

---

# **Synthetic phenylpropanoid pathway for the *in vivo* production of coniferyl alcohol in *Escherichia coli***

---

Synthetischer Phenylpropanoidweg zur *in vivo*-Produktion von Coniferylalkohol in  
*Escherichia coli*

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

vorgelegt von

Anna Christiane Kohl, Master of Science

aus

Wien, Österreich

Berichter: Prof. apl. Dr. rer. nat. Ulrich Commandeur  
Prof. Dr.-Ing. Lars M. Blank

Tag der mündlichen Prüfung: 02.02.2023

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.



## Table of contents

Table of contents .....	I
List of Abbreviations .....	V
Abstract.....	1
Zusammenfassung.....	3
I. Introduction.....	5
I.1 Lignans .....	5
I.2 Lignin and the phenylpropanoid pathway .....	7
I.3 Tyrosine and phenylalanine ammonia lyases .....	11
I.4 Cytochrome P450 monooxygenases (C3H) .....	14
I.4.1 CYP199A2 from <i>Rhodopseudomonas palustris</i> .....	16
I.4.2 C3H alternative: HpaBC from <i>Escherichia coli</i> .....	17
I.5 O-methyltransferases (OMTs) .....	18
I.6 Current microbial production of phenylpropanoids .....	20
I.7 Objectives of the research .....	21
II. Results and Discussion .....	23
II.1 Deamination .....	23
II.1.1 Tyrosine ammonia lyases .....	24
II.1.2 Evaluation of pH change by applied substrates.....	26
II.1.3 Phenylalanine ammonia lyases.....	27
II.2 C3-hydroxylation .....	29
II.2.1 The <i>E. coli</i> hydroxylase complex HpaBC .....	30
II.2.2 The cytochrome P450 monooxygenase from <i>Rhodopseudomonas palustris</i> .....	33
II.3 Upstream pathway .....	43
II.3.1 The combined <i>in vivo</i> activity of TAL and C3H enzymes .....	44
II.3.2 The combined <i>in vivo</i> activity of PAL and C3H/C4H enzyme .....	48
II.4 Downstream pathway .....	51
II.4.1 Different methyltransferases expressed in a genome-integrated <i>E. coli</i> strain ..	51
II.4.2 Different methyltransferases expressed in <i>E. coli</i> BL21(DE3).....	54
II.4.3 Increased substrate feed into the downstream pathway.....	59
II.5 The synthetic pathway towards monolignols .....	61
II.5.1 From L-tyrosine to <i>p</i> -coumaryl alcohol.....	61
II.5.2 From L-tyrosine to coniferyl alcohol .....	64
II.6 Growing cell assay .....	70
II.6.1 Downstream pathway.....	70

II.6.2	Production of coniferyl alcohol in <i>E. coli</i> BL21(DE3).....	72
II.6.3	Production of coniferyl alcohol in genome-integrated <i>E. coli</i> strains .....	77
II.6.4	Increase of coniferyl alcohol production .....	81
III.	General discussion.....	85
III.1	Individual enzyme characterizations .....	85
III.1.1	Deamination.....	86
III.1.2	C3-hydroxylation .....	88
III.2	Upstream pathway .....	91
III.2.1	Combined activity of TAL and C3H enzymes .....	93
III.2.2	Combined activity of a PAL and C3H/C4H enzyme .....	95
III.3	Downstream pathway .....	96
III.4	Complete synthetic pathway for monolignol production .....	98
III.4.1	In resting cells.....	98
III.4.2	In growing cells .....	100
IV.	Conclusion and outlook.....	103
V.	Materials and Methods .....	107
V.1	Materials .....	107
V.1.1	Instruments .....	107
V.1.2	Chemicals and consumables.....	110
V.1.3	Software .....	111
V.1.4	Media, buffers and solutions .....	111
V.1.5	Antibodies.....	116
V.1.6	DNA Oligonucleotides (Primers).....	116
V.1.7	Plasmids .....	119
V.1.8	Bacterial strains .....	123
V.2	Methods .....	123
V.2.1	Molecular Biology Methods.....	123
V.2.2	Microbiological Methods.....	129
V.2.3	Biochemical Methods .....	132
V.2.4	Analytical Methods .....	135
VI.	References .....	137
VII.	Appendix .....	151
VII.1	Used gene sequences .....	151
VII.2	Construction of plasmids via restriction and ligation .....	153
VII.3	Construction of plasmids via other molecular biology methods.....	155



VII.4	Supplementary figures.....	156
VII.5	Toxicity of different substrates on <i>E. coli</i> .....	168
VII.6	Plasmid maps.....	169
VII.7	DNA and protein ladders.....	170
VII.8	List of Figures .....	171
VII.9	List of Tables.....	174
Danksagung .....		175
Declaration.....		177
Curriculum vitae.....		179



## List of Abbreviations

$\alpha$	alpha
°C	degree Celsius (temperature unit)
%	percent
3'	3' end of DNA
5'	5' end of DNA
4CL	4-coumarate:CoA ligase
5/OD <sub>600</sub>	quotient of 5 and OD at 600 nm
A <sub>260</sub>	absorbance at a wavelength of 260 nm
A <sub>260/280</sub>	quotient of absorbance at 260 nm and 280 nm
A <sub>260/230</sub>	quotient of absorbance at 260 nm and 230 nm
AP	alkaline phosphatase
APS	ammonium persulfate
<i>At</i>	<i>Arabidopsis thaliana</i>
BCIP	5-Bromo-4-chloro-3-indolyl phosphate <i>p</i> -toluidine salt
bp	base pair
C3H	C3-hydroxylase, <i>p</i> -coumarate 3-hydroxylase
C4H	C4-hydroxylase, cinnamate 4-hydroxylase
CaCl	calcium chloride
CAD	cinnamyl alcohol dehydrogenase
CCR	cinnamoyl-CoA reductase
cf.	compare (confer in latin)
CIP	calf intestinal phosphatase
CCoAOMT	caffeoyl-CoA 3- <i>O</i> -methyltransferase
co	codon-optimized
CoA	co-enzyme A
COMT	caffeic acid 3- <i>O</i> -methyltransferase
cPCR	colony PCR
CYP	cytochrome P450 enzyme/monooxygenase
DEPC water	nuclease-free water treated with diethyl pyrocarbonate
DMSO	dimethyl sulfoxide
DNA	desoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
dsDNA	double-stranded DNA
<i>E. coli</i>	<i>Escherichia coli</i>
e.g.	for example (exempli gratia in latin)
EDTA	ethylenediaminetetraacetic acid
ER	endoplasmic reticulum
FAD	flavin adenine dinucleotide
FC	fragment crystallizable
<i>Fj</i>	<i>Flavobacterium johnsoniae</i>
FMN	flavin mononucleotide
FpR	reductase from <i>Escherichia coli</i>

## List of Abbreviations

---

Fw	forward (primer)
g	gram (mass unit)
GAM	goat anti-mouse antibody
g/L	gram per liter
$g_{wcw}$	wet cell weight in gram
h	hour(s) (time unit)
<i>Ha</i>	<i>Herpetosiphon aurantiacus</i>
HCl	hydrochloric acid
his <sub>6</sub>	hexa histidine-tag
his <sub>10</sub>	deca histidine-tag
HPLC	high-performance liquid chromatography
IgG	immunoglobulin G
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
IS	internal standard
kb	kilo base pair
KCl	potassium chloride
kDa	kilodalton (atomic mass unit)
K <sub>2</sub> HPO <sub>4</sub>	di-potassium hydrogen phosphate
KH <sub>2</sub> PO <sub>4</sub>	potassium dihydrogen phosphate
KP <sub>i</sub>	potassium phosphate buffer
L	liter (volume unit)
LB	Luria-Bertani broth
L-DOPA	levodopa, L-3,4-dihydroxyphenylalanine
$\mu$ g	microgram
$\mu$ L	microliter
$\mu$ M	micromolar, $10^{-6}$ mol/L
M	molarity in mol/L (concentration unit)
MCS I/II	multiple cloning site I or II
mg	milligram
MgSO <sub>4</sub>	magnesium sulfate
min	minute(s) (time unit)
MIO	3,5-dihydro-5-methylidene-4H-imidazol-4-one
mL	milliliter
mM	millimolar, $10^{-3}$ mol/L
M-MLV	Moloney Murine Leukemia virus reverse transcriptase
MnCl <sub>2</sub>	manganese(II) chloride
mRNA	messenger ribonucleic acid
MQ-H <sub>2</sub> O	MilliQ water, distilled water with low electrolytic conductivity
N $\Delta$ 7	N-terminal deletion of seven amino acids
NaCl	sodium chloride
NAD(P) <sup>+</sup>	oxidized nicotinamide adenine dinucleotide (phosphate)
NAD(P)H	reduced nicotinamide adenine dinucleotide (phosphate)
Na <sub>2</sub> HPO <sub>4</sub>	di-sodium hydrogen phosphate
NBT	nitro blue tetrazolium chloride

---

OD <sub>600</sub>	optical density at a wavelength of 600 nm
OMT	<i>O</i> -methyltransferase
P	pellet
P450	cytochrome P450 monooxygenase
PAL	L-phenylalanine ammonia lyase
PBS	phosphate-buffered saline
<i>Pc</i>	<i>Petroselinum crispum</i>
PCR	polymerase chain reaction
PdR	putidaredoxin reductase
Pdx	putidaredoxin
pH	negative decadic logarithm of the activity of hydrogen ions in solution
PIPES	piperazine- <i>N,N'</i> -bis(2-ethanesulfonic acid), buffering agent
PuR	palustrisredoxin reductase
Pux	palustrisredoxin
RBS	ribosome binding site
<i>Rg</i>	<i>Rhodotorula glutinis</i>
RP-HPLC	reverse-phase HPLC
rpm	revolutions per minute
<i>Rs</i>	<i>Rhodobacter sphaeroides</i>
RT	reverse transcription
Rv	reverse (primer)
S	supernatant
SAH	S-adenosyl-L-homocysteine
SAM	S-adenosyl-L-methionine
Sam5	C3H from <i>Saccharothrix espanaensis</i>
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec	second(s) (time unit)
Sesam8	TAL from <i>Saccharothrix espanaensis</i>
SOC	super optimal broth with catabolite repression
S-tag	S-protein epitope tag
t <sub>0</sub>	point in time of substrate addition
TAE	Tris-acetate-EDTA
TAL	L-tyrosine ammonia lyase
TB	Terrific broth
TEMED	<i>N,N,N',N'</i> -Tetramethylethylenediamine
Tris	Tris-(hydroxymethyl)-amino methane, buffering agent
U	unit(s) (of restriction enzymes and ligase)
V	volt (electric potential unit)
v/v	volume per volume
w/v	weight per volume
xg	multiple of earth's gravitational acceleration
YkuN	flavodoxin from <i>Bacillus subtilis</i>
<i>Zm</i>	<i>Zea mays</i>



## Abstract

Lignans constitute a valuable substance group exhibiting many health-promoting effects on the human body as potential pharmaceuticals and nutraceuticals. One of the most popular lignans is (–)-podophyllotoxin, which is applied in its semi-synthetic derivatives for the treatment of various types of cancer. The industrial demand for alternative production routes remains high, as the main extraction method still depends on endangered natural plant resources. Thus, the controlled microbial fermentation could provide the important lignans in sufficient and scalable amounts in the future. As one of the preconditions, the precursor for lignan synthesis coniferyl alcohol needs to be produced in industrial bulk. For a foremost cost-efficient and economically sustainable green synthesis of coniferyl alcohol, the fermentation from inexpensive resources such as glucose carried out in well-examined production strains like *Escherichia coli* (*E. coli*) can be advantageous also with respect to a possible combinable synthesis of lignans.

The objective of this dissertation was to establish the microbial synthesis of coniferyl alcohol in *E. coli* from L-tyrosine or L-phenylalanine according to the plant phenylpropanoid pathway. In the context of metabolic engineering, different enzyme candidates for crucial central reaction steps concerning deamination, C3- and C4-hydroxylation and methylation were first evaluated. By division of the phenylpropanoid pathway into upstream and downstream pathway, the best enzyme combinations were then investigated for both modules. Furthermore, the complete and partly plasmid-based expression of the phenylpropanoid pathway was realized in *E. coli* BL21(DE3) or its genome-integrated derivatives containing variable gene copies and promoters. At last, two different cell assays were applied and compared.

This thesis describes the first successful application of the TALs from *Flavobacterium johnsoniae* (FjTAL) and *Saccharothrix espanaensis* (Sesam8), the cytochrome P450 enzyme from *Rhodopseudomonas palustris* (CYP199A2) as C3H and the two methyltransferases from *Zea mays* for the production of coniferyl alcohol from L-tyrosine in *E. coli*. Noteworthy, caffeic acid was obtained from L-phenylalanine by using the bifunctionality of the CYP199A2 F185L mutant as C3- and C4-hydroxylase (C3H/C4H) and the TAL from *Rhodotorula glutinis* (RgTAL) for the first time. In the end, the highest amount of coniferyl alcohol (~ 850 µM) was achieved in growing *E. coli* cells of a genome-integrated strain. This coniferyl alcohol titer, produced under non-optimized conditions and in a prototype strain, is comparable to a value in one literature report, in which an L-tyrosine overproduction strain and adjusted cultivation conditions were used. Thus, the results from this dissertation lay a substantial foundation for the further development of a microbial coniferyl alcohol production platform.





## Zusammenfassung

Lignane stellen eine wertvolle Substanzgruppe dar, welche viele gesundheitsfördernde Effekte auf den menschlichen Körper als potenzielle Pharma- und Nutrazeutika besitzen. Eines der bekanntesten Lignane ist (–)-Podophyllotoxin, das als semisynthetisches Derivat für die Behandlung verschiedener Krebsarten eingesetzt wird. Der industrielle Bedarf nach alternativen Produktionswegen ist weiterhin hoch, da die Hauptmethode der Extraktion immer noch von gefährdeten Pflanzen als Ressource abhängt. Deswegen könnte eine kontrollierte mikrobielle Fermentation die wichtigen Lignane in ausreichenden und skalierbaren Mengen zukünftig bereitstellen. Als eine Voraussetzung muss der Vorläufer der Lignansynthese Coniferylalkohol als industrielle Massenware produziert werden. Für eine vor allem kosteneffiziente und wirtschaftlich nachhaltige, grüne Synthese von Coniferylalkohol kann eine Fermentation aus günstigen Rohstoffen wie Glucose, durchgeführt in einem gut untersuchten Produktionsstamm wie *Escherichia coli* (*E. coli*) vorteilhaft sein, auch im Hinblick auf eine mögliche kombinierbare Synthese von Lignanen.

Die Zielsetzung dieser Dissertation war es, eine mikrobielle Synthese von Coniferylalkohol in *E. coli* ausgehend von L-Tyrosin oder L-Phenylalanin nach dem pflanzlichen Phenylpropanoidweg zu etablieren. Im Rahmen von *Metabolic Engineering* wurden zuerst verschiedene Enzymkandidaten für die wesentlichen zentralen Reaktionsschritte hinsichtlich Desaminierung, C3- und C4-Hydroxylierung und Methylierung evaluiert. Durch die Einteilung des Phenylpropanoidweges in vor- und nachgelagerten Syntheseweg wurden die besten Enzymkombinationen für beide Module untersucht. Die gesamte oder teilweise plasmid-basierte Expression des Phenylpropanoidweges wurde zudem in *E. coli* BL21(DE3) oder seinen genomintegrierten Abkömmlingen, welche unterschiedliche Genkopien und Promotoren enthalten, umgesetzt. Zuletzt wurden zwei verschiedene Zellassays angewandt und verglichen.

Diese Doktorarbeit beschreibt die erste erfolgreiche Anwendung der TALs *Flavobacterium johnsoniae* (FjTAL) sowie *Saccharothrix espanaensis* (Sesam8), des Cytochrom P450 Enzyms *Rhodospseudomonas palustris* (CYP199A2) als C3H und den zwei Methyltransferasen aus *Zea mays* für die Produktion von Coniferylalkohol aus L-Tyrosin in *E. coli*. Hervorzuheben ist, dass Kaffeesäure aus L-Phenylalanin zum ersten Mal gewonnen wurde, indem die Bifunktionalität der F185L Mutante von CYP199A2 als C3- und C4-Hydroxylase (C3H/C4H) und die TAL von *Rhodotorula glutinis* (RgTAL) verwendet wurden. Schließlich konnte die höchste Menge an Coniferylalkohol (~ 850 µM) in wachsenden *E. coli*-Zellen eines genomintegrierten Stammes erreicht werden. Dieser Titer an Coniferylalkohol, welcher unter nicht-optimierten

Bedingungen und in einem Prototypenstamm erzeugt wurde, ist mit einem Literaturwert vergleichbar, bei welchem ein L-Tyrosin überproduzierender Stamm und angepasste Kultivierungsbedingungen eingesetzt wurden. Die Ergebnisse dieser Dissertation schaffen somit eine substanzielle Basis für die Weiterentwicklung einer mikrobiellen Coniferylalkohol-Produktionsplattform.

## I. Introduction

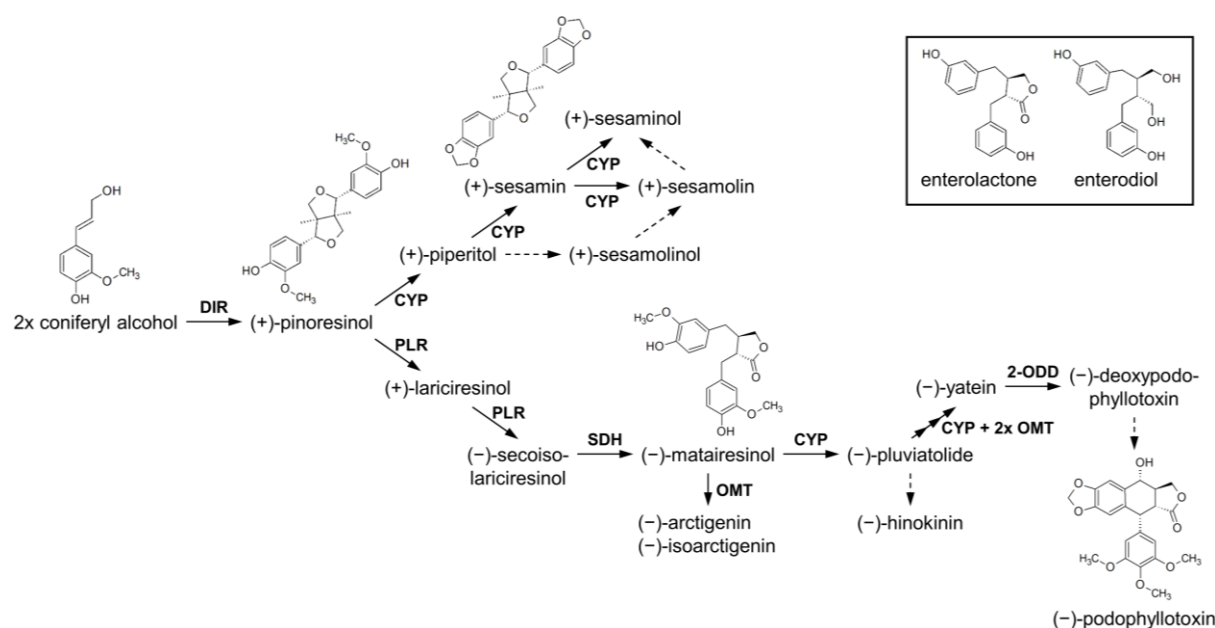
In 2019, coronary heart disease and stroke were the leading causes of death in Europe recording a total of 3.2 million deaths [1]. Further among the top ten causes of death, the incidence for Alzheimer's disease, colon cancer, hypertension and diabetes has been rising since the beginning of the new millennium. One suggestion of causes is supported by the correlation with the Western life-style and high-fat, high-carb low-fiber diets. In contrast, consumption of the Mediterranean or Japanese diet is often associated with a low risk of cardiovascular diseases and a long life expectancy [2]. With regard to that, Adlercreutz and Mazur (1997) compared phytoestrogen plasma levels of men from America, Finland and Japan and found the highest concentration of phytoestrogens in Japanese men [3]. Phytoestrogens are polyphenolic compounds from plant origin which exhibit estrogen-like properties and are assumed to have many health-promoting effects [4]. One group of such phytoestrogens most abundant in the Western diets are lignans [5].

### I.1 Lignans

Lignans are found in fiber-rich plants such as cereal brans and whole grain cereals, in fruits, legumes and other vegetables albeit in generally low titers of less than 2 mg per 100 g fresh weight [6]. The lignan-richest sources are flax and sesame seeds which exceed 130 mg/100 g [6]. Consumed lignans are converted in the human gut to enterolactone and enterodiols by intestinal bacteria (Figure 1.1) [3]. These enterolignans are reported to have protective effects against colon cancer, hormone-dependent cancers like breast and prostate cancer, prevent oxidative stress and damage, as well as lead to a reduced risk of coronary heart disease and osteoporosis [4,7]. In an epidemiological study from 2012, elderly patients with diabetes showed reduced levels of blood plasma glucose and diabetes markers after a three-month long oral administration of flaxseed lignan complexes [8]. Thus, a lignan-rich diet or lignans as dietary supplements could serve as important prevention measures for various life-style related diseases and conditions.

In plants, lignan synthesis is initiated by radical coupling of two coniferyl alcohol moieties by the action of either peroxidases or laccases with the help of dirigent proteins (Figure 1.1). The resulting dimeric compound pinoresinol represents the structurally simplest and basal lignan [9]. Lignans can be divided into eight subgroups according to the way of oxygen incorporation and their cyclization patterns [10]. For instance, the enantiomer (+)-pinoresinol is hydroxylated into (+)-sesamin which both belong to the subgroup of the furofurans. Then again, (+)-pinoresinol can be also converted into (–)-matairesinol, a dibenzylbutyrolactone. Not only

the enterolignans but also the lignans exhibit various biological activities. As another example, (+)-pinoresinol and (+)-sesamin were demonstrated to have anti-oxidant, anti-inflammatory and hepatoprotective properties, while (–)-matairesinol was shown to inhibit the growth of fungi, nematodes and bacteria [11–19]. In accordance with that, lignans are generally considered to play a role in plant defense against herbivores and pathogenic microorganisms, although their complete functions *in planta* have not been elucidated yet [20].



**Figure 1.1: Biosynthesis and examples of lignans.** Several involved enzymes have been identified, but some remain unknown (dashed arrows). DIR– peroxidase or laccase (possibly with dirigent proteins), CYP– specific cytochrome P450 enzyme or hydroxylase, PLR– pinoresinol-lariciresinol reductase, SDH– secoisolariciresinol dehydrogenase, OMT– specific *O*-methyltransferase, 2-ODD– 2-oxoglutarate/Fe(II)-dependent dioxygenase. Adapted from Zálezák *et al.* (2019), Satake *et al.* (2015), Lau and Sattely (2015) and Decembrino *et al.* (2020) [7,9,21,22].

One of the most well-known lignan is the aryltetralin (–)-podophyllotoxin, which exhibits outstanding cytotoxic and antiviral activities [23,24]. Its semi-synthetic derivatives etoposide and teniposide have been applied for over 20 years against various cancer types such as lymphomas, leukemia, lung cancer, neuroblastoma, ovarian, testicular and bladder cancers [7,25,26]. Today, etoposide is on the WHO’s list of essential medicines [21]. Furthermore, (–)-podophyllotoxin is used under the market names Condyllox® and Wartec® for the dermatological treatment of genital warts in Europe [7].

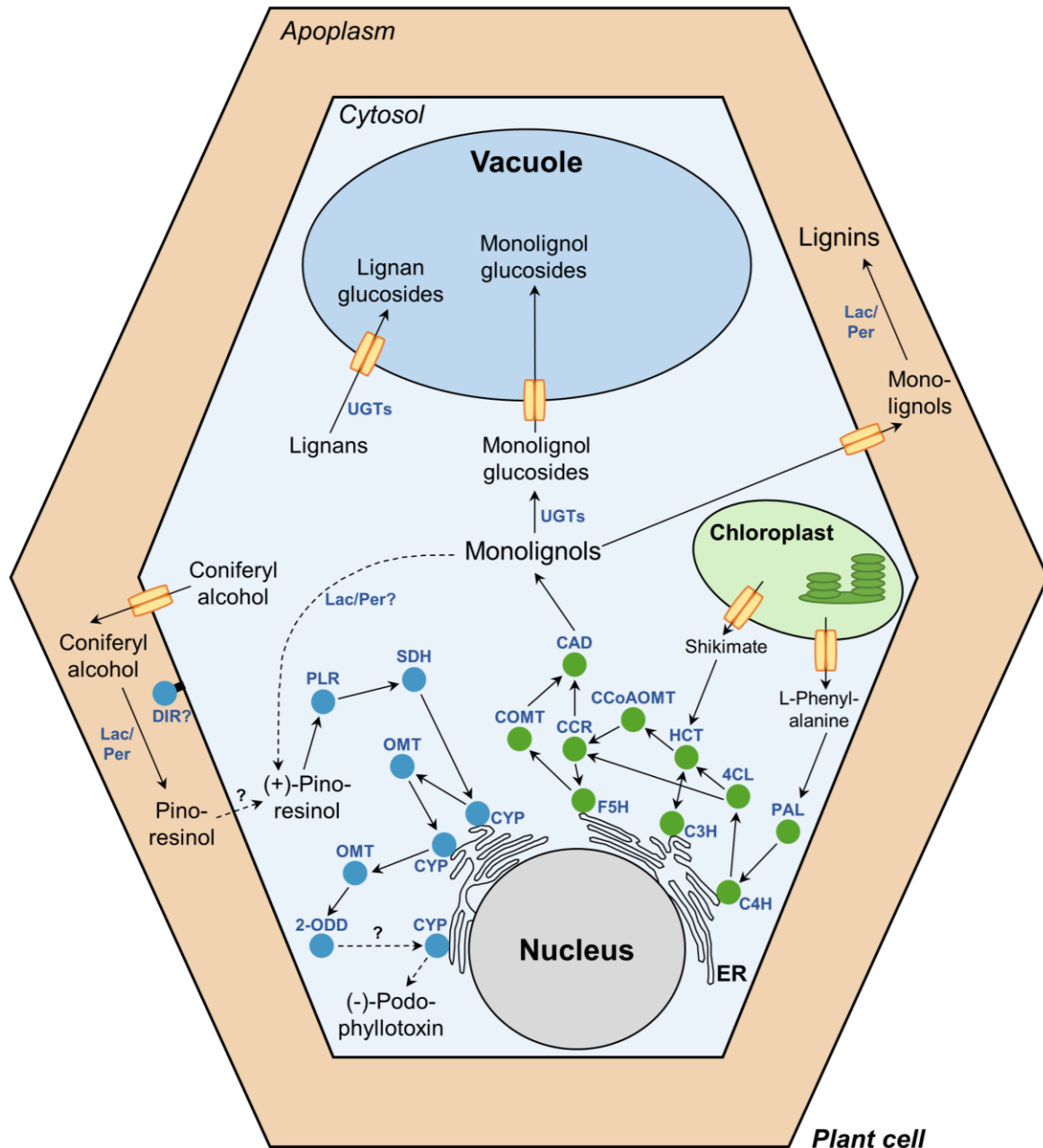
The main natural source of this pharmaceutically interesting compound are the roots and rhizomes of *Podophyllum hexandrum*, a Himalayan herb now endangered due to over-collection and environmental disruption [9,27]. Extraction from the dried plant material using warm ethanol is applied in industry despite the generally low (–)-podophyllotoxin content (4–5 % of dry weight) from the source material [28–30]. So far, neither alternative plant

resources (0.47 % of dry weight from *Juniperus virginiana*) and optimization of extraction methods nor chemical synthesis or plant cell suspension cultures (0.65 % of dry weight) were found to be economically competitive compared to extraction from plants [31–35]. Although many routes of chemical synthesis have been proposed, organic synthesis remains challenging due to the existence of four chiral centers and a base-sensitive trans- $\gamma$ -lactonic ring in (–)-podophyllotoxin [23,24]. Thus, biotechnological production of lignans in a heterologous host such as *Escherichia coli* (*E. coli*) constitutes a promising alternative. In view of a future microbial production from inexpensive resources such as glucose, an important prerequisite is the stable and efficient supply of the lignan precursor coniferyl alcohol.

## I.2 Lignin and the phenylpropanoid pathway

Lignin is part of the secondary plant cell wall apart from cellulose and hemicellulose and represents the second most abundant biopolymer on earth [36]. This large aromatic heteropolymer essentially consists of *p*-hydroxyphenyl, guaiacyl and syringyl units formed from the respective monolignols *p*-coumaryl, coniferyl and sinapyl alcohol [36]. Lignin enables plants to grow vertically and transport water and nutrients by providing strength and rigidity to the cell wall and hydrophobicity to the vascular bundles [37]. Furthermore, lignin seems to play a role in plant defense, because biotic (pathogens such as fungi, bacteria and viruses) and abiotic stress factors (drought, ultraviolet radiation, temperature, mineral deficiency) were found to correlate with several changes in lignin-associated gene expression and in lignin composition [38]. It is generally accepted that the monolignols are transported via the plasma membrane into the apoplast (Figure 1.2). There, the radical polymerization by peroxidases or laccases takes place generating cross-coupled monolignols with different kinds of ether bonds and resulting in the lignin heteropolymer [36,39]. As a form of storage, monolignols as well as lignans can be found as glucosides in the plant vacuoles [40].

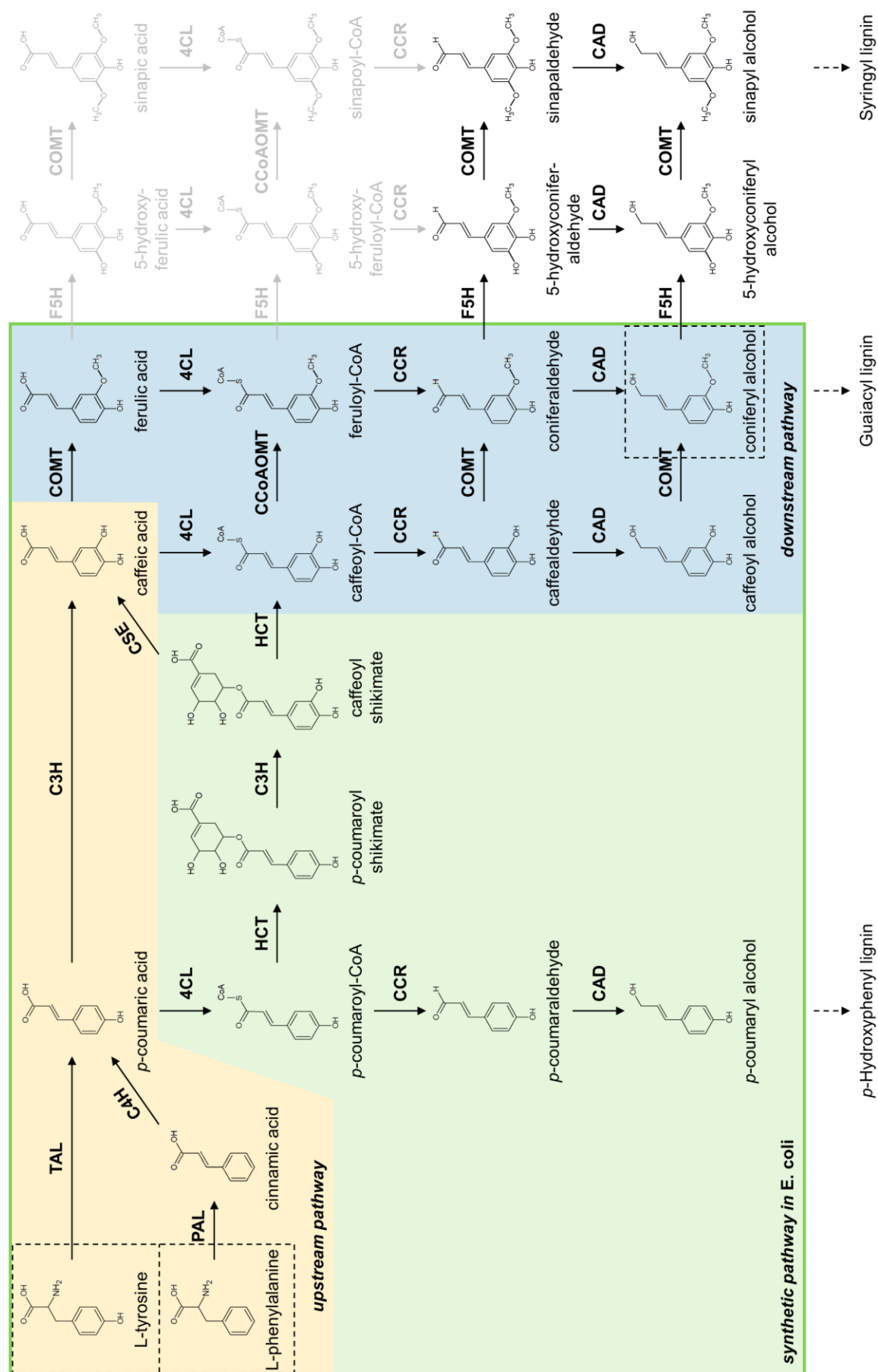
Apart from lignin and lignans, other secondary metabolites are produced via enzymes from the monolignol synthesis pathway as well. Such secondary metabolites are e.g. (iso)flavonoids, curcuminoids, stilbenes, coumarins and cinnamoyl anthranilates serving as pigments, antioxidants, UV protectants, signaling agents, antiherbivory and aroma compounds in plants [41–45].



**Figure 1.2: Schematic view of monolignol and lignan synthesis pathway and the proposed localization of their enzymes in plant cells.** The dots represent enzymes from the monolignol (green filled circles) and apparent lignan synthesis pathway of (-)-podophyllotoxin (blue filled circles). The arrows show known transport and reaction routes (full lines) as well as unidentified routes (dashed lines). The transport of monolignols across the vacuolar and cellular membrane is suspected to occur mainly via ABC transporters (orange pores), while the other transport processes remain less clear. Lac/Per– laccase or peroxidase, UGTs– UDP(uridine diphosphate)-glycosyltransferases, DIR– dirigent protein, ER– endoplasmic reticulum, PAL– phenylalanine ammonia lyase, C4H– cinnamate 4-hydroxylase, 4CL– 4-coumarate:CoA ligase, HCT– *p*-hydroxycinnamoyl-CoA:shikimate *p*-hydroxycinnamoyl-transferase, C3H– *p*-coumaroyl shikimate 3'-hydroxylase, CCR– cinnamoyl-CoA reductase, CAD–cinnamyl alcohol dehydrogenase, CCoAOMT– caffeoyl-CoA *O*-methyltransferase, COMT– caffeic acid *O*-methyltransferase, F5H– ferulate 5-hydroxylase, PLR– pinoresinol-lariciresinol reductase, SDH– secoisolariciresinol dehydrogenase, CYP– specific cytochrome P450 enzyme or hydroxylase, OMT– specific *O*-methyltransferase, 2-ODD– 2-oxoglutarate/Fe(II)-dependent dioxygenase. Adapted from Wang *et al.* (2013) and Seegers *et al.* (2017) [35,40].

In plants, the synthesis of the monolignols occurs in the phenylpropanoid pathway starting from L-phenylalanine (Figure 1.3). First, via the action of a phenylalanine ammonia lyase (PAL), cinnamic acid is formed and further hydroxylated at the C4-position into *p*-coumaric acid by a cytochrome P450 monooxygenase (C4H). Subsequently, the intermediate is activated by the 4-coumarate:CoA ligase (4CL) with the simultaneous cleavage of adenosine triphosphate (ATP) into *p*-coumaroyl-CoA. The CoA thioester is further reduced by a cinnamoyl-CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) using nicotinamide adenine dinucleotide phosphate (NADPH) as cofactor for both enzymes. The resulting *p*-coumaryl alcohol constitutes the simplest monolignol. In contrast, coniferyl and sinapyl alcohol differ in their degree of methoxylation on the phenyl ring. The hydroxylation at the C3-position has been accepted to occur on the level of quinate or shikimate esters catalyzed by another cytochrome P450 monooxygenase (C3H) for many years, until the recent discovery of a direct hydroxylation from *p*-coumaric acid to caffeic acid by a cytosolic ascorbate peroxidase in monocots [46–48]. For methylations in the C3- and C5-position of the phenyl ring, two different types of methyltransferases COMT (caffeic acid *O*-methyltransferase) and CCoAOMT (caffeoyl-CoA *O*-methyltransferase) are responsible for yielding coniferyl and sinapyl alcohol. Lastly, the ferulate 5-hydroxylase (F5H) introduces a hydroxyl group at the C5-position (Figure 1.3). In several studies, the F5H enzyme exhibited superior kinetic properties towards coniferaldehyde and coniferyl alcohol compared to ferulic acid [49,50]. Also due to the lack of substrate specificity of F5H towards feruloyl-CoA, the phenylpropanoid pathway is generally acknowledged to proceed via the aldehyde and alcohol intermediate regarding the F5H catalyzed reaction [51]. In the end, the phenylpropanoid pathway adopts an overall metabolic grid-like structure with several possible reaction routes towards the formation of the monolignols [52].

Usually, the plant cytochrome P450 enzymes (C4H, C3H, F5H) are localized to the ER membrane and assumed to be associated with the soluble PAL, 4CL, CCR, CAD, COMT and CCoAOMT enzymes enabling metabolic channeling *in planta* (Figure 1.2) [53]. Thus, the production of plant cytochrome P450 monooxygenases in a prokaryotic host such as *E. coli* remains challenging due to the hydrophobic N-terminal transmembrane domain and the lack of an endoplasmic reticulum. Moreover, some specialized enzymes from bacteria and fungi, so-called tyrosine ammonia lyases (TAL), are able to directly deaminate L-tyrosine into *p*-coumaric acid despite the absence of the phenylpropanoid pathway in these organisms (Figure 1.3). These TAL enzymes together with alternative, soluble C3H candidates could lead to a simplified synthetic phenylpropanoid pathway in *E. coli*.





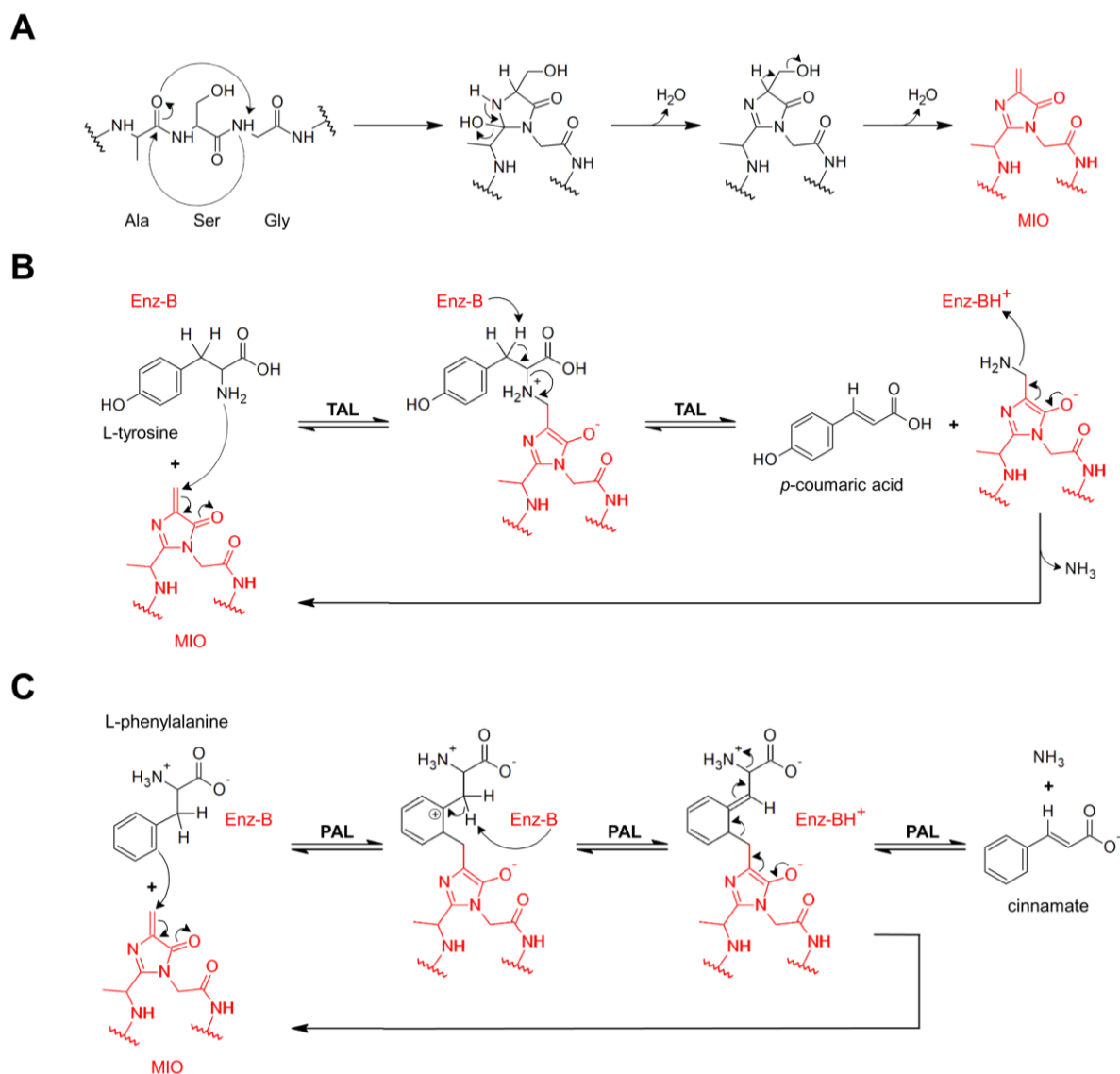
**Figure 1.3: Phenylpropanoid pathway.** The synthetic pathway investigated in *E. coli* in this thesis is encircled by a green box and further divided into upstream (light yellow) and downstream pathway (light blue). Chemical structures surrounded by dashed lines indicate the respective starting substances and the final product of the herein investigated synthetic pathway in *E. coli*. Substances and reaction arrows in grey represent theoretical phenylpropanoid intermediates that are supposed to not occur *in planta* nowadays. TAL– tyrosine ammonia lyase, PAL– phenylalanine ammonia lyase, C4H– cinnamate 4-hydroxylase, C3H– *p*-coumarate 3-hydroxylase/*p*-coumaroyl shikimate 3'-hydroxylase, 4CL– 4-coumarate:CoA ligase, CCR– cinnamoyl-CoA reductase, CAD– cinnamyl alcohol dehydrogenase, HCT– *p*-hydroxycinnamoyl-CoA:shikimate *p*-hydroxycinnamoyl-transferase, CSE– caffeoyl shikimate esterase, COMT– caffeic acid *O*-methyltransferase, CCoAOMT– caffeoyl-CoA *O*-methyltransferase, F5H– ferulate 5-hydroxylase. Adapted from Vanholme *et al.* (2019) [47].

### I.3 Tyrosine and phenylalanine ammonia lyases

Aromatic amino acid deaminases constitute the enzyme family initiating the phenylpropanoid pathway and thus connecting primary with secondary metabolism in plants [53,54]. PAL enzymes are ubiquitous in the plant kingdom and catalyze the specific and non-oxidative deamination of L-phenylalanine to cinnamic acid [55]. The enzyme family of aromatic amino acid deaminases also exists in bacteria as TAL enzymes and as histidine ammonia lyases (HAL) with the latter being absent from plants [53]. Additionally, the PALs from some monocots and yeasts were characterized as bifunctional enzymes, which not only accept L-phenylalanine but also L-tyrosine as substrate and thus exhibit a simultaneous TAL activity [56–58].

Regarding the biological function of TAL in the bacterium *Rhodobacter capsulatus*, the enzyme is thought to be involved in the synthesis of photoactive yellow protein, which is assumed to act as light sensor and contains *p*-coumaric acid as chromophore [59]. Further PAL and TAL functions in bacteria include the production of the antibiotics saccharomicin and stilbene, the bacteriostatic agent enterocin and supposedly the antifungicide soraphen A [55,60–62]. Similar to HALs in bacteria and animals serving as degradative enzymes for the conversion of histidine, PALs from yeasts play a role in the fungal catabolism of L-phenylalanine and L-tyrosine for the supply of carbon and nitrogen [59,60,63,64].

With the aim to characterize and compare substrate specificities of an ammonia lyase enzyme, the PAL/TAL ratio is usually calculated as the quotient of specific activities or catalytic efficiencies from the individual enzyme reactions towards L-tyrosine and L-phenylalanine as substrates [58,65]. In general, Barros and Dixon (2020) assigned a PAL/TAL ratio lower than 0.01 to bacterial TAL enzymes and a ratio between 0.2 and 6 for bifunctional PAL enzymes from monocots and yeast origin [53]. PALs from dicotyledonous plants such as *Arabidopsis thaliana* and *Petroselinum crispum* were found to have a PAL/TAL ratio in the magnitude of hundreds to several thousands [66–69].



**Figure 1.4: Formation of the catalytic MIO group (A) and proposed reaction mechanisms for tyrosine ammonia lyase (B) and phenylalanine ammonia lyase (C).** Reacting amino acids and groups of the TAL and PAL enzyme are indicated in red. The MIO (3,5-dihydro-5-methylidene-4H-imidazol-4-one) group is formed spontaneously from a motif consisting of alanine (Ala), serine (Ser) and glycine (Gly) (A). Adapted from Tan *et al.* (2020) and Calabrese *et al.* (2004) [70,71].

In general, PAL enzymes are homotetrameric proteins with each subunit containing an active site with the cofactor 3,5-dihydro-5-methylidene-4H-imidazol-4-one (MIO), which is formed posttranslationally by autocatalytic cyclization and dehydration from the active site tripeptide Ala-Ser-Gly (Figure 1.4 A) [72–74]. For the biocatalytic deamination, two reaction mechanism have been proposed for ammonia lyases. Several studies discussed and indicated that TAL enzymes catalyze the conversion of L-tyrosine into *p*-coumaric acid via an E<sub>1</sub>cB- or E<sub>2</sub>-elimination reaction, while the deamination of L-phenylalanine by PAL was reported to occur via a Friedel-Crafts-type reaction (Figure 1.4 B and C) [53,75]. In an E<sub>1</sub>cB-elimination reaction, the amino group of the substrate nucleophilically attacks the potent electrophilic MIO group (Figure 1.4 B). Then, a general base abstracts the  $\beta$ -proton and generates a carboanion

which collapses quickly and leads to the elimination of the  $\alpha$ -amino group and formation of a double bond in the product [67]. Lastly, the MIO-bound amine is released as ammonium after its protonation [70]. If a strong base is present in the active site, the elimination can also proceed as a one-step simultaneous E<sub>2</sub>-type mechanism [53].

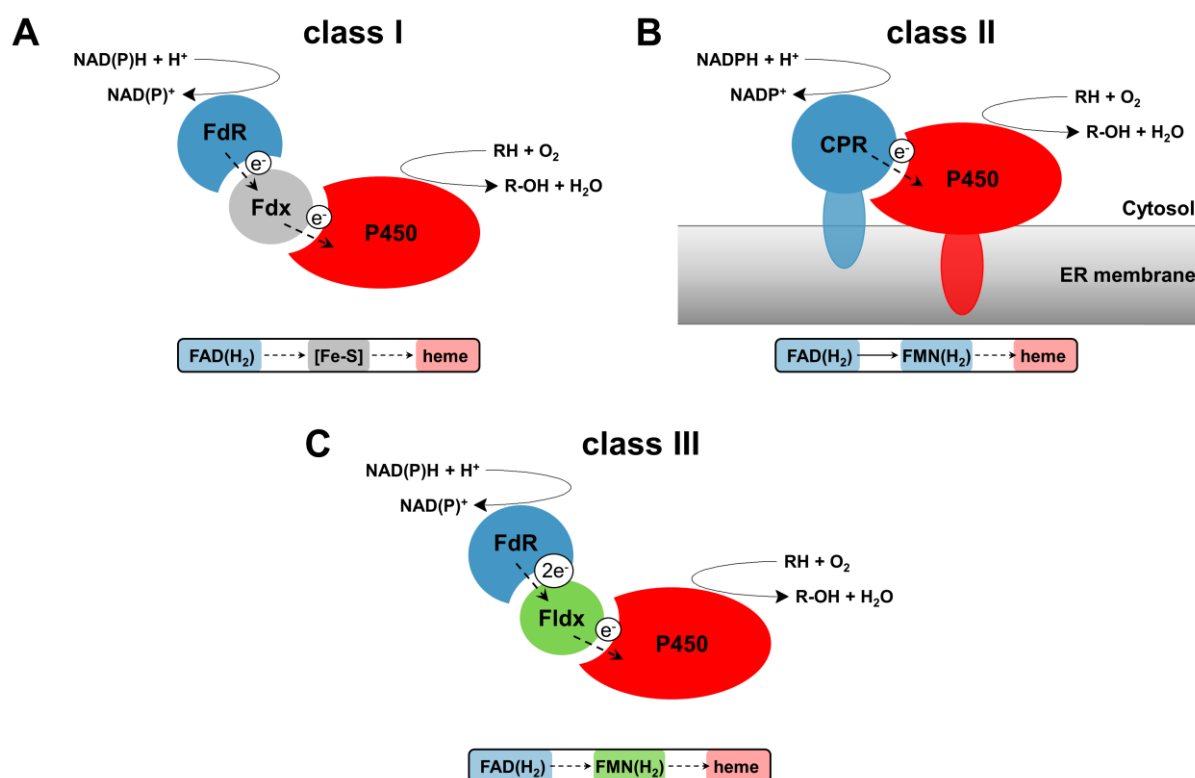
In the alternative Friedel-Crafts mechanism, the MIO group performs an electrophilic attack on the aromatic ring of the substrate (Figure 1.4 C). The generated delocalized carbocation increases the acidity of the  $\beta$ -proton, which is thus abstracted by a general base [67,76]. After electronic rearrangement, the amino group is eliminated and the  $\alpha,\beta$ -unsaturated product formed.

Eukaryotic PAL enzymes are larger proteins containing a shielding domain over the active site and an N-terminal extension suspected to function as anchorage to other cell components *in planta* [75,77]. By contrast, HALs and prokaryotic TALs lack these characteristics [74]. All PAL and TAL enzymes are known to have two flexible loops near the active site, which adapt different conformations among species [74,75]. The inner lid-loop caps the active site and is reported to be displaced upon substrate entrance [77,78]. In TAL from the bacterium *Rhodobacter sphaeroides* (RsTAL), the outer lid-loop seems to press down on the inner lid-loop. This leads to a shift of a tyrosine residue closer to the substrate probably serving as a second proton acceptor in an E<sub>2</sub> reaction mechanism [74,75]. This outer-lid loop was found to be shorter in PALs from the dicot *Petroselinum crispum* and the monocot *Sorghum bicolor* which questions its function in eukaryotic PALs [75,77].

A few reports aimed at elucidating which amino acid residues determine the substrate specificity of PAL and TAL enzymes. As a prominent example, the histidine residue His89 in RsTAL was assumed to form a hydrogen bond to the *para*-hydroxyl group of L-tyrosine. When His89 was mutated to a phenylalanine residue, the deamination activity towards L-tyrosine was lost and the RsTAL enzyme became a PAL enzyme instead [67,74]. An equivalent substrate selectivity switch was observed for PAL from *A. thaliana* (AtPAL) and *Sorghum bicolor* (SbPAL), when the corresponding phenylalanine residue was changed to a histidine in AtPAL or vice versa in SbPAL [67,75]. Further, Bartsch and Bornscheuer (2009) mutated a glutamate residue in PAL from *P. crispum*, which they suspected to prevent the MIO cofactor from an attack on the amino group [68]. This yielded an around eight times lower PAL/TAL ratio of the mutant compared to the wild-type enzyme [68]. Consequently, the substrate specificity of individual PAL and TAL enzymes seems to be determined by other residues and factors as well [75,79].

## I.4 Cytochrome P450 monooxygenases (C3H)

Cytochrome P450 monooxygenases (P450 enzymes) form a large superfamily of heme *b*-containing enzymes distributed in all three kingdoms of life [80]. P450 enzymes are generally involved in drug metabolism, the degradation of xenobiotics, the biosynthesis of steroid hormones and diverse plant secondary metabolites such as lignins and lignans [81]. The enzyme superfamily is well-known for the regio- and stereospecific oxidation of non-activated C-H bonds under mild reaction conditions, which remains difficult to perform using chemical synthesis [80]. Cytochrome P450 monooxygenases are also considered as the most versatile biocatalysts in nature, as they catalyze a multitude of other reactions including epoxidation, dealkylation, C-C bond cleavage, deamination, decarboxylation, N-oxidation and several more [82,83].



**Figure 1.5: Schematic organization of different classes of cytochrome P450 monooxygenases with their redox partners.** The transfer of electrons between the protein components from reducing equivalents (NADH or NADPH) to the terminal substrate (RH) is indicated. Each lower box shows the involved cofactors and prosthetic groups. FdR– ferredoxin/ flavodoxin reductase, Fdx– ferredoxin, CPR– cytochrome P450 reductase, Fldx– flavodoxin. Adapted from Hannemann *et al.* (2007) and Girhard *et al.* (2010) [84,85].

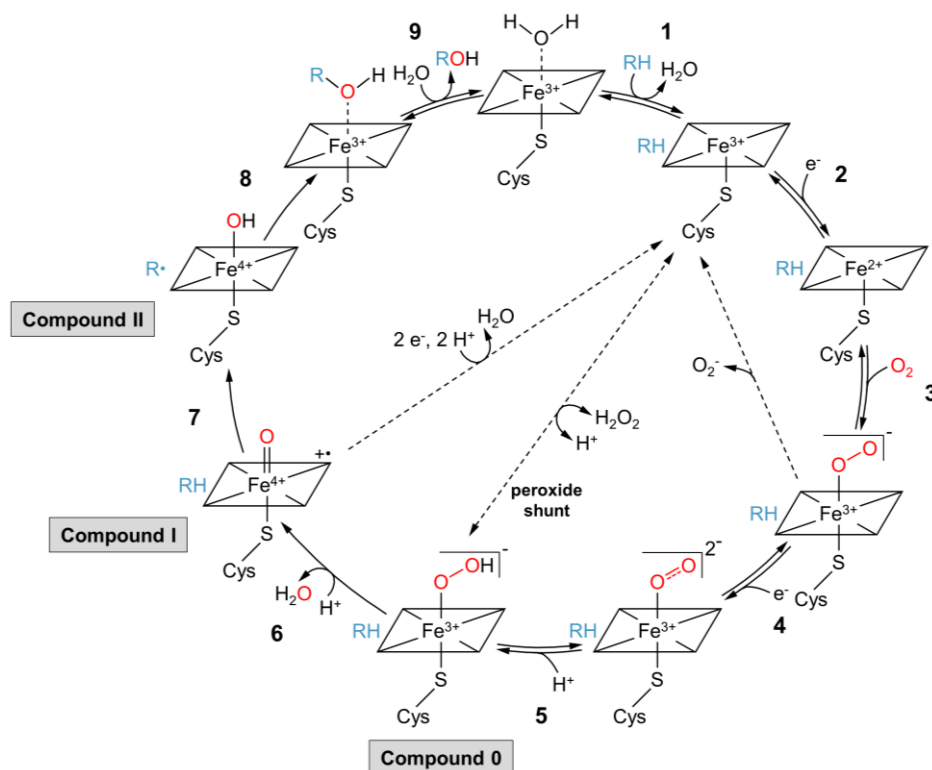
Most P450 enzymes are external monooxygenases, which implicates that they receive electrons from the reducing equivalents NADH or NADPH and via the action of an appropriate reductase. With respect to the composition of the electron transport chain from the reducing equivalent to the P450 heme center, cytochrome P450 monooxygenases are categorized into different classes

(Figure 1.5). P450 enzymes of class I consist of an NAD(P)H-dependent flavin adenine dinucleotide (FAD)-containing ferredoxin reductase and an iron-sulfur [2Fe-2S]-ferredoxin serving as electron shuttle protein (Figure 1.5 A). While most bacterial cytochrome P450 monooxygenases belong to class I and are soluble three-protein systems, some class I mitochondrial P450 enzymes exhibit membrane association of the reductase and the P450 protein [84]. However, class II contains the majority of eukaryotic P450 enzymes (Figure 1.5 B). Apart from the P450 protein, a single NADPH-dependent cytochrome P450 reductase (CPR), containing FAD and flavin mononucleotide (FMN), is responsible for the electron abstraction and transfer. Furthermore, both P450 system components are N-terminally integrated into the membrane of the endoplasmic reticulum. In accordance with the classification by Hannemann and co-workers (2007), an additional class III was characterized, which comprises a soluble three-protein system with an FAD-containing flavodoxin reductase and an FMN-containing flavodoxin (Figure 1.5 C) [84]. Another difference between P450 enzymes of class I and class III constitutes the strict one-electron transfer to ferredoxins, while flavodoxins can accept up to two electrons and thus adopt different redox states [86].

The catalytic center of cytochrome P450 monooxygenases is formed by heme *b* as prosthetic group. Heme *b* is a porphyrin ring with a coordinated iron ion and is coupled to the apoprotein via a conserved, proximal cysteine residue [80]. In the resting catalytic state, a water molecule is coordinated to the iron ion as distal sixth ligand, which is displaced upon substrate binding in the active site cavity (Figure 1.6) [87]. In the ensuing reaction cycle, the ferric, high-spin iron ( $\text{Fe}^{3+}$ ) is reduced by one electron transferred from the redox partner, molecular oxygen binds to the ferrous ion ( $\text{Fe}^{2+}$ ) and the complex is reduced by the second electron forming a ferric-peroxo intermediate [88]. After a single protonation step forming a ferric-hydroperoxo intermediary complex known as Compound 0, a second proton leads to the cleavage of the O-O bond resulting in the release of a water molecule and the generation of the highly reactive ferryl-oxo species referred to as Compound I [88,89]. In subsequent rapid, irreversible steps, a hydrogen atom is abstracted from the substrate forming a ferryl-hydroxo complex (Compound II). The complex-bound hydroxyl group recombines with the substrate radical in accordance with a rebound mechanism [87,90]. At last, the hydroxylated product is released from the active site and the heme center returns to its resting state by binding a water molecule.

Some cytochrome P450 monooxygenases also known as peroxygenases are capable of utilizing hydrogen peroxide to directly form Compound 0 and this reaction is thus referred to as the peroxide shunt pathway (Figure 1.6) [88]. Uncoupling reactions are another characteristic of

P450 enzymes. They result from the non-productive decay of cycle intermediates and lead to the release of water or the generation of reactive oxygen species such as superoxide anion or hydrogen peroxide [91].



**Figure 1.6: The catalytic cycle of P450 enzymes.** Reactions on the schematic heme *b* prosthetic group are shown in nine steps (full arrows) and the peroxide shunt pathway as well as uncoupling reactions (dashed arrows) are indicated. After substrate binding (1) and reduction of the iron ion by one electron (2), dioxygen can bind to the heme center (3). The second transferred electron (4) leads to a ferric-peroxo complex, which is further protonated into Compound 0 (5). The highly reactive Compound I is the result of the addition of a second proton leading to the release of a water molecule from the complex (6). Subsequently, a hydrogen atom is abstracted (7) forming Compound II and the substrate radical reacts with the complexed hydroxyl group in an oxygen rebound mechanism (8). Lastly, the resting state is restored by removal of the oxygenated product and binding of a water molecule (9). Adapted from Li *et al.* (2020), Munro *et al.* (2007) and Guengerich (2018) [88,91,92].

Regarding a synthetic phenylpropanoid pathway, one chosen enzyme candidate for the C3-hydroxylation is a cytochrome P450 monooxygenase.

#### I.4.1 CYP199A2 from *Rhodospseudomonas palustris*

For a long time, the enzyme catalyzing the natural hydroxylation in the C3-position of phenylpropanoids remained unidentified in plants, until Schoch and co-workers (2001) characterized the first possible C3H from *A. thaliana*. Enzyme *in vitro* investigations revealed that neither free *p*-coumaric acid nor *p*-coumaroyl-CoA served as substrates, but the respective shikimate and quinate esters were shown to be converted (cf. Figure 1.3) [46]. As all plant cytochrome P450 monooxygenases belong to class II and contain an N-terminal transmembrane domain, their functional expression in a bacterial host is challenging. In view of these

expression obstacles and reduction of necessary pathway enzymes, a bacterial C3H catalyzing the direct conversion of *p*-coumaric acid seems more advantageous.

The cytochrome P450 monooxygenase CYP199A2 from the soil and water bacterium *Rhodopseudomonas palustris* was first characterized by Bell and co-workers (2006) for the conversion of aromatic carboxylic acids [93]. Later, Furuya *et al.* (2012) generated mutants of the position 185, which equals a phenylalanine residue in the wild-type enzyme and is located above the heme center [94]. In an *E. coli* whole-cell biocatalysis, the mutant CYP199A2 F185L exhibited a 5.5 times higher activity towards the conversion of *p*-coumaric acid into caffeic acid than the wild-type and was further able to convert cinnamic acid directly into caffeic acid [94]. With a substrate feed of 20 mM *p*-coumaric acid and in the presence of glycerol, the CYP199A2 F185L mutant (from now on referred to as CYP199A2) produced 15 mM caffeic acid in resting *E. coli* whole-cells [94].

The bifunctionality as C3H and C4H of the soluble CYP199A2 makes this enzyme candidate attractive for the application in the synthetic phenylpropanoid pathway in *E. coli* starting from either L-tyrosine or L-phenylalanine. As a class I P450 enzyme, the physiological redox partners palustrisredoxin (Pux) and the NADH-dependent palustrisredoxin reductase (PuR) will be used.

Furthermore, alternative redox partners are worth evaluation, such as the redox partner pair putidaredoxin (Pdx) and the NADH-dependent putidaredoxin reductase (PdR), which is often used for the investigation of orphaned class I P450 enzymes [95]. In addition, the class III redox pair consisting of the flavodoxin YkuN from *Bacillus subtilis* and the NADPH-dependent flavodoxin reductase from *E. coli* (FpR) has been successfully applied as surrogate redox pair for the activity of certain P450 enzymes [85,96].

#### **1.4.2 C3H alternative: HpaBC from *Escherichia coli***

In the genome of *E. coli*, a non-cytochrome P450 hydroxylase complex has been discovered several years ago originally being assigned a function in the degradation of 4-hydroxy-phenylacetate (4-HPA) [97]. The hydroxylase complex consists of a larger FAD-containing monooxygenase component HpaB and a smaller flavin:NADH oxidoreductase HpaC [98,99]. The HpaBC complex is described to be not natively expressed in *E. coli* and thus heterologous overexpression is necessary [100]. Apart from pointing out its general broad substrate specificity, Lin and Yan (2012) showed that the HpaBC complex also hydroxylated *p*-coumaric acid at the C3-position and exhibited a two-fold higher specific activity for *p*-coumaric acid over L-tyrosine as substrate [101]. Using the HpaBC complex in growing *E. coli* cells, almost

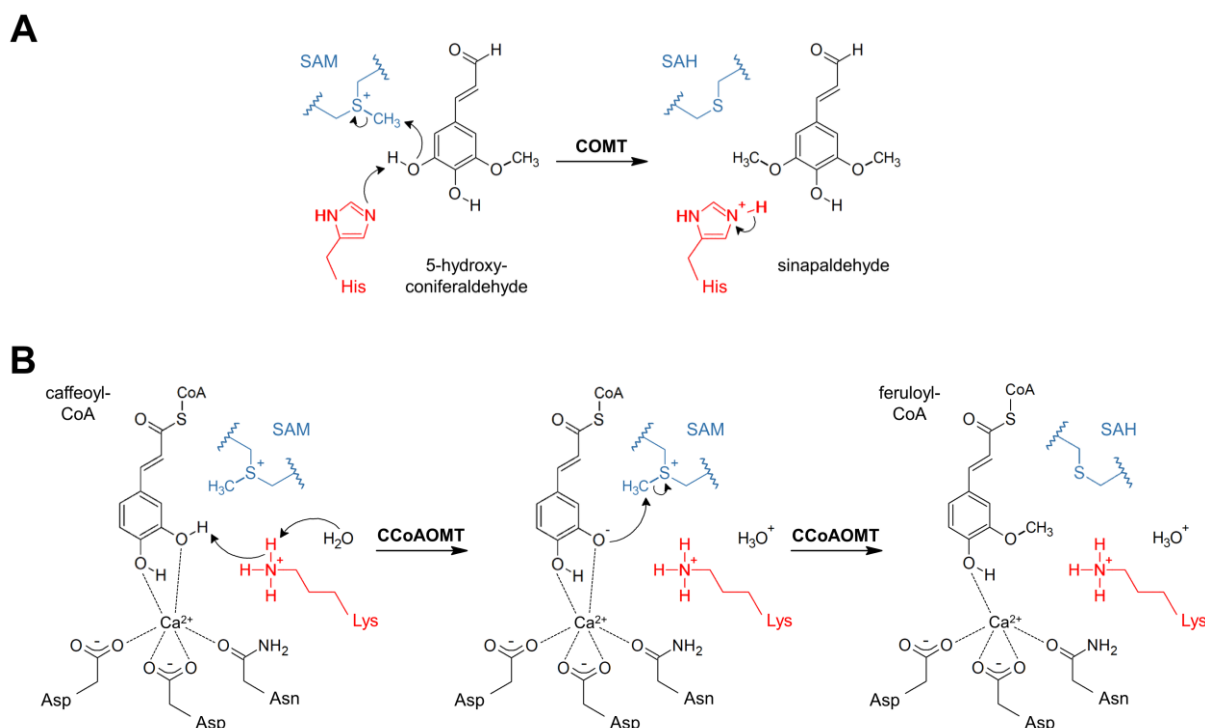
complete conversion of 21.3 mM *p*-coumaric acid into caffeic acid was achieved [102]. So far and to the best of the author's knowledge, a final titer of 56.5 mM caffeic acid constitutes the highest direct production from *p*-coumaric acid in a whole-cell biotransformation with the HpaBC complex from *Pseudomonas aeruginosa* [103].

## **I.5    *O*-methyltransferases (OMTs)**

The *O*-methyltransferases (OMTs) involved in the plant phenylpropanoid pathway catalyze the methylation of the hydroxyl groups at the 3- and 5-positions of the phenolic ring leading to guaiacyl and syringyl lignin monomers (cf. Figure 1.3) [104]. In general, two different classes as well as structural types of plant methyltransferases have been characterized: the caffeic acid *O*-methyltransferase (COMT) and the caffeoyl-CoA *O*-methyltransferase (CCoAOMT) [105]. The COMT enzymes belong to type 1 *O*-methyltransferases and are generally larger proteins with molecular weights ranging from 38 to 43 kDa compared to the smaller type 2 CCoAOMT enzymes (23-29 kDa) [106,107]. COMT enzymes are reported to have an N-terminal dimerization domain and a C-terminal catalytic domain involved in S-adenosyl-L-methionine (SAM) binding and phenylpropanoid binding [108]. On the other hand, the CCoAOMT proteins consist of two tightly associated subunits and are similar in structure and function to the mammalian catechol OMTs, which serve in the metabolism of catecholamine neurotransmitters in animals [105,109,110]. Like catechol OMTs, the activity of CCoAOMT is dependent on divalent cations like magnesium, manganese or calcium ions which are not necessary for the catalytic mechanism of COMT enzymes [105,111].

According to a structural analysis of COMT from *Sorghum bicolor*, a histidine residue was identified in the active site, which functions as a general base by abstracting a proton from the hydroxyl group of the substrate (Figure 1.7 A) [112]. Simultaneously, the oxygen atom nucleophilically attacks the methyl group from the cofactor SAM generating the methoxylated product. The abstracted proton bound to the imidazolium ring of histidine is removed by a bound water molecule afterwards [112]. In contrast, the catalytic residue in the CCoAOMT from the identical plant origin is a lysine residue which also deprotonates the hydroxyl group of the substrate (Figure 1.7 B) [105]. The substrate caffeoyl-CoA is thereby coordinated in a bidentate manner to a calcium ion in the active site and the resulting oxyanion stabilized prior to its nucleophilic attack on the methyl group from SAM. The loss in coordinating bonds from the product feruloyl-CoA towards the metal ion drives the release of the product from the enzyme [105].





**Figure 1.7: Proposed reaction mechanisms of COMT (A) and CCoAOMT (B) from *Sorghum bicolor*.** The reacting amino acid of the OMT enzyme being a histidine (His, **A**) or lysine residue (Lys, **B**) is indicated in red and the methyl group donor SAM/SAH (S-adenosyl-L-methionine/S-adenosyl-L-homocysteine) as cofactor in blue. In the CCoAOMT enzyme, the  $\text{Ca}^{2+}$  ion is coordinated by two aspartate (Asp) and one asparagine residue (Asn) (**B**). Adapted from and Green *et al.* (2014) and Walker *et al.* (2016) [105,112].

Naaz and co-workers (2013) performed a sequence analysis of different plant COMTs and CCoAOMTs and found that many conserved amino acid residues are associated with the binding of the SAM cofactor. A Rossmann-like fold, which constitutes a common structure from nucleotide binding proteins, is thereby involved in mediating the binding of SAM [113,114]. It has been suggested that CCoAOMTs act as monomers in solution due to the absence of N-terminal dimerization domain unlike in COMTs, where N-terminal residues participate in the formation of the other monomer's substrate binding cavity [108,109,114]. Though, crystal structures of CCoAOMT from alfalfa (*Medica sativa*) and *Sorghum bicolor* indicated an association as homodimer as well [105,114].

The substrate specificity represents another difference between both OMT families. COMT enzymes are reported to accept a broad range of substrates, while CCoAOMTs are more specific for *p*-hydroxycinnamoyl-CoA esters, in particular caffeoyl-CoA over 5-hydroxyferuloyl-CoA as substrate [104,105]. In contrast to the catechol OMTs and COMTs, CCoAOMT enzymes have an insertion loop in their active site conferring substrate specificity for CoA esters and its modification showed to alter substrate preference as well as catalytic efficiency [115]. Further, the N-terminus of CCoAOMT was found to be an essential part of the substrate-binding pocket [109,114]. Many COMT enzymes exhibited a preference for phenylpropanoid substrates

following the order of aldehyde over alcohol and both over acids *in vitro* [108,112]. A high catalytic efficiency for COMT enzymes from *A. thaliana*, aspen (*Populus tremuloides*), wheat (*Triticum aestivum*) and rice (*Oryza sativa*) was reported for 5-hydroxyconiferaldehyde, which was also shown to competitively inhibit the methylation of other substrates like 5-hydroxyferulic acid and 5-hydroxyconiferyl alcohol [116–119].

In addition to that, a downregulation of COMT *in planta* led to a lignin composition almost absent of syringyl units and only to a small or no reduction in total lignin content (cf. Figure 1.3) [120–122]. Plants with downregulated CCoAOMT had a reduced total lignin content with less total G-monomeric units and unchanged S-unit content [120,123]. Furthermore, an OMT double-knockout *A. thaliana* plant stopped growth in the plantlet stage where its lignin consisted mainly of *p*-hydroxyphenyl units [124]. Taking together the *in vitro* and *in vivo* findings, 5-hydroxyconiferaldehyde is supposed to be the physiological substrate of plant COMTs [117]. At the same time, the CCoAOMT enzymes primarily act as 3-*O*-methyltransferases, while COMTs are suspected to methylate both 3- and 5-hydroxyl positions *in vivo* (cf. Figure 1.3) [117].

## **I.6 Current microbial production of phenylpropanoids**

The COMT from *A. thaliana* (*AtCOMT*) was shown to accept substrates such as N-acetylserotonin, the stilbenoid resveratrol and the flavonoids quercetin, myricetin and luteolin, thus demonstrating broad substrate specificity [117,125–127]. Furthermore, *AtCOMT* has been successfully implemented in the microbial production of ferulic acid and vanillin from L-tyrosine or *de novo* in *E. coli* [128–130]. Also, a total concentration of 1.33 mM in the supernatant constitutes the highest ferulic acid titer to date in growing *E. coli* cells by using *AtCOMT* with TAL from *Rhodotorula glutinis* (*RgTAL*), a bacterial C3H and a substrate supplementation of 3 mM L-tyrosine [131].

Rodrigues and co-workers (2015) have also previously shown that CYP199A2 with *RgTAL* can produce 1.56 mM caffeic acid from 3 mM L-tyrosine in a growing cell assay [132]. Using the HpaBC complex from *E. coli*, the currently highest *de novo* production of caffeic acid (5.7 mM) was achieved in combination with *RgTAL* in a growing L-tyrosine overproducing *E. coli* strain and under optimized fermentation conditions [133].

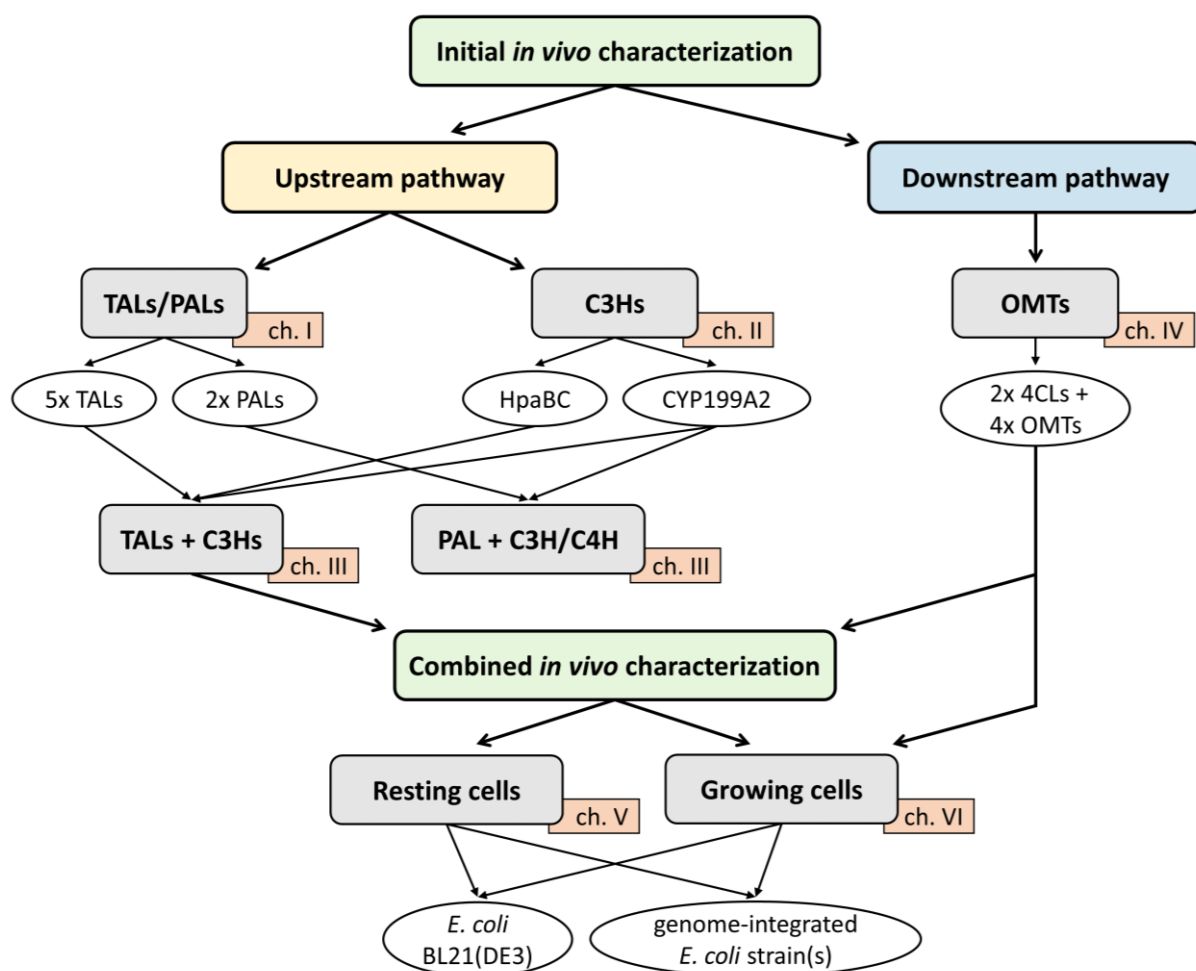
As first report about an active plant C4H enzyme in *E. coli*, a fusion protein of C4H from *Lycoris aurea* and ATR2 as redox partner was used with PAL1 from *A. thaliana* for the *de novo*

generation of 156  $\mu\text{M}$  *p*-coumaric acid from L-phenylalanine after optimization of the intracellular NADPH pool [134].

