Enzymatic Synthesis of
Plant Phenylpropanoids

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät II
Chemie, Physik und Mathematik

der Martin-Luther-Universität
Halle-Wittenberg

vorgelegt von

Frau

Diplom-Lebensmittelchemikerin

Susann Herrmann
geb. am 20. Dezember 1988 in Cottbus

Die vorliegende Arbeit wurde im Zeitraum von Juli 2013 bis Juni 2017 am Leibniz-Institut für Pflanzenbiochemie in der Abteilung Natur- und Wirkstoffchemie (Abteilungsleiter Prof. Dr. Ludger A. Wessjohann) angefertigt.

Tag der Verteidigung: 17. April 2018

1. Gutachter: Professor Dr. Ludger A. Wessjohann

2. Gutachter: Professor Dr. Jörg Pietruszka

Hiermit widme ich diese wissenschaftliche Arbeit meinem Opili, Lothar Riemer, der die Fertigstellung dieser leider nicht mehr miterleben konnte. Jedoch war er stets meine größte Motivation… Du fehlst!

„Leicht ist das Leben für keinen von uns.

Doch was nützt das, man muss Ausdauer haben

und vor allem Zutrauen zu sich selbst.

Man muss daran glauben, für eine bestimmte Sache begabt zu sein,

und diese Sache muss man erreichen, koste es, was es wolle."

MARIE CURIE (1867-1934)

Danksagung

Table of contents

Abbreviations

δ	chemical shift
$(NH_4)_2SO_4$	ammonium sulfate
(v/v)	concentration (volume per volume)
(w/v)	concentration (weight per volume)
4CL	4-coumarate:CoA ligase
4HPA3H	4-hydroxyphenylacetate 3-hydroxylase
approx.	Approximately
ADP	adenosine diphosphate
AMP	adenosine monophosphate
ATP	adenosine triphosphate
APS	ammonium persulfate
At4CL2	4-coumarate CoA-ligase from *Arabidopsis thaliana*
BM3	CYP102A1 from *Bacillus megaterium*
C3H	*p*-coumarate 3-hydroxylase
C4H	cinnamate 4-hydroxylase
CAD	cinnamyl alcohol dehydrogenase
calc	calculated
CCoAOMT	caffeoyl-CoA *O*-methyltransferase
CCR	cinnamoyl-CoA reductase
CD	circular dichroism
C_3D_6O	deuterated acetone
CD_3OD	deuterated methanol

CDCl$_3$	deuterated chloroform
CoA	Coenzyme A
COMT	caffeic acid/5-hydroxyconiferaldeyhde O-methyltransferase
DMAPP	dimethylallyl pyrophosphate
dNTPs	desoxynucleotides
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
e.g.	for example
ESI	electron spray ionization
F5H	ferulate 5-hydroxylase
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
fw	forward (primer)
GMO	genetically modified organisms
G6P	glucose-6-phosphate
G6PDH	glucose-6-phosphate dehydrogenase
HCD	higher-energy collisional dissociation
HCl	hydrogen chloride
HCT	p-hydroxycinnamoyl-CoA: quinate shikimate p-hydroxycinnamoyltransferase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His$_6$-tag	polyhistidine-tag
HpaC	flavin reductase from *Escherichia coli*

HPLC	high performance liquid chromatography
Hz	Hertz
IMAC	Immobilized Metal Ion Affinity Chromatography
IPP	isopentenyl pyrophosphate
IPTG	isopropyl β-D-1-thiogalactopyranoside
K_2HPO_4	dipotassium phosphate
KCl	potassium chloride
KOAc	potassium acetate
LB	lysogeny broth
LeOPR1	12-oxophytodienoate reductase from *Lycopersicon esculentum*
LlCCR	cinnamoyl-CoA reductase from *Leucaena leucocephala*
m/z	mass-to-charge ratio
MeCN	acetonitrile
MeOH	methanol
$MgCl_2$	magnesium chloride
MpCAR	carboxylic acid reductase from *Mycobacterium phlei*
MS	mass spectrometry
Na_2SO_4	sodium sulfate
NaCl	sodium chloride
NAD^+	nicotinamide adenine dinucleotide, oxidized form
NADH	nicotinamide adenine dinucleotide, reduced form
$NADP^+$	nicotinamide adenine dinucleotide phosphate, oxidized form

NADPH	nicotinamide adenine dinucleotide phosphate, reduced form
NaOH	sodium hydroxide
NiCAR	carboxylic acid reductase from *Nocardia iowensis*
NiPPTase	phosphopantetheinyl transferase from *Nocardia iowensis*
NMR	nuclear magnetic resonance
OD_{600}	optical density at 600 nm
OYE	old yellow enzyme
PAL	phenylalanine ammonia lyase
PAP	polyphosphate-AMP phosphotransferase
PFOMT	phenylpropanoid and flavonoid *O*-methyltransferase
pHBH	*para*-hydroxybenzoate hydroxylase
PPK	polyphosphate kinase
ppm	parts per million
PrnF	putative flavin:NAD(P)H reductase from *Pseudomonas protegens*
Qq-TOF	tandem quadrupole/time-of-flight
rev	revers (primer)
Rt	retention time
S. cerevisiae	*Saccharomyces cerevisiae*
S.O.C.	super optimal broth with catabolite repression
SAH	*S*-adenosyl-L-homocysteine
SAHH	*S*-adenosyl-L-homocysteine hydrolase
SAHN	*S*-adenosyl-L-homocysteine nucleosidase
SAM	*S*-adenosyl methionine

SAMS   *S*-adenoylmethionine synthase

ScFDH   formate dehydrogenase from *Saccharomyces cerevisiae*

SDS   sodium dodecyl sulfate

SDS-PAGE   sodium dodecyl sulfate polyacrylamide gel electrophoresis

TaCAD1   cinnamyl-alcohol dehydrogenase from *Triticum aestivum*

TEMED   *N,N,N',N'*-tetramethylethylenediamine

TMS   tetramethylsilane

Tris   tris(hydroxymethyl)aminomethane

UGT   UDP-glucosyltransferase

UV   ultraviolet (light)

Vis   visible (light)

ZY   N-Z-amine and yeast extract

Abstract

Plant secondary metabolites constitute an enormous chemical diversity and are divided into three chemically distinct main groups: terpenoids, alkaloids and phenolic compounds (particularly phenylpropanoids and polyketides). Phenylpropanoids are the most common types of natural products. They are biosynthetically formed from the amino acid L-phenylalanine by deamination catalyzed by L-phenylalanine ammonia lyase. The simplest examples for phenylpropanoids are the hydroxycinnamic acids, such as ferulic acid, and the monolignols, such as coniferyl alcohol. More complex phenylpropanoids comprise e.g. flavonoids, isoflavonoids and stilbenes. Phenylpropanoids have a wide range of important functions in plants, including as structural components (such as lignin), protectants against biotic and abiotic stresses and pigments (particularly anthocyanins).

In particular, monolignols are plant-derived aromatic compounds with many potential applications as they are building blocks for biologically active compounds like lignans, flavor and fragrances or precursors for pharmaceuticals and cosmetics. Even though monolignols are present in high amounts in lignin, its depolymerization and the isolation of the individual monolignols from this renewable resource is difficult due to its complex and highly crosslinked structure. Therefore, the use of enzymes could be an advantage for a specific synthesis of the individual monolignols. In this work, a route for the preparation of various methylated and hydroxylated cinnamaldehyde and cinnamyl alcohol derivatives was developed, starting from their corresponding acids. On the one hand, a flavin-dependent monooxygenase from *Escherichia coli* (*E. coli*) was optimized by rational design for hydroxylation of bulky substrates. On the other hand, a system consisting of two enzymes could be replaced by an enzyme with two domains that allow the formation of cinnamaldehyde derivatives.

Furthermore, a cell-free modular system for the synthesis of phenylpropanoids was developed. As an example, the synthesis of vanillin – the world`s most important flavoring agent – was constructed *in vitro*. A set of eight enzymes (including enzymes for the regeneration of cofactors) was used for a continuous conversion of *p*-coumaric acid to vanillin. This method was used for the successful formation of further phenylpropanoids (e.g. hydroxylation of naringenin to eriodictyol) as well.

Zusammenfassung

Pflanzliche Sekundärmetabolite stellen eine enorme chemische Vielfalt dar und sind in drei verschiedene Hauptgruppen unterteilt: Terpenoide, Alkaloide und phenolische Verbindungen (insbesondere Phenylpropanoide und Polyketide). Phenylpropanoide sind die am weitesten verbreiteten Arten von Naturprodukten. Sie werden biosynthetisch aus der Aminosäure L-Phenylalanin durch Deaminierung geformt, was durch eine L-Phenylalanin-Ammoniak-Lyase katalysiert wird. Die einfachsten Beispiele für Phenylpropanoide sind die Hydroxyzimtsäuren, wie Ferulasäure, und die Monolignole wie Coniferylalkohol. Komplexere Phenylpropanoide umfassen z.B. Flavonoide, Isoflavonoide und Stilbene. Phenylpropanoide weisen ein breites Spektrum an wichtigen Funktionen in Pflanzen auf, unter anderem als Strukturkomponenten (wie Lignin), Schutz vor biotischen und abiotischen Belastungen und als Farbstoffe (speziell Anthocyanine).

Insbesondere Monolignole sind pflanzliche aromatische Verbindungen mit vielen potentiellen Anwendungsmöglichkeiten, da sie als Bausteine für biologische Wirkstoffe, wie Lignane, Geschmacks- und Duftstoffe oder als Grundbausteine für Pharmazeutika und Kosmetika verwendet werden können. Obwohl Monolignole in hohen Mengen in Lignin vorliegen, ist die Depolymerisation und die darauf folgende Isolierung der einzelnen Monolignole aus diesen erneuerbaren Ressourcen aufgrund ihrer komplexen und hochvernetzten Struktur schwierig. Daher könnte die Verwendung von Enzymen ein Vorteil für eine spezifische Synthese der einzelnen Monolignole sein. In dieser Arbeit wurde aus ihren entsprechenden Säuren ein Weg zur Herstellung von verschiedenen methylierten und hydroxylierten Zimtaldehyd- und Zimtalkoholderivaten entwickelt. Einerseits wurde eine Flavin-abhängige Monooxygenase aus *Escherichia coli* (*E. coli*) mittels rationalem Design für die Hydroxylierung von sperrigen Substraten optimiert. Weiterhin konnte ein bestehendes System aus zwei Enzymen durch ein alternatives Enzym mit zwei Domänen ersetzt werden, welches die Bildung von Zimtaldehydderivaten ermöglicht.

Ferner konnte ein zellfreies Modulsystem zur Synthese von Phenylpropanoiden entwickelt werden. Als Beispiel wurde die Synthese von Vanillin - dem weltweit wichtigsten Aromastoff - *in vitro* rekonstruiert. Ein Satz von acht Enzymen (einschließlich Enzymen für die Regeneration von Cofaktoren) wurde für eine kontinuierliche Umsetzung von *p*-Cumarsäure zu Vanillin verwendet. Diese Methode konnte auch für die erfolgreiche Darstellung weiterer Phenylpropanoide (z. B. Hydroxylierung von Naringenin zu Eriodictyol) eingesetzt werden.

List of Figures

Figures in the Appendix

Figure D 1: SDS-PAGE of cells expressing 4HPA3H (lane 1), the variants Y301F (lane 2), S462A (lane 3), M293P (lane 4), I157V (lane 5), Y301L (lane 6), Y301I (lane 7), Y301F/S462A (lane 8), Y301L/S462A (lane 9), Y301I/S462A (lane 10),

List of Tables

Tables in the Appendix

Overview of relevant substances

number (odd)	name and structure	number (even)	name and structure
1	*p*-coumaric acid	2	caffeic acid
3	ferulic acid	4	5-hydroxyferulic acid
5	resveratrol	6	piceatannol
7	naringenin	8	eriodictyol
9	*p*-hydroxybenzoic acid	10	3,4-dihydroxybenzoic acid
11	2-hydroxycarbazol	12	2,3-dihydroxycarbazole
13	umbelliferone	14	esculetin
15	rheosmin	16	4-(3,4-dihydroxyphenyl)-butan-2-one

number (odd)	name and structure	number (even)	name and structure
17	sinapic acid	18	*p*-coumaraldehyde
19	caffeyl aldehyde	20	coniferaldehyde
21	5-hydroxyconiferaldehyde	22	sinapaldehyde
23	*p*-coumaryl alcohol	24	caffeoyl alcohol
25	coniferyl alcohol	26	5-hydroxyconiferyl alcohol
27	sinapyl alcohol	28	vanillin
29	4-(4-hydroxyphenyl)-3-buten-2-one ≙ dehydrorheosmin	30	4-(4-hydroxy-3-methoxyphenyl)-3-buten-2-one ≙ dehydrozingerone
31	zingerone		

1   Introduction

1.1   Natural products

Plants produce a broad range of organic compounds which are not directly involved in growth, development or reproduction of plants, but have important functions in plant defense and signaling, or serve as pigments or fragrances.[1,2] They also protect plants against herbivores (insects and mammals), plant competitors and abiotic stresses, e.g. UV light, ozone and herbicides.[3] These compounds, known as secondary metabolites, originate from primary metabolites (Fig. 1.1), such as common plant metabolic intermediates, e.g. phytosterols, acyl lipids, nucleotides, amino acids and organic acids, through methylation, hydroxylation and glycosylation in biochemical pathways.[1-3] Primary metabolites perform metabolic roles that are essential and usually obvious.[1,2] They are essential for cell survival and dissemination.[3]

The synthesis of a specific secondary metabolite is not observed in all species of the plant kingdom, can vary from species to species and comprise a diverse array of complex chemical structures.[1,4] For a long time the importance of secondary metabolites was not fully understood. Due to progress in research, technology and in genome sequencing (e.g. of crop plants such as barley, rice and tomato), many questions about the role of these substances in the evolution of the main grown species were enlightened in the last decades.[1] The secondary metabolism in plant plays a major role for its survival in the environment.[5]

