression is often accomplished using episomal (plasmid-based) systems. Such systems generally have several advantages like easy cloning and gene manipulation including protein engineering. Nevertheless, there are also a few drawbacks, particularly when antibiotic resistance genes are used as selectable markers to maintain plasmids within the host. Apart from the negative impact on further emergence of antibiotic resistance, the use of antibiotics often affects cell growth. Furthermore, multiple high copy plasmids may lead to the expression heterogeneity and contribute to the high transcriptional metabolic burden (Ajikumar et al., 2010; Birnbaum & Bailey, 1991; Mairhofer et al., 2013). In contrast, chromosomal integration of heterologous genes eliminates the need for antibiotics while maintaining stable expression and reducing the transcriptional metabolic burden which can result in similar or even higher enzyme activities and increased metabolite production compared to plasmid-based systems (Englaender et al., 2017; Mairhofer et al., 2013; Wang et al., 2016). The adaption of the CRISPR/Cas system for *E. coli* has expedited genome editing including chromosomal integration (Jiang et al., 2015), and its combination with λ-Red recombineering has been reported several times for *E. coli* (Dong et al., 2021; Ou et al., 2018). In contrast to the λ-Red recombineering method originally described by Datsenko and Wanner (2000), CRISPR/Cas-assisted recombineering can be designed "scarless" which facilitates iterative genome editing, because it excludes the risk of undesired genomic rearrangements (Reisch & Prather, 2015).

While examples of the use of *E. coli* strains with chromosomally integrated genes or even metabolic pathways are still less common than plasmid-based systems, to our knowledge, no direct comparison between plasmid-based and chromosomal expression for P450-catalyzed biotransformations has been described yet. In this regard, it is not only important to achieve high P450 expression levels, but also to balance its concentration with those of the redox partners because their ratios affect the conversion efficiency (Bakkes et al., 2017; Khatri et al., 2017; Liu et al., 2022). To perform such a comparison, we chose a variant of CYP154E1 from *Thermobifida fusca* YX (further referred to as CYP154E1 QAA) supported by two redox partners—the flavodoxin YkuN from *Bacillus subtilis* and the FAD-containing reductase FdR from *E. coli*—as a three-component model P450 system. In our previous study, the YkuN/FdR couple was selected as the most appropriate for reconstituting CYP154E1 activity among several heterologous redox partners (von Bühler et al., 2013). CYP154E1 was recently engineered by mutagenesis to catalyze a two-step conversion of (*S*)-ketamine via the highly chemoselective *N*-demethylation and the regio- and stereoselective C6-hydroxylation to (2*S*,6*S*)-hydroxynorketamine, which is considered a potential antidepressant (Bokel et al., 2020; Zanos et al., 2016, 2018) (Scheme 1).

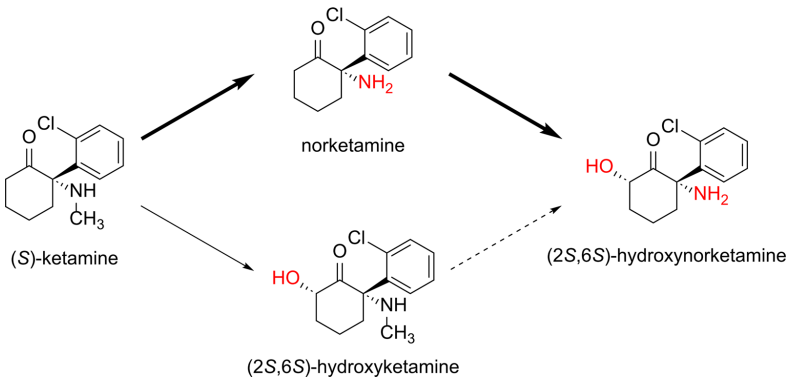
In our previous work, a two-plasmid system harboring all three genes under control of the T7 promoter was established, and whole cells were used for this biotransformation (Bokel et al., 2020). However, T7 expression requires the presence of the phage T7 RNA polymerase (T7RNAP) which thus limits the choice of expression strains. Therefore, the commercial, engineered *E. coli* B strain BL21(DE3) as well as its derivatives, among others, are commonly used (Dumon-Seignovert et al., 2004; Miroux & Walker, 1996). In contrast, the number of commercially available K-12 strains which are suitable for T7 expression is, however, very limited. Hence, the first step of this study was the CRISPR/Cas-assisted chromosomal integration of *T7 gene 1* which encodes the T7RNAP into different K-12 strains.

Further, several plasmid-free recombinant strains were constructed, in which the three genes encoding CYP154E1 QAA, YkuN and FdR were integrated combinatorially into different genomic loci. Particular focus was placed on the experimental design for fast PCR-based synthesis of the linear donor DNA template for homologous recombination. This method facilitates the step from plasmid-based to plasmid-free biotransformations using existing pET vector-based libraries as a starting point. Finally, episomal and chromosomal expression for efficient two-step whole-cell biotransformation of (*S*)-ketamine to (2*S*,6*S*)-hydroxynorketamine were compared.

2 | MATERIALS AND METHODS

2.1 | Strains and plasmids

E. coli DH5α (Clontech) was used for cloning and plasmid propagation. *E. coli* BL21(DE3) (Novagen), *E. coli* JM109 (NEB) and *E. coli*



SCHEME 1 Consecutive two-step oxidation of (S)-ketamine to (2S,6S)-hydroxynorketamine catalyzed by the CYP154E1 QAA variant. Time-course studies have revealed a high preference (indicated by bold arrows) of oxidative *N*-demethylation followed by hydroxylation rather than the other way around (Bokel et al., 2020).

W3110 (DSMZ) were used for genome engineering, expression and whole-cell conversion. pET-16b YkuN, pET-28a(+) CYP154E1 QAA (includes N-terminal 6xHis tag), pET-22b(+) CYP154E1 QAA (without His tag) and pCOLADuet YkuN (I) FdR (II) were described in previous studies (Bokel et al., 2020; Girhard et al., 2010) and served as templates to construct donor DNA for chromosomal integration or were used for episomal expression, respectively. Plasmids pCas (Addgene plasmid #62225) (Jiang et al., 2015), pEcCas (Addgene plasmid #73227) (Li et al., 2021), and pgRNA-bacteria (hereinafter referred to as "pgRNA," Addgene plasmid #44251) (Qi et al., 2013) were derived from Addgene. A complete list of all primers, strains, and plasmids used in this study can be found in the Supporting Information (Tables S1–S3).

2.2 | Verification of Lon protease deficiency in *E. coli* JM109

Primers binding upstream and downstream of the region which expected to harbor the *lopA* gene (encoding Lon protease) in *E. coli* JM109 (cf. Table S1) were designed based on the genome sequence of *E. coli* W3110 (GenBank accession no. AP009048.1). The PCR product was then sequenced using Sanger sequencing (Eurofins Genomics).

2.3 | Chromosomal integration

CRISPR/Cas-assisted chromosomal integration was carried out by adapting the protocol originally described by Jiang et al. (2015). Evaluation and design of guide RNAs (gRNAs) for CRISPR/Cas-assisted genome engineering were performed using the CHOPCHOP v3 web tool (Labun et al., 2019) except for the gRNA targeting the locus *nupG* in K-12 strains which was previously described by Bassalo et al. (2016). Due to a single mutation, a slightly different targeting sequence had to be used in B strains as compared to K-12 strains for this locus (cf. Table S1). To substitute the N_{20} targeting sequence of the aforementioned plasmid pgRNA, a PCR followed by restriction digestion with FastDigest™ BcuI (SpeI) and DpnI (Thermo Scientific) and a self-circularization ligation step was carried out as described by Seo et al. (2017). Successful gRNA substitution was verified by Sanger sequencing (Eurofins Genomics).

Homology arms (~500 bp) for λ-Red-mediated homologous recombination were designed using the genome sequences of *E. coli* W3110 (GenBank accession no. AP009048.1) and *E. coli* BL21(DE3) (GenBank accession no. CP001509.3) and were amplified by PCR from boiled cells. *T7 gene 1* (encodes T7 RNA polymerase for integration in K-12 strains) was amplified from boiled cells of *E. coli* BL21(DE3) as well. Other genes for integration were amplified from pET-28a(+) CYP154E1 QAA, pET-16b YkuN, and pCOLADuet YkuN (I) FdR (II), respectively. Double-stranded donor DNA (dsDonorDNA) for homologous recombination was generated by fusion PCR of homology arms and the gene(s) of interest (cf. Figure S1). Fusion PCR contained 50 ng of the amplified gene and equimolar amounts of the homology arms as well as 0.5 μM of the outermost primers (cf. Table S1) in 50 μL total reaction volume. The further composition of the samples was prepared according to the manufacturer's protocol for Phusion™ DNA polymerase (Thermo Scientific).

For preparation of electrocompetent cells, 5 mL 2xYT medium (30 μg/mL kanamycin) were inoculated with 100 μL of an overnight culture harboring either pCas (in case of W3110) or the updated plasmid version pEcCas (in BL21(DE3) and JM109), which was designed to improve genome editing in BL21(DE3) (Li et al., 2021). For induction of λ-Red genes, 100 μL 1 M L-arabinose were added and the cultures were incubated at 30°C (pCas) or 37°C (pEcCas), 200 revolutions per minute (rpm) for 2 h (pEcCas) or 4 h (pCas),

respectively. After cell harvesting by centrifugation, the cells were washed twice with 1 mL ice-cold 10% (vol/vol) glycerol (cell pellet resuspended, 10 min incubation on ice, centrifugation). Finally, the cell pellet was resuspended in 100 µL ice-cold 10% (vol/vol) glycerol. After addition of 100 ng pgRNA and approximately 500 ng donor dsDonorDNA, electroporation was performed in a 2 mm gap electroporation cuvette at 2.5 kV (MicroPulser™ electroporator from Bio-Rad). The cells were suspended immediately in 1 mL of ice-cold 2xYT or SOC medium and recovered for 1–2 h at 30°C (pCas) or 37°C (pEcCas), respectively. Subsequently, the cells were plated out on LB agar plates (100 µg/mL ampicillin, 30 µg/mL kanamycin) and incubated overnight. Successful genome engineering was verified by colony PCR. Curing of pgRNA plasmid, while maintaining pCas/pEcCas for iterative genome engineering, was performed in 5 mL LB medium either by addition of 0.5 mM IPTG (pCas) or 10 mM L-rhamnose (pEcCas). For curing of pCas/pEcCas plasmids, incubation at 42°C, 180 rpm (pCas) or addition of 5% (wt/vol) sucrose (pEcCas) and subsequent incubation at 37°C, 180 rpm were performed (in some cases this step needed to be repeated at 42°C). After streaking, single colonies were tested for antibiotic resistance to obtain cured strains.

2.4 | Expression and biotransformation

For evaluation purposes, P450 expression was carried out in 50 mL Terrific Broth (TB) medium in 500 mL Erlenmeyer flasks. Episomal expression required 30 µg/mL kanamycin for maintaining the pET-28a(+) plasmid. After inoculation with 1 mL of an overnight culture, the expression cultures were incubated at 37°C, 180 rpm until an optical density (OD_{600}) of 0.6–0.8 was reached. Transcription was induced by adding 0.1 mM IPTG. To support heme production, 0.1 mM $FeSO_4$ as well as 0.5 mM 5-aminolevulinic acid were supplemented. After incubation at 25°C, 140 rpm for 24 h, the cells were harvested by centrifugation. The pellet was resuspended in 4 mL 100 mM potassium phosphate buffer (pH 7.5) per gram cell wet weight for cell disruption by sonication. Cell debris was separated by centrifugation and the soluble fraction was used for P450 quantification by carbon monoxide (CO) difference spectra as described by Omura & Sato, (1964a, 1964b). In addition, sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) analysis of both the soluble and insoluble fractions after cell disruption was performed.

Coexpression of P450 and redox partners was carried out in autoinduction medium (100 mL TB medium containing 0.1 mM $FeSO_4$ and 0.5 mM 5-aminolevulinic acid, 2 mL "50 × 5052," 2 mL "50 x M" as described by Studier, 2014). Episomal expression required 100 µg/mL ampicillin and 30 µg/mL kanamycin for maintaining the plasmids pET-22b(+) and pCOLADuet. After inoculation with 2 mL of an overnight culture, expression cultures were incubated at 37°C, 180 rpm until an OD_{600} of approximately 1.0 was reached. After incubation at 25°C, 120 rpm for 20 h, the cultures were split in half and harvested by centrifugation. One half was treated as described above. The other half was washed once with PSE buffer (6.75 g/L KH_2PO_4, 85.5 g/L sucrose, 0.93 g/L EDTA-Na_2·2 H_2O, pH 7.5), adjusted to a cell density of 50 mg/mL and stored at −80°C for at least 24 h.

Whole-cell conversion was carried out in 500 µL scale for which 465 µL of cell suspension were thawed. After addition of 25 µL nutrient solution (0.12 M glucose, 0.12 M lactose, 0.24 M sodium citrate in PSE buffer) and 10 µL 100 mM (S)-ketamine (2 mM final concentration), the reaction mixture was incubated at 25°C, 1200 rpm for 20 h. Extraction was performed with ethyl acetate after adding sodium carbonate and the internal standard xylazine as described by Bokel et al. (2020). The organic solvent was evaporated, and the residue was resuspended in a mixture of acetonitrile and water. Product formation was quantified by LC/MS analysis (LCMS-2020, Shimadzu; Chromolith Performance RP-8e 100–4.6 mm column and Chromolith RP-8e 5–4.6 mm guard cartridge, Merck Millipore) as described before (Bokel et al., 2020).

3 | RESULTS AND DISCUSSION

3.1 | Enabling T7 expression in K-12 strains

The number of commercially available K-12 strains which are suitable for T7-based expression is very limited, although *E. coli* K-12 strains appear to be more easily manipulated by genome editing methods and show occasionally beneficial characteristics in comparison to B strains like lower sensitivity to stress and tolerance to lower oxygen levels (Kang et al., 2002; Yoon et al., 2012). Thus, we chose the *E. coli* K-12 strains W3110 and JM109 for the integration of *T7 gene 1*. W3110 is an industrially used strain which is closely related to the ancestral K-12 strain (Hayashi et al., 2006; Kang et al., 2002). By selecting JM109, we aimed to combine the high editing efficiency of K-12 strains with improved heterologous protein production due to the lack of Lon protease. The Lon protease encoded by *lopA* gene contributes to degradation of foreign proteins, making Lon-protease-deficient strains such as BL21(DE3) commonly preferred for heterologous gene expression (SaiSree et al., 2001). Lon protease deficiency in JM109 due to IS*186* transposon insertion was first mentioned at the OpenWetWare website (Richter et al., 2019) and confirmed in this study by sequencing the region of the *lopA* gene promoter (Figure S2). Very recently, the complete genome sequence of the JM109 strain was uploaded to GenBank (accession no. CP117962.1) and revealed this transposon insertion and the absence of a second copy of *lopA*.

The BL21(DE3) strain for T7 expression was originally constructed by cloning *T7 gene 1* into a lambda phage derivative and subsequent lysogenization of BL21 (Studier & Moffatt, 1986). Nowadays, advances in genome editing allow targeted integration without lysogenization which avoids the consequent integration of irrelevant phage genes. Since expression levels of chromosomally integrated genes depend on their integration site (Bryant et al., 2014; Englaender et al., 2017), *T7 gene 1* was integrated between the genes

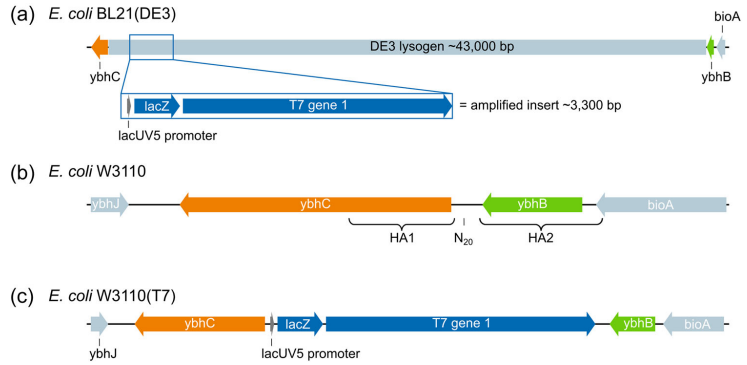


FIGURE 1 Construction of the *Escherichia coli* W3110(T7) strain. (a) *T7 gene 1* (encoding T7 RNA polymerase) under control of lacUV5 promoter and a lacZ fragment located between promoter and *T7 gene 1* were amplified from the DE3 lysogen of *E. coli* BL21(DE3). (b) Gene map of the insertion site in *E. coli* W3110 and design of the 500 bp homology arms (HA1 and HA2) for homologous recombination upstream and downstream of the N_{20} target sequence for CRISPR/Cas cleavage. (c) Gene map of *E. coli* W3110(T7) showing the integration of *T7 gene 1* between the genes *ybhC* (orange) and *ybhB* (green) comparable to *E. coli* BL21(DE3).

ybhB and *ybhC* as seen in commercial BL21(DE3) to ensure comparability (Figure 1). To this end, *T7 gene 1* under control of the lacUV5 promoter was amplified from the genome of *E. coli* BL21(DE3) and integrated into both K-12 strains using CRISPR/Cas9-assisted recombineering. Integration of a second copy of the *lacI* gene, which is also located in the DE3 lysogen of BL21(DE3), was omitted because we suggested that the copy present in the lac operon is sufficient to repress transcription of single chromosomal copies of heterologous genes. This approach was also described by another group and did not result in higher basal plasmid-based expression (Ting et al., 2020). The constructed K-12 strains harboring a chromosomal copy of *T7 gene 1* are hereafter referred to as W3110(T7) and JM109(T7).

3.2 | Effect of Lon protease deficiency on CRISPR/Cas-assisted genome engineering

Besides our efforts to enable T7 expression in K-12 strains, the use of an updated version of the plasmid pCas—so-called pEcCas—improved genome editing in BL21(DE3) (Li et al., 2021), so that in the following all three *E. coli* strains—W3110(T7) and JM109(T7) as well as BL21(DE3)—were compared. However, it should be noted that the integration efficiency in BL21(DE3) remains low compared to W3110. Surprisingly, the observed integration efficiency in JM109 was equally low (data not shown) as in BL21(DE3), although K-12 strains are reported to show higher integration efficiencies than B strains (Chung et al., 2017; Li et al., 2019). This might be due to a negative effect of the Lon protease deficiency on the efficiency of CRISPR/Cas-assisted chromosomal integration, which was very recently reported by Okshevsky et al. (2023). In addition to its contribution to the degradation of improperly folded proteins, the Lon protease plays a critical role in the SOS response to DNA damage in cells (Gottesman, 1996). It is likely that the expression of the *lopA* gene explains the differences between BL21(DE3), JM109, and W3110 in our study, rather than the fact that one is a B strain, and the others K-12 strains. When choosing a strain for chromosomal integration and subsequent expression of heterologous genes, it should be considered whether the lack of Lon protease is crucial to produce recombinant proteins or whether high integration efficiency is more important.

3.3 | Integration procedure, integration loci, and donor DNA design

For integration of the *cyp154e1m* gene coding for the target CYP154E1 QAA enzyme, two genomic loci previously described in literature and located similarly close to the origin of replication were evaluated. The gene-disrupting integration of the green fluorescent protein gene (*gfp*) into *nupG* (encoding a nucleoside permease) described by Bassalo et al. (2016) resulted in high fluorescence and high integration efficiency. High expression levels after integration in this region were also observed by Bryant et al. (2014) and Goormans et al. (2020). The second integration site, an intergenic region between the genes *atpI* and *gidB* (also known as *rsmG*) was previously described by Englaender et al. (2017) and Goormans et al. (2020) among others.

For chromosomal integration, the entire cassette consisting of the *cyp154e1m* gene under control of the T7 promoter as well as the T7 terminator was amplified from the pET-based plasmid and fused with homology arms for homologous recombination (Figure S1). The primers listed in the supporting information were specially designed for PCR amplification from any pET vector

which contains a f1 origin of replication (indicated by the (+) following the name, e.g. in pET-28a(+)) (Merck KGaA, 2011). In deviation from this, different primers for amplification of the cassettes of pET-16b and pCOLADuet were designed. Thus, these primer sets can be used for donor DNA synthesis and subsequent chromosomal integration of any desired gene available on most pET or pCOLADuet vectors which facilitates the step from plasmid-based to chromosomal expression in *E. coli*. Hence, after successful optimization of properties such as activity, selectivity, and stability of an enzyme by mutagenesis using well-established plasmid-based engineering techniques, stable chromosomal expression of the optimized mutant can be achieved with little effort using the herein described method and primers. It is worthy of note, however, when choosing new integration loci for simultaneous expression of more genes in reaction cascades, it is important to keep in mind that the T7 terminator (Tφ) is only approximately 70% effective (Merck KGaA, 2011). Potential read-through and overexpression of homologous genes downstream of the integration loci might be toxic to the cells.

3.4 | Comparison of episomal and chromosomal expression of *cyp154e1m*

Expression studies revealed that chromosomal expression of *cyp154e1m* was improved by an N-terminal polyhistidine (6xHis) tag (data not shown). After integration of the *cyp154e1m* gene including the 6xHis tag into the chosen integration loci (*atpI_gidB* and *nupG*) in all three strains, high discrepancies between episomal and chromosomal expression as well as between the different strains were observed (Figure 2). The highest P450 expression level of 246 nmol per gram cell wet weight (cww) or 6600 nmol per liter culture was achieved in strain BL21(DE3) for chromosomal expression. Similar P450 amounts (also for other P450s) are usually only achieved in high cell density cultivations rather than simple flask cultivation (Pflug et al., 2007). More importantly, the chromosomal integration of *cyp154e1m* increased the titer of soluble P450 by 40% (from 176 ± 23 to 246 ± 23 nmol/g cell wet weight) in strain BL21(DE3) as compared to episomal expression based on pET-28a (+). Englaender et al. (2017) even described a fivefold increase in recombinant mCherry production after integration into the *atpI_gidB* locus compared to a modified pET vector. In our case, this increase can be explained by a stronger episomal expression leading to a high portion of inactive P450 in the insoluble fraction, as revealed by SDS-PAGE analysis (Figure 2b). It is commonly known that overexpression under control of the strong T7 promoter using a high-copy vector may lead to misfolding and aggregation of improperly folded proteins to inclusion bodies in *E. coli* (Hoang et al., 1999). In addition, accumulation of P450 apoprotein which does not contain the prosthetic heme group can occur if protein folding outpaces heme loading (Sudhamsu et al., 2010). A reduction of the gene copy number from up to approximately 40 copies per cell as expected for a pET vector (Merck KGaA, 2011) to a single chromosomal copy reduced the amount of potentially improperly folded, insoluble P450 and, in turn, increased the amount of heme-loaded, soluble enzyme in BL21(DE3).

A moderate increase in P450 amount compared to the episomal system was also observed after chromosomal expression in W3110(T7). The P450 amounts in soluble and insoluble fractions were generally lower compared to BL21(DE3) (Figure 2 and Figure S3). In the literature, integration of *T7 gene 1* into W3110 resulted in higher fluorescence of sfGFP compared to BL21(DE3) (Ting et al., 2020). However, the integration site of the *T7 gene 1* used in our study is different, which does not allow direct comparison. Why P450 expression worked better in BL21(DE3) was not elucidated further in the frame of this study, however, the expression level generally depends on the gene chosen and might vary for different proteins. Furthermore, possible differences in heme

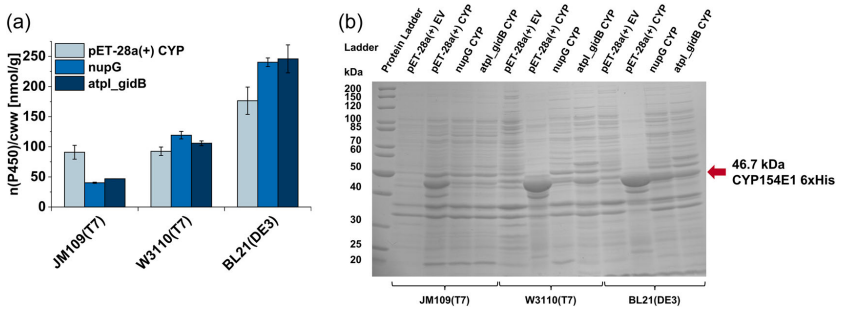


FIGURE 2 Comparison of episomal and chromosomal expression of *cyp154e1m* gene in three *E. coli* strains after induction with 0.1 mM IPTG. Episomal expression was carried out using a pET-28a(+) vector, and chromosomal expression was performed after integration into the loci *nupG* and *atpI_gidB*. (a) Amount of soluble P450 per gram cell wet weight and (b) SDS-PAGE analysis of the insoluble protein fractions after expression and cell disruption. Empty vector (EV) was used as negative control. The expected molecular weight of CYP154E1 QAA (6xHis) is 46.7 kDa.

synthesis or loading in different strains, to the best of our knowledge, have not been investigated so far.

The lowest amount of soluble P450 was measured after chromosomal expression in JM109(T7) (Figure 2a and Figure S3), while after episomal expression, P450 concentration was as high in JM109(T7) as in W3110(T7). Moreover, only in this strain episomal expression resulted in higher concentration of soluble heme-loaded P450 than chromosomal expression. Unexpectedly, less insoluble protein was produced in Lon-deficient JM109(T7) than in W3110(T7), although as mentioned before, Lon protease degrades misfolded proteins. Obviously, Lon protease deficiency is only one feature that can influence heterologous expression in these strains and other factors should be investigated to explain all observed differences between the expression strains used.

Both genomic loci resulted in similar expression levels within the respective strains. In comparison, a previous study by Goormans et al. (2020) reported a slightly higher expression level of GFP in case of the *atpI_gidB* locus compared to the *nupG* locus.

3.5 | Coexpression of redox partners and biotransformation of (*S*)-ketamine

The strains BL21(DE3) and W3110(T7) were chosen for further experiments. The previously described plasmid-based system consisted of a pET-22b(+) plasmid for *cyp154e1m* expression and a pCOLADuet plasmid for coexpression of *ykuN* and *fdr* (Bokel et al., 2020). This system was not altered and served as the reference system for episomal expression. As described for the integration of the *cyp154e1m* gene, the whole bicistronic operon was amplified from the pCOLADuet vector carrying the *ykuN* and *fdr* genes, and integrated into the locus that was not already occupied by the previously integrated *cyp154e1m* gene (Table 1).

Since the FAD-containing reductase FdR is originating from *E. coli* and in vitro setups for P450-catalyzed reactions usually contain an excess of the flavodoxin/ferredoxin but not the reductase, strains harboring only the *cyp154e1*m and *ykuN* genes without additional integration of the *fdr* gene were also created to investigate whether the endogenous expression of *fdr* is sufficient for the whole-cell biotransformation. After expression, SDS-PAGE analysis of the proteins, biotransformation of 2 mM (*S*)-ketamine, and quantification of the P450 concentration were performed.

TABLE 1 Bacterial strains with combinatorial integration of *cyp154e1m*, *ykuN*, or *ykuN/fdr*.

Native strain	Combinatorially integrated genes		Label
	***atpI_gidB* Locus**	***nupG* Locus**	**Label**
BL21(DE3) or W3110(T7)	*cyp154e1m*	*ykuN*	P450-Y
	cyp154e1m	*ykuN/fdr*	P450-Y/R
	ykuN	*cyp154e1m*	Y-P450
	ykuN/fdr	*cyp154e1m*	Y/R-P450

Similar to the case when *cyp154e1m* was expressed alone (Figure 2), an increased P450 expression was observed in BL21(DE3) compared to W3110(T7) (Figure 3). Not only the amount of protein visible on the gel but also the P450 concentrations measured after cell disruption were higher. Higher conversion of (*S*)-ketamine by the BL21(DE3) strains was observed in comparison to W3110(T7). In this context, lower conversion correlates with high levels of the intermediate norketamine (Scheme 1).

The norketamine intermediate is the main product formed from ketamine by W3110(T7), while in case of BL21(DE3) more of the desired (2*S*,6*S*)-HNK was formed (Figure 3c). Remarkably, the turnover numbers of 880–1100 oxidation reactions per P450 calculated for both strains seem comparable, so that differences in product formation and distribution between the strains may be due to different P450 concentrations (Figure 3d).

Consistent with the results shown in Figure 2, which implied high similarity of the two integration loci, it made almost no difference whether the *cyp154e1m* gene was integrated into the *nupG* locus and the redox partners into the *atpI_gidB* locus, or vice versa. Interestingly, in contrast to our previous results, the chromosomal (co)expression of P450 and redox partners resulted in lower P450 concentrations compared to the episomal expression systems in both strains (from 37% up to 64% reduction dependent on the chromosomal system). However, in all cases, simultaneous chromosomal coexpression of *ykuN* and *fdr* resulted in slightly higher P450 amounts compared to the strains where no additional copy of *fdr* was integrated, although the bicistronic operon was expected to increase the transcriptional burden. In addition, a high portion of P450 was found in the insoluble fraction after cell disruption for BL21(DE3), but not in W3110(T7). However, both redox partner proteins were rather soluble in both strains (Figure S4). Even though the highest P450 concentration was achieved after episomal expression (123 ± 7 nmol/g cell wet weight), the highest (*S*)-ketamine conversion was observed for both strains with all three chromosomally integrated genes (up to 90% (*S*)-ketamine conversion for BL21(DE3) and up to 46% conversion for W3110(T7)). For strains without additionally integrated copy of *fdr*, lower turnover was observed which implies that the endogenous *fdr* is not sufficient for this biotransformation (Figure 3).

SDS-PAGE analysis revealed higher amounts of the redox partner proteins after chromosomal expression compared to the episomal coexpression, which might explain the lower conversions in the latter despite higher P450 concentrations (Figure 3 and Figure S4). It is commonly observed for P450-catalyzed reactions that the ratio of P450:flavodoxin:reductase has a great influence on the catalytic performance (Bakkes et al., 2017; Khatri et al., 2017). While in in vitro reactions with purified enzymes this ratio can easily be adjusted and fine-tuned, there are only limited possibilities to do so in recombinant microbial hosts. For future optimization, the ratio could either be adjusted by altering the transcription strength using different promoters or by altering the individual copy numbers of the genes. In addition, in case of chromosomal expression, the use of different loci might result in a different expression pattern and subsequent altered catalytic performance.

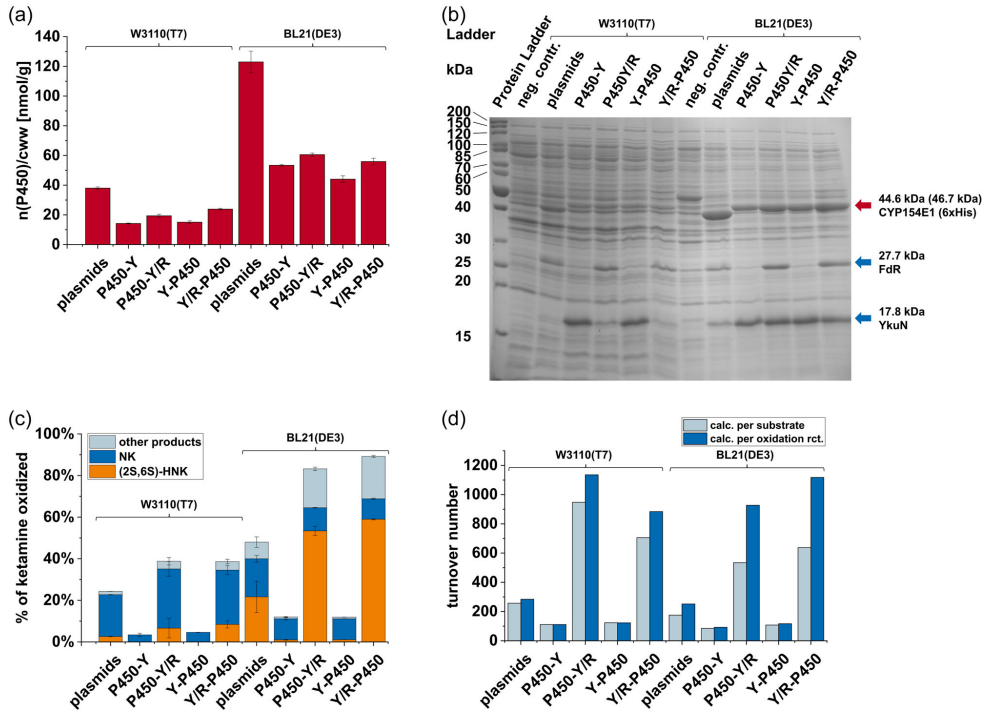


FIGURE 3 Coexpression of P450 and redox partner genes after combinatorial chromosomal integration into two genomic loci (labeling according to Table 1) in comparison to the episomal two-plasmid expression system and whole-cell conversion of 2 mM (*S*)-ketamine. (a) Amount of soluble P450 per gram cell wet weight and (b) SDS-PAGE analysis of whole cells after expression in autoinduction medium. Strains without plasmids or integrated genes were used as negative control. (c) Comparison of product formation ((2*S*,6*S*)-hydroxynorketamine [(2*S*,6*S*)-HNK]), norketamine (NK), and other oxidation products (several hydroxyketamines [HK] and -norketamines [HNK]). (d) Turnover number (TON) was calculated as the number of substrate molecules converted per enzyme (light blue). Since the number of oxidation reactions varies for different products (HK and NK = one oxidation; HNK and dehydronorketamine = two oxidations), an adjusted TON representing the number of oxidation steps catalyzed per enzyme was calculated to account for different product distributions (dark blue).

4 | CONCLUSIONS

In summary, we demonstrated that P450-catalyzed biotransformation can be performed using *E. coli* cells with chromosomally integrated recombinant genes. When expressed alone, the chromosomally integrated *cyp154e1m* resulted in higher P450 concentration compared to the plasmid-based expression. More importantly, a better-balanced electron transport chain improved (*S*)-ketamine conversion to the potential antidepressant (2*S*,6*S*)-hydroxynorketamine. The two integration loci *nupG* and *atpI_gidB* proved to be reliable and comparable. In addition, the chromosomal integration of T7 RNA polymerase gene enables an easy extension of possible T7 expression strains so that specific advantages of different strains can be combined with strong T7-based expression. In our particular case, the BL21(DE3) strain widely used for protein expression still performed best. This is probably different for other genes, as it has been described in literature (Ting et al., 2020). The influence of the Lon protease deficiency not only on the production of recombinant proteins, but also on the efficiency of genome engineering should be considered when performing comparable experiments. The primers designed in this study allow easy chromosomal integration of any genes from pET vectors. Chromosomal integration can be regarded as the final step after successful optimization of biocatalysts by mutagenesis to a stable and antibiotic-free expression and biotransformation.

AUTHOR CONTRIBUTIONS

Vlada B. Urlacher and U. Joost Luelf: conceptualized the study. **U. Joost Luelf**: designed and conducted most of the experiments, analyzed the data and prepared the first draft. **Lisa M. Böhmer**: contributed to data acquisition and analysis. **Vlada B. Urlacher and Shengying Li**: contributed to conceiving experiments, interpretation

of data, and writing of the manuscript. All authors edited the manuscript and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We thank Dr. Marco Girhard and Dr. Katja Koschorreck for fruitful discussions. Financial support was provided by the state of North Rhine-Westphalia (NRW) and the "European Regional Development Fund (EFRE)", Project "ClusterIndustrial Biotechnology (CLIB) Kompetenzzentrum Biotechnologie (CKB)" (34.EFRE-0300095/1703FI04). Open Access funding enabled and organized by Projekt DEAL.

CONFLICTS OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request. Data are available from the corresponding author on reasonable request.