Regarding the microbial production of monolignols, Jansen *et al.* (2014) yielded 146  $\mu\text{M}$  *p*-coumaryl alcohol in growing *E. coli* cells, which heterologously expressed *RsTAL*, 4CL1 from *Petroselinum crispum* (*Pc4CL*) as well as CCR1 and CAD enzymes from *Zea mays* (*ZmCCR* and *ZmCAD*) [135]. Later, the titer has been increased to 346  $\mu\text{M}$  *p*-coumaryl alcohol by engineering the spacing residues between ribosome binding site (RBS) and start codon on a tetracistronic plasmid [136]. In 2017, Chen and co-workers reported the application of *RgTAL* with 4CL1 from *A. thaliana* and alternative enzyme candidates of CCR and CAD for the synthesis of 3.3 mM *p*-coumaryl alcohol *de novo* in an optimized and L-tyrosine overproducing *E. coli* strain [137]. The further addition of HpaBC from *E. coli* and COMT or CCoAOMT from *A. thaliana* resulted in the production of caffeoyl and coniferyl alcohol with maximal titers of 5.1 mM and 693  $\mu\text{M}$ , respectively [137]. So far, a higher coniferyl alcohol titer (1.1 mM) has only been achieved in *Saccharomyces cerevisiae* with TAL from *Herpetosiphon aurantiacus* (*HaTAL*), an HpaBC complex of mixed components, and with *AtCOMT* as well as 4CL, CCR and CAD isoenzymes from *A. thaliana* [138].

## I.7 Objectives of the research

High-level precursor amounts of coniferyl alcohol are required in order to meet the demanded microbial production of valuable lignans purposed to potentially counteract the rising incidences of life-style related diseases. Hence, the main objective of this thesis is to implement an efficient synthetic phenylpropanoid pathway towards coniferyl alcohol into *E. coli* as expression host. In the course of this, different enzyme candidates for the central steps around TAL, C3H and OMT as well as 4CL from different bacterial, fungal and plant origins are selected and evaluated *in vivo*. As part of a successive concept, the TAL and C3H candidates with the highest individual activities are combined in the examination for the upstream pathway (Figure 1.8). The resulting best TAL + C3H combination is united with the separately tested downstream pathway afterwards. The selectively constructed complete synthetic pathway is investigated with respect to further optimization. The novel genome-integrated *E. coli* strains, which contain the *p*-coumaryl alcohol pathway and are assumed to have a reduced metabolic burden, are compared to the original *E. coli* BL21(DE3) strain in two different *in vivo* cell assays.



**Figure 1.8: Schematic overview of the work described in this dissertation.** The complete synthetic pathway (green boxes) is divided into upstream (yellow box) and downstream pathway (blue box), which are first characterized separately. Different enzyme candidates for TALs/PALs and C3Hs are evaluated individually, while various OMTs are examined simultaneously with 4CL enzymes. Finally, both pathways are combined and investigated via two cell assays as well as in two *E. coli* strains. The respective chapters (ch.) in this thesis are depicted in the red boxes.

## II. Results and Discussion

All used vector constructs in this thesis are summarized in Table 6 (chapter V.1.7) and their respective performed cloning procedures are described in detail in Table 16 and 17 (Appendix) including all intermediary plasmids. Most of the constructs were cloned using PCR, restriction and ligation, for which the applied primer pairs (Table 5, chapter V.1.6) and restriction enzyme sites for vectors and inserts are depicted as well (Table 16, Appendix). All constructs were sequenced for correct gene insertions and mutagenesis. Furthermore, colony PCR (cPCR) was performed for all retransformations of plasmids and co-transformations using gene-specific primers in order to ensure properly transformed, picked colonies for further investigations regarding expression and both types of cell assays (chapter V).

In general, whole-cell biotransformation (resting cell assay) is applied for examination of the upstream and downstream pathway and the composed pathways towards *p*-coumaryl and coniferyl alcohol (chapters II.1-II.5). At first, the single enzyme reactions of the upstream pathway encompassing TAL/PAL and C3H/C4H are characterized individually (chapters II.1 and II.2) and together (chapter II.3). Then, the downstream pathway consisting of 4CL, CCR, CAD and OMT is evaluated regarding different methyltransferase and 4CL enzyme candidates (chapter II.4). The complete synthetic pathways towards *p*-coumaryl and coniferyl alcohol are finally examined in chapter II.5. At last, the growing cell assay is utilized for the production of coniferyl alcohol from caffeic acid, L-tyrosine and cinnamic acid (chapter II.6).

### II.1 Deamination

As the first reaction step in a multi-enzyme pathway is often crucial, this enzymatic reaction needs to be fast, efficient and tolerant towards high substrate and product concentrations in order to prevent early inhibition and inactivation effects in the pathway performance. Thus, evaluation of enzyme candidates for the TAL and PAL reaction should give the most suitable enzymes for both entry points of the synthetic pathway. In light of improving the previously constructed pathway towards *p*-coumaryl alcohol, the alternative TAL enzymes are tested in comparison to the TAL from *R. sphaeroides* [135]. The tyrosine ammonia lyases from *Flavobacterium johnsoniae* (FjTAL), *Saccharothrix espanaensis* (Sesam8) and HaTAL were chosen due to their higher *in vitro* catalytic efficiencies and *in vivo* activities towards L-tyrosine in *E. coli* compared to RsTAL [65]. In addition, another TAL from the yeast *R. glutinis* was included in the activity screening towards L-tyrosine due to its previous application in many multi-enzyme pathways and a twelve-fold higher reported catalytic efficiency compared to

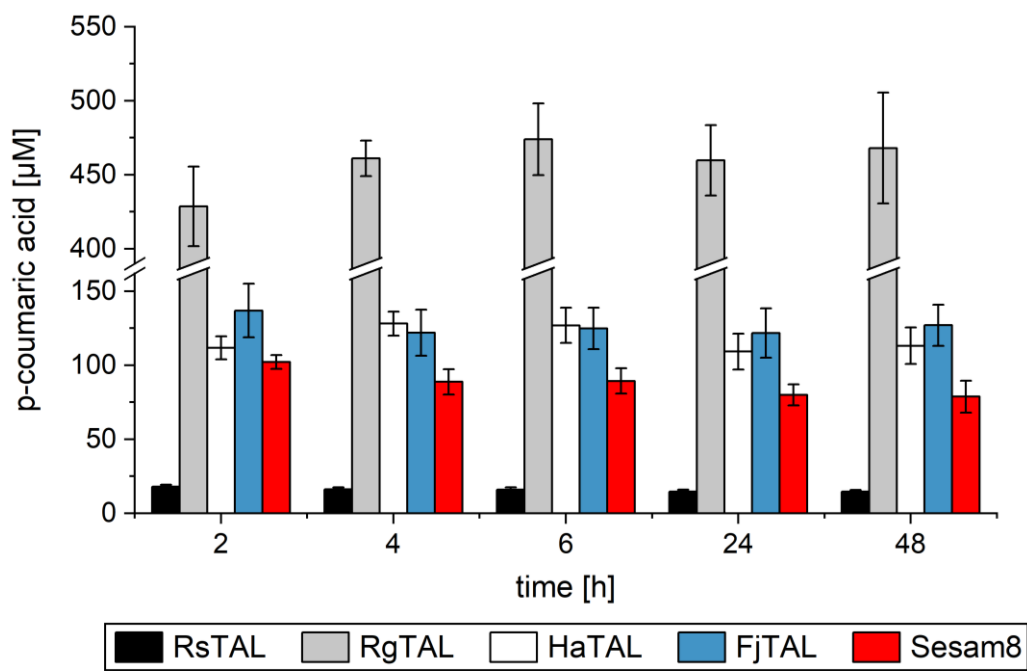
*RsTAL* [133,139–141]. Also, *RgTAL* exhibits rather small PAL/TAL activity ratios, which is why its dual functionality as a PAL is compared to the PAL from *Z. mays* (*ZmPAL*) for the synthetic pathway starting from L-phenylalanine [58,141].

For the determination of *in vivo* activities of the individual enzymes and also for later combined enzymatic activities, whole-cell biotransformation was mainly applied. Thereby, the TAL proteins were expressed from the vector pETM6\_TAL in *E. coli* BL21(DE3) in TB medium, the cells were harvested after 20 h, washed and resuspended as 70 g<sub>wet</sub>/L cell suspension with 2 % (w/v) glucose and 1 mM IPTG prior to substrate addition. In this resting cell assay, glucose was added to ensure a maintained cell metabolism in order to provide reduced equivalents of NADPH and NADH as well as ATP for the individual synthetic pathway reactions [142,143]. Furthermore, a neutral pH of 7.5 and a temperature of 26 °C were chosen based on the model of large-scale industrial fermentations in order to save operational costs.

### II.1.1 Tyrosine ammonia lyases

The five investigated tyrosine ammonia lyases (*RsTAL*, *RgTAL*, *HaTAL*, *FjTAL*, *Sesam8*) already exhibited deamination activity at 2 h after addition of 1 mM L-tyrosine in the whole-cell biotransformation (Figure 2.1). Interestingly, no further significant increase in the production of *p*-coumaric acid was determined after that time point. The highest measured concentration of *p*-coumaric acid was  $473.9 \pm 24.2 \mu\text{M}$  at 6 h for *RgTAL*,  $128.1 \pm 8.1 \mu\text{M}$  at 4 h for *HaTAL*,  $136.9 \pm 18.1 \mu\text{M}$  at 2 h for *FjTAL* and  $102.2 \pm 4.8 \mu\text{M}$  at 2 h for *Sesam8*. The highest *p*-coumaric acid concentration for *RsTAL* was below the standard curve ( $17.8 \pm 1.5 \mu\text{M}$ , 2 h) resulting in the lowest measured TAL activity in the resting cell assay. In Figure 2.1, *RgTAL* exhibited an almost five times higher production of *p*-coumaric acid, whereas the three enzymes *HaTAL*, *FjTAL* and *Sesam8* showed comparable *in vivo* activities in *E. coli* BL21(DE3). No production of *p*-coumaric acid was detected in the activity control consisting of resting *E. coli* cells with the empty vector (data not shown).

With the aim to evaluate expression of the individual TAL enzymes in *E. coli*, 5/OD samples were used for normalization of cell density during expression. Furthermore, the harvested cell pellet was divided into a soluble supernatant and an insoluble pellet fraction after sonication in order to assess the expression behavior of the protein. If the protein is observed in the soluble fraction, the expressed protein is likely located in the cytosol of *E. coli* as sonication disrupts the cell membrane. An accumulation in the insoluble pellet fraction of *E. coli* lysate can not only indicate the formation of inclusion bodies as a result of protein overexpression but also incorrect folding or membrane association of the particular protein [144,145].

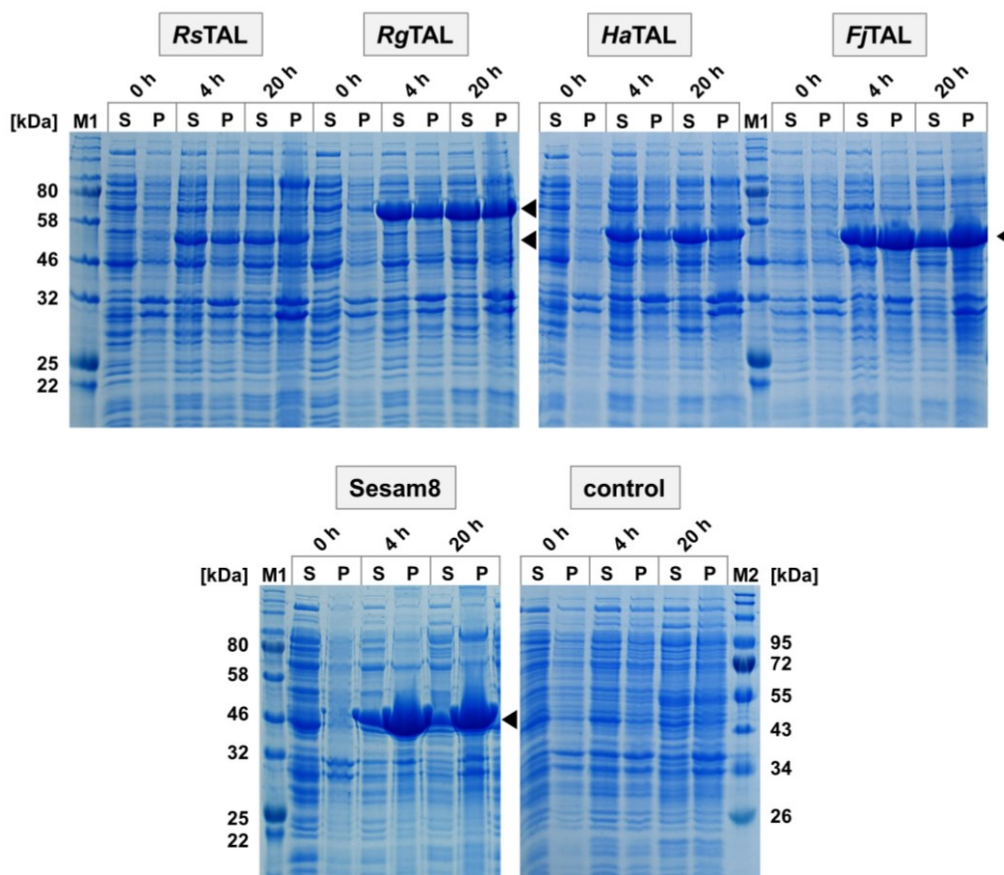


**Figure 2.1:** *In vivo* activities of different TAL enzymes in *E. coli* BL21(DE3). Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM L-tyrosine as substrate added at time  $t_0$ . The product concentration of *p*-coumaric acid was measured over time (2 h, 4 h, 6 h, 24 h, 48 h). The TAL enzymes *RsTAL* (black), *RgTAL* (grey), *HaTAL* (white), *FjTAL* (blue) or *Sesam8* (red) were expressed from the pETM6 vector prior to application in whole-cell biotransformation. The shown data represent mean values with standard deviation from biological triplicates.

All TAL enzymes were found to be expressed in both fractions indicating the soluble protein character of the TALs and the formation of inclusion bodies in the insoluble fraction (Figure 2.2). The production of the individual proteins was stable over time because the band intensities at 4 h and 20 h were equally pronounced. On the SDS-PAGE, the enzymes *RgTAL* (74.8 kDa), *FjTAL* (56.7 kDa) and *Sesam8* (54.0 kDa) showed pronounced expression bands, while the expression bands of *HaTAL* (60.3 kDa) were less pronounced. For *RsTAL*, thin expression bands could be observed around 55.5 kDa, which were not visible in the expression of the empty vector control.

*HaTAL* had the third highest measured activity but showed relatively thin expression bands. Taking into account that both expression and activity evaluation were normalized to cell density, it could be suggested that *HaTAL* might be the second most active TAL enzyme among the tested ones in whole-cell biotransformation. By comparing the pH optima of the TAL enzymes, *FjTAL* and *Sesam8* are found to have more basic pH optima (pH 10) than *RsTAL*, *RgTAL* and *HaTAL* (pH 9.0-9.5) [65,141]. It is supposed that the similar deamination activity of *FjTAL* and *Sesam8* with *HaTAL* originated from the further distance from their optimal pH. However, a pH of 7.5 is also not the optimal pH for all TALs. Thus, the measured TAL activities

are probably a result of improved enzyme stability due to the whole-cell environment compared to free purified enzymes and reduced activity at pH 7.5 [146,147].



**Figure 2.2: SDS-PAGE analysis of the different TAL enzymes.** Expression was performed from pETM6\_TAL with *RsTAL*, *RgTAL*, *HaTAL*, *FjTAL* and *Sesam8* in *E. coli*/BL21(DE3) at 26 °C. 5/OD samples were taken at the point of induction with 1 mM IPTG (0 h) and during the expression (4 h and 20 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE analysis. An expression of the empty vector pETM6 served as **control**. The black arrowheads indicate the relevant expression bands. M1 – P7712 protein ladder (NEB), M2 – P7719 protein ladder (NEB), S – supernatant, P – pellet fraction, *RsTAL* – 55.5 kDa, *RgTAL* – 74.8 kDa, *HaTAL* – 60.3 kDa, *FjTAL* – 56.7 kDa, *Sesam8* – 54.0 kDa.

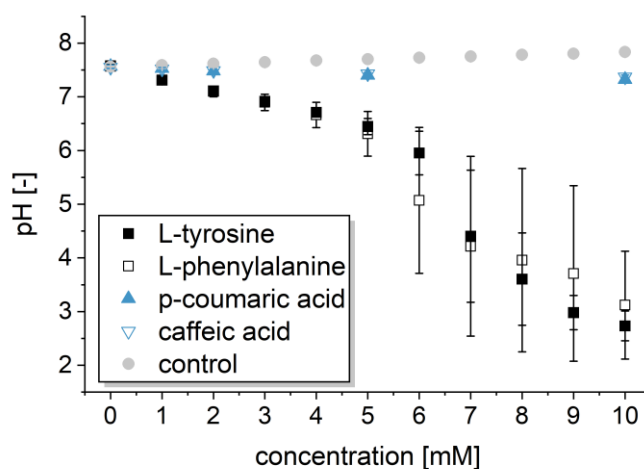
Nonetheless, expression of multiple pathway enzymes in one host cell might become a bottleneck later on. Therefore, *FjTAL* and *Sesam8* were chosen for further investigation with the complete synthetic pathway due to their pronounced expression bands as well as *RgTAL* because of its high detected *in vivo* activity.

### II.1.2 Evaluation of pH change by applied substrates

The whole-cell biotransformation experiments showed that an increasing concentration of L-phenylalanine or L-tyrosine in KPi buffer led to a significant decrease in pH (Figure 2.3). With 1 mM L-phenylalanine, the pH decreased from a value of 7.6 to 7.3 and with 3 mM L-phenylalanine to a pH of  $6.9 \pm 0.2$ . Using higher concentrations of L-phenylalanine, the pH dropped more rapidly probably due to the exceedance of the buffer capacity. The addition of



L-tyrosine showed a similar decrease in pH, while concentrations of *p*-coumaric acid and caffeic acid had only negligible effects on the pH. With 10 mM of *p*-coumaric acid or caffeic acid, the pH stayed at 7.3 or at 7.4, respectively, in comparison to a pH of 7.8 in the control.



**Figure 2.3: Influence of different substrates on pH.** The different substances L-tyrosine (black, filled square), L-phenylalanine (black, open square), *p*-coumaric acid (blue, closed triangle) and caffeic acid (blue, open triangle) were successively added to 10 mL KPi buffer (50 mM, pH 7.5) and the pH measured. Stocks of substances were prepared as 100 mM in DMSO. Addition of only DMSO served as control (grey, filled circle). The measurements were performed in triplicates and the error bars indicate the standard deviation from the mean.

The main reason for the significant decrease in pH with L-tyrosine and L-phenylalanine is their preparation by acidification with HCl due to insolubility of these substances at a neutral pH value [148]. Substance preparations through basification using NaOH resulted in an equivalently strong increase in pH with higher concentrations (data not shown). In the following enzyme characterizations using whole-cell biotransformation, a maximal concentration of 3 mM for L-tyrosine and L-phenylalanine seemed to be applicable with respect to maintaining a neutral pH (chapters II.1.3 and II.3). Meanwhile higher concentrations of caffeic acid are applied in the resting cell assay for evaluation of the downstream pathway (chapter II.4.3).

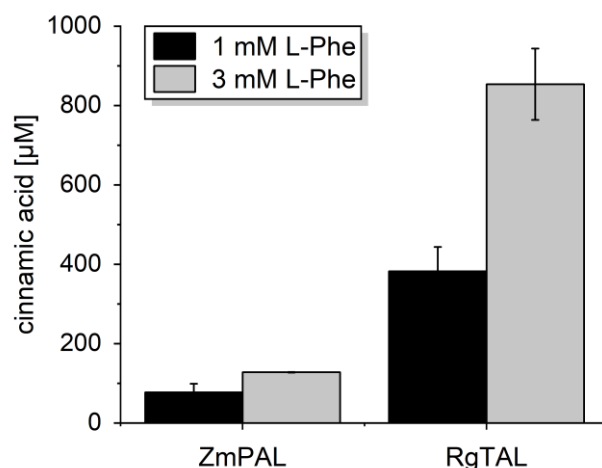
### II.1.3 Phenylalanine ammonia lyases

For an economically efficient process low input costs can be a decision-making factor in industry. Judging from the prices of L-tyrosine (324 €/kg)<sup>1</sup> and L-phenylalanine (283 €/kg)<sup>1</sup>, the latter would be the more preferred substrate [149]. Thus, it is essential to investigate the production starting from L-phenylalanine as well. In plants, the PAL enzyme constitutes the entry point for the synthesis of monolignols. Most PAL enzymes are also known to exhibit a

<sup>1</sup> Calculated from the largest available quantity from L-tyrosine (order number: T207.3) and L-phenylalanine (order number: 4491.2) from Carl Roth GmbH & Co. KG (Karlsruhe, Germany).

TAL activity probably due to the structural similarity between the substrates L-tyrosine and L-phenylalanine.

The PAL/TAL ratio has been determined for the TAL from the yeast *R. glutinis* to be 1.68 using cell extracts [58]. *RgTAL* showed the lowest PAL/TAL ratio in that study indicating this enzyme possessed sufficient PAL activity, too. Indeed, in a whole-cell biotransformation higher concentrations of cinnamic acid were measured for *RgTAL* compared to the PAL from *Z. mays* with both substrate feeds of 1 mM and 3 mM L-phenylalanine (Figure 2.4). The highest cinnamic acid productions in whole-cell biotransformation using *ZmPAL* were  $78.9 \pm 20.2 \mu\text{M}$  with a 1 mM feed of L-phenylalanine and  $130.7 \pm 0.7 \mu\text{M}$  with a 3 mM feed of L-phenylalanine after 48 h. *RgTAL* produced around five times higher amounts of cinnamic acid for 1 mM initial L-phenylalanine concentration,  $382.1 \pm 61.4 \mu\text{M}$  (24 h), and around 6.5 times higher for a feed of 3 mM L-phenylalanine,  $853.6 \pm 90.1 \mu\text{M}$  (24 h).

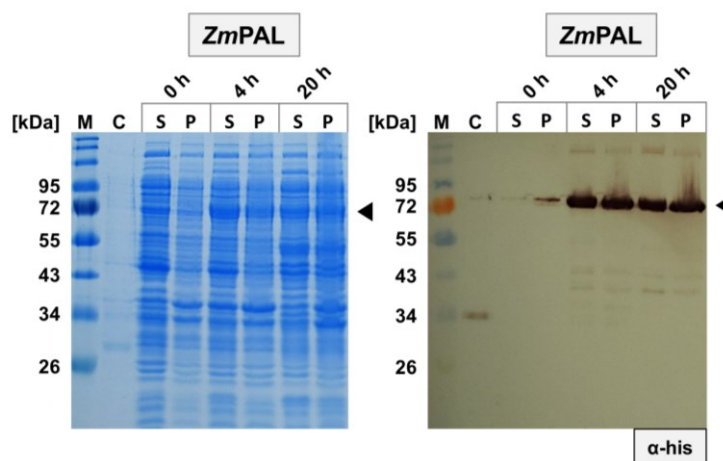


**Figure 2.4:** *In vivo* activities of different PAL enzymes in *E. coli* BL21(DE3). Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM (black) or 3 mM L-phenylalanine (L-Phe, grey) as substrate added at time  $t_0$ . The product concentration of cinnamic acid was measured after 48 h. *ZmPAL* and *RgTAL* were expressed from the pETM6 vector prior to application in whole-cell biotransformation. The shown data represent mean values with standard deviation from biological duplicates.

Kinetic properties of *RgTAL* have been already investigated in the past and the PAL/TAL ratio was found to be dependent on pH and substrate concentration [58,141]. The PAL/TAL ratio of 1 mM L-tyrosine and L-phenylalanine from the whole-cell biotransformation led to a value between 0.66-1.02 *in vivo* and at neutral pH (calculated from  $382.1 \pm 61.4 \mu\text{M}/459.7 \pm 23.7 \mu\text{M}$ , at 24 h). Thus, *RgTAL* seems to be a suitable candidate for catalysis of both deamination reaction steps and enables synthesis from two pathway entry points.

With regard to expression levels of *ZmPAL* (78.6 kDa), thin bands were visible on the SDS gel slightly above 72 kDa at 4 h and 20 h (Figure 2.5). However, the His<sub>6</sub>-tag of the *ZmPAL* protein could be distinctly detected in western blot analysis suggesting its sufficient expression. In

general, western blot analysis applies specific antibodies for detection of a particular protein and allows detection in a lower protein concentration range in comparison to visible bands in SDS-PAGE analysis [150]. For this reason, the sensitive detection resulted in pronounced bands of *ZmPAL* on the western blot, although the protein bands were only poorly visible on the SDS gel. It has to be noted that a semi-quantitative comparison of protein bands between SDS-PAGE and western blot analysis can not be performed because of the above mentioned high sensitivity of protein detection via antibodies [151].



**Figure 2.5: SDS-PAGE and western blot analysis of PAL from *Z. mays*.** Expression was performed from pETDuet<sub>his6</sub>-*ZmPAL* (*ZmPAL*) in *E. coli*/BL21(DE3) at 26 °C. 5/OD samples were taken at the point of induction with 1 mM IPTG (0 h) and during the expression (4 h and 20 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE and western blot analysis. For detection in western blot analysis the anti-His ( $\alpha$ -his) and GAM<sub>FC</sub><sup>AP</sup> antibodies were used. The black arrowheads indicate the relevant expression bands. M– P7719 protein ladder (NEB), C– mCherry-his<sub>6</sub> as His<sub>6</sub>-tag control (0.5  $\mu$ g, 27.7 kDa), S – supernatant, P – pellet fraction, his<sub>6</sub>-*ZmPAL* – 78.6 kDa.

In conjunction with the presented data, *ZmPAL* seems to be unsuitable as PAL enzyme *in vivo* in *E. coli* because of its low conversion of L-phenylalanine. Due to the superior *in vivo* PAL activity of *RgTAL*, this tyrosine ammonia lyase is used in the further combination of a PAL and a C4H enzyme in order to yield *p*-coumaric acid as central metabolite and complete the synthetic pathway starting from L-phenylalanine as substrate.

## II.2 C3-hydroxylation

The hydroxylation on the C3 atom represents the second most important reaction step in the synthetic phenylpropanoid pathway and creates a junction between the simple *p*-hydroxyphenyl units and the diversified guaiacyl and syringyl units (Figure 1.3). The C3-hydroxylation step in the plant phenylpropanoid pathway is assumed to occur mainly on the level of shikimate or quinate esters of *p*-coumaric acid. For the generation of these esters, a transferase enzyme activity such as from a hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase

enzyme is necessary. In the end, three to four enzymes from plant origin are necessary for the realization of the natural C3-hydroxylation step in a synthetic microbial pathway.

With the intention to generate a facile C3-hydroxylation step, two bacterial enzymes were chosen, which are known to catalyze the direct conversion of *p*-coumaric acid into caffeic acid. These enzymes are the non-P450 hydroxylase complex HpaBC from *E. coli* and the class I cytochrome P450 monooxygenase CYP199A2 from *R. palustris*. The expression and activity of the bacterial enzymes in *E. coli* are examined and compared.

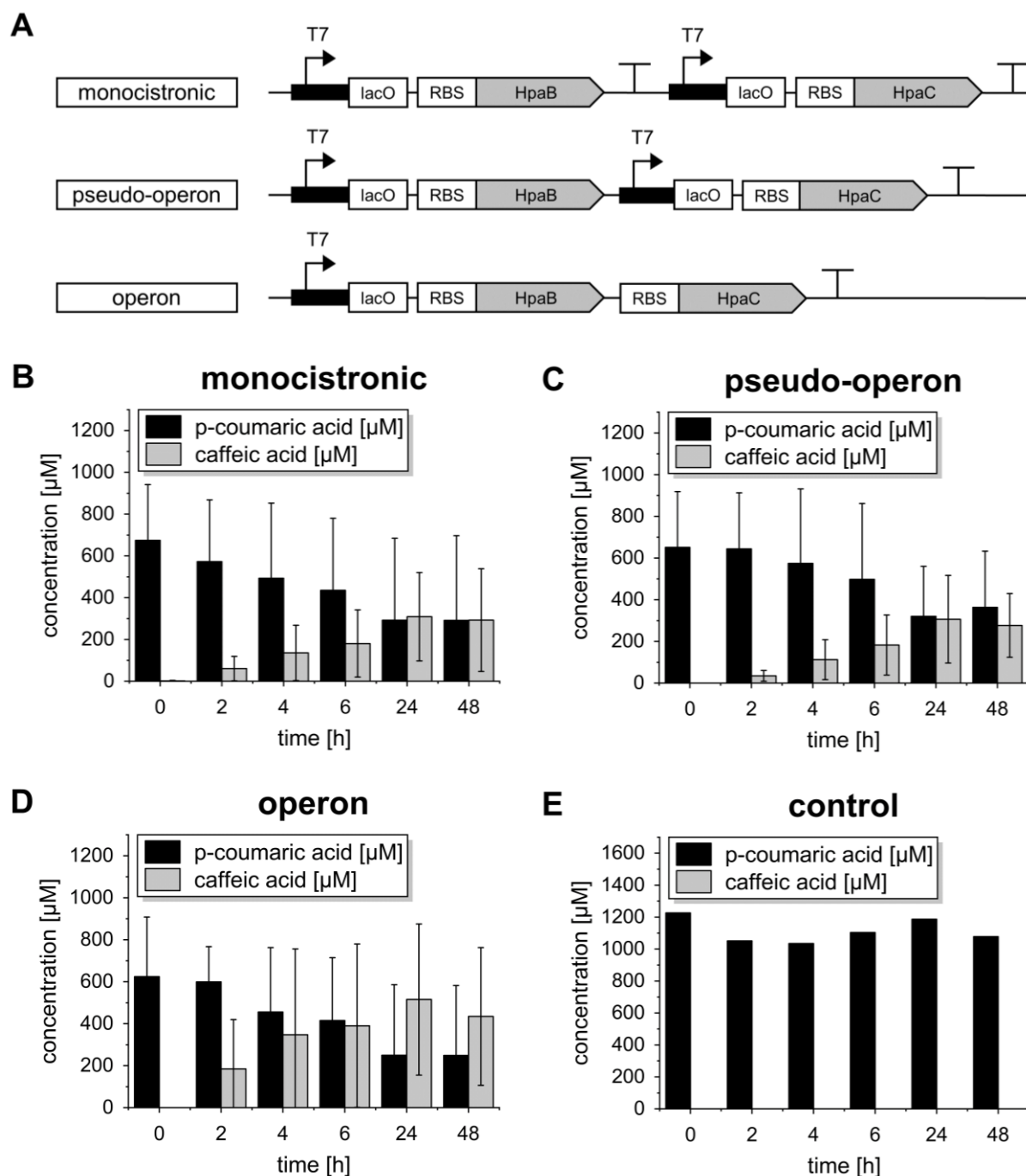
### **II.2.1 The *E. coli* hydroxylase complex HpaBC**

In contrast to the Duet vectors, the ePathBrick vectors contain four additional restriction sites that are recognized by so called isocaudamers (Figures A.20 and A.21, Appendix) [152]. These comprise a group of restriction enzymes that bind to specific individual sequences, but generate one identical overhang by cutting. Such restricted gene fragments can easily be ligated into a vector, which has been cut with the same isocaudamers. In the case of the ePathBrick vectors, the isocaudamer sites are located in a way that allows for unidirectional and successive assembly of genes in three different gene configurations: monocistronic, pseudo-operon and operon. Xu *et al.* (2012) and He *et al.* (2015) demonstrated that evaluation of the appropriate configuration for a fixed gene order was crucial for further pathway engineering, because the performance of multi-enzyme cascades might differ substantially [152,153].

Hence, different configurations for the two components HpaB and HpaC were tested on the pETM6 vector in order to evaluate the highest activity of the non-P450 hydroxylase complex HpaBC towards *p*-coumaric acid in a resting cell assay. In the monocistronic gene organization, both HpaB and HpaC contain a T7 promoter, lac operator sequence (lacO) and an RBS in front of the individual gene sequence and a T7 termination sequence after the gene sequence (Figure 2.6 A). In comparison, in a pseudo-operon, only a single T7 terminator is located after the last gene. An operon gene configuration is present if further a single T7 promoter with lacO is located in front of the first gene. These three gene configurations lead to different levels and compositions of mRNA und in this way influence protein expression levels in the end [153].

In general, hydroxylation activity of *p*-coumaric acid into caffeic acid was observed for all analyzed configurations in whole-cell biotransformations (Figure 2.6 B-D). The highest concentrations of caffeic acid were located at  $308.8 \pm 211.6 \mu\text{M}$  for monocistronic,  $306.6 \pm 210.2 \mu\text{M}$  for pseudo-operon and  $515.5 \pm 360 \mu\text{M}$  for operon gene configuration after 24 h. With regard to mean values, HpaBC exhibited a 1.7 times higher hydroxylation activity, when organized as an operon (conversion yield 82.5 %) compared to the monocistronic (45.8 %) and

pseudo-operon configuration (47.1 %), which showed quite comparable *in vivo* activities. The high calculated standard deviations may be attributed to the performance of two individual expression experiments with induction at different OD<sub>600</sub> values probably leading to variations in protein production of both HpaBC components.



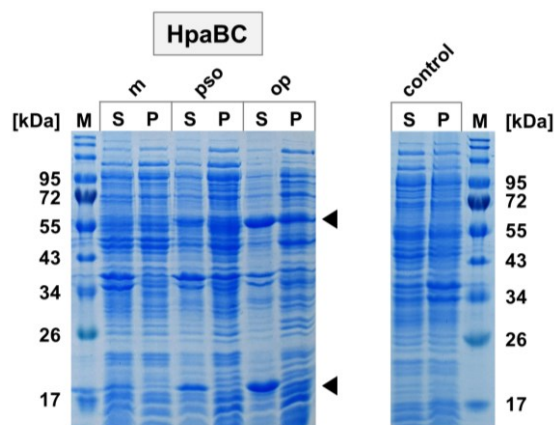
**Figure 2.6:** *In vivo* C3H activity of the hydroxylase complex HpaBC from *E. coli*. **A** Schematic representation of the three different gene configurations of HpaBC on the pETM6 plasmid. **B-E** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM *p*-coumaric acid as substrate added at time  $t_0$  (0 h). The concentration of substances was measured over time (2 h, 4 h, 6 h, 24 h, 48 h) (**B-D**). The expression of HpaBC was performed at 37 °C from pETM6 in three different configurations: monocistronic (**B**), pseudo-operon (**C**) or operon (**D**). The shown data represent mean values with standard deviation from biological triplicates. A single whole-cell biotransformation using the empty vector served as activity control (**E**).

In 2012, Lin and Yan were the first group who reported activity of HpaBC towards *p*-coumaric acid, which was preferred as substrate over L-tyrosine *in vivo* and *in vitro* [101]. The hydroxylase complex was presumably organized as an operon considering that they cloned the genes directly from the *E. coli* genome using a single primer pair and that about half of *E. coli* proteins are organized as operons [154]. In another study, pETM6\_HpaBC in the monocistronic configuration was subjected to random assembly with promoters of different strengths to potentially improve activity towards the substrate naringenin [100]. But no alternative promoter arrangement was identified that performed equally to the construct with the two T7 promoters. These findings together with the activity results support the usage of HpaBC in operon configuration with the strong T7 promoter for the combined investigation with different TALs (chapter II.3).

It has to be noted that HpaBC has been already applied in the *in vivo* production of caffeoyl alcohol by using *p*-coumaryl alcohol as substrate [137]. A single whole-cell biotransformation with a feed of 1 mM *p*-coumaryl alcohol resulted in a one-third lower product concentration of caffeoyl alcohol compared to the production of caffeic acid by HpaBC in operon configuration (Figure A.1, Appendix). This could indicate that *p*-coumaric acid is the slightly more preferred substrate for HpaBC *in vivo*.

No activity in whole-cell biotransformation using the empty vector control was detected, which was in accordance with reports that the hydroxylase complex is natively silenced in *E. coli* (Figure 2.6 E) [133]. Being consistent with that, no protein bands for the single components HpaB (58.9 kDa) and HpaC (18.5 kDa) were observed in the SDS-PAGE analysis, either (Figure 2.7). HpaB and HpaC showed distinct expression bands on the SDS gel, when the proteins were expressed from the pETM6 vector in operon configuration, while the protein bands of the pseudo-operon gene organization appeared weaker. For the HpaBC components expressed from the monocistronic gene configuration, no distinct protein bands for HpaB and only faint bands for HpaC were observed. In a second expression experiment, the band intensities for the protein components were similar using the different gene configurations except that the expression bands of the smaller protein HpaC were more pronounced with regard to all three constructs (Figure A.2, Appendix).

These results point out that the expression of both protein components is more balanced when expressed from an operon gene configuration. At last, HpaBC is applied as an operon in the combined enzyme evaluation with the TAL enzymes also due to the superior hydroxylation activity using this gene configuration (chapter II.3).



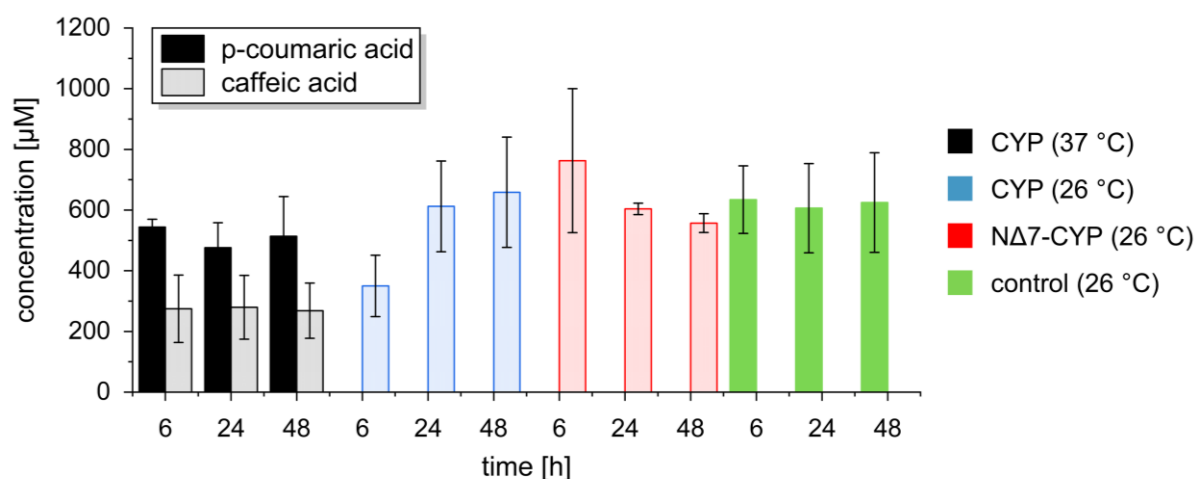
**Figure 2.7:** SDS-PAGE analysis of the HpaBC hydroxylase complex from *E. coli*. Expression was performed from pETM6\_HpaB\_HpaC (HpaBC) in monocistronic (m), pseudo-operon (pso) or operon (op) configuration in *E. coli* BL21(DE3) at 37 °C (left). 5/OD samples were taken at 20 h and divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE analysis. The black arrowheads indicate the relevant expression bands. The empty vector pETM6 expressed at 26 °C and at 20 h served as **control** (right). M – P7719 protein ladder (NEB), S – supernatant, P – pellet fraction, HpaB – 58.9 kDa, HpaC – 18.5 kDa.

## II.2.2 The cytochrome P450 monooxygenase from *Rhodopseudomonas palustris*

### *In vivo* activity using the native redox partners

Another C3H reported to directly convert *p*-coumaric acid into caffeic acid is CYP199A2, a cytochrome P450 monooxygenase from *R. palustris*. Palustrisredoxin (Pux) and palustrisredoxin reductase (PuR) were applied as native redox partners to assess the *in vivo* activity of CYP199A2 in *E. coli* BL21(DE3) (Figure 2.8). Expression was performed at different temperatures and generally higher caffeic acid concentrations were observed in resting cells, when the proteins were expressed at 26 °C.

At 37 °C and with 1 mM *p*-coumaric acid, the substrate was left unconverted and caffeic acid formation did not significantly increase after 6 h of whole-cell biotransformation (Figure 2.8). The highest measured concentration of caffeic acid was  $279.8 \pm 105 \mu\text{M}$  after 24 h for the expression of CYP199A2 at 37 °C. The respective conversion yield was 43.2 %, which equals the C3H activity of the HpaBC complex expressed from monocistronic and pseudo-operon gene configurations (cf. chapter II.2.1). When expression was performed at 26 °C, no residual *p*-coumaric acid could be detected after 6 h. The caffeic acid titers were found to be higher at every point of time with an expression temperature at 26 °C instead of 37 °C, with the highest titer being  $658.7 \pm 181.6 \mu\text{M}$  caffeic acid after 48 h. So the activity of CYP199A2 towards *p*-coumaric acid was increased by 135 % simply by lowering the expression temperature. These results are comparable to previous studies, in which higher expression temperatures for TAL and C3H enzymes at 37 °C and 30 °C resulted in lower titers compared to an expression at 26 °C [132].



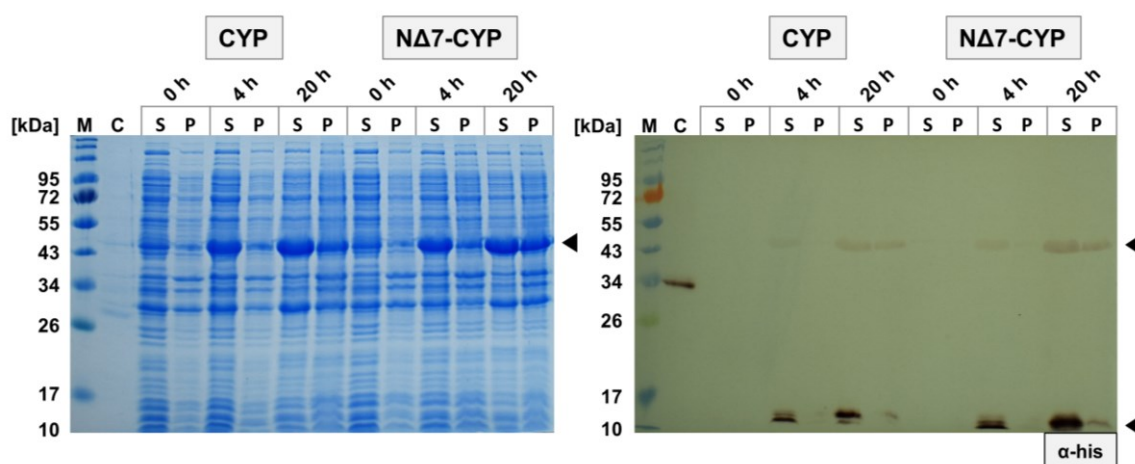
**Figure 2.8:** *In vivo* C3H activity of CYP199A2 and its NΔ7 variant in *E. coli* BL21(DE3). Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM *p*-coumaric acid as substrate added at time  $t_0$ . The concentration of substances was measured over time (6 h, 24 h, 48 h). CYP199A2 (CYP, black and blue) or NΔ7-CYP199A2 (NΔ7-CYP, red) were expressed from the pCDFDuet vector together with the native redox pair Pux/PuR from pRSFDuet prior to application in whole-cell biotransformation. The individual expression temperature was either 37 °C or 26 °C and is indicated in brackets. A whole-cell biotransformation using the empty vectors served as activity control (control, green). The shown data represent mean values with standard deviation from biological triplicates.

Furthermore, a variant NΔ7-CYP199A2 without the first seven amino acids at the N-terminus was constructed as it had been reported to exhibit a higher *in vivo* activity than the full length P450 monooxygenase [94,132]. In the resting cell assay, the NΔ7-CYP199A2 variant produced the highest measured concentration of caffeic acid at 6 h with a titer of  $762.7 \pm 237.3$  μM. In comparison to the highest mean product concentration of CYP199A2, this value equals a 15.8 % increase in caffeic acid titer. Interestingly, the product concentration decreased after that time point to lower values in comparison to CYP199A2 ( $557 \pm 31.1$  μM at 48 h). Nonetheless, with regard to the standard deviations of the highest mean values of caffeic acid, it can be stated that both P450 enzymes showed similar activities *in vivo* in the end. The superior titers of caffeic acid for NΔ7-CYP199A2 at 6 h might indicate that this variant catalyzed the conversion of *p*-coumaric acid faster compared to the original P450 enzyme. In a previous comparative analysis of CYP199A2 and NΔ7-CYP199A2, the latter had demonstrated a higher *in vivo* activity, although in co-expression with a TAL enzyme the differences in activity had been rather small [132]. The empty vector control of *E. coli* BL21(DE3) did not show any conversion of *p*-coumaric acid over time as expected (cf. Figure 2.8).

In single expressions of the individual proteins, it was observed that all proteins were expressed at 37 °C and 26 °C (data not shown). In the co-expression of all three proteins, CYP199A2 or NΔ7-CYP199A2 with Pux and PuR at 26 °C, protein bands were visible around the 43 kDa band of the protein standard ladder in SDS-PAGE analysis as well as around 43 kDa and 10 kDa



in western blot analysis (Figure 2.9). Both redox partners Pux and PuR contain an N-terminal His<sub>6</sub>-tag, which is detectable by antibodies in western blot analysis. The stained bands of Pux at 12.6 kDa were more pronounced compared to the bands of PuR around 44.5 kDa on the blot. A similar observation was made for the expression at 37 °C with the exception that bands for PuR were undetectable (Figure A.3, Appendix). Furthermore, expression bands of Pux at 37 °C were predominantly detected in the insoluble fraction in western blot analysis, while pronounced bands for Pux and PuR were mainly located in the soluble fractions, where the expression was performed at 26 °C (Figure 2.9).



**Figure 2.9: SDS-PAGE and western blot analysis of the cytochrome P450 monooxygenase from *R. palustris* co-expressed with Pux and PuR at 26 °C.** Co-expression was performed using pRSFDuet\_his<sub>6</sub>-Pux\_his<sub>6</sub>-PuR and pCDFDuet\_CYP199A2 (CYP) or pCDFDuet\_NΔ7-CYP199A2 (NΔ7-CYP) in *E. coli* BL21(DE3) at 26 °C. 5/OD samples were taken at the point of induction with 1 mM IPTG (0 h) and during the expression (4 h and 20 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE and western blot analysis. For detection in western blot analysis the anti-His (α-his) and GAM<sub>FC</sub><sup>AP</sup> antibodies were used. The black arrowheads indicate the relevant expression bands. M – P7719 protein ladder (NEB), C – mCherry-his<sub>6</sub> as His<sub>6</sub>-tag control (0.5 μg, 27.7 kDa), S – supernatant, P – pellet fraction, CYP199A2 – 44.6 kDa, NΔ7-CYP199A2 – 43.9 kDa, his<sub>6</sub>-Pux – 12.6 kDa, his<sub>6</sub>-PuR – 44.5 kDa.

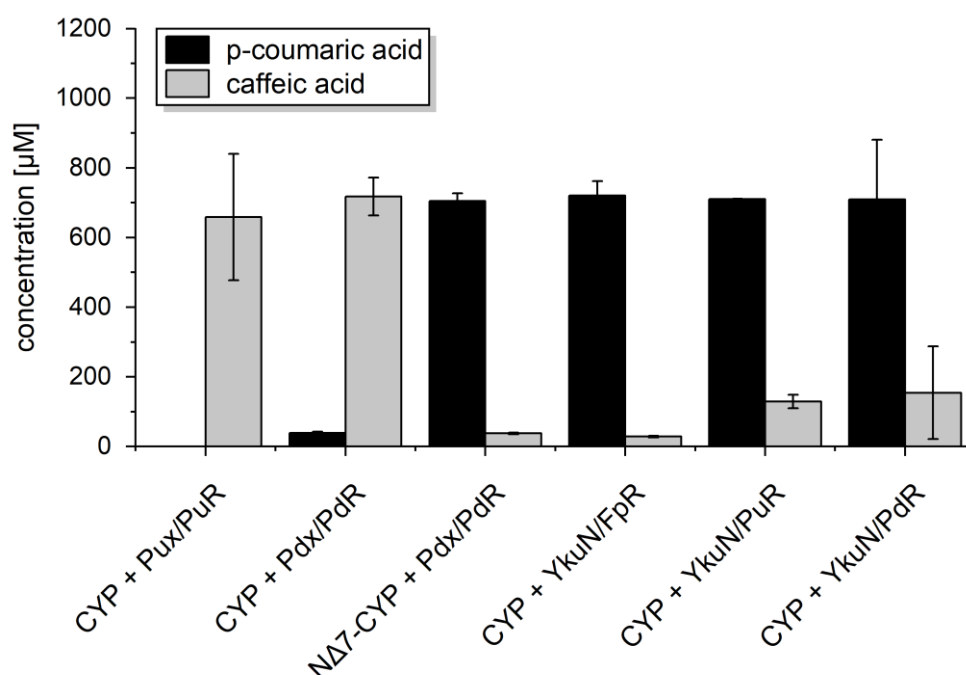
The similar band intensities at 4 h and 20 h demonstrated that the expression of most proteins was stable over time and no leaky expression occurred before induction with IPTG (0 h, Figure 2.9). It can be noted that only the band intensities of PuR improved slightly from 4 h to 20 h probably contributing to the overall expression of the reductase at 26 °C. The unlabeled P450 monooxygenases were probably expressed at 44.6 kDa (CYP199A2) or 43.9 kDa (NΔ7-CYP199A2), as the bands for PuR on the blot did not seem as pronounced as the bands in the SDS-PAGE analysis.

In general, no differences are observed between the co-expressions of Pux/PuR with both P450 enzymes at 26 °C. The expression of the redox partners is found to be more favored at a lower expression temperature leading to the formation of soluble proteins and an improved expression of the palustrisredoxin reductase. This probably contributes to the enhanced activity of the

cytochrome P450 monooxygenase at 26 °C, while the other minor differences in activity may be attributed to the use of the particular P450 variant (Figure 2.8).

### ***In vivo* activity using alternative redox partners**

As a consequence to the activity results, CYP199A2 was further investigated with regard to alternative redox partners to potentially improve the C3-hydroxylation reaction in terms of an efficient and fast electron transfer chain from reducing equivalents to the P450 reaction center. A substitution of the native redox pair with the well-known and mostly used surrogate ferredoxin and ferredoxin reductase from *P. putida* (Pdx/PdR) was found to exhibit a comparable *in vivo* activity ( $717.9 \pm 54.3$   $\mu$ M caffeic acid after 48 h) (Figure 2.10). With view on an earlier point in time (6 h), the production of caffeic acid was  $826 \pm 9.7$   $\mu$ M and thus 136 % higher compared to CYP199A2 supported by Pux/PuR ( $350 \pm 101.1$   $\mu$ M caffeic acid at 6 h, data not shown). Surprisingly, the redox pair Pdx/PdR was found to be insufficient to sustain the activity of the variant N $\Delta$ 7-CYP199A2 ( $37.6 \pm 2.6$   $\mu$ M caffeic acid at 48 h) and, at the same time, exhibited only 5.2 % of the conversion activity of CYP199A2 (Figure 2.10).

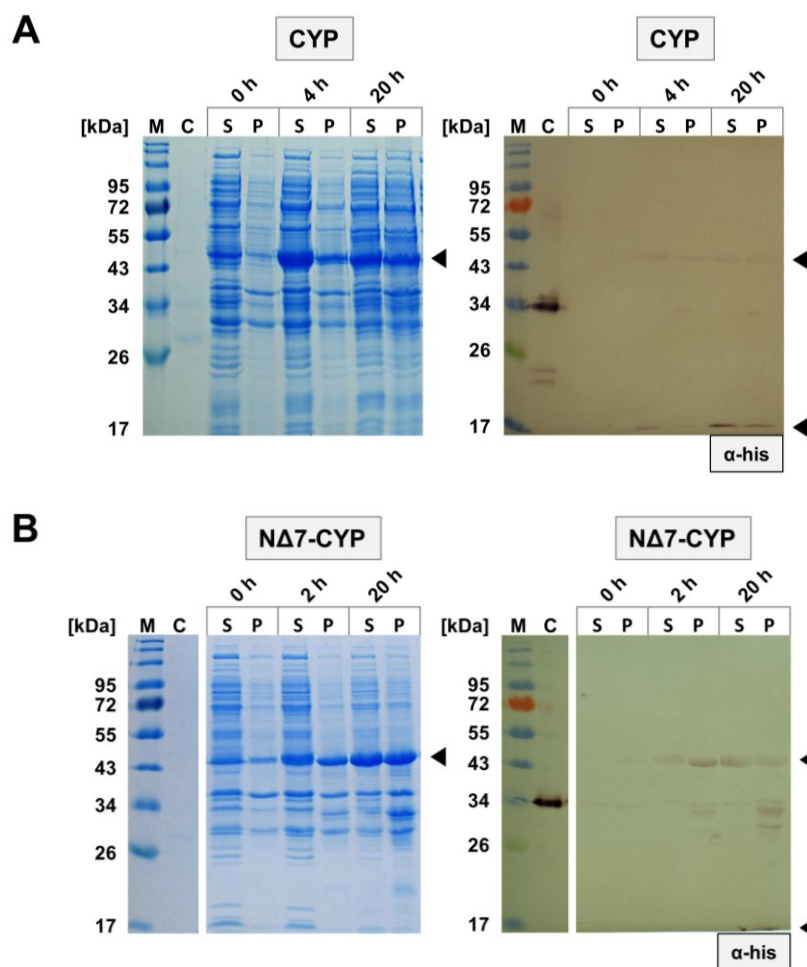


**Figure 2.10: *In vivo* C3H activity of CYP199A2 in *E. coli*/BL21(DE3) using alternative redox partners and mixed combinations.** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM *p*-coumaric acid as substrate. The concentration of substances was measured after 48 h. Expression of CYP199A2 (CYP) or N $\Delta$ 7-CYP199A2 (N $\Delta$ 7-CYP) was performed from the pCDFDuet vector at 26 °C. The different redox partner pairs (Pux/PuR, Pdx/PdR, YkuN/FpR, YkuN/PuR, YkuN/PdR) were co-expressed from the pRSFDuet vector. The shown data represent mean values with standard deviation from at least biological duplicates. Pux– palustrisredoxin from *R. palustris*, PuR– palustrisredoxin reductase from *R. palustris*, Pdx– putidaredoxin from *P. putida*, PdR– putidaredoxin reductase from *P. putida*, YkuN– flavodoxin from *B. subtilis*, FpR– flavodoxin reductase from *E. coli*.

Another redox pair involving a flavodoxin from *B. subtilis* (YkuN) and a flavodoxin reductase from *E. coli* (FpR) has been successfully used to improve *in vitro* P450 activity in former reports [85,96,155,156]. Application of these redox partners YkuN/FpR with CYP199A2 led to no efficient conversion of *p*-coumaric acid ( $28.4 \pm 2.7 \mu\text{M}$  caffeic acid at 48 h) (Figure 2.10). To identify the redox protein responsible for the ceased activity, YkuN was combined with the other two reductases PuR and PdR generating mixed redox partner combinations and reexamined with CYP199A2. The respective resulting caffeic acid concentrations with YkuN/PuR and YkuN/PdR being  $129.2 \pm 19.5 \mu\text{M}$  and  $154.2 \pm 133 \mu\text{M}$  were slightly higher compared to YkuN/FpR as redox partner pair.

These findings suggest that the flavodoxin YkuN is able to serve as electron shuttle and that the reductases PdR and PuR are more appropriate to support the activity of the P450 monooxygenase in comparison to the reductase FpR. A potential explanation can be that FpR uses intracellular NADPH unlike PuR and PdR, which both utilize NADH, and that the availability of this cofactor could have been more limiting compared to the NADH pool (cf. chapter I.4.1). Furthermore, the overall lower concentrations of caffeic acid using YkuN in a redox pair combination indicate that Pux and Pdx are capable of a more efficient electron transfer to the P450 enzyme and that the ferredoxins are generally preferred over the flavodoxin for yielding a high C3H activity of CYP199A2. Other mixed combinations of redox partners like Pux/PdR have already been applied in initial testings with N $\Delta$ 7-CYP199A2 and have been found to support activity of the P450 enzyme although not as efficient as the native redox pair [94,132,157].

Examination of the protein expression levels of CYP199A2 and N $\Delta$ 7-CYP199A2 with Pdx/PdR yielded that all three proteins showed comparable protein band intensities in SDS-PAGE and western blot analysis (Figure 2.11). Likewise, all proteins were permanently produced over time after induction. Expression bands around 44 kDa on the SDS gel demonstrated expression of the cytochrome P450 monooxygenase variants and putidaredoxin reductase as a result of the similar molecular weights of both proteins. The His<sub>6</sub>-tagged redox partners were detectable in western blot analysis with bands around 48.3 kDa for PdR and 13.7 kDa for Pdx. As before with Pux, the expression bands of the ferredoxin Pdx could not be observed on the SDS gel, but were distinctly detected in western blot analysis. Intriguingly, additional bands of putidaredoxin reductase below 34 kDa and in the pellet fraction on the blot suggested a slight degradation of PdR possibly resulting from protein overexpression.



**Figure 2.11: SDS-PAGE and western blot analysis of the cytochrome P450 monooxygenase from *R. palustris* co-expressed with Pdx and PdR at 26 °C.** Co-expression was performed using pRSFDuet\_his<sub>6</sub>-Pdx\_his<sub>6</sub>-PdR and pCDFDuet\_CYP199A2 (**CYP**, **A**) or pCDFDuet\_NΔ7-CYP199A2 (**NΔ7-CYP**, **B**) in *E. coli*/BL21(DE3) at 26 °C. 5/OD samples were taken at the point of induction with 1 mM IPTG (0 h) and during the expression (2 h or 4 h and 20 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE and western blot analysis. For detection in western blot analysis the anti-His (α-his) and GAM<sub>FC</sub><sup>AP</sup> antibodies were used. The black arrowheads indicate the relevant expression bands. M – P7719 protein ladder (NEB), C – mCherry-his<sub>6</sub> as His<sub>6</sub>-tag control (0.5 μg, 27.7 kDa), S – supernatant, P – pellet fraction, CYP199A2 – 44.6 kDa, NΔ7-CYP199A2 – 43.9 kDa, his<sub>6</sub>-Pdx – 13.7 kDa, his<sub>6</sub>-PdR – 48.3 kDa.

Due to the similar expression of the redox pair Pdx/PdR in combination with both CYP199A2 enzymes, the significant differences in activity originated from the choice of cytochrome P450 enzyme variant (Figure 2.10). The ferredoxin accepts two electrons from the FADH<sub>2</sub> of the ferredoxin reductase and transfers these to the heme *b* group of the P450 enzyme (cf. Figure 1.5). Consequently, this interaction between iron-sulfur protein and cytochrome P450 monooxygenase is probably the limiting step based on previous dissociation constants of P450 complexes [157]. The binding of the ferredoxin likely occurs around the heme proximal face of CYP199A2, where the iron is closest to the enzyme surface [158]. It was suspected that the slightly more positively charged electrostatic potential surface of NΔ7-CYP199A2 might influence the binding of Pdx. The introduction of a C-terminal point mutation into Pdx, which

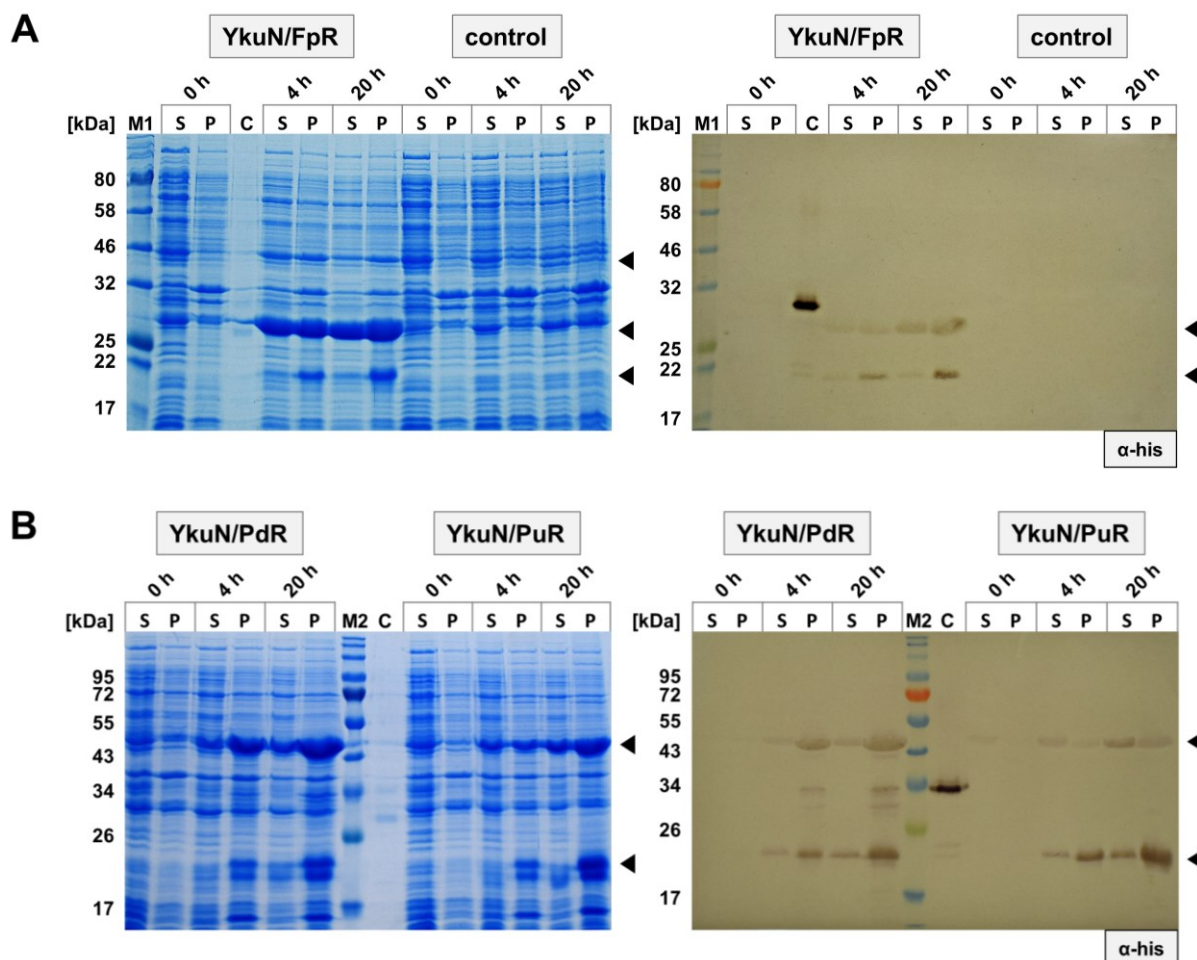
resembled the position in Pux, was found to be ineffective in improving the interaction with CYP199A2 [158]. Other efforts from Haslinger and Prather (2020) focused on complex assembly of all three protein components via individual fusions to different PCNA domains (Proliferating Cell Nuclear Antigen, from *Sulfolobus solfataricus*) [157]. These artificial constructs were able to increase the activity of N $\Delta$ 7-CYP199A2 with Pdx/PdR, which could be attributed to the generated closer proximity of ferredoxin and P450 enzyme [157].

The above presented results point out that the N-terminus of CYP199A2 could be involved in binding Pdx and reconstituted the P450 activity using the free redox pair Pdx/PdR (Figure 2.10). One possible explanation might be that the N-terminus contains hydrophobic and polar amino acids. These residues could facilitate binding of Pdx to the surface of the P450 protein via either hydrophobic and van der Waals interactions or via hydrogen bonds [159]. Alternatively, Pdx could bind to CYP199A2 in a different manner compared to Pux, which might also explain the insignificant differences in activity observed for Pux/PuR with both P450 variants (Figure 2.8). The N-terminal amino acids of CYP199A2 were found to be disordered and thus not resolved in the crystal structure [158]. So the distinct effect of the N-terminal seven amino acids on the surface structure and the involvement in potential binding of Pdx remains unknown.

Concerning the alternative redox pair YkuN/FpR, the flavodoxin reductase FpR exhibited distinct and pronounced bands around 28.6 kDa in SDS-PAGE and western blot analysis (Figure 2.12 A). Opposite to the expression of the ferredoxins Pux and Pdx, the flavodoxin YkuN already showed distinct expression bands at 20.7 kDa on the SDS gel and its His<sub>10</sub>-tag was detectable in western blot analysis. However, in both cases the bands were more pronounced in the insoluble fractions. Due to the different molecular weights between P450 enzyme and reductase, thin expression bands for CYP199A2 were visible around 44.6 kDa in SDS-PAGE analysis for the first time in co-expression experiments. This observation could be indicative that CYP199A2 is expressed in all co-expression experiments although in a lower quantity compared to a single protein expression.

For the detection of cytochrome P450 class I/III activity *in vitro*, several studies utilized an excess of ferredoxin or flavodoxin, because the transfer from the electron shuttle to the P450 enzyme is considered rate-limiting [85,86,159,160]. Thus, the dominant overexpression of FpR may not be ideal for balanced expression of all three components *in vivo* and may have contributed to the low activity using the redox pair YkuN/FpR (Figure 2.10). Furthermore, expression of the flavodoxin could be clearly detected in mixed combinations with PdR or PuR (Figure 2.12 B). The reductases PdR and PuR showed distinct and pronounced expression bands

in SDS-PAGE and western blot analysis. As before, the degradation bands for PdR were detectable on the blot around 34 kDa (cf. Figure 2.11). It generally seems that YkuN may be incompatible with CYP199A2, although all proteins of the mixed combinations were expressed and the activity was found to be slightly improved with alternative reductases. Thus, Pdx and Pux would be more suitable as electron shuttles *in vivo*.



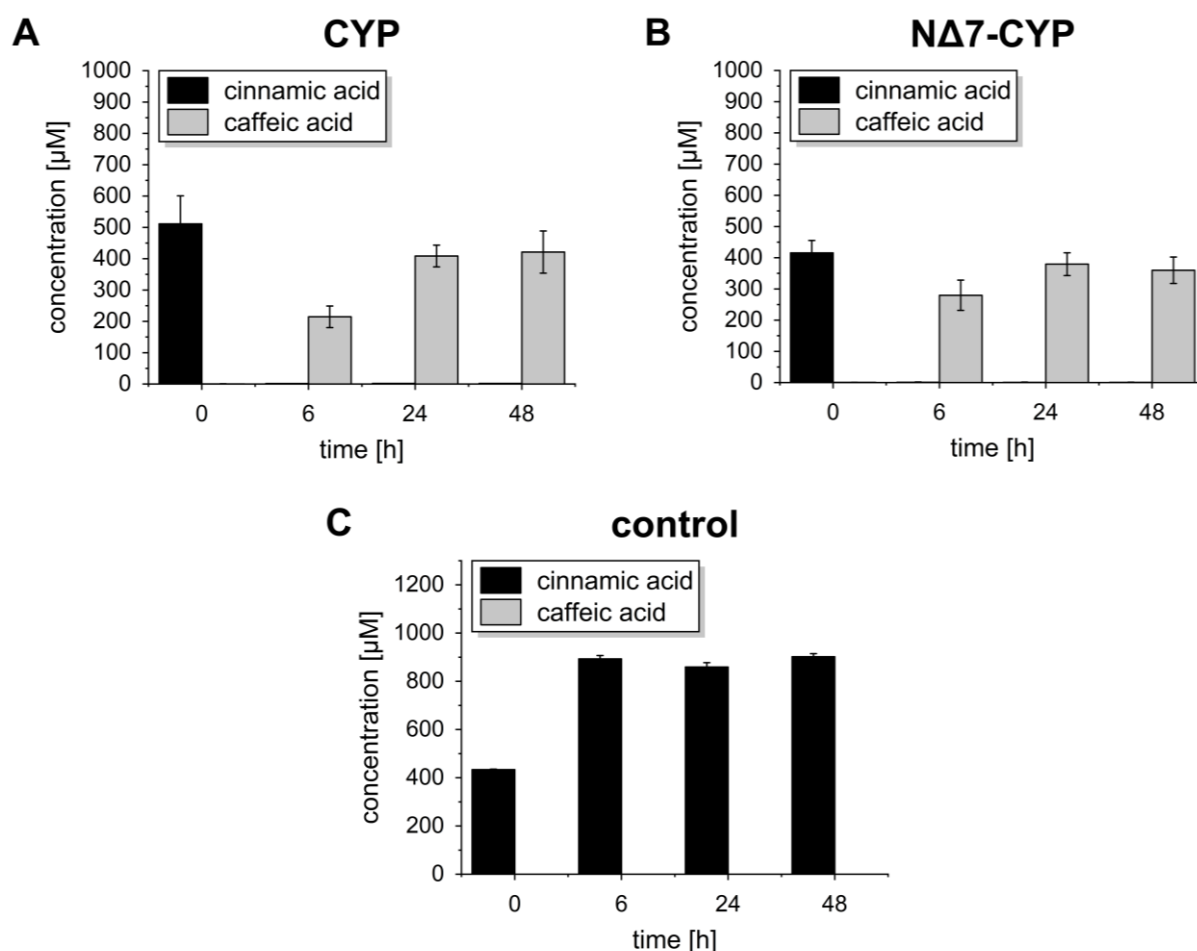
**Figure 2.12: SDS-PAGE and western blot analysis of CYP199A2 from *R. palustris* co-expressed with alternative redox partners (A) and mixed combinations (B).** Co-expression was performed using pCDFDuet\_CYP199A2 and pRSFDuet\_his<sub>10</sub>-YkuN\_his<sub>6</sub>-FpR (YkuN/FpR, A) or pRSFDuet\_his<sub>10</sub>-YkuN\_his<sub>6</sub>-PdR (YkuN/PdR, B) or pRSFDuet\_his<sub>10</sub>-YkuN\_his<sub>6</sub>-PuR (YkuN/PuR, B) in *E. coli* BL21(DE3) at 26 °C. 5/OD samples were taken at the point of induction with 1 mM IPTG (0 h) and during the expression (4 h and 20 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE and western blot analysis. For detection in western blot analysis the anti-His (α-his) and GAM<sub>FC</sub><sup>AP</sup> antibodies were used. An expression of the empty vectors served as **control** (A). The black arrowheads indicate the relevant expression bands. M1 – P7712 protein ladder (NEB), M2 – P7719 protein ladder (NEB), C – mCherry-his<sub>6</sub> as His<sub>6</sub>-tag control (0.5 µg, 27.7 kDa), S – supernatant, P – pellet fraction, CYP199A2 – 44.6 kDa, his<sub>10</sub>-YkuN – 20.7 kDa, his<sub>6</sub>-FpR – 28.6 kDa, his<sub>6</sub>-PdR – 48.3 kDa, his<sub>6</sub>-PuR – 44.5 kDa.

In conclusion, the cytochrome P450 monooxygenase CYP199A2 with Pdx/PdR as redox partners exhibited the highest *in vivo* activity. With regard to further pathway engineering, testing of both cytochrome P450 variants shall be possible. As a consequence, CYP199A2 and NΔ7-CYP199A2 are applied with the native redox pair Pux/PuR for the following evaluations.



### *In vivo* activity towards cinnamic acid

The general phenylpropanoid pathway in plants starts from L-phenylalanine. In order to yield *p*-coumaric acid as the central metabolite, a PAL enzyme and a plant cytochrome P450 monooxygenase are utilized with the latter performing the C4-hydroxylation of cinnamic acid. So far, only a few plant C4H enzymes were demonstrated to be expressed in *E. coli* and to catalyze cinnamic acid conversion *in vitro* and *in vivo* [134,161]. Still, functional expression of the C4H enzyme remains a challenge due to the required suitable redox partners to support activity and anchorage to the endoplasmic reticulum membrane. Thus, a bacterial alternative such as CYP199A2 could be able to eliminate this obstacle, as its F185L mutant was reported to accept cinnamic acid as substrate, too [94].



**Figure 2.13:** *In vivo* C4H activity of CYP199A2 (A) and its NΔ7 variant (B) in *E. coli* BL21(DE3) using its native redox pair Pux/PuR. Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM cinnamic acid as substrate added at time  $t_0$ . The concentration of substances was measured over time (6 h, 24 h, 48 h). CYP199A2 (CYP, A) or NΔ7-CYP199A2 (NΔ7-CYP, B) were expressed from the pCDFDuet vector together with the native redox pair Pux/PuR from pRSFDuet at 26 °C prior to application in whole-cell biotransformation. A whole-cell biotransformation using the empty vectors served as activity control (control, C). The shown data represent mean values with standard deviation from biological duplicates.