Until today, more than two hundred thousand different secondary metabolites have been identified in plants.[3] Based on their biosynthetic origins, plant natural products can be divided into at least three chemically distinct main groups: terpenoids, alkaloids and phenolic compounds (with phenylpropanoids and polyketides as the dominant subgroups).[1,2]
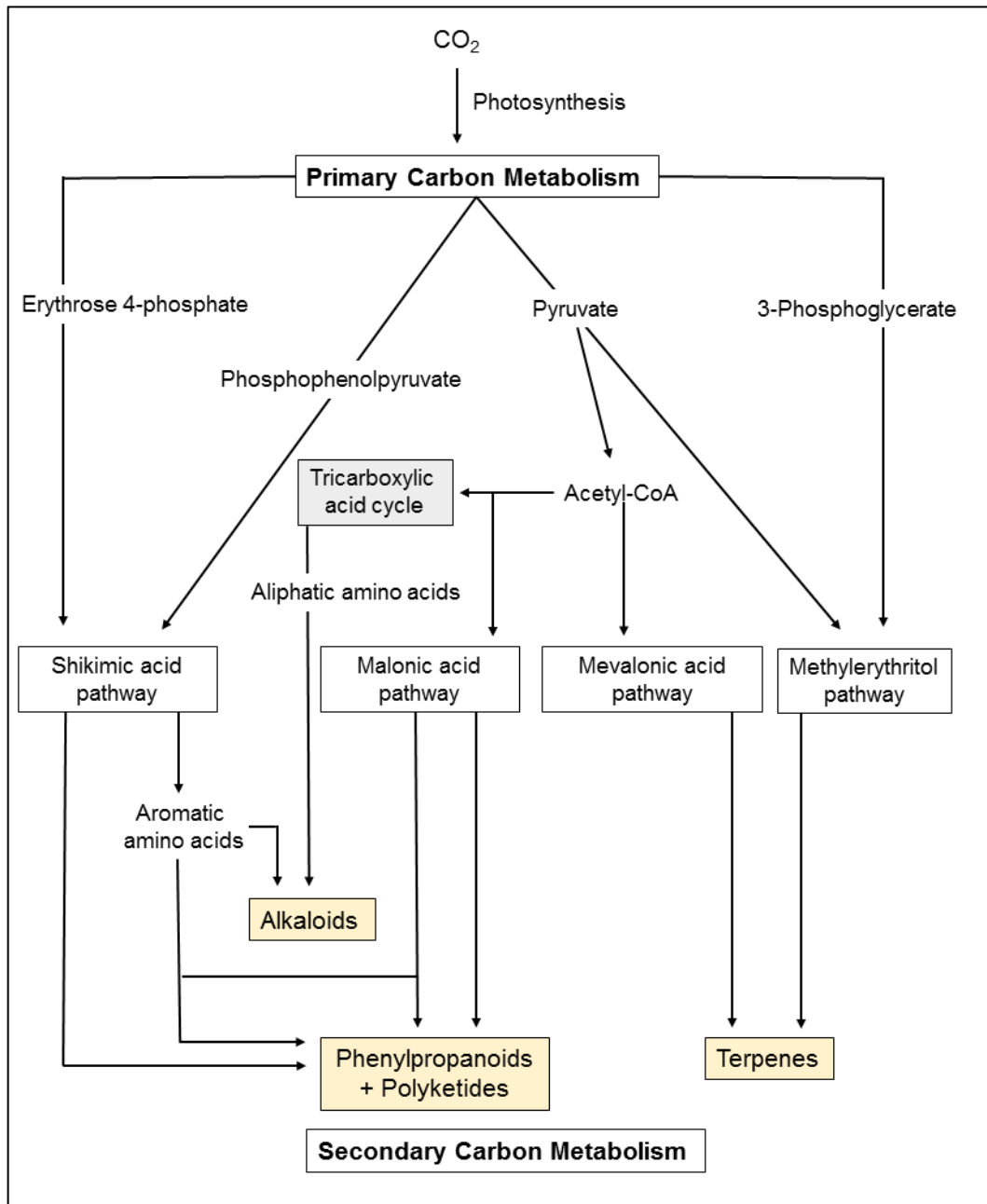
Figure 1.1: General overview of the biosynthetic pathways involved in the biosynthesis of the three main classes of secondary metabolites starting with products of the primary metabolism. Secondary metabolites are presented in yellow boxes. The pathways of the secondary metabolism are presented in frames, the grey boxes presents a part of the primary metabolism. Figure is modified according to Ncube et al.[6]

Terpenoids, also known as isoprenoids, are the most structurally diverse group of plant natural products containing 50.000 identified substances.[1,2,6] All terpenoids are constructed by repetitive fusion of branched five-carbon units based on the isoprene skeleton.[1] The isoprene monomers generally are described as isoprene units because a thermal decomposition of many terpenoid substances yields the alkene gas isoprene as one product. Furthermore with suitable chemical conditions isoprene units can be polymerized into oligomers.[2] Terpenoid biosynthesis in plants, animals and microorganisms involves similar classes of enzymes, but there are important differences between the corresponding pathways.[2] Basically, plants produce a much wider variety of terpenoids, a difference which is reflected in the complex organization of plant terpenoid biosynthesis at the tissue, cellular and genetic levels.[2] They are synthesized from primary metabolites by two independent pathways (Fig. 1.1): the cytosolic mevalonate pathway and the plastidal methylerythriol pathway, which both produce isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP).[1,6] The presence of acetyl-CoA (pathway of IPP) and pyruvate as well as glyceraldehyde 3-phosphate (pathway of DMAPP) are essential for the synthesis of terpenoids in plants.[1,6] Because of their chemical diversity, terpenoids serve numerous biochemical functions that include, e.g. electron transport chains, photosynthetic pigments, hormones and attractants for pollinators.[6] Many plant terpenoids are toxins and serve as repellants to herbivorous animals. Furthermore there are some volatile compounds (e.g. menthol in mint or limonene in lemon) with a specific odor appearing as attractants.[1,2] These compounds can be isolated as essential oils which are important ingredients in the pharmaceutical, cosmetic and food industry.[1]

Plant extracts have been used as ingredients in potions and poisons for a long time in human history.[7] The principal similarity of the enormous class of **alkaloids** is the presence of a basic nitrogen atom at any position in the molecule, which does not include nitrogens in an amide or peptide bond.[6] Many of these alkaloids have a complex chemical structure with multiple asymmetric centers, which complicate their structural elucidation. This makes studies of the biosynthesis of alkaloids quite difficult. For example, nicotine was discovered in 1828, but its structure was not known until it was synthesized in 1904. Likewise nicotine, the structure of morphine was not completely elucidated until 1952, 150 years after Sertürner isolated it from *Papaver somniferum* (the opium poppy).[7,8] With the availability of radioactively labeled organic molecules, alkaloid biosynthesis became available for examination of biosynthetic hypothesis.[2] Feeding

experiments with labeled precursors demonstrated that alkaloids are in most cases formed from L-amino acids (e.g. tryptophan, tyrosine or phenylalanine) as starting materials (Fig. 1.1), although some purine-derived alkaloids are also known.[2,6] In general, alkaloids are often substances with distinct biological activity in low doses and protect plants from a variety of herbivores.[9] Owing to their potent biological activity, many of the approximately 12.000 known alkaloids possess important pharmacological activity and have been exploited as pharmaceuticals, stimulants, narcotics and poisons.[6,7]

1.2 Phenylpropanoids

The third large group of natural products is the group of phenylpropanoids. As plants originated in an aquatic environment, their successful evolutionary adaption to land was accompanied by formation of "plant phenolic" compounds.[2,6] This leads to the assumption that phenylpropanoids are the most widespread types of natural products occurring in nature.[10] Although the majority of these substances assumed cell wall structural roles, a vast array of nonstructural constituents is also known. These substances have various essential roles as defendants, determining certain distinguishing features of different woods and barks, in establishment of coloration and considerably contribute to the flavor of foods (taste and odor).[2]

Phenylpropanoids are generally characterized as aromatic metabolites that possess one or more "acidic" hydroxyl group(s) attached to the aromatic phenyl ring, which can be modified to yield functional derivatives such as esters, methyl ethers and glycosides.[2,6] Phenylpropanoid biosynthesis is developmentally activated in specific tissues and cell types, but can also be triggered in response to environmental stress such as wounding, pathogen infection or irradiation with ultraviolet light.[11] The phenylpropanoid pathway is one of the most important metabolic pathways in plants. Primarily, the phenolic compounds are synthesized by products of the shikimic acid pathway or by products of the malonic acid pathway (Fig. 1.1).[1,6] Most of the phenolic compounds are derived from phenylalanine.[1,6] The name "phenylpropanoids" originates from the aromatic phenyl ring and the three-carbon tail of this amino acid.[12] L-Phenylalanine, product of the shikimic acid pathway, is then bioconverted into cinnamic acid by a non-oxidative deamination process, which is mediated by the key enzyme phenylalanine ammonia lyase (PAL). PAL is one of the most studied enzymes in plant secondary metabolism (Fig. 1.2).[1,6] Subsequently, the phenylpropanoid core can be modified by the actions of reductases, hydroxylases and

O-methyltransferases to get access to various substitution patterns.[4] One of the most involved enzymes is the cinnamate 4-hydroxylase (C4H) which is a member of the ubiquitous enzyme family of oxygenases known as the cytochrome P450 hydroxylases, hydroxylate cinnamic acid into *p*-coumaric acid. This large and evolutionary old gene superfamily is characterized by the characteristic absorbance of the catalytic iron-containing heme cofactors at 450 nm.[13,14] Plant P450 enzymes play important roles in phenylpropanoid metabolism leading e.g. to the production of lignin and flavonoids. The enzyme 4-coumarate CoA-ligase (4CL) catalyzes the formation of CoA esters of the produced hydroxycinnamic acids by activation of the carboxyl group (formation of a thioester bond with CoA).[11,15] These activated intermediates, e.g. *p*-coumaroyl CoA, are used in the biosynthesis of diverse phenylpropanoids (Fig. 1.2).[11,16]
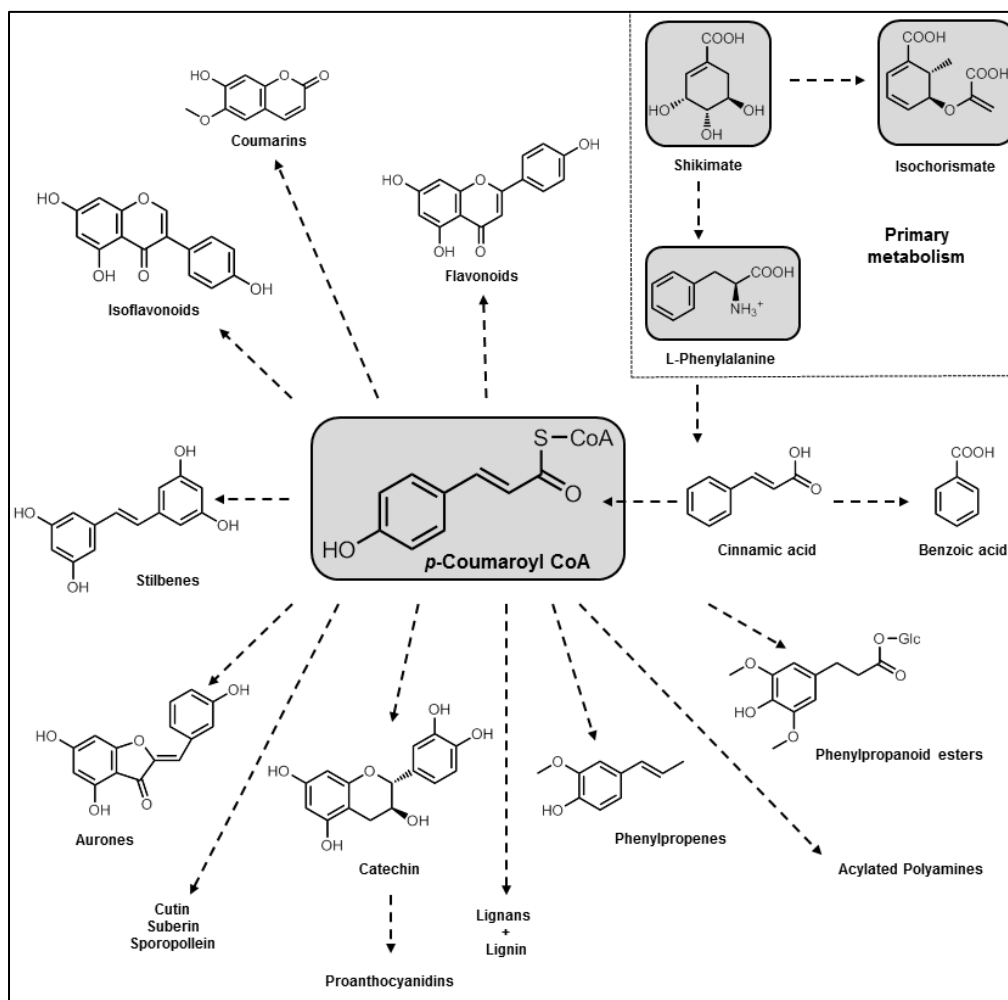
Figure 1.2: Diversity of phenylpropanoids based on the phenylpropanoid pathway starting with L-phenylalanine. Metabolites of the shikimate pathway and 4-coumaroyl CoA as the central metabolites are depicted in grey. Figure is modified according to Vogt (2010).[16]

Generally phenylpropanoids are a large group of compounds, which can be divided into five subgroups: coumarins, flavonoids, phenolic acids, tannins and lignins. Among these, lignins and flavonoids are the most ubiquitous phenolic compounds in plants.[1]

Coumarins are derived from 1,2-benzopyrones and widespread in higher plants, especially in grasses, orchids and umbellifers.[1,10,17] Coumarins receive attention for their diverse bioactivity, e.g. they have antibacterial (umbelliferon and scopoletin), antitumor (mammein and neomammein) and vasodilatory (dicumarol and warfarin) activities.[17,18] Despite the pharmaceutical importance of coumarins, less information is available regarding the biosynthesis of the simplest coumarin skeleton.[17,19] Nevertheless, it is known that the *ortho*-hydroxylation (P450 reactions) of hydroxycinnamic acids is a crucial step for the biosynthesis of simple coumarin, umbelliferone and other hydroxylated coumarins in plants.[17]