ORCID

Vlada B. Urlacher http://orcid.org/0000-0003-1312-4574

REFERENCES

Ajikumar, P. K., Xiao, W. H., Tyo, K. E. J., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T. H., Pfeifer, B., & Stephanopoulos, G. (2010). Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science*, *330*(6000), 70–74. https://doi.org/10.1126/science.1191652

Aranda, C., Carro, J., González-Benjumea, A., Babot, E. D., Olmedo, A., Linde, D., Martínez, A. T., & Gutiérrez, A. (2021). Advances in enzymatic oxyfunctionalization of aliphatic compounds. *Biotechnology Advances*, *51*, 107703. https://doi.org/10.1016/j.biotechadv.2021.107703

Bakkes, P. J., Riehm, J. L., Sagadin, T., Rühlmann, A., Schubert, P., Biemann, S., Girhard, M., Hutter, M. C., Bernhardt, R., & Urlacher, V. B. (2017). Engineering of versatile redox partner fusions that support monooxygenase activity of functionally diverse cytochrome P450s. *Scientific Reports*, *7*(1), 9570. https://doi.org/10.1038/s41598-017-10075-w

Bassalo, M. C., Garst, A. D., Halweg-Edwards, A. L., Grau, W. C., Domaille, D. W., Mutalik, V. K., Arkin, A. P., & Gill, R. T. (2016). Rapid and efficient one-step metabolic pathway integration in *E. coli*. *ACS Synthetic Biology*, *5*(7), 561–568. https://doi.org/10.1021/acssynbio.5b00187

Belcher, J., McLean, K. J., Matthews, S., Woodward, L. S., Fisher, K., Rigby, S. E. J., Nelson, D. R., Potts, D., Baynham, M. T., Parker, D. A., Leys, D., & Munro, A. W. (2014). Structure and biochemical properties of the alkene producing cytochrome P450 OleTJE (CYP152L1) from the *Jeotgalicoccus* sp. 8456 bacterium. *Journal of Biological Chemistry*, *289*(10), 6535–6550. https://doi.org/10.1074/jbc.M113.527325

Birnbaum, S., & Bailey, J. E. (1991). Plasmid presence changes the relative levels of many host cell proteins and ribosome components in recombinant *Escherichia coli*. *Biotechnology and Bioengineering*, *37*(8), 736–745. https://doi.org/10.1002/bit.260370808

Bokel, A., Rühlmann, A., Hutter, M. C., & Urlacher, V. B. (2020). Enzyme-mediated two-step regio- and stereoselective synthesis of potential rapid-acting antidepressant (2*S*,6*S*)-hydroxynorketamine. *ACS Catalysis*, *10*(7), 4151–4159. https://doi.org/10.1021/acscatal.9b05384

Brixius-Anderko, S., Schiffer, L., Hannemann, F., Janocha, B., & Bernhardt, R. (2015). A CYP21A2 based whole-cell system in *Escherichia coli* for the biotechnological production of premedrol. *Microbial Cell Factories*, *14*, 135. https://doi.org/10.1186/s12934-015-0333-2

Bryant, J. A., Sellars, L. E., Busby, S. J. W., & Lee, D. J. (2014). Chromosome position effects on gene expression in *Escherichia coli* K-12. *Nucleic Acids Research*, *42*(18), 11383–11392. https://doi.org/10.1093/nar/gku828

Chung, M. E., Yeh, I. H., Sung, L. Y., Wu, M. Y., Chao, Y. P., Ng, I. S., & Hu, Y. C. (2017). Enhanced integration of large DNA into *E. coli* chromosome by CRISPR/Cas9. *Biotechnology and Bioengineering*, *114*(1), 172–183. https://doi.org/10.1002/bit.26056

Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, *97*(12), 6640–6645. https://doi.org/10.1073/pnas.120163297

Dong, H., Cui, Y., & Zhang, D. (2021). CRISPR/Cas technologies and their applications in *Escherichia coli*. *Frontiers in Bioengineering and Biotechnology*, *9*, 762676. https://doi.org/10.3389/fbioe.2021.762676

Dumon-Seignovert, L., Cariot, G., & Vuillard, L. (2004). The toxicity of recombinant proteins in *Escherichia coli*: A comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expression and Purification*, *37*(1), 203–206. https://doi.org/10.1016/j.pep.2004.04.025

Englaender, J. A., Jones, J. A., Cress, B. F., Kuhlman, T. E., Linhardt, R. J., & Koffas, M. A. G. (2017). Effect of genomic integration location on heterologous protein expression and metabolic engineering in *E. coli*. *ACS Synthetic Biology*, *6*(4), 710–720. https://doi.org/10.1021/acssynbio.6b00350

Fessner, N. D. (2019). P450 monooxygenases enable rapid late-stage diversification of natural products *via* C-H bond activation. *ChemCatChem*, *11*(9), 2226–2242. https://doi.org/10.1002/cctc.201801829

Girhard, M., Klaus, T., Khatri, Y., Bernhardt, R., & Urlacher, V. B. (2010). Characterization of the versatile monooxygenase CYP109B1 from *Bacillus subtilis*. *Applied Microbiology and Biotechnology*, *87*(2), 595–607. https://doi.org/10.1007/s00253-010-2472-z

Goormans, A. R., Snoeck, N., Decadt, H., Vermeulen, K., Peters, G., Coussement, P., Van Herpe, D., Beauprez, J. J., De Maeseneire, S. L., & Soetaert, W. K. (2020). Comprehensive study on *Escherichia coli* genomic expression: Does position really matter? *Metabolic Engineering*, *62*, 10–19. https://doi.org/10.1016/j.ymben.2020.07.007

Gottesman, S. (1996). Proteases and their targets in *Escherichia coli*. *Annual Review of Genetics*, *30*(1), 465–506. https://doi.org/10.1146/annurev.genet.30.1.465

Hannemann, F., Bichet, A., Ewen, K. M., & Bernhardt, R. (2007). Cytochrome P450 systems—Biological variations of electron transport chains. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1770*(3), 330–344. https://doi.org/10.1016/j.bbagen.2006.07.017

Hayashi, K., Morooka, N., Yamamoto, Y., Fujita, K., Isono, K., Choi, S., Ohtsubo, E., Baba, T., Wanner, B. L., Mori, H., & Horiuchi, T. (2006). Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. *Molecular Systems Biology*, *2*, 20060007. https://doi.org/10.1038/msb4100049

Hoang, T. T., Ma, Y., Stern, R. J., McNeil, M. R., & Schweizer, H. P. (1999). Construction and use of low-copy number T7 expression vectors for purification of problem proteins: Purification of *Mycobacterium tuberculosis* RmlD and *Pseudomonas aeruginosa* LasI and RhlI proteins, and functional analysis of purified RhlI. *Gene*, *237*(2), 361–371. https://doi.org/10.1016/S0378-1119(99)00331-5

Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., & Yang, S. (2015). Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system.

Applied and Environmental Microbiology, *81*(7), 2506–2514. https://doi.org/10.1128/AEM.04023-14

Kang, D. G., Kim Y. K., & Cha, H. J. (2002). Comparison of green fluorescent protein expression in two industrial *Escherichia coli* strains, BL21 and W3110, under co-expression of bacterial hemoglobin. *Applied Microbiology and Biotechnology*, *59*(4-5), 523–528. https://doi.org/10.1007/s00253-002-1043-3

Khatri, Y., Schifrin, A., & Bernhardt, R. (2017). Investigating the effect of available redox protein ratios for the conversion of a steroid by a myxobacterial CYP260A1. *FEBS Letters*, *591*(8), 1126–1140. https://doi.org/10.1002/1873-3468.12619

Labun, K., Montague, T. G., Krause, M., Torres Cleuren, Y. N., Tjeldnes, H., & Valen, E. (2019). CHOPCHOP v3: Expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research*, *47*(W1), W171–W174. https://doi.org/10.1093/nar/gkz365

Li, Q., Sun, B., Chen, J., Zhang, Y., Jiang, Y., & Yang, S. (2021). A modified pCas/pTargetF system for CRISPR-Cas9-assisted genome editing in *Escherichia coli*. *Acta Biochimica et Biophysica Sinica*, *53*(5), 620–627. https://doi.org/10.1093/abbs/gmab036

Li, S., Du, L., & Bernhardt, R. (2020). Redox partners: Function modulators of bacterial P450 enzymes. *Trends in Microbiology*, *28*(6), 445–454. https://doi.org/10.1016/j.tim.2020.02.012

Li, Y., Yan, F., Wu, H., Li, G., Han, Y., Ma, Q., Fan, X., Zhang, C., Xu, Q., Xie, X., & Chen, N. (2019). Multiple-step chromosomal integration of divided segments from a large DNA fragment via CRISPR/Cas9 in *Escherichia coli*. *Journal of Industrial Microbiology and Biotechnology*, *46*(1), 81–90. https://doi.org/10.1007/s10295-018-2114-5

Liu, X., Li, F., Sun, T., Guo, J., Zhang, X., Zheng, X., Du, L., Zhang, W., Ma, L., & Li, S. (2022). Three pairs of surrogate redox partners comparison for Class I cytochrome P450 enzyme activity reconstitution. *Communications Biology*, *5*(1), 791. https://doi.org/10.1038/s42003-022-03764-4

Mairhofer, J., Scharl, T., Marisch, K., Cserjan-Puschmann, M., & Striedner, G. (2013). Comparative transcription profiling and in-depth characterization of plasmid-based and plasmid-free *Escherichia coli* expression systems under production conditions. *Applied and Environmental Microbiology*, *79*(12), 3802–3812. https://doi.org/10.1128/AEM.00365-13

Merck KGaA. (2011). *pET System Manual* (*11th Edition*). Merck KGaA.

Miroux, B., & Walker, J. E. (1996). Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of Molecular Biology*, *260*(3), 289–298. https://doi.org/10.1006/jmbi.1996.0399

Okshevsky, M., Xu, Y., Masson, L., & Arbour, M. (2023). Efficiency of CRISPR-Cas9 genetic engineering in *Escherichia coli* BL21 is impaired by lack of Lon protease. *Journal of Microbiological Methods*, *204*, 106648. https://doi.org/10.1016/j.mimet.2022.106648

Omura, T., & Sato, R. (1964a). The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. *Journal of Biological Chemistry*, *239*(7), 2370–2378. https://www.ncbi.nlm.nih.gov/pubmed/14209971

Omura, T., & Sato, R. (1964b). The carbon monoxide-binding pigment of liver microsomes. II. Solubilization, purification, and properties. *Journal of Biological Chemistry*, *239*(7), 2379–2385. https://www.ncbi.nlm.nih.gov/pubmed/14209972

Ou, B., Garcia, C., Wang, Y., Zhang, W., & Zhu, G. (2018). Techniques for chromosomal integration and expression optimization in *Escherichia coli*. *Biotechnology and Bioengineering*, *115*(10), 2467–2478. https://doi.org/10.1002/bit.26790

Park, J. W., Lee, J. K., Kwon, T. J., Yi, D. H., Kim, Y. J., Moon, S. H., Suh, H. H., Kang, S. M., & Park, Y. I. (2003). Bioconversion of compactin into pravastatin by *Streptomyces* sp. *Biotechnology Letters*, *25*(21), 1827–1831. https://doi.org/10.1023/A:1026281914301

Petzoldt, K., Annen, K., Laurent, H., & Wiechert, R. (1982). Process for the preparation of 11 β-hydroxy steroidsUS Pat. 4,353,985, Schering Aktiengesellschaft.

Pflug, S., Richter, S. M., & Urlacher, V. B. (2007). Development of a fed-batch process for the production of the cytochrome P450 monooxygenase CYP102A1 from *Bacillus megaterium* in *E. coli*. *Journal of Biotechnology*, *129*(3), 481–488. https://doi.org/10.1016/j.jbiotec.2007.01.013

Pontrelli, S., Chiu, T. Y., Lan, E. I., Chen, F. Y. H., Chang, P., & Liao, J. C. (2018). *Escherichia coli* as a host for metabolic engineering. *Metabolic Engineering*, *50*, 16–46. https://doi.org/10.1016/j.ymben.2018.04.008

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, *152*(5), 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022

Reisch, C. R., & Prather, K. L. J. (2015). The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in *Escherichia coli*. *Scientific Reports*, *5*, 15096. https://doi.org/10.1038/srep15096

Richter, D., Bayer, K., Toesko, T., & Schuster, S. (2019). ZeBRα a universal, multi-fragment DNA-assembly-system with minimal hands-on time requirement. *Scientific Reports*, *9*(1), 2980. https://doi.org/10.1038/s41598-019-39768-0

Ro, D. K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Kirby, J., Chang, M. C. Y., Withers, S. T., Shiba, Y., Sarpong, R., & Keasling, J. D. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, *440*(7086), 940–943. https://doi.org/10.1038/nature04640

SaiSree, L., Reddy, M., & Gowrishankar, J. (2001). IS*186* insertion at a hot spot in the *lon* promoter as a basis for lon protease deficiency of *Escherichia coli* B: Identification of a consensus target sequence for IS*186* transposition. *Journal of Bacteriology*, *183*(23), 6943–6946. https://doi.org/10.1128/JB.183.23.6943-6946.2001

Seo, J.-H., Baek, S.-W., Lee, J., & Park, J.-B. (2017). Engineering *Escherichia coli* BL21 genome to improve the heptanoic acid tolerance by using CRISPR-Cas9 system. *Biotechnology and Bioprocess Engineering*, *22*(3), 231–238. https://doi.org/10.1007/s12257-017-0158-4

Studier, F. W. (2014). Stable expression clones and auto-induction for protein production in *E. coli*. In *Structural Genomics: General Applications* (pp. 17–32). Humana Press. https://doi.org/10.1007/978-1-62703-691-7_2

Studier, F. W., & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology*, *189*(1), 113–130. https://doi.org/10.1016/0022-2836(86)90385-2

Sudhamsu, J., Kabir, M., Airola, M. V., Patel, B. A., Yeh, S. R., Rousseau, D. L., & Crane, B. R. (2010). Co-expression of ferrochelatase allows for complete heme incorporation into recombinant proteins produced in *E. coli*. *Protein Expression and Purification*, *73*(1), 78–82. https://doi.org/10.1016/j.pep.2010.03.010

Ting, W.-W., Tan, S.-I., & Ng, I. S. (2020). Development of chromosome-based T7 RNA polymerase and orthogonal T7 promoter circuit in *Escherichia coli* W3110 as a cell factory. *Bioresources and Bioprocessing*, *7*(1), 54. https://doi.org/10.1186/s40643-020-00342-6

Urlacher, V. B., & Girhard, M. (2019). Cytochrome P450 monooxygenases in biotechnology and synthetic biology. *Trends in Biotechnology*, *37*(8), 882–897. https://doi.org/10.1016/j.tibtech.2019.01.001

von Bühler, C., Le-Huu, P., & Urlacher, V. B. (2013). Cluster screening: An effective approach for probing the substrate space of uncharacterized cytochrome P450s. *ChemBioChem*, *14*(16), 2189–2198. https://doi.org/10.1002/cbic.201300271

Wang, J., Niyompanich, S., Tai, Y. S., Wang, J., Bai, W., Mahida, P., Gao, T., & Zhang, K. (2016). Engineering of a highly efficient *Escherichia coli* strain for mevalonate fermentation through chromosomal integration. *Applied and Environmental Microbiology*, *82*(24), 7176–7184. https://doi.org/10.1128/AEM.02178-16

Yoon, S., Han, M.-J., Jeong, H., Lee, C., Xia, X.-X., Lee, D.-H., Shim, J., Lee, S., Oh, T., & Kim, J. F. (2012). Comparative multi-omics systems analysis of *Escherichia coli* strains B and K-12. *Genome Biology*, *13*(5), R37. https://doi.org/10.1186/gb-2012-13-5-r37

Zanos, P., Moaddel, R., Morris, P. J., Georgiou, P., Fischell, J., Elmer, G. I., Alkondon, M., Yuan, P., Pribut, H. J., Singh, N. S., Dossou, K. S. S., Fang, Y., Huang, X. P., Mayo, C. L., Wainer, I. W., Albuquerque, E. X., Thompson, S. M., Thomas, C. J., Zarate Jr., C. A., & Gould, T. D. (2016). NMDAR inhibition-independent antidepressant actions of ketamine metabolites. *Nature*, *533*(7604), 481–486. https://doi.org/10.1038/nature17998

Zanos, P., Moaddel, R., Morris, P. J., Riggs, L. M., Highland, J. N., Georgiou, P., Pereira, E. F. R., Albuquerque, E. X., Thomas, C. J., & Zarate, Jr. C. A., & Gould, T. D. (2018). Ketamine and ketamine metabolite pharmacology: Insights into therapeutic mechanisms. *Pharmacological Reviews*, *70*(3), 621–660. https://doi.org/10.1124/pr.117.015198

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Luelf, U. J., Böhmer, L. M., Li, S., & Urlacher, V. B. (2023). Effect of chromosomal integration on catalytic performance of a multi-component P450 system in *Escherichia coli*. *Biotechnology and Bioengineering*, 120, 1762–1772. https://doi.org/10.1002/bit.28404

2.2.1 Supporting information

Supporting information

Effect of chromosomal integration on catalytic performance of a multi-component P450 system in *E. coli*

U. Joost Luelf[1], Lisa M. Böhmer[1], Shengying Li[2], Vlada B. Urlacher[1]*

[1]Institute of Biochemistry, Heinrich-Heine University Düsseldorf, 40225 Düsseldorf, Germany

[2]State Key Laboratory of Microbial Technology, Shandong University, Qingdao, Shandong, 266237, China

*Corresponding Author

Email: vlada.urlacher@uni-duesseldorf.de

ORCID: Vlada B. Urlacher 0000-0003-1312-4574

Running title: Chromosomally integrated P450 systems

Supporting Tables

Table S1. Primer sequences. Restriction sites are underlined. N_{20} targeting sequence are shown in italic.

Name	Sequence (5‘ to 3‘)	Reference
gRNA Primer		
gRNA rev	ATATATACTAGTATTATACCTAGGACTGAGCTAG	general design (Seo et al., 2017)
gRNA_XXX fw (general primer design)	ATATATACTAGT-*N_{20}*-GTTTTAGAGCTAGAAATAGC	general design (Seo et al., 2017)
gRNA _001 fw (nupG BL21(DE3))	ATATATACTAGT*ACCAGCAGTACGGTGCCGAA*GTTTTAGAGCTAGAAATAGC	
gRNA_002 fw (nupG K-12)	ATATATACTAGT*ACCAGCAGTACAGTACCGAA*GTTTTAGAGCTAGAAATAGC	N_{20} sequence (Bassalo et al., 2016)
gRNA_003 fw (ybhB_ybhC K-12)	ATATATACTAGT*CTAACTTGAGCGAAACGGGA*GTTTTAGAGCTAGAAATAGC	
gRNA_004 fw (atpI_gidB)	ATATATACTAGT*TTATTAAAAATGTCAATGGG*GTTTTAGAGCTAGAAATAGC	
pgRNA seq	GCCACCTGACGTCTAAG	
Primers for amplification and sequencing of *lopA* region		
lopA fw	GAAGCGCAACAGGCATC	
lopA rev	AAGCCCGAATTAGCCTGC	
Primers for synthesis of dsDonorDNA		
HA1 fw (ybhB_ybhC)	GCTCATGCCACCATCAAG	
HA1 rev (ybhB_ybhC)	TTCATAGCCCGGAGCAAC	
HA2 fw (ybhB_ybhC)	GCCATCTGGCAGAGTGATTAAC	
HA2 rev (ybhB_ybhC)	GCGCGGTACAGGATGAAAC	
T7RNAP fw + Overlap HA1 (ybhB_ybhC)	GTTGCTCCGGGCTATGAAATACGCAAACCGCCTCTC	
T7RNAP rev + Overlap HA2 (ybhB_ybhC)	GTTAATCACTCTGCCAGATGGCTTACGCGAACGCGAAGTC	
HA1 fw (atpI_gidB)	GCCACCACCAGTAACAC	
HA1 rev (atpI_gidB)	TGATCGAACAGGGTTAGC	
HA2 fw (atpI_gidB)	CAGCCAATGATGGTTCTTAGC	
HA2 rev (atpI_gidB)	CGTCAGGTGCAACATGAG	
pET(+) fw + Overlap HA2 (atpI_gidB)	GCTAACCCTGTTCGATCATTAATGCGCCGCTACAGG	
pET(+) rev + Overlap HA1 (atpI_gidB)	GCTAAGAACCATCATTGGCTGTATAGGCGCCAGCAACC	
pET rev + Overlap HA1 (atpI_gidB)	GCTAACCCTGTTCGATCATGGTTATGCCGGTACTGC	
pCOLADuet fw + HA2 fw (atpI_gidB)	GCTAAGAACCATCATTGGCTGATGCGACTCCTGCATTAG	
pCOLADuet rev + HA1 rev (atpI_gidB)	GCTAACCCTGTTCGATCATGACCGTGTGCTTCTC	
HA1 fw (nupG)	TGCCAACGCTTGGGTTAATC	

Name	Sequence (5‘ to 3‘)	Reference
HA1 rev (nupG)	GGATGGTCAGAATGAACAGGG	
HA2 fw (nupG)	CTAACGGCTTCGGCTGTATC	
HA2 rev (nupG)	TAGCGCTTATATTCGCGGTG	
pET(+) fw + Overlap HA2 (nupG)	GATACAGCCGAAGCCGTTAGTTAATGCGCCGCTACAGG	
pET(+) rev + Overlap HA1 (nupG)	CCCTGTTCATTCTGACCATCCTATAGGCGCCAGCAACC	
pET fw + HA2 (nupG)	GATACAGCCGAAGCCGTTAGTGGTTATGCCGGTACTGC	
pCOLA Duet fw + HA1 (nupG)	CCCTGTTCATTCTGACCATCCATGCGACTCCTGCATTAG	
pCOLA Duet rev + HA2 (nupG)	GATACAGCCGAAGCCGTTAGTGACCGTGTGCTTCTC	

Table S2. Strains used and created. Triple mutant I238Q G239A M388A (QAA) of *cyp154e1* was used for all experiments.

Strain	Purpose	Source
E. coli BL21(DE3)	expression and genome engineering	Novagen
E. coli DH5α	cloning procedures	Clontech
E. coli JM109	genome engineering	NEB
E. coli W3110	genome engineering	DSMZ
E. coli JM109(T7)	expression and genome engineering	this study
E. coli W3110(T7)	expression and genome engineering	this study
E. coli BL21(DE3) CYP154E1 QAA (nupG)	expression and genome engineering	this study
E. coli BL21(DE3) CYP154E1 QAA (atpI_gidB)	expression and genome engineering	this study
E. coli JM109(T7) CYP154E1 QAA (nupG)	expression and genome engineering	this study
E. coli JM109(T7) CYP154E1 QAA (atpI_gidB)	expression and genome engineering	this study
E. coli W3110(T7) CYP154E1 QAA (nupG)	expression and genome engineering	this study
E. coli W3110(T7) CYP154E1 QAA (atpI_gidB)	expression and genome engineering	this study
E. coli BL21(DE3) CYP154E1 QAA (nupG) YkuN (atpI_gidB)	expression	this study
E. coli BL21(DE3) CYP154E1 QAA (nupG) YkuN+FdR (atpI_gidB)	expression	this study
E. coli BL21(DE3) CYP154E1 QAA (atpI_gidB) YkuN (nupG)	expression	this study
E. coli BL21(DE3) CYP154E1 QAA (atpI_gidB) YkuN+FdR (nupG)	expression	this study
E. coli W3110(T7) CYP154E1 QAA (nupG) YkuN (atpI_gidB)	expression	this study
E. coli W3110(T7) CYP154E1 QAA (nupG) YkuN+FdR (atpI_gidB)	expression	this study
E. coli W3110(T7) CYP154E1 QAA (atpI_gidB) YkuN (nupG)	expression	this study
E. coli W3110(T7) CYP154E1 QAA (atpI_gidB) YkuN+FdR (nupG)	expression	this study

Table S3. Plasmids used and created. Triple mutant I238Q G239A M388A (QAA) of *cyp154e1* was used for all experiments.

Plasmid	Purpose	Source
pET-28a(+) CYP154E1 QAA	expression and PCR template	(Bokel et al., 2020)
pET-22b(+) CYP154E1 QAA	expression and PCR template	(Bokel et al., 2020)
pET-16b YkuN	PCR template	(Girhard et al., 2010)
pCOLADuet YkuN(I) FdR(II)	expression and PCR template	(Bokel et al., 2020)
pCas	genome engineering in K-12 strains	(Jiang et al., 2015)
pEcCas	genome engineering in K-12 and B strains	(Li et al., 2021)
pgRNA-bacteria	template for N_{20} exchange	(Qi et al., 2013)
pgRNA_001	targets *nupG* locus in B strains	this study
pgRNA_002	targets *nupG* locus in K-12 strains	this study
pgRNA_003	targets *ybhB_ybhC* locus for integration of *T7 gene 1*	this study
pgRNA_004	targets *atpI_gidB* locus in both B and K-12 strains	this study

Supporting Figures

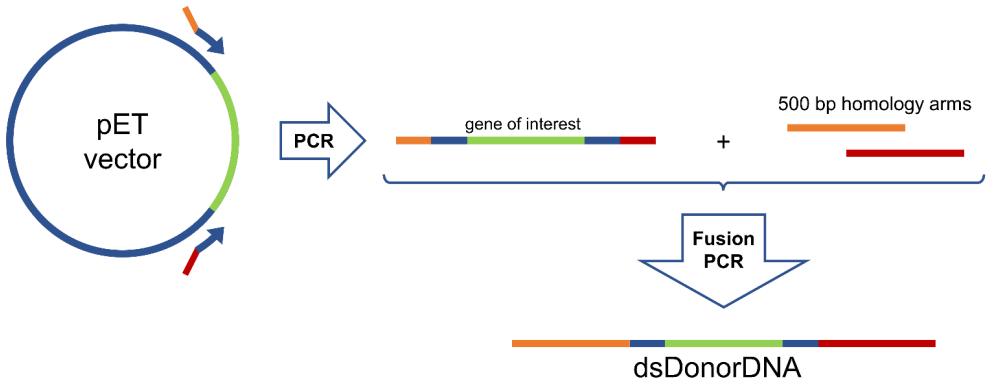


Figure S1. PCR-based synthesis of double-stranded donor DNA (dsDonorDNA) for homologous recombination. Primers (blue) were designed to bind to any pET vector (containing a f1 origin of replication) upstream and downstream of the multiple cloning site which harbors the gene to be integrated in order to amplify the whole operon including promoter and terminator (green). Introduced overlaps match the locus-dependent homology arms for homologous recombination amplified from the genome (orange/red) and enable fusion PCR of all three fragments.

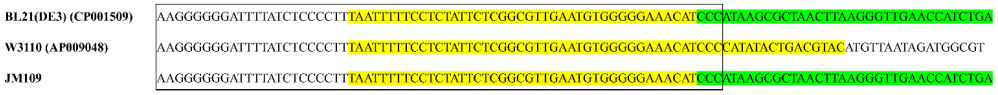


Figure S2: Sequence alignment of *lon* transcription site of strains *E. coli* BL21(DE3) (GenBank Accession no. CP001509.3), *E. coli* W3110 (GenBank accession no. AP009048.1), and *E. coli* JM109 (sequencing results from this study). Box represents sequence homology in all three strains. Lon promoter (yellow) is disrupted in BL21(DE3) and JM109 by IS*186* transposon insertion (green) which leads to protease deficiency (SaiSree et al., 2001).

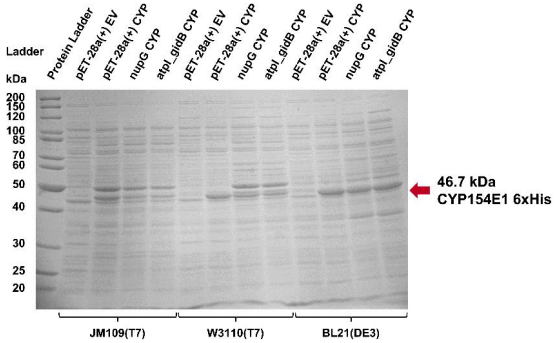


Figure S3: SDS-PAGE analysis of the soluble protein fractions after expression of *cyp154e1m* (cf. Figure 2).

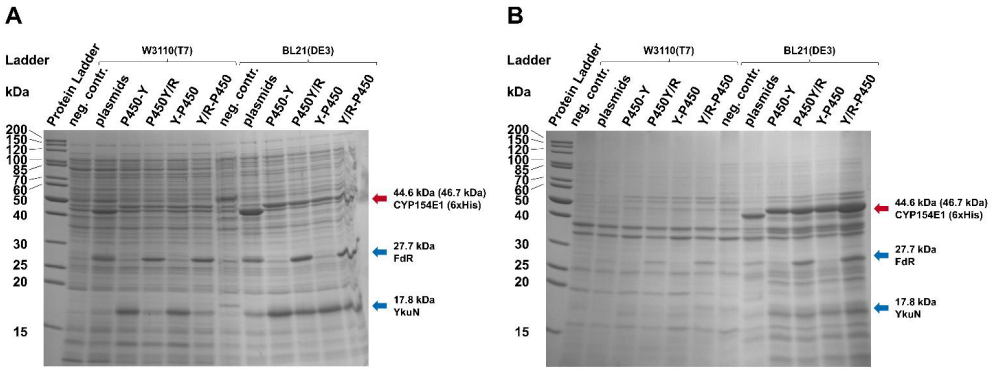


Figure S4: SDS-PAGE analysis of the soluble (A) and insoluble (B) protein fractions after coexpression of *cyp154e1m*, *ykuN* and *fdr* (cf. Figure 3).

References

Bassalo, M. C., Garst, A. D., Halweg-Edwards, A. L., Grau, W. C., Domaille, D. W., Mutalik, V. K., Arkin, A. P., & Gill, R. T. (2016). Rapid and Efficient One-Step Metabolic Pathway Integration in *E. coli*. *ACS Synthetic Biology*, *5*(7), 561-568. https://doi.org/10.1021/acssynbio.5b00187

Bokel, A., Rühlmann, A., Hutter, M. C., & Urlacher, V. B. (2020). Enzyme-Mediated Two-Step Regio- and Stereoselective Synthesis of Potential Rapid-Acting Antidepressant (2*S*,6*S*)-Hydroxynorketamine. *ACS Catalysis*, *10*(7), 4151-4159. https://doi.org/10.1021/acscatal.9b05384

Girhard, M., Klaus, T., Khatri, Y., Bernhardt, R., & Urlacher, V. B. (2010). Characterization of the versatile monooxygenase CYP109B1 from *Bacillus subtilis*. *Applied Microbiology and Biotechnology*, *87*(2), 595-607. https://doi.org/10.1007/s00253-010-2472-z

Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., & Yang, S. (2015). Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Applied and Environmental Microbiology*, *81*(7), 2506-2514. https://doi.org/10.1128/AEM.04023-14

Li, Q., Sun, B., Chen, J., Zhang, Y., Jiang, Y., & Yang, S. (2021). A modified pCas/pTargetF system for CRISPR-Cas9-assisted genome editing in *Escherichia coli*. *Acta Biochimica et Biophysica Sinica*, *53*(5), 620-627. https://doi.org/10.1093/abbs/gmab036

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, *152*(5), 1173-1183. https://doi.org/10.1016/j.cell.2013.02.022

SaiSree, L., Reddy, M., & Gowrishankar, J. (2001). IS186 insertion at a hot spot in the lon promoter as a basis for lon protease deficiency of *Escherichia coli* B: identification of a consensus target sequence for IS186 transposition. *Journal of Bacteriology*, *183*(23), 6943-6946. https://doi.org/10.1128/JB.183.23.6943-6946.2001

Seo, J.-H., Baek, S.-W., Lee, J., & Park, J.-B. (2017). Engineering *Escherichia coli* BL21 genome to improve the heptanoic acid tolerance by using CRISPR-Cas9 system. *Biotechnology and Bioprocess Engineering*, *22*(3), 231-238. https://doi.org/10.1007/s12257-017-0158-4

2.3 Chromosomal integration for the synthesis of pinoresinol in growing cells

Title: Plasmid-free production of the plant lignan pinoresinol in growing *Escherichia coli* cells

Authors: **U. J. Luelf**, A. Wassing, L. M. Böhmer, V. B. Urlacher

Published in: Microbial Cell Factories, 23, 289, 1-12

DOI: 10.1186/s12934-024-02562-3

Contribution: Conceptualized the study, designed and conducted most of the experiments, analyzed the data and prepared the first draft.

Luelf *et al. Microbial Cell Factories* (2024) 23:289
https://doi.org/10.1186/s12934-024-02562-3

Microbial Cell Factories

RESEARCH Open Access

Plasmid-free production of the plant lignan pinoresinol in growing *Escherichia coli* cells

U. Joost Luelf[1], Alexander Wassing[1], Lisa M. Böhmer[1] and Vlada B. Urlacher[1*]

Abstract

Background The high-value aryl tetralin lignan (+)-pinoresinol is the main precursor of many plant lignans including (-)-podophyllotoxin, which is used for the synthesis of chemotherapeutics. As (-)-podophyllotoxin is traditionally isolated from endangered and therefore limited natural sources, there is a particular need for biotechnological production. Recently, we developed a reconstituted biosynthetic pathway from (+)-pinoresinol to (-)-deoxypodophyllotoxin, the direct precursor of (-)-podophyllotoxin, in the recombinant host *Escherichia coli*. However, the use of the expensive substrate (+)-pinoresinol limits its application from the economic viewpoint. In addition, the simultaneous expression of multiple heterologous genes from different plasmids for a multi-enzyme cascade can be challenging and limits large-scale use.

Results In this study, recombinant plasmid-free *E. coli* strains for the multi-step synthesis of pinoresinol from ferulic acid were constructed. To this end, a simple and versatile plasmid toolbox for CRISPR/Cas9-assisted chromosomal integration has been developed, which allows the easy transfer of genes from the pET vector series into the *E. coli* chromosome. Two versions of the developed toolbox enable the efficient integration of either one or two genes into intergenic high expression loci in both *E. coli* K-12 and B strains. After evaluation of this toolbox using the fluorescent reporter mCherry, genes from *Petroselinum crispum* and *Zea mays* for the synthesis of the monolignol coniferyl alcohol were integrated into different *E. coli* strains. The product titers achieved with plasmid-free *E. coli* W3110(T7) were comparable to those of the plasmid-based expression system. For the subsequent oxidative coupling of coniferyl alcohol to pinoresinol, a laccase from *Corynebacterium glutamicum* was selected. Testing of different culture media as well as optimization of gene copy number and copper availability for laccase activity resulted in the synthesis of 100 mg/L pinoresinol using growing *E. coli* cells.

Conclusions For efficient and simple transfer of genes from pET vectors into the *E. coli* chromosome, an easy-to-handle molecular toolbox was developed and successfully tested on several *E. coli* strains. By combining heterologous and endogenous enzymes of the host, a plasmid-free recombinant *E. coli* growing cell system has been established that enables the synthesis of the key lignan pinoresinol.

Keywords Phenylpropanoid, Coniferyl alcohol, Ferulic acid, Lignan, Pinoresinol, Oxidative coupling, *E. coli*, Chromosome, CRISPR/Cas

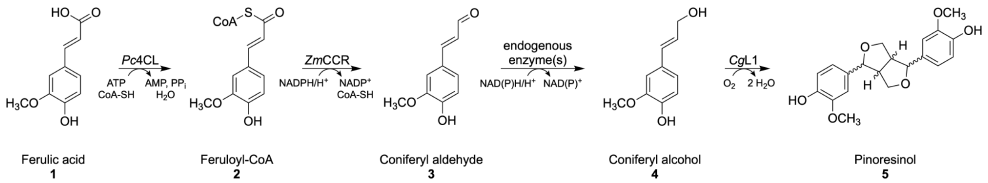
*Correspondence:
Vlada B. Urlacher
vlada.urlacher@uni-duesseldorf.de
[1]Institute of Biochemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

Background

Lignans are a large and structurally diverse group of pharmacologically active secondary plant metabolites which are formed via oxidative dimerization of phenylpropanoids [1, 2]. In mammals, dietary intake of lignans from plants or plant-derived foods leads to the formation of enterolignans by the gut microbiome, which are classified as phytoestrogens and show health-promoting effects [3]. Furthermore, the chemotherapeutic drugs etoposide and teniposide are semi-synthetically produced from the aryl tetralin lignan (-)-podophyllotoxin which is isolated from the endangered plant *Podophyllum hexandrum* Royle [4]. The low content of lignans in plants combined with a high demand and their structural complexity leads to a particular need for biotechnological solutions to produce these natural products [5, 6]. The biosynthesis starts from cinnamic acid derivatives which originate from the shikimate and phenylpropanoid pathways. Depending on the phenylpropanoid monomer, different lignans are formed. The major subclass of furofuran lignans, including the title compound pinoresinol **5**, results from the dimerization of the monolignol coniferyl alcohol **4** [1]. Besides being a starting precursor for higher lignans, pinoresinol **5** possesses multiple biological activities including anti-inflammatory [7] and antioxidant activity, which also explains its hepatoprotective [8] and chemopreventive [9] properties. Furthermore, antifungal [10] and putative hypoglycemic properties [11] add to the list of pharmacological activities of interest. The above-mentioned chemotherapeutics etoposide and teniposide are derived from pinoresinol **5** via (-)-podophyllotoxin, indicating potentially high demand of this high-value compound. A first step towards the biotechnological production of these drugs in recombinant microorganisms was recently taken by our group by synthesizing (-)-deoxypodophyllotoxin from (+)-pinoresinol in the heterologous host *Escherichia coli* [12]. However, the use of the high-value compound (+)-pinoresinol as a starting material remains an economic challenge. (+)-Pinoresinol has been isolated from plants with high enantiopurity but low efficiency [13]. As the subsequent reactions from pinoresinol **5** towards (-)-deoxypodophyllotoxin are catalyzed by enantioselective enzymes, racemic pinoresinol **5** instead of (+)-pinoresinol could be used as starting substrate [14, 15]. Thus, in this study, we engineered and optimized growing *E. coli* cells for the production of racemic pinoresinol **5**. Starting from relatively inexpensive and readily available ferulic acid **1**, we reconstituted an artificial phenylpropanoid path in *E. coli* to first produce the monolignol coniferyl alcohol **4**. Ferulic acid **1** is reduced to coniferyl alcohol **4** via coniferyl aldehyde **3** in three enzymatic steps (Scheme 1). Two of the three steps are catalyzed by enzymes derived from the plants *Petroselinum crispum* and *Zea mays* [16, 17]. The final reduction step is catalyzed by endogenous *E. coli* alcohol dehydrogenases or aldo-keto reductases [18, 19]. For the synthesis of pinoresinol **5** from **4**, laccases [20–22] and peroxidases [23] have been successfully applied in the past. Previously, we have demonstrated that a laccase from *Corynebacterium glutamicum* gave the highest pinoresinol concentration in vitro and thus was used in this study [21]. In the reported reactions with isolated enzymes or in resting *E. coli* cells, pinoresinol concentrations ranged between several milligrams to grams per liter, depending on the complexity of the system, the starting substrate molecule, its concentration, and the number of reaction steps [14, 21–24].