Whole-cell biotransformations of *E. coli* BL21(DE3) with the P450 enzyme from *R. palustris* and Pux/PuR, led to the direct production of caffeic acid from 1 mM cinnamic acid without the accumulation of *p*-coumaric acid as intermediate (Figure 2.13 A and B). The highest concentrations of caffeic acid were  $421.1 \pm 67.2 \mu\text{M}$  for CYP199A2 (48 h) and  $379.4 \pm 36.5 \mu\text{M}$  for NΔ7-CYP199A2 (24 h). The total titer of caffeic acid for CYP199A2 was slightly greater, while conversion yields indicated that this full length P450 enzyme showed a lower yield (82.6 %) in contrast to the NΔ7 variant (93 %). The empty vector control showed no consumption of cinnamic acid by resting cells of *E. coli* BL21(DE3) (Figure 2.13 C). These results indicate that both P450 monooxygenases have a dual and consecutive activity as C3H/C4H and are equally suited for application in the monolignol synthetic pathway starting from L-phenylalanine via cinnamic acid.

### **Influence of L-phenylalanine on C3H and C4H *in vivo* activities**

It is suspected that the large quantities of L-phenylalanine present in TB medium (6-8 mM)<sup>1</sup> may exhibit some kind of inhibiting effect on the enzymes from the phenylpropanoid pathway (cf. chapter II.6). In order to pursue the assumption, both cytochrome P450 monooxygenases variants were evaluated for a possible influence of L-phenylalanine on hydroxylation activities. By addition of 1 mM L-phenylalanine to whole-cell biotransformations of CYP199A2 and NΔ7-CYP199A2 with Pux/PuR using 1 mM *p*-coumaric acid as substrate, the C3H activities of both enzymes were found to be intact (Figure 2.14 A). The measured concentrations of caffeic acid amounted to  $726.4 \pm 11.0 \mu\text{M}$  for CYP199A2 and  $364.4 \pm 24.9 \mu\text{M}$  for NΔ7-CYP199A2 after 24 h. The caffeic acid production using NΔ7-CYP199A2 was 39.7 % lower compared to the value in the absence of L-phenylalanine, while the caffeic acid titer for CYP199A2 was 18.6 % higher but still within the interval of the previous standard deviation around the mean (cf. Figure 2.8). Another observation was that the substrate *p*-coumaric acid was not completely converted by the NΔ7 variant, which contradicted the previous results. This shows that the presence of L-phenylalanine influenced the C3H activity of NΔ7-CYP199A2.

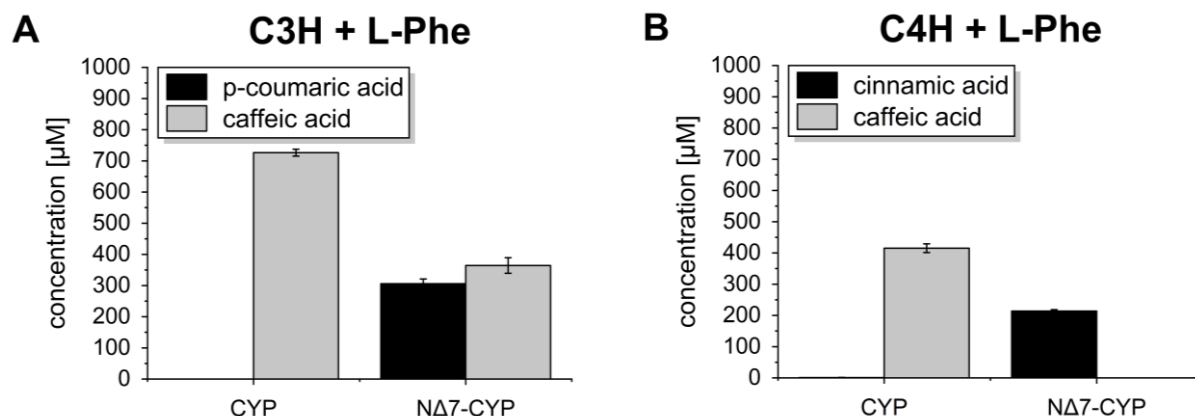
With view on the C4H activities, CYP199A2 converted the substrate 1 mM cinnamic acid in the presence of L-phenylalanine completely into a caffeic acid concentration of  $415.1 \pm 13.9 \mu\text{M}$ , which equaled the activity without addition of L-phenylalanine (Figure 2.14 B, cf.

---

<sup>1</sup> Calculated from specifications of three individual tryptones and two yeast extracts [162–166].



Figure 2.13). For NΔ7-CYP199A2, the C4H activity was abolished by the addition of 1 mM L-phenylalanine.



**Figure 2.14: Influence of L-phenylalanine on the C3H (A) and C4H activity (B) of the cytochrome P450 monooxygenase from *R. palustris*.** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM L-phenylalanine (L-Phe) and either 1 mM *p*-coumaric acid (A) or 1 mM cinnamic acid (B) as substrate. The concentration of substances was measured after 24 h. CYP199A2 (CYP, A) or NΔ7-CYP199A2 (NΔ7-CYP, B) were expressed from the pCDFDuet vector together with the native redox pair Pux/PuR from pRSFDuet at 26 °C prior to application in whole-cell biotransformation. The shown data represent mean values with standard deviation from at least biological duplicates.

It is noteworthy that Furuya and co-workers (2012) found *m*-coumaric acid as main accumulating intermediate apart from *p*-coumaric acid during the time course of the reaction with cinnamic acid as substrate [94]. Due to the chemical similarity between *p*- and *m*-coumaric acid, it is possible that both cannot be distinguished with the established HPLC method. Besides, such intermediates were not observed to accumulate in high concentrations and thus were efficiently and consecutively converted into caffeic acid. Following notation from the pathway overview in Figure 1.3, the general activity of the CYP199A2 variants on cinnamic acid will be referred to as C4H activity.

In the end, the NΔ7-CYP199A2 variant is found to be susceptible to inhibition by L-phenylalanine resulting in reduced C3H activity and completely inhibited C4H activity. With respect to usage of L-phenylalanine as starting substrate in the synthetic pathway and also to later feeding experiments in complex medium, which contains high amounts of L-phenylalanine, CYP199A2 is the C3H/C4H candidate of choice, as this enzyme is not inhibited in the presence of L-phenylalanine.

## II.3 Upstream pathway

The synthetic phenylpropanoid pathway is divided into upstream and downstream synthetic pathway for separate individual characterizations with regard to the most efficient substrate to product conversion in *E. coli* (Figure 1.3). Both, the bacterial cytochrome P450

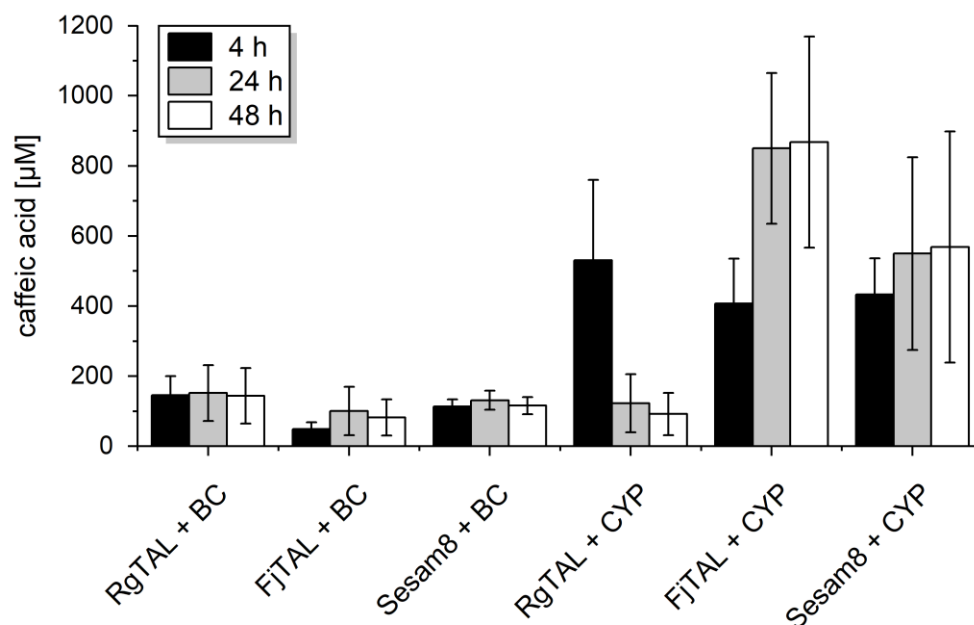
monooxygenase CYP199A2 and the non-P450 hydroxylase HpaBC complex were chosen for evaluations in the upstream pathway. The direct production of caffeic acid is investigated from the initial substrates L-tyrosine (chapter II.3.1) and L-phenylalanine (chapter II.3.2) using different TAL/PAL enzymes.

### **II.3.1 The combined *in vivo* activity of TAL and C3H enzymes**

Numerous reports had focused on the usage of *Rg*TAL and HpaBC as enzyme combination for the conversion of L-tyrosine into caffeic acid [102,133,137,167,168]. With the aim to potentially find a more efficient combined activity of TAL and C3H enzyme, two other TALs, *Fj*TAL and Sesam8, as well as CYP199A2 with its native redox partners Pux/PuR were investigated. For the combination of CYP199A2 with TAL, both proteins were expressed from separate plasmids, as this setup had shown an improved production of caffeic acid compared to an expression from a single Duet vector in either pseudo-operon or operon gene configuration [132]. Thus, the TAL enzymes were expressed from the pETM6 vector and the genes HpaB and HpaC were transferred to the pCDM4 vector in operon configuration. The pCDM4 vector as a derivative of pCDFDuet has the same copy plasmid number and presumably yields similar *in vivo* expression levels of both C3H enzymes in the end [152,169].

In whole-cell biotransformation with 1 mM L-tyrosine, all six combinations of TAL and C3H led to the production of caffeic acid although the values with HpaBC were generally 2.6 to 9.5 times lower compared to CYP199A2 (Figure 2.15). After 4 h of substrate addition, the caffeic acid titer only improved minimally to concentrations of  $151.9 \pm 79.7 \mu\text{M}$  for *Rg*TAL,  $100.9 \pm 68.9 \mu\text{M}$  for *Fj*TAL and  $131.2 \pm 27.2 \mu\text{M}$  for Sesam8 with HpaBC at 24 h. So, the maximal mean caffeic acid titer when using the HpaBC complex was obtained in combination with *Rg*TAL.

With CYP199A2 as C3H, the concentration of caffeic acid increased more significantly from 4 h to 48 h in the resting cell assay. The highest titer of caffeic acid for the bacterial cytochrome P450 monooxygenase was determined in combination with *Fj*TAL after 48 h ( $858.0 \pm 301.5 \mu\text{M}$ ), which equaled almost full substrate conversion. The mean caffeic acid concentration formed by CYP199A2 and *Fj*TAL was thereby 63.6 % larger than the corresponding highest measured titer with *Rg*TAL after 4 h ( $530.6 \pm 229.8 \mu\text{M}$ ). Sesam8 and CYP199A2 exhibited a slightly higher production of  $568.0 \pm 329.7 \mu\text{M}$  caffeic acid after 48 h in comparison to *Rg*TAL with the P450 enzyme.

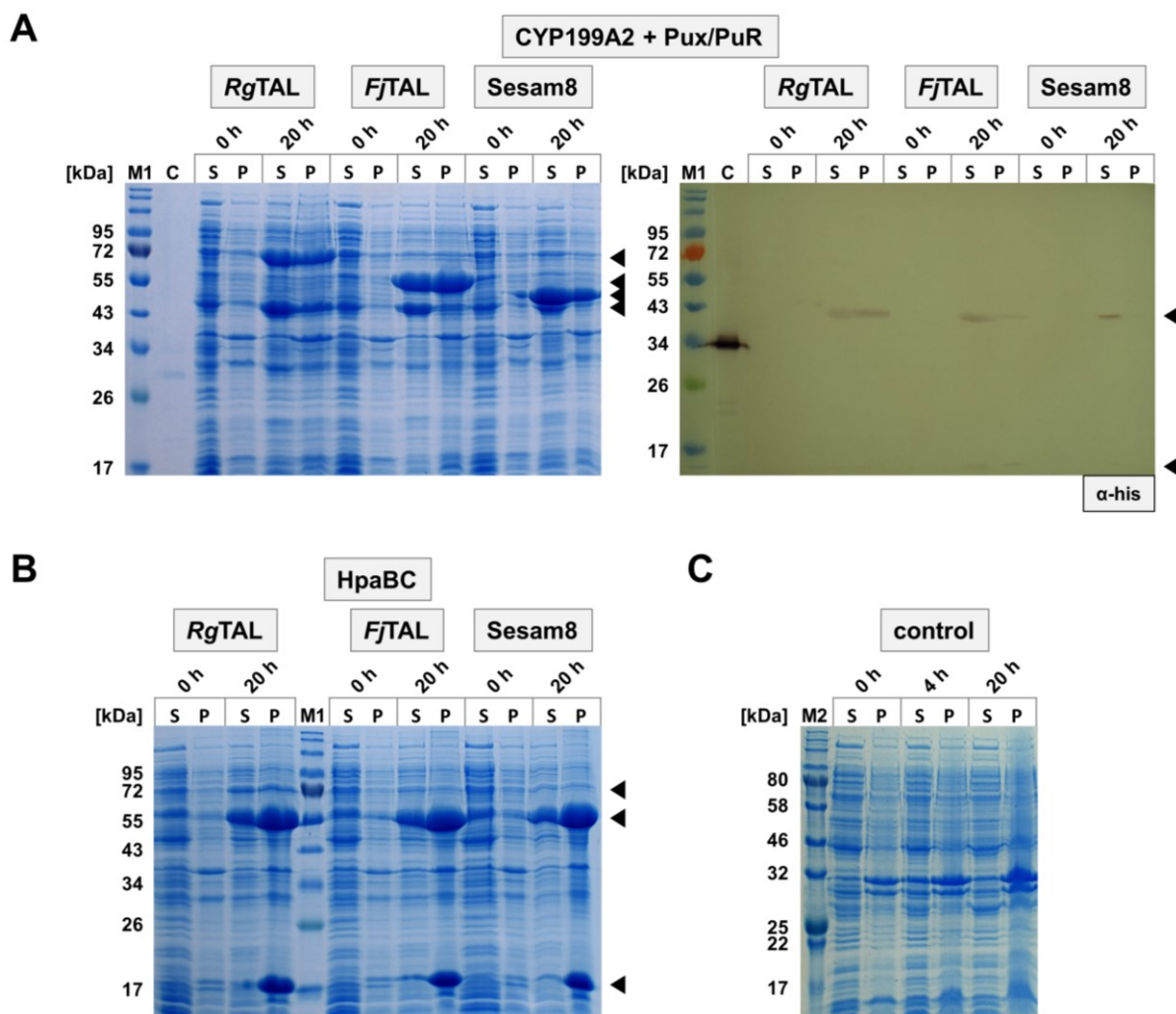


**Figure 2.15: Combined *in vivo* activity of TAL and C3H enzymes in *E. coli* BL21(DE3).** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM L-tyrosine as substrate added at time  $t_0$ . The product concentration of caffeic acid was measured over time (4 h, 24 h, 48 h). Co-expression was performed at 26 °C prior to application in whole-cell biotransformation using pETM6 with TAL (*RgTAL*, *FjTAL*, *Sesam8*) and either pCDM4 with HpaBC in operon gene configuration (BC) or pCDFDuet with CYP199A2 (CYP) and pRSFDuet with Pux/PuR as redox pair. The shown data represent mean values with standard deviation from biological triplicates.

In contrast to the other two combinations, the caffeic acid titer decreased significantly after 4 h, when *RgTAL* was used with CYP199A2, and resulted in values which were similar in the TAL combinations with HpaBC. A possible explanation for this decrease could be that caffeic acid was consumed by endogenous enzymes of metabolically active *E. coli* cells. This hypothesis originates from the observation in a toxicity test, in which growing cultures of *E. coli* turned dark-brown over time in the presence of 10 mM caffeic acid (Figure A.19 A and F, Appendix). Interestingly, these cultures were able to grow after 24 h, which was not observed for the addition of 10 mM *p*-coumaric or cinnamic acid. As an alternative explanation, *RgTAL* could have converted caffeic acid into another side-product such as L-DOPA, which had been reported as a substrate for some TALs [57,74]. It is noteworthy that *p*-coumaric acid accumulated in low amounts exclusively in the combination *RgTAL* and CYP199A2 after 24 h and 48 h (Figure A.4, Appendix). This observation indicates residual activity of *RgTAL* on L-tyrosine at 24 h. Regarding the final titers at 48 h, the caffeic acid production by CYP199A2 with *FjTAL* and *Sesam8* were 9.4- and 6.2-fold higher with respect to the titers when using *RgTAL*.

In the SDS-PAGE analysis from the co-expressions of the cytochrome P450 enzyme with all three TALs, distinct protein bands could be observed for *RgTAL* around 74.8 kDa, for *FjTAL*

at 56.7 kDa and for Sesam8 at 54.0 kDa after induction of protein synthesis (Figure 2.16 A). The expression band slightly above 43 kDa corresponded to the overlay of the proteins CYP199A2 and the reductase PuR. The latter along with the ferredoxin Pux were detected in western blot analysis around 44.5 kDa and 12.6 kDa, respectively, but with weak stained bands for Pux.



**Figure 2.16: Co-expression analysis of the TAL enzymes *RgTAL*, *FjTAL* and *Sesam8* in combination with the C3Hs CYP199A2 (A) and HpaBC (B).** Co-expression was performed using pETM6\_TAL with *RgTAL*, *FjTAL* or *Sesam8*, pCDFDuet\_CYP199A2 and pRSFDuet\_his<sub>6</sub>-Pux\_his<sub>6</sub>-PuR (CYP199A2, A) or with pCDM4\_HpaB\_HpaC\_op (HpaBC, B) in *E. coli* BL21(DE3) at 26 °C. 5/OD samples were taken at the point of induction with 1 mM IPTG (0 h) and at the end of expression (20 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE and western blot analysis. For detection in western blot analysis the anti-His ( $\alpha$ -his) and GAM<sub>FC</sub><sup>AP</sup> antibodies were used. An expression of the empty vectors (pETM6, pCDFDuet, pRSFDuet) served as **control** (C). The black arrowheads indicate the relevant expression bands. M1 – P7719 protein ladder (NEB), M2 – P7712 protein ladder (NEB), C – mCherry-his<sub>6</sub> as His<sub>6</sub>-tag control (0.5  $\mu$ g, 27.7 kDa), S – supernatant, P – pellet fraction, *RgTAL* – 74.8 kDa, *FjTAL* – 56.7 kDa, *Sesam8* – 54.0 kDa, CYP199A2 – 44.6 kDa, his<sub>6</sub>-Pux – 12.6 kDa, his<sub>6</sub>-PuR – 44.5 kDa, HpaB – 58.9 kDa, HpaC – 18.5 kDa.

In combination with the HpaBC complex, only *RgTAL* showed weak expression bands in the SDS-PAGE analysis (Figure 2.16 B). Due to no overexpression bands from *Sesam8* and *FjTAL*

as in the co-expressions with CYP199A2, it was suspected that the bands around 55 kDa mainly correspond to HpaB (58.9 kDa) and not *Fj*TAL. Thus, both C3H components HpaB and HpaC (18.5 kDa) were expressed in pronounced bands, even if the HpaC protein seemed to be less expressed in the soluble fraction.

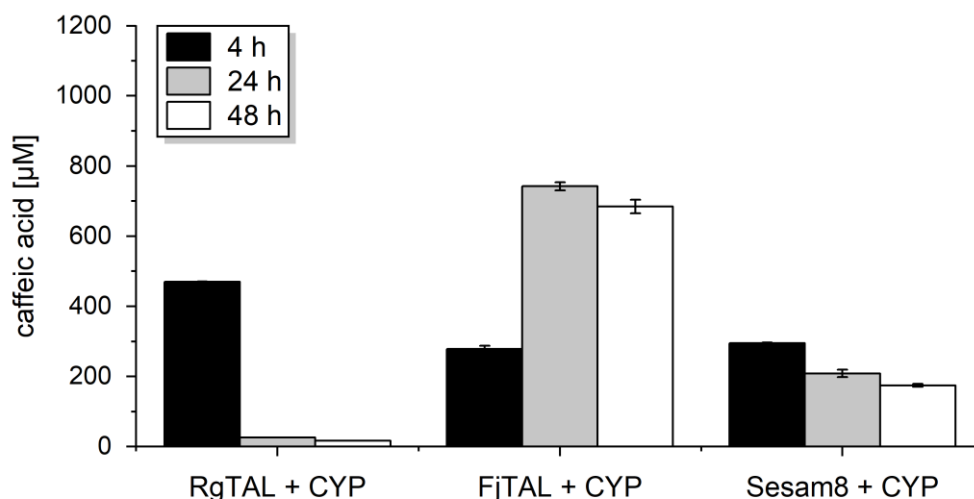
Consequently, the HpaBC complex was overexpressed in large quantities, although the plasmid copy number of the used pCDM4 vector is lower and the resulting T7 promoter strength was reported to be only 38 % of that of the previously used pETM6 vector [152]. This high production of the HpaBC complex presumably led to a marginal co-expression of the TAL enzyme from the pETM6 plasmid. Nonetheless, the transcription efficiency of all genes should be equal, as all used ePathBrick and Duet vectors solely contain T7 promoters. The observed significant difference suggest that the translation of mRNA transcripts might have been more pronounced for HpaBC than for the TAL proteins.

In the end, the imbalance in produced protein amounts between the HpaBC and TALs in co-expressions is likely to be responsible for the low activity of these combinations in whole-cell biotransformations. As a result of that, the majority of L-tyrosine was probably not converted due to the lack of sufficient TAL activity. In the past, the usage of HpaBC as C3H was accompanied by the enzyme's side activity to accept L-tyrosine as substrate and convert it into L-DOPA [101]. Thus, L-tyrosine could have been lost for supply into the synthetic pathway due to the possible generation of L-DOPA as by-product. There are many reports describing a dark brown coloration of growing *E. coli* cell cultures expressing HpaBC, which is indicative of polymeric melanin formation as a result of L-DOPA overoxidation [102,133,170–172]. The mentioned browning of *E. coli* cultures containing the HpaBC complex has been also observed in this work and pointed towards side-product formation.

Consequently, the later combined synthetic pathways and growing cell assays focus on CYP199A2 instead of HpaBC as C3H (chapters II.5 and II.6), as the aim is to yield efficient production levels of caffeic acid, balanced expression levels of TAL and C3H, and to avoid potential side-products.

In order to evaluate a potential increase in substrate feed, 3 mM of L-tyrosine were added to the combinations of CYP199A2 with all three TAL enzymes in whole-cell biotransformations (Figure 2.17). Instead of a further product increase, the concentrations for caffeic acid were determined to be lower compared to a substrate feed of 1 mM L-tyrosine. The highest production of caffeic acid using *Rg*TAL was at  $469.6 \pm 1.6 \mu\text{M}$  after 4 h and for *Fj*TAL at  $742.2 \pm 11.3 \mu\text{M}$  after 24 h. Interestingly, the titer of caffeic acid for Sesam8 was highest at 4 h

with a value of  $295.4 \pm 1.2 \mu\text{M}$ , which equaled a decrease of 31.8 % compared to a feed of 1 mM L-tyrosine. In contrast to the results of Sesam8 with 1 mM L-tyrosine, the concentration of caffeic acid decreased after this time point until a final difference in titers of 69.3 % was reached after 48 h. For both other TALs, the decreases in caffeic acid production were less pronounced being 11.5 % for *RgTAL* at 4 h and 12.7 % for *FjTAL* at 24 h.



**Figure 2.17:** Combined *in vivo* activity of TAL enzymes with CYP199A2 in *E. coli* BL21(DE3) using a higher substrate feed. Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 3 mM L-tyrosine as substrate added at time  $t_0$ . The product concentration of caffeic acid was measured over time (4 h, 24 h, 48 h). Co-expression was performed at 26 °C prior to application in whole-cell biotransformation using pETM6 with TAL (*RgTAL*, *FjTAL*, Sesam8), pCDFDuet with CYP199A2 (CYP) and pRSFDuet with the redox pair Pux/PuR. The shown data represent mean values with standard deviation from biological duplicates.

The results indicate that a higher substrate supply can inhibit the upstream pathway of TAL and C3H. However, the significant drop in pH, when 3 mM L-tyrosine were applied, might have contributed to the reduced *in vivo* enzyme activities (chapter II.1.2). The intermediate *p*-coumaric acid was barely measurable for all three resting cell assays and this could indicate that the shift in pH resulted in a decreased activity of the TAL enzymes (Figure A.5, Appendix).

In the end, the expression of multiple pathway enzymes may become a detrimental factor in the final shake-flask expressions of the complete synthetic pathway, as it was observed with the above described co-expression of the HpaBC complex and the TALs. Thus, for further investigations regarding the production towards coniferyl alcohol, CYP199A2 is applied together with *FjTAL* or Sesam8 due to the measured high *in vivo* activities as well as sufficient and balanced protein production levels in these enzyme combinations.

### II.3.2 The combined *in vivo* activity of PAL and C3H/C4H enzyme

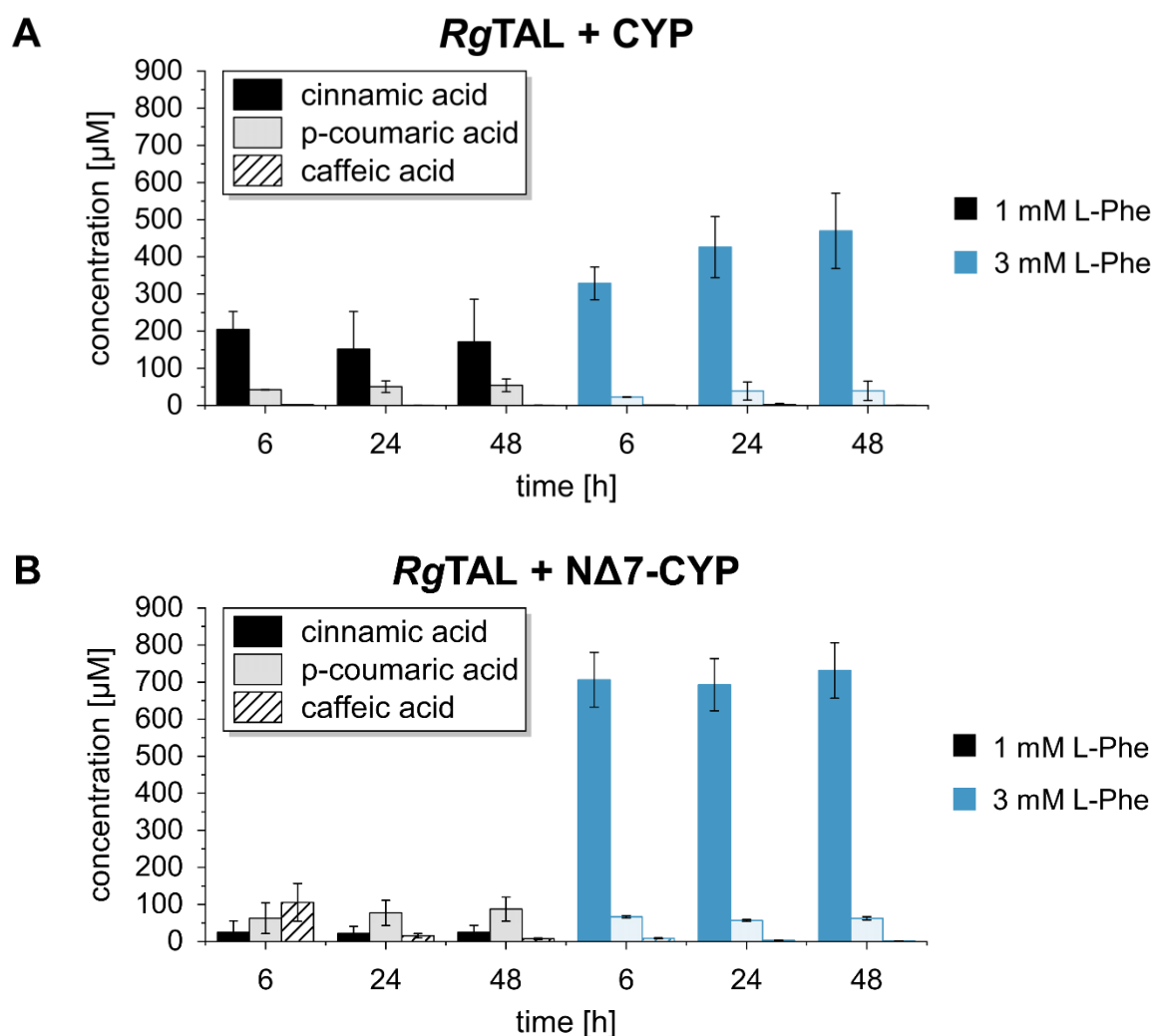
In the previous results, *RgTAL* also functioned as PAL in *E. coli in vivo* (cf. chapter II.1.3). Hence, this enzyme candidate was combined with CYP199A2 and NΔ7-CYP199A2 as

C3H/C4H with the intention to complete the upstream pathway from L-phenylalanine to caffeic acid. As a substrate, 3 mM L-phenylalanine were fed in addition to the whole-cell biotransformation performed with 1 mM L-phenylalanine in order to potentially improve the production of cinnamic acid by RgTAL.

In general, the accumulation of cinnamic acid and small amounts of *p*-coumaric acid were observed over time (Figure 2.18). The cinnamic acid production increased along with a substrate feed of 3 mM L-phenylalanine instead of 1 mM, as observed before in the PAL *in vivo* assay with RgTAL (chapter II.1.3). But the sum of all product concentrations at a certain time point gave a less total concentration compared to the previously determined cinnamic acid concentrations for RgTAL alone. The total product concentration for the combined PAL and C3H/C4H reactions at 24 h was decreased by 46.9 % with 1 mM L-phenylalanine and by 45.2 % with 3 mM L-phenylalanine using CYP199A2 and by 70.1 % and 11.8 % using NΔ7-CYP199A2, respectively (Figure 2.18).

RgTAL in combination with the NΔ7 variant was notably the sole combination, which equaled the product concentration of the single RgTAL reaction with 3 mM L-phenylalanine after 48 h (Figure 2.18 B). At the same time, this combination of enzymes exhibited the only production of caffeic acid with a value of  $105.5 \pm 51.0 \mu\text{M}$  after 6 h and a feed of 1 mM L-phenylalanine. The total product concentrations at 6 h between CYP199A2 and NΔ7-CYP199A2 were similar and thus only the ratio of products was shifted. The titer for caffeic acid decreased after that time point and probably led to the apparent strong reduction in total product concentrations. The loss of caffeic acid had been reported before and the observed decrease in titer might be attributed again to the endogenous metabolization by *E. coli* (chapter II.3.1) [173].

As another remark, the measured concentrations for *p*-coumaric acid with NΔ7-CYP199A2 were higher than those with the full length P450 enzyme. Apparently, the NΔ7 variant of CYP199A2 was able to convert *p*-coumaric acid partly into caffeic acid, although this enzyme had been previously identified to be inhibited by 1 mM L-phenylalanine (chapter II.2.2). It seems that both P450 enzymes were mainly inactive especially with the higher substrate concentration, the simultaneous low caffeic acid titer and the pronounced accumulation of cinnamic acid. This can be a result of unspecific inhibition effects by L-phenylalanine on CYP199A2 and its NΔ7 variant. Regarding protein levels of the individual co-expressions, no differences were observed between TAL, redox partners Pux and PuR and the applied P450 enzyme (cf. Figure 2.16 A, Figure A.6, Appendix).



**Figure 2.18: Combination of *RgTAL* as PAL and C3H/C4H activities of CYP199A2 (A) and NΔ7-CYP199A2 (B) in *E. coli* BL21(DE3).** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM (black) or 3 mM L-phenylalanine (blue, L-Phe) as substrate added at time  $t_0$ . The concentration of substances was measured over time (6 h, 24 h, 48 h). Co-expression was performed at 26 °C prior to application in whole-cell biotransformation using pETM6 with *RgTAL*, pCDFDuet with either CYP199A2 (CYP, A) or NΔ7-CYP199A2 (NΔ7-CYP, B) and pRSFDuet with the redox pair Pux/PuR. The shown data represent mean values with standard deviation from at least biological duplicates.

Taken together, these results show that the combination of *RgTAL* and the NΔ7 variant of the cytochrome P450 monooxygenase performed slightly superior compared to the combination with CYP199A2 due to higher total product concentrations and the intended formation of caffeic acid. At the same time, a later synthetic pathway starting from L-phenylalanine seems unlikely, as insufficient product titers were already attained in the tested upstream pathway of PAL and C3H/C4H. Consequently, the complete synthetic pathway with cinnamic acid as starting substance is investigated in chapter II.6 in growing cells instead.

It is noteworthy that the presented results of *RgTAL* with NΔ7-CYP199A2 as bacterial C3H/C4H constitute the first report of L-phenylalanine conversion into caffeic acid in *E. coli* to date and to the best of the author's knowledge.



## II.4 Downstream pathway

After evaluation of the upstream module from the synthetic coniferyl alcohol pathway consisting of TAL and C3H, the downstream module comprising 4CL, CCR, CAD and OMT is evaluated in more detail. Thereby, the main focus lays on the choice of methyltransferase regarding plant origin and substrate preference. The caffeic acid *O*-methyltransferase (*ZmCOMT*) and caffeoyl-CoA *O*-methyltransferase (*ZmCCoAOMT*) from *Z. mays* are investigated, as the applied CCR and CAD enzyme are also from *Z. mays* and it is supposed that enzymes from the same plant origin might exhibit a more favorable overall activity. For comparison, the COMT and CCoAOMT from *A. thaliana* are evaluated as well, because these two methyltransferases have been already applied successfully for the production of phenylpropanoids in *E. coli* [137].

In general, the enzymes 4CL, CCR and CAD are responsible for the conversion from phenylpropanoic acid via CoA ester and aldehyde to the alcohol form. The 4CL, CCR and CAD in combination were not only applied for phenylpropanoic acids but had been also tested for a multitude of alternative substrates and found to be functionally active [173,174]. Depending on the individual 4CL enzyme, significant differences in substrate specificity have been observed before and thus two 4CL enzyme candidates are evaluated in the synthetic downstream pathway as well.

In the following sections, *in vivo* production of coniferyl alcohol is examined from caffeic acid and in two different *E. coli* strains.

### II.4.1 Different methyltransferases expressed in a genome-integrated *E. coli* strain

The *E. coli* BL21(DE3).G213 strain was a gift from Phytowelt GreenTechnologies GmbH (Nettetal, Germany) and contains the genome-integrated *p*-coumaryl alcohol pathway. Thereby, the genes for *RsTAL-Pc4CL* and *ZmCCR-ZmCAD* were integrated as two individual operons into the genome of *E. coli* BL21(DE3) and put under the control of weak constitutive promoters. This *E. coli* BL21(DE3).G213 strain was used for initial evaluation of the four methyltransferases. The additional plasmid-based expression of the methyltransferases from the pCDFDuet vector and the resulting downstream pathway activity were investigated at two different temperatures.

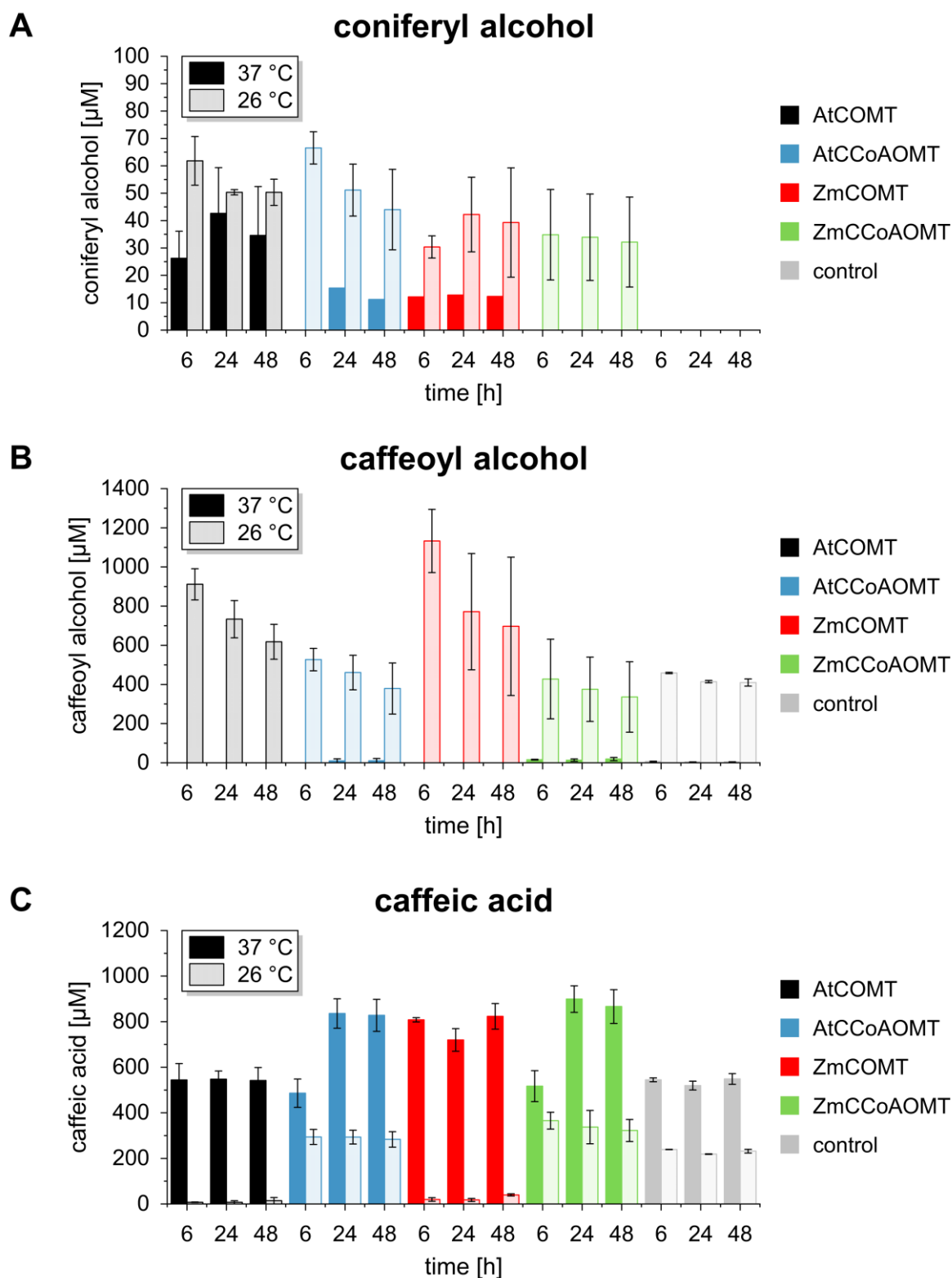
In the whole-cell biotransformation with 1 mM caffeic acid, all methyltransferases were found to be active when expressed at 26 °C and generated higher concentrations of coniferyl alcohol

compared to an expression temperature of 37 °C (Figure 2.19 A). The highest coniferyl alcohol titer was measured for *AtCOMT* with a value of  $42.7 \pm 16.7 \mu\text{M}$  after 24 h, when expression was performed at 37 °C. Apart from a marginal coniferyl alcohol production when using *AtCCoAOMT* and *ZmCOMT* expressed at 37 °C, no coniferyl alcohol was detected with *ZmCCoAOMT*. For an expression at 26 °C, the highest titers for coniferyl alcohol were  $61.8 \pm 8.9 \mu\text{M}$  for *AtCOMT* (6 h),  $66.5 \pm 5.9 \mu\text{M}$  for *AtCCoAOMT* (6 h),  $42.2 \pm 13.6 \mu\text{M}$  for *ZmCOMT* (24 h) and  $34.9 \pm 16.5 \mu\text{M}$  for *ZmCCoAOMT* (6 h). As expected, the empty vector control did not produce any coniferyl alcohol, though  $458.8 \pm 3.3 \mu\text{M}$  caffeoyl alcohol (6 h) were generated due to the activity of the genome-integrated enzymes *Pc4CL*, *ZmCCR* and *ZmCAD*.

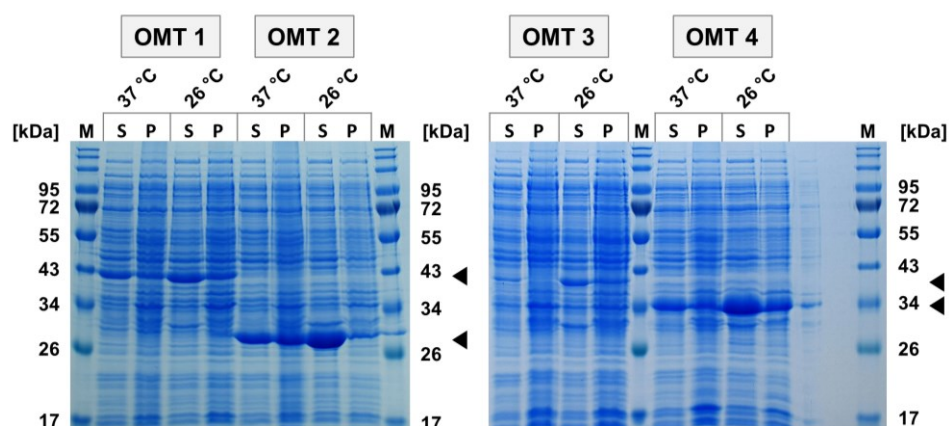
Interestingly, caffeoyl alcohol was solely measured for an expression at 26 °C but not at 37 °C at which caffeic acid accumulated in high titers instead (Figure 2.19 B and C). This observation was generally made for all tested G213 strains with either empty vector control or expressed methyltransferase. The accumulation of caffeic acid indicates that the substrate was not converted and that *Pc4CL* was probably not expressed as an active enzyme at 37 °C (Figure 2.19 C). Regarding an expression at 26 °C, caffeic acid was completely consumed except for the G213 strain containing the CCoAOMT enzymes and the empty vector control.

Accordingly, the strains with complete substrate consumption also showed the highest titers for caffeoyl alcohol with  $911.6 \pm 79.5 \mu\text{M}$  for *AtCOMT* and  $1132.7 \pm 161.6 \mu\text{M}$  for *ZmCOMT* after 6 h (Figure 2.19 B). In contrast, the G213 strain with *AtCCoAOMT* ( $527.0 \pm 57.1 \mu\text{M}$ ) exhibited a 42 % lower caffeoyl alcohol production and with *ZmCCoAOMT* ( $427.7 \pm 203.4 \mu\text{M}$ ) a 63 % lower titer compared to the respective COMT enzymes. All caffeoyl alcohol titers decreased after 6 h of whole-cell biotransformation and no increase in coniferyl alcohol concentrations was observed. This indicates that the 4CL enzyme may have become inactive after 6 h as a result of no further conversion of residual caffeic acid.

With view on the expression of the methyltransferases in the genome-integrated G213 strain, protein bands were observed for all OMTs at both analyzed expression temperatures (Figure 2.20). All methyltransferases showed more pronounced expression bands in the soluble fractions and at 26 °C in the SDS-PAGE analysis. The only exception was formed by *AtCOMT*, which seemed to be expressed in similar distinct bands around 39.6 kDa at both temperatures (OMT 1). The protein bands for *ZmCOMT* (OMT 3) located at the same molecular weight were less pronounced compared to the *AtCOMT*. In general, the CCoAOMT enzymes at 25.9 kDa (OMT 2) and 29.4 kDa (OMT 4) showed bands of greater intensity than the COMT proteins.



**Figure 2.19: *In vivo* activity of different methyltransferases in *E. coli* BL21(DE3).G213 at two different expression temperatures.** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM caffeic acid as substrate added at time  $t_0$ . The concentration of coniferyl alcohol (A), caffeoyl alcohol (B) and caffeic acid (C) were measured over time (6 h, 24 h, 48 h). Expression was performed either at 37 °C (filled) or at 26 °C (pale) prior to application in whole-cell biotransformation using pCDFDuet\_OMT with *At*COMT (black), *At*CCoAOMT (blue), *Zm*COMT (red) or *Zm*CCoAOMT (green). A whole-cell biotransformation using the empty vector served as activity control (control, grey). The shown data represent mean values with standard deviation from at least biological duplicates.



**Figure 2.20: SDS-PAGE analysis of different methyltransferases in the *E. coli* BL21(DE3).G213 strain.** Expression was performed from pCDFDuet\_OMT with *At*COMT (OMT 1), *At*CCoAOMT (OMT 2), *Zm*COMT (OMT 3) and *Zm*CCoAOMT (OMT 4) in *E. coli* BL21(DE3).G213 at either 37 °C or 26 °C. 5/OD samples were taken at 20 h and divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE analysis. The black arrowheads indicate the relevant expression bands. M – P7719 protein ladder (NEB), S – supernatant, P – pellet fraction, *At*COMT – 39.6 kDa, *At*CCoAOMT – 25.9 kDa, *Zm*COMT – 39.6 kDa, *Zm*CCoAOMT – 29.4 kDa.

The expression of the genome-integrated enzymes was not visible as distinct protein bands in the SDS-PAGE analysis and not observed in the empty vector control, either (data not shown). The results show that 26 °C is the more favorable expression temperature for the majority of the methyltransferases, regarding both sufficient protein production and combined downstream pathway activity with 4CL, CCR and CAD.

It has to be noted that coniferaldehyde accumulated in small amounts from 78-360  $\mu$ M after 24 h to 48 h for expressions performed at 37 °C (data not shown). Control experiments did not detect any influence of expression temperature on the fast *in vivo* *Zm*CAD activity and residual activity was measureable even after 24 h (data not shown). It is noteworthy that *E. coli* possesses endogenous alcohol dehydrogenases (ADH), which are known to exhibit a broad substrate specificity and function in detoxification reactions including the removal of aldehydes [175,176]. It has been reported that these ADHs were able to accept coniferaldehyde as well as cinnamaldehyde and convert these into the corresponding alcohols [177–180]. Consequently, the general conversion of coniferaldehyde into coniferyl alcohol is considered to be efficient in *E. coli* and not pose a primary bottleneck for the later overall synthetic pathway at 26 °C.

#### II.4.2 Different methyltransferases expressed in *E. coli* BL21(DE3)

With the intention to increase coniferyl alcohol titers of the downstream pathway, plasmid-based production of all enzymes was performed under the control of T7 promoters in *E. coli* BL21(DE3). Thereby, both 4CL enzyme candidates, *Pc*4CL and 4CL from *A. thaliana* (*At*4CL), were examined with all four methyltransferases via expression from the vectors

pETM6\_4CL and pCDFDuet\_OMT. The two other downstream pathway enzymes *ZmCCR* and *ZmCAD* were codon-optimized for *E. coli* and co-expressed from a previously constructed pRSFDuet vector [135].

Indeed, the coniferyl alcohol titers from seven out of eight enzyme combinations of 4CL and OMT were found to be higher than the concentrations obtained in the G213 strain at 26 °C before (cf. Figure 2.19). The highest production of coniferyl alcohol was  $258.6 \pm 31.3 \mu\text{M}$  yielded by the application of *At4CL* and *ZmCOMT*, which exhibited a titer at least twice as high as all other enzyme combinations (Figure 2.21 A). Furthermore, this combination as well as *At4CL* with *AtCOMT* were the only constructs that did consume caffeic acid completely and showed the highest caffeoyl alcohol productions of  $684.4 \pm 140.2 \mu\text{M}$  for *At4CL* + *ZmCOMT* and  $788.7 \pm 104.1 \mu\text{M}$  for *At4CL* + *AtCOMT* (Figure 2.21 B and C). At the same time, both enzyme combinations exhibited the lowest standard deviations for the production of caffeoyl alcohol indicating a balanced activity of the applied 4CL and OMT enzymes.

In all other enzyme combinations, the substrate caffeic acid was partly left unconverted and the caffeoyl alcohol titers fluctuated substantially. In accordance with the high standard deviations for the substrate, the activity performance of the 4CL enzyme might have been suboptimal in the individual combinations with OMT resulting in different levels of consumption of caffeic acid. Furthermore, this may indicate sensitivity of both 4CL enzymes to inhibition or instability of the 4CL activity in particular combinations with OMT over time. Nevertheless, both 4CL candidates were found to accept caffeic acid as substrate equally well in co-expressions with OMT.

The subsequent most active combinations regarding coniferyl alcohol production were *Pc4CL* + *ZmCCoAOMT* ( $127.1 \pm 43.1 \mu\text{M}$ ), *At4CL* + *AtCOMT* ( $122.6 \pm 24.2 \mu\text{M}$ ), *Pc4CL* + *AtCCoAOMT* ( $120.0 \pm 52.3 \mu\text{M}$ ) and *At4CL* + *ZmCCoAOMT* ( $117.3 \pm 31.4 \mu\text{M}$ ) (Figure 2.21 A). In general, it seemed that the combination of CCoAOMT with *Pc4CL* as well as COMT with *At4CL* were most productive for generation of coniferyl alcohol in the end. Interestingly, the combination of *ZmCOMT* with *Pc4CL* instead of *At4CL* exhibited a 8.5-fold lower coniferyl alcohol concentration and represented the most unproductive of all eight variations. This result emphasizes that exchange of a single enzyme candidate can have an enormous impact on the performance of a synthetic pathway. As expected, the empty vector control exhibited no conversion activity towards caffeic acid (Figure 2.21 C).

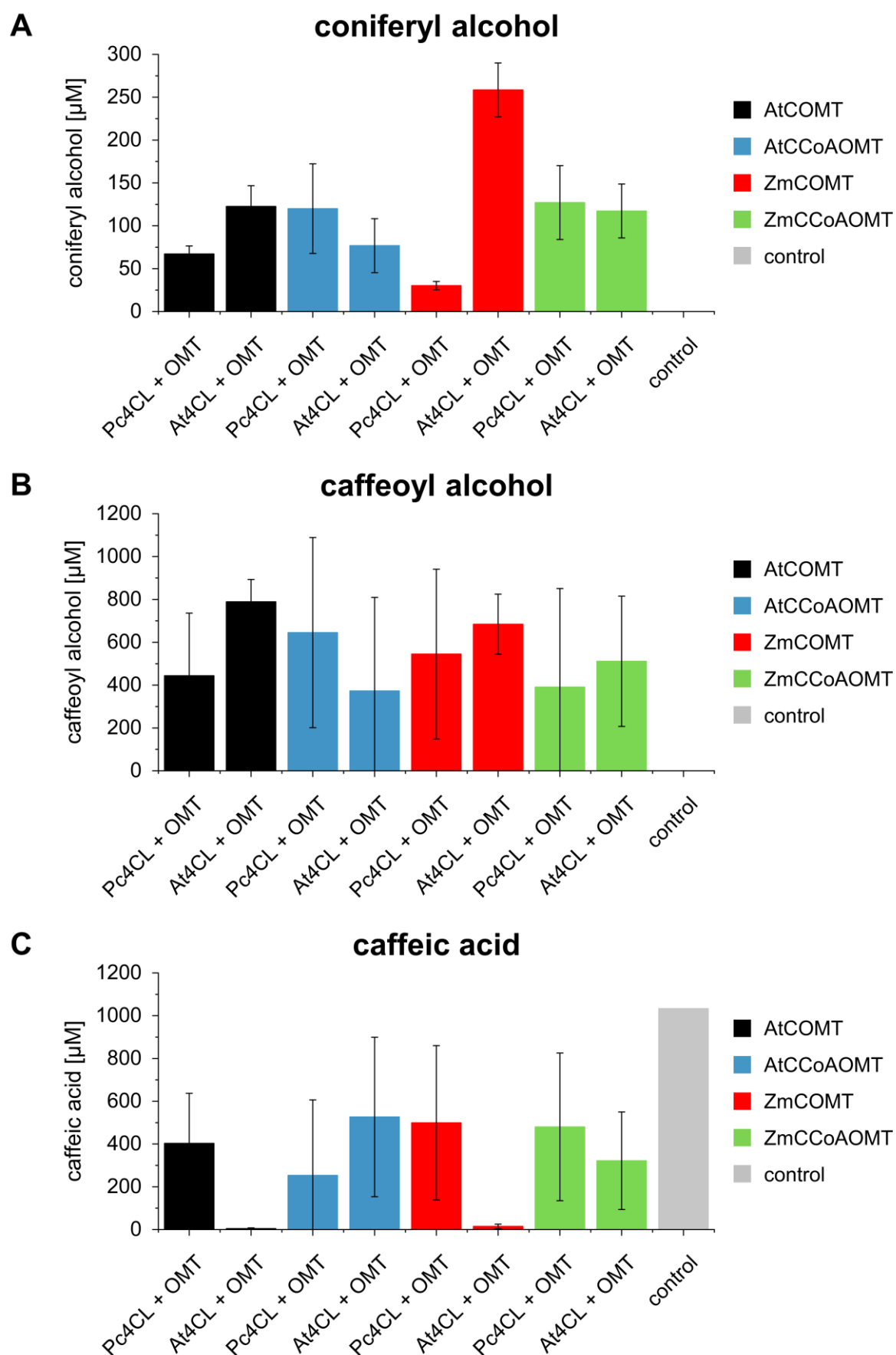


Figure 2.21: Downstream synthetic pathway from caffeic acid towards coniferyl alcohol using plasmid-based enzyme expression in *E. coli* BL21(DE3). Whole-cell biotransformation was performed in 10 mL KPi buffer

(50 mM, pH 7.5) at 26 °C with 1 mM caffeic acid as substrate added at time  $t_0$ . The concentration of coniferyl alcohol (**A**), caffeoyl alcohol (**B**) and caffeic acid (**C**) were measured after 48 h. Expression was performed at 26 °C prior to application in whole-cell biotransformation. Different OMTs (pCDFDuet\_OMT) being *At*COMT (black), *At*CCoAOMT (blue), *Zm*COMT (red) or *Zm*CCoAOMT (green) and both 4CL enzymes (pETM6\_4CL) were combined with the pRSFDuet\_4CL\_ZmCCR\_ZmCAD vector. A single whole-cell biotransformation using the empty vectors served as activity control (control, grey). The shown data represent mean values with standard deviation from at least biological triplicates.

Regarding the individual *in vivo* expression levels of the applied enzymes in the eight combinations, distinct differences between 4CL and OMT proteins could be observed. Firstly, *Pc*4CL was expressed in pronounced bands around 60.3 kDa in co-expression with every methyltransferase assessed from the SDS-PAGE analysis (Figure 2.22). An additional band above 46 kDa in the insoluble fractions is observed in expressions with *Pc*4CL, possibly indicating partly degradation of this 4CL enzyme candidate. Secondly, co-expressions with *At*4CL showed prominent bands at 61.1 kDa for 4CL in combination with both COMT enzymes and after 20 h on the one hand (Figure 2.22, 1 and 3). On the other hand, thin protein bands for *At*4CL were visible on the SDS gel, when the 4CL enzyme was co-expressed with both CCoAOMT proteins (Figure 2.22, 2 and 4). Thirdly, the COMT enzymes exhibited more pronounced bands at 39.6 kDa in co-expression with *At*4CL, while both seemed to be barely expressed in combination with *Pc*4CL (Figure 2.22, 1 and 3). *At*COMT was thereby expressed in more prominent bands compared to the COMT enzyme from *Z. mays*. And lastly, *Zm*CCoAOMT seemed to be expressed in similar amounts with both 4CL enzymes judging from the band intensities around 29.4 kDa (Figure 2.22, 4). However, co-expressions with *At*CCoAOMT showed more pronounced bands at 25.9 kDa with *At*4CL compared to *Pc*4CL (Figure 2.22, 2).

Detection in western blot analysis was performed by using a specific primary antibody for the N-terminal His<sub>6</sub>-tag of *Zm*CCR and the C-terminal S-tag of *Zm*CAD. In general, *Zm*CCR and *Zm*CAD proteins seemed to be produced at similar levels in all eight 4CL and OMT combinations. Distinct bands were detected in western blot analysis around 41.9 kDa for his<sub>6</sub>-*Zm*CCR and at 41.3 kDa for *Zm*CAD-S-tag, while only thin bands were visible for both proteins in SDS-PAGE analysis. Additionally, two bands were co-detected around and below 32 kDa in the insoluble fractions on the blot stained against the S-peptide. These bands suggest protein degradation of *Zm*CAD and were found to occur more pronounced in the co-expressions with *Pc*4CL compared to those with *At*4CL.



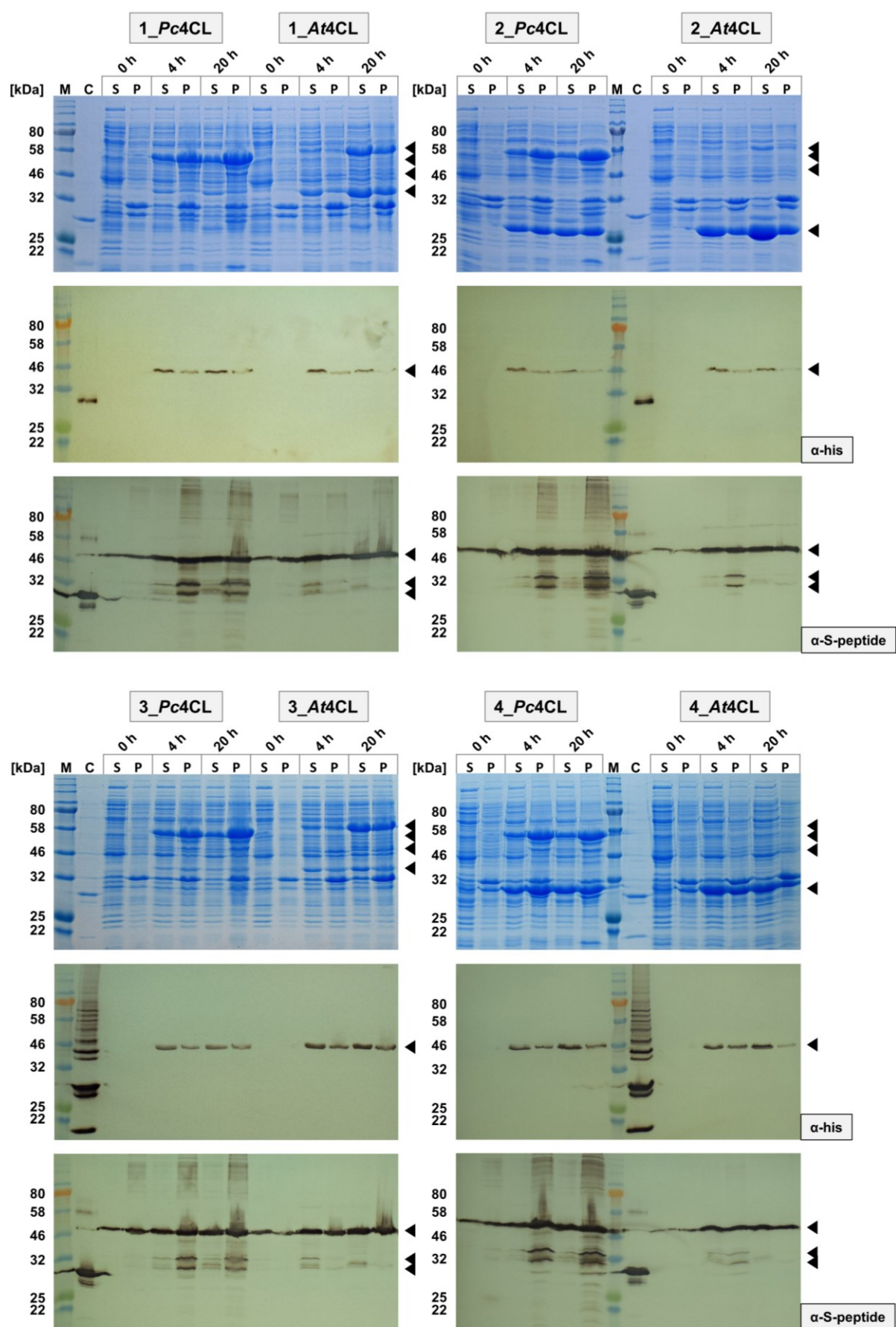


Figure 2.22: Co-expression analysis of the downstream pathway using the four methyltransferases (1-4) in combination with a 4CL enzyme (*Pc4CL* or *At4CL*) and *ZmCCR* and *ZmCAD*. Co-expression was performed using



pETM6\_4CL (*Pc4CL* or *At4CL*), pRSFDuet\_his<sub>6</sub>-*ZmCCR\_ZmCAD*-S-tag and pCDFDuet\_OMT with *AtCOMT* (1), *AtCCoAOMT* (2), *ZmCOMT* (3) or *ZmCCoAOMT* (4) in *E. coli*/BL21(DE3) at 26 °C. 5/OD samples were taken at the point of induction with 1 mM IPTG (0 h) and during the expression (4 h and 20 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE and western blot analysis. For detection in western blot analysis the anti-His (α-his) or S-Peptide Epitope tag (α-S-peptide) and GAM<sub>FC</sub><sup>AP</sup> antibodies were used. The black arrowheads indicate the relevant expression bands. M – P7712 protein ladder (NEB), C – (impure) mCherry-his<sub>6</sub> as His<sub>6</sub>-tag control (0.5 µg, 27.7 kDa) or PVX-S-tag as S-tag control (0.25 µg, 27 kDa), S – supernatant, P – pellet fraction, *AtCOMT* – 39.6 kDa, *AtCCoAOMT* – 25.9 kDa, *ZmCOMT* – 39.6 kDa, *ZmCCoAOMT* – 29.4 kDa, *Pc4CL* – 60.3 kDa, *At4CL* – 61.1 kDa, his<sub>6</sub>-*ZmCCR* – 41.9 kDa, *ZmCAD*-S-tag – 41.3 kDa.

It shall be remarked that an impure mCherry-his<sub>6</sub> solution was used as His<sub>6</sub>-tag control in some experiments. This control showed multiple bands assumably resulting from an unpurified cell lysate containing unspecific degradations of mCherry and protein associations of his<sub>6</sub>-tagged fragments. Nonetheless, this impure protein lysate was found applicable for serving as His<sub>6</sub>-tag control.

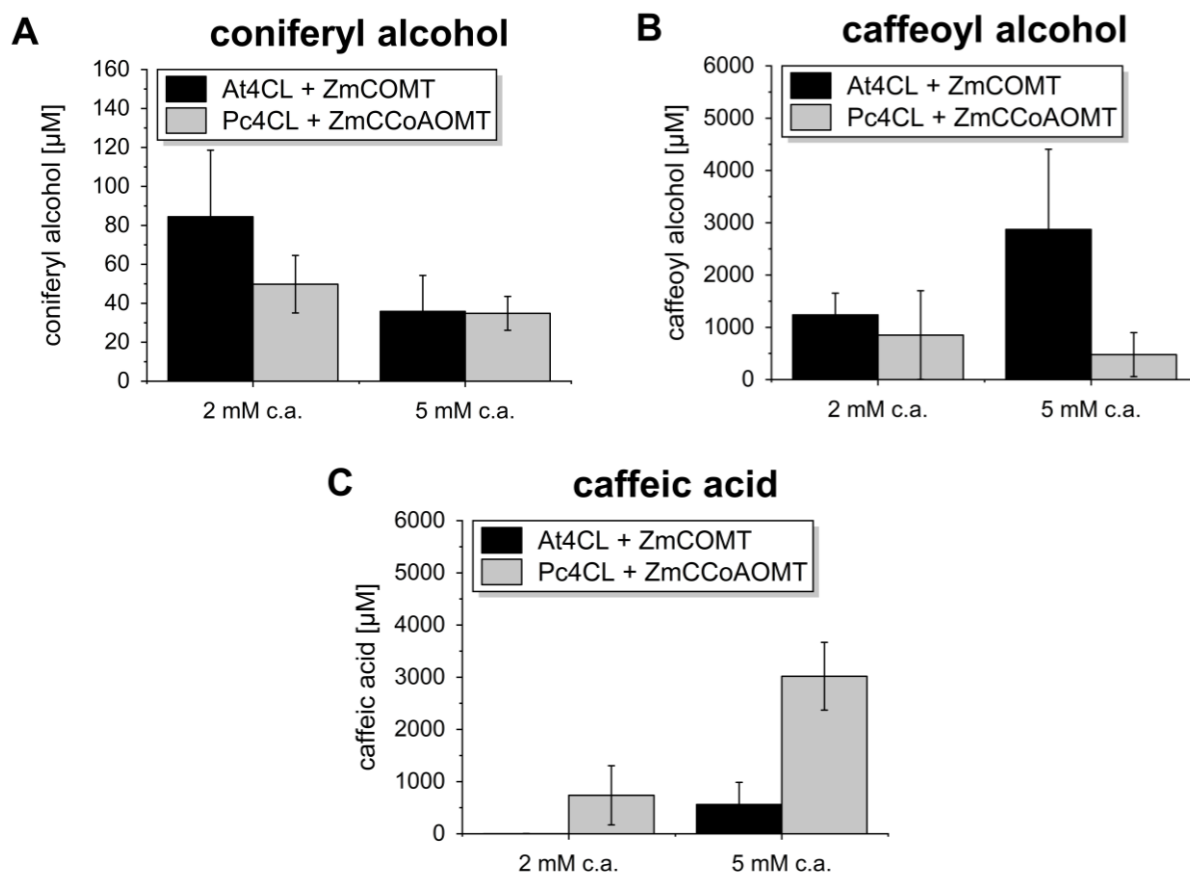
### II.4.3 Increased substrate feed into the downstream pathway

The plasmid-based production of *ZmCCR* and *ZmCAD* in combination with either *At4CL* and *ZmCOMT* or *Pc4CL* and *ZmCCoAOMT* in *E. coli* BL21(DE3) is investigated further with respect to higher substrate feeds. Both enzyme combinations are tested, because they exhibited the highest coniferyl alcohol titers in the previous whole-cell biotransformation with 1 mM caffeic acid (chapter II.4.2). Furthermore, the evaluation serves for comparison of the two novel methyltransferases from the same origin, but with different substrate specificities, and of both 4CL enzyme candidates as well.

In comparison to a substrate feed of 1 mM, the coniferyl alcohol concentration dropped three-fold with 2 mM caffeic acid for *At4CL* + *ZmCOMT* ( $84.4 \pm 34.2$  µM) and 2.6-fold using *Pc4CL* + *ZmCCoAOMT* ( $49.8 \pm 14.8$  µM), respectively (Figure 2.23 A). With 5 mM caffeic acid, both coniferyl alcohol titers decreased even more to values of 36-37 µM, which were slightly below the standard curve. As before, the downstream pathway containing *At4CL* in combination with *ZmCOMT* consumed the substrate completely for a feed of 2 mM and to a large extent of 5 mM caffeic acid (Figure 2.23 C). Whereas with *Pc4CL* and *ZmCCoAOMT*, caffeic acid was left unconverted in concentrations of around  $0.74 \pm 0.57$  mM with a 2 mM and of  $3.02 \pm 0.65$  mM with a 5 mM substrate feed.

The enzyme combination with *At4CL* and *ZmCOMT* produced more caffeoyl alcohol compared to the usage of *Pc4CL* and *ZmCCoAOMT* (Figure 2.23 B). The concentration of caffeoyl alcohol decreased by 43.9 % for *Pc4CL* + *ZmCCoAOMT* with an increase in substrate feed from 2 mM to 5 mM, while the production of caffeoyl alcohol improved by 131.8 % for *At4CL*

+ *ZmCOMT*. A single performed empty vector control with 2 mM and 5 mM caffeic acid did not lead to any production of pathway products and intermediates (data not shown).



**Figure 2.23: Increase of substrate feed into the downstream synthetic pathway using plasmid-based enzyme expression in *E. coli* BL21(DE3).** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 2 mM and 5 mM caffeic acid as substrate added at time  $t_0$ . The concentration of coniferyl alcohol (A), caffeoyl alcohol (B) and caffeic acid (C) were measured after 48 h. Expression was performed at 26 °C prior to application in whole-cell biotransformation. The tested enzyme combinations were *ZmCCR* and *ZmCAD* with either *At4CL* and *ZmCOMT* (black) or *Pc4CL* and *ZmCCoAOMT* (grey). The shown data represent mean values with standard deviation from biological triplicates.

In the end, the consumption of caffeic acid as well as the production of high caffeoyl alcohol concentrations in the enzyme combination with *At4CL* indicates that this 4CL enzyme candidate is more stable concerning substrate and product inhibition *in vivo* compared to *Pc4CL*. Combinations with *Pc4CL* were found to be unable to convert substrate concentrations above 2 mM caffeic acid. Contrary to expectations, the amount of coniferyl alcohol decreased upon addition of higher substrate concentrations. With a parallel accumulation of sufficient amounts of caffeoyl alcohol, the *in vivo* activity of both methyltransferases seemed to be low in whole-cell biotransformation when applying higher substrate feeds. Furthermore, the possibility that *ZmCOMT* and *ZmCCoAOMT* were prone to substrate inhibition by caffeoyl alcohol cannot be excluded.

## II.5 The synthetic pathway towards monolignols

This chapter deals with the production of monolignols performed in resting cell assays. First the combined functionality of the four pathway enzymes (TAL, 4CL, CCR, CAD) towards the synthesis of *p*-coumaryl alcohol as the simplest monolignol is investigated (chapters II.5.1). Thereby, the establishment of a single vector containing the whole basic pathway is intended, since the subsequent complementation by other enzymes (C3H, OMT) produced from plasmids is considered to be beneficial for sufficient co-expression levels (chapter II.5.1, cf. chapter II.6.2). At last, upstream and downstream pathways are combined for the production of coniferyl alcohol in the strains *E. coli* BL21(DE3).G213 and *E. coli* BL21(DE3) (chapter II.5.2).

### II.5.1 From L-tyrosine to *p*-coumaryl alcohol

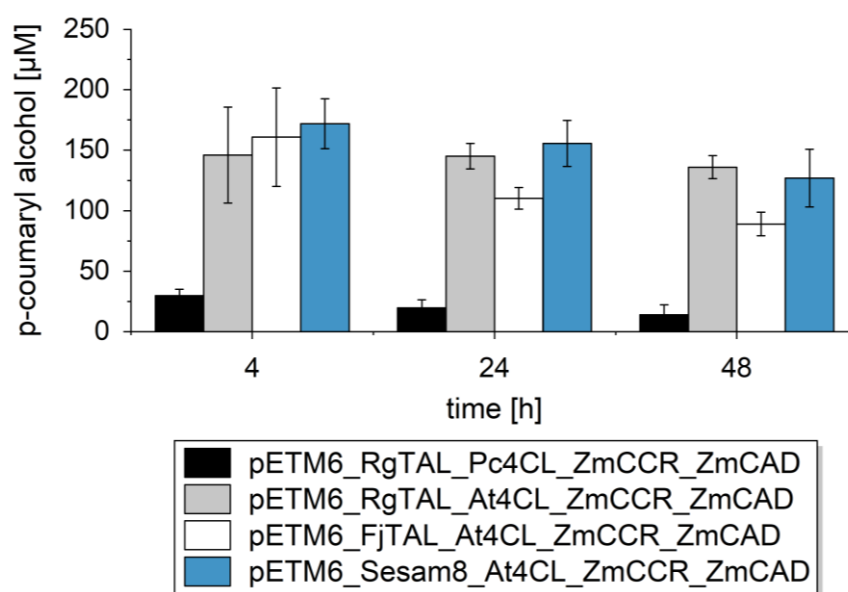
The chosen two 4CL enzyme candidates were shown to have varying co-expression levels depending on the choice of methyltransferase enzyme and are known to exhibit different substrate specificities (cf. Figure 2.22). The enzyme 4CL1 from *A. thaliana* demonstrated the highest catalytic efficiency towards *p*-coumaric and caffeic acid compared to its other isoforms and also converted ferulic and 5-hydroxyferulic acid as substrates [181]. On the other hand, the 4CL1 from *P. crispum* showed stronger activity towards *p*-coumaric and ferulic acid but low activity towards caffeic acid [182].

Aiming at an efficient *in vivo* production of *p*-coumaryl alcohol, the most appropriate 4CL candidate as well as TAL and 4CL combination were investigated. For initial characterization, RgTAL, which had exhibited the highest individual *in vivo* activity before, was combined with either the Pc4CL or At4CL enzyme (cf. chapter II.1.1). The pathway towards *p*-coumaryl alcohol was completed by the co-expression of ZmCCR and ZmCAD. Via expression from two separate plasmids in *E. coli* BL21(DE3), the *in vivo* production of *p*-coumaryl alcohol was 24-fold higher with At4CL compared to Pc4CL (data not shown). The superior activity of At4CL was mostly attributed to the more pronounced co-expression bands of RgTAL in comparison to Pc4CL (data not shown).

In general, the Duet vectors from Novagen® (Merck KGaA, Darmstadt, Germany) enable facile expression of up to eight genes using four vectors with different antibiotic selection markers and compatible replication origins [152,169]. An unspecific side-peak formation was observed in initial tests, when pACYCDuet was used as an empty vector control in resting cells (data not shown). In personal communication with Stefan Wohlgemuth (Institute of Biochemistry II,

HHU Düsseldorf) similar side-products were identified as acetylated products from the phenylpropanoid pathway by LC-MS (data not shown). Consequently, it was suspected that the chloramphenicol acetyltransferase from the pACYCDuet backbone might exhibit acetylation activity on the pathway intermediates. The application of the pACYCDuet was thus omitted due to this unfavorable side-activity as well as due to the lowest copy number of the pACYCDuet plasmid in comparison to the other three vectors.

Therefore, an ePathBrick vector pETM6 with TAL, 4CL, *ZmCCR* and *ZmCAD* was constructed using a monocistronic gene configuration. In the past, a few studies concerning multi-enzyme pathways had applied the monocistronic configuration based on the model from Xu *et al.* (2012) [152,183,184]. Accordingly, each gene on the pETM6 vector was constructed with its own T7 promoter and T7 terminator. Utilizing less plasmids simultaneously gives the advantage of a lower metabolic burden for the *E. coli* host that results from plasmid maintenance and antibiotic stress [131,185,186].



**Figure 2.24: Synthetic pathway towards *p*-coumaryl alcohol using a single ePathBrick vector in *E. coli* BL21(DE3).** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM L-tyrosine as substrate added at time  $t_0$ . The product concentration of *p*-coumaryl alcohol was measured over time (4 h, 24 h, 48 h). All enzymes were expressed from pETM6 vector in a monocistronic gene configuration at 26 °C prior to application in whole-cell biotransformation. Three different TAL enzymes being *RgTAL* (black, grey), *FjTAL* (white) or *Sesam8* (blue) were combined with the 4CL enzyme *Pc4CL* (black) or *At4CL* (grey, white, blue). The shown data represent mean values with standard deviation from biological triplicates.