Constructed of a flavone skeleton, **flavonoids** are biosynthetically derived from intermediates from the shikimic acid and the malonate pathway by polyketide type III enzymes (Fig. 1.1).[6,10] The first committed step of the flavonoid pathway is catalyzed by a chalcone synthase leading to the condensation of three molecules of acetate-derived malonyl-CoA and one molecule of *p*-coumaryl-CoA (Fig. 1.2) to generate a tetrahydroxychalcone. This widespread and diverse class of low molecular weight phenolic compounds can occur as monomers, dimers and higher oligomers. They are constituents of a variety of plant parts, e.g. leaves, fruits, flowers and roots.[6] Differences in basic substitution patterns define several subclasses, such as flavonols, chalcones, catechins and anthocyanidins. Their diverse properties and roles in plants are related to their structural and chemical diversity.[6] Many flavonoids provide plants with bioactive compounds that were used against viral infections or exhibit antitumor and have general anti-inflammatory activity.[10]

Phenolic acids are characterized by being formed by a benzene ring, a carboxylic acid group and one or more hydroxyl groups at the aromatic core. This structural features give this compounds antioxidative properties (as used in foods or for the treatment and prevention of diseases).[1] In contrast to the synthesis of other phenylpropanoids, the formation of phenolic acids appears complex and is still fragmentary.[16]

Tannins along with lignins are considered as one of the most important groups of secondary metabolites in the defense of plants. Tannins are phenolic compounds with a complex and highly variable chemical structure and are divided into two categories: condensed tannins, compounds formed by the addition of a flavonoid

constituent of woody plants, and hydrolysable tannins, polymers formed from phenolic acids and simple sugars.[1,6] Both classes form complexes with proteins (including enzymes) and biopolymers such as cellulose and hemicellulose.[6] Tannins have been exploited in the discovery of medicinal agents due to their antibiotic, antifeedant or biostatic effects.[6] They are used as constituents of drugs because of their interaction with enzymes (proteins) within cell systems.[6]

The last large group of phenylpropanoids are **lignins**[a], which are very important for plants as they are main constituents of a stable cell wall.[1] Lignin is a durable polymer, which is linked in three dimensions and can be formed from three different monomers (*p*-hydroxyphenyl (H-unit), guaiacyl (G-unit) and/or syringyl monomers (S-unit)) via oxidative coupling that is probably catalyzed by both peroxidases and laccases.[11,20] However, the composition of these monomers is not simply random, it varies according to plant species, cell type and stage of tissue development.[11,20] While dicotyledonous angiosperm lignin mostly consists of G- and S-units, softwood gymnosperm lignin does not possess S-units and is highly condensed. In general, gymnosperm lignin is less suitable for chemical pulping than hardwood lignins.[20]

1.3   Green chemistry and the importance of biocatalysis

Green chemistry is defined by the Environmental Protection Agency as "the design of chemical products and processes that reduce or eliminate the generation of hazardous substances".[21] Furthermore, green chemistry is concerned with the prevention of pollution by waste minimization and the avoidance of toxic substances in production and application of chemical products to become more eco-friendly.[22,23] Since in recent years the tendency shifted from coal- and petroleum-based chemicals, an enhanced pursuit for innovative utilization of bio-based products becomes of more interest for society, politics and economics.[7] In the European Union, new chemical legislations were introduced in 2007 to improve the levels of protection of human health and environment. This regulation, so called REACH (registration, evaluation and authorization of chemicals[24]), demands registration and safety testing of all produced or imported chemicals.[23] This leads to the assumption that chemists and chemical engineers should produce greener and more sustainable chemical processes.[25] In 1998, Anastas and Warner published the 12 Principles of Green Chemistry, in which the overall guideline is a 'benign by design' (Fig. 1.3[b]).[26] This principles suggest a

---

[a] Further information about lignin and lignin biosynthesis will be given in chapter II.
[b] source illustration: http://swapsushias.blogspot.de/2013/10/green-chemistry-sustainable-chemistry.html

decrease in the environmental impact of a chemical product by considering aspects of its entire life cycle, starting from raw material to product use and destination.[23]

Point nine of these principles (Fig. 1.3), the principle of catalysis, can be described as the heart of green chemistry philosophy. Catalysts are molecules that reduce the scale of the energy barrier essential to be overcome for a substance to be transformed into another chemically.[27] The catalytic version of a process generates less waste than the version using a stoichiometric amount of activating reagents. The development of efficient greener catalytic alternatives represents one of the most investigated areas in green chemistry.[22,28]
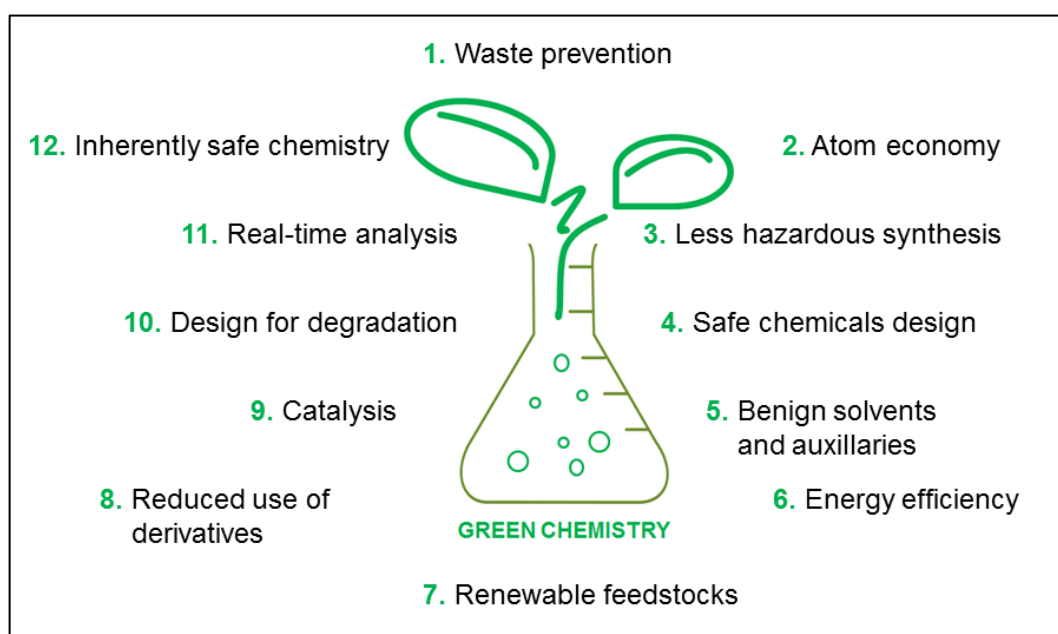
Figure 1.3: The twelve principles of green chemistry published by Anastas and Warner 1998. Figure was modified according to Cioc et al.[28]

One approach is the use of engineered microorganisms for catalytic conversions. With new improvements in biological methods and tools, microorganisms can be engineered for the production of many chemicals based on inexpensive renewable materials such as sugars.[29] The catalyst (an enzyme) itself is derived from renewable resources, is biocompatible, biodegradable and basically non-hazardous and non-toxic as well.[22] In the context of Green Chemistry, biocatalysis can display many attractive features: reactions often are performed in water under mild conditions (physiological pH, ambient temperature and pressure), they are compatible with each other and not restricted to their natural role, as many enzymes can catalyze a broad spectrum of reactions.[22,25,30] Further

advantages of enzymes are the high chemo-, regio- and stereoselectivity and that protection and deprotection sequences are not necessary which are often required in traditional chemical synthesis.[22,25,31] Nevertheless, enzymes possess disadvantages as well. Although they are extremely flexible by accepting non-natural substrates, enzymes are bound to their natural cofactors which serve e.g. as molecular shuttles of redox equivalents (e.g. flavin or NAD(P)H) or as a storage for chemical energy (ATP). The majorities of these cofactors are relatively unstable and are expensive to be used in stoichiometric amounts. As a result, cofactor recycling is an important requirement for biocatalytic approaches.[30]

Genetic engineering can be utilized to increase the yields of flavor and fragrances and other natural products of commercial value (pharmaceuticals, insecticides, fungicides and many others).[2] Metabolic engineering of microorganisms for plant natural products production is a powerful approach to gain access to bioactive compounds which accumulate only at low quantities in their naturally producing plant or are too complex to be synthesized by total chemical synthesis.[4]

Biotechnological approaches can selectively increase the amounts of defense compounds in crop plants, thereby reducing the need for costly and potentially toxic pesticides.[2] However, if a new product should be launched on the market, the achievement of the new bio-based product must equal that of the existing and well established chemical ones.[7] But it is not surprising that in the last 10 years, biocatalysis has been more and more integrated into established organic synthesis.[22] The attributes of enzymes have resulted in a variety of applications, where high reaction selectivity on complex substrates is critical. One example is the preparation of semisynthetic penicillin catalyzed by a penicillin amidase.[31]

1.4 Biotechnology for production of specific natural products

Besides their roles as attractors for pollinators or their defense against predators and diseases in plants, secondary metabolites also have important traits enhancing the quality of our food plants (e.g. taste, color and scent) and of ornamental plants.[5] Additionally, secondary metabolites such as alkaloids, flavonoids, steroids and terpenoids in medicinal and aromatic plants are reliable source for drug discovery and drug development or are commercially applied as e.g. dyes or flavors and fragrances (Fig. 1.4).[5,6,18] The majority of these aromatic compounds are produced by chemical synthesis based on petroleum as raw material.[32]

Figure 1.4: Illustration of different specific phenylpropanoids found in fruits, vegetables, herbs and spices. Most of them are of high interest for industry because of their application as flavors, fragrances or bitter-masking compound. Figure was created by Dr. Martin Dippe.

1.4.1 Metabolic engineering in plants

Some natural products can be isolated and purified from plant materials in large amounts by techniques such as simple distillation, steam distillation or extraction with addition of an acid or base.[5,9] One example are hop bitter acids with a world market of ca. 7,500 tons per year. In many cases, the native productivity of natural products by plants is not satisfying, because of its limitation to species or their production limited to a particular developmental stage or difficult cultivation conditions.[5] Further problems for isolation and purification of target molecules are the low abundance and season or region-depending sourcing.[18]

For pharmaceutical processes the adherence to GMP-rules (good manufacturing practice) is important. Therefore, large efforts have been put into plant cell cultures as a possible production system. Nevertheless, despite extensive studies on optimization of growth and production media, the production of pharmaceuticals from cell culture is generally low. Therefore, research has focused more and more on the regulation of biosynthesis as a starting point for metabolic engineering in these cultures.[5] Metabolic engineering is the directed enhancement of cellular characteristics through the modification of specific biochemical reactions or the introduction of novel reactions by recombinant DNA technologies.[33] In this case, biotechnology employs the tools of genetic engineering to increase conventional methods of producing important food ingredients, drugs and biopolymers.[34] Thereby, the range of possible substrates should be broadened.[34] This may be used for production of special compounds in cell culture or in the plant itself.[5]

In principle, plant cell suspension cultures can be obtained from any plant, although some plants are easier than others. But in comparison to microorganisms, plant cells are much bigger mainly due to the presence of large vacuoles. Even if the technology for production of secondary metabolites is possible, the economy of a plant cell biotechnological production is the other important question. For a commercial process the productivity of the cells has to be improved.[5]

The most common approach for enhancement is screening and selection for high producing cell lines as well as optimization of growth and production media. These approaches are common for microbial production systems, but they have been successful for plant cells as well.[5] The knowledge about fundamental research on plant biochemistry can be expanded by combinatorial biosynthesis and retro-biosynthesis based on the massive and scattered enzyme data that has been accumulated so far.[18]

1.4.2 Metabolic engineering in bacteria and yeast

In general the replacement of traditional chemical synthesis with a bio-based production leads to reduced environmental stress, e.g. in terms of energy use and emission. However, the primary point for the chemical industry is the production of chemicals that have either better properties than traditional chemicals or chemicals with new applications.[35] In some cases the isolation or purification of natural products from plant sources does not result in quantities large enough for the market demands (e.g. vanillin from *Vanilla* pods[36]). Furthermore, the processes are expensive, wasteful and sometimes environmentally unsafe due to the use of problematic solvents.[12] For the industrial processes to produce specific aromatic compounds using simple carbon sources as raw materials, bacterial strains need to be developed.[32,37] Strain improvement was achieved by metabolic engineering (including the basic elements: pathway design, construction and optimization[38]) or random strategies.[32,37] In addition, it could be very difficult to identify all enzymes involved in a certain biosynthesis.[35] Therefore, combinatorial biosynthesis has become an attractive tool: enzymes from different sources (plant, bacteria or fungus) can be recombined to create artificial pathways to desired molecules.[39]

Traditionally, classical strains, e.g. *E. coli* as major prokaryotic genetic model[37,38], *Saccharomyces cerevisiae* (*S. cerevisiae*) as predominant eukaryote[38], *Aspergillus niger*, *Bacillus subtilis* or *Corynebacterium glutaminicum*, are available as improved versions as cell factories. There are many advantages of using cell factories: they are well characterized; it is easier to obtain product approval if they have been used for production already; many bioinformatics tools for genome editing are present; and many gene expression tools are available.[35,38] For the production of fuels and chemicals, there is an increasing focus on the use of *E. coli* and *S. cerevisiae* strains. For the production of the molecule of interest, the biosynthetic pathway is reconstructed on the platform of the cell factory.[35] Since 50 years, *E. coli* has been used for molecular cloning methodologies and as host for the production of primary and secondary metabolites.[37]