In order to simplify the bioprocess design, growing cells were preferred instead of resting cells used for the production of (-)-deoxypodophyllotoxin [12] and its precursors [15] in our previous studies. In addition to the facts that enzymes are steadily produced during the reactions and that only one reaction vessel is required, the combination of cell growth and substrate conversion eliminates the need to prepare the cells for a resting cell approach and is therefore significantly less time-consuming.

Furthermore, plasmids were used for the co-expression of all genes in our previous work. Plasmids offer many advantages like easy cloning, fast and efficient transformation of different host strains, and high and tunable expression levels. However, there are also some shortcomings. As plasmids are extrachromosomal, non-essential DNA, a form of selectable markers has to be used to maintain the plasmids in the host cell and prevent



Scheme 1 Recombinant biosynthetic path from ferulic acid **1** to pinoresinol **5** in *E. coli*. The substrate is sequentially converted by a 4-coumarate-CoA ligase from *P. crispum* (*Pc*4CL), a cinnamoyl-CoA reductase from *Z. mays* (*Zm*CCR), *E. coli* alcohol dehydrogenases or aldo-keto reductases and a laccase from *C. glutamicum* (*Cg*L1)

plasmid loss. For this purpose, antibiotics are often utilized, which are cost-intensive on industrial scale, can affect the cell growth, and contribute to the development of multidrug-resistant human pathogens [25–28]. In recombinant strains with plasmids, the number of gene copies can vary from cell to cell even within stable populations, leading to altered expression levels [25, 28]. It should also be noted that the plasmid copy number can increase drastically after induction resulting in an increased metabolic burden [25, 27]. The transition from plasmid-based to chromosomal expression helps to avoid most of these issues. To this end, we developed an easy-to-handle toolbox that allows the simple transfer of genes from the often-used pET vector series into the *E. coli* chromosome based on CRISPR/Cas-assisted λ-Red-mediated homologous recombination [29, 30].

Since the expression levels of chromosomally integrated genes depend on their position in the genome, different integration sites can be used to tune expression levels. One general rule is the so-called 'gene dosage effect', which states that genes located closer to the chromosome's origin of replication (*oriC*) are statistically more abundant in the cell and therefore more highly expressed [31, 32]. In addition, there are regional effects that affect the transcription of integrated genes, resulting in expression levels that vary on average by 2-4-fold depending on the integration site, with rare outliers of ~ 300-fold deviation [31, 33, 34]. To achieve the highest possible integration efficiencies and expression levels, several previously uncharacterized integration loci present in both *E. coli* K-12 and B strains were first identified and analyzed using the fluorescent reporter mCherry.

Further, we aimed to apply a method that does not require the separate synthesis of linear donor DNA but instead allows the use of the same type II restriction enzymes as used for the cloning of pET vectors. The use of circular instead of linear donor DNA for the CRISPR/Cas-based recombination circumvents potential degradation by exonucleases [35]. In addition, the transformation of supercoiled plasmids is highly efficient and high-fidelity replication within the cell results in a higher intracellular repair template concentration compared to linear PCR-synthesized donor DNA [36]. The combination of easy cloning and characterized integration loci in both *E. coli* K-12 and B strains aims to address the biocatalysis community that already has pET-based libraries in hand and wants to take the step towards plasmid-free whole-cell biocatalysis.

In this study, the application of the developed toolbox led to a highly productive whole-cell system for pinoresinol **5** synthesis which was finally obtained after testing *E. coli* K-12 and B strains and optimizing culture media, gene copy number, and copper availability for laccase activity.

Results and discussion

Design of pgRNA-ET and pgRNA-DUET vector series for chromosomal integration

First, we identified several new intergenic integration sites present in both *E. coli* K-12 and B strains. To this end, information on the prototype *E. coli* K-12 strain MG1655 from its GenBank genome file (accession no. U00096.3), the information in RegulonDB [37] for promoter and terminator sets of this strain, and EcoCyc database [38] for detailed operon information, were combined. Additionally, a study by Vora et al. [39], who described the protein landscape of the genome and identified so-called *highly-expressed extended protein occupancy domains* (heEPODs) was included into the analysis as well. All selected integration sites should match the following criteria: (i) the influence on the expression of neighboring genes is minimized by avoiding integration downstream of a promoter or upstream of a terminator listed in RegulonDB [37], (ii) they can be assigned to the heEPODs described by Vora et al. [39], and (iii) they are at least 100 base pairs long to facilitate design of CRISPR/Cas targeting sequences as well as primer design for amplification of the homology arms for homologous recombination. A similar approach to identifying new integration loci, albeit with different search parameters, was described by Goormans et al. [34]. The homology arms of the donor DNA template were designed to be 500 bp in length, which results in higher integration efficiency compared to shorter variants [40, 41]. However, as our toolbox was designed to be used in both K-12 and B strains, 500 bp homology arms without single nucleotide polymorphisms had to be identified by comparing sequences of *E. coli* BL21(DE3) (GenBank accession no. CP001509.3) and the industrially relevant *E. coli* W3110 (GenBank accession no. AP009048.1). This led to the exclusion of several integration sites that met all previous search criteria. Finally, six integration loci were chosen for further experiments (Table S1), of which only one had been previously described in the literature [26, 34]. After cloning of the appropriate targeting sequences for CRISPR/Cas into the pgRNA plasmids, linear copies of the plasmids were amplified by PCR and fused with corresponding homology arms as well as the multiple cloning sites derived from either pET-28a(+) or pETDuet-1 to yield the pgRNA-ET or pgRNA-DUET vector series, respectively (Fig. 1).

Evaluation of selected integration loci

The selected integration sites were evaluated using the fluorescent reporter protein mCherry. The *mCherry* gene was cloned into the pgRNA-ET plasmid series and subsequently integrated into the chromosomes of *E. coli* BL21(DE3) and *E. coli* W3110(T7) to compare possible differences between these representative B and K-12

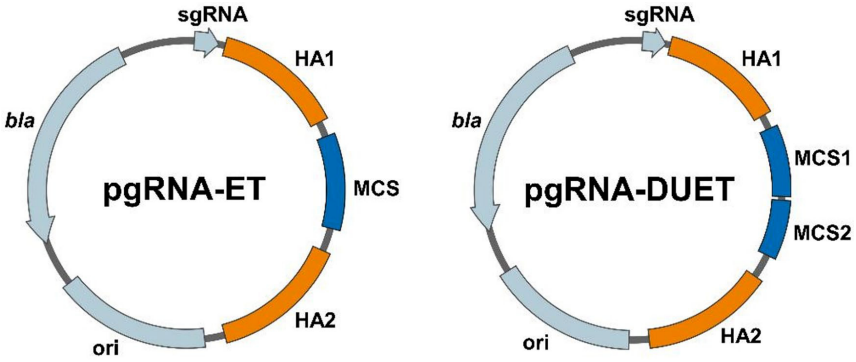


Fig. 1 Plasmid series pgRNA-ET (left) and pgRNA-DUET (right) for easy CRISPR/Cas-assisted transfer of genes from pET vectors into the *E. coli* chromosome. Each vector contains a pBR322 origin of replication (ori), a β-lactamase encoding gene (*bla*), a locus-specific single-guide RNA (sgRNA) as well as the corresponding homology arms (HA1 & HA2) for homologous recombination, and multiple cloning sites originating from pET-28a(+) or pETDuet-1 for pgRNAET and pgRNA-DUET, respectively

strains. Although *E. coli* BL21(DE3) is considered to be hard-to-edit [42, 43], high integration efficiencies of up to 100% of all screened colonies (n=20) were achieved (Table S1 & Figure S1). For reference purposes, *mCherry* was also integrated into the gene-disrupting *lacZ* locus which was previously described as a medium-to-low-expression locus [26, 34].

The expression of the chromosomally integrated *mCherry* gene showed no noticeable effect on cell growth (Figure S2). Compared to the *lacZ* reference locus, all *mCherry* integrations into the selected intergenic loci resulted in higher or at least similar fluorescence for both strains (Fig. 2).

In *E. coli* BL21(DE3), a 30–70% increase in fluorescence compared to *lacZ* was observed for all selected intergenic loci (p<0.02). For *E. coli* W3110(T7), only integration into *flgF_flgG* (70% increase, p<0.01) and *glmU_atpC* (20% increase, p=0.04) loci resulted in higher expression compared to *lacZ*. It should be noted that after gene integration into the six intergenic loci, only low basal non-induced expression was observed, while after integration into *lacZ*, high fluorescence was observed without addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) (Figure S3). A correlation between the fluorescence output and the distance from *oriC* - the above mentioned 'gene dosage effect' due to chromosome replication - could not be observed in our case. However, Bryant et

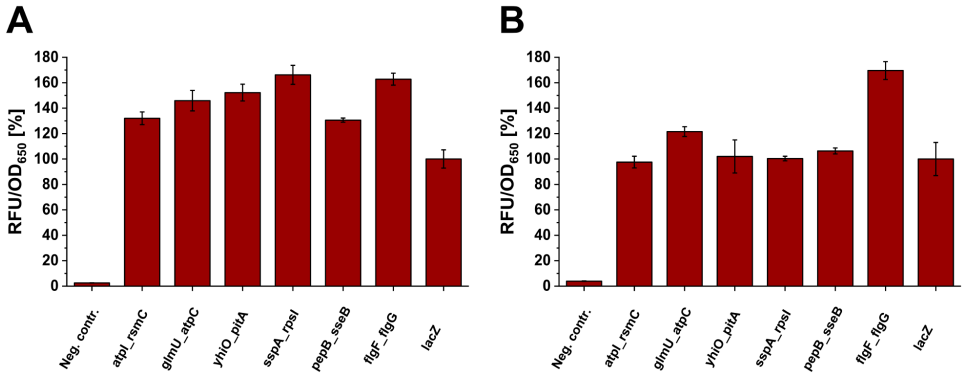


Fig. 2 Relative fluorescence normalized to optical density (OD_{650}) of cell cultures after chromosomal expression of *mCherry* integrated into different genomic loci in (A) *E. coli* BL21(DE3) and (B) *E. coli* W3110(T7). Strains harboring no copy of *mCherry* were used as negative controls to quantify intrinsic fluorescence of *E. coli*. Relative fluorescence was calculated in relation to *lacZ*. Means and standard deviations were calculated from three biological replicates

al. reported that the gene dosage effect only accounted for 1.4-fold differences in expression [31]. Out of the six intergenic loci tested, *atpI_rsmG* had been described as one of the loci leading to the highest expression levels [26, 34]. Three of the newly characterized loci reached even higher fluorescence in *E. coli* BL21(DE3) ($p<0.05$). It should be noted that unlike the gene-disrupting integration into the flagellar region 1 described by Juhas et al. [44], the integration between the genes *flgF* and *flgG* in *E. coli* W3110(T7) led to high expression combined with no loss of motility. Furthermore, plasmid-based expression using the pET-28a(+) vector showed that a single chromosomal copy reached up to 20% fluorescence compared to the vector which is reported to give about 40 gene copies per cell [45]. The *flgF_flgG* locus was chosen as the most promising integration site for further experiments due to the higher fluorescence observed in both strains and particularly in *E. coli* W3110(T7).

Synthesis of coniferyl alcohol

First, the genes coding for the 4-coumarate-CoA ligase of *P. crispum* (*Pc*4CL) and the cinnamoyl-CoA reductase of *Z. mays* (*Zm*CCR) were cloned into a pETDuet vector for plasmid-based co-expression and preliminary tests for the synthesis of coniferyl alcohol **4.** *E. coli* W3110(T7), a K-12 derivative and industrially used strain, carrying this plasmid, was tested in two different media, terrific broth (TB) and lysogeny broth (LB), for the conversion of 5 mM ferulic acid **1** with growing cells. The substrate was added 2 h after induction. It was found that higher coniferyl alcohol **4** titers were reached in TB medium compared to LB medium 18 h after substrate addition (Figure S4). Thus, all following experiments were carried out using TB medium.

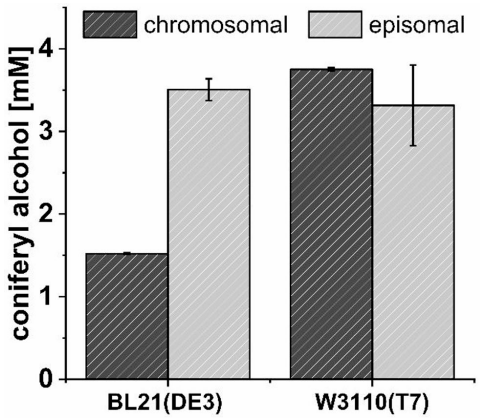


Fig. 3 Coniferyl alcohol **4** synthesized by different *E. coli* strains 18 h after addition of 5 mM ferulic acid **1**. Expression of chromosomally integrated genes was compared to a pETDuet vector in *E. coli* BL21(DE3) and W3110(T7). Means and standard deviation were calculated from three biological replicates

The same plasmid was then introduced into *E. coli* BL21(DE3) which is considered more suitable for heterologous expression due to its lack of Lon and OmpT proteases. Next, both heterologous genes were integrated into the *flgF_flgG* locus of both *E. coli* strains in a single step using the pgRNA-DUET vector. This allowed comparison of plasmid-based and plasmid-free expression as well as biotransformation in both strains.

Interestingly, the plasmid-based expression system performed equally well in both strains and yielded 3.5±0.1 mM and 3.3±0.5 mM coniferyl alcohol **4** (Fig. 3).

After chromosomal integration in BL21(DE3), however, only 1.52±0.01 mM **4** was obtained. In general, lower expression levels and a therefore lower overall activity of the growing *E. coli* cells were expected due to the reduced copy number compared to the plasmid-based system. A striking exception for this was *E. coli* W3110(T7), which showed a coniferyl alcohol **4** titer comparable to the plasmid-based expression (3.75±0.02 mM, 75% theoretical yield). The differences in product titers between plasmid-based and chromosomal expression systems can be mainly attributed to differences in heterologous expression of *ZmCCR* as visualized via Western Blot (Figure S5). Although *ZmCCR* appeared to be better expressed in *E. coli* BL21(DE3) compared to *E. coli* W3110(T7) when using plasmids, it was strongly affected by chromosomal integration, which reduced expression in this strain. In contrast, chromosomal integration in *E. coli* W3110(T7) did not result in lower protein concentration which may explain similar levels after plasmid-based and chromosomal expressions in this strain. In addition, minor differences in cell growth of the chromosomal expression strains were observed ($p=0.02$): whereas *E. coli* W3110(T7) grew to an OD_{600} of 5.1±0.2, *E. coli* BL21(DE3) only reached 4.3±0.3 (Figure S6). Without the addition of the substrate, ferulic acid **1**, both strains reached significantly higher ($p<0.01$) optical densities of 8.6±0.1 and 7.8±0.2, respectively. Since a concentration of 5 mM ferulic acid **1** does not affect the cell growth of *E. coli* as demonstrated by Yoon et al. [46], this difference may be due to either the metabolic burden caused by the consumption of cofactors during the reaction or the toxicity of the by-product coniferyl aldehyde **3**. Like many related phenolic aldehydes, **3** can be considered toxic to the cells and is therefore suspected to be reduced to the less toxic coniferyl alcohol **4** by multiple aldo-keto reductases or dehydrogenases in *E. coli* [18, 47]. In the literature, the combination of 4CL, CCR and heterologous cinnamyl alcohol dehydrogenases (CAD) has been used to produce 0.6 mM **4** from 1 mM **1** (60% yield) [48] and 1.82 mM **4** from 2.5 mM **4** (73% yield) [49] with

plasmid-based whole-cell systems, clearly emphasizing the significance of our yield of 75% using single gene copies in the chromosome. However, an alternative plasmid-based approach using a carboxylic acid reductase instead of 4CL/CCR in combination with a heterologous aldo-keto reductase reached even 97% yield from 5 mM **1** [19].

To further investigate the differences between both strains after chromosomal integration and to gain insight into the metabolic flux, time course experiments were performed. Samples were taken every 2 h after substrate addition and the concentration of both the product coniferyl alcohol **4** and the intermediate coniferyl aldehyde **3** as well as the optical density were quantified (Fig. 4).

The higher productivity of *E. coli* W3110(T7) compared to *E. coli* BL21(DE3), as shown in Fig. 2, becomes already evident within the first few hours of conversion, despite no significant differences in cell growth. After 8 h, the coniferyl alcohol **4** concentration reached 0.60±0.01 mM in *E. coli* BL21(DE3) and 1.58±0.05 mM in *E. coli* W3110(T7). Coniferyl aldehyde **3** remained at constantly low levels and its highest concentrations of 0.17±0.01 mM and 0.48±0.02 mM, respectively, were reached only two hours after substrate addition. Interestingly, we noted an intense orange coloration and a characteristic odor of the cultures after substrate addition. The intensity of the orange color could be correlated with the detection of **3**, an orange substance, via LC/MS.

Pinoresinol production

The *E. coli* W3110(T7) strain was chosen to extend the cascade starting from ferulic acid **1** to pinoresinol **5** due to its higher productivity for coniferyl alcohol synthesis compared to *E. coli* BL21(DE3). Based on our previous results [21], the laccase *Cg*L1 from *C. glutamicum* was selected as a promising candidate. Laccase-catalyzed oxidative coupling of two coniferyl alcohol **4** molecules proceeds via a radical mechanism. The phenol radical is stabilized by delocalization and can be represented by three different resonance forms (Figure S7). Thus, different coupling products are formed, of which pinoresinol **5** results from 8,8'-coupling and accounts for ~28% of all possible products in vitro [50]. Laccases contain four copper ions responsible for enzyme activity, and the reconstitution of the activity of *Cg*L1 in recombinant hosts like *E. coli* depends on copper availability. For this purpose, copper ions, for instance in form of $CuSO_4$, are usually added at the time point of induction of laccase expression [51]. However, in the case of the reaction described here, there are a few aspects to consider. In fact, coniferyl alcohol **4** is not the only potential substrate of the laccase, because other phenolic compounds like ferulic acid **1** and coniferyl aldehyde **3** are confirmed or suspected substrates for homo- and cross-coupling reactions as well [51]. Indeed, during preliminary tests for laccase-mediated phenol coupling, cross-coupled dimers of coniferyl aldehyde **3** and coniferyl alcohol **4** were detected along with dimers of coniferyl alcohol **4**. The assignment of possible respective coupling products was achieved using in vitro reactions with $FeCl_3$ as oxidizing agent (Figures S8-S10).

The equilibrium of the reduction of the aldehyde **3** to the alcohol **4** seemed to depend on the redox state of the cells. The addition of $CuSO_4$ increased the oxidative stress [52], which in turn resulted in high aldehyde **3** titers and a higher amount of coniferyl aldehyde-alcohol

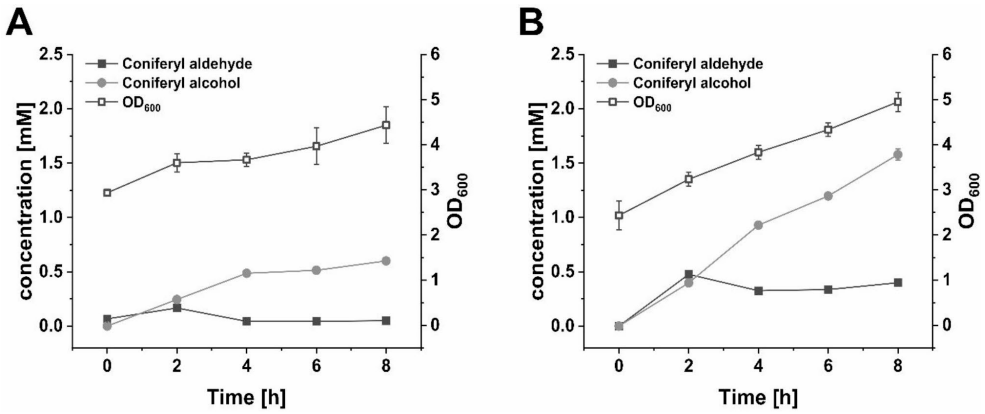


Fig. 4 Time course of coniferyl aldehyde **3** and alcohol **4** formation as well as cell growth (OD_{600}) during conversion of 5 mM ferulic acid **1** by two different plasmid-free whole-cell systems: (A) *E. coli* BL21(DE3) and (B) *E. coli* W3110(T7). Means and standard deviation were calculated from three biological replicates

cross-coupled products (data not shown). Unfortunately, one of these coupling products showed the same retention time in LC/MS as pinoresinol **5** and could not be resolved even by changing the analytical conditions (Figure S10). In our previous study, the addition of glucose or glycerol to the resting cells increased pinoresinol **5** formation [14]. As TB medium used in this study already contained glycerol, we tested additional glucose supplementation. The addition of 5 g/L glucose 0 h, 18 h and 42 h after substrate addition completely prevented aldehyde **3** accumulation and the subsequent formation of cross-coupling products (Figure S11). Apart from the NADPH-dependent reduction of **2** to **3** (Scheme 1), various endogenous NAD(P)H-dependent aldo-keto reductases and dehydrogenases like YqhD and YahK contribute to the subsequent reduction of **3** to **4** [18, 19, 53]. The addition of glucose strongly minimized the reoxidation of coniferyl alcohol **4** back to coniferyl aldehyde **3**, probably due to the change in the redox state in the cell and the increased availability of reduced nicotinamide cofactors [14]. In addition, changes in transcriptional regulation and expression of these endogenous enzymes might contribute to the observed effects of glucose addition.

Furthermore, the final product pinoresinol **5** also represents a substrate for laccases. In our previous in vitro studies, decrease of pinoresinol **5** was observed at a certain point of reaction dependent on the activity and the redox potential of the chosen laccase [21].

For plasmid-free pinoresinol **5** synthesis, the *Cg*L1-coding gene was cloned into the pgRNA-ET vector series for subsequent chromosomal integration at two different sites. The first gene copy was integrated into *pepB_sseB* locus of *E. coli* W3110(T7). An additional gene copy was integrated into *atpI_rsmG* locus to investigate the effects of the gene copy number on pinoresinol **5** production. To determine whether the timing of $CuSO_4$ addition provides advantages in the formation and preservation of pinoresinol **5**, addition at the time of induction (0 h) was compared to delayed addition after 18 h for both one and two integrated copies of *cgl1* (Fig. 5).

Interestingly, the strain harboring two laccase gene copies in the chromosome showed a slower ferulic acid **1** conversion and lower coniferyl alcohol **4** concentrations, but no influence on pinoresinol **5** formation was observed. Apparently, the integration of the second laccase gene copy affected the conversion of **1** to **4** negatively, which might be explained by altered expression levels of *Pc*4CL. Unfortunately, this hypothesis could not be confirmed via SDS-PAGE since the protein band on the gel did not appear prominent enough (Figure S12). However, the increased expression level of the laccase *Cg*L1 could be confirmed, but the titer of **5** did not increase further compared to a single gene copy. Pinoresinol **5** is a known substrate of *Cg*L1 and can therefore be degraded over time [21]. This might also explain why an equilibrium was reached between coniferyl alcohol **4** coupling and pinoresinol **5** degradation during the observation period: coniferyl alcohol **4** titers decreased but the concentration of **5** remained constant. In addition, insufficient incorporation of copper ions into the laccase protein has often been reported as a limiting factor upon laccase expression in recombinant *E. coli*, which may explain similar laccase activity regardless of gene copy number [51, 54, 55].

The time of $CuSO_4$ addition clearly affected the laccase-catalyzed oxidative coupling: When $CuSO_4$ was added at the time of induction, pinoresinol **5** could already be detected after 18 h and reached its maximum after 42 h. In cases where $CuSO_4$ was added 18 h after induction, a comparable titer was only reached after 66 h. This difference was also reflected by the time course of the coniferyl alcohol **4** titers: When $CuSO_4$ was added at the time of induction, lower titers of **4** were observed as it was already converted to its dimers by *Cg*L1. However, the gene expression is induced by IPTG and the same amount of proteins are therefore visible on the

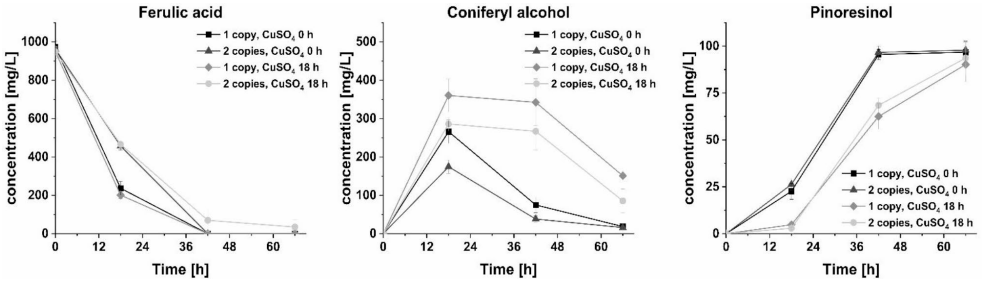


Fig. 5 Time course of ferulic acid **1**, coniferyl alcohol **4** and pinoresinol **5** titers. The effect of delayed $CuSO_4$ addition (at time of induction compared to 18 h after induction) and copy number of *cgl1* (one or two copies) were compared after addition of 970 mg/L (5 mM) ferulic acid **1**. Exemplary HPLC chromatograms are shown in Figure S11

SDS-PAGE gel regardless of the time of $CuSO_4$ addition (Figure S12).

Finally, a maximum concentration of ~ 100 mg/L (~ 270 µM) pinoresinol **5** was achieved using *E. coli* W3110(T7). Considering that 5 mM of substrate could yield a maximum of 2.5 mM dimeric coupling products, this corresponds to a theoretical yield of 10.8% with respect to the amount of ferulic acid **1** used in the reaction. Since the maximum titer of coniferyl alcohol **4** from 5 mM **1** was only 3.75 mM (Fig. 3), the yield of pinoresinol **5** with respect to coniferyl alcohol **4** was 14.4%. Our previously reported approach using recombinant *E. coli* resting cells and the endogenous *E. coli* multicopper oxidase CueO for coniferyl alcohol **4** coupling reached 240 µM pinoresinol **5** starting from 5 mM substrate (9.6% yield) on 0.5 mL scale [22]. In this case, resting cells were applied at higher cell densities. With purified enzymes, even 4.4 mM pinoresinol **5** starting from 10 mM eugenol could be achieved [21]. The higher yield of 17.6% might be attributed to the use of a biphasic in vitro system that might have prevented degradation of **5**. Nevertheless, it should be noted that an *in vitro* setup with isolated/purified enzymes as well as the use of resting cells require significantly more work compared to the growing cell approach presented here.

Since upscaling attempts with 10 mM ferulic acid **1** led to lower substrate conversion (data not shown), for preparative synthesis of pinoresinol **5**, 5 mM ferulic acid **1** was converted using the strain harboring one *cgl1* copy at 400 mL scale. $CuSO_4$ was added at the time of induction and the products were extracted after 66 h of conversion. After flash chromatography using both normal and reversed phase silica, 43 mg of pinoresinol (>98% purity quantified via HPLC, Figure S13) were isolated. Product identity was verified by LC/MS (ESI(-) m/z 357 $[M-H]^-$) and nuclear magnetic resonance (NMR) spectroscopy. The results of ^{1}H-NMR (Figure S14) and ^{13}C-NMR (Figure S15) were consistent with published data [23].

Conclusions

In this study, the evaluation of six intergenic integration loci using the fluorescent reporter mCherry allowed the selection of suitable loci for plasmid-free heterologous expression of plant and bacterial genes and subsequent biotransformation in growing *E. coli* cells. A versatile plasmid toolbox for the transfer of genes from pET vectors into these intergenic loci in both *E. coli* B and K-12 strains was established and evaluated. The application of this toolbox was demonstrated by integrating the genes encoding enzymes for the conversion of ferulic acid **1** to pinoresinol **5** in *E. coli* BL21(DE3) and W3110(T7). In future studies, this toolbox could also be used to extend the developed biosynthetic path and introduce it into a tyrosine-overproducing *E. coli* strain that use the cheaper substrate glucose. The synthesis of coniferyl alcohol **4** revealed substantial differences between the two strains tested. In *E. coli* W3110(T7), the chromosomal expression system enabled the same coniferyl alcohol **4** titer (3.75 mM) as the corresponding plasmid-based expression system. The reduction of coniferyl aldehyde **3** to alcohol **4** was catalyzed by endogenous enzymes of the host. Here, the supplementation of glucose to the medium shifted the equilibrium towards the formation of **4**. Finally, a maximal concentration of 100 mg/L pinoresinol **5** was reached. This study not only provided a nature-inspired route for the synthesis of pinoresinol **5**, but also helped to facilitate the progress towards plasmid-free enzymatic cascades in *E. coli*.

Methods

Strains and plasmids

E. coli DH5α was used for cloning and for plasmid propagation. *E. coli* BL21(DE3) and W3110(T7) [56] were used for expression, genome engineering and biotransformation. A list of all strains used and created in this study can be found in the supporting information (Table S2). Plasmids used for expression, as template for gene amplification or chromosomal integration are listed in Table S3.

Construction of pET-28a(+)_mCherry and pETDuet_ZmCCR_Pc4CL

The gene *mCherry* was amplified from pSEVA227R and cloned into the pET-28a(+) empty vector (Novagen) using FastDigest NdeI and EcoRI restriction enzymes and T4 DNA Ligase (Thermo Fisher Scientific). Digestion and subsequent ligation of the digested genes and vectors was carried out according to the manufacturer's protocols. A list of all primers used in this study can be found in the supporting information (Table S4).

The genes encoding the 4-coumarate-CoA ligase from *P. crispum* (*Pc*4CL) and the cinnamoyl-CoA reductase from *Z. mays* (*Zm*CCR) were amplified from pETDuet-TAL-4CL and pRSFDuet-CCR-CAD [16], respectively, and cloned into the pETDuet-1 empty vector (Novagen). Restriction digestion of the amplified genes and vector was performed using FastDigest NdeI and XhoI for cloning of *Pc4CL* and FastDigest EcoRI and NotI for cloning of *ZmCCR* (Thermo Fisher Scientific).

Chemically competent *E. coli* DH5α cells were transformed with the respective ligation product. The constructed plasmids were isolated from 4 ml of overnight cultures using the ZR Plasmid Miniprep – Classic Kit from ZymoResearch. Sequence integrity was verified by Sanger sequencing (Eurofins Genomics).

Construction of the pgRNA-ET and pgRNA-DUET vector series and chromosomal integration

The basic steps for CRISPR/Cas-assisted chromosomal integration were carried out as described before for linear donor DNA [56]. In brief summary, the design and evaluation of locus-specific sgRNAs was performed using the CHOPCHOP v3 web tool [57] (for details see Table S5). The N_{20} targeting sequence of the plasmid pgRNA was substituted for each of the chosen loci by PCR. The primers contained the locus-specific N_{20} as well as a BcuI (SpeI) restriction site for subsequent circularization of the linear PCR products (Table S4). Digestion with FastDigest BcuI and DpnI (Thermo Scientific) to remove the methylated template DNA was followed by a self-circularization ligation step as described by Seo et al. [43]. Successful substitution was verified by Sanger sequencing (Eurofins Genomics).

For λ-Red-mediated homologous recombination, homology arms of ~500 bp suitable for both *E. coli* K-12 and B strains were designed using the genome sequences of *E. coli* W3110 (GenBank accession no. AP009048.1) and *E. coli* BL21(DE3) (GenBank accession no. CP001509.3). These sequences were amplified by PCR from boiled cells of *E. coli* BL21(DE3). Linear double-stranded donor DNA (dsDonorDNA) for chromosomal integration into the *lacZ* reference locus was generated by fusion PCR of homology arms and *mCherry* as described before [56].

For construction of the pgRNA-ET and pgRNA-DUET vector series, the different pgRNA plasmids containing the locus-specific targeting sequences were linearized by PCR. Corresponding amplified homology arms and expression cassettes from pET-28a(+) (for construction of the pgRNA-ET series) or pETDuet-1 (for construction of the pgRNA-DUET series) were recombined as described by Gibson et al. [58]. In some cases, fusion PCR of the homology arms and the expression cassettes prior to Gibson assembly was used to improve the cloning efficiency. An illustration of the entire cloning procedure can be found in Figure S16.

The expression cassette of pgRNA-ET consists of the T7 promoter, a ribosome binding site, one multiple cloning site containing different optional tags, and the T7 terminator. In contrast, the expression cassette of pgRNA-DUET contains two multiple cloning sites, each with an upstream T7 promoter and ribosome binding site, allowing for two transcripts: a monocistronic transcript of the gene located in multiple cloning site no. 2, and a bicistronic transcript of genes of both multiple cloning sites.

For chromosomal integration of heterologous genes using the created toolbox, the genes to be integrated were amplified from the respective pET vectors by standard PCR using Phusion High-Fidelity DNA Polymerase (Thermo Scientific) according to the manufacturer's protocol. After purification by agarose electrophoresis, the amplicons as well as pgRNAET vectors were digested using FastDigest NdeI/XhoI for *mCherry* and FastDigest BamHI/NcoI for the laccase gene *cgl1* as described in the manufacturer's manual. Vectors and inserts were ligated using T4 DNA Ligase (Thermo Scientific) and subsequently used for transformation of chemically competent *E. coli* DH5α. Plasmids of selected colonies were isolated using the ZR Plasmid Miniprep – Classic kit (Zymo Research) and sequenced using standard T7/T7term primers (Eurofins Genomics).

The laboratory steps for chromosomal integration were carried out as previously described [56]. In brief, electrocompetent cells harboring pEcCas for expression of cas9 and λ-Red genes were freshly prepared. For this purpose, 5 ml 2xYT medium (containing 30 µg/ml kanamycin and 20 mM L-arabinose) was inoculated with 100 µl of an overnight culture and incubated at 37 °C with shaking at 180 revolutions per minute (rpm) for 2 h. Cells were harvested by centrifugation and washed twice with 1 ml ice-cold 10%(v/v) glycerol. After the cell pellet was resuspended in 100 µl 10%(v/v) glycerol, 100 ng of the circular pgRNA-ET or pgRNA-DUET plasmids were added which provided both the sgRNA and the necessary repair template for homologous recombination. After electroporation (2.5 kV, MicroPulser™ electroporator from Bio-Rad), the cells were immediately transferred to 1 ml of SOC medium and incubated at 37 °C, 180 rpm for one hour. Finally, the cells were plated out on LB agar plates (30 µg/ml kanamycin, 100 µg/ml ampicillin) and incubated at 37 °C overnight. Successful integration was verified by colony PCR using the primers listed in Table S4. Primers were designed to amplify both with and without successful integration, so that integration could be identified by the length of the amplicons (as seen in Figure S1). Plasmid curing was performed iteratively in 5 ml LB medium with the addition of 10 mM L-rhamnose for removing pgRNA-ET or pgRNA-DUET and 5%(w/v) sucrose for pEcCas, respectively.

Chromosomal expression and quantification of mCherry

After chromosomal integration into all loci in *E. coli* BL21(DE3) and *E. coli* W3110(T7), the fluorescent reporter *mCherry* was expressed in 50 ml LB medium, which was inoculated with 1 ml of an overnight culture. The cultures were grown at 37 °C, 180 rpm until an OD_{600} of 0.6–0.7 was reached. After addition of 0.1 mM IPTG for induction, the expression was performed at 30 °C, 180 rpm for 20 h. The optical density was now quantified at 650 nm to avoid interference with the fluorescence of mCherry [59]. Fluorescence measurements of 200 µl culture aliquots were performed using a Tecan Infinite

M200 Pro microplate reader (excitation at 570 nm, emission at 625 nm).