In a whole-cell biotransformation with 1 mM L-tyrosine, the application of the ePathBrick vector with all four enzymes led to the production of  $29.8 \pm 5.3$  μM *p*-coumaryl alcohol after 4 h, when using *RgTAL* and *Pc4CL*, and of  $146.0 \pm 39.7$  μM with *RgTAL* and *At4CL* (Figure 2.24). In this way, the enzyme combination with *At4CL* exhibited a near five-fold

higher *in vivo* activity compared to using the downstream pathway enzyme *Pc4CL*. The relatively low titers could be associated with thin expression bands for TAL and 4CL and thus attributed to poor expression levels of all pathway enzymes, when the proteins were co-expressed from a single pETM6 plasmid in the monocistronic gene configuration (Figure A.7, Appendix).

Moreover, only small amounts for *p*-coumaric acid were measured as intermediate indicating an efficient and consecutive conversion into *p*-coumaryl alcohol by the 4CL, *ZmCCR* and *ZmCAD* enzymes (data not shown).

In addition to *RgTAL*, the other two TALs, *FjTAL* and *Sesam8*, which showed high activities when co-expressed with CYP199A2 as C3H, were applied in the pETM6 vector with *At4CL* as well (cf. chapter II.3.1). In the pathway towards *p*-coumaryl alcohol, similar titers compared to *RgTAL* were yielded using both alternative TALs, in particular  $160.8 \pm 40.6 \mu\text{M}$  with *FjTAL* and  $171.8 \pm 20.6 \mu\text{M}$  with *Sesam8* after 4 h (Figure 2.24).

No further increases in *p*-coumaryl alcohol concentration were observed after 4 h for any enzyme combination. The protein bands for *FjTAL* and *Sesam8* were also found to be less pronounced compared to the previous single enzyme expressions and co-expressions with CYP199A2 (Figure A.7, Appendix). Xu *et al.* (2012) suggested for a monocistronic gene configuration that terminator and subsequent promoter regions might interact with each other so that transcription could be hindered [152]. This may be considered a plausible explanation for the constructed pETM6 vector in this work as well.

In a growing cell assay, the pathway enzymes *RgTAL*, *At4CL*, *ZmCCR* and *ZmCAD* expressed from the single pETM6 plasmid in *E. coli* BL21(DE3) led to a *p*-coumaryl alcohol concentration of  $260.7 \mu\text{M}$  after 6 h without any precursor supplementation (Figure A.8, Appendix). This titer of *p*-coumaryl alcohol was 79 % higher compared to the value in the corresponding resting cell assay. The higher titer not only indicates that the growing cell assay provides an environment for improved pathway performance, but also that the L-tyrosine present in TB medium is sufficiently consumed as substrate.

The findings in this chapter generally support that production of *p*-coumaryl alcohol can be improved by applying alternative enzyme candidates for TAL and 4CL reaction steps. Thereby, the combinations of all three TAL enzymes with *At4CL* exhibited similar *in vivo* activities. In the end, the construct containing *FjTAL* was chosen for the later growing cell assay due to the previous high activity in co-expression with CYP199A2 as well as due to its reported superior

*in vivo* activity over *RgTAL* and *Sesam8* without substrate supplementation in the past (chapter II.3.1, cf. chapter II.6.2) [157].

## II.5.2 From L-tyrosine to coniferyl alcohol

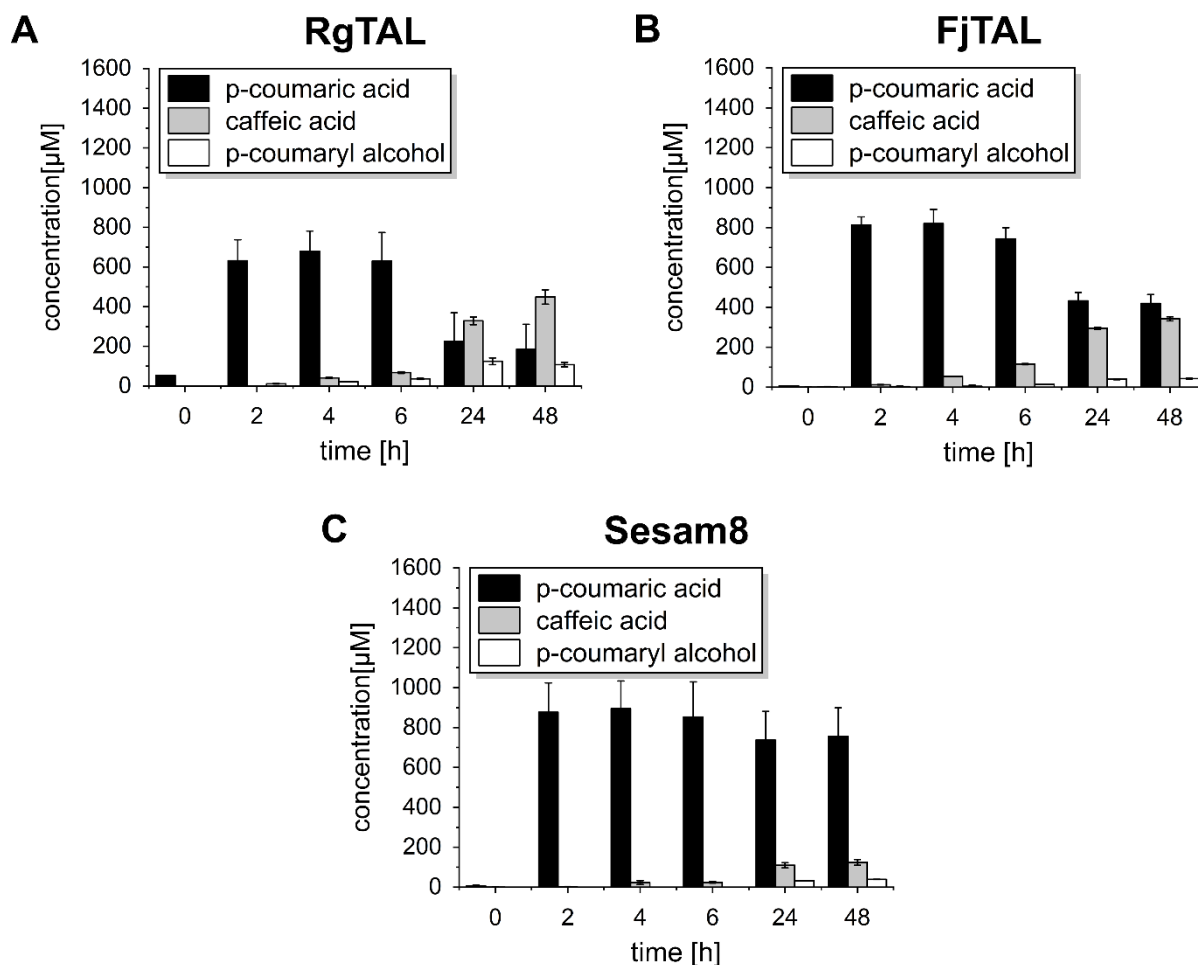
### In a genome-integrated *E. coli* strain

The G213 strain of *E. coli* BL21(DE3) showed no conversion of L-tyrosine into *p*-coumaric acid in a resting cell assay, although this strain contains the genome-integrated *RsTAL* (data not shown). Thus, a plasmid-based gene copy of a TAL enzyme on the pETM6 vector was co-transformed into the G213 strain in addition to the plasmids pCDFDuet\_*AtCCoAOMT\_CYP199A2* and pRSFDuet\_Pux\_PuR. The caffeoyl-CoA methyltransferase from *A. thaliana* was chosen, because this enzyme in combination with the G213 strain had led to the highest production of coniferyl alcohol in the evaluation of the downstream pathway before (chapter II.4.1). The construction of C3H and OMT on the same vector was made in accordance with a previous study, in which co-expression of these enzymes from a single plasmid led to an improved ferulic acid titer compared to the expression from separate plasmids [131].

In whole-cell biotransformations of the G213 strain with 1 mM L-tyrosine, *p*-coumaric acid was rapidly produced after 2 h and the concentration stayed at the same level during the first time (Figure 2.25). This indicated L-tyrosine consumption by TAL enzymes. Besides, small amounts of caffeic acid measured after 4 h, the titer for caffeic acid was higher after 24 h and 48 h. At the same time, minor amounts of *p*-coumaryl alcohol were detected and the concentration of *p*-coumaric acid decreased, which suggested that *p*-coumaric acid as intermediate had been converted further either by the 4CL or CYP199A2 enzyme.

Interestingly, neither coniferyl alcohol nor caffeoyl alcohol were produced in the resting cell assays. The late accumulation of caffeic acid and no further conversion of *p*-coumaric acid indicates that the downstream pathway enzymes in particular the 4CL enzyme became mostly inactive after 24 h. This can also be supported by the constant concentrations of *p*-coumaryl alcohol after 24 h. Furthermore, the only minor increase in caffeic acid titer points towards a low activity of the CYP199A2 as C3H after 24 h as well. The G213 strain with *RgTAL* exhibited the highest production of *p*-coumaryl alcohol with  $124.6 \pm 16.8 \mu\text{M}$  at 24 h. This strain also had the highest titer of caffeic acid with  $448.7 \pm 36.2 \mu\text{M}$  at 48 h, while caffeic acid was produced 24 % less with *FjTAL* ( $342.7 \pm 10.4 \mu\text{M}$ ) and 72 % less with *Sesam8* ( $124.2 \pm 14.4 \mu\text{M}$ ) after 48 h.

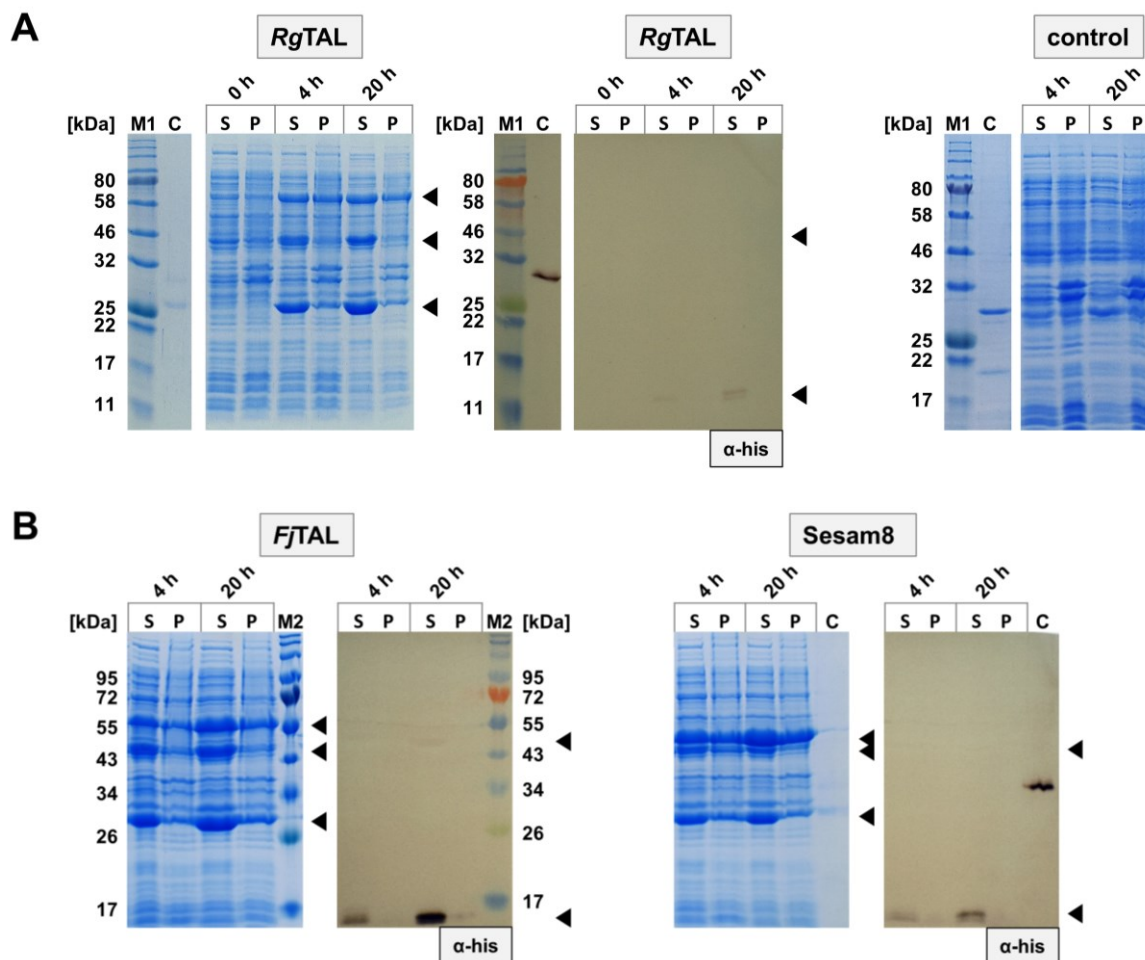
In general, all three TAL enzymes and the *AtCCoAOMT* showed pronounced bands in SDS-PAGE analysis, when expressed with CYP199A2 and redox partners in the G213 strain (Figure 2.26). The third distinct protein band around 46 kDa on the SDS gel corresponded to the expression of CYP199A2 and the reductase PuR. The latter protein PuR was detected with faint bands in western blot analysis, while the small ferredoxin Pux exhibited stronger stained bands around 12 kDa on the blot. Thus, all proteins produced from plasmids were expressed in the *E. coli* BL21(DE3).G213 strain.



**Figure 2.25: Synthetic pathway towards coniferyl alcohol from L-tyrosine in *E. coli*/BL21(DE3).G213.** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM L-tyrosine as substrate added at time  $t_0$  (0 h). The concentration of products and intermediates was measured over time (2 h, 4 h, 6 h, 24 h, 48 h). Expression was performed at 26 °C prior to application in whole-cell biotransformation. Thereby, the constructs pCDFDuet\_AtCCoAOMT\_CYP199A2 and pRSFDuet\_Pux\_PuR were combined with one TAL enzyme expressed from the pETM6 vector: *RgTAL* (A), *FjTAL* (B) or *Sesam8* (C). The shown data represent mean values with standard deviation from biological duplicates.

For the genome-integrated genes, no expression bands for the untagged *RsTAL*, *Pc4CL*, *ZmCCR* or *ZmCAD* could be observed in the SDS-PAGE analysis. Thus, total mRNA was isolated from the co-transformed G213 strains, subjected to reverse transcription and qualitatively assessed for the enzymes *RsTAL*, *Pc4CL*, *ZmCCR* and *ZmCAD* via PCR in order

to verify the expression of the genome-integrated genes. Specific bands visualized in agarose gel electrophoresis of the cDNA from the four proteins proved their transcription from constitutive promoters in the G213 strain and subsequently suggested their expression *in vivo* (data not shown). These results indicate that all nine pathway proteins were expressed and exhibited combined activity to produce caffeic acid and *p*-coumaryl alcohol.



**Figure 2.26: Co-expression analysis of the synthetic pathway towards coniferyl alcohol using *AtCCoAOMT* and *CYP199A2* with three different TAL enzymes in *E. coli* BL21(DE3).G213.** Co-expression was performed using pCDFDuet\_AtCCoAOMT\_CYP199A2, pRSFDuet\_his<sub>6</sub>-Pux\_his<sub>6</sub>-PuR and pETM6\_TAL with *RgTAL* (A), *FjTAL* or *Sesam8* (B) in *E. coli* BL21(DE3).G213 at 26 °C. 5/OD samples were taken at the point of induction with 1 mM IPTG (0 h) and during the expression (4 h and 20 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE and western blot analysis. For detection in western blot analysis the anti-His ( $\alpha$ -his) and GAM<sub>RC</sub><sup>AP</sup> antibodies were used. An expression of the empty vectors (pETM6, pCDFDuet, pRSFDuet) served as **control** (A). The black arrowheads indicate the relevant expression bands. M1 – P7712 protein ladder (NEB), M2 – P7719 protein ladder (NEB), C – mCherry-his<sub>6</sub> as His<sub>6</sub>-tag control (0.5  $\mu$ g, 27.7 kDa), S – supernatant, P – pellet fraction, *RgTAL* – 74.8 kDa, *FjTAL* – 56.7 kDa, *Sesam8* – 54.0 kDa, *CYP199A2* – 44.6 kDa, his<sub>6</sub>-Pux – 12.6 kDa, his<sub>6</sub>-PuR – 44.5 kDa, *AtCCoAOMT* – 25.9 kDa, *RsTAL* – 55.5 kDa, *Pc4CL* – 60.3 kDa, *ZmCCR* – 40.1 kDa, *ZmCAD* – 38.7 kDa.

As stated above, the downstream pathway enzymes as well as the C3H probably became inactive over time, which led to the accumulation of *p*-coumaric and caffeic acid. Regarding the differences in production of caffeic acid between the applied TAL enzymes, it might be that



the expression of Sesam8 was too excessive so that less resources were available for the production of the other enzymes such as the C3H. This would also explain the less distinct protein bands for CYP199A2 and PuR with Sesam8 compared to the co-expression with *FjTAL*, which were performed under the same conditions (Figure 2.26 B). Furthermore, both TAL proteins *RgTAL* and *HaTAL*, which showed weaker expression bands, were evaluated in the co-transformed G213 strain as well, but the results equaled the above mentioned observations and the measured *p*-coumaric acid concentrations were generally lower (Figure A.9, Appendix).

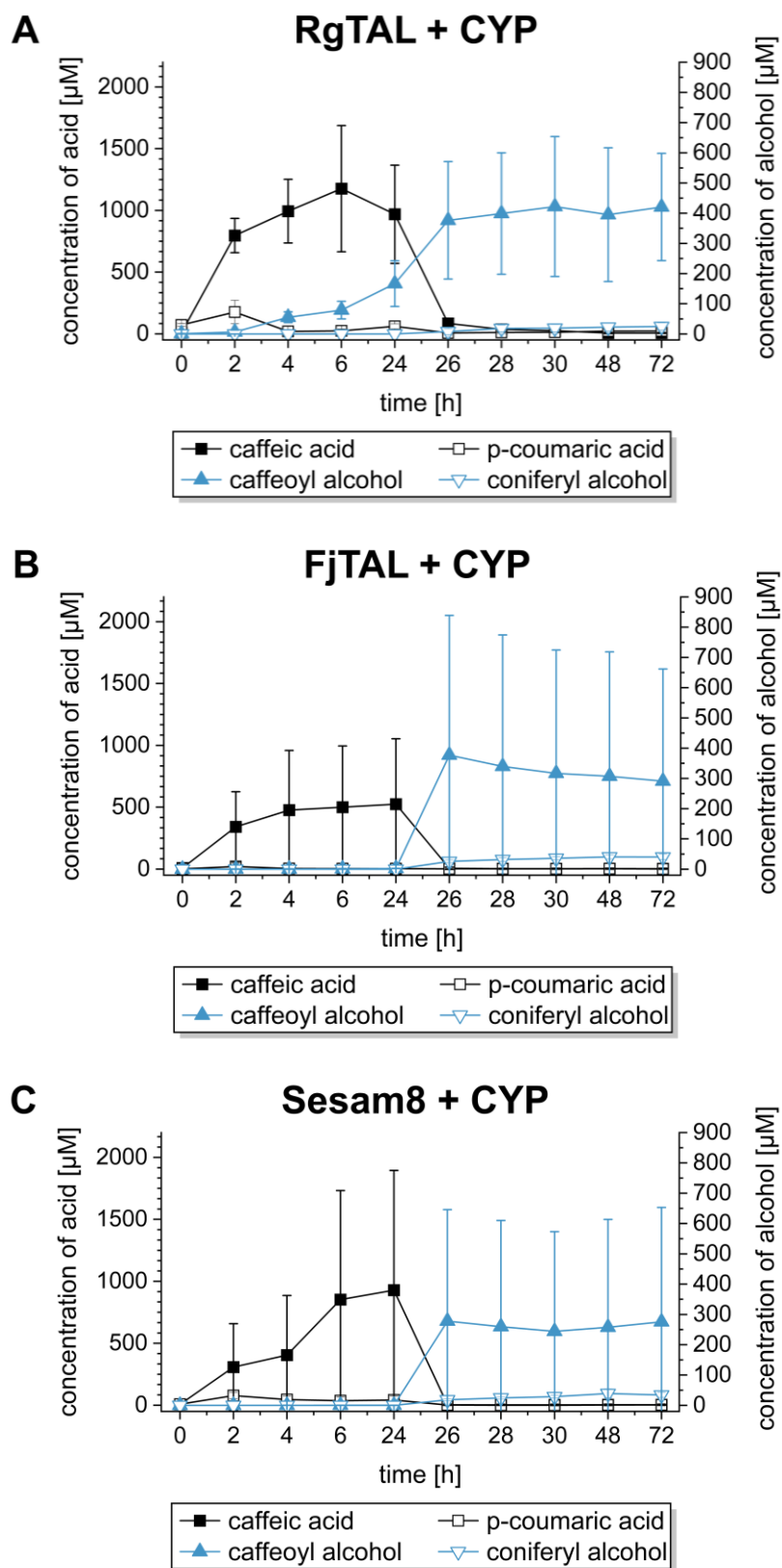
### Two-cells one-pot approach in *E. coli* BL21(DE3)

Both, the upstream and downstream pathway, were individually evaluated to be functional with respect to the formation of their end products. Consequently, a two-cells one-pot approach is evaluated next. This experimental structure offers a solely plasmid-based production of all pathway enzymes, which had been shown to produce proteins in higher quantity and to be more beneficial for the performance of the downstream pathway before (cf. chapter II.4.2). The two-cells one-pot experiment started with a first set of cells containing TAL + CYP199A2 combinations as 140 g<sub>wcw</sub>/L cell suspension and the substrate addition of 2 mM L-tyrosine. After an incubation for 24 h, the second set of cells with the equal cell concentration was mixed in a 1:1 ratio of volume with the first cell suspension yielding a concentration of 70 g<sub>wcw</sub>/L for each cell subset. Corresponding to the previous results, *At4CL* and *ZmCOMT* were used in combination with *ZmCCR* and *ZmCAD* as second set of cells, because these four enzymes had exhibited the highest production of coniferyl alcohol from caffeic acid (chapter II.4.2).

In the first time period, caffeic acid was produced at high levels in the resting cell assay and with low concentrations of *p*-coumaric acid as intermediate, which is consistent with the expectations based on TAL + C3H combined activity (Figure 2.27). The measured concentrations of caffeic acid after 24 h were  $968.5 \pm 397.9 \mu\text{M}$  with *RgTAL*,  $524.5 \pm 528.7 \mu\text{M}$  with *FjTAL* and  $928.8 \pm 966.2 \mu\text{M}$  with Sesam8. Like before, when *RgTAL* was used with CYP199A2, the produced caffeic acid decreased slightly from 6 h to 24 h (cf. chapter II.3.1).

After the addition of the second set of cells, caffeic acid was rapidly consumed within 2 h and caffeoyl and coniferyl alcohol were produced (Figure 2.27). The coniferyl alcohol titers between all three experiments were similarly low with  $24.9 \pm 12.1 \mu\text{M}$  (72 h) using *RgTAL*,  $36.3 \pm 4.8 \mu\text{M}$  (30 h) with *FjTAL* and  $40.7 \pm 7.3 \mu\text{M}$  (48 h) with Sesam8. These values were 6.6- to 10.4-fold lower compared to the previous production of coniferyl alcohol from 1 mM

caffeic acid by applying the downstream pathway alone (chapter II.4.2). A prolonged incubation time beyond 72 h did not lead to improved titers (data not shown).



**Figure 2.27: Two-cells one-pot experiment.** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 2 mM L-tyrosine as substrate added at time  $t_0$  (0 h). The concentration of products

and intermediates was measured over time (from 2 h to 72 h). Expression was performed at 26 °C prior to application in whole-cell biotransformation. Thereby, the first set of cells expressed proteins from pCDFDuet\_CYP199A2 (CYP), pRSFDuet\_Pux\_PuR and pETM6\_TAL with *Rg*TAL (A), *Fj*TAL (B) or Sesam8 (C), while the second set of cells produced enzymes from pETM6\_*At*4CL, pCDFDuet\_*Zm*COMT and pRSFDuet\_*Zm*CCR\_*Zm*CAD and were added at 24 h. The shown data represent mean values with standard deviation from biological triplicates.

Interestingly, a control experiment, in which the first subset of cells containing the empty vectors was mixed with 2 mM caffeic acid after 24 h and prior to addition of the downstream pathway cells, produced an equally low concentration of coniferyl alcohol (Figure A.10, Appendix). With these results being also in contrast to the previous findings, the differences may be attributed to the experimental setup of the two-cells one-pot approach. The higher cell concentration of 140 g<sub>wc</sub>/L might have favored the dilution of the substrate, as it could distribute over the doubled amount of cells compared to the original resting cell assay.

Nonetheless, the empty vector control showed that caffeic acid was completely converted into ~ 1.1 mM caffeoyl alcohol (Figure A.10, Appendix). Thus, the downstream pathway enzymes *At*4CL, *Zm*CCR and *Zm*CAD were active, but the activity of the methyltransferase *Zm*COMT was likely impaired. A possible inhibition of COMT by caffealdehyde or coniferaldehyde can be ruled out, because no accumulating aldehyde compounds were detected as intermediates during the two-cells one-pot experiment. Thus, the activity of *Zm*COMT was probably affected by other factors than substrate or product inhibition, because comparable high caffeoyl alcohol titers had accumulated in the resting cell assay of the downstream pathway before without disturbing the OMTs' methylation activity (chapter II.4.2).

The expected value for caffeic acid in the samples at 24 h, collected prior to addition of the second set of cells, should be around 2 mM. Instead, mean values around 1 mM caffeic acid with very high standard deviations were measured for the two-cells one-pot experiments. Likewise, the caffeoyl alcohol concentrations fluctuated enormously, as it had been observed in the previous evaluation of the downstream pathway (cf. chapter II.4.2). However, the high standard deviations of caffeic acid occurred prior to the addition of the second set of cells. This observation can point towards consumption of caffeic acid by the *E. coli* cells, which was indicated in the toxicity test and discussed before (cf. chapter II.3.1, cf. Figure A.19 A and F, Appendix). Besides, the aforementioned instable and thus variable activity of the 4CL enzyme could have influenced the strong deviation of caffeoyl alcohol titers (chapters II.4.2 and II.5.2). It is further noteworthy that the standard solution of caffeoyl alcohol in DMSO was observed to be susceptible to decomposition due to decreased measured area values in HPLC

chromatograms with respect to multiple freeze-thaw cycles (data not shown). This observation implied that caffeoyl alcohol may be generally instable in solution.

By taking these results together, it was possible to produce coniferyl alcohol from L-tyrosine using the two-cells one-pot approach though only in marginal concentrations. It is suspected that the experimental setup of two-cells one-pot was responsible for the insufficient methyltransferase activity of *ZmCOMT*. As a consequence, a different whole-cell assay is investigated for the complete synthetic pathway in the final chapter.

## II.6 Growing cell assay

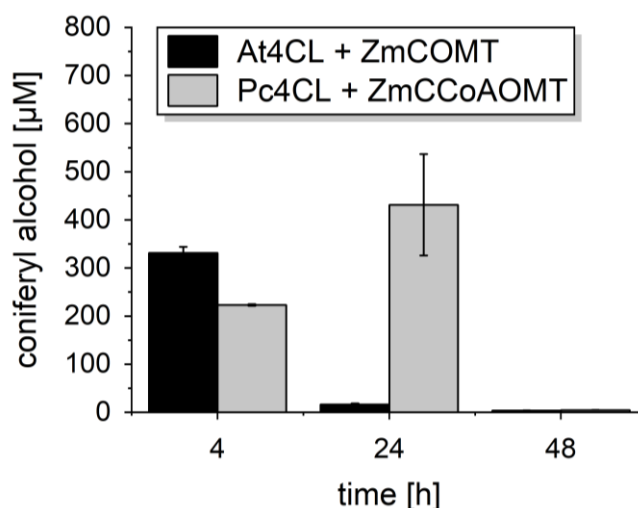
In general, it was noticed that the pH became more acidic with the progress of the whole-cell biotransformations (data not shown). Moreover, Kunjapur and co-workers (2014) supposed that synthetic pathway enzymes in resting cells lose their activity more rapidly and exhibit a lower overall productivity compared to growing cells [187]. Thus, a growing cell assay is performed for evaluation of the downstream pathway (chapter II.6.1) and different combined pathway activities for the production of coniferyl alcohol (chapters II.6.2-6.4).

In this experiment, a culture of *E. coli* cells was grown in TB medium and induced at an OD<sub>600</sub> of around 2 with 0.5 mM IPTG. At the same time, the substrate was added, so that the production of enzymes and the pathway activity occurred simultaneously. The lowered concentration of IPTG was chosen, as higher titers had been achieved from pathway enzymes in a previous study, when the protein expression was induced with amounts below 1 mM IPTG [130]. The higher induction OD<sub>600</sub> was adopted in order to avoid a possible reduction of cell viability at the induction time point and to yield substantial concentrations of all pathway enzymes concurrently with the addition of the substrate. Furthermore, complex TB medium was applied because of the higher cell growth and reported superior pathway product titers compared to the usage of synthetic media [130].

### II.6.1 Downstream pathway

In the resting cell assay, the plasmid-based production of downstream pathway enzymes yielded higher coniferyl alcohol titers compared to a plasmid-based expression of OMT in the *E. coli* BL21(DE3).G213 strain. Thus, the two best performing 4CL and OMT combinations from chapter II.4.2 were evaluated in the growing cell assay with 1 mM caffeic acid. The enzyme combination of *At4CL* + *ZmCOMT* produced  $331.0 \pm 12.7 \mu\text{M}$  coniferyl alcohol after 4 h, which corresponded to a conversion yield of 50.4 % (Figure 2.28). The second combination of *Pc4CL* + *ZmCCoAOMT* led to a lower production of  $223.2 \pm 1.7 \mu\text{M}$  coniferyl alcohol at 4 h

(conversion yield 34.0 %). Interestingly for this latter enzyme combination, the titer of coniferyl alcohol increased further to  $431.2 \pm 105.4 \mu\text{M}$  after 24 h corresponding to a conversion yield of 65.7 %, while coniferyl alcohol decreased enormously for the other combination *At4CL* + *ZmCOMT* at 24 h. After 48 h in the growing cell assay, almost no coniferyl alcohol was detected in both strains containing the downstream pathway.



**Figure 2.28: Growing cells of *E. coli* BL21(DE3) with the downstream synthetic pathway using plasmid-based protein expression.** The investigated enzyme combinations were *At4CL* and *ZmCOMT* (black) as well as *Pc4CL* and *ZmCCoAOMT* (grey) together with *ZmCCR* and *ZmCAD*. The growing cell assay was performed in 10 mL TB medium at 26 °C with supplementation of 1 mM caffeic acid at time  $t_0$ . The concentration of coniferyl alcohol was measured over time (4 h, 24 h, 48 h). The shown data represent mean values with standard deviation from biological duplicates.

Over the time course of the growing cell assay, the substrate caffeic acid was consumed completely and no caffeoyl alcohol accumulated as intermediate (data not shown), which contradicted the previous observations in resting cells (cf. chapter II.4.2). In addition to that, growing cells without supplemented substrate did not lead to the formation of any intermediates and confirmed that TB medium does not contain caffeic acid (data not shown). Furthermore, caffeic acid was not converted into any pathway product in the empty vector control as expected (data not shown).

Altogether, the production of coniferyl alcohol from caffeic acid by the downstream pathway was more efficient and led to higher concentrations in growing cells compared to the whole-cell biotransformation. By comparing the highest individual values of coniferyl alcohol, the enzyme combination *At4CL* + *ZmCOMT* produced 28 % more and *Pc4CL* + *ZmCCoAOMT* generated a 2.4-fold higher titer in growing than in resting cells. With respect to the standard deviations, both 4CL + OMT combinations exhibited around the same activity in growing cells in the end. The only exception was the faster conversion of caffeic acid by *At4CL* + *ZmCOMT* in comparison to *Pc4CL* + *ZmCCoAOMT*.

The observed significant decrease in coniferyl alcohol titers over time indicates that coniferyl alcohol could have been degraded or consumed by the metabolically active cells. It is further possible that coniferyl alcohol dimerized spontaneously and escaped detection with the applied HPLC method [179]. As an additional minor experiment, M9 minimal medium was applied for these growing cells as well, but resulted in no complete substrate consumption, in the accumulation of caffeoyl alcohol and in two- to six-fold lower concentrations of coniferyl alcohol (data not shown). Thus, these results confirm that the usage of complex medium leads to an improved overall performance of the pathway enzymes.

In conclusion, the growing cell assay with the downstream pathway seems to be beneficial for OMT activity, as no caffeoyl alcohol and instead higher coniferyl alcohol titers were detected with both classes of methyltransferases from *Z. mays*.

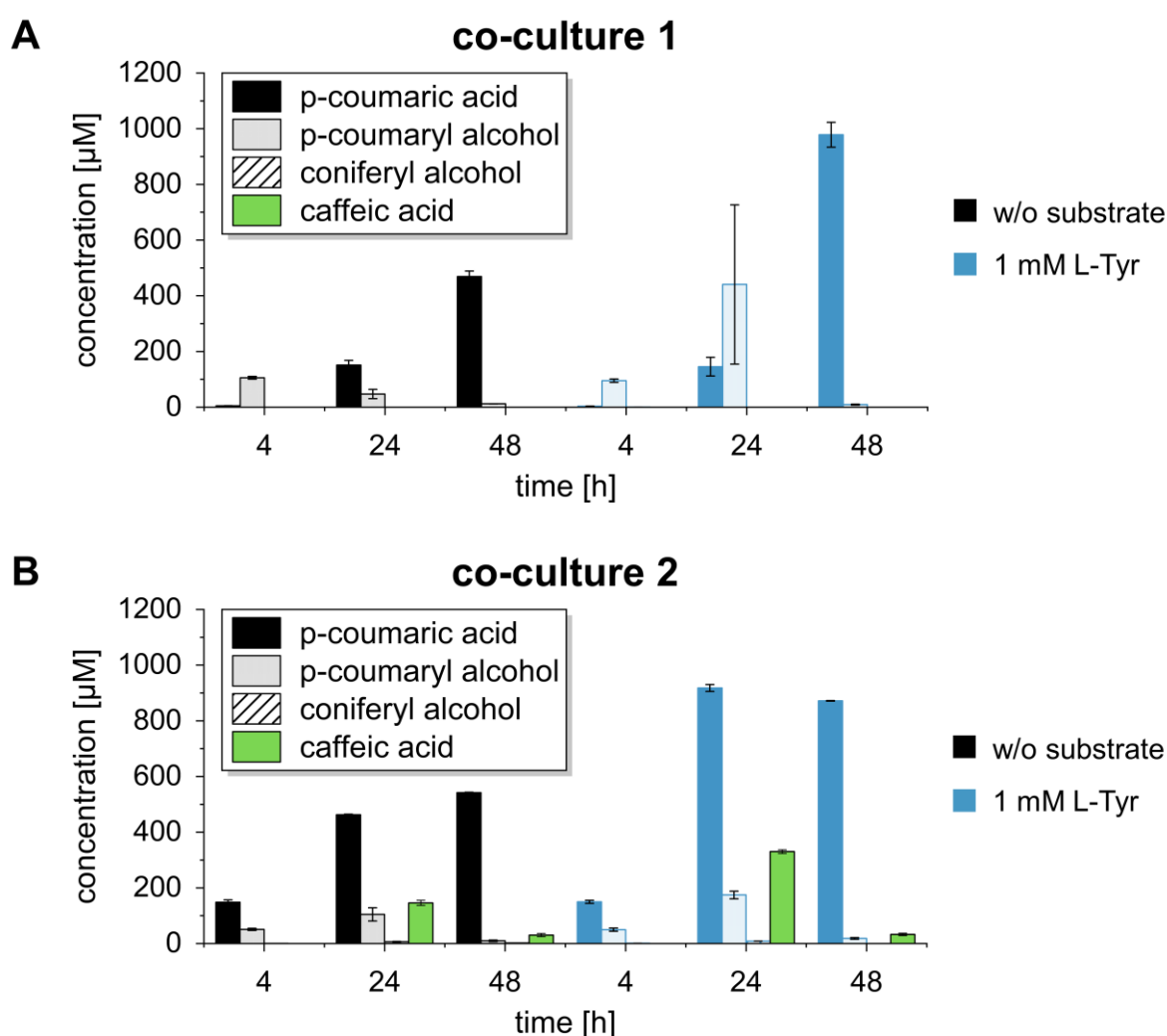
## **II.6.2 Production of coniferyl alcohol in *E. coli* BL21(DE3)**

### **Growing co-culture**

As a growing cell assay equivalent to the two-cells one-pot approach, growing co-cultures using *FjTAL* and the NΔ7-CYP199A2 variant with Pux/PuR in one *E. coli* BL21(DE3) strain were performed (cf. chapter II.5.2). *FjTAL* was chosen due to its sufficient activity in co-expression with other pathway enzymes (chapters II.3.1, II.5.1 and II.5.2). Besides, the co-activity of the NΔ7 variant of the P450 enzyme was investigated for comparison and because of its equal activities with regard to CYP199A2 as C3H (chapters II.2.2). The second *E. coli* BL21(DE3) strain consisted of the downstream pathway with either *At4CL* and *ZmCOMT* or *Pc4CL* and *ZmCCoAOMT* in combination with *ZmCCR* and *ZmCAD*.

In contrast to the two-cells one-pot approach, both strains were inoculated together and at the same time into 10 mL TB medium for the growing co-culture. Thereby, overnight cultures of both strains were mixed in a 1:1 ratio with respect to OD prior to inoculation. The induction with 0.5 mM IPTG was also applied at an OD<sub>600</sub> around 2 and without or with the addition of 1 mM L-tyrosine as substrate. Here, the collected samples over time consisted of the whole cell broth. Such extracted samples containing supernatant and cells had been compared to samples with only supernatant in the past and no differences in concentrations of substances could be determined (data not shown). Thus, it is assumed that all measureable intermediates of the synthetic pathway are well membrane-permeable and that the cells have no influence on extraction.

After 24 h and 48 h of the growing co-culture using *At4CL* + *ZmCOMT*, the accumulation of *p*-coumaric acid was observed (co-culture 1, Figure 2.29 A). Without substrate, the concentration of *p*-coumaric acid was  $469.1 \pm 19.8 \mu\text{M}$  after 48 h, while the titer was twice as high with 1 mM L-tyrosine ( $978.5 \pm 44.7 \mu\text{M}$ ). Apart from *p*-coumaric acid, only *p*-coumaryl alcohol was measured as another product in small amounts without substrate ( $105.7 \pm 5.0 \mu\text{M}$  at 4 h) and in concentrations of  $440.7 \pm 285.8 \mu\text{M}$  with substrate at 24 h. Interestingly, the titer for *p*-coumaryl alcohol decreased greatly to values below the standard curve after 48 h. Neither coniferyl alcohol nor any C3-hydroxylated phenylpropanoids were detectable in this growing co-culture experiment (co-culture 1, Figure 2.29 A).



**Figure 2.29: Growing co-culture of *E. coli* BL21(DE3) using *FjTAL* and NΔ7-CYP199A2 in the upstream and two different 4CL and OMT combinations in the downstream pathway strain.** The investigated enzyme combinations of the downstream pathway were *At4CL* and *ZmCOMT* (A, co-culture 1) as well as *Pc4CL* and *ZmCCoAOMT* (B, co-culture 2) together with *ZmCCR* and *ZmCAD*. The growing cell assay was performed in 10 mL TB medium at 26 °C without substrate addition (w/o, black) or with supplementation of 1 mM L-tyrosine (L-Tyr, blue) at time  $t_0$ . The concentration of products and intermediates was measured over time (4 h, 24 h, 48 h). Caffeic acid (green) was only produced in co-culture 2 (B). The shown data represent mean values with standard deviation from biological duplicates.

The formation of *p*-coumaric acid clearly demonstrates activity of the *Fj*TAL, which was able to efficiently utilize L-tyrosine from the TB medium. Also, the concentrations of *p*-coumaric acid and *p*-coumaryl alcohol increased further upon an additional feed of 1 mM L-tyrosine, which can be attributed to the sufficient *Fj*TAL activity as first enzyme in the synthetic pathway. The observation that the *p*-coumaryl alcohol titer did not increase after 24 h point towards inactivity of the downstream pathway enzymes after that point in time. Due to the further accumulation of *p*-coumaric acid, the 4CL enzyme was presumably inactive, as it had been observed in a resting cell assay before (chapter II.5.2).

In an alternative second strain, the enzyme combination of *Pc*4CL + *Zm*CCoAOMT was also applied in the growing co-culture assay and showed similar results (co-culture 2, Figure 2.29 B). Two exceptions with this strain were that, firstly, *p*-coumaric acid accumulated in higher amounts at 4 h and 24 h. This can be explained by the previously mentioned lower performance and apparent more sensitivity to substrate inhibition of *Pc*4CL in contrast to *At*4CL in whole-cell biotransformations (cf. chapters II.5.1 and II.4.3). Secondly, small concentrations of caffeic acid were detected after 24 h being  $146 \pm 9.4 \mu\text{M}$  without added substrate and  $330 \pm 6.7 \mu\text{M}$  with 1 mM L-tyrosine. The production of caffeic acid in the alternative second strain indicates that NΔ7-CYP199A2 as C3H was active in the growing co-culture, but the overall *in vivo* activity was low and apparently insufficient.

A possible explanation could be the different growth behavior between upstream and downstream pathway strain. The upstream pathway strain might have exhibited growth up to greater cell densities in the combination with *Pc*4CL + *Zm*CCoAOMT as the second strain. This could have resulted in improved and balanced expression levels of TAL and C3H in co-culture 2, as more *p*-coumaric acid and the intermediate caffeic acid were produced compared to the growing co-culture with the other downstream pathway enzymes (co-culture 1). Thus, it can be suspected that the upstream pathway strain exhibited growth in a possibly lower cell density with the co-cultured second strain containing *At*4CL + *Zm*COMT. An experimental evidence of the suggested growth behavior could not be confirmed, as both strains are indistinguishable from each other with regard to OD<sub>600</sub> measurement and phenotypic characteristics.

It is noteworthy that the OD<sub>600</sub> of the inoculation culture was slightly larger for the growing co-culture 2 (*Pc*4CL + *Zm*CCoAOMT) compared to co-culture 1 (*At*4CL + *Zm*COMT). This may have influenced the growth behavior of both co-cultured strains, too. Additionally, *p*-coumaryl

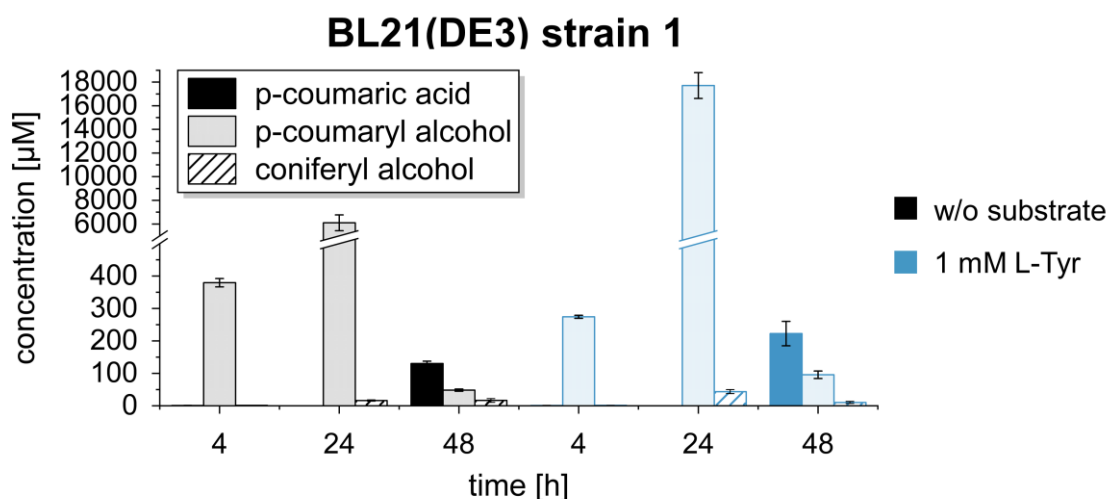


alcohol and caffeic acid were only marginally present in the samples at 48 h, which suspects their degradation or consumption in growing *E. coli* cells.

The previously performed two-cells one-pot approach led to the production of caffeoyl and coniferyl alcohol, but the tested growing co-culture predominantly produced *p*-coumaric acid and *p*-coumaryl alcohol (cf. chapter II.5.2). For a possible optimization with respect to efficient substrate channeling from upstream into the downstream pathway, the whole synthetic pathway is evaluated in a single strain next.

### Production of coniferyl alcohol in a single strain

For the co-expression of the complete synthetic pathway starting from L-tyrosine, the characterized ePathBrick vector pETM6 containing *FjTAL*, *At4CL*, *ZmCCR* and *ZmCAD* was applied in *E. coli* BL21(DE3) (chapter II.5.1). The strain was co-transformed further with pRSFDuet\_his<sub>6</sub>-Pux\_his<sub>6</sub>-PuR and pCDFDuet\_OMT\_NΔ7-CYP199A2 using *ZmCOMT* or *ZmCCoAOMT*, respectively.



**Figure 2.30: Growing cells of *E. coli* BL21(DE3) with the combined upstream and downstream synthetic pathway using plasmid-based enzyme expression.** The co-expression of enzymes was performed by combining the single ePathBrick vector pETM6 containing *FjTAL*, *At4CL*, *ZmCCR* and *ZmCAD* with pCDFDuet\_*ZmCOMT*\_NΔ7-CYP199A2 and pRSFDuet\_his<sub>6</sub>-Pux\_his<sub>6</sub>-PuR. The growing cell assay was performed in 10 mL TB medium at 26 °C without substrate addition (w/o, black) or with supplementation of 1 mM L-tyrosine (L-Tyr, blue) at time *t*<sub>0</sub>. The concentration of products and intermediates was measured over time (4 h, 24 h, 48 h). The shown data represent mean values with standard deviation from biological triplicates.

The growing cell assay of these *E. coli* BL21(DE3) strains resulted in the production of coniferyl alcohol albeit in very low titers. With *ZmCOMT* (BL21(DE3) strain 1), the maximal concentrations for coniferyl alcohol were  $16.5 \pm 5.6 \mu\text{M}$  (48 h) without added substrate and  $43.8 \pm 6.1 \mu\text{M}$  (24 h) with 1 mM L-tyrosine (Figure 2.30). Despite these values being below the standard curve, a minor 1.7-fold improvement in coniferyl alcohol titer could be observed

with an additional substrate feed. Apart from that, *p*-coumaric acid accumulated in low amounts after 48 h, while *p*-coumaryl alcohol started to be measurable earlier in the growing cell assay and only marginally at 48 h. The determined concentrations of *p*-coumaryl alcohol were  $379.5 \pm 12.8 \mu\text{M}$  (without substrate feed) and  $274.5 \pm 4.7 \mu\text{M}$  (with substrate feed) at 4 h. These results suggest that L-tyrosine from the TB medium was readily converted by TAL, 4CL, CCR and CAD into *p*-coumaryl alcohol without any other accumulating intermediates and that the activity of the 4CL enzyme was probably impaired after 48 h.

At 24 h, extraordinary high concentrations for *p*-coumaryl alcohol were measured, whose titers improved from 6.1 mM to 17.7 mM when L-tyrosine was fed to the medium. These calculated values are assumed to be incorrect, as the concentrations exceed the available amount of L-tyrosine in TB medium (around 2-3.9 mM)<sup>1</sup>. This significant increase cannot be explained by the application of an additional feed of L-tyrosine. As expected, growing cells of *E. coli* BL21(DE3), which were co-transformed with the three empty vectors as control, produced no products or intermediates (data not shown). Furthermore, no unusual peak with a retention time similar to *p*-coumaryl alcohol was observed in the HPLC chromatograms as well. Thus, the possible detection of unspecific product peaks from *E. coli* is excluded.

It is noteworthy that the HPLC chromatograms of samples at 48 h exhibited a generally low area value for the internal standard (IS), which led to a considerably high correction of all other area values. A decrease of the IS area value was also observed in the sample at 24 h though not as pronounced as at 48 h. However, it is generally assumed that the area correction with IS is probably not resulting in these pronounced differences observed for *p*-coumaryl alcohol. In the end, the measured values for *p*-coumaryl alcohol cannot be causally explained without any further investigations. It may be that the increase in area value occurred from a possible overlay of other unidentified product peaks. This hypothesis originates from the fact that some of the applied enzymes exhibit broad substrate specificities (4CL, CCR, CAD) and that the complex TB medium may contain many potential alternative substrates. Still, this hypothesis has not been experimentally confirmed.

The addition of 1 mM L-tyrosine mainly led to an increase of product titers and showed that higher amounts of coniferyl alcohol could be generated. In general, similar results were obtained when using *ZmCCoAOMT* in the *E. coli* BL21(DE3) strain (Figure A.11, Appendix). Nevertheless, the low titers of coniferyl alcohol and the high accumulation of *p*-coumaryl

---

<sup>1</sup> Calculated from specifications of three individual tryptones and two yeast extracts [162–166].

alcohol suggest that the C3H was insufficiently active. The expression of all proteins was thereby assumed to be low, since none or only thin protein bands had been observed in SDS-PAGE and western blot analysis (Figure A.12 and A.13 A, Appendix).

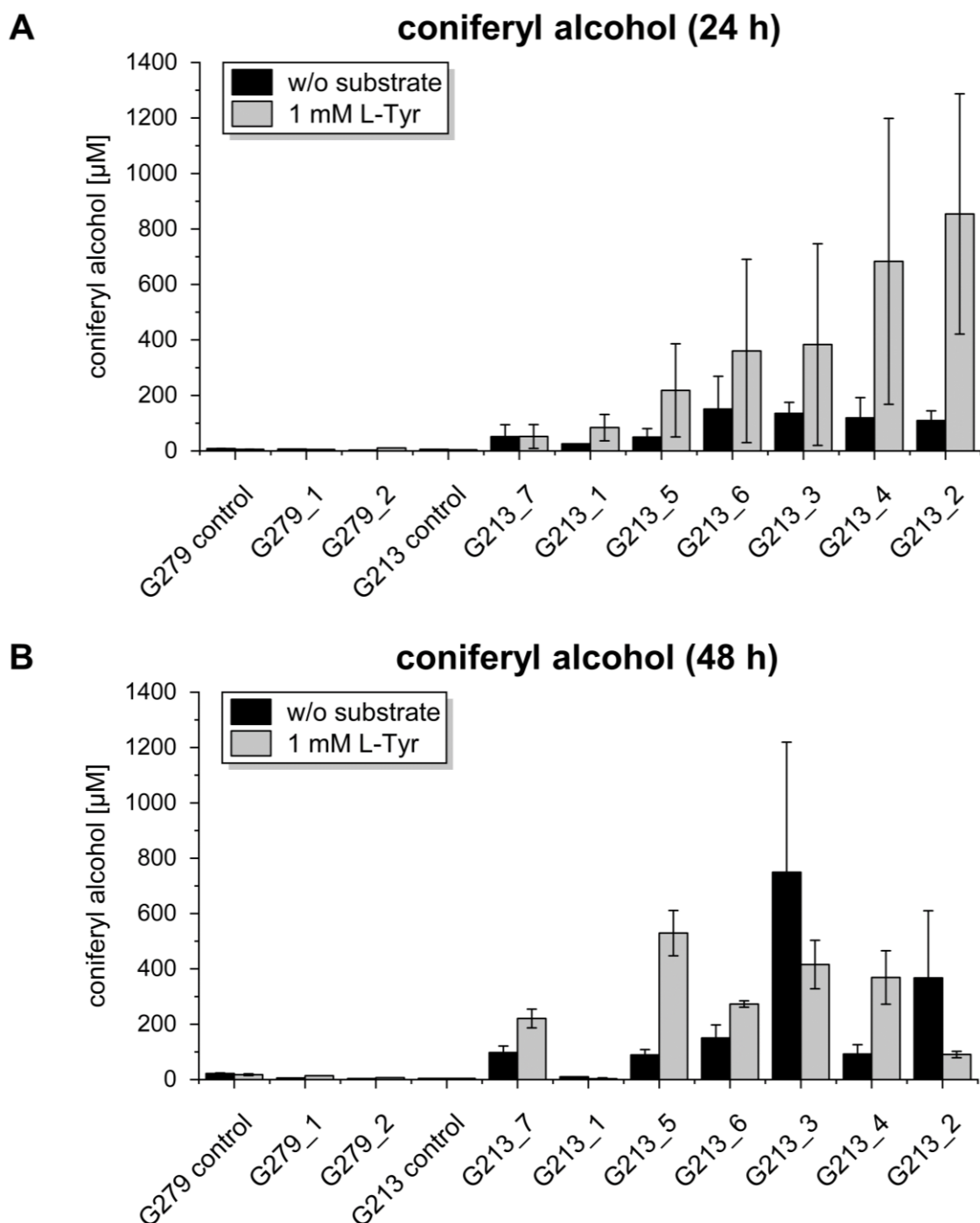
A possible obstacle regarding the applied plasmid-based expression is the transcription of all eight genes using the strong T7 promoters. The plasmid-based simultaneous production of eight proteins might impose metabolic stress on the *E. coli* host. Before, rather low expression levels with respect to the utilized ePathBrick plasmid had been observed in resting cells, too (cf. Figure A.7, Appendix). In growing cells, the enzymes *At4CL*, *ZmCCR* and *ZmCAD* seemed to be reasonably active in combination in contrast to the C3H. This suggests their predominant expression over the applied NΔ7-CYP199A2 variant, Pux and PuR.

Thus, it can be concluded that the expression levels of all enzymes in *E. coli* BL21(DE3) were likely imbalanced for a sufficient synthetic pathway performance towards the production of coniferyl alcohol. In order to reduce metabolic stress and possibly improve C3H expression, the two genome-integrated strains of *E. coli* BL21(DE3), which contain the 4CL, CCR and CAD downstream pathway enzymes under the control of weak constitutive promoters, are investigated in the following chapter.

### II.6.3 Production of coniferyl alcohol in genome-integrated *E. coli* strains

The two applied genome-integrated strains of *E. coli* BL21(DE3) are *E. coli* BL21(DE3).G213 and *E. coli* BL21(DE3).G279. The latter strain contains an additional integrated operon with the genes of *At4CL*, *ZmCCR* and *ZmCAD* in comparison to the first strain, which is also controlled by a weak constitutive promoter. In the tested strains, different TAL and OMT candidates were combined with both variants of the CYP199A2 enzyme and Pux/PuR as redox partners on expression plasmids.

For assessment of the performance of *E. coli* BL21(DE3).G279, this strain was co-transformed with pETM6\_*FjTAL*, pRSFDuet\_his6-Pux\_his6-PuR and pCDFDuet\_OMT\_NΔ7-CYP199A2 with *ZmCOMT* (G279\_1) or *ZmCCoAOMT* (G279\_2), respectively. The growing cells of these co-transformed G279 strains generally led to a marginal production of coniferyl alcohol (Figure 2.31). Instead, the concentration of *p*-coumaryl alcohol increased with the additional plasmid-based expression of *FjTAL* and was found to be higher in both G279 strains compared to the empty vector control of growing *E. coli* BL21(DE3).G213 (G213 control) and *E. coli* BL21(DE3).G279 cells (G279 control) (Figure A.14 B, Appendix). Also, only minor amounts of *p*-coumaric acid accumulated as intermediate in the G279 strains (Figure A.14 A, Appendix).



**Figure 2.31: Growing cells of *E. coli* BL21(DE3).G213 and *E. coli* BL21(DE3).G279.** The growing cell assay was performed in 10 mL TB medium at 26 °C without substrate addition (w/o, black) or with supplementation of 1 mM L-tyrosine (L-Tyr, grey) at time  $t_0$ . The concentration of coniferyl alcohol was measured at 24 h (A) and 48 h (B). The shown data represent mean values with standard deviation from at least biological duplicates. As exceptions, the data for G279\_1, G279\_2, the G213 control and G213\_1 (w/o substrate) originated from single measurements. The investigated enzyme combinations in the G279 strain consisted of *F<sub>7</sub>*TAL, NA7-CYP199A2 with Pux/PuR and *Zm*COMT (G279\_1) or *Zm*CCoAOMT (G279\_2). Growing cells using the three empty vectors served as activity controls (G213/G279 control). In the G213 strain, Pux/PuR was combined with the following: G213\_1 – *R<sub>g</sub>*TAL + CYP199A2 + *At*CCoAOMT, G213\_2 – *F<sub>7</sub>*TAL + CYP199A2 + *At*CCoAOMT, G213\_3 – Sesam8 + CYP199A2 + *At*CCoAOMT, G213\_4 – *F<sub>7</sub>*TAL + NA7-CYP199A2 + *Zm*COMT, G213\_5 – Sesam8 + NA7-CYP199A2 + *Zm*COMT, G213\_6 – *F<sub>7</sub>*TAL + NA7-CYP199A2 + *Zm*CCoAOMT, G213\_7 – Sesam8 + NA7-CYP199A2 + *Zm*CCoAOMT.

Apart from that, significant coniferyl alcohol production was observed in growing cells of *E. coli* BL21(DE3).G213. The strains G213\_1 to G213\_3 correspond to the previously investigated G213 strains in the resting cell assay, which expressed the enzymes *AtCCoAOMT*, *CYP199A2*, *Pux/PuR* and either *RgTAL* (G213\_1), *FjTAL* (G213\_2) or *Sesam8* (G213\_3) (cf. chapter II.5.2). When re-evaluated in the growing cell assay, the G213\_2 strain exhibited the generally highest production of coniferyl alcohol with a concentration of  $854.1 \pm 432.8 \mu\text{M}$  at 24 h and with the addition of 1 mM L-tyrosine (Figure 2.31 A). The determined high standard deviation was probably the result of three independently performed growing cell assays, which differed in induction  $\text{OD}_{600}$  (1.73, 1.80 and 2.37) and could have been also influenced by other factors.

Differences in coniferyl alcohol titer were observed, even if multiple growing cell assays were performed simultaneously and their samples extracted and measured under the same conditions and at the same time (e.g.  $902.5 \mu\text{M}$  and  $764.6 \mu\text{M}$  coniferyl alcohol for G213\_2 at 24 h). In this context, high standard deviations that occur in the following can be attributed to this variability in assay performance and presumably the overall expression of the individual pathway enzymes.

Without any added substrate, the coniferyl alcohol titer in the G213\_2 strain was found to be around eight times lower at 24 h ( $109.7 \pm 34.8 \mu\text{M}$ ) compared to the addition of 1 mM L-tyrosine (Figure 2.31 A). But the titer without substrate increased around three-fold to  $367.5 \pm 242.6 \mu\text{M}$  coniferyl alcohol after 48 h (Figure 2.31 B). An even higher production of coniferyl alcohol without externally added substrate was achieved by G213\_3 with a concentration of  $749.6 \pm 469.8 \mu\text{M}$  at 48 h.

In contrast to the strains G213\_2 and G213\_3, the G213\_1 strain with *RgTAL* showed low coniferyl alcohol titers, although this strain had produced the highest concentration of caffeic acid in the corresponding whole-cell biotransformation (chapter II.5.2). A simultaneous low production of *p*-coumaryl alcohol suggests further that the *in vivo* activity of *RgTAL* was insufficient in this strain under growing cell assay conditions (Figure A.14 B, Appendix).

*FjTAL* and *Sesam8*, which led to improved product titers in the upstream pathway evaluations, were evaluated in combination with the  $\text{N}\Delta 7$ -*CYP199A2* variant and the methyltransferases from *Z. mays* *ZmCOMT* (G213\_4 (*FjTAL*), G213\_5 (*Sesam8*)) and *ZmCCoAOMT* (G213\_6 (*FjTAL*), G213\_7 (*Sesam8*)), respectively (cf. chapter II.3.1). The G213\_4 strain thereby produced the second highest concentration of coniferyl alcohol ( $683.5 \pm 515. \mu\text{M}$ ) under the supplementation of 1 mM L-tyrosine at 24 h. This value decreased afterwards, which was

observed for some titers from the growing cell assays of *E. coli* BL21(DE3).G213, but other values showed an increase from 24 h to 48 h instead. In this way, the coniferyl alcohol concentration of  $529.6 \pm 81.9 \mu\text{M}$  was generated with the G213\_5 strain after 48 h and was found to be the third highest titer for coniferyl alcohol by using growing cells and an additional feed of 1 mM L-tyrosine (Figure 2.31 B).

In general, *p*-coumaric acid only accumulated in low amounts and was found to reach higher concentrations in enzyme combinations, which produced less coniferyl alcohol at 24 h (Figure A.14 A, Appendix). As another side-product, *p*-coumaryl alcohol was detected in a wide concentration range in the different tested strains and similar observations regarding high values and standard deviations were made again, which had been discussed previously (chapter II.6.2) (Figure A.14 B, Appendix). Other intermediates such as caffeoyl alcohol did not accumulate during the growing cell assays and only traces of caffeic acid were detected with the strains G213\_5-G213\_7 (data not shown).

The results indicate that the downstream pathway enzymes comprising 4CL, CCR and CAD exhibited a combined high activity in the G279 strains and expression of *FjTAL* generally favored *p*-coumaryl alcohol production. With respect to the empty vector control of the G213 strain, which generated marginal *p*-coumaryl alcohol titers, the low activity of the genome-integrated *RsTAL* is assumed to be the bottleneck in this strain. Nonetheless, the overall pathway performance towards *p*-coumaryl alcohol was shown to be improved by the introduction of additional gene copies of 4CL, CCR and CAD in the *E. coli* BL21(DE3).G279 compared to the G213 strain. However, it is assessed that the co-transformed C3H displayed insignificant activity in the G279 strains, although SDS-PAGE and western blot analysis indicated the expression of NΔ7-CYP199A2, Pux and PuR under growing cell assay conditions (Figure A.13 B and A.15, Appendix).

The higher titers of coniferyl alcohol with the *E. coli* BL21(DE3).G213 strains demonstrate that a total lower expression of the genome-integrated downstream pathway enzymes 4CL, CCR and CAD favors the production of the intended monolignol. The newly co-transformed G213 strains (G213\_4-G213\_7) showed distinct expression bands for the plasmid-based produced TAL, C3H and OMT enzymes (Figure A.16, Appendix). Thereby, it seemed that the expression bands of NΔ7-CYP199A2 and PuR were more pronounced in co-expression with *FjTAL* compared to Sesam8. The apparent higher expression level of the used C3H could be a reason for the superior performance of the strains G213\_4 and G213\_6, which both expressed *FjTAL*, compared to the strains G213\_5 and G213\_7 with Sesam8 at 24 h (Figure 2.31 A). At the end

of the growing cell assay, G213\_4 yielded a 35.1 % higher titer of coniferyl alcohol compared to G213\_6 and G213\_5 a 140 % higher concentration compared to G213\_7. (Figure 2.31 B). Thus, a slightly higher production of coniferyl alcohol was achieved after 48 h by utilizing *ZmCOMT* instead of *ZmCCoAOMT*.

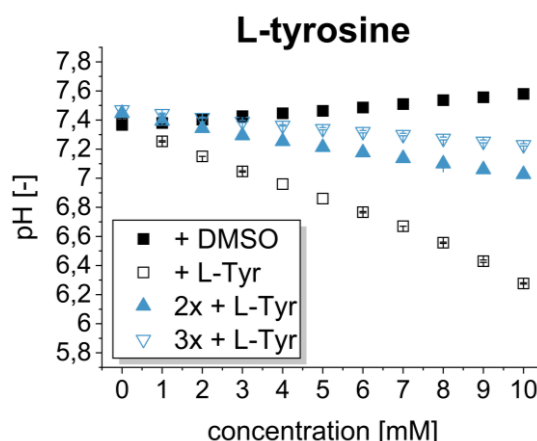
Interestingly, the enzyme combination of *FjTAL* with *AtCCoAOMT* and CYP199A2 (G213\_2) produced a similar concentration of coniferyl alcohol at 24 h as the G213 strain with *FjTAL*,  $\Delta 7$ -CYP199A2 and *ZmCOMT* (G213\_4). These observations may imply that *AtCCoAOMT* and *ZmCOMT* exhibit equal activities while co-expressed with seven other pathway enzymes in the *E. coli* BL21(DE3).G213 strain.

Due to the comparable activities of the CYP199A2 variants in the previous whole-cell biotransformations, the variants of C3H from *R. palustris* are not suspected to be the determining factor for the overall pathway performances (chapter II.2.2). Though, further investigations will be necessary for confirmation. The expression of the plasmid-based genes and the transcription of all genes had been demonstrated for the strains G213\_1-G213\_3 before (cf. chapter II.5.2). Thus, expression analysis of these strains and mRNA isolation with reverse transcription and cPCR, in general, were not repeated here. The 4.6- to 10.1-fold lower coniferyl alcohol titers of the G213\_1 strain with 1 mM L-tyrosine at 24 h compared to the strains G213\_3 and G213\_2, respectively, demonstrated the insufficient pathway performance when using *RgTAL* instead of *FjTAL* or Sesam8. Regarding an improvement of coniferyl alcohol titers, the presented growing cell assays are examined with higher substrate feeds in the last chapter (chapter II.6.4).

## II.6.4 Increase of coniferyl alcohol production

### Evaluation of pH from TB medium

For whole-cell biotransformation, the pH was found to increase or decrease with applied higher substrate concentrations of L-tyrosine (chapter II.1.2). With respect to an increased substrate feed in the growing cell assay, the pH of TB media with simple, doubled (2x) or tripled (3x) concentration of phosphate buffer salts was determined for different L-tyrosine concentrations (Figure 2.32). With 5 mM L-tyrosine, the pH dropped from 7.38 to 6.86 in the normal TB medium, while the pH decreased less in 2x (pH 7.21) or 3x TB medium (pH 7.34). Thus, the pH drop in simple TB medium was two to four times more pronounced compared to the 2x and 3x TB medium, respectively. Due to the already significant improvement with two-fold salt concentration compared to the common TB medium, 2x TB medium was used for the growing cells with 5 mM L-tyrosine.



**Figure 2.32: Influence of L-tyrosine on pH of different TB media.** 10 mL of TB medium (black) and TB media with two-fold (2x) or three-fold (3x) higher salt concentration (blue) were mixed with L-tyrosine (L-Tyr) in different concentrations and the pH was measured. The stock solutions were prepared as 100 mM L-Tyr in DMSO and acidified with HCl. As control only DMSO was added to TB medium (black filled square). All individual experiments were performed in triplicates the error bars indicate the standard deviation from the mean.

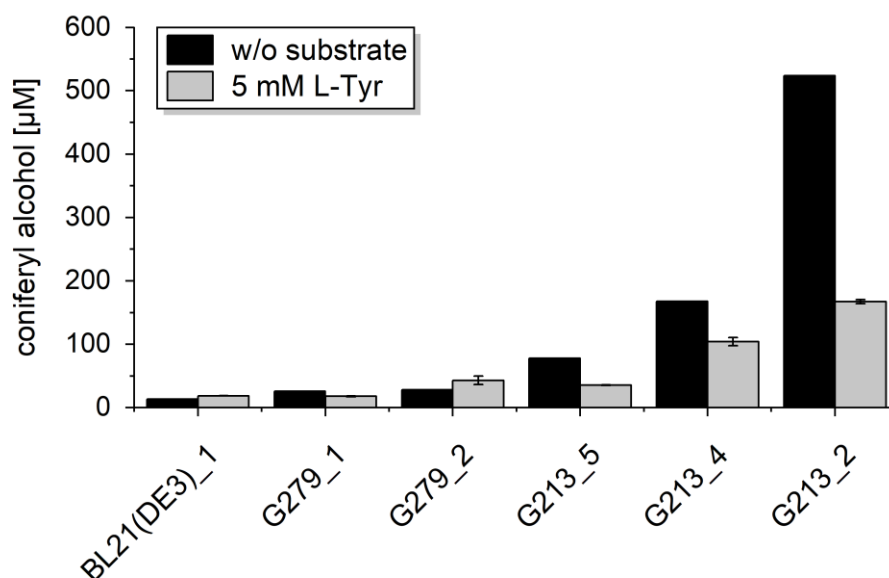
### Increased substrate feeds

For the investigation of an elevated substrate concentration, the three *E. coli* BL21(DE3).G213 strains G213\_2, G213\_4 and G213\_5 were chosen, which had exhibited the highest titers of coniferyl alcohol in growing cell assays (chapter II.6.3). In addition to these G213 strains, G279\_1 and G279\_2 as well as BL21(DE3) strain 1 were also tested in growing cells for potentially improved pathway performances towards coniferyl alcohol (cf. chapter II.6.2).

In growing cells in 2x TB medium, all coniferyl alcohol and most intermediate concentrations were generally found to be improved from 24 h to 48 h, while some highest measured titers had been determined at 24 h previously (cf. Figure 2.31). When growing cells were fed 5 mM L-tyrosine, the highest yielded titers for coniferyl alcohol using the G213 strains were far lower compared to the former growing cell assay with 1 mM L-tyrosine (Figure 2.33). In this way, the concentrations for coniferyl alcohol were around five- and seven-fold lower for the strains G213\_2 ( $167.3 \pm 3.2 \mu\text{M}$ ) and G213\_4 ( $104 \pm 6.4 \mu\text{M}$ ), respectively, compared to the highest titers with the addition of 1 mM L-tyrosine (cf. Figure 2.31). The production of coniferyl alcohol was found to be even 15-fold lower for the strain G213\_5 ( $35.6 \pm 0.4 \mu\text{M}$ ) (Figure 2.33).

Without externally added substrate, the coniferyl alcohol titers in growing cells were determined to be higher compared to the addition of 5 mM L-tyrosine. Thus, the strain G213\_2 produced a three-fold higher amount of coniferyl alcohol ( $523.5 \mu\text{M}$ ) with no added substrate compared to the supplementation of 5 mM L-tyrosine. The measured coniferyl alcohol titers without substrate after 48 h were thereby comparable to the previous results, when simple TB medium was used as sole substrate source (cf. Figure 2.31).





**Figure 2.33: Growing cells of different *E. coli* BL21(DE3) strains with higher L-tyrosine feed.** The growing cell assay was performed in 10 mL 2x TB medium at 26 °C without substrate addition (w/o, black) or with supplementation of 5 mM L-tyrosine (L-Tyr, grey) at time  $t_0$ . The concentration of coniferyl alcohol was measured at 48 h. The shown data with added substrate represent mean values with standard deviation from biological duplicates, while the data w/o substrate originated from a single measurement. The *E. coli* BL21(DE3) strain 1 (BL21(DE3)\_1) contained the enzymes *Zm*COMT and *NΔ7*-CYP199A2 with Pux/PuR and was co-transformed further with pETM6 containing *Fj*TAL, *At*4CL, *Zm*CCR and *Zm*CAD. The investigated enzyme combinations in the G279 strain consisted of *Fj*TAL, *NΔ7*-CYP199A2 with Pux/PuR and *Zm*COMT (G279\_1) or *Zm*CCoAOMT (G279\_2). In the G213 strain, Pux/PuR was combined with the following: G213\_2 – *Fj*TAL + CYP199A2 + *At*CCoAOMT, G213\_4 – *Fj*TAL + *NΔ7*-CYP199A2 + *Zm*COMT, G213\_5 – Sesam8 + *NΔ7*-CYP199A2 + *Zm*COMT.

All other strains exhibited low coniferyl alcohol titers, which were below the standard curve and comparable to the preceding production values. Thus, the strains BL21(DE3)\_1 ( $18.7 \pm 0.2 \mu\text{M}$ ), G279\_1 ( $17.9 \pm 0.5 \mu\text{M}$ ) and G279\_2 ( $43.2 \pm 6.5 \mu\text{M}$ ) showed no further improved coniferyl alcohol titer from the addition of 5 mM L-tyrosine. With the higher substrate feed, *p*-coumaric acid accumulated in high concentrations instead in these strains, which had not been observed before (Figure A.17 A, Appendix). Likewise, the titers for *p*-coumaryl alcohol as monolignol side-product increased in the strains of *E. coli* BL21(DE3).G279 in comparison to the former growing cell assay, while the titer was unchanged for *E. coli* BL21(DE3) strain 1 (Figure A.17 B, Appendix).

However, the determined *p*-coumaric acid concentrations were found to be three-fold higher in the strain G213\_2, when 5 mM L-tyrosine was applied as substrate instead of 1 mM (Figure A.17 A, Appendix). At the same time, the titer for *p*-coumaric acid increased around eight-fold for the strains G213\_4 and G213\_5. The produced amounts of *p*-coumaryl alcohol without added substrate were comparable to the previously measured values, while the *p*-coumaryl

alcohol titer decreased significantly with the high L-tyrosine supplementation (Figure A.17 B, Appendix).

These observations indicate that the co-transformed strains of *E. coli* BL21(DE3).G279 channeled the substrate L-tyrosine more significantly into the production of *p*-coumaryl alcohol rather than coniferyl alcohol. Thus, the activity of the C3H was still impaired in the G279 strains and likely also in *E. coli* BL21(DE3) strain 1 as discussed elsewhere (chapters II.6.2 and II.6.3). The high accumulation of *p*-coumaric acid in all tested strains with 5 mM L-tyrosine not only demonstrated that the 4CL enzyme was possibly inhibited, but also that C3H was insufficiently active. With respect to the observed inactivity of 4CL in whole-cell biotransformations, it is possible that this enzyme also lost its activity after 48 h in growing cells (cf. chapter II.4.3, II.5.2 and II.6.2).

Furthermore, the low amounts of *p*-coumaryl alcohol with 5 mM L-tyrosine in the G213 strains and the simultaneous lower coniferyl alcohol titers suggest an inhibition of the pathway in general. Neither caffeic acid nor caffeoyl alcohol were detected as intermediates, which leads to the assumption that the inhibited enzymes are the C3H and the downstream pathway enzyme 4CL. Regarding the decreased product concentrations, a possible inhibition of 4CL and C3H *in vivo* by L-tyrosine has not been investigated and can thus not be excluded.

Additionally, a feed of 2 mM cinnamic acid was investigated in the growing cell assay. The concentration of cinnamic acid was chosen based on the low toxicity on *E. coli* cells (Figure A.19 D, Appendix). All cinnamic acid was consumed after 48 h, but only marginal amounts of coniferyl alcohol and *p*-coumaric acid as well as low amounts of *p*-coumaryl alcohol were detected (Figure A.18, Appendix). All product titers were found to be far lower compared to any respective concentration in the previous growing cell assays. Consequently, cinnamic acid could have inhibited all pathway enzymes, as the production of all products and intermediates was impaired. When applying cinnamic acid as substrate, the highest titers for coniferyl alcohol were detected with the three G213 strains, which could presumably result from residual activity of the pathway enzymes using L-tyrosine from the TB medium. The direct conversion of cinnamic acid into some of the detected products and intermediates could not be confirmed. In the end, an increase of coniferyl alcohol production with respect to higher substrate supplementation of L-tyrosine and cinnamic acid was not successful.