Some *E. coli* strains were already successfully engineered for the production of aromatic metabolites, e.g. L-tyrosine or L-phenylalanine, or chemicals, e.g. *p*-coumaric acid.[32] While substitution of PAL enzyme allows metabolic engineering to bypass the initial inconveniently insoluble P450 pathway enzyme, subsequent P450 catalyzed steps in downstream pathways of monolignols and flavonoids are so easily circumvented for industrial processes.[15,40] However,

some P450 monooxygenases as well as some *O*-methyltransferases could be engineered by e.g. rational design to be functional in microbial hosts.[4] As mentioned above, the reconstruction of the whole biosynthetic pathway is difficult. For instance, the precursor taxadiene of the anti-cancer drug taxol could be produced by a combination of a native upstream methylerythritol-phosphate pathway forming isopentenyl pyrophosphate and a heterologous downstream terpenoid-forming pathway.[35,41] In some cases, there is a single step limiting the overexpression of an array of heterologous pathways, e.g. in the synthesis of flavonoids the production of the precursor malonyl-CoA is the limiting step.[12] To avoid this, Xu et al. developed a sensor that manipulates two cellular pathways to regulate the production of malonyl-CoA. By using this sensor, researchers were able to maximize the production of the precursor malonyl-CoA while minimizing any damage to the cell.[42]

The common yeast strain *S. cerevisiae* has also been engineered for the high-level production of natural plant products as stilbenoids and flavonoids. By reconstructing of the complex pathways using aromatic amino acids as precursor, the production of the antioxidants resveratrol and naringenin could be achieved.[35]

Doubtless, *E. coli* has long been the workhorse of molecular biology and is one of the preferred host organisms for applications in the field of engineering in industry.[37] However, not all pathways express well in *E. coli* or *S. cerevisiae*. Furthermore, it may be also necessary to have strains that can operate at extreme conditions, e.g. high temperature, high or low pH values and extreme salt conditions.[35]

1.4.3 *In vitro* biosynthesis

The application of genetically modified organisms (GMO) reveals disadvantages as well. One of the major problems is the consumer acceptance. People are persuaded that transgenes of GMOs could have adverse effects on the environment and the human health.[43] Furthermore, the purification of the reaction mixture and the associated removal of the GMO host (plant, fungal or bacterial strain) is very costly and in most cases can be used only once. As another disadvantage, substrate/product or byproduct can possess cytotoxic activities.[44,45] Additionally, in food technology it is important to avoid consumer allergies due to enzyme residues in food. But so far there have been no reports, because the level of enzyme residues appearing in foods is so low. However, like all proteins, enzymes can cause allergic reactions. Because of this, companies use a variety

of protective arrangements, e.g. produced as liquids, granules, in capsules or as immobilized preparations (e.g. catalase in milk systems).[46]

Because of this, the *in vitro* reconstruction is an attractive alternative to generate pure biosynthetic intermediates.[45] Like in chemical catalysis, one can be chosen between homogeneous (dissolved in aqueous liquid phase) and heterogeneous (water-insoluble) biocatalysis. The separation and recycling of the catalyst is more simple, when enzymes are not soluble in the reaction matrices.[47] Moreover, cell-free biosynthesis allows the synthesis of intermediates and products avoiding, e.g. post-translational limitations or protein-protein interactions, and gives a perspective for generating novel chemicals with more potent biological activities.[45] Nevertheless, the production and purification of recombinant enzymes with high stability and activity is one of the main challenges.[44] In addition, the use of expensive cofactors (e.g. NADPH) has to be minimized.

For optimization of any given pathway it is important to realize that each case has its own system-specific sensitivities. Optimization protocols that have shown good results for one system may not perform well for another. Accordingly, it is difficult to predict which strategies will yield the most successful results. Although, there is a high interest in biotechnological approaches, for industrial processes most of systems have to be optimized in relation to their parameters like titers, yields, productivity and economic considerations. There are still many proof-of-concept systems available but transfer to real production is still rare.[48] A combination between metabolic engineering and *in vitro* synthesis could facilitate the production of non-natural products.[45]

1.5 Objectives of this work

Aim of this work is to provide novel access routes for the production of central metabolites of the biosynthesis of lignin (e.g. *p*-coumaryl, coniferyl and sinapyl alcohol) and flavors (e.g. vanillin) starting from e.g. *p*-coumaric acid which is available from primary metabolites that are highly abundant in cells (e.g. tyrosin).

One main topic in this work is to improve the hydroxylation of aromatic compounds, e.g. ferulic acid or naringenin. It shall be shown that by substitutions at amino acid positions the functionality of a flavin-dependent monooxygenases can be changed.

Because some metabolites can be toxic for cells (e.g. vanillin), a further objective will be on studies for the cell-free access to such compounds.

Moreover, a steady focus shall be on minimizing the use of expensive cofactors. If possible, regeneration systems for the cofactors shall be introduced.

Most natural biosynthetic pathways are very complex and not in all cases all enzymes of a pathway can be identified or can be overexpressed in bacterial cells. In this case, it is necessary to find alternative enzymes for the reconstruction of the pathway and for production of the metabolites, which ideally even can yield in simplifications of the whole pathway.

Finally, a toolbox of enzymes for catalysis of different reaction types shall be developed to provide a versatile access to different substitution patterns in phenylpropanoids.

2   Experimental section

2.1   Materials

2.1.1   Consumables (materials) and chemicals

Chemical reagents and consumable materials used are listed in Table 2.1. Bidestilled water for buffers and solutions was produced with a MilliQ-biocel apparatus from Millipore (Billerica, USA); methanol and ethyl acetate was used after in-house distillation of commercial solvents.

Table 2.1: List of all used chemicals.

chemicals and materials	distributor
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Roth, Karlsruhe, Germany
acetonitrile, HPLC grade (MeCN)	Merck, Darmstadt, Germany
acetic acid, 100 %	Roth, Karlsruhe, Germany
adenosine 5'-triphosphate (ATP) disodium salt hydrate	Sigma-Aldrich, St. Louis, USA
agar-agar	Roth, Karlsruhe, Germany
agarose	Roth, Karlsruhe, Germany
ammonium persulfate (APS)	Roth, Karlsruhe, Germany
ammonium sulfate ($(NH_4)_2SO_4$)	Roth, Karlsruhe, Germany
bovine serum albumin	Thermo Fisher, Waltham, USA
coenzyme A trilithium salt dihydrate	AppliChem, Darmstadt, Germany
Coomassie brilliant blue G-250	Roth, Karlsruhe, Germany
deuterated acetone (C_3D_6O)	VWR, Darmstadt, Germany
deuterated chloroform ($CDCl_3$)	VWR, Darmstadt, Germany
deuterated methanol (CD_3OD)	VWR, Darmstadt, Germany
dipotassium phosphate (K_2HPO_4)	Merck, Darmstadt, Germany
desoxynucleotide triphosphate (dNTPs, 10 mM per nucleotide)	Thermo Fisher, Waltham, USA
ethidium bromide (10 mg ml^{-1})	VWR, Darmstadt, Germany
flavin adenine dinucleotide (FAD)	Roth, Karlsruhe, Germany
flavin mononucleotide (FMN)	Roth, Karlsruhe, Germany
formic acid, > 98 %	Roth, Karlsruhe, Germany
Gene-Ruler™ 1 kb DNA Ladder	Thermo Fisher, Waltham, USA
glucose-6-phosphate (G6P)	Sigma-Aldrich, St. Louis, USA
glycerol	Roth, Karlsruhe, Germany
hydrochloric acid (HCl)	Roth, Karlsruhe, Germany
imidazole	Roth, Karlsruhe, Germany
isopropyl β-D-1-thiogalactopyranoside (IPTG)	Roth, Karlsruhe, Germany
kanamycin sulfate	Roth, Karlsruhe, Germany
α-lactose	Sigma-Aldrich, St. Louis, USA
magnesium chloride hexahydrate ($MgCl_2$ x 6 H_2O)	Roth, Karlsruhe, Germany
magnesium sulfate ($MgSO_4$)	VEB Laborchemie, Apolda, (GDR)
monopotassium phosphate (KH_2PO_4)	Sigma-Aldrich, St. Louis, USA
nicotinamide adenine dinucleotide	Roth, Karlsruhe, Germany

disodium salt (NADH)	
nicotinamide adenine dinucleotide phosphate tetrasodium salt (NADPH)	AppliChem, Darmstadt, Germany
PageRuler™ Plus Prestained Protein Ladder	Thermo Fisher, Waltham, USA
potassium acetate (KOAc)	Merck, Darmstadt, Germany
potassium chloride (KCl)	Sigma-Aldrich, St. Louis, USA
Roti®-Quant (Bradford reagent)	Roth, Karlsruhe, Germany
sodium chloride (NaCl)	Roth, Karlsruhe, Germany
sodium dodecylsulfate (SDS)	Roth, Karlsruhe, Germany
sodium hydroxide (NaOH)	Roth, Karlsruhe, Germany
sodium sulfate (Na_2SO_4)	Sigma-Aldrich, St. Louis, USA
tetramethylethylenediamine (TEMED)	Roth, Karlsruhe, Germany
TRIS hydrochloride (Tris-HCl)	Roth, Karlsruhe, Germany
tryptone / peptone ex casein (tryptone)	Roth, Karlsruhe, Germany
yeast extract	Roth, Karlsruhe, Germany

2.1.2 Substrates and reference substances

Substrates for enzymatic reactions and reference substances used as analytical standards are listed in Table 2.2. Each substance used had the highest purity commercially available.

Table 2.2: Applied substrates and reference substances.

substances	distributor
2-hydroxycarbazole	Sigma-Aldrich, St. Louis, USA
3,4-dihydrobenzoic acid	Sigma-Aldrich, St. Louis, USA
5-hydroxyferulic acid	Sigma-Aldrich, St. Louis, USA
caffeic acid	Sigma-Aldrich, St. Louis, USA
caffeoyl alcohol	PhytoLab, Vestenbergsgreuth
cinnamic acid	Sigma-Aldrich, St. Louis, USA
coniferyl alcohol	Sigma-Aldrich, St. Louis, USA
coniferaldehyde	Sigma-Aldrich, St. Louis, USA
eriodictyol	Sigma-Aldrich, St. Louis, USA
esculetin	Sigma-Aldrich, St. Louis, USA
ferulic acid	Sigma-Aldrich, St. Louis, USA
naringenin	Sigma-Aldrich, St. Louis, USA
p-coumaric acid	Sigma-Aldrich, St. Louis, USA
p-coumaryl alcohol	PhytoLab, Vestenbergsgreuth
p-hydroxybenzoic acid	Sigma-Aldrich, St. Louis, USA
piceatannol	Sigma-Aldrich, St. Louis, USA
resveratrol	ABCR, Karlsruhe, Germany
rheosmin	Sigma-Aldrich, St. Louis, USA
sinapaldehyde	Sigma-Aldrich, St. Louis, USA
sinapic acid	Sigma-Aldrich, St. Louis, USA
sinapyl alcohol	Sigma-Aldrich, St. Louis, USA
umbelliferone	Alpha Aesar, Karlsruhe, Germany
vanillin	VEB Laborchemie, Apolda, (GDR)

substances	distributor
dehydrozingerone	IPB compound library, DIM028
dehydrorheosmin	IPB compound library, DIM026

2.1.3 Preparative synthesis of SAM

S-Adenosylmethionine (SAM) was synthesized and purified as described by a standard procedure according to Dippe et al.[49]

2.1.4 Enzymes

Purchased enzymes are listed in Table 2.3. They were stored as described in the instructions of the manufacturers.

Table 2.3: Purchased enzymes.

enzyme type	enzyme	distributor
lysozyme	lysozyme from chicken egg white (> 40,000 U mg^{-1})	Sigma-Aldrich, St. Louis, USA
dehydrogenase	glucose-6-phosphate dehydrogenase from *S. cerevisiae* (140 U ml^{-1}) formate dehydrogenase from *Candida boidinii* (15 U mg^{-1})	Roche, Indianapolis, USA Sigma-Aldrich, St. Louis, USA
DNAse	DNAse I (3,000 U mg^{-1}) from bovine pancreas	AppliChem, Darmstadt, Germany
ligase	T4 DNA ligase (5 U µl^{-1}) purified from *E. coli* œ Lysogen NM989	Thermo Fisher, Waltham, USA
polymerase	Phusion High-Fidelity DNA Polymerase (2 U µl^{-1}) PfuUltra High-Fidelity DNA Polymerase (2.5 U µl^{-1})	Thermo Fisher, Waltham, USA Agilent, Santa Clara, USA
restriction enzyme	BsaI (10 U µl^{-1}) DpnI (10 U µl^{-1}) XhoI (10 U µl^{-1}) NcoI (10 U µl^{-1}) NdeI (10 U µl^{-1})	Thermo Fisher, Waltham, USA

2.1.5 Media

E. coli cells were grown in auto-inducing medium[50] or in lysogeny broth medium.

Lysogeny broth medium (LB medium)

For overexpression of pET-28a(+)-NiPPTase in BL21(DE3) cells and inoculation of pre-cultures lysogeny broth medium (10 g l^{-1} trypton, 10 g l^{-1} NaCl, 5 g l^{-1} yeast extract, pH 7.0) containing kanamycin (50 µg ml^{-1}) was used.

Auto-inducing medium (AI medium)[50]

Auto-inducing medium, which is a mixture of ZY (N-Z-amine and yeast extract) medium, Mg_2SO_4 (1 M), trace metal mix (1000x), 5052 solution (50x) and NPS (20x) containing kanamycin (50 µg ml^{-1}). The compositions of solutions are shown in Table 2.4.

Table 2.4: Compositions of solutions generated for AI medium (1 liter).

solution (volume)	composition
ZY medium (925 ml)	10 g l^{-1} tryptone, 5 g l^{-1} yeast extract
trace metal mix (100 ml)	50 ml of 0.1 M $FeCl_3$ in 0.1 M HCl, 20 mM $CaCl_2$, 10 mM $ZnSO_4$, 2 mM $CoCl_2$, 2 mM CuCl2, 2 mM $NiCl_2$, 2 mM Na_2MoO_4, 2 mM Na_2SeO_3, 2 mM H_3BO_3, add H_2O
5052 solution (20 ml)	25 % (w/v) glycerin, 2.5 % (w/v) glucose, 10 % (w/v) α-lactose
NPS (50 ml)	0.5 M $(NH_4)_2SO_4$, 1 M KH_4PO_4, 1 M Na_2HPO_4

Super optimal broth with catabolite repression medium (S.O.C. medium)

S.O.C. medium contained 20 g l^{-1} trypton, 5 g l^{-1} yeast extract and 0.5 g l^{-1} NaCl. After autoclaving, sterile filtrated salt solutions were added to achieve the final concentrations of 10 mM $MgSO_4$, 10 mM $MgCl_2$ and 2.5 mM KCl. After addition of sterile filtrated 20 mM glucose, S.O.C. medium was aliquoted and stored at -20 °C.