Heterologous expression and biotransformation of ferulic acid

The genes encoding *Pc*4CL and *Zm*CCR were integrated into the *flgF_flgG* locus of the chromosomes of *E. coli* strains BL21(DE3) and W3110(T7). To compare chromosomal and plasmid-based expression, both strains (without integrated genes) were also transformed with pETDuet_ZmCCR_Pc4CL. The laccase gene *cgl1* was subsequently integrated into the loci *pepB_sseB* and *atpI_rsmG* of *E. coli* W3110(T7) already harboring copies of *Pc4CL* and *ZmCCR* in the *flgF_flgG* locus.

For biotransformation, the respective *E. coli* cells were grown in TB or LB medium. Cultures were inoculated with 2%(v/v) of an overnight culture in either 50 ml for analytical reactions or 400 ml scale for the preparative reaction (cf. 4.7) and incubated at 37 °C, 180 rpm. For plasmid-based expression, 30 µg/ml kanamycin was used to maintain the pETDuet plasmid in the cell. At an OD_{600} of 0.6 to 0.8, transcription was induced by addition of 0.1 mM IPTG and the cells were subsequently incubated at 30 °C, 140 rpm. After 2 h, 5 mM ferulic acid **1** was added using a stock solution of 250 mM in dimethyl sulfoxide (DMSO), resulting in a final concentration of 2%(v/v) DMSO. To prevent aldehyde **3** formation, 5 g/L glucose was supplemented at time of substrate addition, after 18 h and after 42 h. For copper loading of the laccases, 3 mM $CuSO_4$ was added either at time of induction or 18 h after substrate addition. For quantification of substrate and products as well as SDS-PAGE analysis and Western Blot, 500 µl samples were taken at different time points.

For product analysis, 125 µl of culture medium was diluted with 115 µl water. Cinnamic acid (62.5 mM dissolved in DMSO) was added as an internal standard (10 µl, 2.5 mM final concentration). Extraction was performed twice using 500 µl ethyl acetate each. The combined organic phase (2×400 µl) was evaporated and the residue was dissolved in 100 µl methanol. Product analysis was performed via HPLC and LC/MS analysis using a gradient of water with 0.1%(v/v) formic acid (A) and methanol (B) at a flow rate of 0.5 ml/min (LCMS-2020, Shimadzu; Chromolith Performance RP-18e 100–4.6 mm column and Chromolith RP-18e 10–4.6 mm guard cartridge, Merck Millipore). Details of the HPLC gradient settings are provided in Table S6. The LCMS-2020 was set to scan mode to detect both positive and negative ions using the dual ion source (m/z 159–1000, scan speed 3750 u/s, event time 0.25 s). Products were identified by comparison of retention time and mass to charge ratio (m/z) with commercial standards. Substrate and products were quantified based on their UV absorption using a photodiode array detector (SPD-M20A, Shimadzu) at different wavelengths (ferulic acid **1** 325 nm, coniferyl aldehyde **3** 340 nm, coniferyl alcohol **4** 260 nm, pinoresinol **5** 280 nm). Linear internal calibration curves of all three compounds were used to calculate the substrate and product titers from peak areas at the indicated wavelengths for each sample.

For SDS-PAGE analysis, aliquots of the cell suspension equivalent to an OD_{600} of 0.25 were centrifuged. The cell pellets were resuspended in 50 µl 1 x SDS-PAGE loading buffer and loaded onto 12.5% polyacrylamide gels [60]. For Western blot analysis, proteins were blotted onto a nitrocellulose membrane which was subsequently washed twice with TBS (20 mM Tris-HCl, 150 mM NaCl, pH 7.6) and blocked with 3%(w/v) bovine serum albumin (BSA) in TBS. After washing with TBST (0.1%(v/v) Tween 20 in TBS) and TBS, the membrane was incubated with a 1:1000 dilution of the primary antibody (mouse 6x-His tag monoclonal antibody HIS.H8 from Thermo Fisher Scientific) in TBS with 3%(w/v) BSA for one hour. Prior to the incubation with the secondary antibody (Peroxidase AffiniPure™ Goat Anti-Mouse IgG (Jackson ImmunoResearch), 1:10,000 diluted in TBS with 10%(w/v) milk powder), the membrane was washed twice with TBST and once with TBS. Finally, the membrane was washed thoroughly with TBS and the color reaction was initiated by the addition of 3,3′,5,5′-tetramethylbenzidine (TMB) Liquid Substrate System (Sigma Aldrich).

Oxidative coupling using $FeCl_3$

For identification of cross-coupling products of coniferyl aldehyde **3** and alcohol **4**, 10 µl of 5 mM solutions in DMSO for homo-coupling or 5 µl each for cross-coupling were diluted in 230 µl ddH_2O. Addition of 10 µl 5 mM $FeCl_3$ (equimolar amount) at room temperature led to the formation of coupling products within minutes. The mixture was extracted twice with 500 µl ethyl acetate. The combined organic phases were evaporated and the residue was dissolved in 50 µl methanol for LC/MS analysis.

Isolation of pinoresinol

For product isolation and purification, two 400 ml cultures for conversion of 5 mM ferulic acid **1** were incubated for 66 h (with glucose addition 0 h, 18 h, and 42 h after substrate addition). The reaction was monitored as described above. The cells were harvested by centrifugation and the supernatant was used for product extraction in a separation funnel. The aqueous solution was extracted with 2×400 ml ethyl acetate. The combined organic phases were dried with $MgSO_4$, filtered and evaporated under reduced pressure. The residue was subjected to flash chromatography (Silica 60 M 0.04–0.063 mm, Macherey-Nagel) using a mixture of ethyl acetate (70%(v/v)) and n-heptane (30%(v/v)) as eluent. The addition of 2%(v/v) triethylamine to the eluent improved

separation. Progress of flash chromatography was monitored using thin layer chromatography. Fractions containing a high amount of pinoresinol **5** were combined and analyzed via LC/MS. Since purification via normal phase was not sufficient, reversed-phase flash chromatography (C_{18}-Reversed-Phase Silica, Sigma Aldrich) was performed. The mobile phase consisted of 50%(v/v) methanol in water. Combined fractions were first evaporated under reduced pressure to remove methanol. The remaining aqueous phase was freeze-dried. The purity of the isolated compound was analyzed via HPLC and the structure was confirmed by ^{1}H- and ^{13}C-NMR (Bruker Avance III – 300).

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12934-024-02562-3.

Supplementary Material 1

Acknowledgements
We would like to thank Dr. Natalie Trachtmann (Institute of Microbiology, University of Stuttgart) for helpful advices and CeMSA@HHU (Center for Molecular and Structural Analytics @ Heinrich Heine University) for recording the NMR-spectroscopic data.

Author contributions
U.J.L conceptualized the study, designed and conducted most of the experiments, analyzed the data and prepared the first draft. A.W. and L.M.B. contributed to data acquisition and analysis. V.B.U. contributed to conceiving experiments, interpretation of data, and writing of the manuscript. All authors edited the manuscript and approved the final version of the manuscript.

Funding
Financial support was provided by the Federal Ministry of Education and Research to Heinrich-Heine University Düsseldorf [grant number 031B0362A].

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 2 July 2024 / Accepted: 9 October 2024
Published online: 23 October 2024

References

1. Teponno RB, Kusari S, Spiteller M. Recent advances in research on lignans and neolignans. Nat Prod Rep. 2016;33(9):1044–92.
2. Zalesak F, Bon DJD, Pospisil J. Lignans and neolignans: Plant secondary metabolites as a reservoir of biologically active substances. Pharmacol Res. 2019;146:104284.
3. Milder IE, Arts IC, van de Putte B, Venema DP, Hollman PC. Lignan contents of Dutch plant foods: a database including lariciresinol, pinoresinol, secoisolariciresinol and matairesinol. Br J Nutr. 2005;93(3):393–402.
4. Nadeem M, Palni L, Purohit A, Pandey H, Nandi S. Propagation and conservation of *Podophyllum hexandrum* Royle: an important medicinal herb. Biol Conserv. 2000;92(1):121–9.
5. Farkya S, Bisaria VS, Srivastava AK. Biotechnological aspects of the production of the anticancer drug podophyllotoxin. Appl Microbiol Biotechnol. 2004;65(5):504–19.
6. Mark R, Lyu X, Lee JJL, Parra-Saldívar R, Chen WN. Sustainable production of natural phenolics for functional food applications. J Funct Foods. 2019;57:233–54.
7. During A, Debouche C, Raas T, Larondelle Y. Among Plant Lignans, Pinoresinol has the Strongest Antiinflammatory properties in Human intestinal Caco-2 cells. J Nutr. 2012;142(10):1798–805.
8. Kim HY, Kim JK, Choi JH, Jung JY, Oh WY, Kim DC, et al. Hepatoprotective Effect of Pinoresinol on Carbon Tetrachloride-Induced hepatic damage in mice. J Pharmacol Sci. 2010;112(1):105–12.
9. Lopez-Biedma A, Sanchez-Quesada C, Beltran G, Delgado-Rodriguez M, Gaforio JJ. Phytoestrogen (+)-pinoresinol exerts antitumor activity in breast cancer cells with different oestrogen receptor statuses. BMC Complement Altern Med. 2016;16(1):350.
10. Hwang B, Lee J, Liu QH, Woo ER, Lee DG. Antifungal effect of (+)-pinoresinol isolated from *Sambucus williamsii*. Molecules. 2010;15(5):3507–16.
11. Wikul A, Damsud T, Kataoka K, Phuwapraisirisan P. (+)-Pinoresinol is a putative hypoglycemic agent in defatted sesame (*Sesamum indicum*) seeds though inhibiting alpha-glucosidase. Bioorg Med Chem Lett. 2012;22(16):5215–7.
12. Decembrino D, Raffaele A, Knöfel R, Girhard M, Urlacher VB. Synthesis of (-)-deoxypodophyllotoxin and (-)-epipodophyllotoxin via a multi-enzyme cascade in *E. coli*. Microb Cell Fact. 2021;20(1):183.
13. Grougnet R, Magiatis P, Mitaku S, Terzis A, Tillequin F, Skaltsounis A-L. New lignans from the perisperm of *Sesamum indicum*. J Agric Food Chem. 2006;54(20):7570–4.
14. Ricklefs E, Girhard M, Urlacher VB. Three-steps in one-pot: whole-cell biocatalytic synthesis of enantiopure (+)- and (-)-pinoresinol via kinetic resolution. Microb Cell Fact. 2016;15:78.
15. Decembrino D, Ricklefs E, Wohlgemuth S, Girhard M, Schullehner K, Jach G, et al. Assembly of plant enzymes in *E. coli* for the production of the valuable (-)-podophyllotoxin precursor (-)-pluviatolide. ACS Synth Biol. 2020;9(11):3091–103.
16. Jansen F, Gillessen B, Mueller F, Commandeur U, Fischer R, Kreuzaler F. Metabolic engineering for *p*-coumaryl alcohol production in *Escherichia coli* by introducing an artificial phenylpropanoid pathway. Biotechnol Appl Biochem. 2014;61(6):646–54.
17. van Summeren-Wesenhagen PV, Voges R, Dennig A, Sokolowsky S, Noack S, Schwaneberg U, et al. Combinatorial optimization of synthetic operons for the microbial production of *p*-coumaryl alcohol with *Escherichia coli*. Microb Cell Fact. 2015;14:79.
18. Kunjapur AM, Tarasova Y, Prather KL. Synthesis and accumulation of aromatic aldehydes in an engineered strain of *Escherichia coli*. J Am Chem Soc. 2014;136(33):11644–54.
19. Tramontina R, Galman JL, Parmeggiani F, Derrington SR, Bugg TDH, Turner NJ, et al. Consolidated production of coniferol and other high-value aromatic alcohols directly from lignocellulosic biomass. Green Chem. 2020;22(1):144–52.
20. Pickel B, Constantin MA, Pfannstiel J, Conrad J, Beifuss U, Schaller A. An enantiocomplementary dirigent protein for the enantioselective laccase-catalyzed oxidative coupling of phenols. Angew Chem Int Ed Engl. 2010;49(1):202–4.
21. Ricklefs E, Girhard M, Koschorreck K, Smit MS, Urlacher VB. Two-step one-pot synthesis of pinoresinol from eugenol in an enzymatic cascade. ChemCatChem. 2015;7(12):1857–64.
22. Decembrino D, Girhard M, Urlacher VB. Use of copper as a trigger for the *in vivo* activity of *E. coli* Laccase CueO: a simple tool for biosynthetic purposes. ChemBioChem. 2021;22(8):1470–9.
23. Yue F, Lan W, Zhang L, Lu F, Sun R, Ralph J. Efficient synthesis of pinoresinol, an important lignin dimeric model compound. Front Energy Res. 2021;9:640337.
24. Steinmann A, Finger M, Nowacki C, Decembrino D, Hubmann G, Girhard M, et al. Heterologous lignan production in stirred-tank reactors—Metabolomics-assisted bioprocess development for an *in vivo* enzyme cascade. Catalysts. 2022;2(11):1473.

25. Mairhofer J, Scharl T, Marisch K, Cserjan-Puschmann M, Striedner G. Comparative transcription profiling and in-depth characterization of plasmid-based and plasmid-free *Escherichia coli* expression systems under production conditions. Appl Environ Microbiol. 2013;79(12):3802–12.
26. Englaender JA, Jones JA, Cress BF, Kuhlman TE, Linhardt RJ, Koffas MAG. Effect of Genomic Integration Location on Heterologous Protein Expression and Metabolic Engineering in *E. coli*. ACS Synth Biol. 2017;6(4):710–720.
27. Striedner G, Pfaffenzeller I, Markus L, Nemecek S, Grabherr R, Bayer K. Plasmid-free T7-Based *Escherichia coli* expression systems. Biotechnol Bioeng. 2010;105(4):786–94.
28. Saleski TE, Chung MT, Carruthers DN, Khasbaatar A, Kurabayashi K, Lin XN. Optimized gene expression from bacterial chromosome by high-throughput integration and screening. Sci Adv. 2021;7(7):eabe1767.
29. Jiang Y, Chen B, Duan C, Sun B, Yang J, Yang S. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. Appl Environ Microbiol. 2015;81(7):2506–14.
30. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 2013;152(5):1173–83.
31. Bryant JA, Sellars LE, Busby SJ, Lee DJ. Chromosome position effects on gene expression in *Escherichia coli* K-12. Nucleic Acids Res. 2014;42(18):11383–1192.
32. Sousa C, de Lorenzo V, Cebolla A. Modulation of gene expression through chromosomal positioning in *Escherichia coli*. Microbiology. 1997;143(Pt 6):2071–8.
33. Scholz SA, Diao R, Wolfe MB, Fivenson EM, Lin XN, Freddolino PL. High-resolution mapping of the *Escherichia coli* chromosome reveals positions of high and low transcription. Cell Syst. 2019;8(3):212–25.
34. Goormans AR, Snoeck N, Decadt H, Vermeulen K, Peters G, Coussement P, et al. Comprehensive study on *Escherichia coli* genomic expression: does position really matter? Metab Eng. 2020;62:10–9.
35. Huang C, Guo L, Wang J, Wang N, Huo YX. Efficient long fragment editing technique enables large-scale and scarless bacterial genome engineering. Appl Microbiol Biotechnol. 2020;104(18):7943–56.
36. Kuhlman TE, Cox EC. Site-specific chromosomal integration of large synthetic constructs. Nucleic Acids Res. 2010;38(6):e92.
37. Santos-Zavaleta A, Salgado H, Gama-Castro S, Sanchez-Perez M, Gomez-Romero L, Ledezma-Tejeida D, et al. RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in *E. coli* K-12. Nucleic Acids Res. 2019;47(D1):D212–20.
38. Keseler IM, Collado-Vides J, Santos-Zavaleta A, Peralta-Gil M, Gama-Castro S, Muniz-Rascado L, et al. EcoCyc: a comprehensive database of *Escherichia coli* biology. Nucleic Acids Res. 2011;39(Database issue):D583–590.
39. Vora T, Hottes AK, Tavazoie S. Protein occupancy Landscape of a bacterial genome. Mol Cell. 2009;35(2):247–53.
40. Bassalo MC, Garst AD, Halweg-Edwards AL, Grau WC, Domaille DW, Mutalik VK, et al. Rapid and efficient one-step metabolic pathway integration in *E. coli*. ACS Synth Biol. 2016;5(7):561–8.
41. Li Y, Lin Z, Huang C, Zhang Y, Wang Z, Tang Y-j, et al. Metabolic engineering of *Escherichia coli* using CRISPR–Cas9 meditated genome editing. Metab Eng. 2015;31:13–21.
42. Chung ME, Yeh IH, Sung LY, Wu MY, Chao YP, Ng IS, et al. Enhanced integration of large DNA into *E. coli* chromosome by CRISPR/Cas9. Biotechnol Bioeng. 2017;114(1):172–83.
43. Seo J-H, Baek S-W, Lee J, Park J-B. Engineering *Escherichia coli* BL21 genome to improve the heptanoic acid tolerance by using CRISPR-Cas9 system. Biotechnol Bioprocess Eng. 2017;22(3):231–8.
44. Juhas M, Ajioka JW. Identification and validation of novel chromosomal integration and expression loci in *Escherichia coli* Flagellar Region 1. PLoS ONE. 2015;10(3):e0123007.
45. Merck KGaA. pET System Manual 11th Edition. Darmstadt: Merck KGaA; 2011.
46. Yoon SH, Lee EG, Das A, Lee SH, Li C, Ryu HK, et al. Enhanced vanillin production from recombinant *E. coli* using NTG mutagenesis and adsorbent resin. Biotechnol Prog. 2007;23(5):1143–8.
47. Zaldivar J, Martinez A, Ingram LO. Effect of selected aldehydes on the growth and fermentation of ethanologenic *Escherichia coli*. Biotechnol Bioeng. 1999;65(1):24–33.
48. Liu S, Liu J, Hou J, Chao N, Gai Y, Jiang X. Three steps in one pot: biosynthesis of 4-hydroxycinnamyl alcohols using immobilized whole cells of two genetically engineered *Escherichia coli* strains. Microb Cell Fact. 2017;16(1):104.
49. Aschenbrenner J, Marx P, Pietruszka J, Marienhagen J. Microbial Production of Natural and Unnatural Monolignols with *Escherichia coli*. ChemBioChem. 2019;20(7):949–54.
50. Halls SC, Davin LB, Kramer DM, Lewis NG. Kinetic Study of Coniferyl Alcohol Radical Binding to the (+)-Pinoresinol Forming Dirigent Protein. Biochemistry. 2004;43(9):2587–95.
51. Ricklefs E, Winkler N, Koschorreck K, Urlacher VB. Expanding the laccase-toolbox: a laccase from *Corynebacterium glutamicum* with phenol coupling and cuprous oxidase activity. J Biotechnol. 2014;191:46–53.
52. Kimura T, Nishioka H. Intracellular generation of superoxide by copper sulphate in *Escherichia coli*. Mutat Research/Genetic Toxicol Environ Mutagen. 1997;389(2):237–42.
53. Zhao M, Zhang B, Wu X, Xiao Y. Whole-cell bioconversion systems for efficient synthesis of monolignols from L-tyrosine in *Escherichia coli*. J Agric Food Chem. 2024;72(26):14799–808.
54. Durão P, Chen Z, Fernandes AT, Hildebrandt P, Murgida DH, Todorovic S, et al. Copper incorporation into recombinant CotA laccase from *Bacillus subtilis*: characterization of fully copper loaded enzymes. J Biol Inorg Chem. 2008;13(2):183–93.
55. Gunne M, Urlacher VB. Characterization of the alkaline laccase SslI1 from *Streptomyces sviceus* with unusual properties discovered by genome mining. PLoS ONE. 2012;7(12):e52360.
56. Luelf UJ, Böhmer LM, Li S, Urlacher VB. Effect of chromosomal integration on catalytic performance of a multi-component P450 system in *Escherichia coli*. Biotechnol Bioeng. 2023;120(7):1762–72.
57. Labun K, Montague TG, Krause M, Torres Cleuren YN, Tjeldnes H, Valen E. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res. 2019;47(W1):W171–4.
58. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA 3rd, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods. 2009;6(5):343–5.
59. Hecht A, Endy D, Salit M, Munson MS. When wavelengths Collide: Bias in cell abundance measurements due to expressed fluorescent proteins. ACS Synth Biol. 2016;5(9):1024–7.
60. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 1970;227(5259):680–5.

Publisher's note

2.3.1 Supporting information

Supporting information

Plasmid-free production of the plant lignan pinoresinol in growing *Escherichia coli* cells

U. Joost Luelf[1], Alexander Wassing[1], Lisa M. Böhmer[1], Vlada B. Urlacher[1]*

[1]Institute of Biochemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

*Corresponding Author

Email: vlada.urlacher@uni-duesseldorf.de

ORCID: Vlada B. Urlacher 0000-0003-1312-4574

Running title: Plasmid-free pinoresinol synthesis

Supporting Tables

Table S1. Efficiency of *mCherry* integration in *E. coli* BL21(DE3) using the pgRNA-ET vector series.

Locus	Integration efficiency
atpI_rsmG	15/20
glmU_atpC	16/20
yhiO_pitA	18/20
sspA_rpsI	19/20
pepB_sseB	19/20
flgF_flgG	20/20

Table S2. List of strains used in this study

Strain	Purpose	Source
E. coli DH5α	cloning procedures	Clontech
E. coli BL21(DE3)	expression and genome engineering	Novagen
E. coli W3110(T7)	expression and genome engineering	(1)
E. coli BL21(DE3) mCherry (atpI_rsmG)	expression	this study
E. coli BL21(DE3) mCherry (glmU_atpC)	expression	this study
E. coli BL21(DE3) mCherry (yhiO_pitA)	expression	this study
E. coli BL21(DE3) mCherry (sspA_rpsI)	expression	this study
E. coli BL21(DE3) mCherry (pepB_sseB)	expression	this study
E. coli BL21(DE3) mCherry (flgF_flgG)	expression	this study
E. coli BL21(DE3) mCherry (lacZ)	expression	this study
E. coli W3110(T7) mCherry (atpI_rsmG)	expression	this study
E. coli W3110(T7) mCherry (glmU_atpC)	expression	this study
E. coli W3110(T7) mCherry (yhiO_pitA)	expression	this study
E. coli W3110(T7) mCherry (sspA_rpsI)	expression	this study
E. coli W3110(T7) mCherry (pepB_sseB)	expression	this study
E. coli W3110(T7) mCherry (flgF_flgG)	expression	this study
E. coli W3110(T7) mCherry (lacZ)	expression	this study
E. coli BL21(DE3) ZmCCR_Pc4CL (flgF_flgG)	expression and biotransformation	this study
E. coli W3110(T7) ZmCCR_Pc4CL (flgF_flgG)	expression and biotransformation	this study
E. coli W3110(T7) ZmCCR_Pc4CL (flgF_flgG) Cgl1 (pepB_sseB)	expression and biotransformation	this study
E. coli W3110(T7) ZmCCR_Pc4CL (flgF_flgG) 2xCgl1 (pepB_sseB, atpI_rsmG)	expression and biotransformation	this study

Table S3. List of plasmids used in this study.

Plasmid	Purpose	Source
pEcCas	expression of cas9 and λ-Red genes	(2)
pgRNA-bacteria	transcription of sgRNA; template for N_{20} exchange	(3)
pgRNA_001	targets *atpI_rsmG* locus	(1)
pgRNA_002	targets *glmU_atpC* locus	this study
pgRNA_003	targets *yhiO_pitA* locus	this study
pgRNA_004	targets *sspA_rpsI* locus	this study
pgRNA_005	targets *pepB_sseB* locus	this study
pgRNA_006	targets *flgF_flgG* locus	this study
pgRNA_007	targets *lacZ* locus	this study
pgRNA-ET_001	for integration in *atpI_rsmG* locus	this study
pgRNA-ET_002	for integration in *glmU_atpC* locus	this study
pgRNA-ET_003	for integration in *yhiO_pitA* locus	this study
pgRNA-ET_004	for integration in *sspA_rpsI* locus	this study
pgRNA-ET_005	for integration in *pepB_sseB* locus	this study
pgRNA-ET_006	for integration in *flgF_flgG* locus	this study
pSEVA227R	template for amplification of *mCherry*	SEVA
pET-28a(+) mCherry	template for cloning of *mCherry*	this study
pgRNA-ET_001_mCherry	for integration of *mCherry* in *atpI_rsmG* locus	this study
pgRNA-ET_002_mCherry	for integration of *mCherry* in *glmU_atpC* locus	this study
pgRNA-ET_003_mCherry	for integration of *mCherry* in *yhiO_pitA* locus	this study
pgRNA-ET_004_mCherry	for integration of *mCherry* in *sspA_rpsI* locus	this study
pgRNA-ET_005_mCherry	for integration of *mCherry* in *pepB_sseB* locus	this study
pgRNA-ET_006_mCherry	for integration of *mCherry* in *flgF_flgG* locus	this study
pgRNA-Duet_006	for integration of *mCherry* in *flgF_flgG* locus	this study
pRSFDuet_CCR_CAD	template for amplification of *ZmCCR*	(4)
pETDuet-TAL-4CL	template for amplification of *Pc4CL*	(4)
pETDuet_ZmCCR_Pc4CL	expression	this study
pgRNA-Duet_006_ZmCCR_Pc4CL	for integration of *ZmCCR* & *Pc4CL* in *flgF_flgG* locus	this study
pET-16b_cgl1	template for amplification of *cgl1*	(5)
pgRNA-ET_001_cgl1	for integration of cgl1 in *atpI_rsmG* locus	this study
pgRNA-ET_005_cgl1	for integration of cgl1 in *pepB_sseB* locus	this study

Table S4. Primer sequences.

Name	Integration Locus	Sequence (5‘ to 3‘)	Reference
Primers for amplification & cloning of mCherry into pET-28a(+) (Restriction sites are underlined)			
mCherry fw	-	ATATATCATATGATGGTGAGCAAGGGCGAGGAG	-
mCherry rev	-	ATATATGAATTCTTACTTGTACAGCTCGTCC	-
Primers for cloning of *mCherry* & *cgl1* into pgRNA-ET plasmids			
pET-28a(+) fw	-	TATAGGCGCCAGCAACC	-
pET-28a(+) rev	-	TTAATGCGCCGCTACAGG	-
cgl1 fw	-	CACCATGGGTCCCTCCCTTCGCCC	-
cgl1 rev	-	CGGGATCCTCCTTACTCGTAGCGAAGCGAG	-
Primers for amplification & cloning of *ZmCCR* and *Pc4CL* into pETDuet (Restriction sites are underlined)			
ZmCCR fw	-	ATATATGAATTCGATGACGGTGGTTGATGCG	-
ZmCCR rev	-	ATATATGCGGCCGCTTAGGCGCGAATTGCAATG	-
Pc4CL fw	-	ATATATCATATGGGAGATTGTGTAGCACC	-
Pc4CL rev	-	ATATATCTCGAGTTATTTGGGAAGATCACCGG	-
gRNA Primer (Restriction sites are underlined. N20 targeting sequences are shown in italic.)			
gRNA rev	-	ATATATACTAGTATTATACCTAGGACTGAGCTAG	general design (6)
gRNA_XXX fw (general primer design)	-	ATATATACTAGT-N_{20}-GTTTTAGAGCTAGAAATAGC	general design (6)
gRNA_001 fw	atpI_rsmG	ATATATACTAGT*TTATTAAAAATGTCAATGGG*GTTTTAGAGCTAGAAATAGC	(1)
gRNA_002 fw	glmU_atpC	ATATATACTAGT*TATGTGAACGCTATTCAGGA*GTTTTAGAGCTAGAAATAGC	-
gRNA_003 fw	yhiO_pitA	ATATATACTAGT*ACTATGTCACAATCTGAAGC*GTTTTAGAGCTAGAAATAGC	-
gRNA_004 fw	sspA_rpsI	ATATATACTAGT*TGTTTTACCTGTCTGTCAGG*GTTTTAGAGCTAGAAATAGC	-
gRNA_005 fw	pepB_sseB	ATATATACTAGT*ACAAATTGCGAATCCCTTGT*GTTTTAGAGCTAGAAATAGC	-
gRNA_006 fw	flgF_flgG	ATATATACTAGT*TTTTGCGAGCACTTGTAGGC*GTTTTAGAGCTAGAAATAGC	-
gRNA_007 fw	lacZ	ATATATACTAGT*ACTATGTCACAATCTGAAGC*GTTTTAGAGCTAGAAATAGC	-
pgRNA seq		GCCACCTGACGTCTAAG	(1)
Primers for cloning of pgRNA-ET and pgRNA-DUET			
pET fw + Overlap HA1		GCTAAGAACCATCATTGGCTGTATAGGCGCCAGCAACC	-
pET rev + Overlap HA2		GCTAACCCTGTTCGATCATTAATGCGCCGCTACAGG	-
HA1 fw	atpI_rsmG	GCCACCACCAGTAACAC	-
HA1 rev		TGATCGAACAGGGTTAGC	-
HA2 fw		CAGCCAATGATGGTTCTTAGC	-

Name	Integration Locus	Sequence (5‘ to 3‘)	Reference
HA2 rev		CGTCAGGTGCAACATGAG	-
pgRNA fw		GTGTTACTGGTGGTGGCTTCAAAAAAAGCACCGACTC	-
pgRNA rev		CTCATGTTGCACCTGACGGGGATAACGCAGGAAAGAAC	-
pET fw + Overlap HA1	glmU_atpC	GTGTGACCCGTCCTGAATATAGGCGCCAGCAACC	-
pET rev + Overlap HA2		CGTCAGGTGGATGTTTTTGTTAATGCGCCGCTACAGG	-
HA1 fw		GAGGGTGATATGGCAATGAC	-
HA1 rev		TTCAGGACGGGTCACAC	-
HA2 fw		CAAAAACATCCACCTGACGC	-
HA2 rev		CATCGGCAAAGAAAGGTG	-
pgRNA fw		GTCATTGCCATATCACCCTCTTCAAAAAAAGCACCGACTC	-
pgRNA rev		CACCTTTCTTTGCCGATGGGGATAACGCAGGAAAGAAC	-
pET fw + Overlap HA1	yhiO_pitA	CCAGTAAACTGACCCGCTTAATGCGCCGCTACAGG	-
pET rev + Overlap HA2		CTGTAATCGTACGCACCATATAGGCGCCAGCAACC	-
HA1 fw		AAGCGGTTTCTCCAGGAC	-
HA1 rev		GCGGGTCAGTTTACTGG	-
HA2 fw		TGGTGCGTACGATTACAG	-
HA2 rev		CGCCGCATTATGCTGTG	-
pgRNA fw		GTCCTGGAGAAACCGCTTTTCAAAAAAAGCACCGACTC	-
pgRNA rev		CACAGCATAATGCGGCGGGGATAACGCAGGAAAGAAC	-
pET fw + Overlap HA1	sspA_rpsI	CCCTTATTGGCGATGTGGTTTTAATGCGCCGCTACAGG	-
pET rev + Overlap HA2		GAATCAGCGTAAAAACTGGAATATAGGCGCCAGCAACC	-
HA1 fw		ATGATGCGAGATTCCCACAG	-
HA1 rev		AACCACATCGCCAATAAGGG	-
HA2 fw		TTCCAGTTTTTACGCTGATTC	-
HA2 rev		CAATACTACGGCACTGGTC	-
pgRNA fw		CTGTGGGAATCTCGCATCATTTCAAAAAAAGCACCGACTC	-
pgRNA rev		GACCAGTGCCGTAGTATTGGGGATAACGCAGGAAAGAAC	-
pET fw + Overlap HA1	pepB_sseB	CCTGATGCGCTACGCTTATTAATGCGCCGCTACAGG	-
pET rev + Overlap HA2		CTAATATGCCGGATGCGGCTATAGGCGCCAGCAACC	-
HA1 fw		ACCGCCTTCCAGGATTTC	-
HA1 rev		TAAGCGTAGCGCATCAGG	-
HA2 fw		GCCGCATCCGGCATATTAG	-
HA2 rev		AGCTGGGCGATATCATCACC	-
pgRNA fw		GGTGATGATATCGCCCAGCTTTCAAAAAAAGCACCGACTC	-
pgRNA rev		GAAATCCTGGAAGGCGGTGGGATAACGCAGGAAAGAAC	-
pET fw + Overlap HA1	flgF_flgG	CTGTAGTGGATTAGTGACAAAGTTAATGCGCCGCTACAGG	-
pET rev + Overlap HA2		GTATAAGTTGCCCGATGCGTATAGGCGCCAGCAACC	-

Name	Integration Locus	Sequence (5‘ to 3‘)	Reference
HA1 fw		GCGTAATGGCAGCATTC	-
HA1 rev		CGCATCGGGCAACTTATAC	-
HA2 fw		CTTTGTCACTAATCCACTACAG	-
HA2 rev		TCACGACCGATGGTGATAC	-
pgRNA fw		GTATCACCATCGGTCGTGATTCAAAAAAAGCACCGACTC	-
pgRNA rev		GAATGCTGCCATTACGCGGGATAACGCAGGAAAGAAC	-
Primers for synthesis of linear dsDonorDNA (*lacZ* locus)			
pET fw + Overlap HA1	lacZ	GTCTTCATCCACGCGTTAATGCGCCGCTACAGG	-
pET rev + Overlap HA2		CCATGTTGCCACTCGTATAGGCGCCAGCAACC	-
HA1 fw		ACTGTGAGCCAGAGTTG	-
HA1 rev		CGCGTGGATGAAGAC	-
HA2 fw		CGAGTGGCAACATGG	-
HA2 rev		TGGCGTAATAGCGAAG	-
Primers for verification of integration by colony PCR			
atpI_rsmG fw	atpI_rsmG	TCAGCGGCAAGAATACC	-
atpI_rsmG rev		TCCTGAAGCCCATTTCAC	-
glmU_atpC fw	glmU_atpC	GGAAGGCGAATACGATCAC	-
glmU_atpC rev		GCGGTTAAAGGCATGTTG	-
yhiO_pitA fw	yhiO_pitA	TCAATCCGCCTTGCTTAC	-
yhiO_pitA rev		ACCATTAACGCGCTCAAC	-
sspA_rpsI fw	sspA_rpsI	GCGGGTCATATAGCCTTTC	-
sspA_rpsI rev		GGAAGGCGAATACGATCAC	-
pepB_sseB fw	pepB_sseB	GCTTCACCGGCTTAATGG	-
pepB_sseB rev		CGGCTACAGCATCAAACAG	-
flgF_flgG fw	flgF_flgG	GTGGAAGGGCTTTCTCTG	-
flgF_flgG rev		TTCGCCAATGCTCTCC	-
lacZ fw fw	lacZ	TATAGGCGCCAGCAACC	-
lacZ rev rev		ACTGTGAGCCAGAGTTG	-

Table S5: Positions of cut sites in the genome of *E. coli* BL21(DE3) (GenBank accession no. CP001509.3) and *E. coli* W3110 (GenBank accession no. AP009048.1).

Locus	Position of target site in *E. coli* BL21(DE3)	Position of target site in *E. coli* W3110
atpI_rsmG	3,811,627	3,713,854
glmU_atpC	3,804,242	3,721,239
yhiO_pitA	3,501,563	4,001,236
sspA_rpsI	3,238,863	3,377,555
pepB_sseB	2,520,924	2,653,655
flgF_flgG	1,137,219	1,136,229

Table S6: HPLC gradient settings. Mobile phase consists of water with 0.1%(v/v) formic acid (A) and methanol (B). Flow rate was set to 0.5 ml/min and column oven temperature was set to 30 °C.