### III. General discussion

The microbial production of the lignan precursor coniferyl alcohol necessitates the transfer of the plant phenylpropanoid pathway into a bacterial host such as *E. coli*. Some alternative pathway enzymes are required to be investigated, as a consequence of the adaption to the bacterial environment and for the purpose of a simplified pathway in order to enable efficient substrate to product conversions. Thus, different enzyme candidates for the reaction steps of TAL/PAL, C3H, 4CL and OMT were selected for evaluation.

A successful implementation of this multi-enzyme pathway involves the simultaneous expression of all pathway enzymes in sufficient individual amounts. The necessary enzymes encompassed seven to eight proteins for the complete synthetic phenylpropanoid pathway depending on the choice of enzyme candidates. So, in addition to the analysis of some individual enzymatic activity and expression patterns (chapter III.1), the pathway activities and expression profiles of different enzyme combinations needed to be examined (chapters III.2-III.4). Only in this way, a heterologous *E. coli* strain could be generated, which efficiently produces coniferyl alcohol and can serve as a prototype strain for further fermentation studies.

#### III.1 Individual enzyme characterizations

In general, whole-cell biotransformation was chosen for evaluation of the *in vivo* activities, as it has been applied as a common procedure in several literature reports [22,94]. Also, no substrates of TAL and PAL such as L-tyrosine and L-phenylalanine are present in the buffer unlike in a cultivation in TB medium. In the resting cell assay, the *E. coli* cells are in a non-growing state and solely survive by the conversion of added glucose. This experimental setup offers the possibility to investigate *in vivo* activity of an individual enzyme by using a specific substrate and with minimalized influence of potential inhibiting compounds.

In some studies, TAL has been identified as bottleneck, which encouraged the search for alternative TAL candidates with a higher and specific activity [65,135]. Furthermore, both bacterial C3H candidates HpaBC and CYP199A2 were tested in combination with a chosen TAL enzyme in the past. But a simultaneous and comparative analysis of multiple TAL and C3H enzymes individually and combined as upstream pathway was scarce in literature [132,157].

### III.1.1 Deamination

The different TAL and PAL candidates were the bacterial enzymes *RsTAL*, *HaTAL*, *FjTAL* and Sesam8, the fungal *RgTAL* as well as the plant *ZmPAL*. According to their PAL/TAL ratios, the bacterial enzymes were determined to have explicit TAL activities, while the candidates from fungi and monocot with a low PAL/TAL ratio were shown to accept L-tyrosine and L-phenylalanine [56,58,65,141]. By comparing the measured TAL *in vivo* activities, *RgTAL* exhibited the highest production of *p*-coumaric acid followed by the other bacterial TAL enzymes (chapter II.1.1). The exception was the performance by *RsTAL*, which was found to be the lowest of all tested TALs. These observations were in accordance with other studies, which reported the poor deamination activity of *RsTAL* towards L-tyrosine *in vivo* [65,140].

Interestingly, *HaTAL* produced a similar concentration of *p*-coumaric acid in the whole-cell biotransformation as in a previous growing cell assay in M9 medium supplemented with 2 mM L-tyrosine (130  $\mu$ M/OD<sub>600</sub>) [65]. In this particular growing cell assay, *FjTAL* and Sesam8 also led to three-fold higher production values of *p*-coumaric acid in comparison to *HaTAL*. Furthermore, Jendresen and co-workers (2015) observed a six- to seven-fold lower *in vivo* productivity for *HaTAL* compared to *FjTAL* and Sesam8, when L-tyrosine was added in the late exponential phase [65].

The SDS-PAGE analysis showed the soluble expression of all TAL/PAL enzymes in pronounced bands except for *RsTAL* and *HaTAL*, which had comparably thin protein bands (Figures 2.2). As a consequence, *RsTAL* and *HaTAL* were not further evaluated due to low *in vivo* protein amounts and their measured as well as reported low TAL activities.

In another study, *RgTAL* was applied in a growing cell assay and with an increased substrate feed (3 mM L-tyrosine) [132]. The determined *in vivo* activity was moderately higher to the herein reported *p*-coumaric acid titer with respect to fed total L-tyrosine (873  $\mu$ M *p*-coumaric acid/ 1 mM L-tyrosine) (cf. Figure 2.1). Though, it has to be noted that Rodrigues *et al.* (2015) used the pRSFDuet vector, which has an around 2.5 times higher copy plasmid number compared to the used pETM6 vector [132]. With the application of the pETDuet vector, their results were similar to the achieved *p*-coumaric acid titers in this work ( $\sim$  330  $\mu$ M *p*-coumaric acid/ 1 mM L-tyrosine) [132].

In addition, *RgTAL* proved to be capable of converting 50 g/L L-tyrosine (276 mM) into 44.4 g/L *p*-coumaric acid (270 mM) in a whole-cell biotransformation at 30 °C, which equaled almost complete substrate conversion [141].

Regarding the PAL activity of the used bifunctional enzymes, an around five to seven times higher cinnamic acid production was achieved with *RgTAL* instead of *ZmPAL*. These results demonstrated the superior *in vivo* TAL and PAL functionality of *RgTAL* (chapter II.1.3). Furthermore, the calculated PAL/TAL ratio was around 0.66-1.02 at neutral pH and *in vivo*, although the pH optimum of *RgTAL* had been determined to be more basic (pH 9.0) [141]. In *in vitro* studies, Vannelli and co-workers (2007) calculated a PAL/TAL ratio of 0.83 at pH 8.5, while Xue *et al.* (2007) determined a PAL/TAL ratio of 10.12 at pH 8.5 and of 0.86 at pH 9.5 for *RgTAL*. These reported *in vitro* PAL/TAL ratios for *RgTAL* partly correspond to the determined PAL/TAL ratio *in vivo*, although increased pH values were applied in these studies. The differences could probably be attributed to the whole-cell environment. This could have led to the varying results between *in vitro* and *in vivo* experiments and likely stabilized the TAL enzyme at lower pH values *in vivo* compared to the determined optimal pH value *in vitro*.

For determination of the *in vivo* activity, *RgTAL* had been expressed in LB medium before and had been capable to produce 630  $\mu\text{M}$  *p*-coumaric acid and 840  $\mu\text{M}$  cinnamic acid solely from both amino acids being present in the complex medium and partly from *de novo* synthesis [58]. These product concentrations were in accordance with the determined values in the above performed resting cell assays, if the individual concentrations for L-tyrosine (1.0-1.6 mM)<sup>1</sup> and L-phenylalanine (3.3-3.8 mM)<sup>1</sup> in LB medium are taken into account.

Only a few reports characterized the *ZmPAL* enzyme which was shown to have a lower  $K_m$  value towards L-tyrosine (41  $\mu\text{M}$ ) compared to L-phenylalanine (658  $\mu\text{M}$ ), but similar catalytic efficiencies resulting in a PAL/TAL ratio of 0.78 [56]. The suspected slightly higher activity as TAL enzyme was demonstrated in growing cells with 1 mM L-tyrosine, which led to the production of 388  $\mu\text{M}$  *p*-coumaric acid and only 11  $\mu\text{M}$  cinnamic acid by *ZmPAL* [188]. In this study, the resulting *p*-coumaric acid concentration was also found to be three-fold higher compared to the value of *RsTAL* with L-tyrosine as substrate.

These reports illustrated that *ZmPAL* could rather be considered as a TAL enzyme [56,188]. Moreover, *RgTAL* was capable of conversion of both substrates with comparably high efficiency in resting and growing cells [58,141].

---

<sup>1</sup> Calculated from specifications of three individual tryptones and two yeast extracts [162–166].

### III.1.2 C3-hydroxylation

#### The hydroxylation complex HpaBC from *E. coli*

The direct C3-hydroxylation of *p*-coumaric acid into caffeic acid proved to be successful with the usage of the HpaBC complex from *E. coli* and the cytochrome P450 monooxygenase CYP199A2 from *R. palustris* (chapter II.2.1 and II.2.2). An empty vector control demonstrated that the native HpaBC complex is silenced in *E. coli*, as no activity towards *p*-coumaric acid was detected in resting and growing cells [101]. In a whole-cell biotransformation from Lin and Yan (2012), the HpaBC complex produced around 241  $\mu$ M caffeic acid within the first hour [101]. This apparently constitutes a much higher initial conversion rate compared to the measured concentrations from the HpaBC activity in this thesis. The discrepancy between these results can be reasoned with the applied shorter expression time of three hours and the lowered expression temperature of 30 °C probably leading to an improved enzyme stability [101]. It can be speculated that the HpaBC complex could exhibit a higher activity at early expression time points and at a milder expression temperature, as it was demonstrated for CYP199A2. At 37 °C, CYP199A2 with Pux/PuR showed a comparable activity to HpaBC in monocistronic and pseudo-operon gene configuration, while the conversion of *p*-coumaric acid was improved, when the expression temperature had been reduced to 26 °C (cf. Figure 2.6 and 2.8). With respect to the highest production values and their standard deviations, both bacterial C3H candidates were found to be equally active and suited for the phenylpropanoid pathway.

The HpaBC complex accepted further *p*-coumaryl alcohol as substrate, whereby half of the product concentration was generated compared to *p*-coumaric acid (cf. Figure A1, Appendix). These results stand in contrast to the report from Chen *et al.* (2017), who found a 1.5 times higher activity towards *p*-coumaryl alcohol than *p*-coumaric acid [137]. The difference to the here presented study might be attributed to the performance as a growing cell assay. Using growing cells it was also possible to convert *p*-coumaric acid completely into a total of 21.3 mM caffeic acid with HpaBC [102].

Nevertheless, in another whole-cell biotransformation using the HpaBC complex from *P. aeruginosa*, a caffeic acid titer of 56.6 mM was achieved via successive feeding of 80 mM *p*-coumaric acid [103]. This titer constitutes the highest for a direct conversion of *p*-coumaric acid into caffeic acid with an HpaBC complex in *E. coli* to date.

At the same time, the applied large feed of *p*-coumaric acid puts cytotoxicity effects on *E. coli* into question. Regarding cell growth determined by means of OD<sub>600</sub>, a concentration of 10 mM *p*-coumaric acid was found to inhibit growth on the one hand, while this effect has been reported

to occur starting from a value of 21 mM *p*-coumaric acid in another study [102]. On the other hand, Xue and co-workers (2007) observed cytotoxicity of *p*-coumaric acid resulting in cell death at a value of 61 mM [141]. As demonstrated by Furuya and Kino (2014), successive feeding can prevent high substrate concentrations in the medium and promote efficient conversions by HpaBC [103]. Due to the occurrence of comparably lower substrate feeds of up to 5 mM in this thesis, toxicity of the tested phenylpropanoids (except cinnamic acid) on *E. coli* should not be an issue.

### **The cytochrome P450 monooxygenase from *R. palustris***

For the P450 enzyme from *R. palustris*, a full length and a NΔ7 variant were investigated and found to exhibit equal conversion values of *p*-coumaric acid with the native redox partners Pux and PuR (chapter II.2.2). However, NΔ7-CYP199A2 catalyzed the C3-hydroxylation faster than CYP199A2. In literature, the NΔ7 variant was capable of converting 1 mM *p*-coumaric acid in only 40 min and by using a slightly lower cell concentration of 50 g<sub>wcw</sub>/L, which represents an even faster conversion rate compared to the results in this work [94]. A probable reason for the faster conversion could be a higher proportion of active enzyme due to a lower expression time of 12 h and the addition of 5-aminolevulinic acid in order to promote heme synthesis [189]. Notably, Furuya *et al.* (2012) used the redox partners Pux and PdR. In a later study, the native redox pair Pux/PuR was shown to support the C3H activity of NΔ7-CYP199A2 more efficiently in comparison to Pux/PdR [157]. The production of caffeic acid in this growing cell assay with respect to total fed substrate (~ 530 μM caffeic acid/1 mM *p*-coumaric acid) was comparable to the caffeic acid titers determined by whole-cell biotransformations of the CYP199A2 variants in this thesis (cf. chapter II.2.2).

Interestingly, the redox partners Pdx/PdR were able to support the *in vivo* activity of CYP199A2, which resulted in a similar caffeic acid concentration compared to the redox pair Pux/PuR, though supported only marginally the activity of the NΔ7 variant (chapter II.2.2). This observation was in accordance with the report from Haslinger and Prather (2020), who also failed to determine activity of NΔ7-CYP199A2 with Pdx/PdR. Taken together, the here presented analyses of different redox partners and the results from Haslinger and Prather (2020) with mixed redox partners suggested two possible implications [157]. Firstly, the electron transfer within the pairs Pdx/PdR and Pux/PuR could occur more efficient in contrast to the mixed redox pair Pux/PdR. With respect to an *in vitro* study, this hypothesis was not confirmed, as the NADH consumption and product formation rate of wildtype CYP199A2 in combination with Pux/PdR were determined to be considerably higher on the tested alternative substrates

compared to Pdx/PdR [158]. It was also stated that the superior activity of the P450 enzyme with Pux originates more likely from its tighter binding rather than from a faster electron transfer [190]. Secondly, the reductase PdR could be unfavorably oriented in the three protein complexes of NΔ7-CYP199A2 with either Pux/PdR or Pdx/PdR, which could lead to a reduced electron transfer efficiency from the FAD cofactor to the heme center (cf. Figure I.5). After all, the cytochrome P450 monooxygenase from *R. palustris* is dependent on receiving electrons via an electron transfer chain, which is sensitive to interference either between the reductase and redoxin or redoxin and the heme center.

The conversion activity of the full length P450 enzyme with Pdx/PdR demonstrated that the seven additional N-terminal amino acids were likely responsible for binding of the redox partner pair Pdx/PdR. This restored the electron transfer chain in CYP199A2 and thus its C3H activity. The exact reason for the activity recovery of CYP199A2 via its N-terminus remains unknown. It can be speculated that the seven amino acids are involved in van der Waals and hydrogen bond contacts, which were assumed to be present in interactions of Pdx and its natural P450 enzyme CYP101 in previous studies [159,191,192]. Consequently, the absence of these complementary interacting bonds in NΔ7-CYP199A2 could prevent Pdx from proper binding to the P450 enzyme's surface in this case.

The surrogate redox pair YkuN/FpR has been successfully applied with other cytochrome P450 monooxygenases, but was found to be inefficient in supporting activity of CYP199A2 [85,156]. It was suspected that this might have arisen from an impaired electron transfer from the flavodoxin YkuN to the P450 reaction center and from a possibly low intracellular NADPH pool (cf. Figure I.5). If low concentrations of NADPH are the cause, then the repression of transhydrogenases, which normally convert NADPH to NADH, and the simultaneous upregulation of such transhydrogenases, that phosphorylate NADH into NADPH, could improve the caffeic acid production from CYP199A2 with YkuN and the NADPH-dependent FpR. By applying this strategy, Li *et al.* (2018) achieved an eight- to nine-fold increased *de novo* titer of *p*-coumaric acid when using a PAL and a C4H fusion protein with reductase 2 from *A. thaliana* [134]. In the other mentioned case, the flavodoxin YkuN may be incompatible with CYP199A2. This could indicate that the P450 enzyme from *R. palustris* constitutes a true class I cytochrome P450 monooxygenase and can rather not be assigned to a class III redox system [84]. The discussed aspects around CYP199A2 emphasize all the more that the characterization of specific redox partners is necessary for each P450 enzyme individually.



## III.2 Upstream pathway

In the end, the first enzymatic steps in the upstream pathway have to be efficient with respect to high substrate to product conversions. As a result, *RgTAL*, *FjTAL* and *Sesam8* were tested in combination with the C3H enzyme candidates *HpaBC* and *CYP199A2* *in vivo* in *E. coli* (chapter II.3.1). This led to the evaluation of the favorable enzyme combination for yielding the most efficient production of caffeic acid from L-tyrosine. In addition, the upstream pathway starting from L-phenylalanine has been investigated by using the PAL activity of *RgTAL* and the C3H/C4H bifunctionality of *CYP199A2* (chapter II.3.2).

Also, Table 1 summarizes a number of studies, which have already established the *in vivo* production of *p*-coumaric and caffeic acid from different substrates or *de novo* in *E. coli*.

**Table 1: *In vivo* production of phenylpropanoic acids in *E. coli*.**

Substrate/ Strain	Enzymes (organism)	Production conditions → product titer	Reference
<i>p</i> -coumaric acid and cinnamic acid			
LB medium <i>E. coli</i> W3110	<i>RgTAL</i>	<ul style="list-style-type: none"> <li>Growing cells in LB medium at 30 °C.</li> </ul> → 630 µM <i>p</i> -coumaric acid → 840 µM cinnamic acid	Vannelli <i>et al.</i> , 2007 [58]
L-tyrosine (3 mM) <i>E. coli</i> BL21(DE3)	<i>RgTAL</i> (a) <i>RgTAL</i> (mutant, b) <sup>1</sup>	<ul style="list-style-type: none"> <li>Growing cells in M9 medium with glucose at 28 °C.</li> </ul> → 1.06 mM <i>p</i> -coumaric acid (a) → 1.71 mM <i>p</i> -coumaric acid (b)	Zhou <i>et al.</i> , 2016 [193]
L-tyrosine (3 mM) <i>E. coli</i> K-12 MG1655(DE3)	<i>RgTAL</i>	<ul style="list-style-type: none"> <li>Pre-expression in LB medium at 26 °C for 5 h.</li> <li>Growing cells in M9 medium with glucose at 26 °C.</li> </ul> → 2.62 mM <i>p</i> -coumaric acid	Rodrigues <i>et al.</i> , 2015 [132]
L-tyrosine (2 mM) <i>E. coli</i> BL21(DE3) pLysS	<i>FjTAL</i> (a) <i>Sesam8</i> (b)	<ul style="list-style-type: none"> <li>Growing cells in M9 medium with glucose at 30 °C.</li> </ul> → 0.44 mM/OD <sub>600</sub> <i>p</i> -coumaric acid (a) → 0.54 mM/OD <sub>600</sub> <i>p</i> -coumaric acid (b)	Jendresen <i>et al.</i> , 2015 [65]
<i>de novo</i> <i>E. coli</i> W3110 derivative strain <sup>2</sup>	<i>RgTAL</i> (a) PAL2 ( <i>A. thaliana</i> , b)	<ul style="list-style-type: none"> <li>Growing cells in M9 medium with glucose at 37 °C.</li> </ul> → 91.5 µM <i>p</i> -coumaric acid + 196.2 µM cinnamic acid (a) → 530 µM cinnamic acid (b)	Vargas-Tah <i>et al.</i> , 2015 [194]

<i>de novo</i> <i>E. coli</i> W3110 (engineered) <sup>3</sup>	PAL ( <i>S. maritimus</i> )	<ul style="list-style-type: none"> <li>Fed-batch cultivation in modified fermentation medium with glucose (at 37 °C, pH 6.8) and with casamino acid feeding solution.</li> </ul> → 46.6 mM cinnamic acid	Bang <i>et al.</i> , 2018 [195]
<i>de novo</i> <i>E. coli</i> C41 (DE3) ΔCOS1 <sup>4</sup>	TAL ( <i>S. espanaensis</i> )	<ul style="list-style-type: none"> <li>Pre-expression in LB medium at 37 °C for 6 h.</li> <li>Growing cells in modified synthetic medium with glucose at 26 °C.</li> </ul> → 4.04 mM <i>p</i> -coumaric acid	Kang <i>et al.</i> , 2015 [196]
<i>de novo</i> <i>E. coli</i> C41 (DE3) (engineered) <sup>4</sup>	TAL ( <i>S. espanaensis</i> )	<ul style="list-style-type: none"> <li>Pre-expression in LB medium at 37 °C for 5 h.</li> <li>Growing cells in modified M9 medium with glucose at 26 °C.</li> </ul> → 5.94 mM <i>p</i> -coumaric acid	Kang <i>et al.</i> , 2012 [129]
<i>de novo</i> <i>E. coli</i> Rosetta (DE3)	PAL1 ( <i>A. thaliana</i> ) + ATR2-LauC4H <sup>5</sup>	<ul style="list-style-type: none"> <li>Growing cells in fermentation medium with glucose at 30 °C.</li> <li>Intracellular NADPH upregulation.</li> </ul> → 136.5 μM <i>p</i> -coumaric acid	Li <i>et al.</i> , 2018 [134]
<b>caffeic acid</b>			
<i>p</i> -coumaric acid (20 mM, <b>a</b> ) cinnamic acid (1 mM, <b>b</b> ) <i>E. coli</i> BL21(DE3)	NΔ7-CYP199A2 + Pux/PdR	<ul style="list-style-type: none"> <li>Expression in LB medium at 25 °C.</li> <li>Resting cells (50 mM KP<sub>i</sub>, pH 7.5, 9 g<sub>DCW</sub>/L) with glycerol at 30 °C.</li> </ul> → 15 mM caffeic acid ( <b>a</b> ) → 0.72 mM caffeic acid ( <b>b</b> )	Furuya <i>et al.</i> , 2012 [94]
<i>p</i> -coumaric acid (80 mM) <i>E. coli</i> BL21Star (DE3)	HpaBC ( <i>P. aeruginosa</i> )	<ul style="list-style-type: none"> <li>Expression in LB medium at 25 °C.</li> <li>Resting cells (200 mM KP<sub>i</sub>, pH 7.5, 8 g<sub>DCW</sub>/L) with glycerol at 30 °C were fed 20 mM substrate successively.</li> </ul> → 56.6 mM caffeic acid	Furuya and Kino, 2014 [103]
<i>p</i> -coumaric acid (21.3 mM, <b>a</b> ) <i>de novo</i> ( <b>b</b> ) <i>E. coli</i> BW25113 ( <b>a</b> ) <i>E. coli</i> ATCC31884 (engineered, <b>b</b> ) <sup>6</sup>	HpaBC ( <i>E. coli</i> , <b>a</b> ) + RgTAL ( <b>b</b> )	<ul style="list-style-type: none"> <li>Growing cells in M9 medium with yeast extract (0.5 %, w/v) and glycerol at 37 °C. (<b>a</b>)</li> <li>Growing cells in M9 medium with glucose and glycerol at 37 °C and supplemented L-phenylalanine. (<b>b</b>)</li> </ul> → 21.2 mM caffeic acid ( <b>a</b> ) → 4.26 mM caffeic acid ( <b>b</b> )	Huang <i>et al.</i> , 2013 [102]
<i>p</i> -coumaric acid (24.4 mM) <i>E. coli</i> BL21 Star (DE3)	HpaBC ( <i>E. coli</i> )	<ul style="list-style-type: none"> <li>Growing cells in AMM medium with glycerol at 37 °C were fed with substrate after 2 h of induction.</li> </ul> → 19.2 mM caffeic acid	Jones <i>et al.</i> , 2016 [100]

L-tyrosine (3 mM) <i>E. coli</i> K-12 MG1655(DE3)	RgTAL + NΔ7-CYP199A2 + Pux/PdR	<ul style="list-style-type: none"> <li>Pre-expression in LB medium at 26 °C for 5 h.</li> <li>Growing cells in M9 medium with glucose at 26 °C.</li> </ul> → 1.56 mM caffeic acid	Rodrigues <i>et al.</i> , 2015 [132]
<i>de novo</i> (a, c) L-tyrosine (3 mM, b) <i>E. coli</i> K-12 MG1655(DE3)	<i>Fj</i> TAL + NΔ7-CYP199A2 + Pux/PuR + Pux (c)	<ul style="list-style-type: none"> <li>Pre-culture in LB medium at 37 °C for 4 h.</li> <li>Growing cells in M9 medium with glucose at 26 °C.</li> </ul> → ~ 0.1 mM caffeic acid (a) → ~ 1.8 mM caffeic acid (b) → 0.26 mM caffeic acid (c)	Haslinger and Prather, 2020 [157]
<i>de novo</i> <i>E. coli</i> C41 (DE3) (a) <i>E. coli</i> C41 (DE3) (engineered, b) <sup>4</sup>	TAL ( <i>S. espanaensis</i> ) + C3H ( <i>S. espanaensis</i> )	<ul style="list-style-type: none"> <li>Pre-expression in LB medium at 37 °C for 6 h.</li> <li>Growing cells in modified synthetic medium with glucose at 26 °C.</li> </ul> → 233 μM caffeic acid (a) → 833 μM caffeic acid (b)	Kang <i>et al.</i> , 2012 [129]
<i>de novo</i> <i>E. coli</i> rpoA14 (DE3) <sup>7</sup>	HpaBC ( <i>E. coli</i> ) + RgTAL	<p>Growing cells in AMM medium with glycerol at 37 °C.</p> → 5.7 mM caffeic acid	Jones <i>et al.</i> , 2017 [133]

<sup>1</sup> triple mutant containing the mutations S9N/A11T/E518V.

<sup>2</sup> strain with inactive phosphoenolpyruvate(PEP):sugar phosphotransferase system resulting in increased PEP availability. Also overexpression of genes from the shikimate pathway (*aroG*, *tktA*) to increase the intracellular levels of L-phenylalanine and L-tyrosine.

<sup>3</sup> L-phenylalanine overproducing strain: deletion of various genes for avoiding PEP degradation, increasing precursor PEP pool and synthesis of L-tryptophan and L-tyrosine. Overexpression of genes from the shikimate pathway as well as genes for facilitated glucose uptake.

<sup>4</sup> L-tyrosine overproducing strain: overexpression of feedback-resistant genes from the shikimate pathway (*aroG*, *tyrA*) and deletion of *tyrR* (repressor of aromatic amino acid synthesis).

<sup>5</sup> fusion protein of ATR2 and C4H from *Lycoris aurea* without their N-terminal transmembrane domains.

<sup>6</sup> L-phenylalanine overproducing strain was further engineered by disruption of L-phenylalanine synthesis (*pheLA*) and with plasmid-based overexpression of certain genes from the shikimate pathway (e.g. feedback-resistant *tyrA* and *aroG*).

<sup>7</sup> L-tyrosine overproducing strain: overexpression of feedback-resistant genes from the shikimate pathway (*aroG*, *tyrA*), deletion of *tyrR* (repressor of aromatic amino acid synthesis) and *pheA* (L-phenylalanine synthesis gene).

### III.2.1 Combined activity of TAL and C3H enzymes

In whole-cell biotransformations with 1 mM L-tyrosine, all three tested TAL enzymes exhibited generally higher titers of caffeic acid when co-expressed with CYP199A2 instead of the HpaBC complex as C3H (chapter II.3.1). The marginal production of caffeic acid using the HpaBC complex was attributed to the pronounced expression bands of this C3H, which seemed to suppress the expression of the TAL proteins. In contrast, distinct expression bands for the TAL enzymes, CYP199A2 as well as the redox partners Pux and PuR were noticed in combinations with the P450 enzyme from *R. palustris* indicating sufficient produced amounts of all proteins.

The observation, that the TAL enzymes exhibited a generally faster conversion of L-tyrosine compared to the C3Hs CYP199A2 and HpaBC for *p*-coumaric acid, is consistent with literature [132]. With view on the different TALs in combination with CYP199A2, the highest caffeic acid titer was measured when applying *Fj*TAL followed by Sesam8 in whole-cell biotransformation. In a previous growing cell assay, NΔ7-CYP199A2 with both mentioned TALs led to *de novo* production of caffeic acid, while no caffeic acid was detected with *Rg*TAL [157]. By applying an additional feed of 3 mM L-tyrosine, equal caffeic acid concentrations were generated in all three TAL enzyme combinations. These reported caffeic acid titers from Haslinger and Prather (2020) were around two to three times higher compared to the determined values in this work [157]. Considering the total feeds of L-tyrosine though, the results from the herein presented whole-cell biotransformation are in a comparable range with the referenced growing cell assay in M9 medium [157]. This comparison also emphasizes that the caffeic acid production in the upstream pathway is independent from the application of the applied CYP199A2 variant, as has been demonstrated by Rodrigues and co-workers (2015) [132].

A later and separate induction of the P450 redox partners as well as the expression of an additional gene copy of Pux were found to exhibit optimization potential regarding an increased production of caffeic acid [132,157]. It is worth mentioning that 5-aminolevulinic acid was supplemented for the production of the cytochrome P450 enzyme from *R. palustris* in some studies [94,157]. However, it can be concluded that the addition of this heme precursor is not necessary due to the comparability of the presented results in this thesis with the literature. This will be also advantageous in terms of costs for a later large-scale fermentation.

Both tested C3H enzymes have been applied in several previous studies and partly compared to the first identified C3H from *S. espanaensis* (Sam5) [60]. In two reports with growing cells of *E. coli*, the activity of Sam5 was determined to be superior over the performance of the HpaBC complex from *E. coli* in LB as well as in M9 minimal medium [44,130]. Rodrigues *et al.* (2015) compared the production of caffeic acid from *p*-coumaric acid by Sam5 and NΔ7-CYP199A2 in growing cells and found a 1.8 times higher titer for the P450 enzyme, when its redox partners were induced separately at a later time point [132]. Combining both results indicated that CYP199A2 would also be the more optimal C3H regarding caffeic acid production in growing cells compared to the HpaBC complex. Hence, a comparative growing cell assay of the upstream pathway was not performed in this thesis. Another reason was the imbalanced expression between TAL and HpaBC and, because of this, it was assumed that the performance of TAL would not be improved in a combined growing cell assay. With the

exclusion of the HpaBC complex, the possible formation of other side-product such as L-DOPA, which originates from the enzyme's broad substrate specificity, can also be diminished.

In the whole-cell biotransformation, CYP199A2 with *RgTAL* produced caffeic acid in the same range as the other two TALs first, but the titer dropped significantly after 24 h. Two main reasons can be concluded for this decrease in titer. Firstly, it was suspected that *RgTAL* is more prone to product inhibition compared to *FjTAL* and *Sesam8*. In a study, *RgTAL* was competitively inhibited by *p*-coumaric acid and this effect had already occurred at a concentration of 100  $\mu$ M *p*-coumaric acid [141]. This can explain the low concentrations of *p*-coumaric acid accumulating in the combination with *RgTAL* and CYP199A2 at 24 h, which has been observed far less for *FjTAL* and *Sesam8*. Secondly, *RgTAL* was shown to accept L-DOPA as substrate [44]. Consequently, it is possible that caffeic acid was partly converted into L-DOPA as reverse reaction, which would coincide with the observed decreased concentrations of caffeic acid with *RgTAL*. Notably, the extraction of L-DOPA with ethyl acetate is unlikely, as L-DOPA resembles the chemical structure of L-tyrosine only with an additional hydroxyl group. Like L-tyrosine, L-DOPA could have escaped detection with the applied extraction method. It thus remains unclear, whether L-DOPA was generated in the end and whether *FjTAL* and *Sesam8* exhibit this side-activity as well. Due to these stated possible complications and to some studies reporting *RgTAL* as rate-limiting enzyme, eventually, *FjTAL* and *Sesam8* were chosen for the growing cell assay comprising the complete synthetic pathway [44,157].

### III.2.2 Combined activity of a PAL and C3H/C4H enzyme

CYP199A2 was examined for its ability to act as C4H and C3H simultaneously and a complete conversion of cinnamic acid into caffeic acid was observed for both variants of the P450 enzyme (chapter II.3.2). Furuya and co-workers (2012) achieved a conversion yield of 72 % for caffeic acid in 90 min using 1 mM cinnamic acid as substrate in resting cells, which was found to be a higher conversion yield compared to the earliest measured time point in this work (data not shown) [94].

Interestingly, the C3H activity of the  $\Delta 7$ -CYP199A2 variant was decreased and its C4H activity completely abolished in the presence of 1 mM L-phenylalanine, while both substrate activities remained intact for the full length CYP199A2 (cf. chapter II.2.2). Hence, it was suspected that the combination of CYP199A2 and the PAL activity from *RgTAL* would lead to a completion of the upstream pathway starting from L-phenylalanine. Surprisingly, low

concentrations of caffeic acid were solely detected in the enzyme combination of *RgTAL* and  $\Delta 7$ -CYP199A2 after 6 h and caffeic acid was not determined after 24 h anymore (chapter II.3.2). Instead, cinnamic acid accumulated as intermediate and its production values generally corresponded to the expected concentrations from the *in vivo* PAL evaluations with *RgTAL*.

Consequently, the PAL activity of *RgTAL* was not impaired and both variants of the P450 enzyme were insufficiently active or more likely inhibited. It remains questionable whether L-phenylalanine could act as inhibitor of the CYP199A2 variants and what type of inhibition is associated with L-phenylalanine.

According to the mentioned side-activity of *RgTAL*, it is possible that caffeic acid was consumed by *RgTAL* purposed for the generation of L-DOPA [44]. Thereafter, L-DOPA could have acted as inhibitor of both cytochrome P450 monooxygenase variants. This assumption was not further investigated and thus not verified in this thesis.

So far, only Li and co-workers (2018) were able to produce 136.5  $\mu\text{M}$  *p*-coumaric acid *de novo* in a wild-type *E. coli* strain by using *AtPAL1* and a C4H from *Lycoris aurea* [134]. To the best of the author's knowledge, direct synthesis of caffeic acid from L-phenylalanine has not yet been demonstrated in *E. coli*. Thus, the presented results from whole-cell biotransformation with *RgTAL* as PAL and CYP199A2 as bifunctional C3H/C4H enzyme constitute the first report of *in vivo* caffeic acid production from the substrate L-phenylalanine in *E. coli*.

### III.3 Downstream pathway

Methyltransferases from two plant origins and different classes were studied in two *E. coli* strains. In the genome-integrated *E. coli* BL21(DE3).G213 strain, only an expression of 26 °C led to the complete consumption of the substrate caffeic acid in whole-cell biotransformation, which indicated that the integrated *Pc4CL* was rather active at 26 °C than at 37 °C (chapter II.4.1). Also, caffeoyl alcohol formed as main product at an expression temperature of 26 °C and suggested sufficient activity of the genome-integrated enzymes *Pc4CL*, *ZmCCR* and *ZmCAD*. With the use of all four tested methyltransferases, the resulting production of coniferyl alcohol was similarly low in the G213 strain.

Via the plasmid-based expression of the downstream pathway enzymes in *E. coli* BL21(DE3), the production of coniferyl alcohol could be significantly improved (chapter II.4.2). Thereby, the different methyltransferases were evaluated with two 4CL candidates and varying expression levels for individual proteins were observed depending on the combination of OMT and 4CL enzyme. The two combinations yielding the highest coniferyl alcohol titers were

*At4CL* and *ZmCOMT* as well as *Pc4CL* and *ZmCCoAOMT*. Apparently, the expression and the activity of the four downstream enzymes were most balanced in these two combinations. The subsequent application of higher substrate feeds led to no increases in coniferyl alcohol titers and was thus unsuccessful (chapter II.4.3). It was assumed that *Pc4CL* was more sensitive to substrate and product inhibition in comparison to *At4CL* on the basis of higher accumulating caffeic acid and low caffeoyl alcohol concentrations. Besides, the titers of coniferyl alcohol were also lower with the supplemented 2 mM and 5 mM substrate compared to the previous 1 mM substrate. This suggested that the used OMT enzymes could have been inhibited or inactive as well.

When the two best performing enzyme combinations from the downstream pathway were used in the growing cell assay, the overall production of coniferyl alcohol was comparable with the resting cell assay (chapter II.6.1). Apart from this, caffeic acid was completely consumed, no caffeoyl alcohol formed as intermediate, and the conversion into coniferyl alcohol was achieved faster. Therefore, the conditions in the growing cell assay were suspected to be more beneficial for the activity of the methyltransferases. The most plausible reason for this could be the better availability of the co-factor SAM in cells with active cell metabolism. Once the methyl group from SAM is transferred, S-adenosyl-L-homocysteine (SAH) forms and is regenerated into SAM via four intermediary catalytic steps in the so-called activated methyl cycle [197]. Within this cycle, the intermediate substance L-homocysteine is methylated into L-methionine and this subsequently converted into SAM by the SAM synthetase with the consumption of one molecule of ATP. The SAM co-factor was thus likely the limiting component in the resting cell assay and ATP-dependent regeneration not efficient suggesting a concomitant lack of ATP, too. As a result of that, the lower coniferyl alcohol titers with higher substrate feeds in resting cells could have originated from a shortage of ATP, which is also needed for the activity of the 4CL enzymes. Furthermore, the demethylated co-factor SAH is described to inhibit methyltransferases [198,199]. This could have additionally contributed to the low coniferyl alcohol levels in resting cells.

In literature, a few studies were able to boost their *in vivo* pathways containing methyltransferases by the external addition of L-methionine and the deregulation and overexpression of genes from SAM biosynthesis and its regeneration [127,197]. Kunjapur and co-workers (2016) stated further that methyltransferases are rather influenced by limited SAM availability than the loss of OMT activity [197]. Hence, the conditions in growing cells generally seemed to be beneficial for an improved availability of the SAM co-factor.

It was generally noticed that the mean titers of caffeoyl alcohol and caffeic acid slightly decreased over time despite the determined high standard deviations (chapter II.4.2). At least, for caffeic acid, a consumption effect by *E. coli* cells was observed. Previously, Jang *et al.* (2018) had investigated the production of allomelanin in *E. coli* from caffeic acid and observed melanin formation even without addition of their used pathway enzymes feruloyl-CoA synthetase and enoyl-CoA hydratase [200]. This supported the mentioned hypothesis that *E. coli* can consume or degrade caffeic acid [100]. In addition to that, the decrease of caffeoyl alcohol has been reported before and was explained to come from degradation or volatilization of caffeoyl alcohol [137]. In the performed growing cell assays, the products *p*-coumaryl and coniferyl alcohol decreased significantly after their highest production points mostly around 24 h and were partly undetectable later (chapter II.6.1). This observation was consistent with a few studies, which noted the reduction of phenylpropanoic alcohol concentrations and attributed the decrease to degradation caused by starvation of the cells [173,201]. Besides, coniferyl alcohol can dimerize as a result of spontaneously generated radicals or via the action of endogenous peroxidases in *E. coli* [179,202,203].

Nevertheless, the produced titers of coniferyl alcohol from supplemented 1 mM caffeic acid in growing cells were found to be minorly lower compared to literature [137]. In addition, the corresponding conversion yield using *Pc4CL* + *ZmCCoAOMT* was similar to the results from Chen *et al.* (2017), who applied *At4CL1* in combination with *AtCCoAOMT* (~ 63 % conversion yield at 15 h).

### **III.4 Complete synthetic pathway for monolignol production**

#### **III.4.1 In resting cells**

The production of *p*-coumaryl alcohol as simplest monolignol was achieved in higher titers with *RgTAL* and *At4CL* instead of *Pc4CL* in whole-cell biotransformations using the expression from either two plasmids or a single vector (chapter II.5.1). These results were consistent with reports from literature in which *At4CL* exhibited a significant higher activity compared to *Pc4CL* in growing cells [184]. Besides, *At4CL* was shown to exhibit higher *in vitro* specific activities towards *p*-coumaric and caffeic acid compared to the *Pc4CL2* isoform, which shows 99 % sequence identity to the used *Pc4CL1* isoform in this thesis [139].

With both expression variations, *RgTAL* was found to be the rate-limiting enzyme, as no *p*-coumaric acid was detected as intermediate. An exchange of the *RgTAL* with *FjTAL* or Sesam8 did not lead to any increases in *p*-coumaryl alcohol titers when applying the ePathBrick



vector containing the TAL, 4CL, CCR and CAD genes. In one study, Jansen and co-workers (2014) used *R5TAL* and *Pc4CL* in growing *E. coli* cells and produced 146  $\mu\text{M}$  *p*-coumaryl alcohol after 20 h just by using the L-tyrosine from LB medium [135]. This concentration for *p*-coumaryl alcohol was in the same range as the titer yielded with the TALs and *At4CL* in resting cells in this work (cf. chapter II.5.1).

With respect to the *in vivo* synthesis of coniferyl alcohol, no or only marginal amounts of the intended monolignol were produced in whole-cell biotransformations in *E. coli* BL21(DE3).G213 and as two-cells one-pot approach in *E. coli* BL21(DE3) (chapter II.5.2). In the former G213 strain, inactivity of the genome-integrated *Pc4CL* and of CYP199A2 with Pux/PuR as C3H were identified as main causes. Despite distinct expression of all three components of the C3H, the activity seemed to be impaired as accumulating amounts of *p*-coumaric acid were not further converted. Consequently, the general deficiency of co-factors such as ATP for 4CL and NADH for C3H can be regarded as a probable explanation. Furuya and Kino (2014) reported a significant pH drop in their whole-cell biotransformation with a buffer capacity of 200 mM [103]. Although *E. coli* is able to maintain pH homeostasis of the cytosol, this might not apply for longer time periods [204]. Thus, an acidic pH in whole-cell biotransformation could lead to cell death and cell lysis in the long term [205]. It would also pose stress on the cells possibly resulting in gradual depletion of necessary co-factors including those required for transferred synthetic pathways.

A second reason could be the inhibition of these identified enzymes, which applies to *Pc4CL*, as this enzyme was found to be prone to substrate inhibition (cf. chapter II.4.3). Notably, the intracellular accumulation of the respective CoA esters cannot be excluded and these intermediates could have an inhibiting effect on the 4CL enzymes. But, such an inhibition is considered improbable with regard to the results from the evaluation of the downstream pathway, in which direct conversion from caffeic acid to the alcohols was generally ensured (cf. chapter II.4.2).

In the two-cells one-pot approach, caffeic acid was expectedly generated by the first set of cells (chapter II.5.2). The addition of the second set of cells only led to the accumulation of caffeoyl alcohol as main product with traces of coniferyl alcohol. Thus, the chosen yet uncharacterized *ZmCOMT* was attributed to be responsible for the low monolignol product titers. In literature, COMTs generally accept a rather broad spectrum of substrates with preference for 3,5-hydroxylated over 3,4-dihydroxy phenylpropanoid compounds *in vitro* [104,117–119,206]. The COMT from *Z. mays* shares a protein sequence identity of 89 % with COMT from *Sorghum*

*bicolor*, which had been reported to convert preferably 5-hydroxyconiferaldehyde *in vitro*, but had been sensitive to substrate inhibition at the same time [112]. Due to no available kinetic data for the methyltransferases of *Z. mays*, it can be speculated that the substrate preference may equal that of COMT from *Sorghum bicolor* and that substrate inhibition can be considered as a possible explanation, on the one hand.

On the other hand, the control experiment pointed towards the experimental design as the primary obstacle impeding an efficient pathway performance (cf. Figure A.10, Appendix). It could be that the overall glucose concentration and the parallel addition of glucose along with the second cells may have been insufficient for cell survival in the cell suspension of a higher cell density. As a consequence of possible starvation, the low availability of the co-substrate SAM could have been limiting the activity of *ZmCOMT* [197,207].

The degradation of caffeoyl alcohol presumably by *E. coli*, as mentioned before, represents another outcome of possible starvation [137,173]. Despite this, the observed huge standard deviations for the mean caffeoyl alcohol values could have two alternative reasons. Firstly, decay of caffeoyl alcohol itself was observed in DMSO solution with multiple freeze-thaw cycles. Secondly, the storage time of the samples and the measurement at different HPLC devices at different time points could have had an impact, as well. As a reasonable comparison, it was observed in the control experiment that a relatively short time of storage as well as a simultaneous extraction of the samples and the performance of a single measurement series resulted in low standard deviations (Figure A.10, Appendix).

### **III.4.2 In growing cells**

The assay type was switched from resting to growing cells as a result of occurring limitations from the whole-cell biotransformation with respect to accelerated inactivity of some synthetic pathway enzymes (chapter II.6). With *E. coli* BL21(DE3), the pathway performance from L-tyrosine to coniferyl alcohol was assessed in two experiments: a growing co-culture and a single strain cultivation (chapter II.6.2). Both experiments showed the high accumulation of *p*-coumaryl alcohol and *p*-coumaric acid and only marginal concentrations of any 3-hydroxylated products. This led to the conclusion that the 4CL enzyme's activity was impaired after 24 h and C3H not sufficiently active. In both experimental cases, the bottleneck regarding the C3H inefficiency was likely attributed to its insufficient expression caused by reduced strain growth and/or low protein expression levels. It is noteworthy that CYP199A2 as C3H involves the sufficient expression of its two redox partners Pux and PuR, which makes expression of all three proteins a prerequisite for an efficient flux towards caffeic acid.

The performed co-culture in this thesis only represents a first trial, though this strategy generally holds great optimization potential. In accordance with this, Jones and co-workers (2016) achieved a 970-fold improved titer of flavan-3-ols compared to their previous monoculture after multiple rounds of optimization [184]. In another report, a total of 5.1 mM caffeoyl alcohol could be produced *de novo* by applying an *E. coli* fed-batch co-culture [137]. Furthermore, in the single strain cultivation, the maintenance and expression of the *p*-coumaryl alcohol synthesis genes from the ePathBrick vector was suspected to pose metabolic stress on the cells, as only thin expression bands could be observed in contrast to the previous expression from two plasmids.

In the light of reduced *p*-coumaryl alcohol levels and increased C3H expression, two genome-integrated strains of *E. coli* were investigated in the growing cell assay (chapter II.6.3). The usage of the strain *E. coli* BL21(DE3).G213 improved flux towards coniferyl alcohol significantly, while the *E. coli* BL21(DE3).G279 strain produced exclusively *p*-coumaryl alcohol in high amounts. Consequently, a comparably lower constitutive expression of the downstream pathway enzymes 4CL, CCR and CAD was more beneficial for pathway performance towards coniferyl alcohol. The highest values for coniferyl alcohol resulted from the co-expression of *FjTAL* with *AtCCoAOMT* or *ZmCOMT* in *E. coli* BL21(DE3).G213. In comparison, Sesam8 produced slightly lower values with the same methyltransferases. The highest mean coniferyl alcohol titer of  $854.1 \pm 432.8 \mu\text{M}$  (with *FjTAL* and *AtCCoAOMT*) in growing cells with 1 mM externally added L-tyrosine was equal to the highest *de novo* value from literature being around  $693 \mu\text{M}$  [137]. Chen and co-workers (2017) achieved this titer by applying a co-culture strategy, multiple alternative enzyme candidates including *RgTAL*, the HpaBC complex and *AtCOMT* as well as using an L-tyrosine-overproducing *E. coli* strain.

The most pronounced difference to the production in this thesis, represents the strain engineering regarding introduced perturbations in the central shikimate and the pentose phosphate pathway. Such an engineered strain was able to produce 4 mM L-tyrosine *de novo* from glycerol and glucose in minimal medium [102]. These concentrations of L-tyrosine are comparable to the calculated range for L-tyrosine in TB medium (2-3.9 mM), which could be a reason for the similar production titers of coniferyl alcohol. However, Chen *et al.* (2017) additionally supplemented yeast extract to their minimal medium for better growth and expression of foreign proteins, thereby minorly increasing the total L-tyrosine content [137]. Moreover, *AtCOMT* was shown to lead to a higher production of coniferyl alcohol compared

to *AtCCoAOMT* in this previous study, though 1.4 mM caffeoyl alcohol remained as by-product when using *AtCOMT* [137].

Finally, the potential for increased coniferyl alcohol titers of the synthetic phenylpropanoid pathway was evaluated with a higher L-tyrosine feed (chapter II.6.4). The results indicated that the pathway performance was impaired in the presence of additional 5 mM L-tyrosine. Furthermore, the establishment of a synthetic pathway starting from cinnamic acid was not successful which was also attributed to unknown pathway inhibiting effects of the substrate. Finally, the comparison to literature shows that the chosen enzyme candidates for TAL, C3H and OMT in this thesis are competitive with alternatively used pathways enzymes in terms of microbial coniferyl alcohol synthesis from L-tyrosine.

## IV. Conclusion and outlook

In this thesis, the achieved titer of around  $\sim 850 \mu\text{M}$  coniferyl alcohol performed in growing cells of *E. coli* BL21(DE3).G213 under non-optimized culturing and feeding conditions was found to be comparable to the highest reported *de novo* concentration of this monolignol in *E. coli* [137]. The constructed prototype strain with the complete synthetic pathway holds great optimization potential for application in future fermentations.

One of such optimizations includes the efficient supply of substrate into the pathway. In the case of L-tyrosine, the intracellular concentrations are low and either supplementation to the medium or the application of an overproduction host is required [132]. So, the introduction of the enzymes from the synthetic phenylpropanoid pathway in this thesis into an overproduction strain or the transformation of the used *E. coli* BL21(DE3).G213 into an L-tyrosine overproducer could result in further increased titers of coniferyl alcohol. Additionally, a successive substrate feeding strategy is necessary for future large-scale fermentations, as it was indicated by the results from the increased substrate supplementation in the growing cell assay.

Another important factor for an efficient pathway performance is represented by the concomitant, sufficient and balanced expression of all pathway enzymes. Thereby, genetic engineering of promoters and RBS sequences could yield more ideally designed vectors for fine-tuning of combined protein levels of all enzymes. Regarding a sole plasmid-based expression in *E. coli* BL21(DE3), the used ePathBrick vector containing the *p*-coumaryl alcohol pathway could be subjected to an adjustment of the gene expression. This can be realized by exchanging the strong T7 promoters with weaker ones or via RBS engineering to influence translation efficiency. Furthermore, alternative gene orders and gene configurations of this ePathBrick vector can be examined in a systematic approach as another form of genetic engineering.

The presented work on a synthetic phenylpropanoid pathway supports a deeper understanding of substrate flow through the pathway and key bottlenecks. With respect to regulation mechanisms, an improved insight into inhibition kinetics is necessary. Hence, more detailed investigations for the determination of possible inhibitions by products and intermediates should be performed. Foremost, L-DOPA was suspected to occur as intermediate and could act as an inhibitor of CYP199A2. The verification of this hypothesis necessitates further analyses such as the evaluation of C3H and C4H activities of CYP199A2 in the presence of L-DOPA. Also, a two-cells one-pot approach, in which cells containing CYP199A2 are successively added to cells containing RgTAL, can provide further insights. The P450 enzyme from

*R. palustris* could be additionally investigated for its potential capability to bind L-phenylalanine and L-tyrosine as substrates and checked for inhibition of its C3H/C4H activities with higher L-phenylalanine concentrations. Apart from that, various *in vitro* characterizations are necessary to elucidate the type of inhibition from L-phenylalanine and to draw conclusions about the effect of the additional N-terminal amino acids of NΔ7-CYP199A2.

With regard to the observed differences in *in vivo* activity between both CYP199A2 variants and the varying redox partners, a systematic comparable study is required, which aims at investigating NAD(P)H consumption, coupling efficiency and the rate of uncoupling. Furthermore, the performance of protein NMR as well as modelling and co-crystallization of the P450 enzyme with its redox partners could give a more detailed insight into the binding behavior of Pdx and PdR to the CYP199A2 surface.

As a part of metabolic engineering, different enzyme candidates were evaluated for central individual reaction steps in this work. Still, some chosen enzymes can be exchanged with a possibly superior candidate. The utilization of the soluble and bifunctional ascorbate peroxidase from the monocot *B. distachyon* may provide an alternative and direct C3H route from *p*-coumaric acid to caffeic acid [48]. Moreover, the application of another PAL enzyme, which exhibits specific and high activity towards cinnamic acid, could alleviate inhibition problems in the upstream pathway composed of PAL + C3H/C4H and lead to a more efficient production of caffeic acid starting from L-phenylalanine. The PAL1 enzyme from *A. thaliana* represents a probable alternative, which was shown to produce 1.6-fold higher *de novo* titers of cinnamic acid compared to RgTAL [134]. Also, PAL from *Streptomyces maritimus* proved to exhibit high deamination activity and performed even superior *in vitro* than AtPAL1 [208].

The results from the growing cell assays indicate that both classes of methyltransferases from one origin exhibit comparable overall methylation activities in the complete synthetic pathway and were shown to correlate with the choice of the TAL enzyme. Nevertheless, with respect to the highest values of coniferyl alcohol achieved with AtCCoAOMT in this thesis, further investigations using other methyltransferases such as AtCOMT can be beneficial. In view of the later synthesis of complex lignans, the possible unspecific dimerization of coniferyl alcohol necessitates the fast, subsequent and direct coupling of produced coniferyl alcohol into the basal lignan (+)-pinorexinol. This may be achieved by applying the laccase CgL1 from *Corynebacterium glutamicum* or triggering the *E. coli* endogenous CueO laccase [179,201,209].

The applied co-culture strategy in this work was performed for the demonstration of its feasibility and led to variable outputs. As possible improvement measures, the cultivation conditions could be adjusted and variations tested. In addition, a detailed knowledge about the growth and expression behavior of the individual strains could provide more insights into the efficient design of a growing co-culture.

Furthermore, the sufficient supply of cofactors represents an important aspect for optimization approaches of the presented synthetic pathway in this thesis. Thus, engineering of central metabolic pathways can be beneficial with the aim to increase the intracellular pools of particularly NAD(P)H and SAM. Likewise, the application of glycerol instead of glucose as energy source was found to be more effective for NADH generation and beneficial for product titers in multiple studies [94,103,133,184].

In general, several optimizations concerning the cultivation and expression conditions are essential for an improved pathway performance, too. In this way, certain parameters can be examined such as the inoculation ratio, a fine-tuned temperature profile, the IPTG concentration and the induction OD<sub>600</sub>. Alternative cheaper fermentation media can be advantageous for a scale-up of the coniferyl alcohol production.

At last, the metabolic burden resulting from plasmid maintenance should be reduced by further integrating all pathway enzymes into the genome of an *E. coli* strain. The possible usage of different inducible promoters could give the opportunity of sequential expression of e.g. upstream and downstream pathway [210]. However, the weak constitutive promoters in the *E. coli* BL21(DE3).G213 strain were shown to be beneficial for the overall pathway performance. As another advantage, these promoters do not necessitate the supplementation of inducers, which is crucial for an economically feasible industrial fermentation process.

After all completed optimizations, the vision is the establishment of an efficient and high *de novo* coniferyl alcohol production in a fermentation strain with balanced expression of all pathway enzymes and a self-sufficient co-factors regeneration system. With a further successful scale-up, coniferyl alcohol could be produced in bulk and contribute to a green microbial production of important health-promoting lignans in the future.





## V. Materials and Methods

### V.1 Materials

#### V.1.1 Instruments

The following table summarizes all used equipment (Table 2).

**Table 2: List of instruments.**

Model	Manufacturer
<b>Agarose gel electrophoresis systems</b>	
Mini-Sub <sup>®</sup> Cell GT	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
Wide Mini-Sub <sup>®</sup> Cell GT	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
<b>Autoclaves</b>	
Systec DE-45	Systec GmbH (Linden, Germany)
Systec VX-150	Systec GmbH (Linden, Germany)
<b>Camera</b>	
Nikon D5300	Nikon Corporation (Tokio, Japan)
<b>Centrifuges</b>	
Allegra <sup>™</sup> X-15R	Beckman Coulter (Brea, CA, USA)
Avanti <sup>™</sup> 30	Beckman Coulter (Brea, CA, USA)
Avanti <sup>™</sup> J-30I	Beckman Coulter (Brea, CA, USA)
Avanti <sup>™</sup> J-E	Beckman Coulter (Brea, CA, USA)
Centrifuge 5417 R	Eppendorf AG (Hamburg, Germany)
Centrifuge 5425	Eppendorf AG (Hamburg, Germany)
Heraeus Pico <sup>™</sup> 17	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
IKA <sup>®</sup> mini G	IKA <sup>®</sup> -Werke GmbH & Co. KG (Staufen, Germany)
Microfuge <sup>®</sup> 22R	Beckman Coulter (Brea, CA, USA)
Microfuge <sup>®</sup> R	Beckman Coulter (Brea, CA, USA)
<b>Clean Benches</b>	
HERA <sup>®</sup> safe HS 9	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
HERA <sup>®</sup> safe HS 12	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
HERA <sup>®</sup> safe HS 18	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
<b>Electroporation device</b>	
Multiporator <sup>®</sup>	Eppendorf AG (Hamburg, Germany)
<b>Fridges and freezers</b>	

Bosch Economic	Robert Bosch GmbH (Gerlingen-Schillerhöhe, Germany)
HERAFreeze™ HFU 586	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
HERAFreeze™ HFU 686	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Liebherr Comfort	Liebherr AG (Bulle, Switzerland)
Liebherr Gastrolino	Liebherr AG (Bulle, Switzerland)
Liebherr Premium	Liebherr AG (Bulle, Switzerland)
Miele Electronic	Miele & Cie. KG (Gütersloh, Germany)
Siemens Electronic KG36E04	Siemens AG (Munich, Germany)
<b>HPLC reversed phase columns (length x inner diameter)</b>	
Chromolith® Performance RP-18e (100 x 4.6 mm)	Merck KGaA (Darmstadt, Germany) <sup>1</sup>
ISAspher 100-5 C-18 BDS (250 x 4.0 mm)	ISERA (Düren, Germany) <sup>2</sup>
ISAspher 100-5 C-18 ODS (250 x 4.0 mm)	ISERA (Düren, Germany) <sup>3</sup>
NUCLEODUR® 100-5 C18 ec (150 x 4.6 mm)	Macherey-Nagel (Düren, Germany) <sup>4</sup>
NUCLEOSHELL® EC100/2 RP18, 2,7 µm (100 x 2.0 mm)	Macherey-Nagel (Düren, Germany) <sup>5</sup>
<b>HPLC systems and detectors</b>	
1260 Infinity II LC System	Agilent Technologies Inc. (Santa Clara, CA, USA) <sup>5</sup>
Dionex™ UltiMate™ 3000 Diode Array Detector	Thermo Fisher Scientific Inc. (Waltham, MA, USA) <sup>3</sup>
EXTREMA UV-4075 (LC-4000 series)	Jasco Inc. (Easton, MD, USA) <sup>4</sup>
LC/MS 2020	Shimadzu (Kyoto, Japan) <sup>1</sup>
System Gold 168 Diode Array Detector	Beckman Coulter (Brea, CA, USA) <sup>2</sup>
<b>Ice machine</b>	
Scotsman MF 22	Scotsman International (Milan, Italy)
<b>Incubators and shakers</b>	
Heraeus B5060E	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Heraeus Function-Line B12	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
HERAtherm™ OMS 100	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
Innova™ 4340 Illuminated Refrigerated Incubator Shaker	New Brunswick Scientific (Edison, NJ, USA)
Innova™ 4430 Incubator Shaker	New Brunswick Scientific (Edison, NJ, USA)
INCU-Line®	VWR International (Radnor, PA, USA)
Lab-Therm LT-X	Kuhner (Birsfelden, Switzerland)
<b>Magnetic stirrers and hot plates</b>	
Heidolph MR 3001 K	Heidolph Instruments GmbH & Co. KG (Kelheim, Germany)

IKA® Combimag RET	IKA®-Werke GmbH & Co. KG (Staufen, Germany)
IKA® KMO 2 basic	IKA®-Werke GmbH & Co. KG (Staufen, Germany)
IKA® RCT basic	IKA®-Werke GmbH & Co. KG (Staufen, Germany)
Severin	Severin Elektrogeräte GmbH (Sundern, Germany)
<b>Microwave</b>	
Micromat	AEG AG (Berlin, Germany)
<b>Orbital shakers and rotators</b>	
GFL®-3017	LAUDA Dr. R. Wobser GmbH & Co. KG (Lauda-Königshofen, Germany)
CMV-ROM rotary wheel	Fröbel Labortechnik GmbH (Lindau, Germany)
<b>PCR thermocyclers</b>	
Primus	VWR International (Radnor, PA, USA)
Primus 25 advanced®	VWR International (Radnor, PA, USA)
Primus 96 plus	Eurofins Genomics Germany GmbH (Ebersberg, Germany)
Mastercycler® nexus gradient	Eppendorf AG (Hamburg, Germany)
<b>Photometer</b>	
BioPhotometer®	Eppendorf AG (Hamburg, Germany)
<b>pH meter</b>	
HI 221	Hanna Instruments Deutschland GmbH (Vöhringen, Germany)
HI 2211	Hanna Instruments Deutschland GmbH (Vöhringen, Germany)
<b>Power supplies</b>	
PowerPac™ 200	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
PowerPac™ 300	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
PowerPac™ Basic	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
<b>Scales</b>	
Kern PCB 2500-2	KERN & SOHN GmbH (Balingen, Germany)
Sartorius BL12	Sartorius AG (Göttingen, Germany)
Sartorius BP121S	Sartorius AG (Göttingen, Germany)
Sartorius BL310	Sartorius AG (Göttingen, Germany)
<b>SDS-PAGE and western blot systems</b>	
Mini-PROTEAN® 3 Cell	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
Mini Trans-Blot® Cell	Bio-Rad Laboratories GmbH (Hercules, CA, USA)
<b>Sonicator and ultrasonic baths</b>	

Sonorex Digitec DT 510 H	BANDELIN electronic GmbH & Co. KG (Berlin, Germany) <sup>3</sup>
Sonorex RK 102	BANDELIN electronic GmbH & Co. KG (Berlin, Germany)
UP200S	Hielscher Ultrasonics GmbH (Teltow, Germany)
<b>Thermomixer</b>	
ThermoMixer® C	Eppendorf AG (Hamburg, Germany)
ThermoMixer® comfort	Eppendorf AG (Hamburg, Germany)
<b>Ultrapure water systems</b>	
arium® pro UV	Sartorius AG (Göttingen, Germany)
arium® pro VF	Sartorius AG (Göttingen, Germany) <sup>3</sup>
<b>(UV) transilluminator</b>	
Herolab UVT-28 ME	Herolab GmbH (Wiesloch, Germany)
Herolab V-2	Herolab GmbH (Wiesloch, Germany)
REX ‘Leuchtplatte’	Palmed GmbH (Blaustein, Germany)
<b>Vacuum concentrator</b>	
Concentrator 5301	Eppendorf AG (Hamburg, Germany)
<b>Vacuum pump</b>	
LVS Laboratory Vacuum System	Gardner Denver Inc. (Davidson, NC, USA)
<b>Vortex</b>	
Vortex-Genie 2	Scientific Industries Inc. (Bohemia, NY, USA)
Vortex-Genie 2 Turbomix™	Scientific Industries Inc. (Bohemia, NY, USA)

<sup>1</sup> Institute of Biochemistry II, Heinrich-Heine University Düsseldorf (HHU)

<sup>2</sup> Institute of Applied Microbiology, RWTH Aachen University

<sup>3</sup> Department of Plant Biotechnology, Fraunhofer Institute for Molecular Biology and Applied Ecology (Fraunhofer IME)

<sup>4</sup> Institute of Plant Physiology, RWTH Aachen

<sup>5</sup> Department of Biochemical and Chemical Engineering, Technical University Dortmund (TU Dortmund University)

### V.1.2 Chemicals and consumables

All chemicals and consumables were purchased from A. Hartenstein GmbH (Würzburg, Germany), Beckman-Coulter (Brea, CA, USA), Bio-Budget Technologies GmbH (Krefeld, Germany), Bio-Rad Laboratories GmbH (Hercules, CA, USA), Carl Roth GmbH & Co. KG (Karlsruhe, Germany), Cytiva (Marlborough, MA, USA), Cfm Oskar Tropitzsch GmbH (Marktredwitz, Germany), Diagonal GmbH & Co. KG (Münster, Germany), Dianova GmbH (Hamburg, Germany), Eurofins Genomics Germany GmbH (Ebersberg, Germany), Eppendorf AG (Hamburg, Germany), Hermann Metz GmbH & Co. KG (Quickborn, Germany), ISERA (Düren, Germany), Merck KGaA (Darmstadt, Germany), Macherey-Nagel (Düren, Germany),

Merck Millipore (Burlington, MA, USA), MicroCombiChem GmbH (Wiesbaden, Germany), NerbePlus GmbH & Co. KG (Winsen, Germany), New England Biolabs GmbH (Frankfurt am Main, Germany), Paul Hartmann AG (Heidenheim, Germany), Promega Corporation (Madison, WI, USA), Qiagen (Hilden, Germany), Sarstedt (Nümbrecht, Germany), Schott AG (Mainz, Germany), Sigma-Aldrich (St. Louis, MO, USA), Stratagene (La Jolla, CA, USA), Thermo Fisher Scientific Inc. (Waltham, MA, USA), VWR International (Radnor, PA, USA).

### V.1.3 Software

Several types of software were deployed during implementation of this thesis. As agarose gel documentation software, E.A.S.Y. Win32 (Herolab GmbH, Wiesloch, Germany) was used. The software Clone Manager Professional Suite 8 (Sci Ed Software, Westminster, CO, USA) and Chromas (Technelysium Pty. Ltd., Australia) served for virtual demonstration of cloning procedures, generation of plasmid maps and examination of the sequencing results. Furthermore, PhotoScape v3.7 (MOOII Tech, Korea) was applied as editing tool for photos and ACD/ChemSketch (Advanced Chemistry Development Inc., Toronto, ON, Canada) for drawing of chemical structures. Depending on the utilized HPLC system individual chromatography data system software was made use of: LabSolutions LCMS (Shimadzu; Institute of Biochemistry II, HHU Düsseldorf University), 32 Karat Software (Beckman Coulter; Institute of Applied Microbiology, RWTH Aachen), Dionex™ Chromeleon™ CDS software (Thermo Fisher Scientific Inc.; Fraunhofer IME), ChromNAV2 (Jasco Inc.; Institute of Plant Physiology, RWTH Aachen), OpenLab CDS (Agilent Technologies Inc.; Department of Biochemical and Chemical Engineering, TU Dortmund University). Besides, the software of NetOGlyc 4.0 Server and NetNGlyc 1.0 Server (Department of Bio and Health Informatics, Lyngby, Denmark) was applied for identification of possible glycosylation sites.

### V.1.4 Media, buffers and solutions

Media, buffers, reagents and solutions were prepared as indicated in Table 3. All prepared media were autoclaved for sterility. If sterile filtration was performed, a filter with a pore size of 0.2 µm was used.

**Table 3: List of used media, buffers, solutions and reagents and their respective composition.**

Name	Composition
<b>Media</b>	
LB medium (LB agar)	5 g/L yeast extract 10 g/L tryptone

M9 minimal medium [211]	10 g/L NaCl
	(15 g/L agar-agar)
	Dissolved in distilled water and autoclaved.
	1x M9 minimal salts (pH 7.5)
	0.1 mM CaCl <sub>2</sub>
	1 mM MgSO <sub>4</sub>
	0.05 mM FeCl <sub>3</sub>
	0.1 mM citric acid
	1x trace elements solution
	0.3 % (w/v) glucose
SOC medium	All components individually dissolved in distilled water and autoclaved or sterile-filtered.
	Sterile addition of all components and addition of sterile distilled water up to 1 L before usage.
	5 g/L yeast extract
	20 g/L tryptone
	10 mM NaCl
	25 mM KCl
	10 mM MgCl <sub>2</sub>
	10 mM MgSO <sub>4</sub>
	20 mM glucose
	Dissolved in distilled water and autoclaved.
TB medium	Aliquots stored at -20 °C.
	24 g/L yeast extract
	12 g/L tryptone
	5 g/L glycerol
	Dissolved in 900 mL distilled water and autoclaved.
	Sterile addition of 100 mL of 10x TB salts before usage.
2x/3x TB medium	Preparation of 900 mL TB medium.
	Sterile addition of 100 mL of 20x/30x TB salts before usage.

### Buffers

AP buffer	100 mM Tris-HCl
	100 mM NaCl
	5 mM MgCl <sub>2</sub>
	Dissolved in distilled water.

	pH adjusted to 9.6 with 37 % (v/v) HCl.
DNA sample buffer (10x)	0.1 % (w/v) bromophenol blue 0.1 % (w/v) xylene cyanol 50 % glycerol Dissolved in 1x TAE buffer and stored at -20 °C.
KP <sub>i</sub> buffer	50 mM phosphate buffer (pH 7.5): 1.27 g/L KH <sub>2</sub> PO <sub>4</sub> 7.09 g/L K <sub>2</sub> HPO <sub>4</sub> Dissolved in distilled water and pH adjusted with 1 % (v/v) HCl. Filtered, autoclaved and stored at 4 °C.
PBS (10x)	1.37 M NaCl 27 mM KCl 81 mM Na <sub>2</sub> HPO <sub>4</sub> 15 mM KH <sub>2</sub> PO <sub>4</sub> Dissolved in distilled water.
Reducing protein loading buffer (5x)	62.5 mM Tris-HCl 30 % (v/v) glycerol 4 % (w/v) SDS 10 % (v/v) β-Mercaptoethanol 0.05 % (w/v) bromophenol blue Dissolved in distilled water and stored at -20 °C.
Resuspension buffer	50 mM Tris-HCl (pH 8.0) Dissolved in distilled water and pH adjusted with 1 % HCl. Sterile-filtered and stored at 4 °C.
RLT buffer	25 mM sodium citrate (pH 7.0) 4 M guanidinium isothiocyanate 0.5 % (v/v) sarcosyl 1 % (v/v) β-mercaptoethanol (added before usage) Dissolved in distilled water.
SDS running buffer (10x)	250 mM Tris 2 M glycine 1 % (w/v) SDS Dissolved in distilled water.
TAE buffer (50x)	2 M Tris

	5.72 % (v/v) acetic acid
	50 mM EDTA (pH 8.0)
	Dissolved in distilled water.
Tank blot buffer (10x)	250 mM Tris
	1,92 M glycine
	Dissolved in distilled water.
	For 1x addition of 20 % (v/v) methanol and stored at 4 °C.
TB buffer	10 mM PIPES
	55 mM MnCl <sub>2</sub>
	15 mM CaCl
	250 mM KCl
	Mixed together from individual sterile stock solutions with sterile distilled water.
	Stored at 4 °C when used for co-transformation.
TB salts (10x)	0.17 M KH <sub>2</sub> PO <sub>4</sub>
	0.72 M K <sub>2</sub> HPO <sub>4</sub>
	Dissolved in distilled water, filtered and autoclaved.
TB salts (20x)	0.34 M KH <sub>2</sub> PO <sub>4</sub>
	1.44 M K <sub>2</sub> HPO <sub>4</sub>
	Dissolved in distilled water, filtered and autoclaved.
TB salts (30x)	0.51 M KH <sub>2</sub> PO <sub>4</sub>
	2.16 M K <sub>2</sub> HPO <sub>4</sub>
	Dissolved in distilled water, filtered and autoclaved.
TE lysis buffer	10 mM Tris-HCl
	1 mM EDTA
	Dissolved in distilled water and pH adjusted to 8.0.

### Solutions and reagents

Acrylamide (30 %)	Rotiphorese® Gel 30 (Carl Roth GmbH & Co. KG)
	30 % acrylamide, 0.8 % bisacrylamide (ratio 37.5:1)
	Stored at 4 °C.
Agarose (1.2 %)	12 g/L agarose
	Dissolved in 1x TAE buffer and autoclaved or heated up in a microwave.
	Ethidium bromide added at a concentration of 0.3 µg/mL.
	Stored at 65 °C.



Ampicillin stock solution	100 mg/mL D- $\alpha$ -Aminobenzylpenicillin sodium salt Dissolved in 50 % (v/v) ethanol and sterile-filtered. Stored at -20 °C.
APS	20 % (w/v) ammonium peroxodisulfate Dissolved in distilled water and filtered.
Blocking solution	4 % (w/v) milk powder Dissolved in 1x PBS.
Chloramphenicol stock solution	50 mg/mL chloramphenicol Dissolved in ethanol and sterile-filtered. Stored at -20 °C.
Coomassie destaining solution	5 % (v/v) methanol 7.5 % (v/v) acetic acid Mixed with distilled water.
Coomassie staining solution	2.5 g/L Coomassie Brilliant Blue G-250 50 % (v/v) methanol 10 % (v/v) acetic acid Mixed with distilled water and stored in the dark.
Glucose stock solution	40 % (w/v) glucose Dissolved in distilled water and sterile-filtered or autoclaved.
Glycerol stock solution	70 % (v/v) glycerol Mixed with distilled water and autoclaved.
IPTG stock solution	1 M isopropyl- $\beta$ -D-thiogalactopyranoside Dissolved in distilled water and sterile-filtered. Stored at -20 °C.
Kanamycin stock solution	100 mg/mL kanamycin sulfate (kanamycin A) Dissolved in distilled water and sterile-filtered. Stored at -20 °C.
NBT/BCIP solution	3.33 % (w/v) nitro blue tetrazolium chloride 1.65 % (w/v) 5-bromo-4-chloro-3-indolyl phosphate <i>p</i> -toluidine salt Dissolved in dimethylformamide. Stored at -20 °C or 4 °C in the dark.
SDS (10 %)	100 g/L sodium dodecyl sulfate Dissolved in distilled water.
SDS-PAGE resolving gel buffer	1 M Tris-HCl buffer

	Dissolved in distilled water.
	pH adjusted to 8.8 with 37 % (v/v) HCl.
SDS-PAGE stacking gel buffer	1 M Tris-HCl buffer
	Dissolved in distilled water.
	pH adjusted tp 6.8 with 37 % (v/v) HCl.
Streptomycin stock solution	50 mg/mL streptomycin sulfate
	Dissolved in distilled water and sterile-filtered.
	Stored at -20 °C.

### V.1.5 Antibodies

The following table lists the antibodies applied in western blot analysis (Table 4).

**Table 4: List of antibodies.**

Antibody	Type	Isolated from	Antigen	Characteristics	Manufacturer
<b>Primary antibodies</b>					
$\alpha$ -his	monoclonal	mouse	His-tag	0.2 mg/mL	Qiagen (Hilden, Germany)
S-peptide Epitope Tag Antibody	monoclonal	mouse	S-tag	1 mg/mL	Thermo Fisher Scientific Inc. (Waltham, MA, USA)
<b>Secondary antibody</b>					
GAM <sub>FC</sub> <sup>AP</sup>	polyclonal	goat	mouse IgG <sub>FC</sub>	0.8 mg/mL, conjugated with AP	Dianova GmbH (Hamburg, Germany)

### V.1.6 DNA Oligonucleotides (Primers)

The synthetic DNA oligonucleotides were ordered from Eurofins Genomics Germany GmbH (Ebersberg, Germany) and from IDT Integrated DNA Technologies (Coralville, IA, USA). The working concentration of all primer stock solutions was 10  $\mu$ M (Table 5).