2.1.6 Buffers and stock solutions

Compositions of used buffers and stock solutions are listed in Table 2.5.

Table 2.5: Compositions of used buffers and stock solutions.

buffer and stock solution	composition
50x TAE running buffer (pH 8.0)	2 M Tris, 1 M acetic acid, 100 mM EDTA
10x SDS-PAGE running buffer	250 mM Tris, 2 M glycine, 10 g l^{-1} SDS
5x SDS-PAGE loading dye	10 % (w/v) SDS , 10 mM $β$-mercaptoethanol, 200 mM Tris-HCl, 0.05 % (w/v) bromphenol blue, 20 % (v/v) glycerol
DNase I solution	1 mg ml^{-1} DNase I, 20 mM Tris/HCl (pH 7.5), 1 mM $MgCl_2$, 50 % glycerol
Inoue buffer	55 mM $MnCl_2·4H_2O$, 15 mM $CaCl_2·2H_2O$, 250 mM KCl in H_2O, pH 6.5 (by using KOH)
kanamycin, sterile filtrated	50 mg ml^{-1} in H_2O
IPTG, sterile filtrated	1 M in H_2O
Coomassie staining solution	10 % acetic acid (v/v), 50 % methanol (v/v), 40 % H_2O (v/v), 0.25 % Coomassie brilliant blue G-250 (w/v)
Coomassie decolorizing solution	30 % methanol (v/v), 10 % acetic acid, ad 1 l H_2O

buffer and stock solution	composition
extraction buffer	50 mM HEPES/NaOH, 10 % (w/v) glycerol, 10 mM imidazole, 300 mM NaCl, pH 7.5
wash buffer	50 mM HEPES/NaOH, 10 % (w/v) glycerol, 40 mM imidazole, 500 mM NaCl, pH 7.5
elution buffer	50 mM HEPES/NaOH, 10 % (w/v) glycerol, 250 mM imidazole, 300 mM NaCl, pH 7.5
dialysis buffer	20 mM HEPES/NaOH, 100 mM NaCl, 5 % (w/v) glycerol, pH 7.5

2.1.7 Bacterial Strains

For gene cloning, chemically competent *E. coli* DH5α, and for gene expression *E. coli* BL21(DE3) were used.

2.1.8 Plasmids

All plasmids used for gene cloning and expressions are listed in Table 2.6. All plasmids are in a pET28a(+) vector, have a kanamycin resistance, a N-terminal His$_6$-tag for purification and a T7 promotor. Plasmid and protein sequences are presented in the Appendix.

Table 2.6: Used plasmids for gene cloning and production of recombinant enzymes.

variant	gene / donor organism	source
4HPA3H	*Escherichia coli* BL21(DE3)	cloning see section 3.
HpaC	*Escherichia coli* BL21(DE3)	cloning see section 3.
PrnF	*Pseudomonas protegens* CHA0	cloning see section 3.
ScFDH	*Saccharomyces cerevisiae* pJ69-4A	cloning see section 3.
PFOMT	*Mesembryanthemum crystallinum*	clone from Thomas Vogt, IPB, Halle
LlCCR	*Leucaena leucocephala*	cloning see section 4.
NiCAR	*Nocardia iowensis*	cloning see section 4.
NiPPTase	*Nocardia iowensis*	cloning see section 4.
TaCAD1	*Triticum aestivum*	cloning see section 4.
G6PDH	*Saccharomyces cerevisiae* pJ69-4A	cloning see section 4.
LeOPR1	*Lycopersicon esculentum*	cloning see section 5.

2.1.9 Synthetic oligonucleotides

Oligonucleotides designed for DNA amplification or mutagenesis were synthesized by MWG Eurofins Operon (Ebersberg, https://www.eurofinsgenomics.eu).

Specific oligonucleotides for amplification are listed in Table 2.7.

Table 2.7: Sequences of specific oligonucleotides for DNA amplification.

primer	sequence of oligonucleotide (5'-3')
4HPA3H_fw	TATGGTCTCCATATGAAACCAGAAGATTTCCG
4HPA3H_rev	TATGGTCTCCTCGAGTTATTTCAGCAGCTTATCC
HpaC_fw	TATGGTCTCCATATGCAATTAGATGAACAACGCCTG
HpaC_rev	TATGGTCTCCTCGAGTTAAATCGCAGCTTCCATTTC
PrnF_fw	TATGGTCTCCATATGAATGCTGCCACCGAAAC
PrnF_rev	TATGGTCTCCTCGAGCTATTCGTGCTTGAGGACGC
ScFDH_fw	TATGGTCTCCATATGTCGAAGGGAAAGGTTTTG
ScFDH_rev	TATGGTCTCCTCGAGTTATTTCTTCTGTCCATAAG
TaCAD1_fw	TATGGTCTCCATATGGGCAGCGTCGACGCCTCCG
TaCAD1_rev	TATGGTCTCCTCGAGTCAGGCGGCGTCCTCGATGTTGC
LlCCR_fw	TATGGTCTCCATATGCCAGCTGCCGCCCCTG
LlCCR_rev	TATGGTCTCCTCGAGTTACTTCGTCGGCAGCGGTAAG
G6PDH_fw	TATGGTCTCCATATGAGTGAAGGCCCCGTCAA ATTC
G6PDH_rev	TATGGTCTCCTCGAGTTCTAATTATCCTTCGTATCTTC
NiCAR_fw	TATGGTCTCCATATGGCAGTGGATTCACCGGATGAGC
NiCAR_rev	TATGGTCTCCTCGAGTCAGAGCAGCTGAAGCAGTTCC
NiPPTase_fw	TATGGTCTCCATATGATCGAGACAATTTTGCCTGCTG
NiPPTase_rev	TATGGTCTCCTCGAGTCAGGCGTACGCGATCGCGGTGAG
LeOPR1_fw	TATGGTCTCCATATGGAAAATAAAGTCGTTGAAGAG
LeOPR1_rev	TATGGTCTCCTCGAGTCATGTCATGGTTTCTAG

Specific oligonucleotides for mutation studies (chapter I) are shown in Table 2.8. Red colored nucleotides presenting the modification of the respective codon for the generation of 4HPA3H variants.

Table 2.8: Sequences of specific oligonucleotides for mutation studies.

primer	sequence of oligonucleotide (5'-3')	mutation in 4HPA3H
4HPA3H_Y301F_fw	GTTTTGCCCGTATGTTTCCGCTGCAAGCC	Y301F
4HPA3H_Y301F_rev	GGCTTGCAGCGGAAACATACGGGCAAAAC	
4HPA3H_Y301I_fw	GTTTTGCCCGTATGATTCCGCTGCAAGCC	Y301I
4HPA3H_Y301I_rev	GGCTTGCAGCGGAATCATACGGGCAAAAC	
4HPA3H_Y301L_fw	GTTTTGCCCGTATGCTTCCGCTGCAAGCC	Y301L
4HPA3H_Y301L_rev	GGCTTGCAGCGGAAGCATACGGGCAAAAC	
4HPA3H_S462A_fw	GAAATCAACTACGCCGGTAGCCAGG	S462A
4HPA3H_S462A_rev	CCTGGCTACCGGCGTAGTTGATTTC	
4HPA3H_M293P_fw	GTCGCTGGACGCCGGAAGGCGGTTTTG	M293P
4HPA3H_M293P_rev	CAAAACCGCCTTCCGGCGTCCAGCGAC	
4HPA3H_I157V_fw	CTTTAACCACGCGGTTGTTAACCCAC	I157V
4HPA3H_I157V_rev	GTGGGTTAACAACCGCGTGGTTAAAG	

2.2   Methods in molecular biology

2.2.1   Cloning of genes and mutagenesis

Genes were amplified by standard PCR methods from bacterial or yeast genomic DNA, which was prepared according to published protocols by use of Revert Aid cDNA synthesis kit (Fermentas, Canada). For RNA isolation from *Triticum aestivum* and *Lycopersicum esculentum* a standard TRIzol™ protocol (Thermo Fischer, Waltham, USA) for RNA extraction was used according to manufacturer´s instructions.

2.2.1.1   Polymerase Chain Reaction (PCR)

PCR was used for isolation of DNA fragments from genomic DNA by selective amplification of a specific region of DNA. Primers were designed and synthesized as described in section 2.1.8. In Table 2.9 the pipetting scheme and the cycling instructions for a polymerase chain reaction with Phusion High-Fidelity (HF) DNA polymerase are listed. For determination of the optimal annealing temperature the Tm calculator of Thermo Fisher was used (www.thermofisher.com/tmcalculator).

Table 2.9: Pipetting and cycling instructions for PCR with Phusion HF DNA polymerase.

component	20 µl reaction	cycle step	temperature/time	cycles
5x Phusion HF Buffer©	1x	initial denaturation	98 °C/30s	1
10 mM dNTPs	200 µM each	denaturation	98 °C/5s	
forward primer	0.5 µM			
reverse primer	0.5 µM	annealing	X °C/10-30 s	35
template DNA	200 ng			
		extension	72 °C/15-30 s/kb	
Phusion HF DNA polymerase	0.02 U µl^{-1}	Final extension	72 °C/5-10 min	1
H$_2$O	add to 20 µl		4 °C/hold	

2.2.1.2   Site-directed mutagenesis

Enzyme variants were constructed by mutagenesis of pET-28a(+)-4HPA3H by the QuikChange II Site-Directed Mutagenesis Kit (Agilent, Santa Clara, USA) according to the instructions of the manufacturer. In Table 2.10 the pipetting scheme and the cycling instructions are shown. Primers were designed and synthesized as described in section 2.1.9.

Table 2.10: Pipetting and cycling instructions for site-directed mutagenesis.

component	20 µl reaction	cycle step	temperature/time	cycles
10x reaction buffer	5 µl	1	95 °C/30s	1
dNTP mix	1 µl	2	95 °C/30s	
primer 1	125 ng			
primer 2	125 ng		55 °C/1 min	18
template DNA	5-50 ng		68 °C/1 min per kb of plasmid length	
PfuUltra HF DNA polymerase (2.5 U µl^{-1})	1 µl	Final extension	72 °C/5-10 min	1
			4 °C/hold	
H_2O	add to 50 µl			

2.2.2 Plasmid DNA isolation

Plasmid DNA was isolated from *E. coli* preculture (section 2.3.1) with the Stratec *Smarter Nucleic Acid Sample Preparation* (Stratec, Birkenfeld, Germany) according to manufacturer's instructions.

2.2.3 Determination of DNA concentrations

DNA concentration was measured by spectrophotometric measurement using Titertek Berthold UV-Vis spectrophotometer *Colibri*. Difference spectra of 2 µl plasmid DNA were measured in comparison to 2 µl H_2O (blank). Concentration of the nucleic acid was determined according to the instruction manual.

2.2.4 DNA ligation reaction

The DNA fragments were introduced into a modified pET-28a(+) (Merck, Germany) plasmid in frame with its NdeI and XhoI restriction sites by BsaI-directed ligation according to the method of Engler et al.[51] In Table 2.11 the pipetting and the cycling instructions are shown. After ligation, 5 µl of the reaction mixture were used for transformation into chemically competent DH5α cells (see section 2.3.2).

Table 2.11: Pipetting and cycling instructions for DNA ligation.

component	20 µl reaction	cycle step	temperature/time	cycles
vector plasmid	50 ng	1	37 °C/2 min	
10x T4 DNA ligase buffer	1x			30
BsaI (10 U µl^{-1})	2.5 U	2	16 °C/5 min	
DNA fragment	2:1 molar ratio of insert to vector			
T4 DNA ligase (5 U µl^{-1})	1 U	3	80 °C/5 min	1
H$_2$O	add to 20 µl	4	4 °C/hold	1

2.2.5 Digestion of plasmid DNA

To verify introduction of the recombinant DNA fragment into pET-28a(+) plasmid, isolated plasmid DNA (section 2.2.2) was digested with the restriction enzymes XhoI and NdeI. In Table 2.12 the pipetting instructions are listed. The reaction mixture was incubated at 37 °C for one hour and analyzed by agarose gel electrophoresis (see section 2.2.6).

Table 2.12: Pipetting instructions for digestion plasmid DNA.

component	10 µl reaction
plasmid DNA	1 µg
10x Tango buffer	2x
NdeI (10 U µl^{-1})	3 U
XhoI (10 U µl^{-1})	3 U
H$_2$O	add to 10 µl

2.2.6 Agarose gel electrophorese

Agarose gels were used for visualization of gene fragments in PCR reactions (section 2.2.1.1) and to analyze success of digestion and transformation (section 2.3.3). The concentration of agarose in 1x TAE running buffer was 1 % (w/v). Ethidium bromide (0.5 µg ml^{-1}) was added to visualize DNA fragments by UV light. The agarose gel was loaded with samples containing DNA Loading Dye and 5 µl Gene-Ruler™ 1 kb DNA Ladder as a molecular weight marker. Separation of DNA fragments was achieved by running the gel at 100 V in a horizontal gel chamber (Biometra Compact XS/S, Göttingen) in 1x TAE buffer. A BioDocAnalyze (BDA) live system including BDA analysis software (Biometra, Göttingen) was used for documentation. The electrophorograms of PCR products (section 2.2.1) are presented in the Appendix.

2.3 Methods in microbiology

2.3.1 Cultivation of *E. coli* cells

E. coli strains were plated onto LB agar plates containing the antibiotic kanamycin and incubated overnight at 37 °C. One single colony was used for inoculation of 4 ml LB medium (section 2.1.4). The culture was incubated overnight at 37 °C and 210 rpm. The resulting preculture was used for long-term storage at -80 °C and plasmid preparation.