Time [min]	concentration of methanol (B) [%]
0.01	30
10.00	30
10.01	30
18.00	43
18.01	65
19.00	90
19.01	90
24.00	30

Supporting Figures

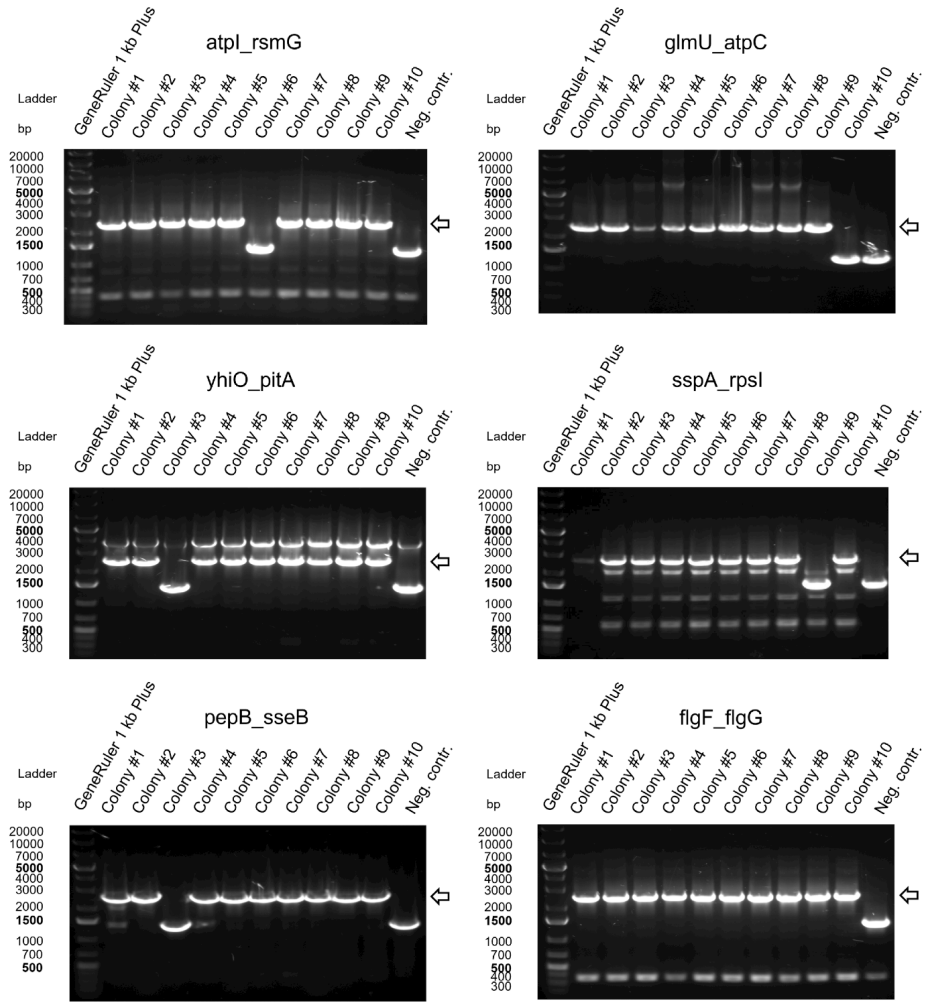


Figure S1: Agarose gels stained with MIDORI Green Advance under UV light. Colony PCR results after chromosomal integration of *mCherry* into different intergenic loci (*E. coli* BL21(DE3)). The amplicon lengths indicating successful integration events are marked by an arrow.

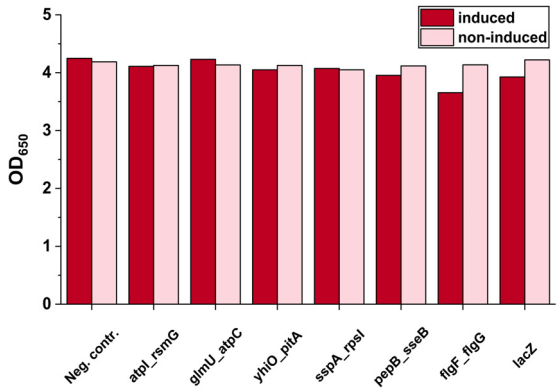


Figure S2. Impact of chromosomal integration and expression of *mCherry* on cell growth of *E. coli* BL21(DE3). *E. coli* BL21(DE3) without integrated genes was used as a negative control. Induced (dark red) and non-induced, basal (light red) expression were compared.

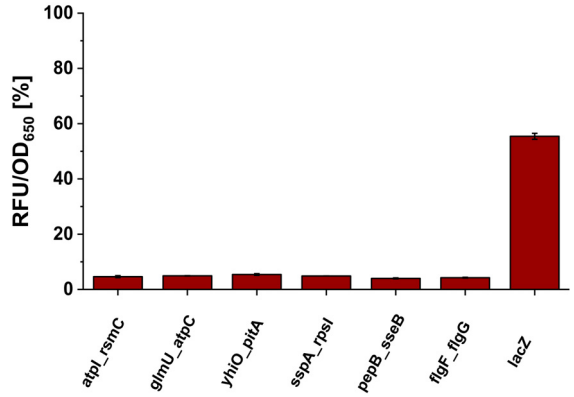


Figure S3. Non-induced (leaky) expression of *mCherry* after chromosomal integration into different genomic loci in *E. coli* BL21(DE3). Relative fluorescence was calculated in relation to induced expression after chromosomal integration into *lacZ* locus shown in Figure 2 in the main manuscript.

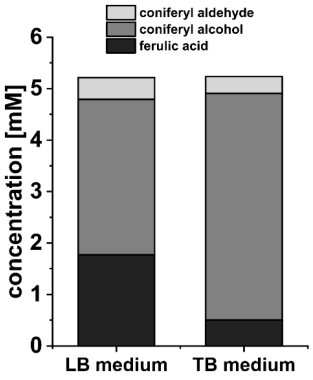


Figure S4: Medium comparison. Lysogeny broth (LB) and terrific broth (TB) were tested for conversion of 5 mM ferulic acid **1** by growing *E. coli* W3110(T7) cells harboring pETDuet_*Zm*CCR_*Pc*4CL. Concentrations of ferulic acid **1**, coniferyl aldehyde **3**, and coniferyl alcohol **4** were quantified 18 h after substrate addition via HPLC.

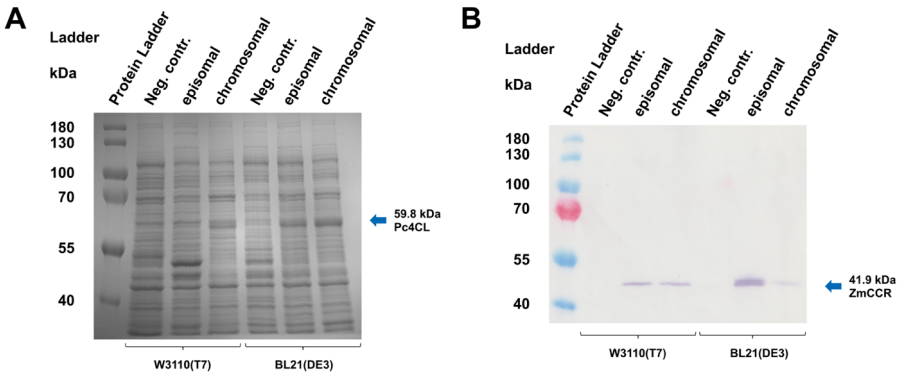


Figure S5: SDS-PAGE (A) and Western Blot (B) of whole cells 18 h after addition of 5 mM ferulic acid **1**. Expression of chromosomally integrated genes was compared to episomal expression using a pETDuet-1 vector. *E. coli* strains harboring no copy of heterologous genes were used as negative controls.

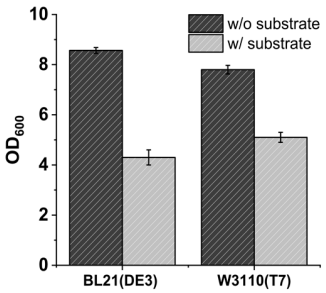


Figure S6: Impact of substrate addition on cell growth of different *E. coli* harboring chromosomally integrated genes encoding *Pc*4CL and *Zm*CCR. Optical density at 600 nm wavelength (OD_{600}) was measured 18 h after addition of 5 mM ferulic acid **1** (dissolved in DMSO). The same amount of DMSO was added to the samples without substrate to ensure comparability. Means and standard deviations were calculated from three biological replicates.

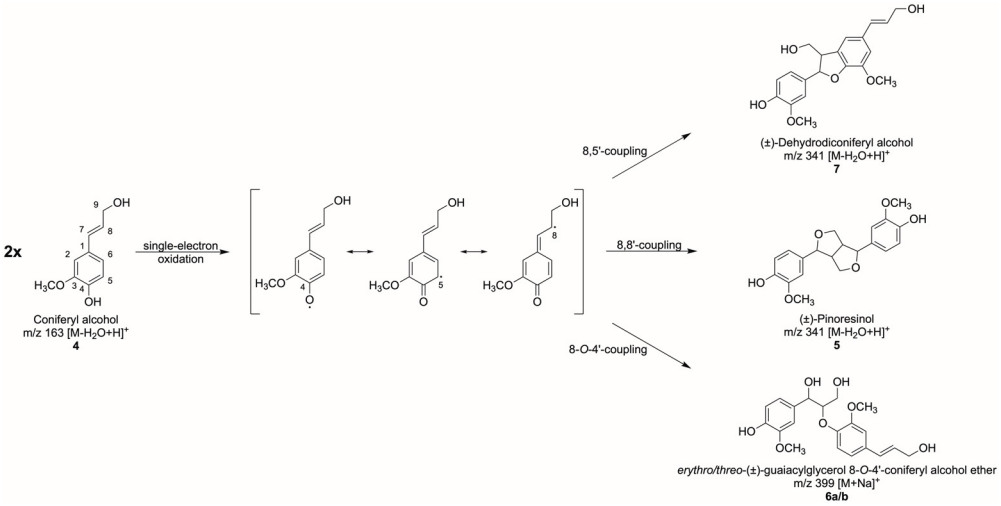


Figure S7: Oxidative radical coupling products of coniferyl alcohol **4**. Pinoresinol **5** originates from the 8,8'-coupling of two monomer radicals. The direction of approach determines which of the two enantiomers of **5** is formed (*si* face recombination results in (+)-pinoresinol, *re* face in (-)-pinoresinol). Thus, a racemic mixture is formed in the absence of dirigent proteins. m/z of base peaks observed in TIC(+) (LC/MS) is annotated to the compounds.

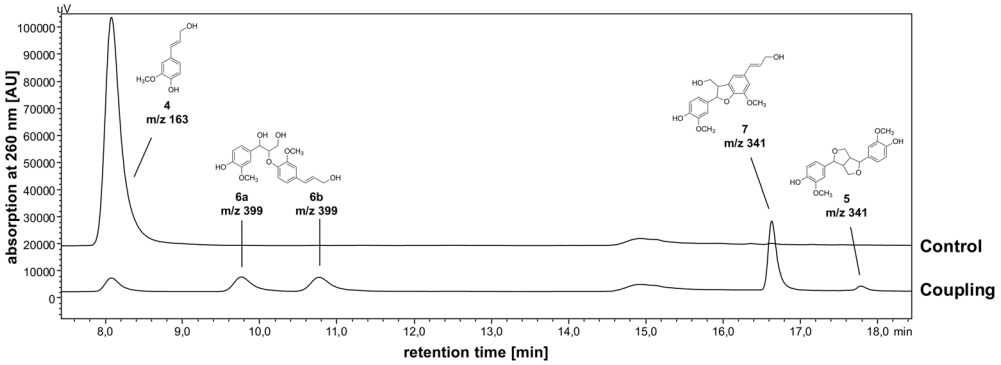


Figure S8: HPLC chromatogram for identification of oxidative coupling products of coniferyl alcohol **4**. Oxidative coupling was performed chemically by using $FeCl_3$ and resulted in the formation of pinoresinol **5**, *erythro*/*threo*-(±)-guaiacylglycerol 8-*O*-4'-coniferyl alcohol ethers **6a/b**, and dehydrodiconiferyl alcohol **7**. Control without addition of $FeCl_3$ shows no coupling products. Products were identified based on m/z according to our previous study (7). m/z of base peaks observed in the corresponding LC/MS TIC(+) are annotated.

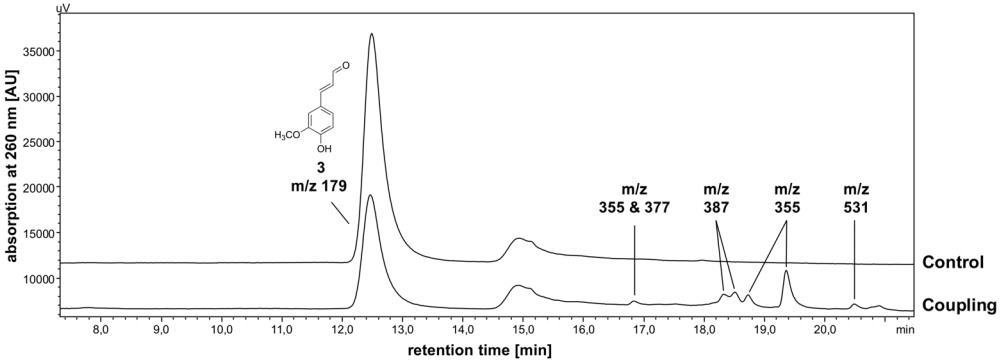


Figure S9: HPLC chromatogram for identification of oxidative coupling products of coniferyl aldehyde **3**. Oxidative coupling was performed chemically by using $FeCl_3$. Control without addition of $FeCl_3$ shows no coupling products. m/z of base peaks observed in the corresponding LC/MS TIC(+) are annotated.

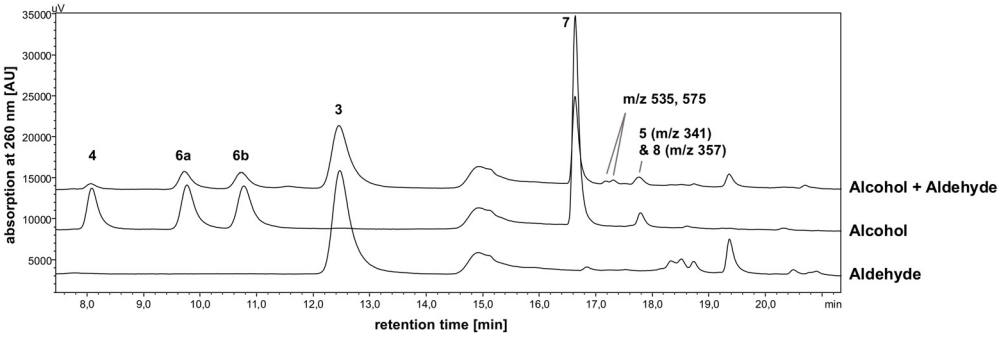


Figure S10: Comparison of HPLC chromatograms after $FeCl_3$-mediated coupling of coniferyl aldehyde **3** and coniferyl alcohol **4** performed separately (cf. Figures S8 & S9) as well as in the same reaction tube in order to identify cross-coupling products. m/z of base peaks observed in the corresponding LC/MS TIC(+) are annotated.

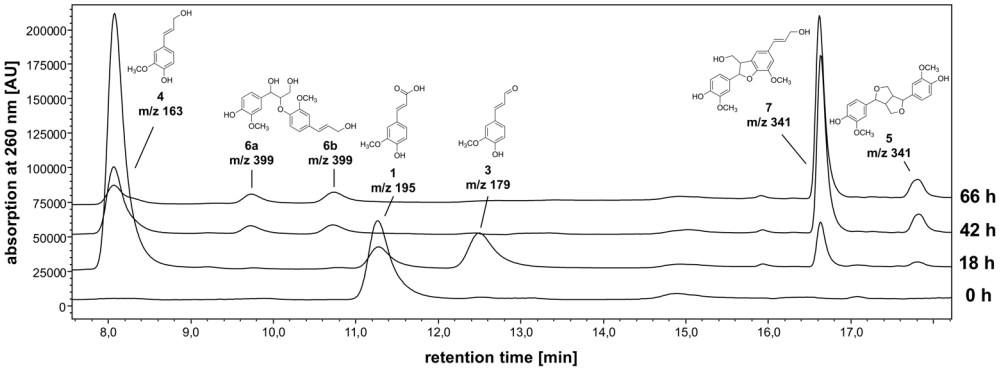


Figure S11: Comparison of HPLC chromatograms at different time points (0 h, 18 h, 42 h, 66 h after substrate addition) during conversion of 5 mM ferulic acid **1** to pinoresinol **5** (cf. Figure 4). The enzymatic cascade proceeds via coniferyl aldehyde **3** and alcohol **4**. Laccase-catalyzed radical C-C-coupling resulted in the formation of pinoresinol **5**, *erythro*/*threo*-(±)-guaiacylglycerol 8-*O*-4'-coniferyl alcohol ethers **6a/b**, and dehydrodiconiferyl alcohol **7**. m/z of base peaks observed in the corresponding LC/MS TIC(+) are annotated.

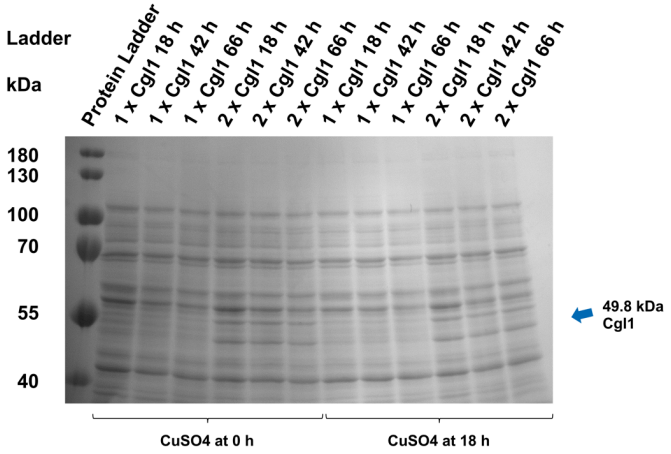


Figure S12: SDS-PAGE of whole cells 18 h, 42 h, and 66 h after addition of 5 mM ferulic acid **1**. $CuSO_4$ addition at time of induction (0 h) was compared to 18 hours after induction. Additionally, different copy numbers of *cgl1* (one or two copies) were compared (cf. Figure 5).

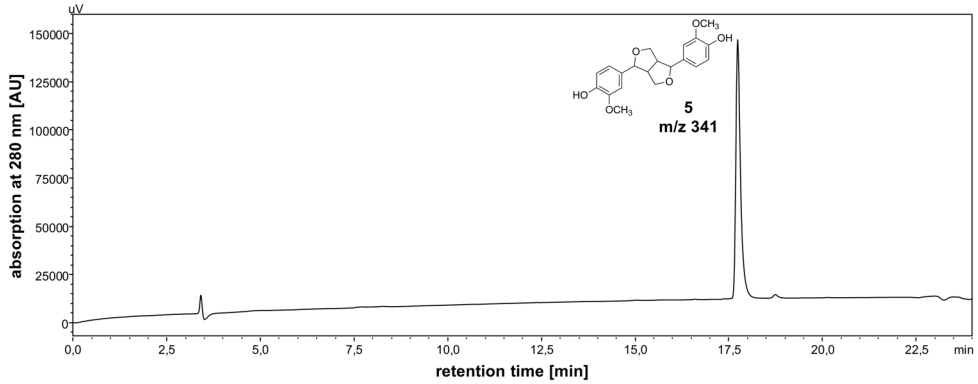


Figure S13: HPLC chromatogram of isolated pinoresinol **5** for determination of purity. m/z of the base peak observed in the corresponding LC/MS TIC(+) is annotated.

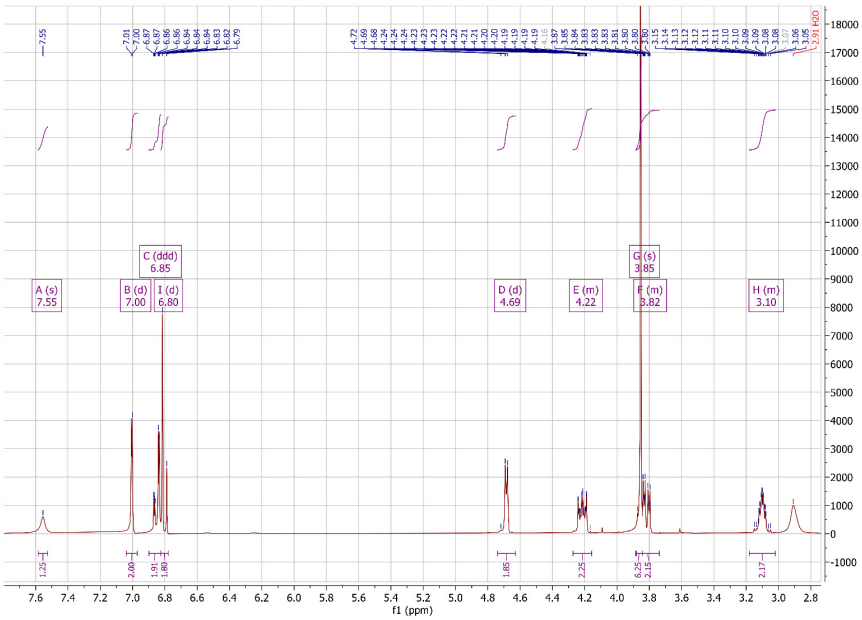


Figure S14: ^{1}H-NMR (300 MHz, Acetone) δ 7.55 (s, 1H), 7.00 (d, J = 1.9 Hz, 2H), 6.85 (ddd, J = 8.1, 1.9, 0.6 Hz, 2H), 6.80 (d, J = 8.0 Hz, 2H), 4.69 (d, J = 4.3 Hz, 2H), 4.27 – 4.16 (m, 2H), 3.85 (s, 6H), 3.88 – 3.74 (m, 2H), 3.18 – 3.02 (m, 2H).

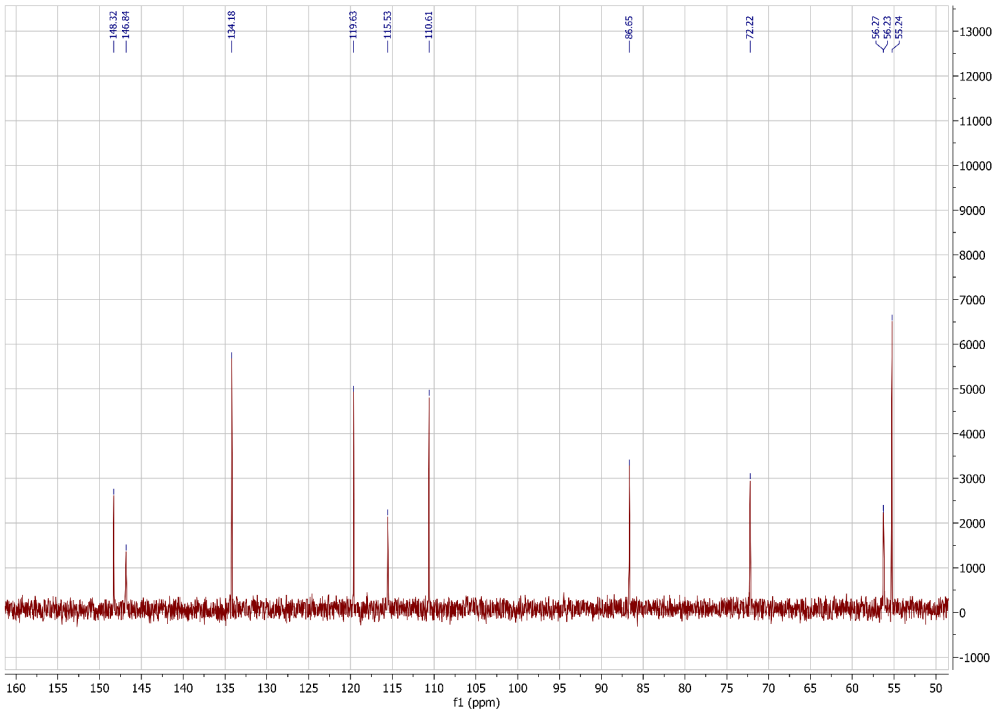


Figure S15: ^{13}C-NMR (75 MHz, Acetone) δ 206.28, 148.32, 146.84, 134.18, 119.63, 115.53, 110.61, 86.65, 72.22, 56.27, 56.23, 55.24.

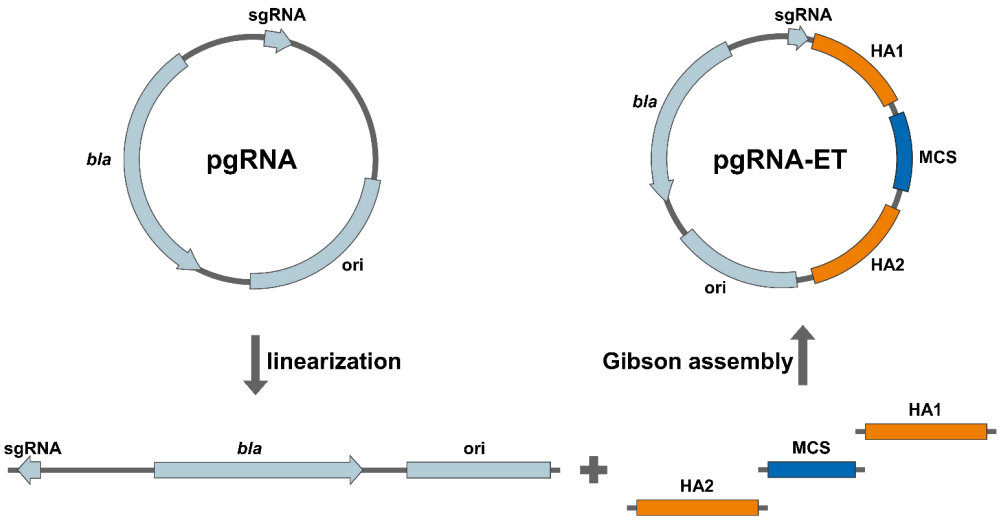


Figure S16. Construction of the plasmid series pgRNA-ET. The plasmid pgRNA (left) contains a locus-specific single-guide RNA (sgRNA) under control of a constitutive promoter, a ColE1 origin of replication (ori) as well as a gene encoding the β-lactamase (*bla*) for ampicillin resistance (3). PCR was used to linearize the plasmid and introduce overlaps for Gibson assembly. Locus-specific homology arms of ~500 bp length (HA1 & HA2) for homologous recombination during the chromosomal integration process were amplified from *E. coli*'s genome. The expression cassette containing the multiple cloning site (MCS) originated from pET-28a(+). All four fragments were fused by Gibson assembly to result in the pgRNA-ET plasmid series.

References

1. Luelf UJ, Böhmer LM, Li S, Urlacher VB. Effect of chromosomal integration on catalytic performance of a multi-component P450 system in *Escherichia coli*. Biotechnol Bioeng. 2023;120(7):1762-72.
2. Li Q, Sun B, Chen J, Zhang Y, Jiang Y, Yang S. A modified pCas/pTargetF system for CRISPR-Cas9-assisted genome editing in *Escherichia coli*. Acta Biochimica et Biophysica Sinica. 2021;53(5):620-7.
3. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell. 2013;152(5):1173-83.
4. Jansen F, Gillessen B, Mueller F, Commandeur U, Fischer R, Kreuzaler F. Metabolic engineering for *p*-coumaryl alcohol production in *Escherichia coli* by introducing an artificial phenylpropanoid pathway. Biotechnol Appl Biochem. 2014;61(6):646-54.
5. Ricklefs E, Girhard M, Koschorreck K, Smit MS, Urlacher VB. Two-Step One-Pot Synthesis of Pinoresinol from Eugenol in an Enzymatic Cascade. ChemCatChem. 2015;7(12):1857-64.
6. Seo J-H, Baek S-W, Lee J, Park J-B. Engineering *Escherichia coli* BL21 Genome to Improve the Heptanoic Acid Tolerance by using CRISPR-Cas9 System. Biotechnology and Bioprocess Engineering. 2017;22(3):231-8.
7. Decembrino D, Girhard M, Urlacher VB. Use of Copper as a Trigger for the *in Vivo* Activity of *E. coli* Laccase CueO: A Simple Tool for Biosynthetic Purposes. ChemBioChem. 2021;22(8):1470-9.

2.4 Combinatorial promoter replacement after chromosomal integration – A proof-of-concept

Title: Combinatorial promoter replacement after integration of heterologous genes into the *Escherichia coli* chromosome

Authors: **U. J. Luelf**, V. B. Urlacher

Manuscript in preparation

Contribution: Conceptualized the study, designed and conducted the experiments, analyzed the data and prepared the first draft.

Combinatorial promoter replacement after integration of heterologous genes into the *Escherichia coli* chromosome

U. Joost Luelf[1], Vlada B. Urlacher[1]*

[1]Institute of Biochemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

*Corresponding Author

Email: vlada.urlacher@uni-duesseldorf.de

ORCID: Vlada B. Urlacher 0000-0003-1312-4574

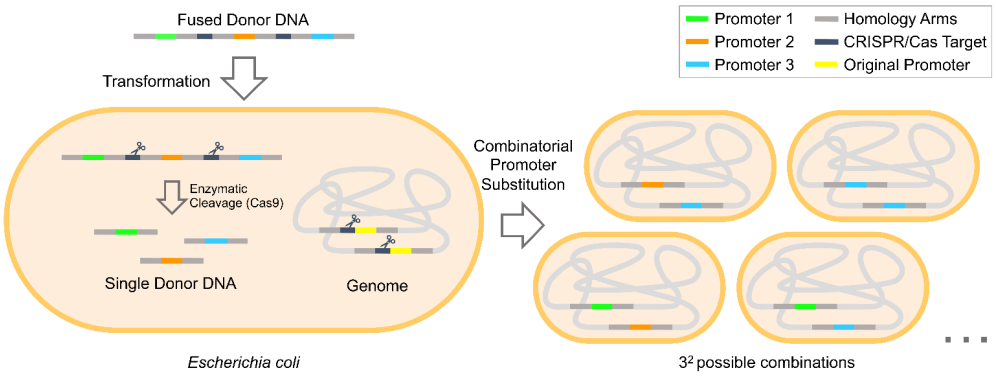
Running title: Promoter substitution

ABSTRACT

Chromosomal integration of heterologous genes into well-studied hosts like *Escherichia coli* is one way to establish stable production strains. However, the metabolic flux in newly constructed or reconstituted pathways as well as the catalytic activity of multi-component enzymatic systems is rarely optimal from the very beginning and depends on both enzyme activity and expression levels of the individual genes. While enzyme activity can be optimized by well-established protein engineering methods using plasmids prior to chromosomal integration, gene expression may be improved, among other things, by varying the promoter strength. The optimal choice of promoters for chromosomal expression systems, however, can often not be predicted based on previous plasmid-based experiments. Moreover, the simultaneous expression of multiple genes can affect each other, which can quickly lead to a high-dimensional optimization problem. Consequently, either all selected promoter combinations must already be considered during integration, or the promoters can be exchanged iteratively after successful gene integration into the genome. Both strategies create an enormous workload. This *Technical Note* presents a simple method that enables a one-step combinatorial CRISPR/Cas-mediated promoter substitution after heterologous genes have been successfully integrated into the genome of *Escherichia coli*.

KEYWORDS: promoter substitution, chromosomal integration, *E. coli*, CRISPR/Cas

1 **Table Of Contents Figure**



2

3

Chromosomal integration of heterologous genes to entire pathways can be advantageous compared to plasmid-based expression in *Escherichia coli*. In addition to eliminating the need for selectable markers such as antibiotics and reducing cell-to-cell variability, chromosomal integration also allows finer tuning of multi-enzymatic systems or cascades.[1-4] The introduction of CRISPR/Cas-assisted recombineering greatly facilitated these integration procedures.[5, 6] There, a Cas endonuclease can be programmed to target almost every region of *E. coli*'s chromosome and introduce lethal double-strand breaks which can be used as a counter-selection tool for recombineering.[7, 8] However, each genome editing procedure still requires a certain amount of time and effort and the iterative integration of multiple heterologous genes may add up to several weeks of work. Unfortunately, the balance in a multi-component system is rarely optimal at the very first attempt. Whether it is the metabolic flux in a reconstituted or newly designed pathway, or the ratios of redox partner proteins necessary for enzyme activity, tuning of the individual components is often required. In this context, both the enzyme activity and the expression level are relevant. Whereas the first is defined by the enzyme's sequence, the latter can be altered, for example, by changing the rate of transcription and testing different promoter designs.[9] Once the heterologous genes have been integrated into the chromosome, iterative promoter substitution results in a significant amount of additional work. Even with only two integrated genes and three promoters to test (cf. table of contents figure), there are nine possible combinations that must be introduced in two iterative steps targeting one gene at a time when using classical methods. Furthermore, only some of these promoter combinations will show positive results with regard to the performance of the multi-component system, making most of the work obsolete in retrospect. This also applies to the approach of considering all promoter combinations already when integrating the individual genes. Thus, in this *Technical Note*, a brief description of a method is presented, that allows fast and simultaneous one-step combinatorial exchange of promoters of previously integrated genes.

The described method relies on the well-known λ-Red-mediated homology-directed repair pathway. Here, the original promoter is removed by the insertion of a new promoter in combination with counterselection by CRISPR/Cas9. Homology-directed repair requires the introduction of one DNA repair template (donor DNA) per promoter into the cells.[5] However, the probability that not one but two or even more different DNA fragments will be incorporated into a cell simultaneously during transformation decreases exponentially with the number of fragments. Consequently, the number of cells that accommodate only one of the repair templates is disproportionately large. For the intended promoter exchange at multiple sites, this would result in the same promoter being integrated at all sites if a mixture of separate templates for each promoter is used (Figure 1A). Existing methods such as multiplex automated genome engineering (MAGE) circumvent this problem by repeating the method in several cycles.[10] In contrast, in the method described here, this repetition can be omitted because a single DNA fragment is used instead of a mixture. This fused donor DNA combines the individual repair templates for each promoter separated by a CRISPR/Cas target sequence and is cleaved in the cells (Figure 1B).

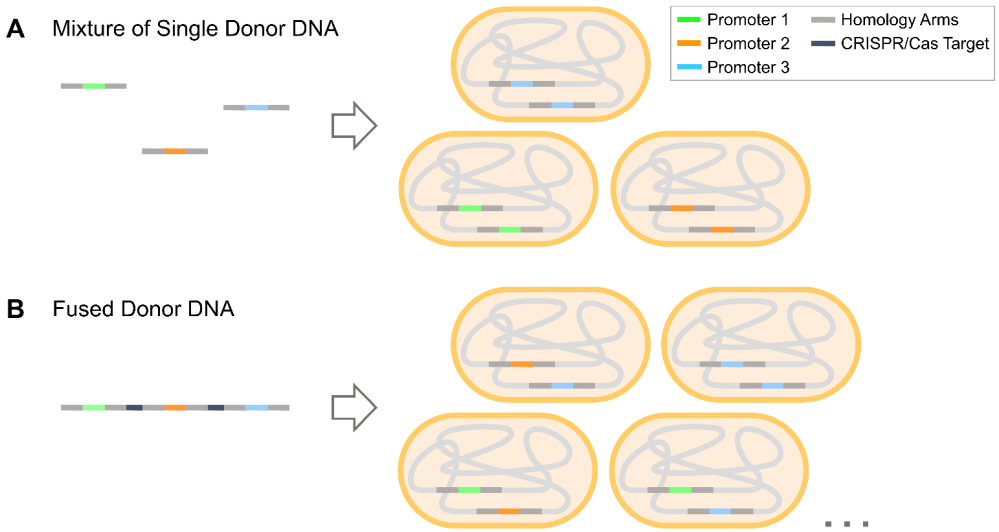


Figure 1: Comparison of promoter substitution approaches. (A) Application of a mixture of different linear DNA fragments for each promoter results in a high likelihood that only one promoter is used to replace the original promoter at all sites. (B) The use of fused donor DNA allows the substitutive integration of different promoters at different genomic loci.

In the following, the individual steps of the method are described in more detail. In a previous study, we provided primer sequences which allow easy amplification and CRISPR/Cas-assisted chromosomal integration of genes located on pET vectors.[11] The integration of several heterologous genes is performed iteratively according to the known CRISPR/Cas-assisted recombineering method using site-specific 500 bp homology arms (Figure 2A). Along with the genes, the T7lac promoter from the pET vector is also amplified and integrated into the chromosome. Although this promoter is known to lead to overexpression of heterologous genes, it is not the optimal choice in all cases and might lead to inclusion bodies or suffer from inactivation of the T7 RNA polymerase gene.[12-14] Thus, in this exemplary demonstration, T7lac was

substituted by tac and T5lac promoter which are recognized by *E. coli*'s own RNA polymerase instead. For CRISPR/Cas-assisted combinatorial promoter replacement, a target (spacer) sequence upstream of the original T7lac promoter was selected and two short homology arms (59 bp and 45 bp) were identified that are present among all integrated heterologous genes (Figure 2B). Although longer homology arms are correlated with higher editing efficiency in general,[6, 15] these short homology arms are sufficient for successful promoter replacement due to the overall shortness of the repair template. Preliminary tests for single substitution of promoters at a single site showed a high editing efficiency of 100% (n = 5) for all three promoters tested (Figure S1). Apart from the T5lac and tac promoters, we decided to also re-introduce T7lac, which would otherwise have been completely excluded from further experiments due to the high targeting efficiency of the gRNA; the chosen target sequence resulted in none to only a few escaper colonies when no DNA repair template was added. In addition, no off-target activity was observed in a test with a strain without the target sequence (data not shown).