**Table 5: List of synthetic DNA oligonucleotides.**

Primer name	Sequence (5'→3')	Application
9_NdeI_HpaC_fw	GGAGATATACATATGCAATTAGATG	PCR
10_HpaC_XhoI_rv	CTTTACCAGACTCGAGTTAAATCGCAGCTTCC	cPCR, PCR, RT
29_Mut2_ZmCCR_SalI_557. bp_fw	GTAGTGGACGAGTCGTG	cPCR

30_Mut2_ZmCCR_SalI_557. bp_rv	CTCGTCCACTACCACATC	cPCR
31_NdeI_CYP199A2(-7aa)_fw	GATATACATATGCCGGTCACGACTCCGAG	PCR
32_NdeI_AtCOMT_fw	GATATACATATGGGGAGCACCGCAGAAAC	cPCR
34_NdeI_AtCCoAOMT_fw	GGAGATATACATATGGCAAAAGACGAAGC	cPCR
35_AtCCoAOMT_XhoI_rv	GACTCGAGTTAGTACAGGCGACGACAG	RT, cPCR
36_NdeI_ZmCOMT_fw	GATATACATATGGGCAGTACGGCCGGTGATG	cPCR, sequencing
38_NdeI_ZmCCoAOMT_fw	GATATACATATGGCGACAACAGCCACAG	cPCR, sequencing
40_EcoRV_his <sub>6</sub> -camA_fw	GCAGATCTCAATTGGATATCGCATCATCATCATCA TC	cPCR, PCR
41_his <sub>6</sub> -camA_AatII_rv	GAGGGTACCGACGTCTCAGGCACTACTCAG	cPCR, PCR
42_NdeI_his <sub>6</sub> -FpR_fw	GAAGGAGATATACATATGCATCATCATCATC	cPCR, PCR
43_his <sub>6</sub> -FpR_XhoI_rv	CTTTACCAGACTCGAGTTACCAGTAATGCTCCGCT GTC	cPCR, PCR
ACYCDuet UP1	GGATCTCGACGCTCTCCCT	cPCR, sequencing
At4CL_end_fw	GTGAAGCAATTCTGTGTCGAAACAG	cPCR
E3_BamHI_his <sub>6</sub> -Pux_fw	CAGGATCCGCCTAGCATTACCTTTATCCTGC	cPCR, PCR
E4_Pux_NotI_rv	CATTATGCGGCCGCTTATACTTGGCGGTC	PCR, RT, cPCR
E8_CYP199A2_XhoI_rv	CTTTACCAGACTCGAGTTACGCCGGGGTCAGTTG	cPCR, PCR, RT
E9_NdeI_his <sub>6</sub> -PuR_fw2	GAAGGAGATATACATATGCATCACCATCATCACC ACAATGATACTGTGCTGATTG	cPCR, PCR
E10_PuR_XhoI_rv2	CTTTACCAGACTCGAGTCATGCTGCTGCTTTC	cPCR, PCR, RT
FjTAL_KpnI_rv	TGGACTCGAGGGTACCTTAATTGTTAATCAGGTG GTCTTTTACTTTC	RT, PCR
HaTAL_KpnI_rv	AGACTCGAGGGTACCTTAGCGAAACAGAATAATA CTACGC	RT, PCR
HpaB_end_fw	GCAGTGTCTGCGCCAGGCACAAAACCTC	cPCR, sequencing
HpaB_internal_fw	GTTCGAGCAGAACGCCCCGTAACCTGG	cPCR
M1_pETM6_fw	CTGAAGATGCTGGCATAACTCGAGTCTGGTAAAG	FastCloning
M1_pETM6_rv	GTCGGGCGAGGCGCCATATGTATATCTCCTTC	FastCloning
M1_RgTAL_fw	GAAGGAGATATACATATGGCGCCTCGCCCCGAC	cPCR, FastCloning
M1_RgTAL_rv	GGTTTCTTTACCAGACTCGAGTTATGCCAGCATCT TC	FastCloning, RT, cPCR

M2_HpaB_fw	GAAGGAGATATACATATGAAACCAGAAGATTTC	cPCR, FastCloning
M2_HpaB_rv	CTTTACCAGACTCGAGTTATTTTCAGCAGCTTATCC AG	FastCloning
M2_pETM6_fw	GATAAGCTGCTGAAATAACTCGAGTCTGGTAAAG AAAC	FastCloning
M2_pETM6_rv	GGAAATCTTCTGGTTTCATATGTATATCTCCTTCTT AAAG	FastCloning
M9_At4CL_Sall_515.bp_fw	GCTCGTTACGTGGACAAAATCAAACCACTTC	cPCR, mutagenesis
M9_At4CL_Sall_515.bp_rv	GTCCACGTAACGAGCTTCGGTGATTATGAG	mutagenesis, sequencing
M9_At4CL_Sall_629.bp_fw	CAGTCCACAACCGAGGCATCAGAAGTCATC	mutagenesis
M9_At4CL_Sall_629.bp_rv	GGTTGTGGACTGAGTCAACTCGGTGAAGC	mutagenesis
M9_At4CL_Sall_776.bp_fw	CAGCAAGTGGACGGCGAGAACCCGAATC	cPCR, mutagenesis
M9_At4CL_Sall_776.bp_rv	CCGTCCACTTGCTGAGCAACGCTCGTGACTAG	cPCR, mutagenesis
MCS1_fw	ATGCGTCCGGCGTAGA	cPCR, sequencing
MCS1_rv	GATTATGCGGCCGTGTACAA	cPCR, sequencing
MCS2_fw	TTGTACACGGCCGCATAATC	cPCR, sequencing
MCS2_rv	GCTAGTTATTGCTCAGCGG	cPCR, sequencing
NcoI_FjTAL_fw	AACCATGGGCAACACCATCAACGAATATCTGAG	cPCR, sequencing
NcoI_HaTAL_fw	AACCATGGGCAGCACACCCTGATTCTGACCGGT G	cPCR
NcoI_Sesam8_fw	AACCATGGGCACCCAGGTTGTTGAACGTCAGGCA G	cPCR, sequencing
NdeI_RsTAL_fw	ATACATATGAGCGCGCAGGATCCGGCTCTGGC	cPCR
Pc4CL_end_fw	GCAATTCCGAAATCACCATCTGGAAAG	cPCR
Pc4CL_internal_fw	CCTTATTCATCGGGTACTACAGGAC	cPCR
Pc4CL_fw	GTCAAGATCTCATGGGAGATTGTGTAGCACC	cPCR
Pc4CL_rv	TAGCACTCGAGTTTGGGAAGATCACCGGATGC	cPCR, RT
pETM6_fw	ATTGTGAGCGGATAACAATTC	cPCR, sequencing
pETM6_fw2	GGTGATGTCGGCGATATAGGCGCCAGCAACC	cPCR, sequencing
pETM6_rv	GCTACAGGGCGCGTCCCATTGCGCCAATCC	cPCR, sequencing

RgTAL_end_fw	GCATTTGTGCGTGAGGAAGTGGGTGTC	cPCR
RgTAL_internal_fw	CAAAGTCTGGAGGGCTCGCGCTTCG	cPCR
RsTAL_EcoRI_rv	AAGAGCTCGAATTCAACTGGACTCTGTTGCAGCA GATGAGTAG	RT, cPCR
RsTAL_internal_fw	GCACTGTCTGATCTGCGTCCTC	cPCR
RsTAL_KpnI_rv	AGACTCGAGGGTACCTTAAACTGGACTCTGTTGC AGC	RT, cPCR
Sesam8_KpnI_rv	AGACTCGAGGGTACCTTAGCCAAAATCTTTACCA TCTG	RT, cPCR
ZmCAD_fw	ATATACATATGGGTAGTCTGGCGTCAG	cPCR, sequencing
ZmCAD_XhoI_TAA_rv	ACCAGACTCGAGTTAATTCGAGGCTGCATCGGCG GCAG	cPCR, RT
ZmCCR_fw	AGGATCCGAATTCGATGACGGTGGTTGATGCG	cPCR
ZmCCR_end_fw	GCACCAACGGCGGAAATGCAACAAGG	Sequencing

### V.1.7 Plasmids

Novagen® Duet vectors from Merck KGaA (Darmstadt, Germany) were pACYCDuet™-1, pCDFDuet™-1, pETDuet™-1 and pRSFDuet™-1 (Figure A.20, Appendix). The so-called ePathBrick vectors namely pCDM4, pETM6 and pETM6\_RgTAL\_HpaB\_HpaC\_m were purchased from Addgene (Watertown, MA, USA) (Figure A.21, Appendix). Synthetic genes were ordered in customized vector backbones from Eurofins Genomics Germany (Ebersberg, Germany) and GENEWIZ (South Plainfield, NJ, USA). The following table summarizes all used plasmids (Table 6). Notably, some plasmids applied in this thesis were cloned by Dr. Christina Dickmeis (Institute for Molecular Biotechnology, RWTH Aachen University) or a gift from the group of Prof. Dr. Vlada Urlacher (Biochemistry II, HHU Düsseldorf). The codon-optimized gene sequences of CYP199A2, Pux, PuR and the four OMTs for the expression in *E. coli* are summarized in Table 15 (Appendix).

**Table 6: List of all plasmids.** co- codon-optimized for the expression in *E. coli*, WT- wild type, I/II- cloned in MCS I or II, m- monocistronic, pso- pseudo-operon, op- operon, NΔ7- N-terminal deletion of 7 amino acids. GenBank accession numbers are indicated in square brackets.

Plasmid	Details	Source
<b>TAL and PAL enzymes</b>		
pETM6_FjTAL	TAL enzyme from <i>Flavobacterium johnsoniae</i> (co) [KR095306.1]	Dr. Christina Dickmeis (Institute for Molecular Biotechnology, RWTH Aachen University)

pETM6_ <i>Ha</i> TAL	TAL1 enzyme from <i>Herpetosiphon aurantiacus</i> (co) [KR095308.1]	Dr. Christina Dickmeis <sup>1</sup>
pETDuet_ <i>his</i> <sub>6</sub> - <i>Zm</i> PAL (I)	PAL1 enzyme from <i>Zea mays</i> [BT054938.1]	Dr. Christina Dickmeis <sup>1</sup>
pETM6_ <i>Rg</i> TAL	TAL enzyme from <i>Rhodotorula glutinis</i> (co) [AGZ04575.1]	This work
pETM6_ <i>Rs</i> TAL	TAL enzyme from <i>Rhodobacter sphaeroides</i> (co) [ABA81174.1]	Dr. Christina Dickmeis <sup>1</sup>
pETM6_ <i>Sesam</i> 8	TAL enzyme from <i>Saccharothrix espanaensis</i> (co) [KR095297.1]	Dr. Christina Dickmeis <sup>1</sup>
<b>4CL enzymes</b>		
pCDM4_ <i>At</i> 4CL	4CL1 from <i>Arabidopsis thaliana</i> , without internal SalI restriction sites	This work
pETM6_ <i>At</i> 4CL	4CL1 from <i>Arabidopsis thaliana</i> , without internal SalI restriction sites	This work
pETM6_ <i>At</i> 4CL (WT)	4CL1 from <i>Arabidopsis thaliana</i> [OAP14948.1]	Dr. Christina Dickmeis <sup>1</sup>
pETM6_ <i>Pc</i> 4CL	4CL1 from <i>Petroselinum crispum</i> [X13324.1]	Dr. Christina Dickmeis <sup>1</sup>
<b>TAL and 4CL enzymes</b>		
pETM6_ <i>Rg</i> TAL_ <i>At</i> 4CL_m	TAL enzyme from <i>Rhodotorula glutinis</i> (co) and 4CL from <i>Arabidopsis thaliana</i> , cloned in a monocistronic gene configuration	This work
pETM6_ <i>Rg</i> TAL_ <i>Pc</i> 4CL_m	TAL enzyme from <i>Rhodotorula glutinis</i> (co) and 4CL from <i>Petroselinum crispum</i> , cloned in a monocistronic gene configuration	This work
<b>CCR and CAD enzyme</b>		
pRSFDuet_ <i>his</i> <sub>6</sub> - <i>Zm</i> CCR_ <i>Zm</i> CAD-S-Tag	CCR1 enzyme from <i>Zea mays</i> (co) [Y13734.1] and CAD enzyme from <i>Zea mays</i> (co) [Y13733.1]	Dr. Christina Dickmeis <sup>1</sup>
pETM6_ <i>Zm</i> CCR_ <i>Zm</i> CAD_m	CCR1 enzyme from <i>Zea mays</i> (co) [Y13734.1] and CAD enzyme from <i>Zea mays</i> (co) [Y13733.1], cloned in a monocistronic gene configuration	Dr. Christina Dickmeis <sup>1</sup>
<b>C3H enzymes and redox partners</b>		
pCDFDuet_ <i>CYP</i> 199A2 (II)	Cytochrome P450 enzyme from <i>Rhodopseudomonas palustris</i> (co)	This work
pCDFDuet_ <i>NΔ</i> 7- <i>CYP</i> 199A2 (II)	Cytochrome P450 enzyme from <i>Rhodopseudomonas palustris</i> (co), without first seven amino acids at the N-terminus	This work
pCDM4_ <i>HpaB</i> _HpaC_op	C3H enzyme complex from <i>Escherichia coli</i> , cloned as an operon	This work

pET11a_his <sub>6</sub> -FpR	Flavodoxin reductase from <i>Escherichia coli</i> [AKF70073.1]	Gift from Prof. Dr. Vlada Urlacher (Biochemistry II, HHU Düsseldorf)
pET16b_his <sub>10</sub> -YkuN	Flavodoxin from <i>Bacillus subtilis</i> [CAA10877.1]	Gift from Prof. Dr. Vlada Urlacher <sup>2</sup>
pET28a(+)_his <sub>6</sub> -PdR	Putidaredoxin reductase from <i>Pseudomonas putida</i> [BAA00413.1]	Gift from Prof. Dr. Vlada Urlacher <sup>2</sup>
pET28a(+)_his <sub>6</sub> -Pdx	Putidaredoxin from <i>Pseudomonas putida</i> [BAA00414.1]	Gift from Prof. Dr. Vlada Urlacher <sup>2</sup>
pETM6_HpaB_HpaC_m	C3H enzyme complex from <i>Escherichia coli</i> , cloned in a monocistronic gene configuration	This work
pETM6_HpaB_HpaC_pso	C3H enzyme complex from <i>Escherichia coli</i> , cloned in a pseudo-operon gene configuration	This work
pETM6_HpaB_HpaC_op	C3H enzyme complex from <i>Escherichia coli</i> , cloned as an operon	This work
pETM6_RgTAL_HpaB_HpaC_m	TAL enzyme from <i>Rhodotorula glutinis</i> (co) and C3H enzyme complex from <i>Escherichia coli</i> , all genes cloned in a monocistronic gene configuration	Addgene (Watertown, MA, USA) [133]
pEX-A258-CYP199A2	Cytochrome P450 enzyme from <i>Rhodopseudomonas palustris</i> (co) [BX572599.1]	Eurofins Genomics Germany GmbH (Ebersberg, Germany)
pRSFDuet_his <sub>6</sub> -Pdx_his <sub>6</sub> -PdR	Putidaredoxin and putidaredoxin reductase from <i>Pseudomonas putida</i>	This work
pRSFDuet_his <sub>6</sub> -Pux_his <sub>6</sub> -PuR	Palustrisredoxin (co) and palustrisredoxin reductase (co) from <i>Rhodopseudomonas palustris</i>	This work
pRSFDuet_his <sub>10</sub> -YkuN_his <sub>6</sub> -PdR	Flavodoxin from <i>Bacillus subtilis</i> and putidaredoxin reductase from <i>Pseudomonas putida</i>	This work
pRSFDuet_his <sub>10</sub> -YkuN_his <sub>6</sub> -PuR	Flavodoxin from <i>Bacillus subtilis</i> and palustrisredoxin reductase (co) from <i>Rhodopseudomonas palustris</i>	This work
pRSFDuet_his <sub>10</sub> -YkuN_his <sub>6</sub> -FpR	Flavodoxin from <i>Bacillus subtilis</i> and flavodoxin reductase from <i>Escherichia coli</i>	This work
pUC57_PuR	Palustrisredoxin reductase from <i>Rhodopseudomonas palustris</i> (co) [BX572605.1]	GENEWIZ Inc. (South Plainfield, NJ, USA)
pUC57_Pux	Palustrisredoxin from <i>Rhodopseudomonas palustris</i> (co) [BX572599.1]	GENEWIZ Inc. (South Plainfield, NJ, USA)
<b>OMT enzymes</b>		
pCDFDuet_AtCCoAOMT (I)	Caffeoyl-CoA <i>O</i> -methyltransferase from <i>Arabidopsis thaliana</i> (co)	This work
pCDFDuet_AtCOMT (I)	Caffeic acid <i>O</i> -methyltransferase from <i>Arabidopsis thaliana</i> (co)	This work

pCDFDuet_ZmCCoAOMT (I)	Caffeoyl-CoA <i>O</i> -methyltransferase from <i>Zea mays</i> (co)	This work
pCDFDuet_ZmCOMT (I)	Caffeic acid <i>O</i> -methyltransferase from <i>Zea mays</i> (co)	This work
pRSFDuet_ZmCCoAOMT (I)	Caffeoyl-CoA <i>O</i> -methyltransferase from <i>Zea mays</i> (co)	This work
pRSFDuet_ZmCOMT (I)	Caffeic acid <i>O</i> -methyltransferase from <i>Zea mays</i> (co)	This work
pUC57_AtCCoAOMT	Caffeoyl-CoA <i>O</i> -methyltransferase from <i>Arabidopsis thaliana</i> (co) [AF360317.1]	GENEWIZ Inc. (South Plainfield, NJ, USA)
pUC57_AtCOMT	Caffeic acid <i>O</i> -methyltransferase from <i>Arabidopsis thaliana</i> (co) [AY062837.1]	GENEWIZ Inc. (South Plainfield, NJ, USA)
pUC57_ZmCCoAOMT	Caffeoyl-CoA <i>O</i> -methyltransferase from <i>Zea mays</i> (co) [AJ242981.1]	GENEWIZ Inc. (South Plainfield, NJ, USA)
pUC57_ZmCOMT	Caffeic acid <i>O</i> -methyltransferase from <i>Zea mays</i> (co) [M73235.1]	GENEWIZ Inc. (South Plainfield, NJ, USA)
<b>C3H and OMT enzymes</b>		
pCDFDuet_AtCCoAOMT_CYP199A2	Caffeoyl-CoA methyltransferase from <i>Arabidopsis thaliana</i> (co) and cytochrome P450 enzyme from <i>Rhodopseudomonas palustris</i> (co)	This work
pCDFDuet_ZmCCoAOMT_CYP199A2	Caffeoyl-CoA methyltransferase from <i>Zea mays</i> (co) and cytochrome P450 enzyme from <i>Rhodopseudomonas palustris</i> (co)	This work
pCDFDuet_ZmCCoAOMT_NΔ7-CYP199A2	Caffeoyl-CoA methyltransferase from <i>Zea mays</i> (co) and cytochrome P450 enzyme from <i>Rhodopseudomonas palustris</i> (co), without first seven amino acids at N-terminus	This work
pCDFDuet_ZmCOMT_CYP199A2	Caffeic acid methyltransferase from <i>Zea mays</i> (co) and cytochrome P450 enzyme from <i>Rhodopseudomonas palustris</i> (co)	This work
pCDFDuet_ZmCOMT_NΔ7-CYP199A2	Caffeic acid methyltransferase from <i>Zea mays</i> (co) and cytochrome P450 enzyme from <i>Rhodopseudomonas palustris</i> (co), without first seven amino acids at N-terminus	This work
<b>Multi-enzyme pathway vectors</b>		
pETM6_FjTAL_At4CL_ZmCCR_ZmCAD_m	All genes cloned in a monocistronic gene configuration	This work
pETM6_RgTAL_At4CL_ZmCCR_ZmCAD_m	All genes cloned in a monocistronic gene configuration	Dr. Christina Dickmeis <sup>1</sup>
pETM6_RgTAL_Pc4CL_ZmCCR_ZmCAD_m	All genes cloned in a monocistronic gene configuration	Dr. Christina Dickmeis <sup>1</sup>
pETM6_Sesam8_At4CL_ZmCCR_ZmCAD_m	All genes cloned in a monocistronic gene configuration	This work



<sup>1</sup> Institute for Molecular Biotechnology, RWTH Aachen University<sup>2</sup> Biochemistry II, HHU Düsseldorf

## V.1.8 Bacterial strains

The following table summarizes all used *Escherichia coli* strains with their respective genotype, and application in this thesis (Table 7). The strains *E. coli* BL21(DE3).G213 and *E. coli* BL21(DE3).G279 were gifts from Phytowelt GreenTechnologies GmbH (Nettetal, Germany) and contained genes of the phenylpropanoid pathway (TAL, 4CL, CCR, CAD), which had been integrated into the genome and put under the control of weak constitutive promoters.

**Table 7: List of all applied strains.**

Strain	Genotype/Details	Application
<i>E. coli</i> BL21(DE3)	F- <i>ompT hsdS<sub>B</sub> (r<sub>B</sub>-m<sub>B</sub><sup>-</sup>) gal dcm</i> (DE3)	Expression Resting and growing cell assay Toxicity test
<i>E. coli</i> BL21(DE3).G213	F- <i>ompT hsdS<sub>B</sub> (r<sub>B</sub>-m<sub>B</sub><sup>-</sup>) gal dcm</i> (DE3) additional genome-integrated genes: <i>RsTAL-Pc4CL</i> , <i>ZmCCR-ZmCAD</i>	Expression Resting and growing cell assay
<i>E. coli</i> BL21(DE3).G279	F- <i>ompT hsdS<sub>B</sub> (r<sub>B</sub>-m<sub>B</sub><sup>-</sup>) gal dcm</i> (DE3) additional genome-integrated genes: <i>RsTAL-Pc4CL</i> , <i>ZmCCR-ZmCAD</i> , <i>At4CL-ZmCCR-ZmCAD</i>	Expression Growing cell assay
<i>E. coli</i> DH5α	<i>fhuA2 Δ(argF-lacZ) U169 phoA glnV44 Φ80</i> <i>lacZΔM15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17</i>	Cloning Retransformation
<i>E. coli</i> TOP10	F- <i>mcrA Δ( mrr-hsdRMS-mcrBC) Φ80 lacZΔM15</i> <i>ΔlacX74 recA1 araD139 Δ(araleu)7697 galU galK</i> <i>rpsL (StrR) endA1 nupG</i>	Cloning (FastCloning, Mutagenesis)

## V.2 Methods

### V.2.1 Molecular Biology Methods

#### Polymerase chain reaction (PCR)

Polymerase chain reaction was used for the *in vitro* amplification of DNA fragments. Thereby, a reaction mixture of MQ-H<sub>2</sub>O, buffer, dNTPs, DNA polymerase, and forward and reverse primer was prepared with a DNA template and put into a thermocycler (Table 8). The thermostable DNA polymerases were either GoTaq<sup>®</sup> DNA Polymerase (5 U/μL) with 5x green GoTaq<sup>®</sup> reaction buffer or *Pfu* DNA polymerase (3 U/μL) with 10x *Pfu* reaction buffer with MgSO<sub>4</sub> (Promega Corporation, Madison, WI, USA). The program in the thermocycler started with an initial denaturation step at a temperature of 95 °C for 3 min followed by 30

amplification cycles. One amplification cycle consisted of heating at 95 °C for 30 sec, primer annealing at 45-55 °C for 30 sec and elongation at 72 °C for 1 min/kb. At last, a final elongation step at 72 °C for 2 min/kb was performed and the reaction mixture was stored at 10 °C in the thermocycler. Afterwards, the PCR mixture was applied to an agarose gel electrophoresis.

**Table 8: Composition of a single PCR reaction mixture.**

Component	PCR reaction mixture
MQ-H <sub>2</sub> O	35.25 / 40.25 µL
DNA polymerase reaction buffer (5x) /(10x)	10 / 5 µL
dNTPs (10 mM)	1.25 µL
Forward primer (10 µM)	1 µL
Reverse primer (10 µM)	1 µL
DNA template	1 µL
DNA polymerase	0.5 µL
Total volume	50 µL

### Restriction and ligation

For the construction of the majority of vectors, classical cloning via restriction endonucleases followed by ligation was performed in accordance with the standard guidelines from Sambrook and Russell [212]. The 10x CutSmart® buffer and the various restriction enzymes were purchased from New England Biolabs GmbH (Frankfurt am Main, Germany). For the construction of most Duet and all ePathBrick vectors, the reaction mixture of restriction 2 was used, while restriction 1 was utilized whenever the DNA concentration of 1 µg could not be reached with restriction 2 (Table 9). Every time when the DNA template was a PCR amplification product, restriction 1 with a standard amount of 30 µL of PCR product was chosen.

**Table 9: Composition of both reaction mixtures for restriction.**

Component	Restriction 1	Restriction 2
MQ-H <sub>2</sub> O	0-13 µL	0-11 µL
DNA	30-43 µL	5-16 µL (1-2 µg)
CutSmart® buffer (10x)	5 µL	2 µL
Enzyme 1 (5-20 U/µL)	1 µL	0.5-1 µL
Enzyme 2 (5-20 U/µL)	1 µL	0.5-1 µL
Total volume	50 µL	20 µL

The reaction mixtures were incubated for 1.5-3 h at 37 °C. Afterwards, 1 µL of calf intestinal phosphatase (CIP or Quick CIP, New England Biolabs) was added to the restriction mixtures of the vectors for dephosphorylation of the 5'-ends and a further incubation step at 37 °C for 30 min followed. The restriction reactions were applied to agarose gel electrophoresis and gel extraction for purification of the cut DNA fragments and vectors (chapter V.2.1). The cut and purified DNA was used in a ligation reaction consisting of 7 µL MQ-H<sub>2</sub>O, 4 µL of vector DNA, 6 µL of insert DNA, 2 µL 10x T4 DNA ligase reaction buffer and 1 µL T4 DNA ligase [213]. Depending on the band intensities of vector and insert in agarose gel electrophoresis, different amounts of vector and insert were used for individual reaction mixtures to yield at least a 1:1 ratio of amount of insert to amount of vector (Table 10). The ligation reaction was incubated at 16 °C overnight or at 4 °C for three to four days.

**Table 10: Composition of a single ligation reaction.**

Component	Ligation	
MQ-H <sub>2</sub> O	7 µL	(0-7 µL)
Vector DNA	4 µL	(4-8 µL)
Insert DNA	6 µL	(6-13 µL)
T4 DNA ligase reaction buffer (10x)	2 µL	
T4 DNA ligase (3 U/µL)	1 µL	
Total volume	20 µL	

### Precipitation of ligation reaction and electrotransformation

An amount of 1-2 µL glycogen and 65 µL of 100 % ethanol were added to the ligation reaction mixture for precipitation of DNA. The mixture was centrifuged at 14,000 rpm (18,000 xg) and 4 °C for 2-4 h. Then, the supernatant was discarded, the DNA pellet resuspended in 250 µL of 70 % (v/v) ethanol and centrifuged at 14,000 rpm (18,000 xg) and 4 °C for 40 min again. After removal of the supernatant, the DNA pellet was dried at 37 °C and dissolved in 5 µL MQ-H<sub>2</sub>O for electrotransformation.

### FastCloning and *DpnI* digestion

For the design of some constructs, FastCloning was used as cloning method [214]. In this method, individual PCR reactions for insert and vector DNA are performed by using primers with overlapping sequences and *Pfu* DNA polymerase (chapter V.2.1, Table 5). After both PCR reactions were run with 18 amplification cycles, 4 µL of amplified vector DNA was mixed with 4 µL of amplified insert DNA and 0.5 µL *DpnI*. This reaction mixture was then incubated for

1 h at 37 °C followed by an inactivation step at 80 °C for 10 min. Afterwards, 2 µL of the reaction mixture was directly used for heat shock transformation into *E. coli* DH5α or TOP10 cells (chapter V.2.2). The enzyme *DpnI* digests parental *dam*-methylated DNA strands, thus eliminating the DNA template from the PCR reactions. The 3'-exonuclease activity of the high fidelity DNA polymerase probably creates sticky ends at the position of the overlapping sequences leading to nicks that are repaired after transformation in *E. coli* [214].

### Site-directed mutagenesis (QuikChange™ method)

Site-directed mutagenesis is performed with a pair of mutagenic primers and *Pfu* DNA polymerase (Table 5). A PCR reaction was prepared accordingly and run with 25 amplification cycles (chapter V.2.1). A volume of 1 µL of *DpnI* was added into the 50 µL reaction and incubated first at 37 °C for 1 h and followed by 10 min at 80 °C. Next, 5 µL of the reaction mixture was directly transformed into *E. coli* DH5α or TOP10 cells via the heat shock method (chapter V.2.2).

### Colony PCR (cPCR)

In order to evaluate the proper insertion of gene fragments into vectors and the correct transfer of plasmids in retransformations and cotransformations colony PCR was performed. The composition of one reaction mixture is listed in Table 11. In contrast to PCR, the DNA templates here were *E. coli* colonies. The colonies were picked from agar plates with a toothpick, stirred in the reaction mixture and put into a thermocycler for performance of the cPCR program (Table 12). An agarose gel electrophoresis was applied afterwards.

**Table 11: Composition of a single colony PCR reaction mixture.**

Component	cPCR reaction mixture
MQ-H <sub>2</sub> O	15 µL
Green GoTaq® reaction buffer (5x)	4 µL
dNTPs (10 mM)	0.4 µL
Forward primer (10 µM)	0.2 µL
Reverse primer (10 µM)	0.2 µL
GoTaq® DNA polymerase (5 U/µL)	0.2 µL
Total volume	20 µL

**Table 12: Program for a colony PCR.**

Step		Temperature and time	
Initial denaturation		95 °C	5 min
Denaturation	} 30	95 °C	30 sec
Primer annealing		50 °C	30 sec
Elongation		72 °C	1 min/kb
Final elongation		72 °C	2 min/kb
Storage		10 °C	∞

### Agarose gel electrophoresis

For separation of DNA fragments agarose gel electrophoresis was applied. Thereby, a 1.2 % (w/v) agarose gel with 0.3 µg/mL ethidium bromide was prepared and put into an agarose gel electrophoresis system filled with 1x TAE buffer. DNA samples were mixed with 3-6 µL 10x DNA sample buffer and a volume of 10-26 µL were pipetted per gel pocket of the agarose gel. As molecular weight marker 10 µL of GeneRuler™ 1 kb or GeneRuler™ 100 bp Plus DNA ladder (Thermo Fisher Scientific Inc., Waltham, MA, USA) was utilized (Figure A.22, Appendix). The agarose gel electrophoresis was run at 120 V for 20-30 min. Thereafter, the agarose gel was illuminated under UV light for documentation and excision of individual DNA bands.

### Gel extraction

The Wizard® SV Gel and PCR Clean-Up System (Promega Corporation, Madison, WI, USA) was used for the purification of individual DNA fragments which resulted from a PCR or restriction reaction from an agarose gel. The principle process steps of this reaction kit consist of dissolving the agarose gel, applying the gel solution on a column, washing the column and eluting the bound DNA with nuclease-free water.

### Plasmid isolation

For the isolation of plasmids the PureYield™ Plasmid Miniprep System and PureYield™ Plasmid Midiprep System (Promega Corporation, Madison, WI, USA) were made use of. For the first system, 5 mL or 20 mL LB culture of *E. coli* DH5α or TOP10 were incubated overnight. For the second system, a 100 mL or 200 mL overnight culture in LB medium was used and also 0.1 mg/mL RNase A was added to the lysis buffer. The elution from the column in the Miniprep system was performed in 30 µL nuclease-free water and in the Midiprep system in a volume of 500 µL. The concentration of the eluates was determined at 260 nm assuming

that 50 ng/ $\mu$ L dsDNA was equivalent to  $A_{260}=1$ . The quotient  $A_{260/280}$  describing the purity of DNA regarding contamination with proteins was documented as well. The value of  $A_{260/280}$  should be in a range between 1.8 to 1.9 [215]. Whereas, the quotient  $A_{260/230}$  as an indicator of other organic contaminants should be above a value of 2.

### Sequencing

Sequencing of cloned plasmids was conducted with the Mix2Seq Kit and the Mix2Seq Kit NightXpress (Eurofins Genomics Germany GmbH, Ebersberg, Germany). Therefore, a volume of 15  $\mu$ L containing a plasmid concentration of at least 50 ng/ $\mu$ L was mixed with 2  $\mu$ L of primer working solution (10  $\mu$ M) and the solution transferred into a sequencing tube. Apart from that, the sequencing service of IME Fraunhofer Institute for Molecular and Applied Ecology (Aachen, Germany) was made use of. Here, a plasmid solution of 30  $\mu$ L with 500 ng vector in 10 mM Tris-HCl buffer (pH 8.0) and 2  $\mu$ L of primer working solution (10  $\mu$ M) was prepared and submitted for sequencing. The results of both sequencing ways were analyzed in the software Clone Manager Professional Suite 8 (Sci Ed Software, Westminster, CO, USA) and in Chromas (Technelysium Pty. Ltd., Australia).

### RNA isolation and reverse transcription (RT)

For the isolation of total mRNA from *E. coli* culture, 10 mL LB medium in a 100 mL shake flask was inoculated with 100  $\mu$ L of an overnight culture and incubated for 1 h at 37 °C and 160 rpm. Then, the culture was induced with 1 mM IPTG, grown for 5 h and 4 mL of culture harvested at 18,000 xg and 4 °C for 5 min. The cell pellet was resuspended 100  $\mu$ L TE lysis buffer containing 400  $\mu$ g/mL lysozyme and incubated for 5 min at room temperature. Afterwards, a volume of 350  $\mu$ L RLT buffer was mixed with the suspension, then 250  $\mu$ L of 100 % (v/v) ethanol added and the solution gently mixed. The total RNA isolation was performed using the RNeasy Mini Kit (Qiagen, Hilden, Germany). In this respect, 700  $\mu$ L of the solution was transferred to an RNeasy column, centrifuged for 1 min at 17,000 xg, the flow through discarded and the column washed with 700  $\mu$ L of RW1 washing buffer (Qiagen). After another centrifugation step, the column was washed three times with 500  $\mu$ L RPE washing buffer (Qiagen) and centrifuged longer in the end in order to remove ethanol. The final elution was performed in 50  $\mu$ L DEPC water.

The RNA concentration of the RNA isolation was determined photometrically at 260 nm and by applying a dilution of 1:25 in DEPC water. The assumption that an  $OD_{260} = 1$  equals 40  $\mu$ g/mL RNA was used for calculations.

**Table 13: Composition of a single RT reaction mixture and incubation procedure.**

Component	RT reaction mixture	Incubation temperature and time
<b>1</b> DNase digested RNA solution	11 $\mu$ L	10 min at 80 °C,
Reverse primer (10 $\mu$ M)	0.5 $\mu$ L	then 5 min on ice.
<b>2</b> RT buffer (5x)	4 $\mu$ L	30 min at 40 °C,
dNTPs (10 mM)	1 $\mu$ L	30 min at 45 °C,
DEPC water	2.5 $\mu$ L	30 min at 50 °C,
M-MLV (200 U/ $\mu$ L)	1 $\mu$ L	30 min at 55 °C,
		20 min at 75 °C,
		cooking of the reaction mixtures for 2 min,
		then on ice or storage at 4 °C.
Total volume	20 $\mu$ L	

Prior to reverse transcription (RT) the RNA isolate was treated with DNase in a 30  $\mu$ L reaction mixture containing 3  $\mu$ g RNA, 3  $\mu$ L 10 x DNase buffer and 3  $\mu$ L DNaseI (1 U/ $\mu$ L). This reaction mixture was incubated at 37 °C for 30 min followed by addition of 3  $\mu$ L EDTA (25 mM) and inactivation at 75 °C for 20 min. The thus prepared RNA solution was subjected to reverse transcription described in Table 13 using M-MLV reverse transcriptase and 5x RT buffer (Promega Corporation, Madison, WI, USA).

Finally, the cDNA samples were analyzed by performing a cPCR reaction with additional 2  $\mu$ L of cDNA solution and a subsequent agarose gel electrophoresis.

## V.2.2 Microbiological Methods

### Cultivation of *E. coli*

The different strains of *Escherichia coli* were generally cultivated in LB medium with appropriate antibiotics concentration. The pETDuet and pETM6 vectors were supplemented with 100  $\mu$ g/mL ampicillin, pRSFDuet with 50  $\mu$ g/mL kanamycin and pCDFDuet as well as pCDM4 with 50  $\mu$ g/mL streptomycin. For the preparation of overnight cultures, either a LB volume of 5 mL in a test tube or 20 mL LB in a 100 mL shake flask were used. Incubation of these cultures took place in a rotary incubator at 37 °C or in a shaker at 37 °C and 160 rpm.

### **Strain maintenance**

In order to store *E. coli* cultures for a longer period of time, glycerol stocks were prepared and stored at -80 °C. Therefore, 900 µL of 70 % (v/v) sterile glycerol were mixed with 900 µL of an *E. coli* overnight culture in LB medium.

### **Preparation of competent cells**

#### ***Chemically competent cells***

An overnight culture of the *E. coli* cells was diluted in a ratio of 1:100 in 100 mL LB medium in a 1 L shake flask and incubated at 37 °C and 160 rpm. As soon as an OD<sub>600</sub> of 0.5 was reached, the cells were harvested at 4,000 xg and 4 °C for 15 min. After discarding the supernatant, the cells were gently washed with 40 mL of cold TB buffer and incubated for at least 10 min on ice. Then, the cells were centrifuged at 4,000 xg and 4 °C for 15 min and the supernatant removed again. A volume of 8 mL cold TB buffer added with 560 µL DMSO was used for the final resuspension of the cell pellet. Aliquots of 200 µL were prepared, directly frozen in liquid nitrogen and stored at -80 °C until further usage.

As a fast preparation of competent cells, 5 mL LB medium was inoculated with the *E. coli* overnight culture in a ratio between 1:50 to 1:20 and incubated at 37 °C in a rotary incubator. When the LB medium turned turbid, 3-4.5 mL of the cells were harvested in an Eppendorf tube at 18,000 xg and 4 °C for 1 min. Again, the supernatant was discarded and the cells washed in 1 mL cold TB buffer. The thus prepared *E. coli* cell pellet was resuspended in 50-200 µL of cold TB buffer and was used for transformation.

#### ***Electrocompetent cells***

An overnight culture of the *E. coli* cells was diluted in a ratio of 1:100 in a total of 2 L LB medium. Thereby, each 2 L shake flask was containing 200 mL of freshly inoculated LB medium and was incubated at 37 °C and 160 rpm. When an OD<sub>600</sub> of 0.5 was reached, the cells were harvested at 4,000 xg and 4 °C for 6 min. Like before, the supernatant was removed and the cells were gently washed with cold 10 % (v/v) glycerol in multiple rounds. In each round, the cells were either combined from two shake flasks or resuspended in a smaller washing volume. After 5-8 rounds, the final resuspension was performed in 4 mL of 10 % (v/v) glycerol. Subsequently, aliquots of 100 µL were prepared, directly frozen in liquid nitrogen and stored at -80 °C until usage for electroporation.



## Transformation

### *Electrotransformation*

Electrotransformation was applied with all ligation products. Thereby, 100  $\mu$ L of electrocompetent *E. coli* DH5 $\alpha$  were thawed on ice for 10 min and then gently mixed with 5  $\mu$ L of precipitated DNA solution. The competent cells were transferred into a cold electroporation cuvette and treated with an impulse of 2500 V for 5 ms in an electroporation device. An amount of 900  $\mu$ L of pre-warmed SOC medium was directly mixed with the cells. After that, the mixture was transferred into a new Eppendorf tube for incubation at 37 °C and 800 rpm for 40 min in a thermoshaker. Then, 100  $\mu$ L and a concentrate of 900  $\mu$ L of the cell solution were plated on antibiotic agar plates. The agar plates were incubated at 37 °C overnight.

### *Heat shock transformation*

In general, heat shock transformation was used for retransformation of plasmids into *E. coli* strains and for co-transformations of multiple plasmids as well as for FastCloning and site-directed mutagenesis. For that, 200  $\mu$ L of chemically competent cells were thawed on ice for 10 min and gently mixed with 2  $\mu$ L of plasmid solution. After an incubation period for 30 min on ice, the cells were subjected to a heat shock in a thermoblock at 42 °C for 45 sec. The cells were chilled on ice for 5 min, then 800  $\mu$ L of pre-warmed SOC medium was added and the cells were incubated at 37 °C and 800 rpm for 40 min in a thermoshaker. Subsequently, 50-100  $\mu$ L of the cell suspension was plated on pre-warmed agar plates containing corresponding antibiotics.

For transformation of multiple plasmids into one *E. coli* strain, the plasmids were added in successive rounds of co-transformations and confirmed with cPCR in between. In order to add the second or third plasmid, the fast preparation protocol for competent cells was followed for the relevant *E. coli* strain (chapter V.2.2). The so prepared *E. coli* cell suspension was directly used in the transformation procedure described above.

## Protein expression

The expression of proteins in different strains of *E. coli* was performed in TB medium with the concentration of corresponding antibiotics as described previously (chapter V.2.2). For investigation of the expression of individual proteins and protein combinations, 50 mL of TB medium in a 250 mL shake flask or 100 mL of TB medium in a 500 mL shake flask were inoculated with the relevant *E. coli* overnight culture in ratios between 1:50 to 1:10. These shake flasks were incubated at 37 °C with 180 rpm. Once an OD<sub>600</sub> between 0.6-1 was reached,

the cultures were induced by the addition of 1 mM IPTG and transferred to the appropriate expression temperature at 180 rpm. At this time point, a 5/OD<sub>600</sub> sample (0 h) was taken. Further 5/OD<sub>600</sub> samples were collected over the time of the expression. Predominantly, these samples were taken at 4 h and 20 h after induction.

With the aim of imitating expression conditions from the growing cell assay, the expression was performed in 50 mL of TB medium in a 500 mL shake flask. After inoculation with the overnight culture, the shake flask was incubated at 37 °C and 160 rpm until an OD<sub>600</sub> of around 2 (OD<sub>600</sub> from 1.5 to 2.5) was reached. The cultures were induced with 0.5 mM IPTG and the expression was performed at 26 °C and 160 rpm for 24 h. 5/OD<sub>600</sub> samples were collected at 0 h, 4 h and at 24 h of expression.

For large-scale expression and whole-cell biotransformation, 200 mL of TB medium in a 1 L shake flask or 400 mL of TB medium in a 2 L shake flask were inoculated with an *E. coli* overnight culture and incubated at 37 °C and 180 rpm. When an OD<sub>600</sub> between 0.6-1 was reached, the cultures were induced by adding 1 mM IPTG and put to 180 rpm and either 37 °C or 26 °C for 20 h. Again 5/OD<sub>600</sub> samples were collected at 0 h, 4 h and 20 h of expression. After 20 h, the cells were harvested at 4,000 xg and 4 °C for 15 min in beakers. Then, the cells were washed twice with 40 mL of cold KP<sub>1</sub> buffer on ice. All liquid was removed from the beakers, the wet cell weight  $g_{wew}$  of the cells was determined and the cell pellets stored at -20 °C until further usage.

### Cell disruption

The harvested 5/OD<sub>600</sub> cell pellets were resuspended in 500 µL resuspension buffer and sonicated using ultrasonic waves with a cycle of 0.5 (0.5 sec power discharge followed by 0.5 sec pause) and an amplitude of 65 % (~ 170 µm) on ice. The procedure consisted of 45 sec ultrasonic treatment, 30 sec pause and another 45 sec sonication. Afterwards, the samples were centrifuged at 18,000 xg and 4 °C for 1 min and separated into supernatant (S) and pellet fraction (P) before their application in SDS-PAGE analysis (chapter V.2.3).

## V.2.3 Biochemical Methods

### SDS-PAGE

The method of sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was applied for separation of proteins according to their molecular weights. Here, a discontinuous SDS-PAGE was performed using an SDS gel consisting of a 12 % (v/v) resolving gel and a 4 % (v/v) stacking gel (Table 14). When the resolving gel was prepared, the top of the gel was

covered with isopropanol, which was then removed after gel solidification for subsequent preparation of the stacking gel.

**Table 14: Composition of one SDS gel.**

Component	12 % resolving gel		4 % stacking gel	
MQ-H <sub>2</sub> O	1.06 mL	(21.1 %, v/v)	1.83 mL	(72.9 %, v/v)
1 M Tris-HCl (pH 8.8)	1.88 mL	(37.5 %, v/v)	–	
1 M Tris-HCl (pH 6.8)	–		313 µL	(12.5 %, v/v)
Acrylamide (30 %, w/v)	2.0 mL	(40 %, v/v)	415 µL	(16.6 %, v/v)
SDS (10 %, w/v)	50 µL	(1 %, v/v)	25 µL	(1 %, v/v)
APS (20 %, w/v)	15 µL	(0.3 %, v/v)	7.5 µL	(0.3 %, v/v)
TEMED	5 µL	(0.1 %, v/v)	2.5 µL	(0.1 %, v/v)
Total volume	5 mL		2.5 mL	

The disrupted and separated 5/OD<sub>600</sub> samples were mixed with 5x protein reducing loading buffer and heated to 95-100 °C for 8 min. Thus, 80 µL of the supernatant was mixed with 20 µL 5x protein reducing loading buffer, while the pellet fraction was resuspended in 200 µL of 1x protein reducing loading buffer, which had been diluted in resuspension buffer before. 15 µL of the samples were applied to the SDS gel pockets and 5 µL of P7712 or P7719 Color Prestained Protein Standards (New England Biolabs) were used (Figure A.22, Appendix). As controls for western blot analysis, PVX-S-tag (25 µg/mL) and mCherry-his<sub>6</sub> (50 µg/mL) were utilized, of which 10 µL were applied per gel pocket. The thus produced SDS gels were either used for western blot analysis or dyed with Coomassie staining solution for at least 30 min and decolorized with Coomassie destaining solution afterwards. In the end, the destained SDS gel was photographed for documentation.

### Western blot

In order to analyze the expression bands of specific proteins, western blot was performed using the Tankblot system. Therefore, a recently prepared SDS gel and the Amersham™ Hybond™-C nitrocellulose membrane (Cytiva/GE Healthcare Life Sciences, Marlborough, MA, USA) were put between several layers of Whatman blotting paper (3.6 mm) and two sponges. All components had been individually equilibrated in cold 1x Tankblot buffer before and were sandwiched in blotting cassettes. These cassettes were assembled with an ice pack into the blotting chambers, which were further cooled by surrounding ice and filled with 1x Tankblot buffer. The transfer of proteins from the SDS gel to the nitrocellulose membrane by an electric

field was carried out at 100 V for 60 min. Afterwards, the membrane was incubated in blocking solution in order to fill all free binding sites and to avoid later unspecific detection. After short washing with 1x PBS, the respective primary antibody was added in a dilution of 1:5,000 in 1x PBS and incubated in an orbital shaker overnight. The applied primary antibodies were ‘ $\alpha$ -his’ for all proteins fused with a His-tag and the ‘S-peptide Epitope Tag Antibody’ for fusion enzymes with an S-protein tag (Table 4). The next day, the membrane was washed three times with 1x PBS and incubated with the secondary antibody in a dilution of 1:5,000 in 1x PBS for at least 3 h. Another three washing steps with 1x PBS were performed, before the membrane was equilibrated in AP buffer and the substrate solution NBT/BCIP was added in a dilution of 1:100 in AP buffer. Subsequently, the membrane was incubated in the dark, until the formation of bands was clearly visible, and a photo was taken for documentation. The principle of detection is based on the activity of the alkaline phosphatase (AP), which is conjugated to the secondary antibody and dephosphorylates BCIP as substrate. The resulting indoxyl intermediate oxidizes under formation of an insoluble dimer of blue color. The substrate NBT is reduced in the process to NBT diformazan, which precipitates with a dark red color as well. Thus, a dark purple to black staining of specific protein bands is obtained in the end.

#### **Whole-cell biotransformation (resting cell assay)**

The stored cell pellets from large-scale expressions were thawed on ice for 30 min (chapter V.2.2). The final resuspension of the cell pellets was performed in  $KP_i$  buffer (50 mM, pH 7.5) with 2 % (w/v) glucose and 1 mM IPTG yielding a cell suspension of 70 g<sub>wcw</sub>/L. Aliquots of 10 mL were used as individual experiments in 100 mL shake flasks. Substrates prepared in DMSO in individual substrate concentrations were added to the aliquots and the cell suspension was incubated at 26 °C and 160 rpm over time. Samples of 500  $\mu$ L were taken at substrate addition and at different time points during the whole-cell biotransformation. The collected samples were centrifuged at 17,000 xg for 1 min and the supernatant was stored at -20 °C until the extraction process. This procedure was mainly adapted from Ricklefs *et al.* (2016) [201].

The two-cells one-pot experiment was performed as a variant of the resting cell assay. Here, the cell suspensions had a cell concentration of 140 g<sub>wcw</sub>/L and two subsets of cells were added successively. A volume of 10 mL of the first set of cells was incubated with 2 mM L-tyrosine added at the start and samples of 500  $\mu$ L were collected up to 24 h. After that time point, a second set of cells was mixed with the first subset in a 1:1 ratio of volumes and samples were taken until 72 h.

### Growing cell assay

An overnight culture of *E. coli* BL21(DE3) or the genome-integrated strains was inoculated in a volume of 10 mL TB medium in a 100 mL shake flask and cultivated at 37 °C and 160 rpm. Once an OD<sub>600</sub> of around 2 was reached, the culture was induced by addition of 0.5 mM IPTG and the substrate L-tyrosine dissolved in DMSO was added in a concentration of 1 mM unless indicated otherwise. Then, the cultures were transferred to a temperature of 26 °C at 160 rpm for 48 h. Samples of 500 µL were collected at the time point of induction (0 h) and at 4 h, 24 h and 48 h, centrifuged at 17,000 xg for 1 min and the supernatant stored at -20 °C until further usage. As control experiment, no substrate was added at the time point of induction.

### Toxicity test of substances

For evaluation of the toxicity of different substrates and intermediates of the phenylpropanoid pathway on *E. coli* cells, growth curves of individual cultures were recorded. Thus, 20 mL of LB medium was inoculated with 500 µL of an *E. coli* BL21(DE3) overnight culture and supplemented with different concentrations between 0-20 mM of the examined substance. The cultures were grown at 37 °C and 180 rpm and the OD<sub>600</sub> value was measured over time. All applied substance stocks were prepared as 1 M in DMSO. As control, growth curves from *E. coli* with solely DMSO were performed.

## V.2.4 Analytical Methods

### Extraction procedure

The samples from the whole-cell biotransformations and growing cell assays were thawed on ice prior to extraction using ethylacetate with internal standard (IS). A volume of 500 µL of ethylacetate was added to a sample, both phases mixed by vortexing for 45 sec and then centrifuged at 4 °C and 14,000 xg for 1 min. The upper ethylacetate phase was collected in a separate Eppendorf tube and the extraction process repeated twice. The resulting pooled ethylacetate phases were either placed under a fume cabinet with open lids overnight or put in a concentrator device for evaporation of the organic ethylacetate phase. The dried extract was finally dissolved with 300 µL of ROTISOLV® HPLC methanol (Carl Roth GmbH & Co. KG), centrifuged at 17,000 xg for 10 min and analyzed in high performance liquid chromatography (HPLC). As general internal standard, 100 µM cinnamic acid was utilized. Unless cinnamic acid was a possible product, then 200 µM coniferaldehyde was used.

### HPLC measurement

The analysis of the extracted and dissolved samples was carried out with different HPLC systems and their respective reversed phase HPLC column indicated in Table 2. The used solvents were MQ-H<sub>2</sub>O with 0.1 % formic acid (A) and HPLC grade methanol (B), which had been degassed in an ultrasonic waterbath prior to the measurement. The HPLC method started with a gradient from 20 % to 35 % B from 0-5 min and stayed at 35 % B within 5-10 min. A second gradient from 35 % to 75 % B followed in minutes 10-25. Afterwards, the concentration of solvent B was increased to 90 % within one minute, stayed at 90 % B for another minute before the content of solvent B was reduced to 20 % within one minute again. The HPLC method was run at 20 % B until 35 min in order to equilibrate the column for the next sample run. Depending on the HPLC system slight adjustments regarding the HPLC method were undertaken. In general, a flow rate of 0.5 or 0.8 mL/min, an oven temperature of 30 °C and a detection at 280 nm were deployed.

At the Department of Biochemical and Chemical Engineering at the Technical University Dortmund, the HPLC method started with 2 % of solvent B for 2 min and a flow rate of 0.3 mL/min. The amount of B was increased to 25 % from minute 2 to 4 and a gradient was run from 25 % to 30 % B until minute 15. Afterwards, the concentration of solvent B increased within 3 min to 90 %, stayed there for 2 min and was decreased within 1 min to 2 %. The method run finished after a total of 26 min.

Apart from the samples from resting and growing cell assay, standard curves and extraction samples containing only the internal standard were measured as well. The mean area value of IS, which was calculated from all IS samples of a single extraction, was applied for normalization of the other samples from the same extraction run. In order to take into account the possible different partition coefficients, the samples containing the standard curves were extracted, too. Thus, the concentration of a specific substance in a sample was calculated by correction of its area value with the quotient of IS area value in the sample to the mean IS area value. Subsequently, the division by the slope of the respective standard curve followed yielding the concentration of a specific substance in the sample. When conversion yields were calculated, the respective mean product concentration at a certain time point was divided by the mean substrate concentration at addition time point  $t_0$  (0 h) and multiplied with 100 %.

## VI. References

1. World Health Organization (2020). – Global Health Estimates 2019: Deaths by cause, age, sex, by country and by region. 2000-2019. Available at: <https://www.who.int/data/stories/leading-causes-of-death-and-disability-2000-2019-a-visual-summary> (accessed on 21 July 2021).
2. Tokudome S., Ichikawa Y., Okuyama H., Tokudome Y., Goto C., Imaeda N., Kuriki K., Suzuki S., Shibata K., Jiang J., Wang J. & Takeda E. (2004). – The Mediterranean vs the Japanese diet. *European Journal of Clinical Nutrition*, **58** (9), 1323–1323. doi:10.1038/sj.ejcn.1601970.
3. Adlercreutz H. & Mazur W. (1997). – Phyto-oestrogens and Western diseases. *Annals of Medicine*, **29** (2), 95–120. doi:10.3109/07853899709113696.
4. Wang L.Q. (2002). – Mammalian phytoestrogens: enterodiol and enterolactone. *Journal of Chromatography B*, **777** (1), 289–309. doi:10.1016/S1570-0232(02)00281-7.
5. Landete J.M. (2012). – Plant and mammalian lignans: A review of source, intake, metabolism, intestinal bacteria and health. *Food Research International*, **46** (1), 410–424. doi:10.1016/j.foodres.2011.12.023.
6. Peterson J., Dwyer J., Adlercreutz H., Scalbert A., Jacques P. & McCullough M.L. (2010). – Dietary lignans: physiology and potential for cardiovascular disease risk reduction. *Nutrition Reviews*, **68** (10), 571–603. doi:10.1111/j.1753-4887.2010.00319.x.
7. Zálešák F., Bon D.J.Y.D. & Pospíšil J. (2019). – Lignans and Neolignans: Plant secondary metabolites as a reservoir of biologically active substances. *Pharmacological Research*, **146** (104284), 1–27. doi:10.1016/j.phrs.2019.104284.
8. Barre D.E., Mizier-Barre K.A., Stelmach E., Hobson J., Griscti O., Rudiuk A. & Muthuthevar D. (2012). – Flaxseed lignan complex administration in older human type 2 diabetics manages central obesity and prothrombosis– An invitation to further investigation into polypharmacy reduction. *Journal of Nutrition and Metabolism*, **2012** (585170), 1–7. doi:10.1155/2012/585170.
9. Satake H., Koyama T., Bahabadi S.E., Matsumoto E., Ono E. & Murata J. (2015). – Essences in metabolic engineering of lignan biosynthesis. *Metabolites*, **5** (2), 270–290. doi:10.3390/metabo5020270.
10. Umezawa T. (2003). – Diversity in lignan biosynthesis. *Phytochemistry Reviews*, **2** (3), 371–390. doi:10.1023/B:PHYT.0000045487.02836.32.
11. Owen R.W., Giacosa A., Hull W.E., Haubner R., Spiegelhalder B. & Bartsch H. (2000). – The antioxidant/anticancer potential of phenolic compounds isolated from olive oil. *European Journal of Cancer*, **36** (10), 1235–1247. doi:10.1016/S0959-8049(00)00103-9.
12. During A., Debouche C., Raas T. & Larondelle Y. (2012). – Among plant lignans, pinorexinol has the strongest antiinflammatory properties in human intestinal Caco-2 cells. *The Journal of Nutrition*, **142** (10), 1798–1805. doi:10.3945/jn.112.162453.
13. Kim H.Y., Kim J.K., Choi J.H., Jung J.Y., Oh W.Y., Kim D.C., Lee H.S., Kim Y.S., Kang S.S., Lee S.H. & Lee S.M. (2010). – Hepatoprotective effect of pinorexinol on carbon tetrachloride–induced hepatic damage in mice. *Journal of Pharmacological Sciences*, **112** (1), 105–112. doi:10.1254/jphs.09234FP.
14. Nakano D., Itoh C., Ishii F., Kawanishi H., Takaoka M., Kiso Y., Tsuruoka N., Tanaka T. & Matsumura Y. (2003). – Effects of sesamin on aortic oxidative stress and endothelial dysfunction in deoxycorticosterone acetate-salt hypertensive rats. *Biological and Pharmaceutical Bulletin*, **26** (12), 1701–1705. doi:10.1248/bpb.26.1701.
15. Utsunomiya T., Chavali S.R., Zhong W.W. & Forse R.A. (2000). – Effects of sesamin-supplemented dietary fat emulsions on the *ex vivo* production of lipopolysaccharide-induced prostanoids and tumor necrosis factor  $\alpha$  in rats. *The American Journal of Clinical Nutrition*, **72** (3), 804–808. doi:10.1093/ajcn/72.3.804.

16. Nakai M., Harada M., Nakahara K., Akimoto K., Shibata H., Miki W. & Kiso Y. (2003). – Novel antioxidative metabolites in rat liver with ingested sesamin. *Journal of Agricultural and Food Chemistry*, **51** (6), 1666–1670. doi:10.1021/jf0258961.
17. Shain L. & Hillis W.E. (1971). – Phenolic extractives in Norway spruce and their effects on *Fomes annosus*. *Phytopathology*, **61**, 841–845. doi:10.1094/Phyto-61-841.
18. González J.A., Estévez-Braun A., Estévez-Reyes R., Bazzocchi I.L., Moujir L., Jimenez I.A., Ravelo A.G. & González A.G. (1995). – Biological activity of secondary metabolites from *Bupleurum salicifolium* (Umbelliferae). *Experientia*, **51** (1), 35–39.
19. Kumarasamy Y., Nahar L., Cox P.J., Dinan L.N., Ferguson C.A., Finnie D.A., Jaspars M. & Sarker S.D. (2003). – Biological activity of lignans from the seeds of *Centaurea scabiosa*. *Pharmaceutical Biology*, **41** (3), 203–206. doi:10.1076/phbi.41.3.203.15099.
20. Markulin L., Corbin C., Renouard S., Drouet S., Gutierrez L., Mateljak I., Auguin D., Hano C., Fuss E. & Lainé E. (2019). – Pinoresinol–lariciresinol reductases, key to the lignan synthesis in plants. *Planta*, **249** (6), 1695–1714. doi:10.1007/s00425-019-03137-y.
21. Lau W. & Sattely E.S. (2015). – Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science*, **349** (6253), 1224–1228. doi:10.1126/science.aac7202.
22. Decembrino D., Ricklefs E., Wohlgemuth S., Girhard M., Schullehner K., Jach G. & Urlacher V.B. (2020). – Assembly of plant enzymes in *E. coli* for the production of the valuable (–)-podophyllotoxin precursor (–)-pluviatolide. *ACS Synthetic Biology*, **9** (11), 3091–3103. doi:10.1021/acssynbio.0c00354.
23. Gordaliza M., García P.A., Miguel del Corral J.M., Castro M.A. & Gómez-Zurita M.A. (2004). – Podophyllotoxin: distribution, sources, applications and new cytotoxic derivatives. *Toxicon*, **44** (4), 441–459. doi:10.1016/j.toxicon.2004.05.008.
24. Canel C., Moraes R.M., Dayan F.E. & Ferreira D. (2000). – Podophyllotoxin. *Phytochemistry*, **54** (2), 115–120. doi:10.1016/S0031-9422(00)00094-7.
25. Srivastava V., Negi A.S., Kumar J.K., Gupta M.M. & Khanuja S.P.S. (2005). – Plant-based anticancer molecules: A chemical and biological profile of some important leads. *Bioorganic & Medicinal Chemistry*, **13** (21), 5892–5908. doi:10.1016/j.bmc.2005.05.066.
26. Courdavault V., O'Connor S.E., Oudin A., Besseau S. & Papon N. (2020). – Towards the microbial production of plant-derived anticancer drugs. *Trends in Cancer*, **6** (6), 444–448. doi:10.1016/j.trecan.2020.02.004.
27. Chaurasia O.P., Ballabh B., Tayade A., Kumar R., Kumar G.P. & Singh S.B. (2012). – Podophyllum L.: An endangered and anticancerous medicinal plant– An overview. *Indian Journal of Traditional Knowledge*, **11** (2), 234–241.
28. Izadifar M. & Baik O.D. (2008). – An optimum ethanol–water solvent system for extraction of podophyllotoxin: Experimental study, diffusivity determination and modeling. *Separation and Purification Technology*, **63** (1), 53–60. doi:10.1016/j.seppur.2008.03.041.
29. Jackson D.E. & Dewick P.M. (1984). – Aryltetralin lignans from *Podophyllum hexandrum* and *Podophyllum peltatum*. *Phytochemistry*, **23** (5), 1147–1152. doi:10.1016/S0031-9422(00)82628-X.
30. Canel C., Dayan F.E., Ganzera M., Khan I.A., Rimando A., Jr C.L.B. & Moraes R.M. (2001). – High yield of podophyllotoxin from leaves of *Podophyllum peltatum* by in situ conversion of podophyllotoxin 4-O-β-D-glucopyranoside. *Planta Medica*, **67** (1), 97–99. doi:10.1055/s-2001-10636.
31. Maqbool M., Cushman K.E., Gerard P.D., Bedir E., Lata H. & Moraes R.M. (2004). – Podophyllotoxin content in leaves of Eastern red cedar (*Juniperus virginiana*). *Acta Horticulturae*, **629**, 87–92. doi:10.17660/ActaHortic.2004.629.11.
32. Kartal M., Konuklugil B., Indrayanto G. & Alfermann A.W. (2004). – Comparison of different extraction methods for the determination of podophyllotoxin and 6-methoxypodophyllotoxin in



- Linum species. *Journal of Pharmaceutical and Biomedical Analysis*, **35** (3), 441–447. doi:10.1016/j.jpba.2004.01.016.
33. Abdullah S.H., Zhao S., Mittal G.S. & Baik O.D. (2012). – Extraction of podophyllotoxin from *Podophyllum peltatum* using pulsed electric field treatment. *Separation and Purification Technology*, **93**, 92–97. doi:10.1016/j.seppur.2012.04.002.
  34. Ionkova I., Antonova I., Momekov G. & Fuss E. (2010). – Production of podophyllotoxin in *Linum linearifolium* in vitro cultures. *Pharmacognosy Magazine*, **6** (23), 180–185. doi:10.4103/0973-1296.66932.
  35. Seegers C.L.C., Setroikromo R. & Quax W.J. (2017). – Towards metabolic engineering of podophyllotoxin production. . In *Natural Products and Cancer Drug Discovery*. Badria, F. A., IntechOpen, London. pp 287–306doi:10.5772/67615.
  36. Boerjan W., Ralph J. & Baucher M. (2003). – Lignin biosynthesis. *Annual Review of Plant Biology*, **54** (1), 519–546. doi:10.1146/annurev.arplant.54.031902.134938.
  37. Vanholme R., Morreel K., Ralph J. & Boerjan W. (2008). – Lignin engineering. *Current Opinion in Plant Biology*, **11** (3), 278–285. doi:10.1016/j.pbi.2008.03.005.
  38. Moura J.C.M.S., Bonine C.A.V., De Oliveira Fernandes Viana J., Dornelas M.C. & Mazzafera P. (2010). – Abiotic and biotic stresses and changes in the lignin content and composition in plants. *Journal of Integrative Plant Biology*, **52** (4), 360–376. doi:10.1111/j.1744-7909.2010.00892.x.
  39. Miao Y.C. & Liu C.J. (2010). – ATP-binding cassette-like transporters are involved in the transport of lignin precursors across plasma and vacuolar membranes. *Proceedings of the National Academy of Sciences*, **107** (52), 22728–22733. doi:10.1073/pnas.1007747108.
  40. Wang Y., Chantreau M., Sibout R. & Hawkins S. (2013). – Plant cell wall lignification and monolignol metabolism. *Frontiers in Plant Science*, **4** (220), 1–14. doi:10.3389/fpls.2013.00220.
  41. Hahlbrock K. & Scheel D. (1989). – Physiology and molecular biology of phenylpropanoid metabolism. *Annual Review of Plant Physiology and Plant Molecular Biology*, **40** (1), 347–369. doi:10.1146/annurev.pp.40.060189.002023.
  42. Rodrigues J.L., Prather K.L.J., Kluskens L.D. & Rodrigues L.R. (2015). – Heterologous Production of Curcuminoids. *Microbiology and Molecular Biology Reviews*, **79** (1), 39–60. doi:10.1128/MMBR.00031-14.
  43. Watts K.T., Lee P.C. & Schmidt-Dannert C. (2006). – Biosynthesis of plant-specific stilbene polyketides in metabolically engineered *Escherichia coli*. *BMC Biotechnology*, **6** (22), 1–12. doi:10.1186/1472-6750-6-22.
  44. Eudes A., Juminaga D., Baidoo E.E.K., Collins F.W., Keasling J.D. & Loqué D. (2013). – Production of hydroxycinnamoyl anthranilates from glucose in *Escherichia coli*. *Microbial Cell Factories*, **12** (62), 1–10. doi:10.1186/1475-2859-12-62.
  45. Humphreys J.M. & Chapple C. (2002). – Rewriting the lignin roadmap. *Current Opinion in Plant Biology*, **5** (3), 224–229. doi:10.1016/S1369-5266(02)00257-1.
  46. Schoch G., Goepfert S., Morant M., Hehn A., Meyer D., Ullmann P. & Werck-Reichhart D. (2001). – CYP98A3 from *Arabidopsis thaliana* Is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *Journal of Biological Chemistry*, **276** (39), 36566–36574. doi:10.1074/jbc.M104047200.
  47. Vanholme R., De Meester B., Ralph J. & Boerjan W. (2019). – Lignin biosynthesis and its integration into metabolism. *Current Opinion in Biotechnology*, **56**, 230–239. doi:10.1016/j.copbio.2019.02.018.
  48. Barros J., Escamilla-Trevino L., Song L., Rao X., Serrani-Yarce J.C., Palacios M.D., Engle N., Choudhury F.K., Tschaplinski T.J., Venables B.J., Mittler R. & Dixon R.A. (2019). – 4-Coumarate 3-hydroxylase in the lignin biosynthesis pathway is a cytosolic ascorbate peroxidase. *Nature Communications*, **10** (1994), 1–11. doi:10.1038/s41467-019-10082-7.

49. Humphreys J.M., Hemm M.R. & Chapple C. (1999). – New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. *Proceedings of the National Academy of Sciences*, **96** (18), 10045–10050. doi:10.1073/pnas.96.18.10045.
50. Wang J.P., Shuford C.M., Li Q., Song J., Lin Y.C., Sun Y.H., Chen H.C., Williams C.M., Muddiman D.C., Sederoff R.R. & Chiang V.L. (2012). – Functional redundancy of the two 5-hydroxylases in monolignol biosynthesis of *Populus trichocarpa*: LC–MS/MS based protein quantification and metabolic flux analysis. *Planta*, **236** (3), 795–808. doi:10.1007/s00425-012-1663-5.
51. Osakabe K., Tsao C.C., Li L., Popko J.L., Umezawa T., Carraway D.T., Smeltzer R.H., Joshi C.P. & Chiang V.L. (1999). – Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms. *Proceedings of the National Academy of Sciences*, **96** (16), 8955–8960. doi:10.1073/pnas.96.16.8955.
52. Dixon R.A., Chen F., Guo D. & Parvathi K. (2001). – The biosynthesis of monolignols: a “metabolic grid”, or independent pathways to guaiacyl and syringyl units? *Phytochemistry*, **57** (7), 1069–1084. doi:10.1016/S0031-9422(01)00092-9.
53. Barros J. & Dixon R.A. (2020). – Plant phenylalanine/tyrosine ammonia-lyases. *Trends in Plant Science*, **25** (1), 66–79. doi:10.1016/j.tplants.2019.09.011.
54. Kong J.Q. (2015). – Phenylalanine ammonia-lyase, a key component used for phenylpropanoids production by metabolic engineering. *RSC Advances*, **5** (77), 62587–62603. doi:10.1039/C5RA08196C.
55. Williams J.S., Thomas M. & Clarke D.J.Y. (2005). – The gene *stlA* encodes a phenylalanine ammonia-lyase that is involved in the production of a stilbene antibiotic in *Photorhabdus luminescens* TT01. *Microbiology*, **151** (8), 2543–2550. doi:10.1099/mic.0.28136-0.
56. Rösler J., Krekel F., Amrhein N. & Schmid J. (1997). – Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. *Plant Physiology*, **113** (1), 175–179. doi:10.1104/pp.113.1.175.
57. Barros J., Serrani-Yarce J.C., Chen F., Baxter D., Venables B.J. & Dixon R.A. (2016). – Role of bifunctional ammonia-lyase in grass cell wall biosynthesis. *Nature Plants*, **2** (6), 1–9. doi:10.1038/nplants.2016.50.
58. Vannelli T., Wei Qi W., Sweigard J., Gatenby A.A. & Sariaslani F.S. (2007). – Production of p-hydroxycinnamic acid from glucose in *Saccharomyces cerevisiae* and *Escherichia coli* by expression of heterologous genes from plants and fungi. *Metabolic Engineering*, **9** (2), 142–151. doi:10.1016/j.ymben.2006.11.001.
59. Kyndt J.A., Meyer T.E., Cusanovich M.A. & Van Beeumen J.J. (2002). – Characterization of a bacterial tyrosine ammonia lyase, a biosynthetic enzyme for the photoactive yellow protein. *FEBS Letters*, **512** (1–3), 240–244. doi:10.1016/S0014-5793(02)02272-X.
60. Berner M., Krug D., Bihlmaier C., Vente A., Müller R. & Bechthold A. (2006). – Genes and enzymes involved in caffeic acid biosynthesis in the actinomycete *Saccharothrix espanaensis*. *Journal of Bacteriology*, **188** (7), 2666–2673. doi:10.1128/JB.188.7.2666-2673.2006.
61. Xiang L. & Moore B.S. (2005). – Biochemical characterization of a prokaryotic phenylalanine ammonia lyase. *Journal of Bacteriology*, **187** (12), 4286–4289. doi:10.1128/JB.187.12.4286-4289.2005.
62. Zirkle R., Ligon J.M. & Molnár I. 2004 (2004). – Heterologous production of the antifungal polyketide antibiotic soraphen A of *Sorangium cellulosum* So ce26 in *Streptomyces lividans*. *Microbiology*, **150** (8), 2761–2774. doi:10.1099/mic.0.27138-0.
63. Hyun M.W., Yun Y.H., Kim J.Y. & Kim S.H. (2011). – Fungal and plant phenylalanine ammonia-lyase. *Mycobiology*, **39** (4), 257–265. doi:10.5941/MYCO.2011.39.4.257.

64. MacDonald M.C., Arivalagan P., Barre D.E., MacInnis J.A. & D'Cunha G.B. (2016). – *Rhodotorula glutinis* phenylalanine/tyrosine ammonia lyase enzyme catalyzed synthesis of the methyl ester of para-hydroxycinnamic acid and its potential antibacterial activity. *Frontiers in Microbiology*, **7** (281), 1–11. doi:10.3389/fmicb.2016.00281.
65. Jendresen C.B., Stahlhut S.G., Li M., Gaspar P., Siedler S., Förster J., Maury J., Borodina I. & Nielsen A.T. (2015). – Highly active and specific tyrosine ammonia-lyases from diverse origins enable enhanced production of aromatic compounds in bacteria and *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, **81** (13), 4458–4476. doi:10.1128/AEM.00405-15.
66. Cochrane F.C., Davin L.B. & Lewis N.G. (2004). – The Arabidopsis phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. *Phytochemistry*, **65** (11), 1557–1564. doi:10.1016/j.phytochem.2004.05.006.
67. Watts K.T., Mijts B.N., Lee P.C., Manning A.J. & Schmidt-Dannert C. (2006). – Discovery of a substrate selectivity switch in tyrosine ammonia-lyase, a member of the aromatic amino acid lyase family. *Chemistry & Biology*, **13** (12), 1317–1326. doi:10.1016/j.chembiol.2006.10.008.
68. Bartsch S. & Bornscheuer U.T. (2009). – A single residue influences the reaction mechanism of ammonia lyases and mutases. *Angewandte Chemie International Edition*, **48** (18), 3362–3365. doi:10.1002/anie.200900337.
69. Appert C., Logemann E., Hahlbrock K., Schmid J. & Amrhein N. (1994). – Structural and catalytic properties of the four phenylalanine ammonia-lyase isoenzymes from parsley (*Petroselinum Crispum* Nym.). *European Journal of Biochemistry*, **225** (1), 491–499. doi:10.1111/j.1432-1033.1994.00491.x.
70. Tan X., Song W., Chen X., Liu L. & Wu J. (2020). – Recent advances in biocatalytic derivatization of L-tyrosine. *Applied Microbiology and Biotechnology*, **104** (23), 9907–9920. doi:10.1007/s00253-020-10949-6.
71. Calabrese J.C., Jordan D.B., Boodhoo A., Sariaslani S. & Vannelli T. (2004). – Crystal structure of phenylalanine ammonia lyase: Multiple helix dipoles implicated in catalysis. *Biochemistry*, **43** (36), 11403–11416. doi:10.1021/bi049053+.
72. Poppe L. (2001). – Methylidene-imidazolone: a novel electrophile for substrate activation. *Current Opinion in Chemical Biology*, **5** (5), 512–524. doi:10.1016/S1367-5931(00)00253-2.
73. Schwede T.F., Rétey J. & Schulz G.E. (1999). – Crystal structure of histidine ammonia-lyase revealing a novel polypeptide modification as the catalytic electrophile. *Biochemistry*, **38** (17), 5355–5361. doi:10.1021/bi982929q.
74. Louie G.V., Bowman M.E., Moffitt M.C., Baiga T.J., Moore B.S. & Noel J.P. (2006). – Structural determinants and modulation of substrate specificity in phenylalanine-tyrosine ammonia-lyases. *Chemistry & Biology*, **13** (12), 1327–1338. doi:10.1016/j.chembiol.2006.11.011.
75. Jun S.Y., Sattler S.A., Cortez G.S., Vermerris W., Sattler S.E. & Kang C. (2018). – Biochemical and structural analysis of substrate specificity of a phenylalanine ammonia-lyase. *Plant Physiology*, **176** (2), 1452–1468. doi:10.1104/pp.17.01608.
76. Schuster B. & Rétey J. (1995). – The mechanism of action of phenylalanine ammonia-lyase: the role of prosthetic dehydroalanine. *Proceedings of the National Academy of Sciences*, **92** (18), 8433–8437. doi:10.1073/pnas.92.18.8433.
77. Ritter H. & Schulz G.E. (2004). – Structural basis for the entrance into the phenylpropanoid metabolism catalyzed by phenylalanine ammonia-lyase. *The Plant Cell*, **16** (12), 3426–3436. doi:10.1105/tpc.104.025288.
78. Wang L., Gamez A., Archer H., Abola E.E., Sarkissian C.N., Fitzpatrick P., Wendt D., Zhang Y., Vellard M., Bliesath J., Bell S.M., Lemontt J.F., Sriver C.R. & Stevens R.C. (2008). – Structural and biochemical characterization of the therapeutic *Anabaena variabilis* phenylalanine ammonia lyase. *Journal of Molecular Biology*, **380** (4), 623–635. doi:10.1016/j.jmb.2008.05.025.