2.3.2 Preparation of chemically competent cells

Chemically competent *E. coli* cells were prepared according to the method of Inoue.[52] A preculture of 25 ml LB medium was inoculated with a single colony from a LB agar plate and incubated at 160 rpm and 37 °C for six hours. In the evening three flasks with 250 ml LB medium were inoculated with different amounts of preculture (1x 10 ml, 1x 4 ml and 1x 2 ml) and incubated at 20 °C overnight. The best culture with an $OD_{600\ nm}$ of 0.55 was chosen for further steps. The culture was centrifuged for 10 min at 5,000 x g and 4 °C. Supernatant was removed. The cell pellet was carefully resuspended in 80 ml Inoue buffer and centrifuged again (10 min at 4 °C). The supernatant was removed and the cell pellet was resuspended in 20 ml Inoue buffer with an addition of 1.5 ml DMSO. The advanced suspension of chemically competent *E. coli* cells was aliquoted (50 µl) and stored at -80 °C. Transformation efficiency was tested using 10 pg test plasmid (pUC19, New England Biolabs, Ipswich, USA) for transformation into one aliquot according to section 2.3.3.

2.3.3 Transformation of plasmid DNA into chemically competent *E. coli* cells

Aliquots (50 µl) of chemical competent *E. coli* cells were incubated for 30 min on ice. After addition of 100 ng plasmid DNA the suspension was carefully mixed and incubated for 10 min on ice. Heat shock transformation was achieved by incubation of the mixture for 30 s at 42 °C in a water bath. 250 µl S.O.C. medium was added after cooling on ice and incubated at 37 °C and 210 rpm for 1 h. Subsequently, 150 µl of the culture was plated on LB agar plate containing kanamycin for selection of successfully transformed cells. The LB agar plate was incubated at 37 °C overnight.

2.4 Protein production

2.4.1 Production of recombinant enzymes

For the production of recombinant enzymes different expression conditions were used. For production of NiPPTase enzyme, expression was induced by adding IPTG to the culture medium at $OD_{600\,nm}$ of 0.7.

The conditions and resulting expression yields are summarized in Table 2.13.

Table 2.13: Conditions and resulted expression yields for production of recombinant enzymes.

variant	expression conditions	expression yield [mg l^{-1}]	predicted size of protein [kDa]
4HPA3H and variants	37 °C, overnight, auto-induced media	4HPA3H: 248 Y301F/S462A:295 Y301I: 112	59
PrnF	37 °C, overnight, auto-induced media	175	20
ScFDH	25 °C, overnight, auto-induced media	162	42
PFOMT	25 °C, overnight, auto-induced media	202	27
G6PDH	37 °C, overnight, auto-induced media	4	58
NiCAR	25 °C, overnight, auto-induced media	83	128
NiPPTase	25 °C, 16 hours, LB media, 3 mM IPTG	73	24
TaCAD1	25 °C, overnight, auto-induced media	192	39
LeOPR1	25 °C, overnight, auto-induced media, spade point tip of FMN (0.5 - 1 g)	56	42

2.4.2 Cell lysis and enzyme purification

After gene expression the cell culture was centrifuged at 4 °C and 8,000 x g for 20 min. The supernatant was removed and the cell pellet was carefully resuspended in 35 ml extraction buffer (composition see section 2.1.5) containing 100 µg ml^{-1} lysozyme and 10 µg ml^{-1} DNAse I. After incubation on ice, proteins were extracted by high-pressure dispersion. Cell debris were collected by centrifugation (15,000 x g, 30 min). The supernatant was applied to a Talon affinity column (HiTrap crude, 5 ml, GE Healthcare, USA) according to instruction manuals with a flow rate of 1 ml min^{-1} using a peristaltic pump. After removal of impurities by washing with extraction buffer (10 ml) and washing buffer (10 ml), the recombinant proteins were eluted by using an elution buffer with high

imidazole concentration (composition of solutions see section 2.1.5). The eluate was collected in 2 ml-fractions and analyzed via protein detection methods (section 2.6.1). Appropriate protein fractions were combined and dialyzed twice against a 100fold volume of dialysis buffer (composition see section 2.1.5) at 4 °C using a dialysis membrane (ZelluTrans, Roth, Karlsruhe, Germany). The resulting enzyme solution was stored in aliquots at -80 °C.

2.5 Enzymatic reactions

2.5.1 Enzyme activity assay

In Table 2.14 and 2.15 compositions and reaction conditions for small scale activity assays are given. For this assays purified enzymes were used.

Table 2.14: Compositions and reaction conditions for activity assays. Compositions and reaction conditions for assaying activity of the listed enzymes.

	4HPA3H, 4HPA3H-Y301I	PFOMT	NiCAR/NiPPTase	TaCAD1
buffer system	100 mM Tris/HCl pH 7.5	100 mM Tris/HCl, 1 mM $MgCl_2$, pH 7.5	10 mM Tris/HCl, 2 mM $MgCl_2$, 0.2 mM DTT, 0.2 mM EDTA, 2 % (v/v) glycerol	100 mM Tris/HCl pH 7.5, 20 mM sodium ascorbate
enzyme concentration	1 mg ml^{-1}	0.25 mg ml^{-1}	1 mg ml^{-1} NiCAR 350 µg ml^{-1} NiPPTase	6.9 µg ml^{-1}
substrates, cosubstrates	for 4HPA3H: 1 mM of **1** for 4HPA3H-Y301I: 1 mM of **3**	1 mM of **2** or **4**, 1 mM SAM, 2.5 mM GSH	1 mM of **1**, **2**, **3**, **4** or **17** 1 mM ATP	100 µM of **20** or **22**
cofactor regeneration system	10 µM FAD, 25 µM NADH, 34 µg ml^{-1} PrnF, 100 mU ml^{-1} ScFDH	none	5 mM glucose-6-phosphate, 0.5 µg ml^{-1} G6PDH, 200 µM NADP^{+}	250 µM glucose-6-phosphate, 0.3 µg ml-1 G6PDH, 125 µM NADP^{+}
final volume	80 µl per well, 96 well-plates	80 µl per well, 96 well-plates	1 ml	150 µl per well, 96 well-plates
Incubation conditions	catechol-assay, reaction is started by addition of 4HPA3H	catechol-assay, reaction is started by addition of SAM	incubation of NiCAR and NiPPTase with 1mM CoA at 28 °C, 250 rpm for 1 hour	continuously measurements at 400 nm, RT, 10 minutes

	4HPA3H, 4HPA3H-Y301I	PFOMT	NiCAR/NiPPTase	TaCAD1
specific activity	4HPA3H: 11.5 ± 0.3 mU mg^{-1} 4HPA3H-Y301I: 4.5 ± 0.2 mU mg^{-1}	**2**: 35.5 ± 0.6 mU mg^{-1} **4**: 86.7 ± 1.57 mU mg^{-1}	not determined because of instability of the provided cinnamaldehydes	**20**: 17.1 ± 2.9 U mg^{-1} **22**: 13.5 ± 2.9 U mg^{-1}

Table 2.15: Compositions and reaction conditions for activity assays, continued. Compositions and reaction conditions for assaying activity of the listed enzymes.

	PrnF	ScFDH	G6PDH	LeOPR1
buffer system	100 mM Tris/HCl pH 7.5	100 mM Tris/HCl, 20 mM MgCl$_2$, 200 mM KCl pH 7.5	100 mM Tris/HCl, 20 mM MgCl$_2$, 200 mM KCl pH 7.5	100 mM Tris/HCl, 20 mM MgCl$_2$, 200 mM KCl pH 7.5
enzyme concentration	12.5 µg ml^{-1}	5 µg ml^{-1}	0.27 µg ml^{-1}	0.8 mg ml^{-1}
substrates, cosubstrates	200 µM NADH, 30 µM FAD	300 mM Na-formate, 100 µM NAD$^+$	5 mM glucose-6-phosphate, 200 µM NADP$^+$	100 µM of **29** or **30**
cofactor regeneration system	none	none	none	5 mM glucose-6-phosphate, 0.5 µg ml^{-1} G6PDH, 200 µM NADP$^+$
final volume	1 ml (cuvette)	500 µl (cuvette)	500 µl (cuvette)	500 µl (cuvette)
incubation conditions	continuously measurements at 339 nm[53], RT	continuously measurements at 339 nm[53], RT	continuously measurements at 339 nm[53], RT	continuously measurements at 410 nm, RT
specific activity	16.66 ± 1.60 U mg^{-1}	1.17 ± 0.06 U mg^{-1}	183.4 ± 20.9 U mg^{-1}	**29**: 6.4 ± 0.56 mU mg^{-1} **30**: 10.5 ± 0.2 mU mg^{-1}

2.5.1.1 Spectrophotometric activity assay – Catechol-assay

The assay was performed in a microplate. Reactions (80 µl) contained reaction buffer (100 mM Tris/HCl, pH 7.5), substrate (1 mM *p*-coumaric acid **1** or ferulic acid **3** which were added from stock solution (10 mM in methanol), FAD (10 – 1000 µM), NADH (25 – 200 µM), sodium formate (300 mM) and formate dehydrogenase from *Candida boidinii* (100 mU ml^{-1}). Reactions were started by

addition of PrnF (566 mU ml^{-1}) and oxidase (1 mg ml^{-1}). The reactions were stopped after distinct times of incubation (0 – 240 min) by mixing with 40 µl of catechol reagent (2 mM FeCl$_3$ in 10 mM HCl) to generate a colored Fe^{3+}/product complex. From the absorbance at 595 nm, measured after 5 min of equilibration at RT, the concentration of formed product was calculated by means of a standard curve taken with mixtures of the substrate (x mM of **1** or **3**) and the corresponding product (1 - x mM of **2** or **4**). One unit of enzyme activity represents the amount of oxidase converting 1 µM of *p*-coumaric acid (**1**) in one minute of reaction time.

This assay can also be used for *O*-methyltransferases (PFOMT), without regeneration system.

2.5.1.2   Activity assay for conversion of naringenin (7)

The catechol-assay could not be used as activity assay for the conversion of naringenin (**7**) to eriodictyol (**8**) with the oxidase variant 4HPA3H-Y301F/S462A, because of the product does not chelate Fe^{3+} under the conditions used. Thus, the following HPLC method was used:

Reactions (2 ml) contained reaction buffer (100 mM Tris/HCl, pH 7.5), substrate (1 mM naringenin **7** which were added from stock solution (10 mM in methanol), FAD (10 µM), NADH (25 µM), sodium formate (300 mM) and formate dehydrogenase from *Candida boidinii* (100 mU ml^{-1}). Reactions were started by addition of PrnF (566 mU ml^{-1}) and oxidase variant 4HPA3H-Y301F/S462A (1 mg ml^{-1}). The reactions were stopped after distinct times of incubation (0 – 240 min) by mixing 100 µl of the reaction mixture with 1 µl formic acid, products were extracted twice with 100 µl ethyl acetate. The combined organic phases were evaporated and the residue was dissolved in 100 µl of acetonitrile. The samples were used for HPLC analysis (see section 2.6.4).

The measured activity for the conversion of **7** with 4HPA3H-Y301F/S462A to **8** was 1.5 mU mg^{-1}.

This assay was further used for the determination of chirality in the conversion of **7** (see section 5.2.2 and for HPLC analysis section 2.6.4). Therefore, the residue was dissolved in 100 µl isopropyl alcohol before HPLC analysis.

2.5.2   Enzymatic reactions for product identification

For identification of product adjusted assay conditions were used.

2.5.2.1   *In vitro* conversion of monophenols

The reactions (200 µl) were performed in 1.5 ml reaction tubes. Reaction buffer (100 mM Tris/HCl, pH 7.5), monophenol substrate (1 mM from a stock (10 mM) in

methanol), FAD (10 µM), NADH (25 µM), sodium formate (300 mM) and formate dehydrogenase from *Candida boidinii* (100 mU ml^{-1}) were mixed with PrnF (566 mU ml^{-1}) and oxidase (1 mg ml^{-1}) (4HPA3H or variants Y301I and Y301F/S462A).

After 16 hours of incubation, the samples were centrifuged at 14,000 rpm for 10 min. The supernatant was acidified by the addition of formic acid (2 µl) and extracted two times with 200 µl of ethyl acetate. The combined organic phases were evaporated and the residue was dissolved in 200 µl of acetonitrile. The samples were used for HPLC analysis (see section 2.6.4). The data shown in this study are the means of three independent experiments (see section 3).

2.5.2.2 *In vitro* conversion of cinnamic acid derivatives – specific aromatic decoration

The reactions (500 µl) were performed in 1.5 ml reaction tubes.

For formation of caffeic acid, *p*-coumaric acid (200 µM from a stock (10 mM in methanol)), FAD (10 µM), NADH (25 µM), sodium formate (300 mM) and FDH (100 mU ml^{-1}) were mixed with PrnF (34 µg ml^{-1}) and 4HPA3H (1 mg ml^{-1}) in reaction buffer (100 mM Tris/HCl, 1 mM MgCl$_2$, pH 7.5). For formation of ferulic acid, L-glutathione (2.5 mM), SAM (0.2 mM) and PFOMT (50 µg ml^{-1}) were added to the reaction mixture as described above. For formation of 5-hydroxyferulic acid, the variant 4HPA3H-Y301I (1 mg ml^{-1}) was added to the reaction mixture. Instead of 200 µM SAM, 1 mM SAM was used for the formation of sinapic acid.

After 16 hours of incubation, the samples were acidified by the addition of formic acid (5 µl) and extracted two times with 500 µl of ethyl acetate. The combined organic phases were evaporated and the residue was dissolved in 200 µl acetonitrile. The samples were used for HPLC analysis (see section 2.6.4).

The data shown in this study are the means of three independent experiments (see section 4).