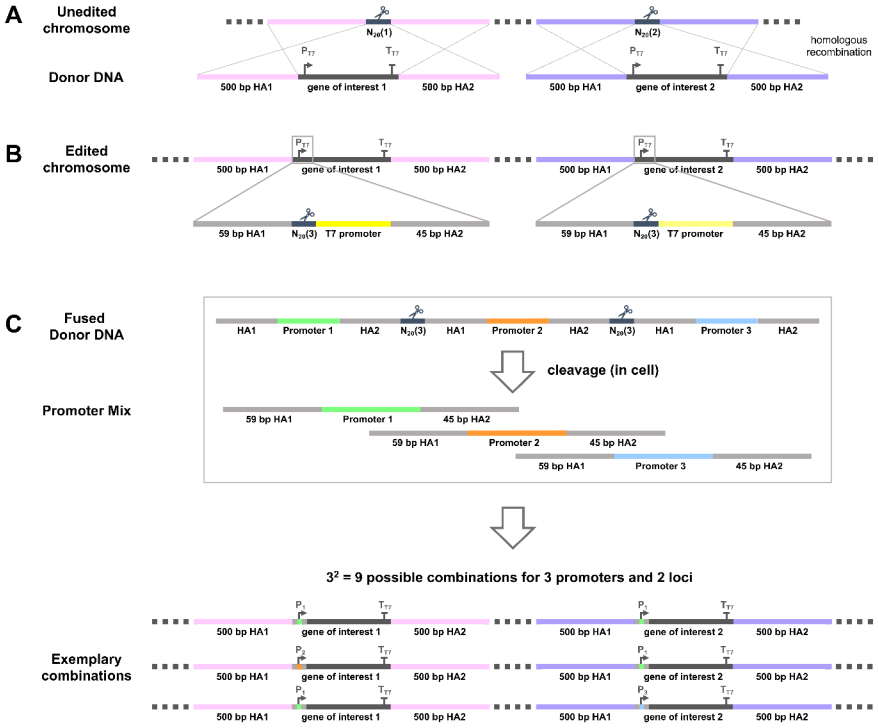


Figure 2: Mechanism of combinatorial promoter substitution after chromosomal integration of heterologous genes. (A) Two different genes of interest amplified from pET vectors are integrated into the chromosome using CRISPR/Cas-assisted recombineering. The donor DNA contains locus-specific 500 bp homology arms flanking the gene of interest for λ-red-mediated homologous recombination. The CRISPR/Cas target sequences (N_{20}(1) and N_{20}(2)) depend on the integration site. (B) Along with the gene, the T7lac promoter originating from the pET vector is introduced as described before.[11] This part of the sequence is therefore identical for all integrated genes at all sites. Upstream of the T7lac promoter, a CRISPR/Cas target sequence (N_{20}(3)) for promoter substitution is selected and surrounding short homology arms (59 bp HA1 & 45 bp HA2) are designed. (C) A fused donor DNA containing all repair templates separated by the same target sequence (N_{20}(3)) is cleaved into individual repair templates (promoter mix) in the cell and yields different promoter combinations.

For combinatorial promoter substitution, the fused donor DNA containing repair templates for T7lac, T5lac and tac insertion was synthesized and initially tested for substitution at one site in the genome. Each promoter repair template was separated by the targeting sequence so that the fused construct is cut in the cells. This design also eliminates the possibility of multiple promoters being inserted in sequence at the same site. If the template is not cleaved before insertion into the chromosome, it would represent a target for CRISPR/Cas cleavage again. Successful promoter substitution was verified by colony PCR with specific primers for each promoter. Of 20 colonies tested, 14 contained the newly integrated T7lac promoter, and only 4 and 2 times the tac or T5lac promoter were found, respectively (Figure S2). This uneven distribution is most likely due to the higher homology of the T7lac repair template to the target site. The fact that the original promoter and the promoter to be inserted are identical (T7lac) results in a longer homology arm (91 bp compared to 45 bp) downstream of the double-strand break. An inefficient substitution and false positive results caused by "wild type" T7lac could be excluded by specific primer design.

Nonetheless, simultaneous promoter substitution at two different sites was tested. For this purpose, a previously described bacterial three-component cytochrome P450 system was chosen as a model case.[11] The activity of CYP154E1 from *Thermobifida fusca* YX was reconstituted in *E. coli* based on two redox partners – the flavodoxin YkuN from *Bacillus subtilis* and the FAD-containing reductase FdR from *E. coli*. Since the latter is already present in the host, only the genes encoding the P450 and the flavodoxin were integrated into the chromosome as described and subject to promoter substitution. For CYP154E1, all three promoters were found, with a distribution of 87% T7lac, 10% tac, and 3% T5lac. For the flavodoxin, however, only T7lac (82%) and tac (18%) and no T5lac were found. In total, only six of the nine possible combinations were identified in a screening of 94 colonies. The statistical distribution

with high preference for T7lac would therefore require the screening of an even larger number of colonies to find all nine desired combinations. However, it should be noted that the PCR-based genotyping used for this proof of concept creates an enormous workload and a high-throughput screening for the desired phenotypic properties (e.g. metabolic flux, conversion, or enzyme concentration) would be highly preferred. In our model case, the replacement of the T7lac promoter by T5lac or tac resulted in lower P450 concentrations (Figure S3). Furthermore, it has been demonstrated that the endogenous copy of the gene encoding flavodoxin reductase is not sufficient for catalysis.[11] Thus, no phenotypic screening based on substrate conversion was established for this system. Nevertheless, the fundamental proof of concept using a fused donor DNA for simultaneous and combinatorial promoter substitution was successfully demonstrated. In further experiments, different promoters for a multi-component system eligible for a phenotypic high-throughput screening should be tested, the limits of multiplexing should be determined, and the extent to which the order of promoters in the fused donor DNA affects the statistical distribution should be investigated.

AUTHOR CONTRIBUTIONS

U.J.L conceptualized the study, designed and conducted the experiments, analyzed the data and prepared the first draft. V.B.U. contributed to conceiving experiments, interpretation of data, and writing of the manuscript. Both authors edited the manuscript and approved the final version of the manuscript.

1 **ACKNOWLEDGMENT**

2 We would like to thank Martin Sosna for assisting in the colony PCR screening. Financial support was
3 provided by the state of North Rhine-Westphalia (NRW) and the “European Regional Development Fund
4 (EFRE)”, Project “ClusterIndustrial Biotechnology (CLIB) Kompetenzzentrum Biotechnologie (CKB)”
5 (34.EFRE-0300095/1703FI04).

6 **CONFLICT OF INTEREST STATEMENT**

7 The authors declare no conflict of interest.

8 **DATA AVAILABILITY STATEMENT**

9 Data are available from the corresponding author on reasonable request.

REFERENCES

(1) Yadav, V. G.; De Mey, M.; Lim, C. G.; Ajikumar, P. K.; Stephanopoulos, G. The future of metabolic engineering and synthetic biology: towards a systematic practice. *Metab Eng* **2012**, *14* (3), 233-241. DOI: 10.1016/j.ymben.2012.02.001.

(2) Mairhofer, J.; Scharl, T.; Marisch, K.; Cserjan-Puschmann, M.; Striedner, G. Comparative Transcription Profiling and In-Depth Characterization of Plasmid-Based and Plasmid-Free *Escherichia coli* Expression Systems under Production Conditions. *Appl Environ Microbiol* **2013**, *79* (12), 3802-3812. DOI: 10.1128/AEM.00365-13.

(3) Englaender, J. A.; Jones, J. A.; Cress, B. F.; Kuhlman, T. E.; Linhardt, R. J.; Koffas, M. A. G. Effect of Genomic Integration Location on Heterologous Protein Expression and Metabolic Engineering in *E. coli*. *ACS Synth Biol* **2017**, *6* (4), 710-720. DOI: 10.1021/acssynbio.6b00350.

(4) Ajikumar, P. K.; Xiao, W. H.; Tyo, K. E.; Wang, Y.; Simeon, F.; Leonard, E.; Mucha, O.; Phon, T. H.; Pfeifer, B.; Stephanopoulos, G. Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* **2010**, *330* (6000), 70-74. DOI: 10.1126/science.1191652.

(5) Jiang, Y.; Chen, B.; Duan, C.; Sun, B.; Yang, J.; Yang, S. Multigene Editing in the *Escherichia coli* Genome via the CRISPR-Cas9 System. *Appl Environ Microbiol* **2015**, *81* (7), 2506-2514. DOI: 10.1128/AEM.04023-14.

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

(7) Jiang, W.; Bikard, D.; Cox, D.; Zhang, F.; Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* **2013**, *31* (3), 233-239. DOI: 10.1038/nbt.2508.

(8) Ao, X.; Yao, Y.; Li, T.; Yang, T.-T.; Dong, X.; Zheng, Z.-T.; Chen, G.-Q.; Wu, Q.; Guo, Y. A Multiplex Genome Editing Method for Escherichia coli Based on CRISPR-Cas12a. *Front Microbiol* **2018**, *9*, 2307-2307. DOI: 10.3389/fmicb.2018.02307.

(9) Xu, N.; Wei, L.; Liu, J. Recent advances in the applications of promoter engineering for the optimization of metabolite biosynthesis. *World J Microbiol Biotechnol* **2019**, *35* (2), 33. DOI: 10.1007/s11274-019-2606-0.

(10) Wang, H. H.; Isaacs, F. J.; Carr, P. A.; Sun, Z. Z.; Xu, G.; Forest, C. R.; Church, G. M. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **2009**, *460* (7257), 894-898. DOI: 10.1038/nature08187.

(11) Luelf, U. J.; Böhmer, L. M.; Li, S.; Urlacher, V. B. Effect of chromosomal integration on catalytic performance of a multi-component P450 system in *Escherichia coli*. *Biotechnol Bioeng* **2023**, *120* (7), 1762-1772. DOI: 10.1002/bit.28404.

(12) Studier, F. W.; Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **1986**, *189* (1), 113-130. DOI: https://doi.org/10.1016/0022-2836(86)90385-2.

(13) Vethanayagam, J. G.; Flower, A. M. Decreased gene expression from T7 promoters may be due to impaired production of active T7 RNA polymerase. *Microb Cell Fact* **2005**, *4* (1), 3. DOI: 10.1186/1475-2859-4-3.

(14) Angius, F.; Ilioaia, O.; Amrani, A.; Suisse, A.; Rosset, L.; Legrand, A.; Abou-Hamdan, A.; Uzan, M.; Zito, F.; Miroux, B. A novel regulation mechanism of the T7 RNA polymerase based expression system improves overproduction and folding of membrane proteins. *Sci Rep* **2018**, *8* (1), 8572. DOI: 10.1038/s41598-018-26668-y.

(15) Bassalo, M. C.; Garst, A. D.; Halweg-Edwards, A. L.; Grau, W. C.; Domaille, D. W.; Mutalik, V. K.; Arkin, A. P.; Gill, R. T. Rapid and Efficient One-Step Metabolic Pathway Integration in *E. coli*. *ACS Synth Biol* **2016**, *5* (7), 561-568. DOI: 10.1021/acssynbio.5b00187.

2.4.1 Supporting information

1 **Supporting information**

2 **Combinatorial promoter replacement after integration of heterologous genes into the**

3 ***Escherichia coli* chromosome**

4 U. Joost Luelf[1], Vlada B. Urlacher[1]*

5 [1]Institute of Biochemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

6

7 *Corresponding Author

8 Email: vlada.urlacher@uni-duesseldorf.de

9 ORCID: Vlada B. Urlacher 0000-0003-1312-4574

10 Running title: Promoter replacement

1. METHODS

1.1. Strains and plasmids

E. coli DH5α (Clontech) was used as cloning host and for plasmid propagation. *E. coli* W3110(T7) mCherry (atpI_rsmG),[1] *E. coli* W3110(T7) CYP154E1 QAA (nupG)[2] and *E. coli* W3110(T7) CYP154E1 QAA (nupG) YkuN (atpI_rsmG)[2] were used for genome engineering and/or heterologous expression of a *cyp154e1* triple mutant described elsewhere.[2] Plasmids pEcCas (Addgene plasmid #73227),[3] and pgRNA-bacteria (Addgene plasmid #44251)[4] were derived from Addgene.

1.2. CRISPR/Cas-assisted promoter substitution

The N_{20} targeting sequence of the sgRNA encoded on pgRNA-bacteria plasmid was exchanged as described before.[2] For single promoter exchange, single-stranded oligonucleotides were ordered and fused in a standard PCR reaction to obtain the individual double-stranded DNA repair templates for each promoter. A list of all oligonucleotides used in this study can be found in Table S1. For combinatorial promoter exchange, the promoter DNA template was synthesized and cloned into a pUC-57 vector by BioCat GmbH (Heidelberg). The annotated sequence can be found in Figure S4. For recombination, the fused promoter DNA template was amplified by standard PCR to obtain a higher amount of linear donor DNA. Promoter substitution was achieved by following the procedure described for chromosomal integration before.[2] Briefly, electrocompetent cells harboring pEcCas for expression of *cas9* and λ-red genes were prepared and 400-500 ng of the fused promoter DNA template as well as the pgRNA plasmid were co-electroporated. Cells were recovered after electroporation and plated onto selective LB agar plates. Promoter substitution was verified by separate colony PCRs using specific primers for each possible promoter and site.

1.3. Expression and quantification of P450

To quantify the effect of different promoters on P450 expression, 50 ml Terrific Broth (TB) medium were inoculated with 1 ml of the overnight cultures and incubated at 37 °C, 180 rpm. The experimental details of expression and cell disruption to obtain a crude extract can be found elsewhere.[2] The quantification of functional, heme-loaded P450 via carbon monoxide difference spectra was carried out as described by Omura and Sato.[5, 6]

2. Supporting Tables

Table S1. Oligonucleotide sequences. Restriction sites are underlined. N_{20} targeting sequence is shown in italic.

Name	Sequence (5‘ to 3‘)	Reference
gRNA Primer		
gRNA rev	ATATATACTAGTATTATACCTAGGACTGAGCTAG	general design [7]
gRNA_XXX fw (general primer design)	ATATATACTAGT-*N_{20}*-GTTTTAGAGCTAGAAATAGC	general design [7]
gRNA _001 fw (nupG BL21(DE3))	ATATATACTAGT*GAGATCTCGATCCTCTACGC*GTTTTAGAGCTAGAAATAGC	
pgRNA seq	GCCACCTGACGTCTAAG	[2]
Primers for synthesis of donor DNA for single promoter exchange		
T5lac fw	TATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGG CGTCATAAAAAATTTATTTGCTTTGTGAGCGGATAAC	
T5lac rev	CCATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGATATTATAATTGTTATC CGCTCACAAAGC	
tac fw	TATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGG CGTTGACAATTAATCATCGGCTCGTATAATGTG	
tac rev	CCATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAAATTGTTATCCGCTCA CAATTCCACACATTATACGAGCCGATG	
T7lac fw	TATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGG CGTAATACGACTCACTATAGGGGAATTGTGAGCGGATAAC	
T7lac rev	CCATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATCC GCTCACAATTCCC	
Primers for amplification of fused promoter DNA template		
pUC57 fw	GTACCTCGCGAATGCATC	
pUC57 rev	GACCATGATTACGCCAAG	
Primers for colony PCR		
T5lac	GTTATCCGCTCACAAAGC	
tac	CACATTATACGAGCCGATG	
T7lac (new)	CTATAGTGAGTCGTATTACGCC	
T7lac (original)	TAATTTCGCGGGATCGAG	
atpI_rsmG locus	TGCCCGAAGAATATCAGG	
nupG locus	AGCACTTTCCGCCATTC	
sspA_rpsI locus	GCGGTTAAAGGCATGTTG	

3. Supporting Figures

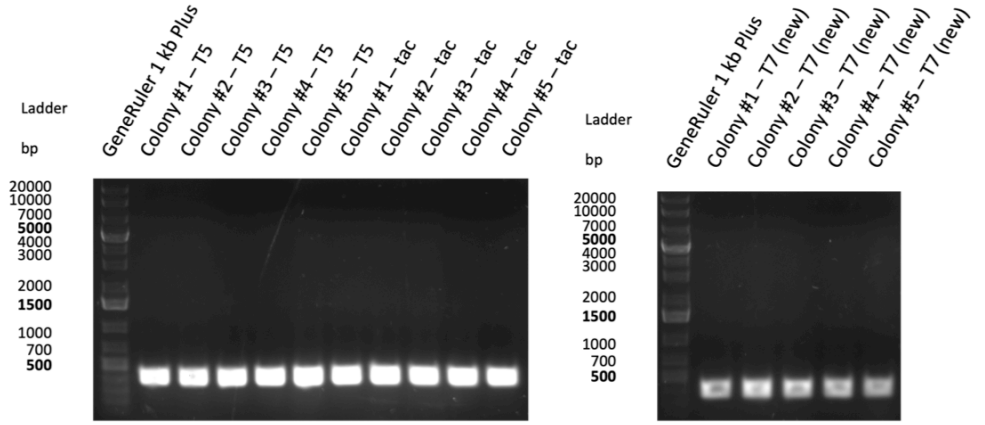


Figure S1: PCR-based verification of promoter substitution using single donor DNA. Promoter-specific primers can be found in Table S1.

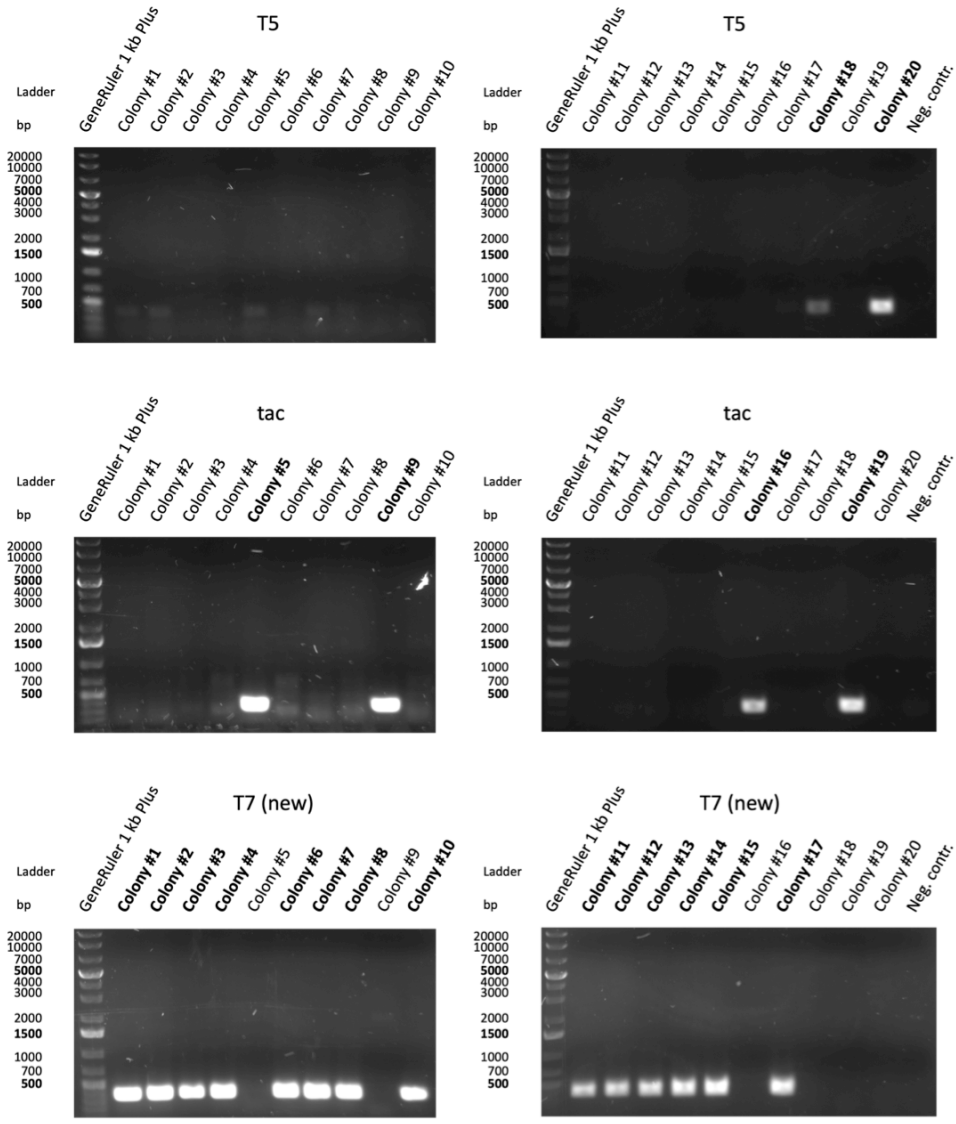


Figure S2: Exemplary PCR-based verification of promoter substitution using fused donor DNA. Promoter-specific primers can be found in Table S1.

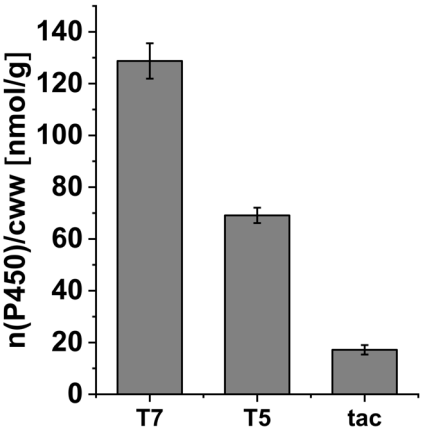


Figure S3: Amount of soluble P450 (n(P450)) per gram cell wet weight (cww) quantified by carbon monoxide difference spectra after chromosomal expression of *cyp154e1* triple mutant under control of three different promoters. Means and standard deviation were calculated from four biological replicates.

N20 targeting sequence Homology arm 1 Promoter Homology arm 2

CCGGCGTAGAGGATCGAGATCTCTATAGGCGCCAGCAACCGCACCTGTGGCGCC
GGTGATGCCGGCCACGATGCGTCCGGCGTCATAAAAAATTTATTTGCTTTGTGAG
CGGATAACAATTATAATATCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACC
ATGGCCGGCGTAGAGGATCGAGATCTCTATAGGCGCCAGCAACCGCACCTGTGG
CGCCGGTGATGCCGGCCACGATGCGTCCGGCGTAATACGACTCACTATAGGGGA
ATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGA
GATATACCATGGCCGGCGTAGAGGATCGAGATCTCTATAGGCGCCAGCAACCGCA
CCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGGCGTTGACAATTAATCAT
CGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCTAGAAATAATTTTGT
TTAACTTTAAGAAGGAGATATACCATGGCCGGCGTAGAGGATCGAGATCTC

Figure S4: Annotated sequence of fused donor DNA with three different promoters (T5lac, T7lac, tac).

References

(1) Luelf, U. J.; Wassing, A.; Böhmer, L. M.; Urlacher, V. B. Plasmid-free production of the plant lignan pinoresinol in growing *Escherichia coli* cells. *Microb Cell Fact* **2024**, *23* (1). DOI: 10.1186/s12934-024-02562-3.

(2) Luelf, U. J.; Böhmer, L. M.; Li, S.; Urlacher, V. B. Effect of chromosomal integration on catalytic performance of a multi-component P450 system in *Escherichia coli*. *Biotechnol Bioeng* **2023**, *120* (7), 1762-1772. DOI: 10.1002/bit.28404.

(3) Li, Q.; Sun, B.; Chen, J.; Zhang, Y.; Jiang, Y.; Yang, S. A modified pCas/pTargetF system for CRISPR-Cas9-assisted genome editing in *Escherichia coli*. *Acta Biochim Biophys Sin* **2021**, *53* (5), 620-627. DOI: 10.1093/abbs/gmab036.

(4) Qi, L. S.; Larson, M. H.; Gilbert, L. A.; Doudna, J. A.; Weissman, J. S.; Arkin, A. P.; Lim, W. A. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **2013**, *152* (5), 1173-1183. DOI: 10.1016/j.cell.2013.02.022.

(5) Omura, T.; Sato, R. The Carbon Monoxide-Binding Pigment of Liver Microsomes. I. Evidence for Its Hemoprotein Nature. *J Biol Chem* **1964**, *239* (7), 2370-2378.

(6) Omura, T.; Sato, R. The Carbon Monoxide-Binding Pigment of Liver Microsomes. II. Solubilization, Purification, and Properties. *J Biol Chem* **1964**, *239* (7), 2379-2385.

(7) Seo, J.-H.; Baek, S.-W.; Lee, J.; Park, J.-B. Engineering *Escherichia coli* BL21 Genome to Improve the Heptanoic Acid Tolerance by using CRISPR-Cas9 System. *Biotechnol Bioprocess Eng* **2017**, *22* (3), 231-238. DOI: 10.1007/s12257-017-0158-4.

3 General Discussion and Outlook

The applicability of biocatalysts depends not only on enzymes with the right properties in terms of activity and selectivity, but also on the overall productivity of the entire biocatalytic system. This productivity is determined, among other things, by the concentration of the enzyme(s). For *in vitro* biocatalysis, the enzyme concentration can be easily adjusted by adding more of the isolated (purified) enzyme. However, in *in vivo* systems, enzyme concentration is set by the expression level which can be adjusted, for example, by altering the gene copy number or choosing a different promoter. In addition, the productivity of multi-component systems like the P450 system described in Chapter 2.2 or the biosynthetic pathway described in Chapter 2.3 is usually limited by their weakest component. Therefore, optimization of the interaction of the individual components is of great importance. When using whole cells or crude extracts, another issue can arise: other enzymes present in the biocatalytic system can degrade (co-)substrates, intermediates or even the product, reducing the yield of the target compound.

To overcome these challenges, within this thesis, CRISPR/Cas-assisted recombineering was applied in various ways for the development and optimization of different *E. coli*-based biocatalysts. These biocatalysts ranged from growing *E. coli* cells in precursor fermentation to resting cells for biotransformation and cell-free extracts for enzyme catalysis. In order to provide a comprehensive and comprehensible discussion of the results, the following discussion is divided into two sub-chapters, one dealing with aspects of the results regarding biocatalysis and the other focusing on genome engineering.

3.1 Biocatalysis

P450s and UPOs are both promising biocatalysts for late-stage oxyfunctionalization of non-activated carbon centers. Their activity depends on either redox partners, molecular oxygen and the cofactor NAD(P)H in case of P450 monooxygenases, or the co-substrate hydrogen peroxide in case of peroxygenases. Although the latter is preferred from an economical perspective, it can affect the enzyme's stability and is quickly degraded by *E. coli*'s catalases katE (hydroperoxidase II) and katG (hydroperoxidase I). To avoid this competitive co-substrate degradation, purified enzymes are usually used to study peroxygenase-catalyzed reactions. However, the purification process involves a considerable amount of work and is associated with enzyme losses and stability issues. Especially in the iterative optimization of peroxygenases using protein engineering, this can quickly become an overwhelming challenge. Thus, in the first peer-reviewed manuscript (Chapter 2.1), a simple in-frame deletion of both catalase genes in *E. coli* BL21-Gold(DE3) was performed. Due to the reduced competitive degradation of H_2O_2, crude extracts and whole cells can now be used for H_2O_2-driven biocatalysis, avoiding the need for enzyme purification. However, in this context, the biological role of catalases and the effects of their knockout on *E. coli*-based biocatalysts should be discussed. Although no or only minor effects on cell growth due to the knockout in the absence of H_2O_2 were reported both in the literature[192, 193] and in the present study, viability is dramatically reduced in the presence of H_2O_2, as expected. Only a minor reduction of heterologous expression of peroxygenases was observed for catalase-deficient *E. coli* BL21-Gold(DE3). Thus, with only a minor effect on biomass and protein concentrations, this engineered strain is well suited for biocatalysis with resting cells or cell-free extracts in which the production of the biocatalyst and the conversion phase are clearly separated. In contrast, an application in precursor fermentation with growing cells is rather unfeasible due to the generally low viability of bacteria in presence of H_2O_2, which is further impaired by the catalase deficiency. However, it may be interesting to test to what extent *in situ* generation of H_2O_2 is compatible with cell growth of catalase-deficient *E. coli*. In this context, it should be noted that strictly coupled cascade systems, in which H_2O_2 is first formed in one reaction step, e.g. by an

oxidase, and then used in another reaction step catalyzed by peroxygenases, are not feasible *in vivo*, as a slight decrease in the peroxide concentration over time could still be observed despite the lack of catalases. This is probably due to the activity of the alkyl hydroperoxide reductase in *E. coli* which shows a high affinity towards H_2O_2 and can scavenge even low peroxide concentrations. An additional knockout of this gene, however, leads to drastically impaired cell growth and is therefore not advisable.[194]

Another issue of catalase deficiency arises with the use of hydrophobic organic solvents in biphasic biocatalytic systems. Organic solvents are often used to improve the substrate availability and facilitate product isolation.[195] However, toxicity of these solvents is a major concern. In this context, catalase-deficient *E. coli* strains are even more sensitive to oxidative stress caused by hydrophobic organic solvents such as n-hexane and cyclohexane than wild-type strains.[193] Moreover, the lack of catalases is inherently associated with a higher susceptibility to oxidative DNA damage, resulting in an enhanced mutation rate.[196] It can therefore be assumed that the performance of a cell-based biocatalyst might change over several generations. Avoiding extensive subcultures should, however, prevent the unintentional transmission of mutations. In this context, one advantage of plasmids over chromosomal expression systems should be emphasized: plasmids can be easily reintroduced into the original catalase-deficient strain. In contrast, extensive and stepwise metabolic engineering might fail in catalase-deficient strains due to the accumulation of mutations over time. In this case, the catalase deficiency should be introduced as the last step at the very end of the engineering process. Despite the challenges arising from catalase deficiency, it will be exciting to see to what extent catalase-deficient *E. coli* can be used in H_2O_2-driven biocatalysis in the future, as it was clearly demonstrated that these can be used for biotransformations with UPOs and P450 peroxygenases in whole cells and crude extracts.

Biocatalytic oxyfunctionalization was also the subject of the second peer-reviewed manuscript (Chapter 2.2). The aim of this study was to investigate the effect of chromosomal integration on the catalytic performance of a three-component model P450 system. Using a bacterial

CYP154E1 variant and the redox partner proteins YkuN and FdR, the human metabolite of the anesthetic (*S*)-ketamine, (2*S*,6*S*)-hydroxynorketamine, was successfully produced in resting *E. coli* cells. The expression of P450s in *E. coli* is particularly interesting for studying the effects of chromosomal integration on the performance of biocatalysts. Firstly, *E. coli* does not contain any intrinsic P450s. Thus, heterologous P450s can be quantified very easily by utilizing their special spectral properties. The reduced, carbon monoxide bound form shows a characteristic Soret band at 450 nm wavelength, hence the name P450 (Pigment 450), which differs from other heme proteins such as hemoglobin, which exhibits a Soret band at 420 nm. The essential difference is due to the fifth ligand of the heme iron, which for P450s is a cysteinate instead of a histidine. In the case of denaturation, however, this cysteinate ligand can also be replaced by a histidine, which can be observed by an increase in absorbance at 420 nm.[197-199] Interestingly, chromosomal expression of CYP154E1 without redox partners resulted in a 40% increase in functional heme-loaded enzyme compared to plasmid-based expression in *E. coli* BL21(DE3). However, plasmid-based expression showed a stronger protein band on the SDS-PAGE gel. According to the SDS-PAGE gel analysis, the largest portion of this protein was localized in the insoluble fraction after cell disruption. It can therefore be assumed, that the reduction of the copy number from approximately 40 gene copies for the plasmid system to a single chromosomal gene copy reduced the overall expression as expected, but increased the amount of correctly folded, heme-loaded and therefore active P450. This result clearly shows that the expression of P450s does not necessarily depend on strong transcription, but rather on correct folding and heme loading. This is in line with the literature, in which, for example, a higher enzymatic activity in combination with a lower amount of total protein was achieved by reducing the inducer concentration and the promoter strength for a CYP102A1 (BM3) variant in *E. coli*. Bertelmann et al. concluded that, under strong expression, a considerable amount of P450 likely did not undergo proper folding or heme incorporation.[200] In this context, it is important to note that CYP154E1 is produced at high levels in *E. coli*, even higher than the well characterized CYP102A1 used in the aforementioned study by Bertelmann et al.[201] Although the heme precursors 5-aminolevulinic acid and ferrous sulfate were supplemented in

the study provided in Chapter 2.2, it is reasonable to assume an overstrain of the heme synthesis. Thus, metabolic engineering of heme biosynthesis is an interesting approach to further improve the expression of heme proteins like P450s.[202, 203]

Likewise, even in other organisms such as *Pseudomonas taiwanensis* VLB120, a moderate translation rate was found to be the optimum for the heterologous expression of a P450. An increase in gene copy number only led to an increase in inactive P450 rather than improved hydroxylation activity.[204]

It should be noted that the expression levels varied for different *E. coli* strains with *E. coli* JM109(T7) showing reduced P450 expression after chromosomal integration. The reason for this remains elusive und necessitates further exploration. However, this also highlights the importance of screening different *E. coli* strains for a specific biocatalytic application. Here, the introduction of T7RNAP into different strains using CRISPR/Cas-assisted recombineering was demonstrated to be a simple method to extend the pool of possible candidates capable of strong T7 expression.

Surprisingly, the advantage of chromosomal expression of CYP154E1 was no longer apparent upon introduction of the redox partners. Instead, the plasmid-based co-expression of CYP154E1, YkuN and FdR resulted in significantly higher P450 concentrations. While the P450 concentration of the chromosomal expression system was reduced by an average of 50% compared to the plasmid-based system, the turnover number increased by 80%. This can be explained by the fact that the catalytic performance of multi-component P450 systems not only depends on the P450 concentration, but the ratios of redox partners. [50, 205] Here, the SDS-PAGE analysis revealed an alteration of the protein concentrations with a stronger protein band of the redox partners after chromosomal integration. With a chromosomal expression system at hand, further optimization of the interaction of the redox partners and upscaling of the reaction in terms of volume and substrate concentration could be the next steps on the way to a larger-scale selective synthesis of (2*S*,6*S*)-hydroxynorketamine. A promising

approach might be to increase the relative amount of the flavodoxin since it has been shown that an excess of flavodoxin improves P450-catalyzed reactions *in vitro*.[50, 205]

In the third manuscript (Chapter 2.3), a biosynthetic pathway for the production of pinoresinol was created. This furofuran lignan is the key precursor in the biosynthesis of derived lignans of interest such as sesamin or (-)-podophyllotoxin.[134, 206] The latter is commonly isolated from endangered plants and used for the synthesis of the chemotherapeutics etoposide and teniposide.[207] In previous work, a biosynthetic pathway for the synthesis of (-)-deoxypodophyllotoxin, the direct precursor of (-)-podophyllotoxin, was reconstituted in *E. coli*.[134, 135] However, its starting substrate (+)-pinoresinol is incredibly expensive, rendering this pathway unfeasible from an economic perspective. Pinoresinol is formed via oxidative radical coupling of two coniferyl alcohol monomers which can be catalyzed by laccases or peroxidases.[208] Due to the radical mechanism of the laccase-catalyzed oxidation, pinoresinol is not the main product of this reaction since the product distribution is determined by the unpaired electron distribution.[209] In addition, although coniferyl alcohol is a less expensive substrate compared to pinoresinol, it is not cheap. Thus, to further improve the economic efficiency, ferulic acid was chosen as the starting point which was transformed in four steps to pinoresinol. These steps were catalyzed by heterologous enzymes originating from *Z. mays*, *P. crispum*, and *C. glutamicum* as well as endogenous enzymes of the host. Finally, after comparison of different strains and optimization of the reaction conditions, 100 mg/L pinoresinol were produced from 1 g/L ferulic acid. A major issue of this reaction is the fact, that pinoresinol is a substrate of the laccase as well and is degraded over time.[208] Besides the perfect timing of the reaction endpoint to maximize the yield, a biphasic system should be tested to separate the product from the catalyst. In addition, the enzyme kinetics should be investigated to determine to which extent a higher coniferyl alcohol concentration could lead to a higher pinoresinol concentration. However, in the whole cell system developed in this thesis, it will be difficult to further increase the concentration of the intermediate coniferyl alcohol before inducing oxidative coupling. This is partly due to the reaction equilibrium between coniferyl alcohol and coniferyl aldehyde, and partly because higher ferulic acid

concentrations (2 g/L) led to an inhibition of the entire pathway. This inhibitory effect was also observed in a very recent study which demonstrated metabolic engineering and *de novo* fermentation for the production of coniferyl alcohol in *E. coli*.[210] It can be assumed that the currently used wild-type enzymes *Pc*4CL and *Zm*CCR cannot achieve much higher concentrations of coniferyl alcohol. However, detailed kinetic studies should be performed *in vitro* to determine which compound inhibits which enzyme. To overcome this inhibitory effect, analogous enzymes from other organisms, protein engineering, or carboxylic acid reductases (CAR) may be applied instead of wild-type *Pc*4CL and *Zm*CCR.[211] Finally, it would be interesting to extent the plasmid-free pathway towards the production of derived lignans like sesamin and podophyllotoxin which are not phenolic substrates of the laccase.