79. Hsieh L.S., Ma G.J., Yang C.C. & Lee P.D. (2010). – Cloning, expression, site-directed mutagenesis and immunolocalization of phenylalanine ammonia-lyase in *Bambusa oldhamii*. *Phytochemistry*, **71** (17), 1999–2009. doi:10.1016/j.phytochem.2010.09.019.
80. Urlacher V.B. & Girhard M. (2012). – Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. *Trends in Biotechnology*, **30** (1), 26–36. doi:10.1016/j.tibtech.2011.06.012.
81. Werck-Reichhart D. & Feyereisen R. (2000). – Cytochromes P450: a success story. *Genome Biology*, **1** (6, reviews 3003.1), 1–9. doi:10.1186/gb-2000-1-6-reviews3003.
82. Sono M., Roach M.P., Coulter E.D. & Dawson J.H. (1996). – Heme-containing oxygenases. *Chemical Reviews*, **96** (7), 2841–2888. doi:10.1021/cr9500500.
83. Guengerich F.P. & Munro A.W. (2013). – Unusual cytochrome P450 enzymes and reactions. *Journal of Biological Chemistry*, **288** (24), 17065–17073. doi:10.1074/jbc.R113.462275.
84. 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. doi:10.1016/j.bbagen.2006.07.017.
85. 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. doi:10.1007/s00253-010-2472-z.
86. Jenkins C.M. & Waterman M.R. (1998). – NADPH-flavodoxin reductase and flavodoxin from *Escherichia coli*: Characteristics as a soluble microsomal P450 reductase. *Biochemistry*, **37** (17), 6106–6113. doi:10.1021/bi973076p.
87. Girvan H.M. & Munro A.W. (2016). – Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology. *Current Opinion in Chemical Biology*, **31**, 136–145. doi:10.1016/j.cbpa.2016.02.018.
88. Li Z., Jiang Y., Guengerich F.P., Ma L., Li S. & Zhang W. (2020). – Engineering cytochrome P450 enzyme systems for biomedical and biotechnological applications. *Journal of Biological Chemistry*, **295** (3), 833–849. doi:10.1016/S0021-9258(17)49939-X.
89. McLean K.J. & Munro A.W. (2018). – Cytochrome P450 (cyp). . In *Encyclopedia of Signaling Molecules* (S. Choi, ed), Springer International Publishing, Cham. pp 1288–1305doi:10.1007/978-3-319-67199-4\_101615.
90. Rittle J. & Green M.T. (2010). – Cytochrome P450 compound I: Capture, characterization, and C–H bond activation kinetics. *Science*, **330** (6006), 933–937. doi:10.1126/science.1193478.
91. Munro A.W., Girvan H.M. & McLean K.J. (2007). – Variations on a (t)heme—novel mechanisms, redox partners and catalytic functions in the cytochrome P450 superfamily. *Natural Product Reports*, **24** (3), 585–609. doi:10.1039/B604190F.
92. Guengerich F.P. (2018). – Mechanisms of cytochrome P450-catalyzed oxidations. *ACS Catalysis*, **8** (12), 10964–10976. doi:10.1021/acscatal.8b03401.
93. Bell S.G., Hoskins N., Xu F., Caprotti D., Rao Z. & Wong L.L. (2006). – Cytochrome P450 enzymes from the metabolically diverse bacterium *Rhodopseudomonas palustris*. *Biochemical and Biophysical Research Communications*, **342** (1), 191–196. doi:10.1016/j.bbrc.2006.01.133.
94. Furuya T., Arai Y. & Kino K. (2012). – Biotechnological production of caffeic acid by bacterial cytochrome P450 CYP199A2. *Applied and Environmental Microbiology*, **78** (17), 6087–6094. doi:10.1128/AEM.01103-12.
95. Furuya T. & Kino K. (2010). – Genome mining approach for the discovery of novel cytochrome P450 biocatalysts. *Applied Microbiology and Biotechnology*, **86** (4), 991–1002. doi:10.1007/s00253-010-2450-5.

96. 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** (9570), 1–13. doi:10.1038/s41598-017-10075-w.
97. Prieto M.A. & Garcia J.L. (1994). – Molecular characterization of 4-hydroxyphenylacetate 3-hydroxylase of *Escherichia coli*. A two-protein component enzyme. *Journal of Biological Chemistry*, **269** (36), 22823–22829. doi:10.1016/S0021-9258(17)31719-2.
98. Xun L. & Sandvik E.R. (2000). – Characterization of 4-hydroxyphenylacetate 3-hydroxylase (HpaB) of *Escherichia coli* as a reduced flavin adenine dinucleotide-utilizing monooxygenase. *Applied and Environmental Microbiology*, **66** (2), 481–486. doi:10.1128/AEM.66.2.481-486.2000.
99. Galán B., Díaz E., Prieto M.A. & García J.L. (2000). – Functional analysis of the small component of the 4-hydroxyphenylacetate 3-monooxygenase of *Escherichia coli* W: a prototype of a new flavin:NAD(P)H reductase subfamily. *Journal of Bacteriology*, **182** (3), 627–636. doi:10.1128/JB.182.3.627-636.2000.
100. Jones J.A., Collins S.M., Vernacchio V.R., Lachance D.M. & Koffas M.A.G. (2016). – Optimization of naringenin and p-coumaric acid hydroxylation using the native *E. coli* hydroxylase complex, HpaBC. *Biotechnology Progress*, **32** (1), 21–25. doi:10.1002/btpr.2185.
101. Lin Y. & Yan Y. (2012). – Biosynthesis of caffeic acid in *Escherichia coli* using its endogenous hydroxylase complex. *Microbial Cell Factories*, **11** (42), 1–9. doi:10.1186/1475-2859-11-42.
102. Huang Q., Lin Y. & Yan Y. (2013). – Caffeic acid production enhancement by engineering a phenylalanine over-producing *Escherichia coli* strain. *Biotechnology and Bioengineering*, **110** (12), 3188–3196. doi:10.1002/bit.24988.
103. Furuya T. & Kino K. (2014). – Catalytic activity of the two-component flavin-dependent monooxygenase from *Pseudomonas aeruginosa* toward cinnamic acid derivatives. *Applied Microbiology and Biotechnology*, **98** (3), 1145–1154. doi:10.1007/s00253-013-4958-y.
104. Parvathi K., Chen F., Guo D., Blount J.W. & Dixon R.A. (2001). – Substrate preferences of O-methyltransferases in alfalfa suggest new pathways for 3-O-methylation of monolignols. *The Plant Journal*, **25** (2), 193–202. doi:10.1111/j.1365-3113X.2001.00956.x.
105. Walker A.M., Sattler S.A., Regner M., Jones J.P., Ralph J., Vermerris W., Sattler S.E. & Kang C. (2016). – The structure and catalytic mechanism of *Sorghum bicolor* caffeoyl-CoA O-methyltransferase. *Plant Physiology*, **172** (1), 78–92. doi:10.1104/pp.16.00845.
106. Naaz H., Pandey V.P., Singh S. & Dwivedi U.N. (2013). – Structure–function analyses and molecular modeling of caffeic acid-O-methyltransferase and caffeoyl-CoA-O-methyltransferase: Revisiting the basis of alternate methylation pathways during monolignol biosynthesis. *Biotechnology and Applied Biochemistry*, **60** (2), 170–189. doi:10.1002/bab.1075.
107. Hong B.S., Kim B.G., Jung N.R., Lee Y.K., Lim Y.H., Chong Y.H. & Ahn J.H. (2009). – Structural modeling and biochemical characterization of flavonoid O-methyltransferase from rice. *Bulletin of the Korean Chemical Society*, **30** (11), 2803–2805. doi:10.5012/bkcs.2009.30.11.2803.
108. Zubieta C., Kota P., Ferrer J.L., Dixon R.A. & Noel J.P. (2002). – Structural basis for the modulation of lignin monomer methylation by caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase. *The Plant Cell*, **14** (6), 1265–1277. doi:10.1105/tpc.001412.
109. Hoffmann L., Maury S., Bergdoll M., Thion L., Erard M. & Legrand M. (2001). – Identification of the enzymatic active site of tobacco caffeoyl-coenzyme A O-methyltransferase by site-directed mutagenesis. *Journal of Biological Chemistry*, **276** (39), 36831–36838. doi:10.1074/jbc.M104977200.
110. Huotari M., Gogos J.A., Karayiorgou M., Koponen O., Forsberg M., Raasmaja A., Hyttinen J. & Männistö P.T. (2002). – Brain catecholamine metabolism in catechol-O-methyltransferase (COMT)-deficient mice. *European Journal of Neuroscience*, **15** (2), 246–256. doi:10.1046/j.0953-816x.2001.01856.x.

111. Wiens B. & De Luca V. (2016). – Molecular and biochemical characterization of a benzenoid/phenylpropanoid meta/para-O-methyltransferase from *Rauwolfia serpentina* roots. *Phytochemistry*, **132**, 5–15. doi:10.1016/j.phytochem.2016.10.004.
112. Green A.R., Lewis K.M., Barr J.T., Jones J.P., Lu F., Ralph J., Vermerris W., Sattler S.E. & Kang C. (2014). – Determination of the structure and catalytic mechanism of *Sorghum bicolor* caffeic acid O-methyltransferase and the structural impact of three brown midrib12 mutations. *Plant Physiology*, **165** (4), 1440–1456. doi:10.1104/pp.114.241729.
113. Rossmann M.G., Moras D. & Olsen K.W. (1974). – Chemical and biological evolution of a nucleotide-binding protein. *Nature*, **250** (5463), 194–199. doi:10.1038/250194a0.
114. Ferrer J.L., Zubieta C., Dixon R.A. & Noel J.P. (2005). – Crystal structures of alfalfa caffeoyl coenzyme A 3-O-methyltransferase. *Plant Physiology*, **137** (3), 1009–1017. doi:10.1104/pp.104.048751.
115. Kopycki J.G., Rauh D., Chumanevich A.A., Neumann P., Vogt T. & Stubbs M.T. (2008). – Biochemical and structural analysis of substrate Promiscuity in plant Mg<sup>2+</sup>-dependent O-methyltransferases. *Journal of Molecular Biology*, **378** (1), 154–164. doi:10.1016/j.jmb.2008.02.019.
116. Li L., Popko J.L., Umezawa T. & Chiang V.L. (2000). – 5-hydroxyconiferyl aldehyde modulates enzymatic methylation for syringyl monolignol formation, a new view of monolignol biosynthesis in angiosperms. *Journal of Biological Chemistry*, **275** (9), 6537–6545. doi:10.1074/jbc.275.9.6537.
117. Nakatsubo T., Kitamura Y., Sakakibara N., Mizutani M., Hattori T., Sakurai N., Shibata D., Suzuki S. & Umezawa T. (2008). – At5g54160 gene encodes *Arabidopsis thaliana* 5-hydroxyconiferaldehyde O-methyltransferase. *Journal of Wood Science*, **54** (4), 312–317. doi:10.1007/s10086-008-0958-4.
118. Ma Q.H. & Xu Y. (2008). – Characterization of a caffeic acid 3-O-methyltransferase from wheat and its function in lignin biosynthesis. *Biochimie*, **90** (3), 515–524. doi:10.1016/j.biochi.2007.09.016.
119. Koshiba T., Hirose N., Mukai M., Yamamura M., Hattori T., Suzuki S., Sakamoto M. & Umezawa T. (2013). – Characterization of 5-hydroxyconiferaldehyde O-methyltransferase in *Oryza sativa*. *Plant Biotechnology*, **30** (2), 157–167. doi:10.5511/plantbiotechnology.13.0219a.
120. Guo D., Chen F., Inoue K., Blount J.W. & Dixon R.A. (2001). – Downregulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa: Impacts on lignin structure and implications for the biosynthesis of G and S lignin. *The Plant Cell*, **13** (1), 73–88. doi:10.1105/tpc.13.1.73.
121. Atanassova R., Favet N., Martz F., Chabbert B., Tollier M.T., Monties B., Fritig B. & Legrand M. (1995). – Altered lignin composition in transgenic tobacco expressing O-methyltransferase sequences in sense and antisense orientation. *The Plant Journal*, **8** (4), 465–477. doi:10.1046/j.1365-313X.1995.8040465.x.
122. Van Doorselaere J., Baucher M., Chognot E., Chabbert B., Tollier M.T., Petit-Conil M., Leplé J.C., Pilate G., Cornu D., Monties B., Van Montagu M., Inzé D., Boerjan W. & Jouanin L. (1995). – A novel lignin in poplar trees with a reduced caffeic acid/5-hydroxyferulic acid O-methyltransferase activity. *The Plant Journal*, **8** (6), 855–864. doi:10.1046/j.1365-313X.1995.8060855.x.
123. Pinçon G., Maury S., Hoffmann L., Geoffroy P., Lapierre C., Pollet B. & Legrand M. (2001). – Repression of O-methyltransferase genes in transgenic tobacco affects lignin synthesis and plant growth. *Phytochemistry*, **57** (7), 1167–1176. doi:10.1016/S0031-9422(01)00098-X.
124. Do C.T., Pollet B., Thévenin J., Sibout R., Denoue D., Barrière Y., Lapierre C. & Jouanin L. (2007). – Both caffeoyl Coenzyme A 3-O-methyltransferase 1 and caffeic acid O-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in *Arabidopsis*. *Planta*, **226** (5), 1117–1129. doi:10.1007/s00425-007-0558-3.

125. Byeon Y., Lee H.Y., Lee K. & Back K. (2014). – Caffeic acid O-methyltransferase is involved in the synthesis of melatonin by methylating N-acetylserotonin in *Arabidopsis*. *Journal of Pineal Research*, **57** (2), 219–227. doi:10.1111/jpi.12160.
126. Muzac I., Wang J., Anzellotti D., Zhang H. & Ibrahim R.K. (2000). – Functional expression of an *Arabidopsis* cDNA clone encoding a flavonol 3'-O-methyltransferase and characterization of the gene product. *Archives of Biochemistry and Biophysics*, **375** (2), 385–388. doi:10.1006/abbi.1999.1681.
127. Heo K.T., Kang S.Y. & Hong Y.S. (2017). – De novo biosynthesis of pterostilbene in an *Escherichia coli* strain using a new resveratrol O-methyltransferase from *Arabidopsis*. *Microbial Cell Factories*, **16** (30), 1–8. doi:10.1186/s12934-017-0644-6.
128. Choi O., Wu C.Z., Kang S.Y., Ahn J.S., Uhm T.B. & Hong Y.S. (2011). – Biosynthesis of plant-specific phenylpropanoids by construction of an artificial biosynthetic pathway in *Escherichia coli*. *Journal of Industrial Microbiology and Biotechnology*, **38** (10), 1657–1665. doi:10.1007/s10295-011-0954-3.
129. Kang S.Y., Choi O., Lee J.K., Hwang B.Y., Uhm T.B. & Hong Y.S. (2012). – Artificial biosynthesis of phenylpropanoic acids in a tyrosine overproducing *Escherichia coli* strain. *Microbial Cell Factories*, **11** (153), 1–9. doi:10.1186/1475-2859-11-153.
130. Ni J., Tao F., Du H. & Xu P. (2015). – Mimicking a natural pathway for *de novo* biosynthesis: natural vanillin production from accessible carbon sources. *Scientific Reports*, **5** (13670), 1–12. doi:10.1038/srep13670.
131. Rodrigues J.L., Gomes D. & Rodrigues L.R. (2020). – A combinatorial approach to optimize the production of curcuminoids from tyrosine in *Escherichia coli*. *Frontiers in Bioengineering and Biotechnology*, **8** (59), 1–15. doi:10.3389/fbioe.2020.00059.
132. Rodrigues J.L., Araújo R.G., Prather K.L.J., Kluskens L.D. & Rodrigues L.R. (2015). – Heterologous production of caffeic acid from tyrosine in *Escherichia coli*. *Enzyme and Microbial Technology*, **71**, 36–44. doi:10.1016/j.enzmictec.2015.01.001.
133. Jones J.A., Vernacchio V.R., Collins S.M., Shirke A.N., Xiu Y., Englaender J.A., Cress B.F., McCutcheon C.C., Linhardt R.J., Gross R.A. & Koffas M.A.G. (2017). – Complete biosynthesis of anthocyanins using *E. coli* polycultures. *mBio*, **8** (3), 1–9. doi:10.1128/mBio.00621-17.
134. Li Y., Li J., Qian B., Cheng L., Xu S. & Wang R. (2018). – *De novo* biosynthesis of p-coumaric acid in *E. coli* with a trans-cinnamic acid 4-hydroxylase from the Amaryllidaceae plant *Lycoris aurea*. *Molecules*, **23** (12), 1–19. doi:10.3390/molecules23123185.
135. Jansen F., Gillesen B., Mueller F., Commandeur U., Fischer R. & Kreuzaler F. (2014). – Metabolic engineering for p-coumaryl alcohol production in *Escherichia coli* by introducing an artificial phenylpropanoid pathway. *Biotechnology and Applied Biochemistry*, **61** (6), 646–654. doi:10.1002/bab.1222.
136. Summeren-Wesenhagen P.V. van, Voges R., Dennig A., Sokolowsky S., Noack S., Schwaneberg U. & Marienhagen J. (2015). – Combinatorial optimization of synthetic operons for the microbial production of p-coumaryl alcohol with *Escherichia coli*. *Microbial Cell Factories*, **14** (79), 1–10. doi:10.1186/s12934-015-0274-9.
137. Chen Z., Sun X., Li Y., Yan Y. & Yuan Q. (2017). – Metabolic engineering of *Escherichia coli* for microbial synthesis of monolignols. *Metabolic Engineering*, **39**, 102–109. doi:10.1016/j.ymben.2016.10.021.
138. Yang J., Liang J., Shao L., Liu L., Gao K., Zhang J.L., Sun Z., Xu W., Lin P., Yu R. & Zi J. (2020). – Green production of silybin and isosilybin by merging metabolic engineering approaches and enzymatic catalysis. *Metabolic Engineering*, **59**, 44–52. doi:10.1016/j.ymben.2020.01.007.
139. Lin Y., Sun X., Yuan Q. & Yan Y. (2013). – Combinatorial biosynthesis of plant-specific coumarins in bacteria. *Metabolic Engineering*, **18**, 69–77. doi:10.1016/j.ymben.2013.04.004.

140. Santos C.N.S., Koffas M. & Stephanopoulos G. (2011). – Optimization of a heterologous pathway for the production of flavonoids from glucose. *Metabolic Engineering*, **13** (4), 392–400. doi:10.1016/j.ymben.2011.02.002.
141. Xue Z., McCluskey M., Cantera K., Ben-Bassat A., Sariaslani F.S. & Huang L. (2007). – Improved production of p-hydroxycinnamic acid from tyrosine using a novel thermostable phenylalanine/tyrosine ammonia lyase enzyme. *Enzyme and Microbial Technology*, **42** (1), 58–64. doi:10.1016/j.enzmictec.2007.07.025.
142. 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. doi:10.1016/j.ymben.2018.04.008.
143. Zhou J., Liu L., Shi Z., Du G. & Chen J. (2009). – ATP in current biotechnology: Regulation, applications and perspectives. *Biotechnology Advances*, **27** (1), 94–101. doi:10.1016/j.biotechadv.2008.10.005.
144. Rogl H., Kosemund K., Kühlbrandt W. & Collinson I. (1998). – Refolding of *Escherichia coli* produced membrane protein inclusion bodies immobilised by nickel chelating chromatography. *FEBS Letters*, **432** (1), 21–26. doi:10.1016/S0014-5793(98)00825-4.
145. Baneyx F. & Mujacic M. (2004). – Recombinant protein folding and misfolding in *Escherichia coli*. *Nature Biotechnology*, **22** (11), 1399–1408. doi:10.1038/nbt1029.
146. Kadisch M., Willrodt C., Hillen M., Bühler B. & Schmid A. (2017). – Maximizing the stability of metabolic engineering-derived whole-cell biocatalysts. *Biotechnology Journal*, **12** (8), 1–27. doi:10.1002/biot.201600170.
147. deCarvalho C.C.C.R. (2016). – Whole-cell biocatalysts: essential workers from nature to the industry. *Microbial Biotechnology*, **10** (2), 250–263. doi:10.1111/1751-7915.12363.
148. Hitchcock D.I. (1924). – The solubility of tyrosine in acid and in alkali. *Journal of General Physiology*, **6** (6), 747–757. doi:10.1085/jgp.6.6.747.
149. Prices of L-tyrosine (T207.3) and L-phenylalanine (4491.2) (Carl Roth webpage) Available at: <https://www.carlroth.com> (accessed on 1 June 2021).
150. Mahmood T. & Yang P.C. (2012). – Western Blot: Technique, theory, and trouble shooting. *North American Journal of Medical Sciences*, **4** (9), 429–434. doi:10.4103/1947-2714.100998.
151. Ghosh R., Gilda J.E. & Gomes A.V. (2014). – The necessity of and strategies for improving confidence in the accuracy of western blots. *Expert Review of Proteomics*, **11** (5), 549–560. doi:10.1586/14789450.2014.939635.
152. Xu P., Vansiri A., Bhan N. & Koffas M.A.G. (2012). – ePathBrick: A synthetic biology platform for engineering metabolic pathways in *E. coli*. *ACS Synthetic Biology*, **1** (7), 256–266. doi:10.1021/sb300016b.
153. He W., Fu L., Li G., Jones J.A., Linhardt R.J. & Koffas M. (2015). – Production of chondroitin in metabolically engineered *E. coli*. *Metabolic Engineering*, **27**, 92–100. doi:10.1016/j.ymben.2014.11.003.
154. Okuda S., Kawashima S., Kobayashi K., Ogasawara N., Kanehisa M. & Goto S. (2007). – Characterization of relationships between transcriptional units and operon structures in *Bacillus subtilis* and *Escherichia coli*. *BMC Genomics*, **8** (48), 1–12. doi:10.1186/1471-2164-8-48.
155. Lawson R.J., vonWachenfeldt C., Haq I., Perkins J. & Munro A.W. (2004). – Expression and characterization of the two flavodoxin proteins of *Bacillus subtilis*, YkuN and YkuP: Biophysical properties and interactions with cytochrome P450 BioI. *Biochemistry*, **43** (39), 12390–12409. doi:10.1021/bi049131t.
156. vonBühler 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. doi:10.1002/cbic.201300271.



157. Haslinger K. & Prather K.L.J. (2020). – Heterologous caffeic acid biosynthesis in *Escherichia coli* is affected by choice of tyrosine ammonia lyase and redox partners for bacterial cytochrome P450. *Microbial Cell Factories*, **19** (26), 1–12. doi:10.1186/s12934-020-01300-9.
158. Bell S.G., Xu F., Forward I., Bartlam M., Rao Z. & Wong L.L. (2008). – Crystal structure of CYP199A2, a para-substituted benzoic acid oxidizing cytochrome P450 from *Rhodopseudomonas palustris*. *Journal of Molecular Biology*, **383** (3), 561–574. doi:10.1016/j.jmb.2008.08.033.
159. Aoki M., Ishimori K., Fukada H., Takahashi K. & Morishima I. (1998). – Isothermal titration calorimetric studies on the associations of putidaredoxin to NADH-putidaredoxin reductase and P450cam. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, **1384** (1), 180–188. doi:10.1016/S0167-4838(98)00017-X.
160. Johnson E.O.D. & Wong L.L. (2016). – Partial fusion of a cytochrome P450 system by carboxy-terminal attachment of putidaredoxin reductase to P450cam (CYP101A1). *Catalysis Science & Technology*, **6** (20), 7549–7560. doi:10.1039/C6CY01042C.
161. Zhang B., Lewis K.M., Abril A., Davydov D.R., Vermerris W., Sattler S.E. & Kang C. (2020). – Structure and function of the cytochrome P450 monooxygenase cinnamate 4-hydroxylase from *Sorghum bicolor*. *Plant Physiology*, **183** (3), 957–973. doi:10.1104/pp.20.00406.
162. Jacob F.F., Hutzler M. & Methner F.J. (2018). – Comparison of various industrially applicable disruption methods to produce yeast extract using spent yeast from top-fermenting beer production: Influence on amino acid and protein content. *European Food Research and Technology*, **245** (1), 95–109. doi:10.1007/s00217-018-3143-z.
163. Podpora B., Świdorski F., Sadowska A., Rakowska R. & Wasiak-Zys G. (2016). – Spent brewer's yeast extracts as a new component of functional food. *Czech Journal of Food Sciences*, **34** (6), 554–563. doi:10.17221/419/2015-CJFS.
164. Tryptone, product info (Khimexpert webpage) Available at: <https://khimexpert.com/wp-content/uploads/2018/12/GCM23-Tryptone.pdf> (accessed on 25 October 2021).
165. Tryptone USP, technical data sheet (Labchem Wako webpage) Available at: [https://labchem-wako.fujifilm.com/jp/product\\_data/docs/03700467\\_doc02.pdf](https://labchem-wako.fujifilm.com/jp/product_data/docs/03700467_doc02.pdf) (accessed on 25 October 2021).
166. T8750 Tryptone, specifications (USBiological Life Sciences webpage) Available at: <https://www.usbio.net/media/T8750/tryptone-bse-free> (accessed on 25 October 2021).
167. Li Z., Wang X. & Zhang H. (2019). – Balancing the non-linear rosmarinic acid biosynthetic pathway by modular co-culture engineering. *Metabolic Engineering*, **54**, 1–11. doi:10.1016/j.ymben.2019.03.002.
168. Dornheim M. (2021). – *Untersuchungen zur biotechnologischen Produktion von Ferula- und Sinapinsäure*. Dissertation, Martin Luther University of Halle-Wittenberg.
169. Tolia N.H. & Joshua-Tor L. (2006). – Strategies for protein coexpression in *Escherichia coli*. *Nature Methods*, **3** (1), 55–64. doi:10.1038/nmeth0106-55.
170. Satoh Y., Tajima K., Munekata M., Keasling J.D. & Lee T.S. (2012). – Engineering of L-tyrosine oxidation in *Escherichia coli* and microbial production of hydroxytyrosol. *Metabolic Engineering*, **14** (6), 603–610. doi:10.1016/j.ymben.2012.08.002.
171. Santos C.N.S. & Stephanopoulos G. (2008). – Melanin-based high-throughput screen for L-Tyrosine production in *Escherichia coli*. *Applied and Environmental Microbiology*, **74** (4), 1190–1197. doi:10.1128/AEM.02448-07.
172. Muñoz A.J., Hernández-Chávez G., deAnda R., Martínez A., Bolívar F. & Gosset G. (2011). – Metabolic engineering of *Escherichia coli* for improving L-3,4-dihydroxyphenylalanine (L-DOPA) synthesis from glucose. *Journal of Industrial Microbiology and Biotechnology*, **38** (11), 1845–1852. doi:10.1007/s10295-011-0973-0.
173. Liu S., Liu J., Hou J., Chao N., Gai Y. & Jiang X. (2017). – Three steps in one pot: biosynthesis of 4-hydroxycinnamyl alcohols using immobilized whole-cells of two genetically engineered

- Escherichia coli* strains. *Microbial Cell Factories*, **16** (104), 1–12. doi:10.1186/s12934-017-0722-9.
174. Aschenbrenner J., Marx P., Pietruszka J. & Marienhagen J. (2019). – Microbial production of natural and unnatural monolignols with *Escherichia coli*. *ChemBioChem*, **20** (7), 949–954. doi:10.1002/cbic.201800673.
175. Atsumi S., Wu T.Y., Eckl E.M., Hawkins S.D., Buelter T. & Liao J.C. (2010). – Engineering the isobutanol biosynthetic pathway in *Escherichia coli* by comparison of three aldehyde reductase/alcohol dehydrogenase genes. *Applied Microbiology and Biotechnology*, **85** (3), 651–657. doi:10.1007/s00253-009-2085-6.
176. Pérez J.M., Arenas F.A., Pradenas G.A., Sandoval J.M. & Vásquez C.C. (2008). – *Escherichia coli* YqhD exhibits aldehyde reductase activity and protects from the harmful effect of lipid peroxidation-derived aldehydes. *Journal of Biological Chemistry*, **283** (12), 7346–7353. doi:10.1074/jbc.M708846200.
177. Visvalingam J., Hernandez-Doria J.D. & Holley R.A. (2013). – Examination of the genome-wide transcriptional response of *Escherichia coli* O157:H7 to cinnamaldehyde exposure. *Applied and Environmental Microbiology*, **79** (3), 942–950. doi:10.1128/AEM.02767-12.
178. Klumbys E., Zebec Z., Weise N.J., Turner N.J. & Scrutton N.S. (2018). – Bio-derived production of cinnamyl alcohol via a three step biocatalytic cascade and metabolic engineering. *Green Chemistry*, **20** (3), 658–663. doi:10.1039/C7GC03325G.
179. Decembrino D., Girhard M. & Urlacher V.B. (2021). – Use of copper as a trigger for the *in vivo* activity of *E. coli* laccase CueO: A simple tool for biosynthetic purposes. *ChemBioChem*, **22** (8), 1470–1479. doi:10.1002/cbic.202000775.
180. Müsgens T. (2021). – *Sinapyl alcohol production by heterologous expression of plant genes in Escherichia coli*. Master thesis, RWTH Aachen University.
181. Costa M.A., Bedgar D.L., Moinuddin S.G.A., Kim K.W., Cardenas C.L., Cochrane F.C., Shockey J.M., Helms G.L., Amakura Y., Takahashi H., Milhollan J.K., Davin L.B., Browse J. & Lewis N.G. (2005). – Characterization *in vitro* and *in vivo* of the putative multigene 4-coumarate:CoA ligase network in *Arabidopsis*: syringyl lignin and sinapate/sinapyl alcohol derivative formation. *Phytochemistry*, **66** (17), 2072–2091. doi:10.1016/j.phytochem.2005.06.022.
182. Lozoya E., Hoffmann H., Douglas C., Schulz W., Scheel D. & Hahlbrock K. (1988). – Primary structures and catalytic properties of isoenzymes encoded by the two 4-coumarate: CoA ligase genes in parsley. *European Journal of Biochemistry*, **176** (3), 661–667. doi:10.1111/j.1432-1033.1988.tb14328.x.
183. Zhao S., Jones J.A., Lachance D.M., Bhan N., Khalidi O., Venkataraman S., Wang Z. & Koffas M.A.G. (2015). – Improvement of catechin production in *Escherichia coli* through combinatorial metabolic engineering. *Metabolic Engineering*, **28**, 43–53. doi:10.1016/j.ymben.2014.12.002.
184. Jones J.A., Vernacchio V.R., Sinkoe A.L., Collins S.M., Ibrahim M.H.A., Lachance D.M., Hahn J. & Koffas M.A.G. (2016). – Experimental and computational optimization of an *Escherichia coli* co-culture for the efficient production of flavonoids. *Metabolic Engineering*, **35**, 55–63. doi:10.1016/j.ymben.2016.01.006.
185. Bentley W.E., Mirjalili N., Andersen D.C., Davis R.H. & Kompala D.S. (1990). – Plasmid-encoded protein: The principal factor in the “metabolic burden” associated with recombinant bacteria. *Biotechnology and Bioengineering*, **35** (7), 668–681. doi:10.1002/bit.260350704.
186. Ma D., Cook D.N., Alberti M., Pon N.G., Nikaido H. & Hearst J.E. (1995). – Genes *acrA* and *acrB* encode a stress-induced efflux system of *Escherichia coli*. *Molecular Microbiology*, **16** (1), 45–55. doi:10.1111/j.1365-2958.1995.tb02390.x.
187. Kunjapur A.M., Tarasova Y. & Prather K.L.J. (2014). – Synthesis and accumulation of aromatic aldehydes in an engineered strain of *Escherichia coli*. *Journal of the American Chemical Society*, **136** (33), 11644–11654. doi:10.1021/ja506664a.

188. Gilleßen B. (2009). – *Rekombinante Biosynthese amino-substituierter Phenylpropanoide in E. coli*. Dissertation, RWTH Aachen University.
189. Kang Z., Wang Y., Gu P., Wang Q. & Qi Q. (2011). – Engineering *Escherichia coli* for efficient production of 5-aminolevulinic acid from glucose. *Metabolic Engineering*, **13** (5), 492–498. doi:10.1016/j.ymben.2011.05.003.
190. Bell S.G., Xu F., Johnson E.O.D., Forward I.M., Bartlam M., Rao Z. & Wong L.L. (2010). – Protein recognition in ferredoxin–P450 electron transfer in the class I CYP199A2 system from *Rhodospseudomonas palustris*. *Journal of Biological Inorganic Chemistry*, **15** (3), 315–328. doi:10.1007/s00775-009-0604-7.
191. Lewis D.F.V. & Hlavica P. (2000). – Interactions between redox partners in various cytochrome P450 systems: Functional and structural aspects. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, **1460** (2), 353–374. doi:10.1016/S0005-2728(00)00202-4.
192. Hollingsworth S.A., Batabyal D., Nguyen B.D. & Poulos T.L. (2016). – Conformational selectivity in cytochrome P450 redox partner interactions. *Proceedings of the National Academy of Sciences*, **113** (31), 8723–8728. doi:10.1073/pnas.1606474113.
193. Zhou S., Liu P., Chen J., Du G., Li H. & Zhou J. (2016). – Characterization of mutants of a tyrosine ammonia-lyase from *Rhodotorula glutinis*. *Applied Microbiology and Biotechnology*, **100** (24), 10443–10452. doi:10.1007/s00253-016-7672-8.
194. Vargas-Tah A., Martínez L.M., Hernández-Chávez G., Rocha M., Martínez A., Bolívar F. & Gosset G. (2015). – Production of cinnamic and p-hydroxycinnamic acid from sugar mixtures with engineered *Escherichia coli*. *Microbial Cell Factories*, **14** (6), 1–12. doi:10.1186/s12934-014-0185-1.
195. Bang H.B., Lee K., Lee Y.J. & Jeong K.J. (2018). – High-level production of trans-cinnamic acid by fed-batch cultivation of *Escherichia coli*. *Process Biochemistry*, **68**, 30–36. doi:10.1016/j.procbio.2018.01.026.
196. Kang S.Y., Choi O., Lee J.K., Ahn J.O., Ahn J.S., Hwang B.Y. & Hong Y.S. (2015). – Artificial de novo biosynthesis of hydroxystyrene derivatives in a tyrosine overproducing *Escherichia coli* strain. *Microbial Cell Factories*, **14** (78), 1–11. doi:10.1186/s12934-015-0268-7.
197. Kunjapur A.M., Hyun J.C. & Prather K.L.J. (2016). – Dereglulation of S-adenosylmethionine biosynthesis and regeneration improves methylation in the *E. coli* de novo vanillin biosynthesis pathway. *Microbial Cell Factories*, **15** (61), 1–17. doi:10.1186/s12934-016-0459-x.
198. Coward J.K., D'Urso-Scott M. & Sweet W.D. (1972). – Inhibition of catechol-O-methyltransferase by S-adenosylhomocysteine and S-adenosylhomocysteine sulfoxide, a potential transition-state analog. *Biochemical Pharmacology*, **21** (8), 1200–1203. doi:10.1016/0006-2952(72)90114-1.
199. Coward J.K., Slisz E.P. & Wu F.Y.H. (1973). – Kinetic studies in catechol O-methyltransferase. Product inhibition and the nature of the catechol binding site. *Biochemistry*, **12** (12), 2291–2297. doi:10.1021/bi00736a017.
200. Jang S., Gang H., Kim B.G. & Choi K.Y. (2018). – FCS and ECH dependent production of phenolic aldehyde and melanin pigment from L-tyrosine in *Escherichia coli*. *Enzyme and Microbial Technology*, **112**, 59–64. doi:10.1016/j.enzmictec.2017.10.011.
201. Ricklefs E., Girhard M. & Urlacher V.B. (2016). – Three-steps in one-pot: whole-cell biocatalytic synthesis of enantiopure (+)- and (–)-pinorexinol via kinetic resolution. *Microbial Cell Factories*, **15** (78), 1–11. doi:10.1186/s12934-016-0472-0.
202. Halls S.C., Davin L.B., Kramer D.M. & Lewis N.G. (2004). – Kinetic study of coniferyl alcohol radical binding to the (+)-pinorexinol forming dirigent protein. *Biochemistry*, **43** (9), 2587–2595. doi:10.1021/bi035959o.

203. Lv Y., Cheng X., Du G., Zhou J. & Chen J. (2017). – Engineering of an H<sub>2</sub>O<sub>2</sub> auto-scavenging in vivo cascade for pinoselin production. *Biotechnology and Bioengineering*, **114** (9), 2066–2074. doi:10.1002/bit.26319.
204. Wilks J.C. & Slonczewski J.L. (2007). – pH of the cytoplasm and periplasm of *Escherichia coli*: Rapid measurement by green fluorescent protein fluorimetry. *Journal of Bacteriology*, **189** (15), 5601–5607. doi:10.1128/JB.00615-07.
205. Chung H.J., Bang W. & Drake M.A. (2006). – Stress response of *Escherichia coli*. *Comprehensive Reviews in Food Science and Food Safety*, **5** (3), 52–64. doi:10.1111/j.1541-4337.2006.00002.x.
206. Chen L., Auh C.K., Dowling P., Bell J., Lehmann D. & Wang Z.Y. (2004). – Transgenic down-regulation of caffeic acid O-methyltransferase (COMT) led to improved digestibility in tall fescue (*Festuca arundinacea*). *Functional Plant Biology*, **31** (3), 235–245. doi:10.1071/fp03254.
207. Li K. & Frost J.W. (1998). – Synthesis of vanillin from glucose. *Journal of the American Chemical Society*, **120** (40), 10545–10546. doi:10.1021/ja9817747.
208. Bang H.B., Lee Y.H., Kim S.C., Sung C.K. & Jeong K.J. (2016). – Metabolic engineering of *Escherichia coli* for the production of cinnamaldehyde. *Microbial Cell Factories*, **15** (16), 1–12. doi:10.1186/s12934-016-0415-9.
209. Ricklefs E., Girhard M., Koschorreck K., Smit M.S. & Urlacher V.B. (2015). – Two-step one-pot synthesis of pinoselin from eugenol in an enzymatic cascade. *ChemCatChem*, **7** (12), 1857–1864. doi:10.1002/cctc.201500182.
210. Meyer A.J., Segall-Shapiro T.H., Glassey E., Zhang J. & Voigt C.A. (2019). – *Escherichia coli* “Marionette” strains with 12 highly optimized small-molecule sensors. *Nature Chemical Biology*, **15** (2), 196–204. doi:10.1038/s41589-018-0168-3.
211. Harwood C.R. & Cutting S.M. (1990). – *Molecular biological methods for Bacillus*. Wiley, Chichester; New York.
212. Sambrook J.F. & Russell D.W. (2001). – *Molecular Cloning: A Laboratory Manual*. 3rd ed., Cold Spring Harbor Laboratory Press.
213. Xu P. & Koffas M.A.G. (2013). – Assembly of Multi-gene Pathways and Combinatorial Pathway Libraries Through ePathBrick Vectors. . In *Synthetic Biology* (K.M. Polizzi & C. Kontoravdi, eds), Humana Press, Totowa, NJ. pp 107–129 doi:10.1007/978-1-62703-625-2\_10.
214. Li C., Wen A., Shen B., Lu J., Huang Y. & Chang Y. (2011). – FastCloning: a highly simplified, purification-free, sequence- and ligation-independent PCR cloning method. *BMC Biotechnology*, **11** (92), 1–10. doi:10.1186/1472-6750-11-92.
215. Gallagher S.R. (2004). – Quantitation of DNA and RNA with absorption and fluorescence spectroscopy. *Current Protocols in Molecular Biology*, **66** (1), A.3D.1–A.3D.12. doi:10.1002/0471142727.mba03ds66.

## VII. Appendix

### VII.1 Used gene sequences

**Table 15: DNA sequences of cloned genes.** All gene sequences were codon-optimized for expression in *E. coli*. The 5'- and 3'-ends are indicated in bold as well as start and stop codons. DNA bases colored in orange indicate the N-terminal deletion sequences (NΔ7-CYP199A2).

<i>AtCCoAOMT</i>
<p>5'- <b>ATGGCAA</b>AAGACGAAGCTAAAGGCCTGTTGAAGAGCGAGGAAGTGTACAAGTACATTCTTG  AAACCAGCGTTTATCCGCGTGAACCGGAAGTGCTGCGTGAATTACGCAATATCACACACAATCA  TCCGCAGGCGGGCATGGCCACGGCGCCTGATGCAGGTCAACTTATGGGCATGTTATTGAACTTG  GTGAACGCCCCGAAAACGATTGAAGTGGGTGTATTTACCGGCTATTCACCTCTGCTTACGGCGCT  GACCTTGCCTGAGGACGGAAGAGTCATTGCAATCGATATGAACCGCGATAGCTACGAGATTGGC  CTGCCAGTCATTAAGAAGGCCGGGGTCGAGCATAAGATTGACTTTAAGGAAAGTGAGGCCCTGC  CGGCGCTGGATGAAGTGTGAATAACAAGGTAAACGAAGGCGGCTTTGATTTTGCCTTCGTTGA  CGCCGATAAACTTAATACTGGAATTACCACGAACGCCTTATCCGTCTGATTAAAGTCGGTGGCA  TTATTGTATACGACAATACATTATGGGGTGGCTCGGTGGCCGAGCCCGATTCAAGCACCCAGCA  ATGGCGCATTGAGGTTAAGAAGGCGACTCTGGAAGTTAATAAGAAGCTGTCAGCGGACCAACGC  GTTCAAATTAGCCAGGCGGCCTTAGGGGACGGGATTACTATCTGTCGTCGCTGTACTAA -3'</p>
<i>AtCOMT</i>
<p>5'- <b>ATGGGGAGC</b>ACCGCAGAACTCAGCTTACCCCTGTTTCAGGTAACCGATGATGAGGCGGCACT  GTTTGCAATGCAGCTGGCAAGCGCCTCTGTCTGCTGCTATGGCCTTGAAGTCGGCGCTGGAAGTGG  ACCTGCTGGAAATCATGGCGAAGAACGGCTCACCGATGAGCCCCACTGAAATTGCGTCAAAGCT  GCCTACAAAGAACCCAGAGGCGCCCGTTATGCTGGACCGCATTTTGCGCCTGCTGACCAGCTATT  CGGTGCTGACGTGTTCTAATCGCAAGCTGTCGGGAGATGGTGTAGAGCGTATCTATGGTCTGGGC  CCAGTGTGTAAATACCTCACTAAGAATGAGGACGGCGTATCGATTGCAGCCTTATGCCTGATGA  ATCAGGATAAAGTACTGATGGAGTCGTGGTATCACCTTAAAGACGCTATCTTAGACGGCGGCAT  CCCTTTTAATAAAGCCTACGGCATGTCCGCCTTTGAATATCATGGCACAGACCCCGCTTTAATA  AAGTATTCAATAACGGCATGAGCAATCACTCTACCATTACGATGAAGAAAATCTTAGAAACGTA  CAAAGGCTTCGAGGGTCTGACAAGTTTAGTGGACGTGGGCGGCGGCATCGGCGCGACCCCTGAAG  ATGATTGTTTCGAAATATCCGAACTTAAAGGGGATTAATTTTCGACCTGCCGCACGTGATTGAGGA  CGCGCCATCACACCCGGGAATCGAACACGTCCGGCGGCGACATGTTCTGTGTCAGTTCCGAAGGGT  GACGCGATTTTCATGAAATGGATTTGCCACGATTGGAGCGATGAGCACTGTGTAAAGTTTCTGAA  GAATTGTTATGAATCGCTCCCGGAAGATGGTAAGGTCATTCTGGCGGAATGCATTCTGCCGGAA  ACGCCGGATAGTTCAATTATCGACTAAGCAGGTGGTGCACGTAGACTGTATCATGTTAGCGCATA  ACCCGGGCGGAAAGGAGCGCACTGAAAAGGAGTTTCAAGCGCTGGCGAAGGCATCCGGTTTTTA  AAGGGATTAAGGTCGTGTGTGATGCCTTCGGCGTCAATCTGATTGAACTGCTGAAGAAATTATG  A -3'</p>
<i>CYP199A2</i>
<p>5'- <b>ATGACCACAGCACCAGCCTTGT</b>CCGGTACGACTCCGAGTCAGCATGGAGCGGGTGTCC  CGCATCTCGGTATCGATCCGTTGCCCTGGATTACTTTGCAGATCCTTATCCGGAGCAGGAAACC  CTGCGTGAAGCGGGCCAGTGGTTTACCTGGACAAGTGGAACGTTTATGGCGTTGCACGTTATGC  GGAAGTGTACGCCGTCTTGAACGATCCGCTGACCTTCTGCTCAAGCCGGGGTGTGGGTTAAGCG  ACTTCAAGAAGGAGAAACCTTGGCGTCCACCGAGCCTGATTCTGGAAGCCGACCCTCCCGCGCA  CACGCGCACTCGGGCCGTGCTGTGCAAGTCTGTCCCCAGCGACAATGAAACGCCTGCGCGAT  GGTTTCGCTGCGGCTGCAGACGCCAAAATCGACGAATTGCTCGCCCGTGGTGGCAACATTGATG  CGATTGCAGACCTTGCTGAGGCGTATCCACTGAGCGTGTTCGGATGCGATGGGCCTCAAACA  GGAAGGTGCGGAAAACCTTGCTCCCTTATGCTGGCCTGGTACTCAATGCGTTTGGCCCTCCCAACG  AACTGCGCCAATCAGCGATTGAGCGTTCTGCTCCGCATCAGGCATACGTTGCCGAACAGTGCCA  ACGCCCGAATTTGGCACCAAGGTGGGTTTGGCGCTTGCAATCACGCGTTCAGCGATACCGGAGAA  ATCACGCCAGAAGAAGCGCCGCTGTTAGTACGCAGTCTGTTGAGTGCAGGCCTGGATACCACGG</p>

TGAATGGCATTGCCGACGCCGTGTACTGTCTGGCTCGCTTTCCGGACGAATTTGCTCGTTTACGT  
GCCGATCCGTCGTTGGCACGCAATGCGTTTGAGGAAGCCGTACGCTTTGAGTCCCCGGTTCAGAC  
CTTCTTTTCGCACTACCACTCGCGATGTCGAACTGGCGGGTGCCACCATTGGTGAGGGTGAGAAA  
GTGCTGATGTTCTGGGCTCTGCCAATCGTGATCCGCGTCGTTGGGACGATCCCGATCGCTATGA  
CATCACGCGGAAACATCGGGTCATGTTGGGTTCCGATCTGGCGTACACATGTGTGTGGGCCAG  
TTAGTGGCGCGCCTGGAAGGGGAAGTCGTGCTTGGCGCTTTAGCGCGCAAAGTGGCCGCCATCG  
AAATTGCAGGCCCGCTGAAACGTCGTTCAACAACACGCTTCGTGGGCTGGAATCCTTACCGATC  
CAACTGACCCCGGCGTAA -3'

#### his<sub>6</sub>-Pux

5'- **ATGGGCAGCAGCCATCACCATCATCACCACAGCCAGGATCCGCCTAGCATTACCTTTATCCT**  
GCCGGACGGTGAACGTCGTACAACAGAAGCGGCCGTAGGGGACACTGCAATGTACGCGGCACT  
GTCGCTTGGCCTGGATGGGGTGGTAGCGGAGTGTGGTGGTAACGCGGTTTGTGCCACTTGCCATG  
TTACGTGGAACATGGGTAGAGAACTCCCCGCGGTAGCGGCGGACGAAGATGACTTGTTGGA  
CGGGACAGCAGCTGAACGTTTACCTAATAGTCGTTTGAAGTTGTCAAATTAAGTTAGTAGTGACC  
TGGACGGTTTAATCCTTCGCATCCCCGACCGCCAAGTATAA -3'

#### his<sub>6</sub>-PuR

5'- **ATGCATCACCATCATCACCACAATGATACTGTGCTGATTGCAGGTGCTGGACATGCGGGTTT**  
CCAAGTGGCAGTGAGCTTACGTCAAGCAAAATATCCGGGTCGTATTGCCTTAATTAACGATGAA  
AAGCACTTACCTTACCAACGTCCACCTCTCAGCAAAGCCTACTTGAAATCGGGTGGCGACCCTA  
ATAGCTTAATGTTCCGCCCTGAGAAATCTTTCAAGACCAAGCAATTGAATTAATCTCTGACCGT  
ATGGTCAGTATCGACCGTGAAGGTCGTAAATTATTGCTGGCATCGGGTACAGCAATTGAATATG  
GACATCTGGTTCTGGCAACAGGGGCACGTAATCGCATGCTGGATGTTCCGAATGCATCCCTTCCC  
GATGTACTGTATCTCCGTACCCTTGATGAAAGCGAGGTCTTCGCCAACGCATGCCCGATAAGAA  
ACACGTTGTCGTTATTGGTGTGCGGTTTATTGGATTAGAGTTTCGCGGCGACCGCACGCGCGAAAG  
GCCTGGAAGTGGATGTGGTGGAAATTAGCACCCCGTGAATGGCTCGTGTTGTTACCCCCGAAATT  
TCTAGCTATTTTCATGACCGTCATTCGGGCGCCGGTATTCGTATGCATTATGGTGTGCGCGCGAC  
GGAGATTGCAGCAGAAGGGGACCGTGTGACGGGTGTGGTGTAAAGTGATGGTAATACCTTACCG  
TGTGACCTGGTTGTTGTAGGGGTTGGTGTATCCCTAATGTGGAAATTGCGGCAGCGGCTGGCCT  
TCCAACGGCTGCGGGTATTATTGTTGACCAACAATTACTGACGTCCGACCCACATATTAGCGCAA  
TTGGAGATTGTGCCCTCTTTGAATCTGTACGTTTTTGGTGAACCATGCGCGTTGAAAGCGTTCAA  
AATGCGACAGACCAAGCCCGTTGTGTGGCGGCCCGCTTACGGGAGATGCGAAGCCTTACGATG  
GATACCCCTGGTTTTGGTCTGACCAAGGAGATGATAAATTACAAATTGTAGGTCTTACTGCGGGG  
TTTGACCAGGTTGTTATTCGTGGTTCCGTAGCTGAACGCTCATTTTCAGCCTTTTGTATAAAGCT  
GGTAAACTTATTGGTATTGAAAGCGTGAATCGTGCGGCAGACCATGTTTTCGGTCGTAAATTTCT  
TCCTTTAGATAAGTCTGTTACCCCTGAACAAGCTGCGGACCTTTCATTTGATTTGAAGAAAGCAG  
CAGCATGA -3'

#### ZmCCoAOMT

5'- **ATGGCGACAACAGCCACAGAAGCAACTAAAACCTACCGCGCCTGCCCAAGAACAACAAGCTA**  
ATGGAAATGGTAATGGGGAACAGAAAACTCGTATAGCGAAGTAGGGCATAAATCGTTGCTGA  
AAAGTGATGATTTGTATCAATATATTTTGGATACATCCGTCTATCCCCGTGAACCAGAATCAATG  
AAAGAACTCCGTGAAATTACTGCGAAACATCCTTGAATCTCATGACGACTTCGGCGGATGAAG  
GACAATTCTTAAATATGCTGATTAAGTTGATTGGAGCGAAGAAAACGATGGAAATTGGTGTTTA  
TACTGGATATTCCTTATTAGCAACTGCCCTGGCGCTTCCTGAAGATGGAACCATCTGGCAATGG  
ACATTAATCGTGAAAATTATGAGTTAGGTTTACCTTGTATTAATAAAGCAGGTGTCGGGCATAAA  
ATCGATTTTCGTGAAGGGCCAGCACTGCCGGTTCTCGATGATTTAGTTGCCGATAAAGAACAACA  
TGGTAGCTTCGATTCGCGTTTGTAGATGCGGACAAAGATAATTACTTAAATTATCATGAGCGTC  
TGTTGAAATTGGTCCGTCCGGTGGTCTGATTGGATATGATAATACCCTCTGGAATGGTAGCGTG  
GTCCTGCCGGATGATGCCCCGATGCGTAAATATATTCGTTTCTATCGTGATTTTGTACTTGCACGTG  
AATAGCGCACTGGCTGCGGATGACCGTGTGGAAATTTGTCAACTGCCAGTGGGTGATGGTGTGA  
CTTTATGTCGTCGTGTGAAATAA -3'

**ZmCOMT**

5'-ATGGGCAGTACGGCCGGTGATGTTGCAGCAGTTGTAGACGAAGAAGCGTGTATGTATGCCAT  
GCAACTTGCCAGCAGTTCGATTTTGCCAATGACCTTGAAGAATGCTATTGAATTAGGATTATTGG  
AAGTCCTGCAGAAAGAAGCGGGTGGCGGAAAAGCTGCCTTAGCCCCGGAAGAAGTAGTCGCAC  
GTATGCCGGCTGCCCCGTCAGATCCTGCGGCGGCGGCGGCGATGGTTGATCGTATGCTGCGTTTG  
CTTGCGAGCTATGATGTGGTTCGCTGTCAAATGGAAGATCGCGATGGGCGCTATGAACGTCGTTA  
TTCGGCGGCCCCGGTTTGTAAATGGCTGACGCCAAATGAAGATGGAGTTTCAATGGCGGCGCTG  
GCCTTAATGAATCAAGATAAAGTACTGATGGAATCCTGGTATTACCTGAAAGATGCCGTATTAG  
ATGGTGGTATTCCATTTAATAAAGCCTATGGTATGACAGCCTTGAATATCATGGTACGGATGCA  
CGCTTTAATCGTGTTTTCAATGAAGGTATGAAGAACCATTCCGTTATCATTACTAAGAAGTTGTT  
AGACTTCTATACTGGGTTTGAAGGTGCTCTACCCTTGTCGATGTAGGTGGCGGGGTTGGGGCGA  
CCTTGCATGCGATTACCTCTCGTCATCCACATATTAGTGGAGTAAATTTTGATCTCCCTCATGTAA  
TTAGCGAAGCACCACCTTTTCCGGGTGTTTCGTCATGTTGGTGGTGATATGTTTGCATCAGTTCCA  
GCTGGAGATGCTATTCTGATGAAATGGATTCTGCATGATTGGTCAGATGCCCATTTGTGCGACTTT  
GCTGAAGAATTGTTATGATGCGCTCCAGAGAACGGGAAAGTTATTGTAGTTGAATGTGTCCTTC  
CTGTGAATACAGAAGCGACGCCGAAAGCCCAAGGTGTTTTCCATGTAGATATGATTATGCTTGCT  
CATAATCCGGGTGGTAAAGAACGCTATGAACGTGAATTCGTGAATTGGCGAAAGGAGCGGGAT  
TTAGCGGCTTTAAAGCGACTTATATTTATGCTAATGCGTGGGCTATTGAATTTATTAAATAA-3'

## VII.2 Construction of plasmids via restriction and ligation

**Table 16: Construction of vectors via restriction and ligation.** The vectors produced by restriction and ligation are listed in the first column with their respective designation. The symbol (I) or (II) indicate the position of the gene in the MCS I or II of the Duet vectors. The second column refers to the template DNA for the insert fragment and the last two column list the applied restriction enzymes used for inserts and vectors. If a PCR was performed prior to restriction and ligation, the corresponding used primer pair is mentioned in the third column.

Plasmid	DNA template (insert)	Fw primer / Rv primer	Restriction enzymes	
			(insert)	(vector)
Duet vectors				
pCDFDuet_ <i>AtCCoAOMT</i> (I)	pUC57_ <i>AtCCoAOMT</i>	–	NcoI, HindIII	NcoI, HindIII
pCDFDuet_ <i>AtCCoAOMT</i> _CYP199A2	pRSFDuet_ <i>AtCCoAOMT</i> (I)	–	NcoI, HindIII	NcoI, HindIII
pCDFDuet_ <i>AtCOMT</i> (I)	pUC57_ <i>AtCOMT</i>	–	NcoI, HindIII	NcoI, HindIII
pCDFDuet_CYP199A2 (II)	pEX-A258- CYP199A2	–	NdeI, EcoRV	NdeI, EcoRV
pCDFDuet_NΔ7- CYP199A2 (II)	pRSFDuet_NΔ7- CYP199A2 (II)	–	NdeI, XhoI	NdeI, XhoI
pCDFDuet_ <i>ZmCCoAOMT</i> (I)	pUC57_ <i>ZmCCoAOMT</i>	–	NcoI, HindIII	NcoI, HindIII
pCDFDuet_ <i>ZmCCoAOMT</i> _CYP199A2	pCDFDuet_ <i>ZmCCoAOMT</i> (I)	–	NcoI, HindIII	NcoI, HindIII
pCDFDuet	pCDFDuet	–	NcoI,	NcoI,

<i>ZmCCoAOMT_NΔ7-CYP199A2</i>	<i>ZmCCoAOMT</i> (I)		HindIII	HindIII
pCDFDuet_ <i>ZmCOMT</i> (I)	pUC57_ <i>ZmCOMT</i>	–	NcoI, HindIII	NcoI, HindIII
pCDFDuet_ <i>ZmCOMT</i> _ <i>CYP199A2</i>	pRSFDuet_ <i>ZmCOMT</i> (I)	–	NcoI, HindIII	NcoI, HindIII
pCDFDuet_ <i>ZmCOMT</i> _ NΔ7-CYP199A2	pCDFDuet_ <i>ZmCOMT</i> (I)	–	NcoI, HindIII	NcoI, HindIII
pRSFDuet_ <i>AtCCoAOMT</i> (I)	pUC57_ <i>AtCCoAOMT</i>	–	NcoI, HindIII	NcoI, HindIII
pRSFDuet_ <i>his</i> <sub>6</sub> -FpR (II)	pET11a_ <i>his</i> <sub>6</sub> -FpR	42_NdeI_ <i>his</i> <sub>6</sub> -FpR_fw / 43_ <i>his</i> <sub>6</sub> -FpR_XhoI_rv	NdeI, XhoI	NdeI, XhoI
pRSFDuet_ <i>his</i> <sub>6</sub> -PdR (II)	pET28a(+)_ <i>his</i> <sub>6</sub> - PdR	40_EcoRV_ <i>his</i> <sub>6</sub> -camA_fw / 41_ <i>his</i> <sub>6</sub> -camA_AatII_rv	EcoRV, AatII	EcoRV, AatII
pRSFDuet_ <i>his</i> <sub>6</sub> -Pdx_ <i>his</i> <sub>6</sub> -PdR	pET28a(+)_ <i>his</i> <sub>6</sub> - Pdx	–	NcoI, EcoRI	NcoI, EcoRI
pRSFDuet_ <i>his</i> <sub>6</sub> -PuR (II)	pUC57_PuR	E9_NdeI_ <i>his</i> <sub>6</sub> -PuR_fw2 / E10_PuR_XhoI_rv2	NdeI, XhoI	NdeI, XhoI
pRSFDuet_ <i>his</i> <sub>6</sub> -Pux (I)	pUC57_Pux	E3_BamHI_ <i>his</i> <sub>6</sub> -Pux_fw / E4_Pux_NotI_rv	BamHI, NotI	BamHI, NotI
pRSFDuet_ <i>his</i> <sub>6</sub> -Pux_ <i>his</i> <sub>6</sub> -PuR	pRSFDuet_ <i>his</i> <sub>6</sub> - Pux	–	NcoI, NotI	NcoI, NotI
pRSFDuet_ <i>his</i> <sub>10</sub> -YkuN_ <i>his</i> <sub>6</sub> -PdR	pET16b_ <i>his</i> <sub>10</sub> - YkuN	–	NcoI, BamHI	NcoI, BamHI
pRSFDuet_ <i>his</i> <sub>10</sub> -YkuN_ <i>his</i> <sub>6</sub> -PuR	pET16b_ <i>his</i> <sub>10</sub> - YkuN	–	NcoI, BamHI	NcoI, BamHI
pRSFDuet_ <i>his</i> <sub>10</sub> -YkuN_ <i>his</i> <sub>6</sub> -FpR	pET16b_ <i>his</i> <sub>10</sub> - YkuN	–	NcoI, BamHI	NcoI, BamHI
pRSFDuet_HpaC (II)	pETM6_RgTAL_ HpaB_HpaC_m	9_NdeI_HpaC_fw / 10_HpaC_XhoI_rv	NdeI, XhoI	NdeI, XhoI
pRSFDuet_NΔ7- CYP199A2 (II)	pEX-A258- CYP199A2	31_NdeI_CYP199A2 (-7aa)_fw / E8_CYP199A2_XhoI_rv	NdeI, XhoI	NdeI, XhoI
pRSFDuet_ <i>ZmCCoAOMT</i> (I)	pUC57_ <i>ZmCCoAOMT</i>	–	NcoI, HindIII	NcoI, HindIII
pRSFDuet_ <i>ZmCOMT</i> (I)	pUC57_ <i>ZmCOMT</i>	–	NcoI, HindIII	NcoI, HindIII
<b>ePathBrick vectors</b>				
pCDM4_ <i>At4CL</i>	pETM6_ <i>At4CL</i>	–	AvrII, SalI	AvrII, SalI



pCDM4_ <i>At4CL_ZmCCR_</i> <i>ZmCAD_m</i>	pCDM4_ <i>ZmCCR_</i> <i>ZmCAD_m</i>	–	AvrII, SalI	NheI, SalI
pCDM4_ <i>HpaB_HpaC_op</i>	pETM6_ <i>HpaB_</i> <i>HpaC_op</i>	–	AvrII, SalI	AvrII, SalI
pCDM4_ <i>ZmCCR_</i> <i>ZmCAD_m</i>	pETM6_ <i>ZmCCR_</i> <i>ZmCAD_m</i>	–	AvrII, SalI	AvrII, SalI
pETM6_ <i>FjTAL_At4CL_</i> <i>ZmCCR_ZmCAD_m</i>	pCDM4_ <i>At4CL_</i> <i>ZmCCR_</i> <i>ZmCAD_m</i>	–	AvrII, SalI	NheI, SalI
pETM6_ <i>HpaC</i>	pRSFDuet_ <i>HpaC</i> (II)	–	NdeI, XhoI	NdeI, XhoI
pETM6_ <i>HpaB_HpaC_m</i>	pETM6_ <i>HpaC</i>	–	AvrII, SalI	NheI, SalI
pETM6_ <i>HpaB_HpaC_pso</i>	pETM6_ <i>HpaC</i>	–	AvrII, SalI	SpeI, SalI
pETM6_ <i>HpaB_HpaC_op</i>	pETM6_ <i>HpaC</i>	–	XbaI, SalI	SpeI, SalI
pETM6_ <i>RgTAL_At4CL_</i> <i>m</i>	pETM6_ <i>At4CL</i>	–	AvrII, SalI	NheI, SalI
pETM6_ <i>RgTAL_Pc4CL_</i> <i>m</i>	pETM6_ <i>Pc4CL</i>	–	AvrII, SalI	NheI, SalI
pETM6_ <i>Sesam8_At4CL_</i> <i>ZmCCR_ZmCAD_m</i>	pCDM4_ <i>At4CL_</i> <i>ZmCCR_</i> <i>ZmCAD_m</i>	–	AvrII, SalI	NheI, SalI

### VII.3 Construction of plasmids via other molecular biology methods

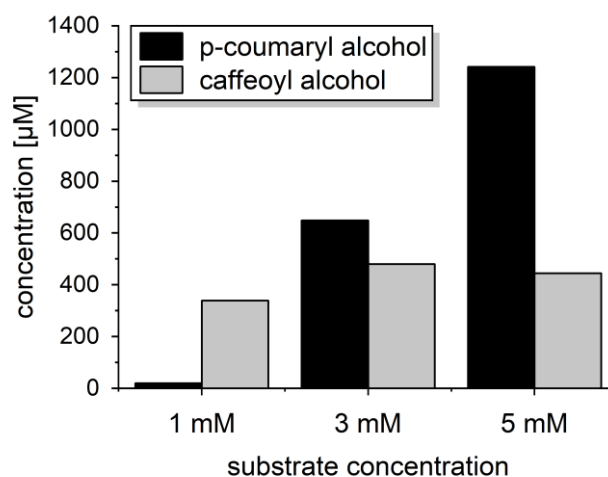
**Table 17: Construction of vectors via FastCloning and site-directed mutagenesis.** The vectors produced by FastCloning and mutagenesis are listed in the first column with their respective designation. For FastCloning individual PCR reactions were performed for insert and vector, the corresponding DNA templates are mentioned in columns 2 and 4 and the used primer pairs in columns 3 and 5. For mutagenesis, only the vector DNA template is relevant.

Plasmid	DNA template (insert)	Fw primer / Rv primer (insert)	DNA template (vector)	Fw primer / Rv primer (vector)
FastCloning				
pETM6_ <i>HpaB</i>	pETM6_ <i>RgTAL_</i> <i>HpaB_HpaC_</i> <i>m</i>	M2_ <i>HpaB_fw</i> / M2_ <i>HpaB_rv</i>	pETM6	M2_pETM6_ <i_fw< i=""> / M2_pETM6_<i_rv< i=""></i_rv<></i_fw<>
pETM6_ <i>RgTAL</i>	pETM6_ <i>RgTAL_</i> <i>HpaB_HpaC_</i> <i>m</i>	M1_ <i>RgTAL_fw</i> / M1_ <i>RgTAL_rv</i>	pETM6	M1_pETM6_ <i_fw< i=""> / M1_pETM6_<i_rv< i=""></i_rv<></i_fw<>

Mutagenesis					
pETM6_ <i>At4CL</i> (515. bp)	–	–	pETM6_ <i>At4CL</i> (WT)	M9_ <i>At4CL</i> _SalI_ 515.bp_fw / M9_ <i>At4CL</i> _SalI_ 515.bp_rv	
pETM6_ <i>At4CL</i> (515. bp, 629. bp)	–	–	pETM6_ <i>At4CL</i> (515. bp)	M9_ <i>At4CL</i> _SalI_ 629.bp_fw / M9_ <i>At4CL</i> _SalI_ 629.bp_rv	
pETM6_ <i>At4CL</i>	–	–	pETM6_ <i>At4CL</i> (515. bp, 629. bp)	M9_ <i>At4CL</i> _SalI_ 776.bp_fw / M9_ <i>At4CL</i> _SalI_ 776.bp_rv	

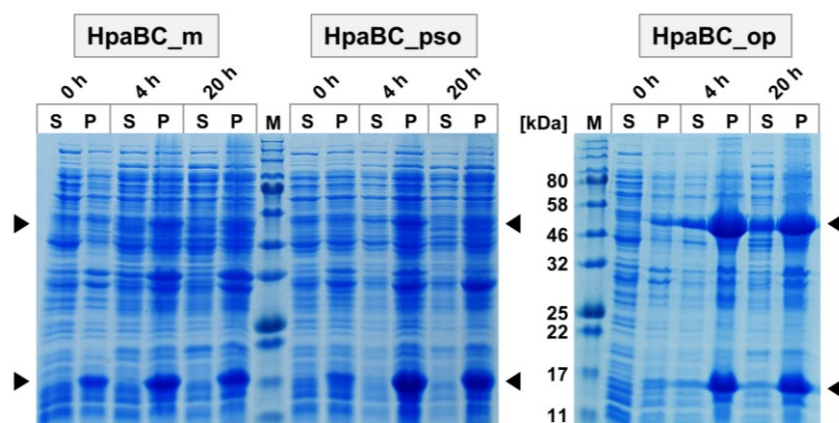
## VII.4 Supplementary figures

### Activity of HpaBC towards *p*-coumaryl alcohol



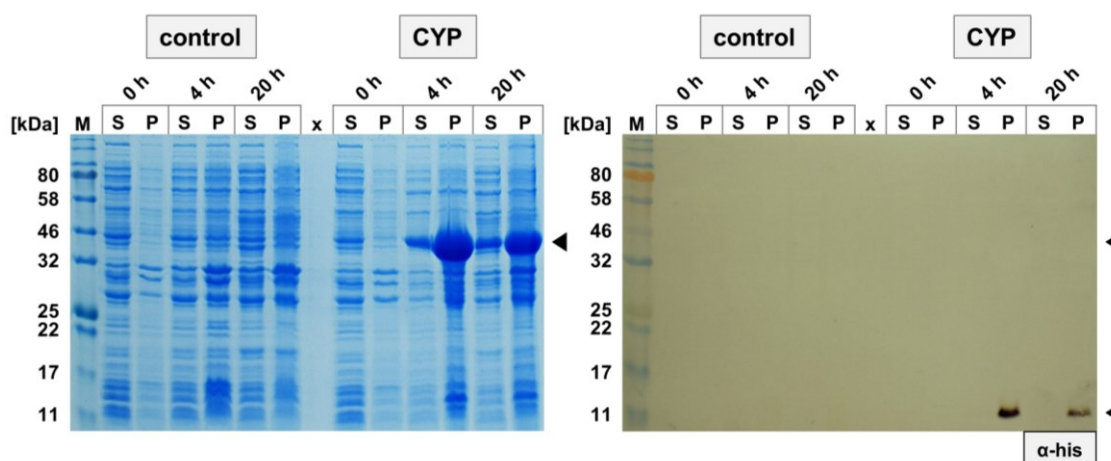
**Figure A.1: HpaBC activity towards *p*-coumaryl alcohol as substrate in *E. coli* BL21(DE3).** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM, 3 mM or 5 mM *p*-coumaryl alcohol (black) as substrate added at time  $t_0$ . The concentration of substances was measured after 24 h (caffeoyl alcohol (grey): 1 mM– 339.2 μM, 3 mM– 479.9 μM, 5 mM– 444.4 μM). The expression of HpaBC was performed at 37 °C from pETM6 in operon configuration prior to application in whole-cell biotransformation. The shown data originated from a single measurement.

## Expression analysis of the HpaBC complex over time



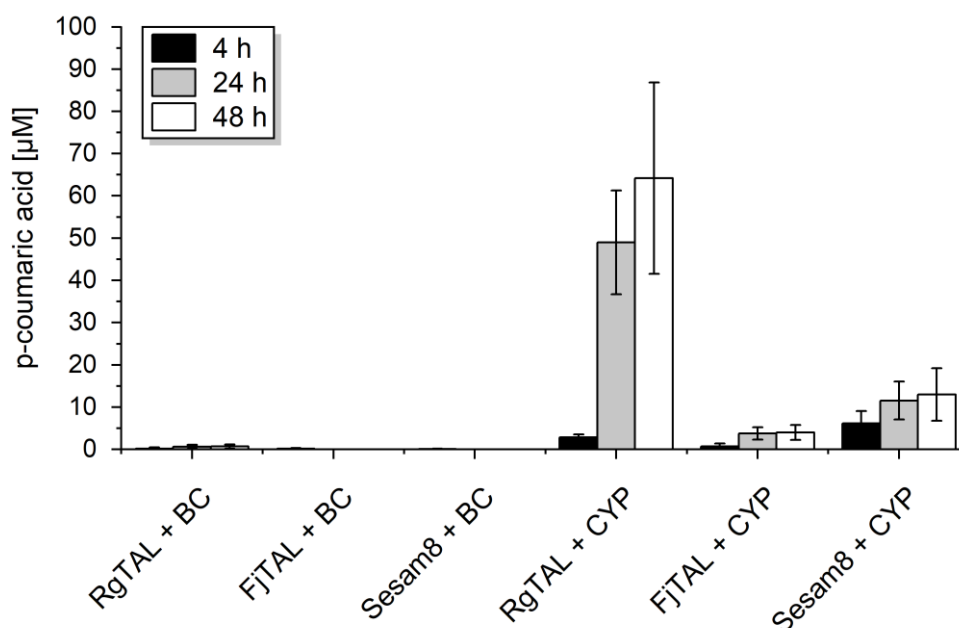
**Figure A.2: Detailed expression of HpaBC from pETM6 in different gene configurations.** Expression was performed from pETM6\_HpaB\_HpaC (HpaBC) in monocistronic (m), pseudo-operon (pso) or operon (op) configuration in *E. coli* BL21(DE3) at 37 °C. 5/OD samples were taken at the point of induction with 1 mM IPTG (0 h) and during the expression (4 h and 20 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE analysis. The black arrows point to the expression bands of both proteins. M – P7712 protein ladder (NEB), S – supernatant, P – pellet fraction, HpaB – 58.9 kDa, HpaC – 18.5 kDa.

## Expression analysis of CYP199A2 with Pux/PuR at 37 °C



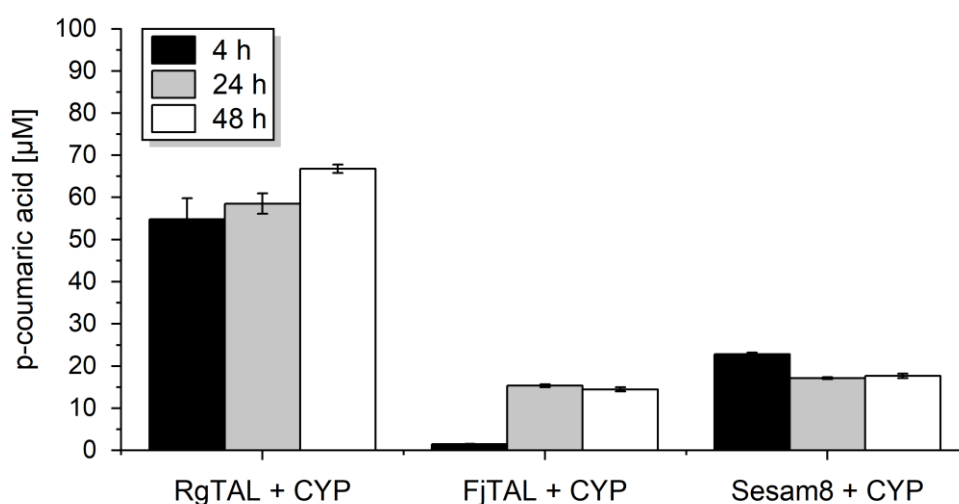
**Figure A.3: SDS-PAGE and western blot analysis of the cytochrome P450 monooxygenase from *R. palustris* co-expressed with Pux and PuR at 37 °C.** Co-expression was performed using pRSFDuet\_his<sub>6</sub>-Pux\_his<sub>6</sub>-PuR and pCDFDuet\_CYP199A2 (CYP) in *E. coli* BL21(DE3) at 37 °C. 5/OD samples were taken at the point of induction with 1 mM IPTG (0 h) and during the expression (4 h and 20 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE and western blot analysis. For detection in western blot analysis, the anti-His (α-his) and GAM<sub>FC</sub><sup>AP</sup> antibodies were used. An expression of the empty vectors (pCDFDuet, pRSFDuet) served as **control**. M – P7712 protein ladder (NEB), S – supernatant, P – pellet fraction, CYP199A2 – 44.6 kDa, his<sub>6</sub>-Pux – 12.6 kDa, his<sub>6</sub>-PuR – 44.5 kDa.