2.5.2.3 *In vitro* conversion of cinnamic acid derivatives – side chain reduction

The reactions (1 ml) were performed in 2 ml reaction tubes. Incubation of CoA (1 mM), NiCAR (1 mg) and 350 µg NiPPTase filled up with reaction buffer (50 mM Tris/HCl, 10 mM MgCl$_2$, 1 mM DTT, 1 mM EDTA, 10 % (v/v) glycerol) for 1 hour at 28 °C and 250 rpm (total volume of 500 µl). After incubation, the activated NiCAR was added to the reaction mixture (500 µl):

For the reduction of carboxylic acids to the respective aldehydes the cinnamic acid derivative (1 mM from a stock (100 mM) in methanol), ATP (1 mM), NADPH

(2 mM), glucose-6-phosphate (5 mM) and G6PDH (0.3 µg ml^{-1}) were mixed. For the reduction of carboxylic acids to the corresponding alcohols, CAD (6.9 µg ml^{-1}) was added to the reaction mixture described above.

After 1 hour (conversion to respective aldehydes **18**-**22**) or 16 hours (conversion to respective cinnamyl alcohols **23**-**27**) of reaction time the samples were extracted by using two different extraction steps: at first the samples were extracted three times with 500 µl of ethyl acetate. Then, the samples were acidified by the addition of formic acid (10 µl) and extracted two times with 500 µl of ethyl acetate afterwards.

The combined organic phases (each extraction step was pooled separately) were evaporated and the residue was dissolved in 250 µl acetonitrile. At this point samples from the two extraction steps were combined. The samples were used for HPLC analysis (see section 2.6.4). The data shown in this study are the means of three independent experiments (see section 4).

2.5.3   Whole-cell bioconversions

Whole-cell bioconversions were performed for reactions with 4HPA3H and or its variants.

Reactions were performed in 50 ml Erlenmeyer flasks. 10 ml of auto-inducing medium containing oxidase substrate (200 µM from stock solutions (200 mM) in methanol) and kanamycin (50 µg ml^{-1}) were inoculated with *E. coli* BL21(DE3) harboring either the plasmids for production of 4HPA3H or its variants (hydroxylation reactions) or the empty vector (controls). Cells were grown at 37 °C with shaking (200 rpm). After 16 hours of incubation, samples (1.5 ml) were taken for HPLC analysis (see section 2.6.4). For all conversions except those of 2-hydroxycarbazole, the samples were centrifuged at 14,000 g for 10 min. The supernatants (1 ml) were acidified by the addition of 10 µl of formic acid and extracted two times with 500 µl of ethyl acetate. For samples containing the poorly soluble 2-hydroxycarbazole, the centrifugation step was omitted, and the cell suspension was directly used in the extraction procedure. After extraction the organic phases were combined, the solvent was removed by vacuum distillation and the residue was re-dissolved in acetonitrile (200 µl). This solution was used for HPLC measurements (see section 2.6.4). The data shown in this study were generated from two independent experiments (see section 3).

2.5.4 Preparative substrate conversion

2.5.4.1 Whole-cell bioconversion – preparative scale

For product identification, preparative transformations were performed similar to the reactions described above (section 2.5.3) but on larger scale, i.e. in 2,000 ml Erlenmeyer flasks containing 500 ml of medium. After incubation at 37 °C for 16 hours, formic acid (5 ml) was added to the cell suspension (conversion of 2-hydroxycarbazole) or to the supernatant from centrifugation (8,000 g, 10 min) (all other conversions). The combined organic phases from extraction with ethyl acetate (2 x 250 ml) were dried with Na_2SO_4, filtrated and the solvent was removed under reduced pressure. The residue was re-dissolved in a minimum volume of methanol and used for preparative HPLC (see section 2.5.3) and ESI-MS measurements (see section 2.6.7.1).

2.5.4.2 Preparative scale reaction for identification of products formed from sinapic acid

The reaction conditions were identical to this described in section 2.5.2.3, but the reactions were performed in twentyfold implementations (20 tubes).

After extraction the organic phases were combined in a 50 ml flask and the solvent removed under reduced pressure. The residue was re-dissolved in deuterated solvent: sinapaldehyde (**22**) and sinapyl alcohol (**27**) in $CDCl_3$.

2.6 Analytics

2.6.1 Protein concentration determination

Proteins were assayed with commercial Bradford reagent (Roth, Germany) in microplates according to manufacturer's instructions. Protein concentrations were calculated using standard curves of bovine serum albumin.

2.6.2 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was used to analyze protein samples from overexpression and purification. The standard procedure was performed according to Laemmli.[54] As molecular weight markers 5 µl PageRuler[TM] Plus Prestained Protein Ladder (10 to 180 kDa) was used. Depending on the molecular weight of the expected proteins, the polyacrylamide content in the running gel was adjusted (10 - 14 % (w/v)). In the stacking gel 4 % acrylamide was used for polymerization. Samples (20 µl) contained 5 µg of purified protein dissolved in 1x SDS-PAGE loading buffer. For denaturation the sample mixture was heated for 6 min at 95 °C. Subsequently, SDS-PAGE ran at 30 mA in the stacking gel and at 60 mA for

separation in the running gel in a vertical gel chamber from SERVA Electrophoresis (SERVA, Heidelberg, Germany). The gels were stained with Coomassie Brilliant Blue R250, and apparent molecular masses were estimated using a molecular weight marker. A Bio-5000$_{plus}$ Mikrotek system including analysis software was used for documentation.

Figures of used purified enzymes are imaged in the Appendix.

2.6.3 Verification of cofactor load of NiCAR by peptide analysis

The verification of the phosphopantetheinylated position at CAR enzyme from *Nocardia iowensis* (position S689) was done in cooperation with Petra Majovsky and Dr. Wolfgang Hoehenwarter at the IPB Halle (Saale).

After incubation (1 hour, 28 °C) of the NiCAR enzyme (256 µg) with 1 mM CoA and 350 µg NiPPTase in reaction buffer (50 mM Tris/HCl, pH 7.5) in a total volume of 500 µl under shaking (250 rpm), the proteins were in-solution digested with a combination of AspN and Glu-C. Therefore, 12 µl of 50 mM ammonium hydrogencarbonate pH 7.6 were mixed with 10 µg of the reaction solution in 12 µl of reaction buffer and 0.4 µg of AspN in 10 µl of AspN stock solution (manufacturer) were added (enzyme to protein ratio 1:25). Digestion was allowed to proceed for 8 h at 37°C. Subsequently, 0.4 µg of Glu-C in 2 µl of distilled H_2O were added and digestion was prolonged overnight at the same temperature. The peptide mix was dried down completely and peptides were desalted on C18 STAGE tips as described previously.[55]

0.4 µg of peptides were injected into an EASY-nLC II liquid chromatography system (Thermo Fisher Scientific) and separated on a reverse phase column of length of 10 cm, a column inner diameter (ID) of 75 µm, a particle size of 3 µm and a 60 min gradient from 5 to 40 % acetonitrile in water and a flow rate of 300 nl min^{-1}. Peptides were electrosprayed on-line into an Orbitrap Velos Pro mass spectrometer (Thermo Fisher, Waltham, USA). MS/MS peptide sequencing was performed using a targeted inclusion list-based scan strategy that isolated and performed CID of ions with m/z on the inclusion list regardless of their precursor intensity.

Peptides were identified and the phosphopantetheine modification was mapped using the Mascot Software v. 2.5.0 (Matrix Science) linked to Proteome Discoverer v1.4. (Thermo Fisher, Waltham, USA) to search the SwissProt database (updated Dec. 2015, 550,116 sequences; 196,219,159 residues). A precursor mass error of 7 ppm and a fragment ion mass error of 0.8 Da were tolerated; methionine oxidation and phosphopantetheinylation of serine residues (mass difference: 340.085794) were set as variable modifications. A false

discovery rate (FDR) was determined with the target-decoy database approach. Peptides with a q-value below 0.01 and an ion score surpassing the Mascot significance threshold (ion score >33) were accepted.

2.6.4 High performance liquid chromatography (HPLC)

2.6.4.1 HPLC with UV/Vis detection

Samples (10 µl) from whole-cell biotransformation (see section 2.5.3), enzyme assays (see section 2.5.2.1 and 2.5.1.2) and samples (20 µl) from multi-enzyme cascade (see section 2.8) were applied to a C18 reversed-phase column (Pack ODS-A, 5 µM, 150 x 4.6 mm, YMC Europe Dinslaken, Germany) on a Merck Hitachi HPLC system. The compounds were separated by elution with acetonitrile-water (containing 0.2 % formic acid). The following gradient was used at a flow of 0.8 ml min^{-1} (percentage of acetonitrile in v/v): 5 % for 4 min, 5 to 100 % for 21 min, 100 % for 5 min, 100 to 5 % for 1 min and 5 % for 5 min.

All products that were generated on a larger scale (see section 2.5.4.1) were separated on a preparative HPLC system by using a C18 reversed-phase column (Column ODS-A, 5 µM, 150 x 20 mm, YMC Europe, Dinslaken, Germany). The following gradient was used at a flow of 15 ml min^{-1} (percentage of acetonitrile in v/v): 50 % for 4 min, 50 to 100 % for 21 min, 100 % for 5 min, 100 to 50 % for 1 min and 50 % for 5 min.

Samples (10 µl) of coupled *in vitro* conversions (see section 2.5.2.2) were applied to a C18 reversed-phase column (Pack ODS-A, 5 µM, 150 x 4.6 mm, YMC Europe Dinslaken, Germany) on a VWR Hitachi HPLC System. The compounds were separated by elution with acetonitrile-water (containing 0.2 % formic acid). The following gradients were used at a flow of 0.8 ml min^{-1} (percentage of acetonitrile in v/v): Separation of *p*-coumaric acid (**1**), ferulic acid (**3**) and sinapic acid (**17**) – 5 % for 4 min, 5 to 20 % for 1 min, 20 % for 20 min, 20 to 100 % for 1 min, 100 % for 4 min, 100 to 5 % for 1 min and 5 % for 5 min; separation of caffeic acid (**2**) and 5-OH-ferulic acid (**4**) – 5 % for 4 min, 5 to 13 % for 1 min, 13 % for 20 min, 13 to 100 % for 1 min, 100 % for 4 min, 100 to 5 % for 1 min and 5 % for 5 min.

For both analytical and preparative HPLC, elution was monitored at a wavelength of 280 nm. Reaction products were identified by means of authentic standards, which were applied in different concentrations for quantification (Calibrations are listed in the Appendix), or in case of 4-(3,4-dihydroxyphenyl)-butan-2-one (**16**), 2,3-dihydroxycarbazole (**12**), *p*-coumaryl aldehyde (**18**), caffeic aldehyde (**19**),

5-hydroxyconiferaldehyde (**21**) and 5-hydroxyconiferyl alcohol (**26**) by ESI-MS analysis.

2.6.4.2   HPLC with chiral detector

Samples (10 µl) from the activity assay (see section 2.5.1.2) were applied to a Chiralpak® column (LTD, ASH, 5 µM, 250 x 4.6 mm, Daicel Corporation, Japan) on an Agilent 1100 & 1260 HPLC system. The compounds were separated by elution with an isocratic gradient of n-hexane/isopropyl alcohol (60:40) for 60 min with a flow of 0.5 ml min^{-1} at 25 °C. Elution was monitored at different wavelengths using a PDA detector (210 nm, 238 nm, 254 nm, 280 nm and 200 – 400 nm).

2.6.5   NMR spectroscopy

The ^{1}H NMR spectra were recorded on a 400 MHz Agilent DD2 400 NMR spectrometer at 25 °C. Substances were dissolved in CD_3OD, and chemical shifts were referenced to internal TMS. The data were evaluated by Mestrelab Research S.L. MestReNova 6.0.2 software.

2.6.6   UV-Vis spectroscopy

The formation of NADH, NADPH and other UV/Vis-active products (for detailed information see section 2.5.1) were determined at selected wavelengths on a SpectraMax M5 Multi-Mode Microplate Reader. This reader could be used for assays in microplates and cuvettes.

2.6.7   MS spectrometry and LC/MS measurements

2.6.7.1   ESI-MS analysis for product formation of whole-cell biotransformation (section 2.4.5.1)

The negative ion high resolution ESI mass spectra and the collision induced dissociation (CID) MS^2 measurements were obtained from a Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with an HESI electrospray ion source (spray voltage 4.0 kV; capillary temperature 275 °C, source heater temperature 40 °C; FTMS resolution 60.000). Nitrogen was used as sheath gas. The sample solutions were introduced continuously via a 500 µl Hamilton syringe pump with a flow rate of 5 µl min^{-1}. The instrument was externally calibrated by the Pierce® LTQ Velos ESI positive ion calibration solution (product number 88323) and Pierce® ESI negative ion calibration solution (product number 88324) from Thermo Fisher Scientific, Rockford, IL, 61105 USA). The data were evaluated by the Xcalibur software 2.7 SP1 by averaging 25 scans.

2.6.7.2 UHPLC/MS analysis for product formation of multi-enzyme cascade

The negative ion high resolution ESI mass spectra, collision-induced dissociation (CID) MS² and higher-energy collisional dissociation (HCD) measurements were obtained from an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a HESI electrospray ion source (spray voltage: 4.0 kV, source heater temperature: 280 °C, capillary temperature: 275 °C, FTMS resolution: 30.000). Nitrogen was used as sheath and auxiliary gas. Samples were either injected directly (samples containing caffeic acid (**2**)) or were separated on a UHPLC system (Dionex UltiMate 3000, Thermo Fisher Scientific), equipped with a RP-C18 column (1.9 µm, 50 x 2,1 mm, Hypersil GOLD, Thermo Fisher Scientific, column temperature: 30 °C), and a photodiode array detector (PDA, Thermo Fisher Scientific) coupled to the MS instrument (product mixture containing **3**, **28**, *p*-hydroxybenzaldehyde and adenine). For the UHPLC separation, water/acetonitrile mixtures containing 0.2 % (v/v) formic acid were used. The following gradient was applied at a flow rate of 150 µl min^{-1} (percentage of acetonitrile in v/v): 5 % for 2.5 min, 5 to 100 % for 10 min, and 100 % for 3 min. From the PDA wavelength spectrum (190 – 400 nm), which was recorded during separations, signals at 280 nm were used for detection. The instrument was externally calibrated with the Pierce® LTQ Velos ESI negative ion calibration solution (Thermo Fisher Scientific, Rockford, IL, 61105 USA). The data were evaluated by the Xcalibur software 2.1 (Thermo Fisher Scientific). Data are listed in the Appendix.