3.2 CRISPR/Cas-assisted recombineering in *E. coli*

Since its first description over a decade ago, CRISPR/Cas-assisted recombineering has been applied many times for metabolic engineering in *E. coli*. The vast majority of studies dealt with the optimization of the carbon flux of primary metabolites for *de novo* syntheses of bulk chemicals including proteinogenic and non-proteinogenic amino acids or aliphatic alcohols.[212-214] However, CRISPR/Cas-assisted recombineering appears to be less established in chemistry-oriented application of biocatalysis, especially for the transformation of non-natural substrates, e.g. oxyfunctionalization. Confronted with the problems of plasmid-based expression like cell-to-cell variation in the laboratory, this method appeared promising for various applications for both *in vitro* and *in vivo* biocatalysis. Besides the transfer of plasmid-based to plasmid-free expression systems, which was a major focus of this thesis, the prevention of side reactions and tuning of multi-component systems were also pursued. In the course, various lessons were learned, which led, among other things, to the development of a toolbox for the easy transfer of genes from plasmid libraries into *E. coli*'s chromosome.

The simplest application of CRISPR/Cas-assisted recombineering is the deletion of DNA segments as demonstrated for the catalase-deficient *E. coli* strain. Here, the advantage of this method over classic λ-Red recombineering becomes clear. Without the integration of an antibiotic resistance cassette as for the classic recombineering approach, the method is scar-less and significantly faster. Since plasmids were subsequently used for heterologous expression in the catalase-deficient strain, no antibiotic resistance cassette would have been allowed to remain, which would have increased the workload of classic recombineering. Also, the iterative procedure is shortened by using CRISPR/Cas9 as a counterselection method. However, the patent disputes around CRISPR/Cas9 are still unresolved and the vast and intricate patent landscape as well as license fees stand in the way of widespread industrial application,[215, 216] resulting in classical methods remaining important for genome engineering to this day.

The deletion of DNA segments is predestined for the use of linear donor DNA, as it is usually only used once. The time and effort required for the preparation of circular donor DNA is

disproportionate for these cases.[163] However, for chromosomal integration, the preparation of linear donor DNA quickly encounters limitations. Although short homology arms can be introduced using overhangs by PCR primers,[150] the integration efficiency decreases rapidly with increasing length of the construct to be integrated.[172] This must be countered by extending the homology arms up to hundreds of base pairs which can no longer be introduced as primer overhangs. Instead, long homology arms can either be synthesized or amplified from *E. coli*'s chromosome and subsequently joined with the construct to be integrated via fusion PCR, as described in Chapter 2.2. This increases the cloning effort considerably, and unexpected problems in annealing the three fragments resulted in no successful donor DNA being obtained sometimes. Redesigning homology arms or splitting the setup into two separate fusion PCRs, with two fragments each, further increases the time and effort required. Furthermore, the integration site and thus the design of homology arms affect the expression level of integrated genes. Therefore, the search for and selection of suitable integration loci is important. In this context, it should be noted that other factors such as ribosomal binding sites, promoters or media composition influence cell growth and expression to a much greater extent.[189] However, the optimization of integration loci should not be ignored as it can improve the expression significantly. Ideally, chromosomal integration should have no unintended influence on the host metabolism. For these reasons, intergenic integration was preferred, and a set of reliable high expression loci were identified in this thesis. With a manageable number of ideal integration sites at hand, the repetitive, tedious, and error-prone construction of linear donor DNA seemed avoidable, and the pgRNA-ET and pgRNA-DUET vector series were created. In contrast to other approaches involving circular donor DNA described in the literature,[163, 174] the idea was to implement multiple cloning sites originating from the widely used pET vectors for the cloning of the genes to be integrated. Thus, large enzyme and variant libraries based on pET vectors can now be easily transferred into plasmid-free expression systems, combining the advantages of strong T7 expression and chromosomal expression. Furthermore, the plasmid design ensures that the appropriate type II restriction enzymes are most likely already available in the laboratory where the template pET vector for plasmid-based expression has been

created, and that no problems with intragenic restriction sites occur, as would be the case with a limited number of restriction sites present in the vector.

Even when long homology regions are used, it appears to be not possible to integrate arbitrarily large fragments using CRISPR/Cas-assisted recombineering. The current limit reported in the literature is around 12 kb of donor DNA size.[181, 191] For this reason, longer pathways are usually divided into several fragments to be integrated.[177] The stepwise extension of the lignan cascade described above should therefore be easily possible using the developed pgRNA-ET and pgRNA-DUET toolbox. Moreover, a slightly different approach described in the literature, which involves a combination of positive and negative selection for recombination and the introduction of a bacterial artificial chromosome from which the donor DNA is excised by CRISPR/Cas9, has proven successful for insertions of over 100 kb. The authors suspected issues in the provision of intact linear double-stranded DNA as the reason for the limits observed in λ-Red recombineering.[217] In view of this, it would be interesting to find out where the size limits of the pgRNA-ET and pgRNA-DUET vector series are. Moreover, the extent to which these vector series are suitable for plasmid-based expression and thus could be used as templates for iterative protein engineering, e.g. via QuikChange, should be investigated and could reduce a cloning step if plasmid-free expression is the aim from the start. However, it may be necessary to add a copy of *lacI* to the vector, as the chromosomal copy may be insufficient for transcriptional control considering the high copy number due to the pMB1 origin of replication. It will be interesting to see, how these vector series will be used in the future and what biocatalysts are created with their help. For example, the problem of a high cell-to-cell variability while using plasmids has been observed in the past with recurrent expressions such as redox partner proteins needed for P450-catalyzed reactions *in vitro*. It would therefore be advantageous to establish stable, plasmid-free strains for fermentation for the most frequently used redox partners.

In the course of this thesis, coincidences have also helped. By fortunate selection of expression strains for the three-component P450 system, the effect of the Lon protease on the editing efficiency of different strains became apparent. It is rather the Lon protease deficiency

than the differences between B and K-12 strains that explains the low editing efficiency observed for *E. coli* BL21(DE3). This finding was confirmed almost simultaneously by another group with an elegant experimental design.[218]

The work with CRISPR/Cas9 and the reported problems with the flux of the lignan biosynthesis pathway[135] eventually led to an idea for a method to optimize pathways or multi-component systems after chromosomal integration. However, this method was only tested as a proof-of-concept and has yet to be applied to an enzymatic pathway or multi-component system as well as needs further optimization and validation. Furthermore, the limits of the system remain unknown, but it is assumed that the promoter can be replaced at a maximum of three or four sites simultaneously. Various multiplexing experiments for gene knock-out or knock-in described in the literature never exceeded this limit,[163, 219] as the editing efficiency and number of colony-forming units decrease drastically with the number of sites targeted in parallel. Only the future will show to what extent this number can be increased through clever experimental design.

In conclusion, CRISPR/Cas9-assisted recombineering is a well-suited method for the development of *E. coli*-based biocatalysts and can be applied in various ways. As for almost all molecular biological methods, some experimental details must be considered for a successful experiment. However, the development of robust tools for chromosomal integration based on the experience gained in this work allows their use without knowledge of these details. The tools and strains developed in this thesis are widely applicable and a good starting point e.g. for the easy conversion of plasmid-based to plasmid-free expression systems, the extension and optimization of the lignan biosynthetic pathway, T7 expression in new *E. coli* strains, or H_2O_2-driven whole-cell biocatalysis.

4 References

(1) Singh, R.; Kumar, M.; Mittal, A.; Mehta, P. K. Microbial enzymes: Industrial progress in 21st century. *3 Biotech* **2016**, *6* (2), 174. DOI: 10.1007/s13205-016-0485-8

(2) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. Engineering the third wave of biocatalysis. *Nature* **2012**, *485* (7397), 185-194. DOI: 10.1038/nature11117

(3) Buchholz, K.; Collins, J. The roots—A short history of industrial microbiology and biotechnology. *Appl Microbiol Biotechnol* **2013**, *97* (9), 3747-3762. DOI: 10.1007/s00253-013-4768-2

(4) Sumner, J. B. The chemical nature of enzymes. *J Nutr* **1933**, *6* (1), 103-112. DOI: 10.1093/jn/6.1.103

(5) Heckmann, C. M.; Paradisi, F. Looking back: A short history of the discovery of enzymes and how they became powerful chemical tools. *ChemCatChem* **2020**, *12* (24), 6082-6102. DOI: 10.1002/cctc.202001107

(6) Raper, K. B.; Alexander, D. F.; Coghill, R. D. Penicillin. *J Bacteriol* **1944**, *48* (6), 639-659. DOI: 10.1128/jb.48.6.639-659.1944

(7) Singh Dhillon, G.; Kaur Brar, S.; Verma, M.; Tyagi, R. D. Recent advances in citric acid bio-production and recovery. *Food Bioprocess Technol* **2010**, *4* (4), 505-529. DOI: 10.1007/s11947-010-0399-0

(8) Wang, Y.; Janssen, H.; Blaschek, H. P. Fermentative biobutanol production: An old topic with remarkable recent advances. In *Bioprocessing of Renewable Resources to Commodity Bioproducts*, 2014; pp 227-260.

(9) Jensen, V. J.; Rugh, S. Industrial-scale production and application of immobilized glucose isomerase. In *Methods in Enzymology*, Vol. 136; Academic Press, 1987; pp 356-370.

(10) Wu, S.; Snajdrova, R.; Moore, J. C.; Baldenius, K.; Bornscheuer, U. T. Biocatalysis: Enzymatic synthesis for industrial applications. *Angew Chem Int Ed Engl* **2021**, *60* (1), 88-119. DOI: 10.1002/anie.202006648

(11) Jackson, D. A.; Symons, R. H.; Berg, P. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proc Natl Acad Sci U S A* **1972**, *69* (10), 2904-2909. DOI: 10.1073/pnas.69.10.2904

(12) Cohen, S. N.; Chang, A. C. Y.; Boyer, H. W.; Helling, R. B. Construction of biologically functional bacterial plasmids *in vitro*. *Proc Natl Acad Sci* **1973**, *70* (11), 3240-3244. DOI: 10.1073/pnas.70.11.3240

(13) Pham, J. V.; Yilma, M. A.; Feliz, A.; Majid, M. T.; Maffetone, N.; Walker, J. R.; Kim, E.; Cho, H. J.; Reynolds, J. M.; Song, M. C.; et al. A review of the microbial production of bioactive natural products and biologics. *Front Microbiol* **2019**, *10*, 1404. DOI: 10.3389/fmicb.2019.01404

(14) Atanasov, A. G.; Waltenberger, B.; Pferschy-Wenzig, E. M.; Linder, T.; Wawrosch, C.; Uhrin, P.; Temml, V.; Wang, L.; Schwaiger, S.; Heiss, E. H.; et al. Discovery and resupply of pharmacologically active plant-derived natural products: A review. *Biotechnol Adv* **2015**, *33* (8), 1582-1614. DOI: 10.1016/j.biotechadv.2015.08.001

(15) Sheldon, R. A.; Brady, D. The limits to biocatalysis: pushing the envelope. *Chem Commun (Camb)* **2018**, *54* (48), 6088-6104. DOI: 10.1039/c8cc02463d

(16) Schoemaker, H. E.; Mink, D.; Wubbolts, M. G. Dispelling the myths—biocatalysis in industrial synthesis. *Science* **2003**, *299* (5613), 1694-1697. DOI: 10.1126/science.1079237

(17) Hanefeld, U.; Hollmann, F.; Paul, C. E. Biocatalysis making waves in organic chemistry. *Chem Soc Rev* **2022**, *51* (2), 594-627. DOI: 10.1039/d1cs00100k

(18) Bottcher, D.; Bornscheuer, U. T. Protein engineering of microbial enzymes. *Curr Opin Microbiol* **2010**, *13* (3), 274-282. DOI: 10.1016/j.mib.2010.01.010

(19) Lutz, S.; Iamurri, S. M. Protein engineering: Past, present, and future. In *Protein engineering: Methods and protocols*, Bornscheuer, U. T., Höhne, M. Eds.; Springer New York, 2018; pp 1-12.

(20) Kortemme, T. *De novo* protein design - From new structures to programmable functions. *Cell* **2024**, *187* (3), 526-544. DOI: 10.1016/j.cell.2023.12.028

(21) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583-589. DOI: 10.1038/s41586-021-03819-2

(22) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. Design of a novel globular protein fold with atomic-level accuracy. *Science* **2003**, *302* (5649), 1364-1368. DOI: 10.1126/science.1089427

(23) *Press release*. https://www.nobelprize.org/prizes/chemistry/2024/press-release/ (accessed 2025-02-03).

(24) Yang, Z.; Zeng, X.; Zhao, Y.; Chen, R. AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduct Target Ther* **2023**, *8* (1), 115. DOI: 10.1038/s41392-023-01381-z

(25) Alcantara, A. R.; Dominguez de Maria, P.; Littlechild, J. A.; Schurmann, M.; Sheldon, R. A.; Wohlgemuth, R. Biocatalysis as key to sustainable industrial chemistry. *ChemSusChem* **2022**, *15* (9), e202102709. DOI: 10.1002/cssc.202102709

(26) Sheldon, R. A.; Woodley, J. M. Role of biocatalysis in sustainable chemistry. *Chem Rev* **2018**, *118* (2), 801-838. DOI: 10.1021/acs.chemrev.7b00203

(27) Bommarius, A. S.; Riebel, B. R. Introduction to biocatalysis. In *Biocatalysis*, Bommarius, A. S., Riebel, B. R. Eds.; 2004; pp 1-18.

(28) Tufvesson, P.; Lima-Ramos, J.; Nordblad, M.; Woodley, J. M. Guidelines and cost analysis for catalyst production in biocatalytic processes. *Org Process Res Dev* **2011**, *15* (1), 266-274. DOI: 10.1021/op1002165

(29) Ajikumar, P. K.; Xiao, W. H.; Tyo, K. E.; Wang, Y.; Simeon, F.; Leonard, E.; Mucha, O.; Phon, T. H.; Pfeifer, B.; Stephanopoulos, G. Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* **2010**, *330* (6000), 70-74. DOI: 10.1126/science.1191652

(30) Roman-Camacho, J. J.; Garcia-Garcia, I.; Santos-Duenas, I. M.; Garcia-Martinez, T.; Mauricio, J. C. Latest trends in industrial vinegar production and the role of acetic acid bacteria:

Classification, metabolism, and applications—A comprehensive review. *Foods* **2023**, *12* (19), 3705. DOI: 10.3390/foods12193705

(31) Lin, B.; Tao, Y. Whole-cell biocatalysts by design. *Microb Cell Fact* **2017**, *16* (1), 106. DOI: 10.1186/s12934-017-0724-7

(32) de Carvalho, C. C. Whole cell biocatalysts: essential workers from nature to the industry. *Microb Biotechnol* **2017**, *10* (2), 250-263. DOI: 10.1111/1751-7915.12363

(33) Zhang, C.; Sultan, S. A.; T, R.; Chen, X. Biotechnological applications of S-adenosyl-methionine-dependent methyltransferases for natural products biosynthesis and diversification. *Bioresour Bioprocess* **2021**, *8* (1), 72. DOI: 10.1186/s40643-021-00425-y

(34) France, S. P.; Hepworth, L. J.; Turner, N. J.; Flitsch, S. L. Constructing biocatalytic cascades: *In vitro* and in *vivo* approaches to *de novo* multi-enzyme pathways. *ACS Catal* **2016**, *7* (1), 710-724. DOI: 10.1021/acscatal.6b02979

(35) de Carvalho, C. C. Enzymatic and whole cell catalysis: Finding new strategies for old processes. *Biotechnol Adv* **2011**, *29* (1), 75-83. DOI: 10.1016/j.biotechadv.2010.09.001

(36) Siedentop, R.; Claaßen, C.; Rother, D.; Lütz, S.; Rosenthal, K. Getting the most out of enzyme cascades: Strategies to optimize *in vitro* multi-enzymatic reactions. *Catalysts* **2021**, *11* (10). DOI: 10.3390/catal11101183

(37) Dong, J.; Fernandez-Fueyo, E.; Hollmann, F.; Paul, C. E.; Pesic, M.; Schmidt, S.; Wang, Y.; Younes, S.; Zhang, W. Biocatalytic oxidation reactions: A chemist's perspective. *Angew Chem Int Ed Engl* **2018**, *57* (30), 9238-9261. DOI: 10.1002/anie.201800343

(38) Sharma, V. K.; Hutchison, J. M.; Allgeier, A. M. Redox biocatalysis: Quantitative comparisons of nicotinamide cofactor regeneration methods. *ChemSusChem* **2022**, *15* (22), e202200888. DOI: 10.1002/cssc.202200888

(39) Xu, X.; Hilberath, T.; Hollmann, F. Selective oxyfunctionalisation reactions catalysed by P450 monooxygenases and peroxygenases – A bright future for sustainable chemical synthesis. *Curr Opin Green Sustainable Chem* **2023**, *39*. DOI: 10.1016/j.cogsc.2022.100745

(40) Fessner, N. D. P450 monooxygenases enable rapid late-stage diversification of natural products *via* C-H bond activation. *ChemCatChem* **2019**, *11* (9), 2226-2242. DOI: 10.1002/cctc.201801829

(41) Zhang, R. K.; Huang, X.; Arnold, F. H. Selective CH bond functionalization with engineered heme proteins: New tools to generate complexity. *Curr Opin Chem Biol* **2019**, *49*, 67-75. DOI: 10.1016/j.cbpa.2018.10.004

(42) Chakrabarty, S.; Wang, Y.; Perkins, J. C.; Narayan, A. R. H. Scalable biocatalytic C-H oxyfunctionalization reactions. *Chem Soc Rev* **2020**, *49* (22), 8137-8155. DOI: 10.1039/d0cs00440e

(43) Bernhardt, R.; Urlacher, V. B. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. *Appl Microbiol Biotechnol* **2014**, *98* (14), 6185-6203. DOI: 10.1007/s00253-014-5767-7

(44) Nelson, D. R. Cytochrome P450 diversity in the tree of life. *Biochim Biophys Acta Proteins Proteom* **2018**, *1866* (1), 141-154. DOI: 10.1016/j.bbapap.2017.05.003

(45) Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H. Heme-containing oxygenases. *Chem Rev* **1996**, *96* (7), 2841-2888. DOI: 10.1021/cr9500500

(46) Bernhardt, R. Cytochromes P450 as versatile biocatalysts. *J Biotechnol* **2006**, *124* (1), 128-145. DOI: 10.1016/j.jbiotec.2006.01.026

(47) Urlacher, V. B.; Girhard, M. Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends Biotechnol* **2012**, *30* (1), 26-36. DOI: 10.1016/j.tibtech.2011.06.012

(48) Hannemann, F.; Bichet, A.; Ewen, K. M.; Bernhardt, R. Cytochrome P450 systems - biological variations of electron transport chains. *Biochim Biophys Acta* **2007**, *1770* (3), 330-344. DOI: 10.1016/j.bbagen.2006.07.017

(49) Urlacher, V. B. Catalysis with cytochrome P450 monooxygenases. In *Handbook of Green Chemistry*, Anastas, P. T. Ed.; 2010; pp 1-25.

(50) Girhard, M.; Klaus, T.; Khatri, Y.; Bernhardt, R.; Urlacher, V. B. Characterization of the versatile monooxygenase CYP109B1 from *Bacillus subtilis*. *Appl Microbiol Biotechnol* **2010**, *87* (2), 595-607. DOI: 10.1007/s00253-010-2472-z

(51) Petzoldt, K.; Annen, K.; Laurent, H.; Wiechert, R. Process for the preparation of 11 beta-hydroxy steroids. 1982.

(52) Park, J. W.; Lee, J. K.; Kwon, T. J.; Yi, D. H.; Kim, Y. J.; Moon, S. H.; Suh, H. H.; Kang, S. M.; Park, Y. I. Bioconversion of compactin into pravastatin by *Streptomyces* sp. *Biotechnol Lett* **2003**, *25* (21), 1827-1831. DOI: 10.1023/A:1026281914301

(53) Ro, D. K.; Paradise, E. M.; Ouellet, M.; Fisher, K. J.; Newman, K. L.; Ndungu, J. M.; Ho, K. A.; Eachus, R. A.; Ham, T. S.; Kirby, J.; et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **2006**, *440* (7086), 940-943. DOI: 10.1038/nature04640

(54) Schroer, K.; Kittelmann, M.; Lutz, S. Recombinant human cytochrome P450 monooxygenases for drug metabolite synthesis. *Biotechnol Bioeng* **2010**, *106* (5), 699-706. DOI: 10.1002/bit.22775

(55) Sawayama, A. M.; Chen, M. M.; Kulanthaivel, P.; Kuo, M. S.; Hemmerle, H.; Arnold, F. H. A panel of cytochrome P450 BM3 variants to produce drug metabolites and diversify lead compounds. *Chemistry* **2009**, *15* (43), 11723-11729. DOI: 10.1002/chem.200900643

(56) Bokel, A.; Hutter, M. C.; Urlacher, V. B. Molecular evolution of a cytochrome P450 for the synthesis of potential antidepressant (2*R*,6*R*)-hydroxynorketamine. *Chem Commun (Camb)* **2021**, *57* (4), 520-523, 10.1039/D0CC06729F. DOI: 10.1039/d0cc06729f

(57) Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I. Structure and chemistry of cytochrome P450. *Chem Rev* **2005**, *105* (6), 2253-2277. DOI: 10.1021/cr0307143

(58) Huang, X.; Groves, J. T. Oxygen activation and radical transformations in heme proteins and metalloporphyrins. *Chem Rev* **2018**, *118* (5), 2491-2553. DOI: 10.1021/acs.chemrev.7b00373

(59) Groves, J. T.; Mcclusky, G. A. Aliphatic hydroxylation *via* oxygen rebound - oxygen-transfer catalyzed by iron. *J Am Chem Soc* **1976**, *98* (3), 859-861. DOI: 10.1021/ja00419a049

(60) Werck-Reichhart, D.; Feyereisen, R. Cytochromes P450: A success story. *Genome Biol* **2000**, *1* (6), REVIEWS3003. DOI: 10.1186/gb-2000-1-6-reviews3003

(61) Montellano, P. R. O. d. *Cytochrome P450: Structure, mechanism, and biochemistry*; Kluwer Academic/Plenum Publishers, 2005. DOI: 10.1007/b139087.

(62) Girhard, M.; Schuster, S.; Dietrich, M.; Durre, P.; Urlacher, V. B. Cytochrome P450 monooxygenase from *Clostridium acetobutylicum*: a new alpha-fatty acid hydroxylase. *Biochem Biophys Res Commun* **2007**, *362* (1), 114-119. DOI: 10.1016/j.bbrc.2007.07.155

(63) Shoji, O.; Watanabe, Y. Peroxygenase reactions catalyzed by cytochromes P450. *J Biol Inorg Chem* **2014**, *19* (4-5), 529-539. DOI: 10.1007/s00775-014-1106-9

(64) Matsunaga, I.; Yokotani, N.; Gotoh, O.; Kusunose, E.; Yamada, M.; Ichihara, K. Molecular cloning and expression of fatty acid alpha-hydroxylase from *Sphingomonas paucimobilis*. *J Biol Chem* **1997**, *272* (38), 23592-23596. DOI: 10.1074/jbc.272.38.23592

(65) Belcher, J.; McLean, K. J.; Matthews, S.; Woodward, L. S.; Fisher, K.; Rigby, S. E. J.; Nelson, D. R.; Potts, D.; Baynham, M. T.; Parker, D. A.; et al. Structure and biochemical properties of the alkene producing cytochrome P450 $OleT_{JE}$ (CYP152L1) from the *Jeotgalicoccus* sp. 8456 bacterium. *J Biol Chem* **2014**, *289* (10), 6535-6550. DOI: 10.1074/jbc.M113.527325

(66) Joo, H.; Lin, Z.; Arnold, F. H. Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation. *Nature* **1999**, *399* (6737), 670-673. DOI: 10.1038/21395

(67) Ebrecht, A. C.; Smit, M. S.; Opperman, D. J. Natural alternative heme-environments allow efficient peroxygenase activity by cytochrome P450 monooxygenases. *Catalysis Science & Technology* **2023**, *13* (21), 6264-6273. DOI: 10.1039/d3cy01207g

(68) Zhao, P.; Kong, F.; Jiang, Y.; Qin, X.; Tian, X.; Cong, Z. Enabling peroxygenase activity in cytochrome P450 monooxygenases by engineering hydrogen peroxide tunnels. *J Am Chem Soc* **2023**, *145* (9), 5506-5511. DOI: 10.1021/jacs.3c00195

(69) Podgorski, M. N.; Akter, J.; Churchman, L. R.; Bruning, J. B.; De Voss, J. J.; Bell, S. G. Engineering peroxygenase activity into cytochrome P450 monooxygenases through modification of the oxygen binding region. *ACS Catal* **2024**, *14* (10), 7426-7443. DOI: 10.1021/acscatal.4c01326

(70) Girhard, M.; Kunigk, E.; Tihovsky, S.; Shumyantseva, V. V.; Urlacher, V. B. Light-driven biocatalysis with cytochrome P450 peroxygenases. *Biotechnol Appl Biochem* **2013**, *60* (1), 111-118. DOI: 10.1002/bab.1063

(71) Albertolle, M. E.; Peter Guengerich, F. The relationships between cytochromes P450 and H_2O_2: Production, reaction, and inhibition. *J Inorg Biochem* **2018**, *186*, 228-234. DOI: 10.1016/j.jinorgbio.2018.05.014

(72) Hofrichter, M.; Ullrich, R. Oxidations catalyzed by fungal peroxygenases. *Curr Opin Chem Biol* **2014**, *19*, 116-125. DOI: 10.1016/j.cbpa.2014.01.015

(73) Hobisch, M.; Holtmann, D.; Gomez de Santos, P.; Alcalde, M.; Hollmann, F.; Kara, S. Recent developments in the use of peroxygenases - Exploring their high potential in selective oxyfunctionalisations. *Biotechnol Adv* **2021**, *51*, 107615. DOI: 10.1016/j.biotechadv.2020.107615

(74) Burek, B. O.; Bormann, S.; Hollmann, F.; Bloh, J. Z.; Holtmann, D. Hydrogen peroxide driven biocatalysis. *Green Chem* **2019**, *21* (12), 3232-3249. DOI: 10.1039/c9gc00633h

(75) Knorrscheidt, A.; Soler, J.; Hunecke, N.; Pullmann, P.; Garcia-Borras, M.; Weissenborn, M. J. Accessing chemo- and regioselective benzylic and aromatic oxidations by protein

engineering of an unspecific peroxygenase. *ACS Catal* **2021**, *11* (12), 7327-7338. DOI: 10.1021/acscatal.1c00847

(76) Püllmann, P.; Weissenborn, M. J. Improving the heterologous production of fungal peroxygenases through an episomal *Pichia pastoris* promoter and signal peptide shuffling system. *ACS Synth Biol* **2021**, *10* (6), 1360-1372. DOI: 10.1021/acssynbio.0c00641

(77) Teetz, N.; Lang, S.; Liese, A.; Holtmann, D. Yeast surface display enables one-step production and immobilization of unspecific peroxygenases. *ChemCatChem* **2024**, *16* (21). DOI: 10.1002/cctc.202400908

(78) Zhao, L. X.; Zou, S. P.; Shen, Q.; Xue, Y. P.; Zheng, Y. G. Enhancing the expression of the unspecific peroxygenase in *Komagataella phaffii* through a combination strategy. *Appl Microbiol Biotechnol* **2024**, *108* (1), 320. DOI: 10.1007/s00253-024-13166-7

(79) Molina-Espeja, P.; Garcia-Ruiz, E.; Gonzalez-Perez, D.; Ullrich, R.; Hofrichter, M.; Alcalde, M. Directed evolution of unspecific peroxygenase from *Agrocybe aegerita*. *Appl Environ Microbiol* **2014**, *80* (11), 3496-3507. DOI: 10.1128/AEM.00490-14

(80) Yan, X.; Zhang, X.; Li, H.; Deng, D.; Guo, Z.; Kang, L.; Li, A. Engineering of unspecific peroxygenases using a superfolder-green-fluorescent-protein-mediated secretion system in *Escherichia coli*. *JACS Au* **2024**, *4* (4), 1654-1663. DOI: 10.1021/jacsau.4c00129

(81) Carro, J.; González-Benjumea, A.; Fernández-Fueyo, E.; Aranda, C.; Guallar, V.; Gutiérrez, A.; Martínez, A. T. Modulating fatty acid epoxidation vs hydroxylation in a fungal peroxygenase. *ACS Catal* **2019**, *9* (7), 6234-6242. DOI: 10.1021/acscatal.9b01454

(82) Linde, D.; Olmedo, A.; González-Benjumea, A.; Estévez, M.; Renau-Mínguez, C.; Carro, J.; Fernández-Fueyo, E.; Gutiérrez, A.; Martínez, A. T. Two new unspecific peroxygenases from heterologous expression of fungal genes in *Escherichia coli*. *Appl Environ Microbiol* **2020**, *86* (7), e02899-02819. DOI: 10.1128/AEM.02899-19

(83) Robinson, W. X. Q.; Mielke, T.; Melling, B.; Cuetos, A.; Parkin, A.; Unsworth, W. P.; Cartwright, J.; Grogan, G. Comparing the catalytic and structural characteristics of a ‚short' unspecific peroxygenase (UPO) expressed in *Pichia pastoris* and *Escherichia coli*. *Chembiochem* **2023**, *24* (1), e202200558. DOI: 10.1002/cbic.202200558

(84) Romero, E. O.; Saucedo, A. T.; Hernandez-Melendez, J. R.; Yang, D.; Chakrabarty, S.; Narayan, A. R. H. Enabling broader adoption of biocatalysis in organic chemistry. *JACS Au* **2023**, *3* (8), 2073-2085. DOI: 10.1021/jacsau.3c00263

(85) Zhao, M.; Ma, J.; Zhang, L.; Qi, H. Engineering strategies for enhanced heterologous protein production by *Saccharomyces cerevisiae*. *Microb Cell Fact* **2024**, *23* (1), 32. DOI: 10.1186/s12934-024-02299-z

(86) Juturu, V.; Wu, J. C. Heterologous protein expression in *Pichia pastoris*: Latest research progress and applications. *Chembiochem* **2018**, *19* (1), 7-21. DOI: 10.1002/cbic.201700460

(87) Degering, C.; Eggert, T.; Puls, M.; Bongaerts, J.; Evers, S.; Maurer, K. H.; Jaeger, K. E. Optimization of protease secretion in *Bacillus subtilis* and *Bacillus licheniformis* by screening of homologous and heterologous signal peptides. *Appl Environ Microbiol* **2010**, *76* (19), 6370-6376. DOI: 10.1128/AEM.01146-10

(88) Tungekar, A. A.; Castillo-Corujo, A.; Ruddock, L. W. So you want to express your protein in *Escherichia coli*? *Essays Biochem* **2021**, *65* (2), 247-260. DOI: 10.1042/EBC20200170

(89) Zajkoska, P.; Rebros, M.; Rosenberg, M. Biocatalysis with immobilized *Escherichia coli*. *Appl Microbiol Biotechnol* **2013**, *97* (4), 1441-1455. DOI: 10.1007/s00253-012-4651-6

(90) Schulze, B.; Wubbolts, M. G. Biocatalysis for industrial production of fine chemicals. *Curr Opin Biotechnol* **1999**, *10* (6), 609-615. DOI: 10.1016/S0958-1669(99)00042-7

(91) Ruiz, N.; Silhavy, T. J. How *Escherichia coli* became the flagship bacterium of molecular biology. *J Bacteriol* **2022**, *204* (9), e0023022. DOI: 10.1128/jb.00230-22

(92) Chang, A. C. Y.; Cohen, S. N. Genome construction between bacterial species *in vitro*: Replication and expression of *Staphylococcus* plasmid genes in *Escherichia coli*. *Proc Natl Acad Sci* **1974**, *71* (4), 1030-1034. DOI: 10.1073/pnas.71.4.1030

(93) Hutchison, C. A.; Phillips, S.; Edgell, M. H.; Gillam, S.; Jahnke, P.; Smith, M. Mutagenesis at a specific position in a DNA sequence. *J Biol Chem* **1978**, *253* (18), 6551-6560. DOI: 10.1016/S0021-9258(19)46967-6

(94) Saiki, R. K.; Scharf, S.; Faloona, F.; Mullis, K. B.; Horn, G. T.; Erlich, H. A.; Arnheim, N. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **1985**, *230* (4732), 1350-1354. DOI: 10.1126/science.2999980

(95) Blattner, F. R.; Plunkett, G.; Bloch, C. A.; Perna, N. T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J. D.; Rode, C. K.; Mayhew, G. F.; et al. The complete genome sequence of *Escherichia coli* K-12. *Science* **1997**, *277* (5331), 1453-1462. DOI: 10.1126/science.277.5331.1453

(96) Huang, C. J.; Lin, H.; Yang, X. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J Ind Microbiol Biotechnol* **2012**, *39* (3), 383-399. DOI: 10.1007/s10295-011-1082-9

(97) Theisen, M.; Liao, J. C. Industrial biotechnology: *Escherichia coli* as a host. In *Industrial Biotechnology*, 2016; pp 149-181.

(98) Monk, J. M.; Koza, A.; Campodonico, M. A.; Machado, D.; Seoane, J. M.; Palsson, B. O.; Herrgard, M. J.; Feist, A. M. Multi-omics quantification of species variation of *Escherichia coli* links molecular features with strain phenotypes. *Cell Syst* **2016**, *3* (3), 238-251 e212. DOI: 10.1016/j.cels.2016.08.013

(99) Kaper, J. B.; Nataro, J. P.; Mobley, H. L. Pathogenic *Escherichia coli*. *Nat Rev Microbiol* **2004**, *2* (2), 123-140. DOI: 10.1038/nrmicro818

(100) Anderson, E. S. Viability of, and transfer of a plasmid from, *E. coli* K12 in the human intestine. *Nature* **1975**, *255* (5508), 502-504. DOI: 10.1038/255502a0

(101) Castonguay, M. H.; van der Schaaf, S.; Koester, W.; Krooneman, J.; van der Meer, W.; Harmsen, H.; Landini, P. Biofilm formation by *Escherichia coli* is stimulated by synergistic interactions and co-adhesion mechanisms with adherence-proficient bacteria. *Res Microbiol* **2006**, *157* (5), 471-478. DOI: 10.1016/j.resmic.2005.10.003

(102) Studier, F. W.; Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **1986**, *189* (1), 113-130. DOI: 10.1016/0022-2836(86)90385-2

(103) Miroux, B.; Walker, J. E. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* **1996**, *260* (3), 289-298. DOI: 10.1006/jmbi.1996.0399

(104) Kwon, S. K.; Kim, S. K.; Lee, D. H.; Kim, J. F. Comparative genomics and experimental evolution of *Escherichia coli* BL21(DE3) strains reveal the landscape of toxicity escape from membrane protein overproduction. *Sci Rep* **2015**, *5*, 16076. DOI: 10.1038/srep16076

(105) Tegel, H.; Tourle, S.; Ottosson, J.; Persson, A. Increased levels of recombinant human proteins with the *Escherichia coli* strain Rosetta(DE3). *Protein Expr Purif* **2010**, *69* (2), 159-167. DOI: 10.1016/j.pep.2009.08.017

(106) Kang, D. G.; Kim, Y. K.; Cha, H. J. Comparison of green fluorescent protein expression in two industrial *Escherichia coli* strains, BL21 and W3110, under co-expression of bacterial hemoglobin. *Appl Microbiol Biotechnol* **2002**, *59* (4-5), 523-528. DOI: 10.1007/s00253-002-1043-3

(107) Marisch, K.; Bayer, K.; Cserjan-Puschmann, M.; Luchner, M.; Striedner, G. Evaluation of three industrial *Escherichia coli* strains in fed-batch cultivations during high-level SOD protein production. *Microb Cell Fact* **2013**, *12* (1), 58. DOI: 10.1186/1475-2859-12-58

(108) Mori, H.; Kataoka, M.; Yang, X. Past, present, and future of genome modification in *Escherichia coli*. *Microorganisms* **2022**, *10* (9). DOI: 10.3390/microorganisms10091835

(109) Lin-Chao, S.; Chen, W.-T.; Wong, T.-T. High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. *Mol Microbiol* **1992**, *6* (22), 3385-3393. DOI: 10.1111/j.1365-2958.1992.tb02206.x

(110) Merck KGaA. *pET system manual 11th edition*; Merck KGaA, 2011.