### Intermediate concentrations in TAL + C3H activities (with 1 mM L-tyrosine)



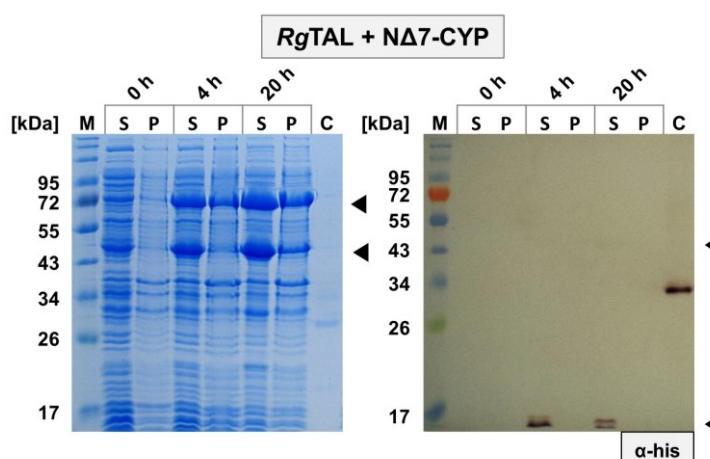
**Figure A.4: Production of *p*-coumaric acid from combined *in vivo* activity of TAL and C3H enzymes in *E. coli* BL21(DE3) using 1 mM L-tyrosine as substrate.** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM L-tyrosine as substrate added at time  $t_0$ . The concentration of *p*-coumaric acid was measured over time (4 h, 24 h, 48 h). Co-expression was performed at 26 °C prior to application in whole-cell biotransformation using pETM6 with TAL (*RgTAL*, *FjTAL*, *Sesam8*) and either pCDM4 with *HpaBC* in operon gene configuration (BC) or pCDFDuet with CYP199A2 (CYP) and pRSFDuet with Pux/PuR as redox pair. The shown data represent mean values with standard deviation from biological triplicates.

### Intermediate concentrations in TAL + C3H activities (with 3 mM L-tyrosine)



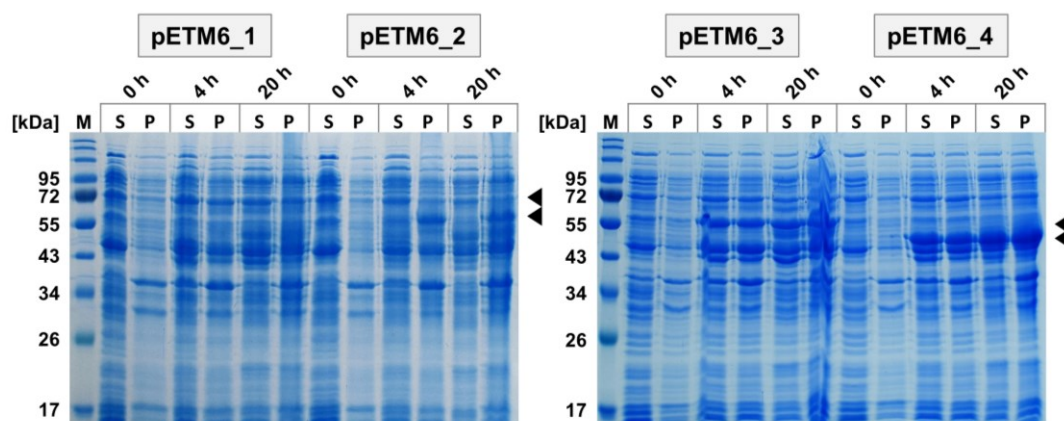
**Figure A.5: Production of *p*-coumaric acid from combined *in vivo* activity of TAL and C3H enzymes in *E. coli* BL21(DE3) using 3 mM L-tyrosine as substrate.** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 3 mM L-tyrosine as substrate added at time  $t_0$ . The concentration of *p*-coumaric acid was measured over time (4 h, 24 h, 48 h). Co-expression was performed at 26 °C prior to application in whole-cell biotransformation using pETM6 with TAL (*RgTAL*, *FjTAL*, *Sesam8*), pCDFDuet with CYP199A2 (CYP) and pRSFDuet with Pux/PuR as redox pair. The shown data represent mean values with standard deviation from biological duplicates.

## Co-expression analysis of *RgTAL* with NΔ7-CYP199A2



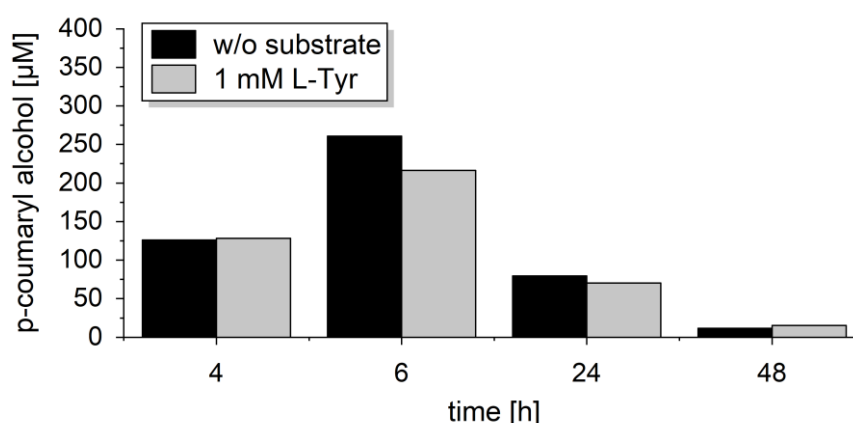
**Figure A.6: Co-expression analysis of *RgTAL* in combination with the NΔ7 variant of CYP199A2.** Co-expression was performed using pETM6\_*RgTAL* (*RgTAL*) with pCDFDuet\_NΔ7-CYP199A2 and pRSFDuet\_his<sub>6</sub>-Pux\_his<sub>6</sub>-PuR (NΔ7-CYP) in *E. coli* BL21(DE3) at 26 °C. 5/OD samples were taken at the point of induction with 1 mM IPTG (0 h) and during the expression (4 h and 20 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE and western blot analysis. For detection in western blot analysis the anti-His (α-his) and GAM<sub>FC</sub><sup>AP</sup> antibodies were used. The black arrowheads indicate the relevant expression bands. M – P7719 protein ladder (NEB), C – mCherry-his<sub>6</sub> as His<sub>6</sub>-tag control (0.5 μg, 27.7 kDa), S – supernatant, P – pellet fraction, *RgTAL* – 74.8 kDa, NΔ7-CYP199A2 – 43.9 kDa, his<sub>6</sub>-Pux – 12.6 kDa, his<sub>6</sub>-PuR – 44.5 kDa.

## Expression analysis of *p*-coumaryl alcohol pathway genes from a single vector



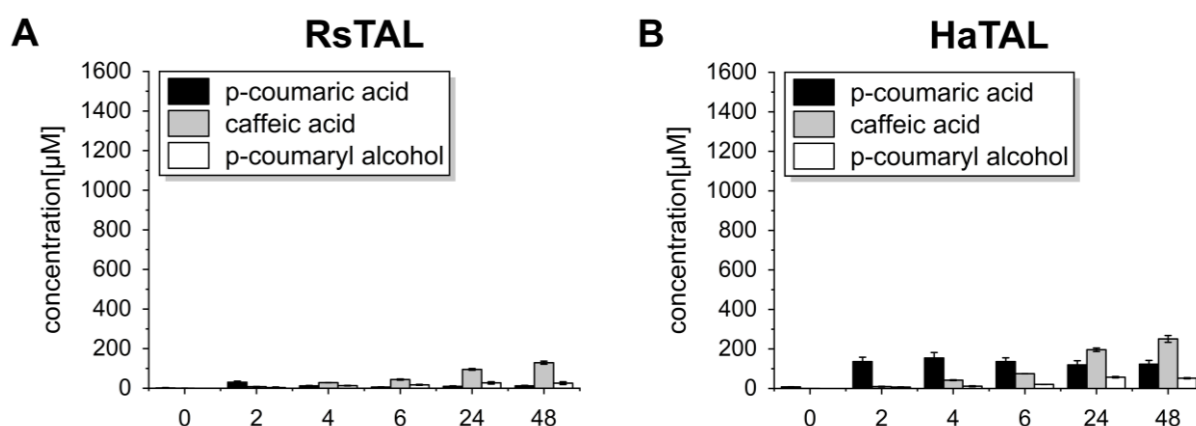
**Figure A.7: SDS-PAGE analysis of TAL, 4CL, *ZmCCR* and *ZmCAD* expressed from a single pETM6 plasmid.** The tested constructs were pETM6\_*RgTAL*\_At4CL\_*ZmCCR*\_ZmCAD\_m (pETM6\_1), pETM6\_*RgTAL*\_Pc4CL\_*ZmCCR*\_ZmCAD\_m (pETM6\_2), pETM6\_*FjTAL*\_At4CL\_*ZmCCR*\_ZmCAD\_m (pETM6\_3) and pETM6\_Sesam8\_At4CL\_*ZmCCR*\_ZmCAD\_m (pETM6\_4). Co-expression was performed using pETM6\_TAL\_4CL\_*ZmCCR*\_ZmCAD\_m in *E. coli* BL21(DE3) at 26 °C. 5/OD samples were taken at the point of induction with 1 mM IPTG (0 h) and during the expression (4 h and 20 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE analysis. M – P7719 protein ladder (NEB), S – supernatant, P – pellet fraction, *RgTAL* – 74.8 kDa, *FjTAL* – 56.7 kDa, Sesam8 – 54.0 kDa, *Pc4CL* – 60.3 kDa, *At4CL* – 61.1 kDa, *ZmCCR* – 40.1 kDa, *ZmCAD* – 38.7 kDa.

### The *p*-coumaryl alcohol pathway on a single plasmid and in growing cells



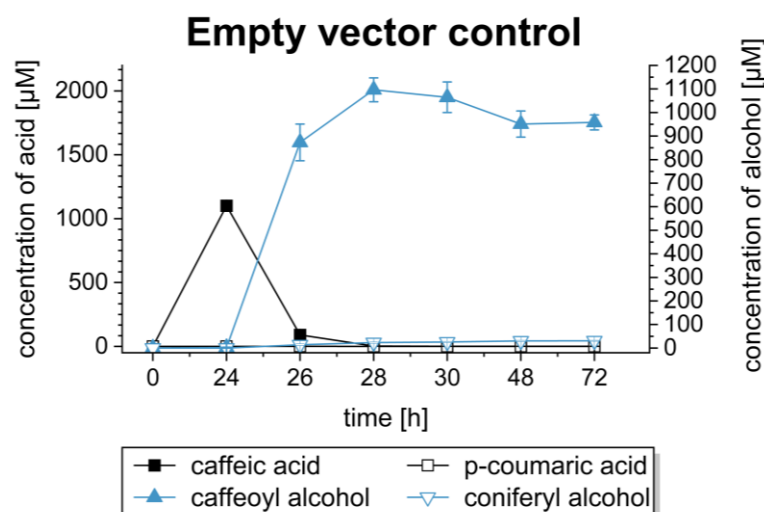
**Figure A.8:** Growing cells of *E. coli* BL21(DE3) with pETM6\_*RgTAL*\_At4CL\_ZmCCR\_ZmCAD\_m. The growing cell assay was performed in 10 mL TB medium at 26 °C without substrate addition (w/o, black) or with supplementation of 1 mM L-tyrosine (L-Tyr, grey) at time  $t_0$ . The concentration of *p*-coumaryl alcohol was measured over time (4 h, 6 h, 24 h, 48 h). The shown data originated from a single measurement.

### Coniferyl alcohol pathway with *RsTAL* and *HaTAL* in *E. coli* BL21(DE3).G213



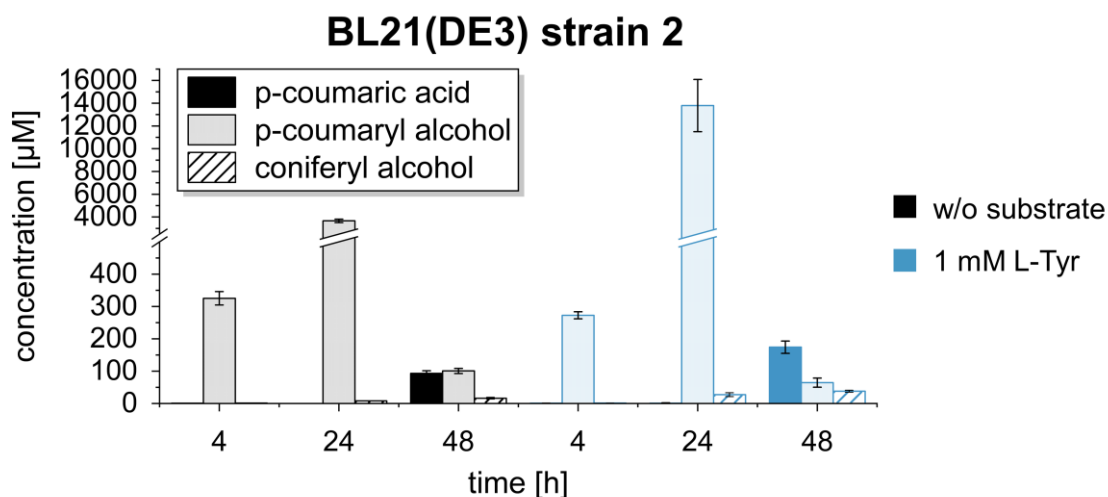
**Figure A.9:** Synthetic pathway towards coniferyl alcohol from L-tyrosine in *E. coli*/BL21(DE3).G213. Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 1 mM L-tyrosine as substrate added at time  $t_0$  (0 h). The concentration of products and intermediates was measured over time (2 h, 4 h, 6 h, 24 h, 48 h). Expression was performed at 26 °C prior to application in whole-cell biotransformation. Thereby, the constructs pCDFDuet\_AtCCoAOMT\_CYP199A2 and pRSFDuet\_Pux\_PuR were combined with one TAL enzyme expressed from pETM6: *RsTAL* (A) or *HaTAL* (B). The shown data represent mean values with standard deviation from biological duplicates.

## Empty vector control of two-cells one-pot



**Figure A.10: Empty vector control of two-cells one-pot experiment.** Whole-cell biotransformation was performed in 10 mL KPi buffer (50 mM, pH 7.5) at 26 °C with 2 mM caffeic acid as substrate added after 24 h. The concentration of products and intermediates was measured over time (from 24 h to 72 h). Expression was performed at 26 °C prior to application in whole-cell biotransformation. Thereby, the first set of cells contained empty vectors, while the second set of cells produced enzymes from pETM6\_At4CL, pCDFDuet\_ZmCOMT and pRSFDuet\_ZmCCR\_ZmCAD and were added at 24 h. The shown data represent mean values with standard deviation from biological duplicates.

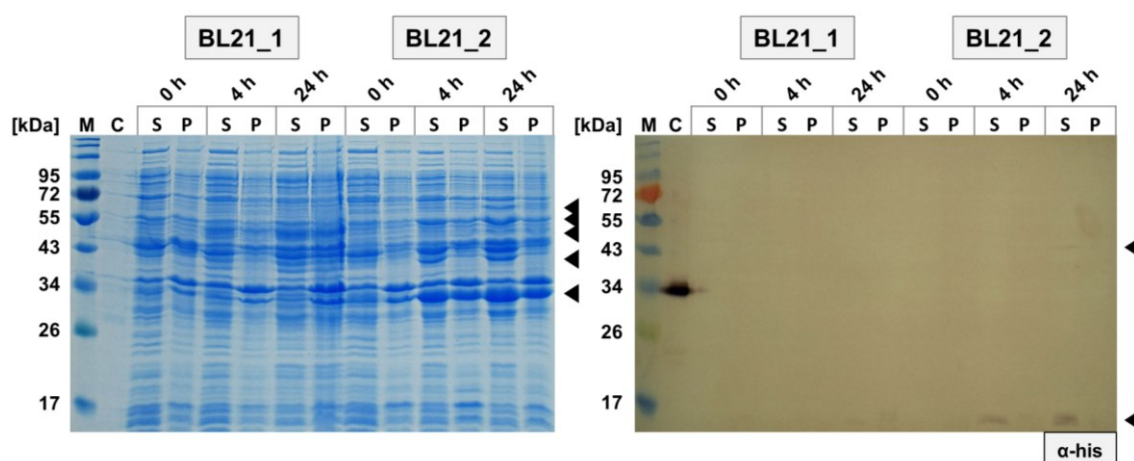
## Growing cell assay of *E. coli* BL21(DE3) with ZmCCoAOMT



**Figure A.11: Growing cells of *E. coli* BL21(DE3) with the combined upstream and downstream synthetic pathway using plasmid-based enzyme expression.** The co-expression of enzymes was performed by combining the single ePathBrick vector pETM6 containing *F<sub>1</sub>TAL*, *At4CL*, *ZmCCR* and *ZmCAD* with pCDFDuet\_ZmCCoAOMT\_NΔ7-CYP199A2 and pRSFDuet\_his<sub>6</sub>-Pux<sub>his</sub>-PuR. The growing cell assay was performed in 10 mL TB medium at 26 °C without substrate addition (w/o, black) or with supplementation of 1 mM L-tyrosine (L-Tyr, blue) at time *t*<sub>0</sub>. The concentration of products and intermediates was measured over time (4 h, 24 h, 48 h). The highest titers for coniferyl alcohol were 16.4 ± 2.7 μM (w/o, 48 h) and 38.1 ± 3.2 μM (with 1 mM L-tyrosine, 48 h). The shown data represent mean values with standard deviation from biological triplicates.

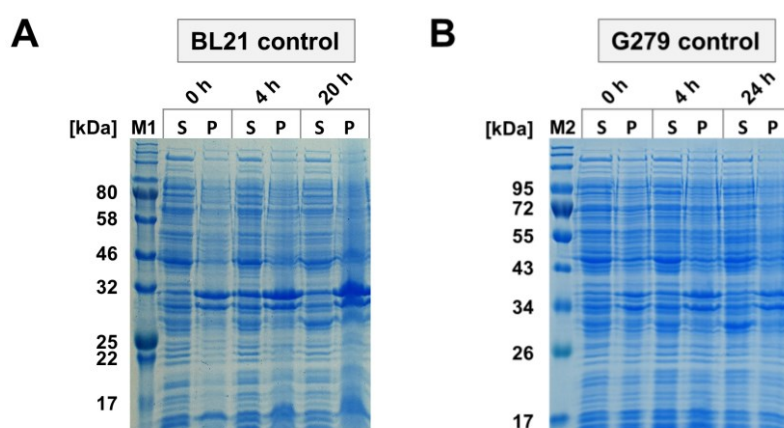


## Expression analysis of *E. coli* BL21(DE3) strains in growing cell assay



**Figure A.12: SDS-PAGE and western blot analysis of *E. coli* BL21(DE3) containing the complete synthetic coniferyl alcohol pathway under growing cell assay conditions.** Co-expression was performed using pETM6\_*FjTAL*\_At4CL\_*ZmCCR*\_ZmCAD\_m with pRSFDuet\_his<sub>6</sub>-Pux\_his<sub>6</sub>-PuR and either pCDFDuet\_*ZmCOMT*\_NΔ7-CYP199A2 (**BL21\_1**) or pCDFDuet\_*ZmCCoAOMT*\_NΔ7-CYP199A2 (**BL21\_2**) in *E. coli* BL21(DE3) at 26 °C. 5/OD samples were taken at the point of induction (OD~2) with 0.5 mM IPTG (0 h) and during the expression (4 h and 24 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE and western blot analysis. For detection in western blot analysis, the anti-His ( $\alpha$ -his) and GAM<sub>FC</sub><sup>AP</sup> antibodies were used. The black arrowheads indicate the relevant expression bands. M – P7719 protein ladder (NEB), C – mCherry-his<sub>6</sub> as His<sub>6</sub>-tag control (0.5  $\mu$ g, 27.7 kDa), S – supernatant, P – pellet fraction, *FjTAL* – 56.7 kDa, At4CL – 61.1 kDa, *ZmCCR* – 40.1 kDa, *ZmCAD* – 38.7 kDa, NΔ7-CYP199A2 – 43.9 kDa, his<sub>6</sub>-Pux – 12.6 kDa, his<sub>6</sub>-PuR – 44.5 kDa, *ZmCOMT* – 39.6 kDa, *ZmCCoAOMT* – 29.4 kDa.

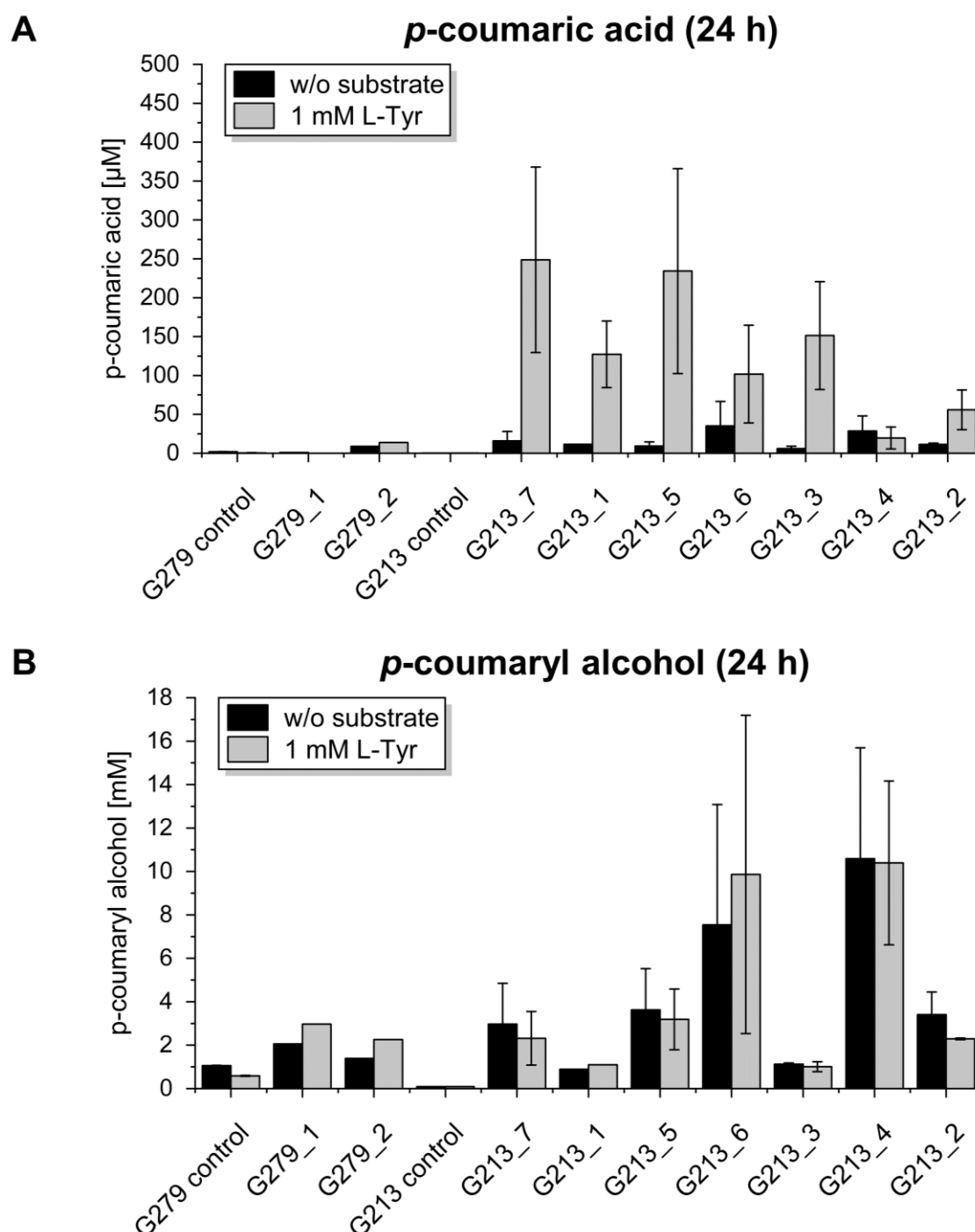
## Expression analysis of empty vector controls in growing cell assays



**Figure A.13: SDS-PAGE analysis of *E. coli* strains containing empty vectors under growing cell assay conditions.** Co-expression was performed using the empty vectors pETM6, pRSFDuet and pCDFDuet in *E. coli* BL21(DE3) (**BL21 control**, A) or *E. coli* BL21(DE3).G279 (**G279 control**, B) at 26 °C. 5/OD samples were taken at the point of induction (OD~2) with 0.5 mM IPTG (0 h) and during the expression (4 h and 24 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE analysis. M1 – P7712 protein ladder (NEB), M2 – P7719 protein ladder (NEB), S – supernatant, P – pellet fraction.

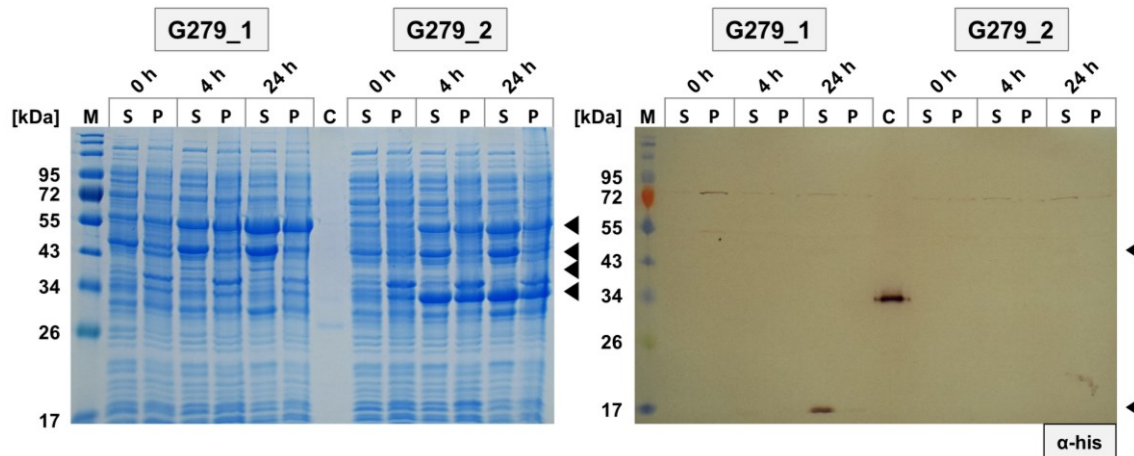


# Growing cell assays of *E. coli* BL21(DE3).G213 and *E. coli* BL21(DE3).G279



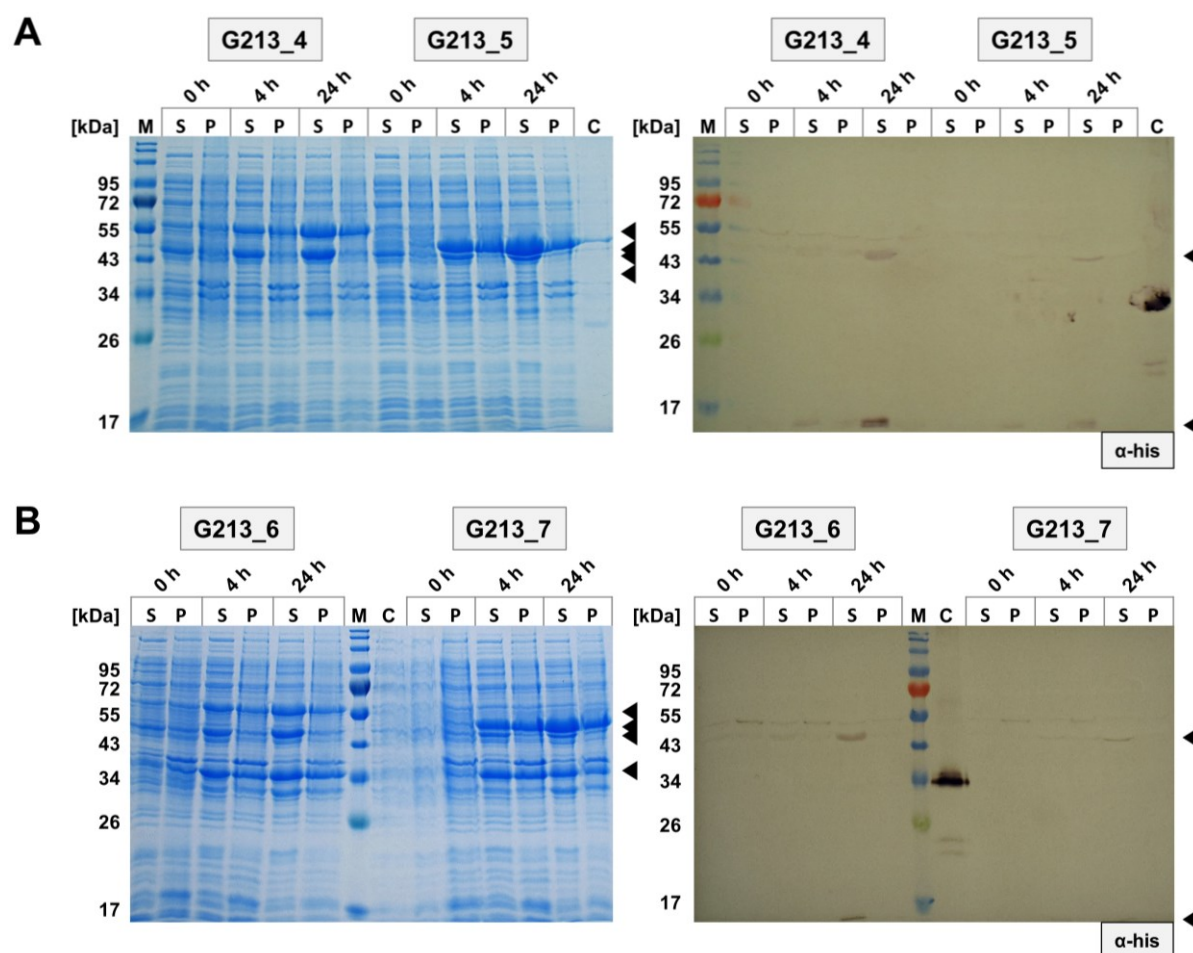
**Figure A.14:** Growing cells of *E. coli* BL21(DE3).G213 and *E. coli* BL21(DE3).G279. The growing cell assay was performed in 10 mL TB medium at 26 °C without substrate addition (w/o, black) or with supplementation of 1 mM L-tyrosine (L-Tyr, grey) at time  $t_0$ . The concentrations of *p*-coumaric acid (A) and *p*-coumaryl alcohol (B) were measured at 24 h. The shown data represent mean values with standard deviation from at least biological duplicates. As exceptions, the data for G279\_1, G279\_2, the G213 control, G213\_1 (w/o substrate) and G213\_1 (1 mM L-tyrosine, B) originated from single measurements. The investigated enzyme combinations in the G279 strain consisted of *Fj*TAL, *ND7*-CYP199A2 with *Pux*/*PuR* and *Zm*COMT (G279\_1) or *Zm*CCoAOMT (G279\_2). Growing cells using the three empty vectors served as activity controls (G213/G279 control). In the G213 strain, *Pux*/*PuR* was combined with the following: G213\_1 – *Rg*TAL + CYP199A2 + *At*CCoAOMT, G213\_2 – *Fj*TAL + CYP199A2 + *At*CCoAOMT, G213\_3 – Sesam8 + CYP199A2 + *At*CCoAOMT, G213\_4 – *Fj*TAL + *ND7*-CYP199A2 + *Zm*COMT, G213\_5 – Sesam8 + *ND7*-CYP199A2 + *Zm*COMT, G213\_6 – *Fj*TAL + *ND7*-CYP199A2 + *Zm*CCoAOMT, G213\_7 – Sesam8 + *ND7*-CYP199A2 + *Zm*CCoAOMT.

### Expression analysis of *E. coli* BL21(DE3).G279 strains in growing cell assay



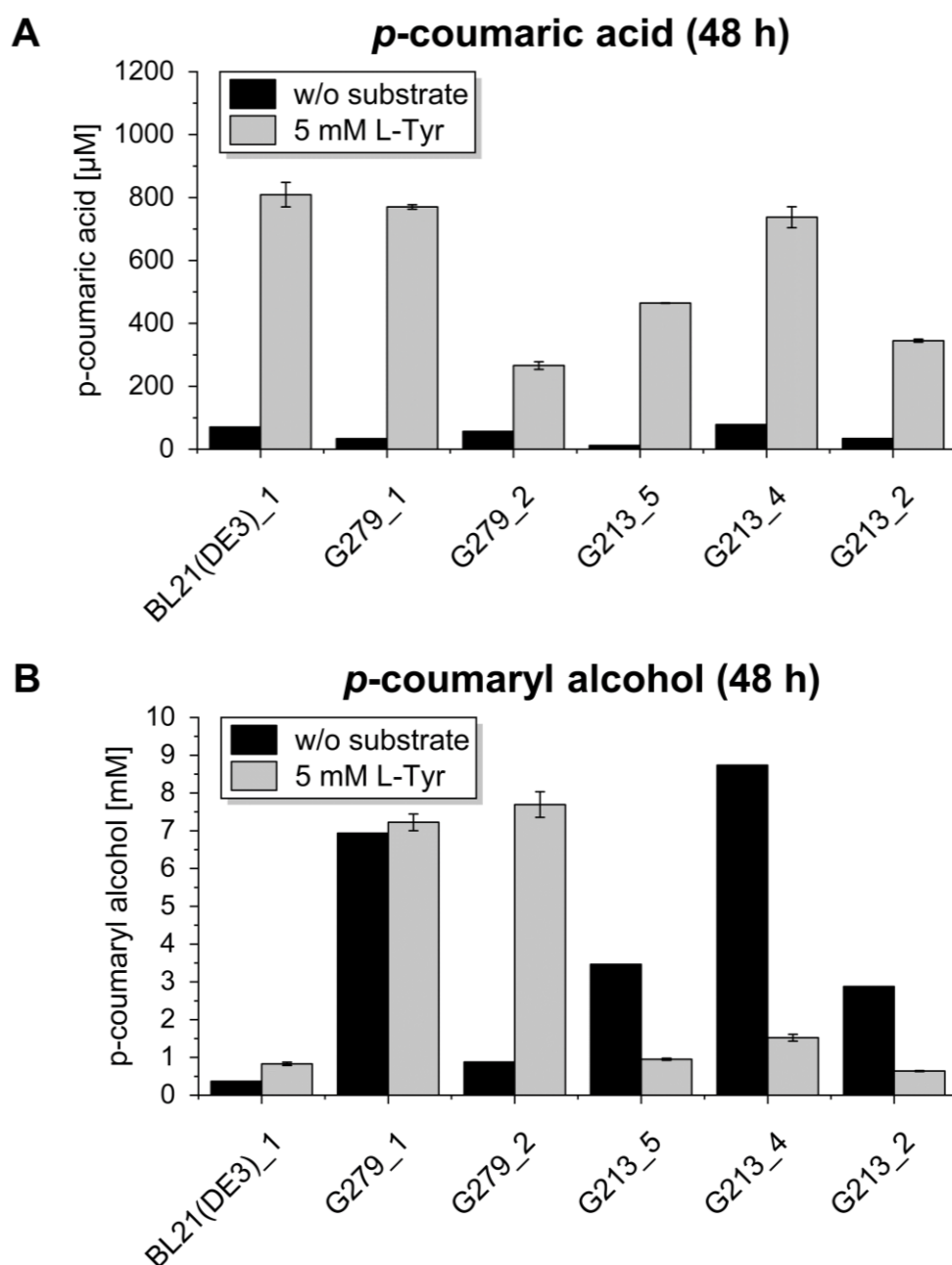
**Figure A.15: SDS-PAGE and western blot analysis of *E. coli* BL21(DE3).G279 strains under growing cell assay conditions.** Co-expression was performed using pETM6\_*FjTAL* with pRSFDuet\_his<sub>6</sub>-Pux\_his<sub>6</sub>-PuR and either pCDFDuet\_*ZmCOMT*\_NΔ7-CYP199A2 (**G279\_1**) or pCDFDuet\_*ZmCCoAOMT*\_NΔ7-CYP199A2 (**G279\_2**) in *E. coli* BL21(DE3).G279 at 26 °C. 5/OD samples were taken at the point of induction (OD~2) with 0.5 mM IPTG (0 h) and during the expression (4 h and 24 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE and western blot analysis. For detection in western blot analysis, the anti-His ( $\alpha$ -his) and GAM<sub>FC</sub><sup>AP</sup> antibodies were used. The black arrowheads indicate the relevant expression bands. M – P7719 protein ladder (NEB), C – mCherry-his<sub>6</sub> as His<sub>6</sub>-tag control (0.5  $\mu$ g, 27.7 kDa), S – supernatant, P – pellet fraction, *FjTAL* – 56.7 kDa, *At4CL* – 61.1 kDa, *ZmCCR* – 40.1 kDa, *ZmCAD* – 38.7 kDa, NΔ7-CYP199A2 – 43.9 kDa, his<sub>6</sub>-Pux – 12.6 kDa, his<sub>6</sub>-PuR – 44.5 kDa, *ZmCOMT* – 39.6 kDa, *ZmCCoAOMT* – 29.4 kDa, *R5TAL* – 55.5 kDa, *Pc4CL* – 60.3 kDa.

## Expression analysis of *E. coli* BL21(DE3).G213 strains in growing cell assay



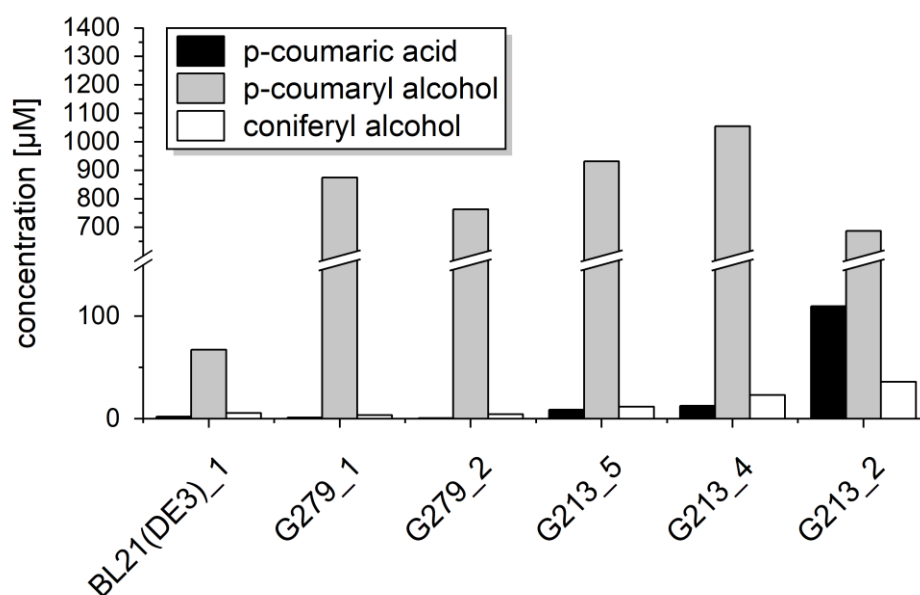
**Figure A.16: SDS-PAGE and western blot analysis of *E. coli* BL21(DE3).G213 strains under growing cell assay conditions.** Co-expression was performed using pETM6\_TAL with pRSFDuet<sub>his6</sub>-Pux<sub>his6</sub>-PuR and pCDFDuet\_OMT\_NΔ7-CYP199A2 in *E. coli* BL21(DE3).G213 at 26 °C. The tested enzyme combinations were *Fj*TAL + *Zm*COMT (G213\_4, A), Sesam8 + *Zm*COMT (G213\_5, A), *Fj*TAL + *Zm*CCoAOMT (G213\_6, B) and Sesam8 + *Zm*CCoAOMT (G213\_7, B). 5/OD samples were taken at the point of induction (OD~2) with 0.5 mM IPTG (0 h) and during the expression (4 h and 24 h). The harvested cells were divided into a soluble (S) and an insoluble fraction (P) prior to application in SDS-PAGE and western blot analysis. For detection in western blot analysis, the anti-His (α-his) and GAM<sub>FC</sub><sup>AP</sup> antibodies were used. The black arrowheads indicate the relevant expression bands. M – P7719 protein ladder (NEB), C – mCherry-his<sub>6</sub> as His<sub>6</sub>-tag control (0.5 μg, 27.7 kDa), S – supernatant, P – pellet fraction, *Fj*TAL – 56.7 kDa, Sesam8 – 54.0 kDa, NΔ7-CYP199A2 – 43.9 kDa, his<sub>6</sub>-Pux – 12.6 kDa, his<sub>6</sub>-PuR – 44.5 kDa, *Zm*COMT – 39.6 kDa, *Zm*CCoAOMT – 29.4 kDa, *R5*TAL – 55.5 kDa, *Pc*4CL – 60.3 kDa, *Zm*CCR – 40.1 kDa, *Zm*CAD – 38.7 kDa.

## Increase of L-tyrosine feed in growing cell assays

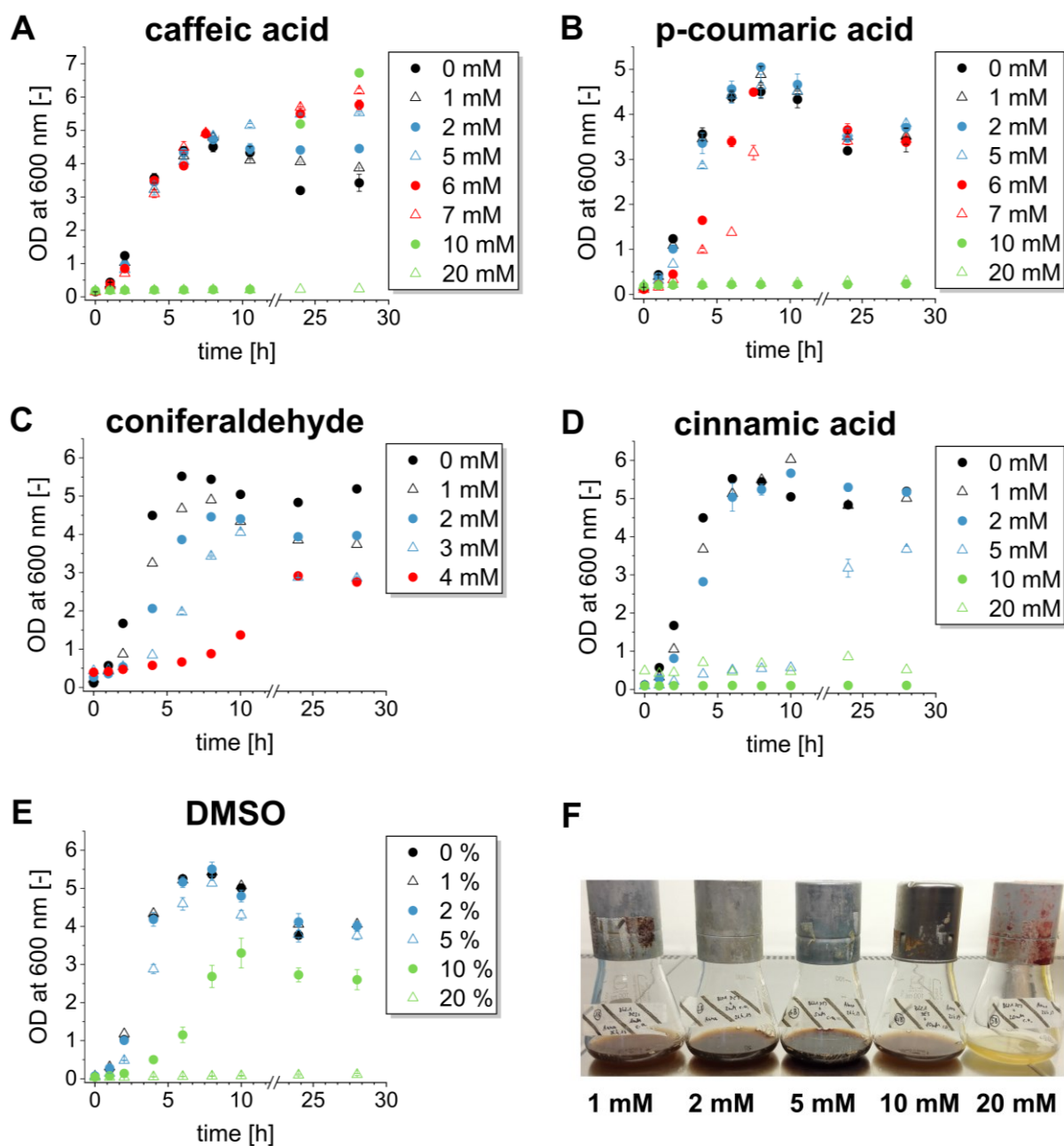


**Figure A.17: Intermediates from growing cells of different *E. coli*/BL21(DE3) strains with higher L-tyrosine feed.** The growing cell assay was performed in 10 mL TB medium at 26 °C without substrate addition (w/o, black) or with supplementation of 5 mM L-tyrosine (L-Tyr, grey) at time  $t_0$ . The concentrations of *p*-coumaric acid (**A**) and *p*-coumaryl alcohol (**B**) were measured at 48 h. The shown data with added substrate represent mean values with standard deviation from biological duplicates, while the data w/o substrate originated from a single measurement. The *E. coli* BL21(DE3) strain 1 (BL21(DE3)\_1) contained the enzymes *Zm*COMT and *NA*7-CYP199A2 with Pux/PuR and was co-transformed further with pETM6 containing *Fj*TAL, *At*4CL, *Zm*CCR and *Zm*CAD. The investigated enzyme combinations in the G279 strain consisted of *Fj*TAL, *NA*7-CYP199A2 with Pux/PuR and *Zm*COMT (G279\_1) or *Zm*CCoAOMT (G279\_2). In the G213 strain, Pux/PuR was combined with the following: G213\_2 – *Fj*TAL + CYP199A2 + *At*CCoAOMT, G213\_4 – *Fj*TAL + *NA*7-CYP199A2 + *Zm*COMT, G213\_6 – *Fj*TAL + *NA*7-CYP199A2 + *Zm*CCoAOMT.

## Growing cell assays with cinnamic acid



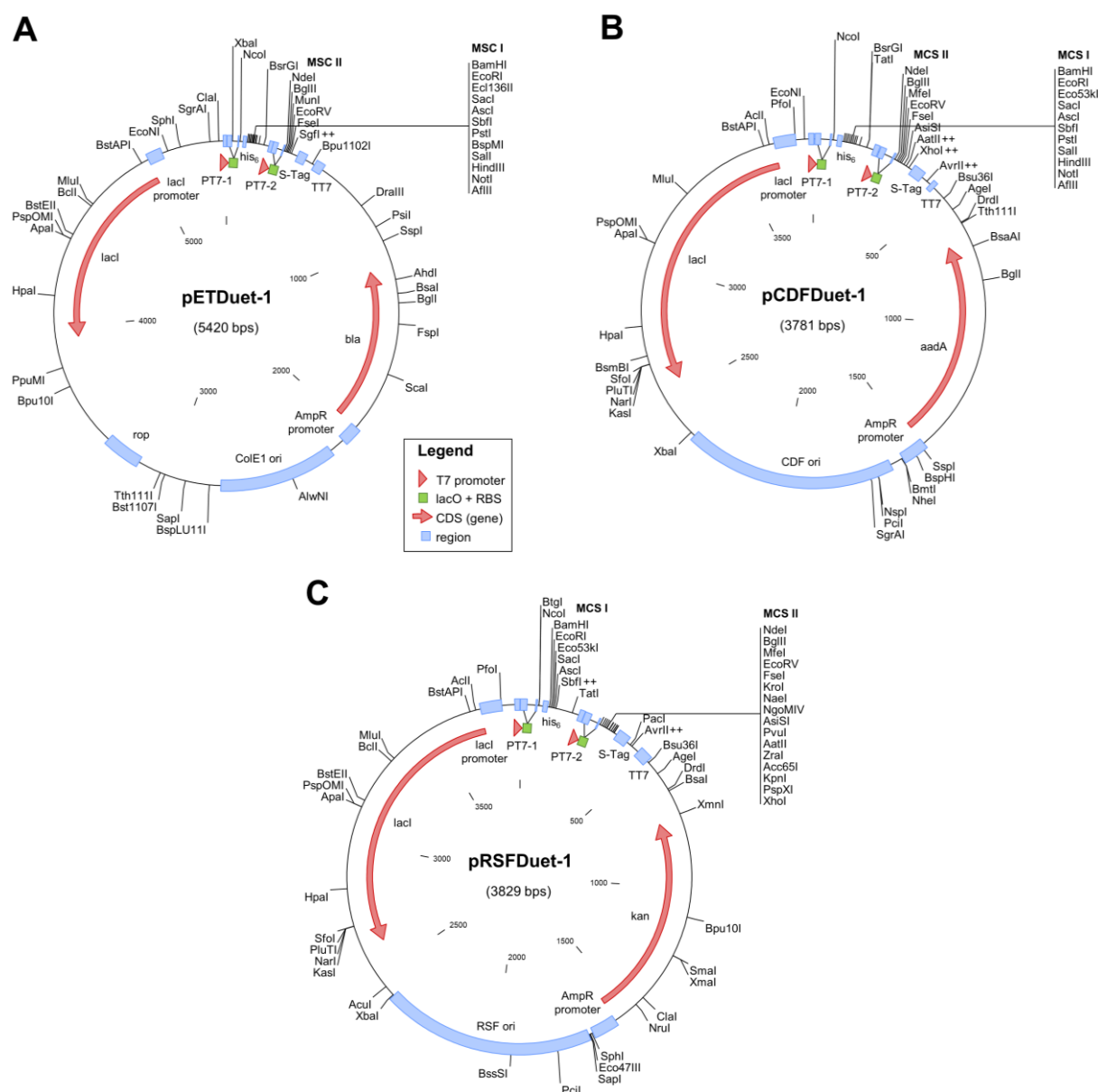
**Figure A.18: Growing cells of different *E. coli*/BL21(DE3) strains with cinnamic acid feed.** The growing cell assay was performed in 10 mL 2x TB medium at 26 °C with supplementation of 2 mM cinnamic acid at time  $t_0$ . The concentration of substances was measured at 48 h. The shown data originated from a single measurement. The *E. coli*/BL21(DE3) strain 1 (BL21(DE3)\_1) contained the enzymes *ZmCOMT* and  $\Delta 7$ -CYP199A2 with Pux/PuR and was co-transformed further with pETM6 containing *FjTAL*, *At4CL*, *ZmCCR* and *ZmCAD*. The investigated enzyme combinations in the G279 strain consisted of *FjTAL*,  $\Delta 7$ -CYP199A2 with Pux/PuR and *ZmCOMT* (G279\_1) or *ZmCCoAOMT* (G279\_2). In the G213 strain, Pux/PuR was combined with the following: G213\_2 – *FjTAL* + CYP199A2 + *AtCCoAOMT*, G213\_4 – *FjTAL* +  $\Delta 7$ -CYP199A2 + *ZmCOMT*, G213\_6 – *FjTAL* +  $\Delta 7$ -CYP199A2 + *ZmCCoAOMT*.

VII.5 Toxicity of different substrates on *E. coli*

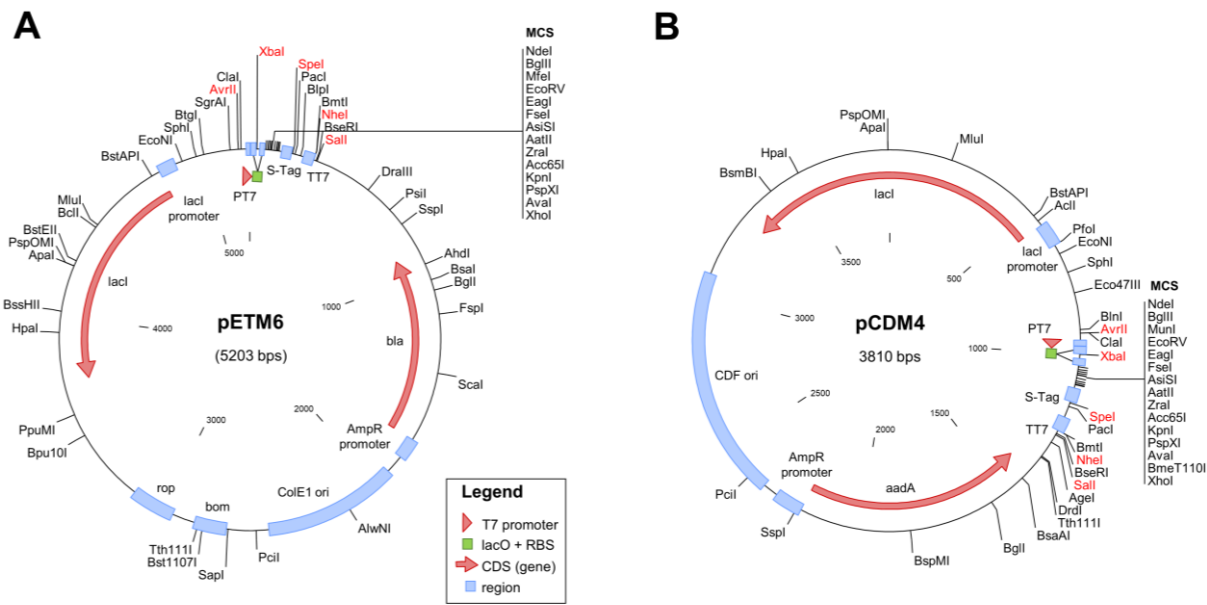
**Figure A.19: Toxicity of different substrates and intermediates as well as DMSO as control for *E. coli*.** 20 mL LB medium in a 100 mL shake flask were inoculated with 500  $\mu$ L overnight culture of *E. coli* BL21(DE3) and incubated at 37  $^{\circ}$ C and 180 rpm for 28 h. At the start of the incubation different concentrations of caffeic acid (A), *p*-coumaric acid (B), coniferaldehyde (C) cinnamic acid (D) and DMSO (E) were added using stock solutions of 1 M in DMSO or pure DMSO. The OD<sub>600</sub> (OD at 600 nm) of the growing cultures was measured over time. For A-D the means and standard deviations originate from the measurement of two dilutions at a certain time point. Only for E a biological duplicate was performed additionally. F shows the shake-flasks of *E. coli* BL21(DE3) cultures with different concentrations of caffeic acid after a cultivation of 24 h.



## VII.6 Plasmid maps

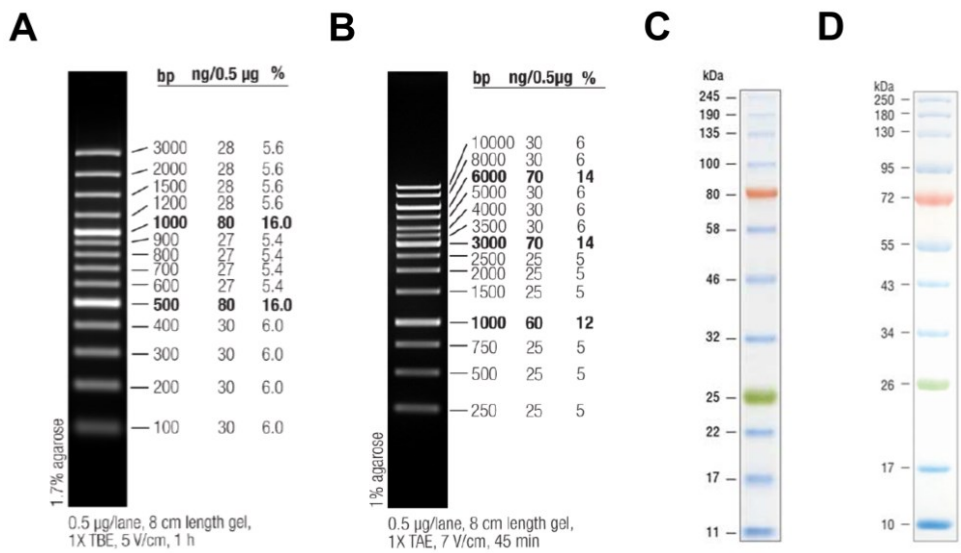


**Figure A.20: Plasmid maps of the Duet vectors pETDuet (A), pCDFDuet (B) and pRSFDuet (C).** Restriction enzyme sites are indicated outside the vector. A local vector site containing many restriction enzymes is referred to as multiple cloning site I or II (MCS I or MCS II). The protein expression is controlled by an IPTG-inducible T7 promoter (PT7) and a T7 terminator (TT7). The ColE1 origin of replication (ori) from pETDuet confers a medium copy plasmid number (~40), while pCDFDuet with a CDF replicon has a low to medium (20-40) and pRSFDuet with the RSF ori a high copy plasmid number (100) [169]. The resistance genes for pETDuet, pCDFDuet and pRSFDuet are  $\beta$ -lactamase (bla), aminoglycoside adenyltransferase (aadA) and aminoglycoside phosphotransferase (kan), respectively. lacO– lac operator, RBS– ribosome binding site, his<sub>6</sub>– His<sub>6</sub>-tag, S-Tag– S-protein epitope tag, lacI– lactose-inducible lac operon transcriptional repressor, CDS– coding sequence.



**Figure A.21: Plasmid maps of the ePathBrick vectors pETM6 (A) and pCDM4 (B).** Restriction enzyme sites are indicated outside the vector. A local vector site containing many restriction enzymes is referred to as multiple cloning site (MCS). In addition to the Duet vectors, the ePathBrick vectors contain the restriction enzyme sites *AvrII*, *XbaI*, *SpeI*, *NheI* (isocaudamers) and *SalI*, which are highlighted in red. The protein expression is controlled by an IPTG-inducible T7 promoter (PT7) and a T7 terminator (TT7). The ColE1 origin of replication (ori) from pETM6 confers a medium copy plasmid number (40), while pCDM4 with a CDF replicon has a low to medium copy plasmid number (20-40) [152]. The resistance genes for pETM6 and pCDM4 are  $\beta$ -lactamase (*bla*) and aminoglycoside adenyltransferase (*aadA*), respectively. lacO– lac operator, RBS– ribosome binding site, S-Tag– S-protein epitope tag, lacI– lactose-inducible lac operon transcriptional repressor, CDS– coding sequence.

VII.7 DNA and protein ladders



**Figure A.22: DNA and protein ladders.** GeneRuler™ 100 bp Plus DNA Ladder (A) (<https://www.thermofisher.com/order/catalog/product/SM0323>). GeneRuler™ 1 kb DNA Ladder (B) (<https://www.thermofisher.com/order/catalog/product/SM0312>). P7712 ColorPlus™ Prestained Protein Standard, Broad Range (11-245 kDa) (C) (<https://international.neb.com/products/p7712-color-prestained-protein-standard-broad-range-11-245-kda>). P7719 ColorPlus™ Prestained Protein Standard, Broad Range (10-250 kDa) (D) (<https://international.neb.com/products/p7719-color-prestained-protein-standard-broad-range-10-250-kda>).



## VII.8 List of Figures

Figure 1.1: Biosynthesis and examples of lignans. ....	6
Figure 1.2: Schematic view of monolignol and lignan synthesis pathway and the proposed localization of their enzymes in plant cells. ....	8
Figure 1.3: Phenylpropanoid pathway. ....	11
Figure 1.4: Formation of the catalytic MIO group (A) and proposed reaction mechanisms for tyrosine ammonia lyase (B) and phenylalanine ammonia lyase (C). ....	12
Figure 1.5: Schematic organization of different classes of cytochrome P450 monooxygenases with their redox partners. ....	14
Figure 1.6: The catalytic cycle of P450 enzymes. ....	16
Figure 1.7: Proposed reaction mechanisms of COMT (A) and CCoAOMT (B) from <i>Sorghum bicolor</i> . ....	19
Figure 1.8: Schematic overview of the work described in this dissertation. ....	22
Figure 2.1: <i>In vivo</i> activities of different TAL enzymes in <i>E. coli</i> BL21(DE3). ....	25
Figure 2.2: SDS-PAGE analysis of the different TAL enzymes. ....	26
Figure 2.3: Influence of different substrates on pH. ....	27
Figure 2.4: <i>In vivo</i> activities of different PAL enzymes in <i>E. coli</i> BL21(DE3). ....	28
Figure 2.5: SDS-PAGE and western blot analysis of PAL from <i>Z. mays</i> . ....	29
Figure 2.6: <i>In vivo</i> C3H activity of the hydroxylase complex HpaBC from <i>E. coli</i> . ....	31
Figure 2.7: SDS-PAGE analysis of the HpaBC hydroxylase complex from <i>E. coli</i> . ....	33
Figure 2.8: <i>In vivo</i> C3H activity of CYP199A2 and its NΔ7 variant in <i>E. coli</i> BL21(DE3). ..	34
Figure 2.9: SDS-PAGE and western blot analysis of the cytochrome P450 monooxygenase from <i>R. palustris</i> co-expressed with Pux and PuR at 26 °C. ....	35
Figure 2.10: <i>In vivo</i> C3H activity of CYP199A2 in <i>E. coli</i> BL21(DE3) using alternative redox partners and mixed combinations. ....	36
Figure 2.11: SDS-PAGE and western blot analysis of the cytochrome P450 monooxygenase from <i>R. palustris</i> co-expressed with Pdx and PdR at 26 °C. ....	38
Figure 2.12: SDS-PAGE and western blot analysis of CYP199A2 from <i>R. palustris</i> co-expressed with alternative redox partners (A) and mixed combinations (B). ....	40
Figure 2.13: <i>In vivo</i> C4H activity of CYP199A2 (A) and its NΔ7 variant (B) in <i>E. coli</i> BL21(DE3) using its native redox pair Pux/PuR. ....	41
Figure 2.14: Influence of L-phenylalanine on the C3H (A) and C4H activity (B) of the cytochrome P450 monooxygenase from <i>R. palustris</i> . ....	43
Figure 2.15: Combined <i>in vivo</i> activity of TAL and C3H enzymes in <i>E. coli</i> BL21(DE3). ....	45
Figure 2.16: Co-expression analysis of the TAL enzymes RgTAL, FjTAL and Sesam8 in combination with the C3Hs CYP199A2 (A) and HpaBC (B). ....	46
Figure 2.17: Combined <i>in vivo</i> activity of TAL enzymes with CYP199A2 in <i>E. coli</i> BL21(DE3) using a higher substrate feed. ....	48

Figure 2.18: Combination of <i>RgTAL</i> as PAL and C3H/C4H activities of CYP199A2 (A) and NΔ7-CYP199A2 (B) in <i>E. coli</i> BL21(DE3).....	50
Figure 2.19: <i>In vivo</i> activity of different methyltransferases in <i>E. coli</i> BL21(DE3).G213 at two different expression temperatures. ....	53
Figure 2.20: SDS-PAGE analysis of different methyltransferases in the <i>E. coli</i> BL21(DE3).G213 strain. ....	54
Figure 2.21: Downstream synthetic pathway from caffeic acid towards coniferyl alcohol using plasmid-based enzyme expression in <i>E. coli</i> BL21(DE3).....	56
Figure 2.22: Co-expression analysis of the downstream pathway using the four methyltransferases (1-4) in combination with a 4CL enzyme ( <i>Pc4CL</i> or <i>At4CL</i> ) and <i>ZmCCR</i> and <i>ZmCAD</i> .....	58
Figure 2.23: Increase of substrate feed into the downstream synthetic pathway using plasmid-based enzyme expression in <i>E. coli</i> BL21(DE3). ....	60
Figure 2.24: Synthetic pathway towards <i>p</i> -coumaryl alcohol using a single ePathBrick vector in <i>E. coli</i> BL21(DE3).....	62
Figure 2.25: Synthetic pathway towards coniferyl alcohol from L-tyrosine in <i>E. coli</i> BL21(DE3).G213.....	65
Figure 2.26: Co-expression analysis of the synthetic pathway towards coniferyl alcohol using <i>AtCCoAOMT</i> and CYP199A2 with three different TAL enzymes in <i>E. coli</i> BL21(DE3).G213. ....	66
Figure 2.27: Two-cells one-pot experiment. ....	68
Figure 2.28: Growing cells of <i>E. coli</i> BL21(DE3) with the downstream synthetic pathway using plasmid-based protein expression. ....	71
Figure 2.29: Growing co-culture of <i>E. coli</i> BL21(DE3) using <i>FjTAL</i> and NΔ7-CYP199A2 in the upstream and two different 4CL and OMT combinations in the downstream pathway strain. ....	73
Figure 2.30: Growing cells of <i>E. coli</i> BL21(DE3) with the combined upstream and downstream synthetic pathway using plasmid-based enzyme expression.....	75
Figure 2.31: Growing cells of <i>E. coli</i> BL21(DE3).G213 and <i>E. coli</i> BL21(DE3).G279. ....	78
Figure 2.32: Influence of L-tyrosine on pH of different TB media. ....	82
Figure 2.33: Growing cells of different <i>E. coli</i> BL21(DE3) strains with higher L-tyrosine feed. ....	83
Figure A.1: HpaBC activity towards <i>p</i> -coumaryl alcohol as substrate in <i>E. coli</i> BL21(DE3). ....	156
Figure A.2: Detailed expression of HpaBC from pETM6 in different gene configurations. .	157
Figure A.3: SDS-PAGE and western blot analysis of the cytochrome P450 monooxygenase from <i>R. palustris</i> co-expressed with Pux and PuR at 37 °C. ....	157
Figure A.4: Production of <i>p</i> -coumaric acid from combined <i>in vivo</i> activity of TAL and C3H enzymes in <i>E. coli</i> BL21(DE3) using 1 mM L-tyrosine as substrate. ....	158
Figure A.5: Production of <i>p</i> -coumaric acid from combined <i>in vivo</i> activity of TAL and C3H enzymes in <i>E. coli</i> BL21(DE3) using 3 mM L-tyrosine as substrate. ....	158

Figure A.6: Co-expression analysis of <i>RgTAL</i> in combination with the NΔ7 variant of CYP199A2. ....	159
Figure A.7: SDS-PAGE analysis of <i>TAL</i> , <i>4CL</i> , <i>ZmCCR</i> and <i>ZmCAD</i> expressed from a single pETM6 plasmid. ....	159
Figure A.8: Growing cells of <i>E. coli</i> BL21(DE3) with pETM6_ <i>RgTAL</i> _ <i>At4CL</i> _ <i>ZmCCR</i> _ <i>ZmCAD</i> _m. ....	160
Figure A.9: Synthetic pathway towards coniferyl alcohol from L-tyrosine in <i>E. coli</i> BL21(DE3).G213. ....	160
Figure A.10: Empty vector control of two-cells one-pot experiment. ....	161
Figure A.11: Growing cells of <i>E. coli</i> BL21(DE3) with the combined upstream and downstream synthetic pathway using plasmid-based enzyme expression. ....	161
Figure A.12: SDS-PAGE and western blot analysis of <i>E. coli</i> BL21(DE3) containing the complete synthetic coniferyl alcohol pathway under growing cell assay conditions. ....	162
Figure A.13: SDS-PAGE analysis of <i>E. coli</i> strains containing empty vectors under growing cell assay conditions. ....	162
Figure A.14: Growing cells of <i>E. coli</i> BL21(DE3).G213 and <i>E. coli</i> BL21(DE3).G279. ....	163
Figure A.15: SDS-PAGE and western blot analysis of <i>E. coli</i> BL21(DE3).G279 strains under growing cell assay conditions. ....	164
Figure A.16: SDS-PAGE and western blot analysis of <i>E. coli</i> BL21(DE3).G213 strains under growing cell assay conditions. ....	165
Figure A.17: Intermediates from growing cells of different <i>E. coli</i> BL21(DE3) strains with higher L-tyrosine feed. ....	166
Figure A.18: Growing cells of different <i>E. coli</i> BL21(DE3) strains with cinnamic acid feed. ....	167
Figure A.19: Toxicity of different substrates and intermediates as well as DMSO as control for <i>E. coli</i> . ....	168
Figure A.20: Plasmid maps of the Duet vectors pETDuet (A), pCDFDuet (B) and pRSFDuet (C). ....	169
Figure A.21: Plasmid maps of the ePathBrick vectors pETM6 (A) and pCDM4 (B). ....	170
Figure A.22: DNA and protein ladders. ....	170

**VII.9 List of Tables**

Table 1: <i>In vivo</i> production of phenylpropanoic acids in <i>E. coli</i> . .....	91
Table 2: List of instruments.....	107
Table 3: List of used media, buffers, solutions and reagents and their respective composition. .....	111
Table 4: List of antibodies. ....	116
Table 5: List of synthetic DNA oligonucleotides. ....	116
Table 6: List of all plasmids. ....	119
Table 7: List of all applied strains.....	123
Table 8: Composition of a single PCR reaction mixture. ....	124
Table 9: Composition of both reaction mixtures for restriction. ....	124
Table 10: Composition of a single ligation reaction. ....	125
Table 11: Composition of a single colony PCR reaction mixture. ....	126
Table 12: Program for a colony PCR.....	127
Table 13: Composition of a single RT reaction mixture and incubation procedure. ....	129
Table 14: Composition of one SDS gel. ....	133
Table 15: DNA sequences of cloned genes. ....	151
Table 16: Construction of vectors via restriction and ligation. ....	153
Table 17: Construction of vectors via FastCloning and site-directed mutagenesis. ....	155

## Danksagung

Zunächst möchte ich einen großen Dank an Herrn Prof. apl. Dr. Ulrich Commandeur aussprechen. Für die Möglichkeit, eine Promotion am Institut für Molekulare Biotechnologie anfertigen zu können, und die Erstbetreuung des interessanten Dissertationsthemas bin ich sehr dankbar. Herr Prof. apl. Dr. Commandeur stand mir stets mit Rat und Tat zur Seite, brachte neue Ideen ein und diskutierte konstruktiv projektspezifische Fragestellungen. Für sein Engagement danke ich Ihm sehr.

Mein zweiter großer Dank richtet sich an Herrn Prof. Dr. Lars Blank für seine Bereitschaft mein Zweitbetreuer zu sein. Zudem danke ich Ihm für die Möglichkeit, die HPLC am Lehrstuhl für Angewandte Mikrobiologie für Messungen benutzt haben zu können.

Des Weiteren danke ich Herrn Prof. Dr. Jan Marienhagen für die Übernahme des dritten Prüfers und Herrn Prof. Dr. Lothar Elling für die Übernahme des Vorsitzenden der Promotionskommission.

Ich möchte mich besonders und äußerst herzlich bei meiner Kollegin und betreuenden Post-Doc Dr. Christina Dickmeis bedanken. Tini hat mir stets bei auftretenden Fragen geholfen sowie Tipps und Ideen eingebracht. Ich konnte mit ihr immer Problemstellungen und Schwachstellen aber auch Methodenprotokolle und Experimente professionell diskutieren. Für dein freundliches Auftreten danke ich Dir sehr und ich wünsche Dir das Beste für die Zukunft.

Ich danke außerdem unseren LignaSyn Projektpartnern für die gute Zusammenarbeit, spannenden Diskussionsrunden und den regen Austausch über die letzten Jahre. In diesem Sinne gilt mein Dank Prof. Vlada Urlacher, Prof. Stephan Lütz, Dr. Guido Jach, Dr. Katrin Schullehner, Dr. Marco Girhard, Stefan, Davide und Andrea.

Dann danke ich dem Fraunhofer Institut für Molekularbiologie und angewandte Ökologie IME und dem Institut für Biologie III (Biochemie und Molekularbiologie der Pflanzen) für die Möglichkeit, ihre HPLCs für eigenständige Messungen zu benutzen. Diesbezüglich geht ein besonderer Dank an Alexander Croon, der mir nicht nur sehr viel an der HPLC gezeigt hat, sondern mit dem man sich auch super unterhalten konnte. Ebenso haben mich David Spencer und Alexander Beesley (Institut für Biologie III) an Ihrer HPLC betreut. Danke euch beiden und der ganzen Bio3 für die freundliche Atmosphäre und das gelegentliche Aushelfen, wenn ich mal wieder einen Schlüssel benötigte. Zuletzt bedanke ich mich bei Dr. Isabel Bator (Institut für Biologie IV, iAMB) für die Einführung und Betreuung an der HPLC in Ihrem Verantwortungsbereich.

Auch besondere Freude hat mir die Zusammenarbeit mit meinen Labornachbarn Malin, Astrid, Kathrin und Katja bereitet. Danke für eure offene Art und die vielen gemeinsamen Mittagessen in der Mensa.

Außerdem danke ich der gesamten Bio7 für die gemeinsame Zeit, die gegenseitige Unterstützung im Laboralltag und die freundliche Atmosphäre. Danke an Juliane, Ingrid, Karo, Ibrahim, Kathi, Louisa und allen Ehemaligen am Institut über die Jahre insbesondere Selin, Hormoz, Esther, Johannes, Rebecca, Christian und Thomas.

Zu guter Letzt geht ein Riesendank an meine Familie und Freunde, die mich moralisch sehr unterstützt haben, und an David. Danke, dass Du immer ein offenes Ohr hast und für mich da bist.

## **Declaration**

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that as required by these rules and conduct I have fully cited and referenced all material and results that are not original to this work.

---

Anna Christiane Kohl, Aachen 31.03.2023





## Curriculum vitae

Name: Anna Christiane Kohl

Date of birth: 13.04.1992

Place of birth: Vienna, Austria

### Education

2002-2011: Marion-Dönhoff-Gymnasium Mölln, Germany

2011-2015: Undergraduate studies: Bachelor of Science in Biochemistry, University of Greifswald

Bachelor thesis: *Synthesis and spectroscopic investigation of a novel G-quadruplex binding indoloquinoline derivative*, Department of Analytical Biochemistry, University of Greifswald

2014-2017: Graduate studies: Master of Science in Biochemistry, University of Greifswald

Master thesis: *Optimization of whole-cell multienzyme cascade reactions*, Department of Biotechnology & Enzyme Catalysis, University of Greifswald

2017-2023: PhD in molecular biotechnology, Institute for Molecular Biotechnology, RWTH Aachen University

Working titel: *Synthetic phenylpropanoid pathway for the in vivo production of coniferyl alcohol in Escherichia coli*

This work was part of the LignaSyn project funded by the Federal Ministry of Education and Research.