(111) Nakamura, T.; Koma, D.; Oshima, M.; Hoshino, H.; Ohmoto, T.; Uegaki, K. Application of chromosomal gene insertion into *Escherichia coli* for expression of recombinant proteins. *J Biosci Bioeng* **2018**, *126* (2), 266-272. DOI: 10.1016/j.jbiosc.2018.02.016

(112) Koma, D.; Yamanaka, H.; Moriyoshi, K.; Ohmoto, T.; Sakai, K. A convenient method for multiple insertions of desired genes into target loci on the *Escherichia coli* chromosome. *Appl Microbiol Biotechnol* **2012**, *93* (2), 815-829. DOI: 10.1007/s00253-011-3735-z

(113) Jones, K. L.; Kim, S. W.; Keasling, J. D. Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. *Metab Eng* **2000**, *2* (4), 328-338. DOI: 10.1006/mben.2000.0161

(114) Striedner, G.; Pfaffenzeller, I.; Markus, L.; Nemecek, S.; Grabherr, R.; Bayer, K. Plasmid-free T7-based *Escherichia coli* expression systems. *Biotechnol Bioeng* **2010**, *105* (4), 786-794. DOI: 10.1002/bit.22598

(115) Sudhamsu, J.; Kabir, M.; Airola, M. V.; Patel, B. A.; Yeh, S. R.; Rousseau, D. L.; Crane, B. R. Co-expression of ferrochelatase allows for complete heme incorporation into recombinant proteins produced in *E. coli*. *Protein Expr Purif* **2010**, *73* (1), 78-82. DOI: 10.1016/j.pep.2010.03.010

(116) Pasini, M.; Fernandez-Castane, A.; Jaramillo, A.; de Mas, C.; Caminal, G.; Ferrer, P. Using promoter libraries to reduce metabolic burden due to plasmid-encoded proteins in recombinant *Escherichia coli*. *N Biotechnol* **2016**, *33* (1), 78-90. DOI: 10.1016/j.nbt.2015.08.003

(117) Liang, L.; Liu, R.; Garst, A. D.; Lee, T.; Nogue, V. S. I.; Beckham, G. T.; Gill, R. T. CRISPR EnAbled Trackable genome Engineering for isopropanol production in *Escherichia coli*. *Metab Eng* **2017**, *41*, 1-10. DOI: 10.1016/j.ymben.2017.02.009

(118) Saleski, T. E.; Chung, M. T.; Carruthers, D. N.; Khasbaatar, A.; Kurabayashi, K.; Lin, X. N. Optimized gene expression from bacterial chromosome by high-throughput integration and screening. *Sci Adv* **2021**, *7* (7), eabe1767. DOI: 10.1126/sciadv.abe1767

(119) Mairhofer, J.; Scharl, T.; Marisch, K.; Cserjan-Puschmann, M.; Striedner, G. Comparative transcription profiling and in-depth characterization of plasmid-based and plasmid-free *Escherichia coli* expression systems under production conditions. *Appl Environ Microbiol* **2013**, *79* (12), 3802-3812. DOI: 10.1128/AEM.00365-13

(120) Englaender, J. A.; Jones, J. A.; Cress, B. F.; Kuhlman, T. E.; Linhardt, R. J.; Koffas, M. A. G. Effect of genomic integration location on heterologous protein expression and metabolic engineering in *E. coli*. *ACS Synth Biol* **2017**, *6* (4), 710-720. DOI: 10.1021/acssynbio.6b00350

(121) Wang, J.; Niyompanich, S.; Tai, Y. S.; Wang, J.; Bai, W.; Mahida, P.; Gao, T.; Zhang, K. Engineering of a highly efficient *Escherichia coli* strain for mevalonate fermentation through chromosomal integration. *Appl Environ Microbiol* **2016**, *82* (24), 7176-7184. DOI: 10.1128/AEM.02178-16

(122) Luelf, U. J.; Böhmer, L. M.; Li, S.; Urlacher, V. B. Effect of chromosomal integration on catalytic performance of a multi-component P450 system in *Escherichia coli*. *Biotechnol Bioeng* **2023**, *120* (7), 1762-1772. DOI: 10.1002/bit.28404

(123) Lopez-Gallego, F.; Schmidt-Dannert, C. Multi-enzymatic synthesis. *Curr Opin Chem Biol* **2010**, *14* (2), 174-183. DOI: 10.1016/j.cbpa.2009.11.023

(124) Schrittwieser, J. H.; Velikogne, S.; Hall, M.; Kroutil, W. Artificial biocatalytic linear cascades for preparation of organic molecules. *Chem Rev* **2018**, *118* (1), 270-348. DOI: 10.1021/acs.chemrev.7b00033

(125) Ricca, E.; Brucher, B.; Schrittwieser, J. H. Multi-enzymatic cascade reactions: Overview and perspectives. *Adv Synth Catal* **2011**, *353* (13), 2239-2262. DOI: 10.1002/adsc.201100256

(126) Huffman, M. A.; Fryszkowska, A.; Alvizo, O.; Borra-Garske, M.; Campos, K. R.; Canada, K. A.; Devine, P. N.; Duan, D.; Forstater, J. H.; Grosser, S. T.; et al. Design of an *in vitro* biocatalytic cascade for the manufacture of islatravir. *Science* **2019**, *366* (6470), 1255-1259. DOI: 10.1126/science.aay8484

(127) Jones, J. A.; Vernacchio, V. R.; Lachance, D. M.; Lebovich, M.; Fu, L.; Shirke, A. N.; Schultz, V. L.; Cress, B.; Linhardt, R. J.; Koffas, M. A. ePathOptimize: A combinatorial approach for transcriptional balancing of metabolic pathways. *Sci Rep* **2015**, *5*, 11301. DOI: 10.1038/srep11301

(128) Luo, Y.; Li, B. Z.; Liu, D.; Zhang, L.; Chen, Y.; Jia, B.; Zeng, B. X.; Zhao, H.; Yuan, Y. J. Engineered biosynthesis of natural products in heterologous hosts. *Chem Soc Rev* **2015**, *44* (15), 5265-5290. DOI: 10.1039/c5cs00025d

(129) Luo, S.; Diehl, C.; He, H.; Bae, Y.; Klose, M.; Claus, P.; Cortina, N. S.; Fernandez, C. A.; Schulz-Mirbach, H.; McLean, R.; et al. Construction and modular implementation of the THETA cycle for synthetic CO2 fixation. *Nat Catal* **2023**, *6* (12), 1228-1240. DOI: 10.1038/s41929-023-01079-z

(130) Milder, I. E.; Arts, I. C.; van de Putte, B.; Venema, D. P.; Hollman, P. C. Lignan contents of Dutch plant foods: a database including lariciresinol, pinoresinol, secoisolariciresinol and matairesinol. *Br. J. Nutr.* **2005**, *93* (3), 393-402. DOI: 10.1079/bjn20051371

(131) Zalesak, F.; Bon, D. J. D.; Pospisil, J. Lignans and neolignans: Plant secondary metabolites as a reservoir of biologically active substances. *Pharmacol Res* **2019**, *146*, 104284. DOI: 10.1016/j.phrs.2019.104284

(132) Farkya, S.; Bisaria, V. S.; Srivastava, A. K. Biotechnological aspects of the production of the anticancer drug podophyllotoxin. *Appl Microbiol Biotechnol* **2004**, *65* (5), 504-519. DOI: 10.1007/s00253-004-1680-9

(133) Schultz, B. J.; Kim, S. Y.; Lau, W.; Sattely, E. S. Total biosynthesis for milligram-scale production of etoposide intermediates in a plant chassis. *J Am Chem Soc* **2019**, *141* (49), 19231-19235. DOI: 10.1021/jacs.9b10717

(134) Decembrino, D.; Raffaele, A.; Knöfel, R.; Girhard, M.; Urlacher, V. B. Synthesis of (-)-deoxypodophyllotoxin and (-)-epipodophyllotoxin via a multi-enzyme cascade in *E. coli*. *Microb Cell Fact* **2021**, *20* (1), 183. DOI: 10.1186/s12934-021-01673-5

(135) Decembrino, D.; Ricklefs, E.; Wohlgemuth, S.; Girhard, M.; Schullehner, K.; Jach, G.; Urlacher, V. B. Assembly of plant enzymes in *E. coli* for the production of the valuable (-)-podophyllotoxin precursor (-)-pluviatolide. *ACS Synth Biol* **2020**, *9* (11), 3091-3103. DOI: 10.1021/acssynbio.0c00354

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

(137) Doudna, J. A. The promise and challenge of therapeutic genome editing. *Nature* **2020**, *578* (7794), 229-236. DOI: 10.1038/s41586-020-1978-5

(138) Matres, J. M.; Hilscher, J.; Datta, A.; Armario-Najera, V.; Baysal, C.; He, W.; Huang, X.; Zhu, C.; Valizadeh-Kamran, R.; Trijatmiko, K. R.; et al. Genome editing in cereal crops: An overview. *Transgenic Res* **2021**, *30* (4), 461-498. DOI: 10.1007/s11248-021-00259-6

(139) Doudna, J. Perspective: Embryo editing needs scrutiny. *Nature* **2015**, *528* (7580), S6-S6. DOI: 10.1038/528S6a

(140) Sanger, F.; Nicklen, S.; Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci* **1977**, *74* (12), 5463-5467. DOI: 10.1073/pnas.74.12.5463

(141) Carr, P. A.; Church, G. M. Genome engineering. *Nat Biotechnol* **2009**, *27* (12), 1151-1162. DOI: 10.1038/nbt.1590

(142) Sternberg, N.; Hamilton, D. Bacteriophage P1 site-specific recombination: I. Recombination between loxP sites. *J Mol Biol* **1981**, *150* (4), 467-486. DOI: 10.1016/0022-2836(81)90375-2

(143) Cox, M. M. The FLP protein of the yeast 2-microns plasmid: Expression of a eukaryotic genetic recombination system in *Escherichia coli*. *Proc Natl Acad Sci* **1983**, *80* (14), 4223-4227. DOI: 10.1073/pnas.80.14.4223

(144) Pines, G.; Freed, E. F.; Winkler, J. D.; Gill, R. T. Bacterial recombineering: Genome engineering via phage-based homologous recombination. *ACS Synth Biol* **2015**, *4* (11), 1176-1185. DOI: 10.1021/acssynbio.5b00009

(145) Hamilton, C. M.; Aldea, M.; Washburn, B. K.; Babitzke, P.; Kushner, S. R. New method for generating deletions and gene replacements in *Escherichia coli*. *J Bacteriol* **1989**, *171* (9), 4617-4622. DOI: 10.1128/jb.171.9.4617-4622.1989

(146) Scherer, S.; Davis, R. W. Replacement of chromosome segments with altered DNA sequences constructed *in vitro*. *Proc Natl Acad Sci* **1979**, *76* (10), 4951-4955. DOI: 10.1073/pnas.76.10.4951

(147) Smithies, O.; Gregg, R. G.; Boggs, S. S.; Koralewski, M. A.; Kucherlapati, R. S. Insertion of DNA sequences into the human chromosomal β-globin locus by homologous recombination. *Nature* **1985**, *317* (6034), 230-234. DOI: 10.1038/317230a0

(148) Kim, J. S. Genome editing comes of age. *Nat Protoc* **2016**, *11* (9), 1573-1578. DOI: 10.1038/nprot.2016.104

(149) Zhang, Y.; Buchholz, F.; Muyrers, J. P. P.; Stewart, A. F. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat Genet* **1998**, *20* (2), 123-128. DOI: 10.1038/2417

(150) Datsenko, K. A.; Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **2000**, *97* (12), 6640-6645. DOI: 10.1073/pnas.120163297

(151) Reisch, C. R.; Prather, K. L. The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in *Escherichia coli*. *Sci Rep* **2015**, *5*, 15096. DOI: 10.1038/srep15096

(152) Rouet, P.; Smih, F.; Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol* **1994**, *14* (12), 8096-8106. DOI: 10.1128/mcb.14.12.8096-8106.1994

(153) Rudin, N.; Haber, J. E. Efficient repair of HO-induced chromosomal breaks in *Saccharomyces cerevisiae* by recombination between flanking homologous sequences. *Mol Cell Biol* **1988**, *8* (9), 3918-3928. DOI: 10.1128/mcb.8.9.3918-3928.1988

(154) Strobel, S. A.; Dervan, P. B. Site-specific cleavage of a yeast chromosome by oligonucleotide-directed triple-helix formation. *Science* **1990**, *249* (4964), 73-75. DOI: 10.1126/science.2195655

(155) Mojica, F. J. M.; Montoliu, L. On the origin of CRISPR-Cas technology: From prokaryotes to mammals. *Trends Microbiol* **2016**, *24* (10), 811-820. DOI: 10.1016/j.tim.2016.06.005

(156) Kim, Y. G.; Cha, J.; Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to *Fok* I cleavage domain. *Proc Natl Acad Sci* **1996**, *93* (3), 1156-1160. DOI: 10.1073/pnas.93.3.1156

(157) Christian, M.; Cermak, T.; Doyle, E. L.; Schmidt, C.; Zhang, F.; Hummel, A.; Bogdanove, A. J.; Voytas, D. F. Targeting DNA double-strand breaks with TAL Effector Nucleases. *Genetics* **2010**, *186* (2), 757-761. DOI: 10.1534/genetics.110.120717 (acccessed 7/30/2024).

(158) Fleischmann, R. D.; Adams, M. D.; White, O.; Clayton, R. A.; Kirkness, E. F.; Kerlavage, A. R.; Bult, C. J.; Tomb, J.-F.; Dougherty, B. A.; Merrick, J. M.; et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **1995**, *269* (5223), 496-512. DOI: 10.1126/science.7542800

(159) Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2006**, *2*, 2006 0008. DOI: 10.1038/msb4100050

(160) Hayashi, K.; Morooka, N.; Yamamoto, Y.; Fujita, K.; Isono, K.; Choi, S.; Ohtsubo, E.; Baba, T.; Wanner, B. L.; Mori, H.; et al. Highly accurate genome sequences of *Escherichia coli*

K-12 strains MG1655 and W3110. *Mol Syst Biol* **2006**, *2*, 2006 0007. DOI: 10.1038/msb4100049

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

(162) Wang, J. Y.; Doudna, J. A. CRISPR technology: A decade of genome editing is only the beginning. *Science* **2023**, *379* (6629), eadd8643. DOI: 10.1126/science.add8643

(163) Jiang, Y.; Chen, B.; Duan, C.; Sun, B.; Yang, J.; Yang, S. Multigene editing in the *Escherichia coli* genome *via* the CRISPR-Cas9 system. *Appl Environ Microbiol* **2015**, *81* (7), 2506-2514. DOI: 10.1128/AEM.04023-14

(164) *Press release*. https://www.nobelprize.org/prizes/chemistry/2020/press-release/ (accessed 2024-10-31).

(165) Marraffini, L. A. CRISPR-Cas immunity in prokaryotes. *Nature* **2015**, *526* (7571), 55-61. DOI: 10.1038/nature15386

(166) Cui, L.; Bikard, D. Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*. *Nucleic Acids Res* **2016**, *44* (9), 4243-4251. DOI: 10.1093/nar/gkw223

(167) Su, T.; Liu, F.; Gu, P.; Jin, H.; Chang, Y.; Wang, Q.; Liang, Q.; Qi, Q. A CRISPR-Cas9 Assisted Non-Homologous End-Joining Strategy for One-step Engineering of Bacterial Genome. *Sci Rep* **2016**, *6*, 37895. DOI: 10.1038/srep37895

(168) Jakociunas, T.; Jensen, M. K.; Keasling, J. D. CRISPR/Cas9 advances engineering of microbial cell factories. *Metab Eng* **2016**, *34*, 44-59. DOI: 10.1016/j.ymben.2015.12.003

(169) Dominguez, A. A.; Lim, W. A.; Qi, L. S. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol* **2016**, *17* (1), 5-15. DOI: 10.1038/nrm.2015.2

(170) Jiang, W.; Bikard, D.; Cox, D.; Zhang, F.; Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* **2013**, *31* (3), 233-239. DOI: 10.1038/nbt.2508

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

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

(173) Xia, J.; Wang, L.; Zhu, J. B.; Sun, C. J.; Zheng, M. G.; Zheng, L.; Lou, Y. H.; Shi, L. Expression of *Shewanella frigidimarina* fatty acid metabolic genes in *E. coli* by CRISPR/cas9-coupled lambda Red recombineering. *Biotechnol Lett* **2016**, *38* (1), 117-122. DOI: 10.1007/s10529-015-1956-4

(174) Bassalo, M. C.; Garst, A. D.; Halweg-Edwards, A. L.; Grau, W. C.; Domaille, D. W.; Mutalik, V. K.; Arkin, A. P.; Gill, R. T. Rapid and efficient one-step metabolic pathway integration in *E. coli*. *ACS Synth Biol* **2016**, *5* (7), 561-568. DOI: 10.1021/acssynbio.5b00187

(175) Ng, I. S.; Hung, Y.-H.; Kao, P.-H.; Zhou, Y.; Zhang, X. CRISPR/Cas9 nuclease cleavage enables marker-free genome editing in *Escherichia coli* : A sequential study. *J Taiwan Inst Chem Eng* **2016**, *68*, 31-39. DOI: 10.1016/j.jtice.2016.08.015

(176) Zhang, H.; Cheng, Q. X.; Liu, A. M.; Zhao, G. P.; Wang, J. A novel and efficient method for bacteria genome editing employing both CRISPR/Cas9 and an antibiotic resistance cassette. *Front Microbiol* **2017**, *8*, 812. DOI: 10.3389/fmicb.2017.00812

(177) Li, Y.; Yan, F.; Wu, H.; Li, G.; Han, Y.; Ma, Q.; Fan, X.; Zhang, C.; Xu, Q.; Xie, X.; et al. Multiple-step chromosomal integration of divided segments from a large DNA fragment via CRISPR/Cas9 in *Escherichia coli*. *J Ind Microbiol Biotechnol* **2019**, *46* (1), 81-90. DOI: 10.1007/s10295-018-2114-5

(178) Seo, J.-H.; Baek, S.-W.; Lee, J.; Park, J.-B. Engineering *Escherichia coli* BL21 genome to improve the heptanoic acid tolerance by using CRISPR-Cas9 system. *Biotechnol Bioprocess Eng* **2017**, *22* (3), 231-238. DOI: 10.1007/s12257-017-0158-4

(179) Ding, W.; Weng, H.; Du, G.; Chen, J.; Kang, Z. 5-Aminolevulinic acid production from inexpensive glucose by engineering the C4 pathway in *Escherichia coli*. *J Ind Microbiol Biotechnol* **2017**, *44* (8), 1127-1135. DOI: 10.1007/s10295-017-1940-1

(180) Zerbini, F.; Zanella, I.; Fraccascia, D.; Konig, E.; Irene, C.; Frattini, L. F.; Tomasi, M.; Fantappie, L.; Ganfini, L.; Caproni, E.; et al. Large scale validation of an efficient CRISPR/Cas-based multi gene editing protocol in *Escherichia coli*. *Microb Cell Fact* **2017**, *16* (1), 68. DOI: 10.1186/s12934-017-0681-1

(181) Chung, M. E.; Yeh, I. H.; Sung, L. Y.; Wu, M. Y.; Chao, Y. P.; Ng, I. S.; Hu, Y. C. Enhanced integration of large DNA into *E. coli* chromosome by CRISPR/Cas9. *Biotechnol Bioeng* **2017**, *114* (1), 172-183. DOI: 10.1002/bit.26056

(182) McCarty, N. S.; Graham, A. E.; Studena, L.; Ledesma-Amaro, R. Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nat Commun* **2020**, *11* (1), 1281. DOI: 10.1038/s41467-020-15053-x

(183) Konstantakos, V.; Nentidis, A.; Krithara, A.; Paliouras, G. CRISPR-Cas9 gRNA efficiency prediction: an overview of predictive tools and the role of deep learning. *Nucleic Acids Res* **2022**, *50* (7), 3616-3637. DOI: 10.1093/nar/gkac192

(184) Tong, Y.; Jorgensen, T. S.; Whitford, C. M.; Weber, T.; Lee, S. Y. A versatile genetic engineering toolkit for *E. coli* based on CRISPR-prime editing. *Nat Commun* **2021**, *12* (1), 5206. DOI: 10.1038/s41467-021-25541-3

(185) Li, Q.; Sun, M.; Lv, L.; Zuo, Y.; Zhang, S.; Zhang, Y.; Yang, S. Improving the editing efficiency of CRISPR-Cas9 by reducing the generation of escapers based on the surviving mechanism. *ACS Synth Biol* **2023**. DOI: 10.1021/acssynbio.2c00619

(186) Ou, X. Y.; Wu, X. L.; Peng, F.; Zeng, Y. J.; Li, H. X.; Xu, P.; Chen, G.; Guo, Z. W.; Yang, J. G.; Zong, M. H.; et al. Metabolic engineering of a robust *Escherichia coli* strain with a dual protection system. *Biotechnol Bioeng* **2019**, *116* (12), 3333-3348. DOI: 10.1002/bit.27165

(187) Mulet, A. P.; Ripoll, M.; Betancor, L. CRISPR tools in bacterial whole-cell biocatalysis. *ACS Sustainable Chem Eng* **2023**, *11* (44), 15765-15788. DOI: 10.1021/acssuschemeng.3c05735

(188) Bryant, J. A.; Sellars, L. E.; Busby, S. J.; Lee, D. J. Chromosome position effects on gene expression in *Escherichia coli* K-12. *Nucleic Acids Res* **2014**, *42* (18), 11383-11392. DOI: 10.1093/nar/gku828

(189) Goormans, A. R.; Snoeck, N.; Decadt, H.; Vermeulen, K.; Peters, G.; Coussement, P.; Van Herpe, D.; Beauprez, J. J.; De Maeseneire, S. L.; Soetaert, W. K. Comprehensive study on *Escherichia coli* genomic expression: Does position really matter? *Metab Eng* **2020**, *62*, 10-19. DOI: 10.1016/j.ymben.2020.07.007

(190) Scholz, S. A.; Diao, R.; Wolfe, M. B.; Fivenson, E. M.; Lin, X. N.; Freddolino, P. L. High-resolution mapping of the *Escherichia coli* chromosome reveals positions of high and low transcription. *Cell Syst* **2019**, *8* (3), 212-225. DOI: 10.1016/j.cels.2019.02.004

(191) Su, B.; Song, D.; Zhu, H. Homology-dependent recombination of large synthetic pathways into *E. coli* genome via lambda-Red and CRISPR/Cas9 dependent selection methodology. *Microb Cell Fact* **2020**, *19* (1), 108. DOI: 10.1186/s12934-020-01360-x

(192) Nakagawa, S.; Ishino, S.; Teshiba, S. Construction of catalase deficient *Escherichia coli* strains for the production of uricase. *Biosci Biotechnol Biochem* **1996**, *60* (3), 415-420. DOI: 10.1271/bbb.60.415

(193) Doukyu, N.; Taguchi, K. Involvement of catalase and superoxide dismutase in hydrophobic organic solvent tolerance of *Escherichia coli*. *AMB Express* **2021**, *11* (1), 97. DOI: 10.1186/s13568-021-01258-w

(194) Seaver, L. C.; Imlay, J. A. Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. *J Bacteriol* **2001**, *183* (24), 7173-7181. DOI: 10.1128/JB.183.24.7173-7181.2001

(195) Garikipati, S. V.; McIver, A. M.; Peeples, T. L. Whole-cell biocatalysis for 1-naphthol production in liquid-liquid biphasic systems. *Appl Environ Microbiol* **2009**, *75* (20), 6545-6552. DOI: 10.1128/AEM.00434-09

(196) Abril, N.; Pueyo, C. Mutagenesis in *Escherichia coli* lacking catalase. *Environ Mol Mutagen* **1990**, *15* (4), 184-189. DOI: 10.1002/em.2850150403

(197) Omura, T.; Sato, R. The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. *J Biol Chem* **1964**, *239* (7), 2370-2378. DOI: 10.1016/S0021-9258(20)82244-3

(198) Omura, T.; Sato, R. The carbon monoxide-binding pigment of liver microsomes. II. Solubilization, purification, and properties. *J Biol Chem* **1964**, *239* (7), 2379-2385. DOI: 10.1016/S0021-9258(20)82245-5

(199) Omura, T. Forty years of cytochrome P450. *Biochem Biophys Res Commun* **1999**, *266* (3), 690-698. DOI: 10.1006/bbrc.1999.1887

(200) Bertelmann, C.; Buhler, B. Strategies found not to be suitable for stabilizing high steroid hydroxylation activities of CYP450 BM3-based whole-cell biocatalysts. *PLoS One* **2024**, *19* (9), e0309965. DOI: 10.1371/journal.pone.0309965

(201) Pflug, S.; Richter, S. M.; Urlacher, V. B. Development of a fed-batch process for the production of the cytochrome P450 monooxygenase CYP102A1 from *Bacillus megaterium* in *E. coli*. *J Biotechnol* **2007**, *129* (3), 481-488. DOI: 10.1016/j.jbiotec.2007.01.013

(202) Hu, B.; Yu, H.; Zhou, J.; Li, J.; Chen, J.; Du, G.; Lee, S. Y.; Zhao, X. Whole-cell P450 biocatalysis using engineered *Escherichia coli* with fine-tuned heme biosynthesis. *Adv Sci (Weinheim, Ger)* **2022**, e2205580. DOI: 10.1002/advs.202205580

(203) Ge, J.; Wang, X.; Bai, Y.; Wang, Y.; Wang, Y.; Tu, T.; Qin, X.; Su, X.; Luo, H.; Yao, B.; et al. Engineering *Escherichia coli* for efficient assembly of heme proteins. *Microb Cell Fact* **2023**, *22* (1), 59. DOI: 10.1186/s12934-023-02067-5

(204) Schäfer, L.; Karande, R.; Bühler, B. Maximizing biocatalytic cyclohexane hydroxylation by modulating cytochrome P450 monooxygenase expression in *P. taiwanensis* VLB120. *Front Bioeng Biotechnol* **2020**, *8*, 140. DOI: 10.3389/fbioe.2020.00140

(205) Bakkes, P. J.; Riehm, J. L.; Sagadin, T.; Ruhlmann, A.; Schubert, P.; Biemann, S.; Girhard, M.; Hutter, M. C.; Bernhardt, R.; Urlacher, V. B. Engineering of versatile redox partner fusions that support monooxygenase activity of functionally diverse cytochrome P450s. *Sci Rep* **2017**, *7* (1), 9570. DOI: 10.1038/s41598-017-10075-w

(206) Murata, J.; Ono, E.; Yoroizuka, S.; Toyonaga, H.; Shiraishi, A.; Mori, S.; Tera, M.; Azuma, T.; Nagano, A. J.; Nakayasu, M.; et al. Oxidative rearrangement of (+)-sesamin by CYP92B14 co-generates twin dietary lignans in sesame. *Nat Commun* **2017**, *8* (1), 2155. DOI: 10.1038/s41467-017-02053-7

(207) Nadeem, M.; Palni, L.; Purohit, A.; Pandey, H.; Nandi, S. Propagation and conservation of *Podophyllum hexandrum* Royle: an important medicinal herb. *Biol Conserv* **2000**, *92* (1), 121-129. DOI: 10.1016/S0006-3207(99)00059-2

(208) Ricklefs, E.; Girhard, M.; Koschorreck, K.; Smit, M. S.; Urlacher, V. B. Two-step one-pot synthesis of pinoresinol from eugenol in an enzymatic cascade. *ChemCatChem* **2015**, *7* (12), 1857-1864. DOI: 10.1002/cctc.201500182

(209) Russell, W. R.; Forrester, A. R.; Chesson, A.; Burkitt, M. J. Oxidative coupling during lignin polymerization is determined by unpaired electron delocalization within parent phenylpropanoid radicals. *Arch Biochem Biophys* **1996**, *332* (2), 357-366. DOI: 10.1006/abbi.1996.0353

(210) Wang, L.; Chen, Q.; Li, D.; Chen, J.; Wang, H.; Hu, M.; Shan, X.; Zeng, W.; Zhou, J. Improving production of coniferyl alcohol in *Escherichia coli* by regulating the multiple cofactors and efflux of intermediates. *ACS Sustainable Chem Eng* **2024**, *12* (26), 9704-9715. DOI: 10.1021/acssuschemeng.4c01190

(211) Tramontina, R.; Galman, J. L.; Parmeggiani, F.; Derrington, S. R.; Bugg, T. D. H.; Turner, N. J.; Squina, F. M.; Dixon, N. Consolidated production of coniferol and other high-value aromatic alcohols directly from lignocellulosic biomass. *Green Chem* **2020**, *22* (1), 144-152. DOI: 10.1039/c9gc02359c

(212) Wu, M. Y.; Sung, L. Y.; Li, H.; Huang, C. H.; Hu, Y. C. Combining CRISPR and CRISPRi systems for metabolic engineering of *E. coli* and 1,4-BDO biosynthesis. *ACS Synth Biol* **2017**, *6* (12), 2350-2361. DOI: 10.1021/acssynbio.7b00251

(213) Liu, P.; Zhang, B.; Yao, Z. H.; Liu, Z. Q.; Zheng, Y. G. Multiplex design of metabolic network for production of L-homoserine in *Escherichia coli*. *Appl Environ Microbiol* **2020**, *86* (20). DOI: 10.1128/AEM.01477-20

(214) Chen, M.; Liang, H.; Han, C.; Zhou, P.; Xing, Z.; Chen, Q.; Liu, Y.; Xie, G. A.; Xie, R. Engineering of global transcription factor FruR to redirect the carbon flow in *Escherichia coli* for enhancing L-phenylalanine biosynthesis. *Microb Cell Fact* **2022**, *21* (1), 222. DOI: 10.1186/s12934-022-01954-7

(215) Contreras, J. L.; Sherkow, J. S. CRISPR, surrogate licensing, and scientific discovery. *Science* **2017**, *355* (6326), 698-700. DOI: 10.1126/science.aal4222

(216) Sheridan, C. Amyloid is everywhere, but new treatments could stop the toxic build up. *Nat Biotechnol* **2024**, *42* (11), 1625-1628. DOI: 10.1038/s41587-024-02471-1

(217) Wang, K.; Fredens, J.; Brunner, S. F.; Kim, S. H.; Chia, T.; Chin, J. W. Defining synonymous codon compression schemes by genome recoding. *Nature* **2016**, *539* (7627), 59-64. DOI: 10.1038/nature20124

(218) Okshevsky, M.; Xu, Y.; Masson, L.; Arbour, M. Efficiency of CRISPR-Cas9 genetic engineering in *Escherichia coli* BL21 is impaired by lack of Lon protease. *J Microbiol Methods* **2022**, *204*, 106648. DOI: 10.1016/j.mimet.2022.106648

(219) Feng, X.; Zhao, D.; Zhang, X.; Ding, X.; Bi, C. CRISPR/Cas9 assisted multiplex genome editing technique in *Escherichia coli*. *Biotechnol J* **2018**, *13* (9), e1700604. DOI: 10.1002/biot.201700604

5 Abbreviations

(s)gRNA	(single-)guide RNA
ABE	acetone-butanol-ethanol
ADH	alcohol dehydrogenase
AKR	aldo-keto reductase
ATP	adenosine triphosphate
BM3	CYP102A1
bp	base pair
BSA	bovine serum albumin
C. glutamicum	*Corynebacterium glutamicum*
CAD	cinnamyl alcohol dehydrogenase
Cas	CRISPR-associated system
Cas9n	Cas9-nickase
CFE	cell-free extract
*Cg*L1	laccase from *Corynebacterium glutamicum*
CH	carbon hydrogen
CO2	carbon dioxide
Cpd 0	compound 0
Cre	recombinase from P1 bacteriophage
CRISPR	clustered regularly interspaced short palindromic repeats
CRISPRa	CRISPR activation
CRISPRa	CRISPR activation
CRISPRi	CRISPR interference
CRISPRi	CRISPR interference
crRNA	CRISPR-RNA
cww	cell wet weight
CYP	cytochrome P450
CYP102A1	CYP from *Bacillus megaterium*
CYP154E1	CYP from *Thermobifidum fusca* YX
dCas9	catalytically inactive Cas9
*Dca*UPO	short UPO from *Daldinia caldariorum*
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DSB	double strand break
E. coli	*Escherichia coli*
FAD	flavin adenine dicnucleotide
FdR	flavodoxin reductase from *Escherichia coli*
Flp	recombinase from *Saccharomyces cerevisiae*
FRT	recognition sequence of Flp
GC-FID	gas chromatography - flame ionization detector
Gox	glucose oxidase
HA	homology arm (for recombineering)
HDR	homology-directed repair
HK	hydroxyketamine
HNH	nuclease domain of Cas9
HNK	hydroxynorketamine
HPLC	high performance liquid chromatography
IPTG	isopropyl β-d-1-thiogalactopyranoside
ISTD	internal standard
katE	gene encoding hydroperoxidase II from *E. coli*
katG	gene encoding hydroperoxidase I from *E. coli*
kDa	kilodalton
LB	lysogeny broth
LC/MS	liquid chromatography / mass spectrometry
lopA	gene encoding Lon protease
loxP	recognition sequence of Cre
m/z	mass-to-charge ratio
mM	millimolar
N_{20}	targeting sequence of Cas9
NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
NHEJ	non-homologous end joining
NMR	nuclear magnetic resonance
OD_{600} / OD_{650}	optical density at 600 nm or 650 nm
P. crispum	*Petroselinum crispum*
P. pastoris	*Pichia pastoris*
P450	cytochrome P450
PAM	protospacer adjacent motif
*Pc*4CL	4-coumarate-CoA ligase from *Petroselinum crispum*
PCR	polymerase chain reaction
PDA	photodiode array detector
pgRNA-ET / pgRNA-DUET	vector series for chromosomal integration
rpm	revolutions per minute
RuvC	nuclease domain of Cas9
S. cerevisiae	*Saccharomyces cerevisiae*
SAM	S-adenosyl methionine
SDS-PAGE	sodium dodecyl sulfate - polyacrylamide gel electrophoresis
*Ssca*CYP	CYP peroxygenase from *Streptomyces scabiei*
T. fusca YX	*Thermobifida fusca* YX
T7RNAP	T7 RNA polymerase
TALEN	transcription-activator-like effector nucleases
TB	terrific broth
TIC	total ion chromatogram
TON	turnover number
tracrRNA	trans-activation CRISPR-RNA
UPO	unspecific peroxygenase
UV	ultraviolet
WCs	whole cells
YkuN	flavodoxin from *Bacillus subtilis*
Z. mays	*Zea mays*
ZFN	zinc finger nuclease
*Zm*CCR	cinnamoyl-CoA reductase from *Zea mays*

6 Acknowledgements

After four years of lab time at the Institute of Biochemistry, only briefly interrupted by a full-scale pandemic and the relocation of the institute to a new building, it is overdue to thank a few people without whom this work would not have been possible.

First and foremost, I would like to thank Prof. Dr. Vlada B. Urlacher for her trust, her support, and the liberties she granted me. Thank you for your time, patience, and the opportunity to follow my dreams while paying my rent. I very much enjoyed our meetings and the time at your institute, learned a lot during that time and am looking forward to catching up in the future.

I also want to express my gratitude to the co-examiner and mentor Prof. Lutz Schmitt.

A very special thanks goes to PD Dr. Ulrich Schulte, who, in his role as student counselor and course coordinator, not only enabled but also strongly supported my switch from studying chemistry to biochemistry. Without him, I probably would not have found my way into biochemistry. Rest in peace.

I would especially like to thank Dr. Marco Girhard for his straightforward support in all questions and problems in everyday lab work and beyond. It was a pleasure to share an office with you!

Thanks also go to Dr. Katja Koschorreck for her constant support in cloning and fermentation questions and the management of the plasmid and strain collections.

Furthermore, I would like to thank all the students whose theses I had the pleasure to supervise: Christian Dreiz, Marieke Willing, Lisa M. Böhmer, and Alexander Wassing. I have learned a lot from and along with you. By pursuing your own goals with your theses, you have greatly helped the progress of my projects. In addition, the lab work would definitely not have been as much fun without you!

I would like to thank the technical assistance crew, Sebastian Hölzel, Martin Sosna and Melanie Wachtmeister, especially for the preparation of all the practical courses that we had the pleasure of carrying out together.

Moreover, I would like to thank my co-authors and all former and current members of the Institute of Biochemistry, Chair II for their support and the pleasant working atmosphere.

Finally, I would like to thank my family, my fiancée, and all my friends who supported me on my way.