2.6.7.3 LC/MS analysis for side chain reduction

Mass spectrometric measurements for side chain reduction (see section 4.2.4) were performed with a Qq-TOF mass spectrometer that was coupled to a Waters Aquity UltraPerformance LC system. Samples (2 µl) were applied to a C18 reverse-phase column (Nucleoshel EC, 150 x 2 mm, 2.7 µM, heated at 40 °C, partial loop with 10 µl loop size) and separated by elution with acetonitrile (eluent A) and 3 mM ammonium formate (eluent B). The following gradient was used at a flow of 400 µl min^{-1}: isocratic 5 % of eluent A for 2 min, linear gradient from 5 to 95 % of eluent A in 17 min, isocratic 95 % of eluent A for 2 min, linear gradient from 95 to 5 % of eluent A in 0.1 min. Ionization was carried out via positive TOF MS (ion source gas 1: 60, ion source gas 2: 70, curtain gas: 35, temperature: 450 °C, ion spray voltage 5500 V in positive ionization mode with different collision energies (10 to 60 eV), pulser frequency: 16.384 kHz, minimal recorded intensity: 1 cps, *m/z* range: 360–1300).

2.7 Homology modeling

The three-dimensional structures of 4HPA3H from *E. coli* and *Pseudomonas aeruginosa* were modeled by means of the online protein fold recognition server Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index, accession at 01.08.2013).[56] The model with the highest score (confidence: 100 %, coverage: 93 %), which is based on crystallographic data of the 4HPA3H enzyme from *Thermus thermophilus* (PDB accession 2YYJ), was used for alignments and active site studies.

2.8 Multi-enzyme cascade

Overexpression, purification and characterization of enzymes used in modul M and modul L was done by Dr. Martin Dippe, IPB Halle (Saale). Data are listed in the Appendix.

2.8.1 Immobilization

Protein solutions containing the amount of enzymes specified in Table 2.16 were mixed with 50 mM Tris/HCl, pH 7.5 to yield a final volume of 3 ml. The mixture was repeatedly applied to a Talon affinity column (HiTrap crude, 1 ml, GE Healthcare, USA). The progress of immobilization was checked by determination of protein in the flow-through by the Bradford assay (see section 2.6.1). Loading was continued until the binding of the enzymes was complete (i.e. until the flow-through was negative for proteins). After washing with 5 ml of 50 mM Tris/HCl, pH 7.5 to remove components of storage buffer, the enzyme modules were immediately used for biocatalytic conversions to ensure full enzyme activity.

Table 2.16: Applied amount of enzymes for immobilization.

module	enzyme	applied activity in modules [mU]
O1 for formation of **2**	4HPA3H	20
	PrnF	5000
	ScFDH	640
M	PFOMT	70
	SAMS-I317V	70
	SAHN	10
	SAHH	-
L	4CL2	70
	FCoAHL	70
O2 for formation of **8**	4HPA3H-Y301F/S462A	35
	Prnf	5000
	ScFDH	640
R	LeOPR1	35
	G6PDH	2000

2.8.2 Synthesis of caffeic acid/eriodictyol

The reaction (10 ml) contained reaction buffer (O1: 100 mM Tris/HCl, 20 mM MgCl$_2$, 200 mM KCl, pH 7.5 or O2: 100 mM Tris/HCl, pH 7.5), the enzyme substrates p-coumaric acid (**1**, substrate O1, 200 µM) or naringenin (**7**, substrate O2, 200 µM), sodium formate (300 mM), and the cofactors NADH (25 µM) and FAD (10 µM). The reagents were mixed in a round-bottom flask (25 ml) connected to a peristaltic pump (P-1, GE Healthcare, USA), which was set to a flow rate of 1 ml min^{-1}. After integration of module O1 or module O2, the reaction was incubated at 25 °C for 120 min (module O1) and 360 min (module O2) under continuous stirring. For analysis of product formation, samples (35 µl) were taken after distinct times of incubation from the round-bottom flask, mixed with an equal volume of acetonitrile and analyzed by HPLC (see section 2.6.4). The residual reaction mixture of module O1 was extracted twice with 10 ml of ethyl acetate containing 1 % (v/v) formic acid. After removal of the solvent from the pooled organic phases by rotary evaporation, the sample was dissolved in methanol and used for identification of caffeic acid (**2**) by electrospray ionization mass spectrometry (ESI-MS) (see section 2.6.7).

2.8.3 Synthesis of ferulic acid/vanillin

The experimental set-up was similar to that in synthesis of caffeic acid (**2**, see section 2.8.2), but the reaction mixture additionally contained L-methionine (1 mM) and ATP (2.5 mM). The conversion to ferulic acid (**3**) was started by integration of modules O1 and M. The reaction was incubated for 360 min under continuous stirring. For subsequent transformation of ferulic acid (**3**) into vanillin (**28**), module O1 was replaced by module L, and CoA (200 µM) was added. After distinct times of incubation, samples (35 µl) were taken from the round-bottom flask and prepared for HPLC analysis (see section 2.6.4) similar as in the hydroxylation experiment (see above). The residual reaction mixture was extracted twice with 10 ml of ethyl acetate containing 1 % (v/v) formic acid. After removal of the solvent from the pooled organic phases by rotary evaporation, the sample was dissolved in methanol and used for identification of the products by ultra-high performance liquid chromatography (UHPLC)/ESI-MS (see section 2.6.7).

2.8.4 Synthesis of rheosmin and zingerone

The experimental set-up was similar to section 2.8.2. The mixture (10 ml) contained reaction buffer (100 mM Tris/HCl, 20 mM MgCl$_2$, 200 mM KCl, pH 7.5),

the enzyme substrates dehydrorheosmin (**29**, 200 µM) or dehydrozingerone (**30**, 200 µM) and glucose-6-phosphate (5 mM), and the cofactor NADPH (200 µM). For both conversions a discoloration of the yellow reaction mixture can be observed. After distinct times of incubation (total reaction time: 180 min for the synthesis of zingerone, 240 min for the synthesis of rheosmin), samples (35 µl) were taken from the round-bottom flask and prepared for HPLC analysis (see section 2.6.4) similar as in the hydroxylation experiment (see section 2.8.2).

3 Chapter I: Rational engineering of the substrate spectrum of a FAD-dependent monooxygenase from *E. coli*

3.1 Introduction

Hydroxylations are oxygen transfer reactions, which functionalize organic molecules. The hydroxyl group is introduced by substitution of a hydrogen atom or – less commonly – of another functionality. Although these reactions are an essential part of the synthetic routes to a multiplicity of natural products, the selective introduction of a hydroxyl group into alkyl or, in particular, aryl moieties is one of the most challenging fields in chemical synthesis. Hence, the number of methods for direct (aromatic) hydroxylation is still limited.[57] The hydroxylation of aromatics (e.g. benzene, toluene) can be achieved with hydrogen peroxide, alkyl hydroperoxide, monopersulfates or other oxidants. Most of them are usually activated by metal catalysts, often with restricted selectivity. One industrial example is the production of phenol from cumene.[58] In contrast to these chemosynthetic approaches, enzyme-mediated hydroxylation often proved to be more effecient due to high regio- and/or stereoselectivity, and is increasingly favored as it can be environmentally more compatible (due to the omission of reactive chemicals and the potential reduction in organic solvents).[59] Because of this the selective oxygenation of aromatic compounds by oxidoreductases has received much attention.[60]

Indeed, biocatalytic oxidative transformations have been used for hydroxylation of structurally divergent compounds, including steroids[61,62], alkaloids[61], terpenoids[63] or fatty acids[61]. A wide range of such reactions have been documented using enzymes from microbial sources.[60] These enzymatic processes nearly exclusively rely on cytochrome P450-type enzymes, certainly because a multitude of these proteins has been produced recombinantly. In addition, many cytochrome enzymes are well characterized with respect to substrate scope, such as the CYP102A1 enzyme from *Bacillus megateriu*m (BM3) and its variants.[14,61,62] Alternatively, flavin-dependent monooxygenases are increasingly used in hydroxylation reactions in recent years.[64,65] These monooxygenases belong to a superfamily of enzymes that is involved in key metabolic processes in both prokaryotic and eukaryotic cells, such as the biosynthesis of polyketides, cholesterol or antibiotics.[65,66] A distinction is made between one- and two-component flavoenzymes (Fig. 3.1).
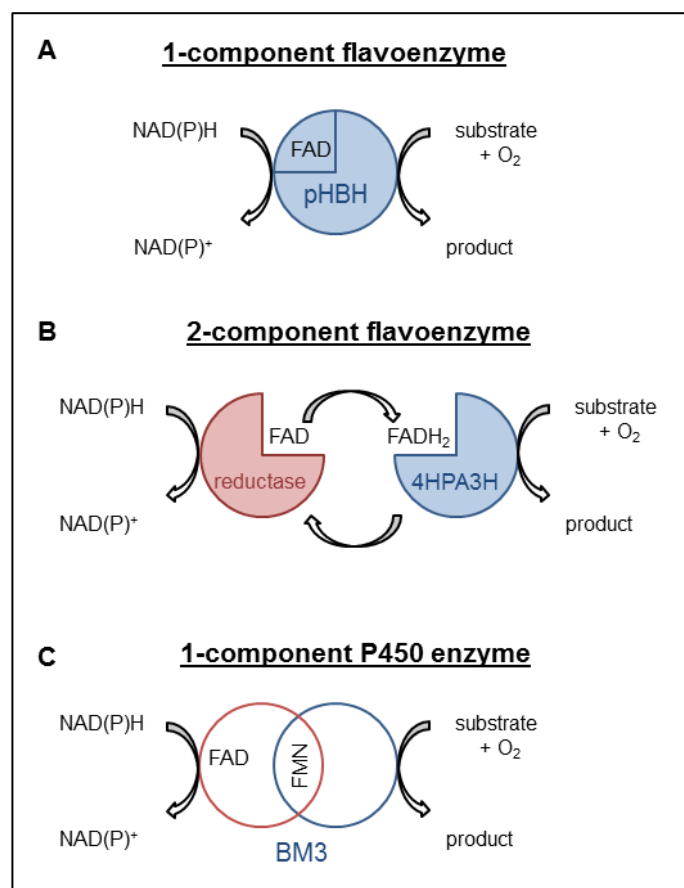
Figure 3.1: Comparison of one- and two-component flavoenzymes (A and B) and 1-component P450 enzymes (C). A) *para*-hydroxybenzoate hydroxylase (pHBH) from *Pseudomonas fluorescence*, B) 4-hydroxyphenylacetate-3-hydroxylase (4HPA3H) from *E. coli* and C) CYP102A1 from *Bacillus megaterium* (BM3) are presented as examples for each type of oxidases.

The used flavins of flavin-dependent monooxygenases are either flavin adenine dinucleotide (FAD) or flavin mononucleotide (FMN). This can be bound tightly to the enzyme or function as a cosubstrate.[40] Flavoenzymes can be classified by different criteria, e.g. the chemical reaction that is catalyzed, reducing and oxidizing substrates, or based on homology in sequence. In Table 3.1. flavoenzymes are grouped in six sub-classes classified by structure data and their amino acid sequence similarity. [40,67]

Table 3.1: Classification of flavoprotein monooxygenases according to van Berkel et al.[67]
The most common reactions of *in vivo* oxidation activities are given.

sub-classes	prototype	reactions	cofactor	coenzyme
A	4-OH-benzoate hydroxylase	hydroxylation epoxidation	FAD	NAD(P)H
B	cyclohexanone monooxygenase	Bayer-Villiger *N*-oxidation	FAD	NADPH
C	luciferase	light emissions *S*-oxidation Bayer-Villiger	–	FMN NAD(P)H
D	4-OH-phenylacetate 3-hydroxylase	hydroxylation	–	FAD NAD(P)H
E	styrene monooxygenase	epoxidation	–	FAD NAD(P)H
F	tryptophan 7-halogenase	halogenation	–	FAD NAD(P)H

The classification of FAD-dependent monooxygenases is based on structural features as well as on the type of electron donor and oxygen transfer.[65,67] The prototypic enzyme 4-hydroxyphenylacetate 3-hydroxylase (4HPA3H or HpaB, EC 1.14.14.9) catalyzes the reaction of 4-hydroxyphenylacetate with oxygen to 3,4-dihydroxyphenylacetate. During the reaction, water is formed by consumption of the cofactor FADH2, which is produced from FAD by an additional reductase component (HpaC, E.C. 1.5.1.36).[68]

Because its highly specific reaction is not restricted to the natural substrate 4-hydroxyphenylacetate, 4HPA3H is an attractive biocatalyst for challenging oxidation of monophenols to σ-diphenols (catechols).[66,69] In 2010 Coloumbel and co-workers published transformations of 4-halophenols to 4-catechols by 4HPA3H from *Escherichia coli* as living whole-cell biocatalysts.[70] The same enzyme was used for the quantitative oxidation of *p*-coumaric acid (**1**) to caffeic acid (**2**) as part of a reconstructed plant biosynthetic pathway by Lin and Yan[71], especially because all corresponding enzymes identified in plants are cytochrome P450-dependent monooxygenases which proved to be difficult to express functionally in *E. coli*.

In our own studies, we have applied 4HPA3H-mediated oxidation of *p*-coumaric acid (**1**) for *in vitro* synthesis of vanillin by a cascade of immobilized enzymes (chapter III). Notably, 4HPA3H-mediated hydroxylation proceeds without side reactions to undesired by-products. However, despite its efficiency and reaction specificity, applicability of 4HPA3H in biotechnology is still limited because a toolbox of enzymes with extended substrate scope, i.e. including natural products

other than simple monophenols (such as flavonoids or alkaloids), is not yet available.

In this chapter, the rational optimization of 4HPA3H from *E. coli* for conversion of a broad range of monophenols, and an application of the improved enzymes in *in vitro* and *in vivo* production of the corresponding catechols is presented.

3.2 Results and discussion

3.2.1 Cloning and transformation of genes

The 4HPA3H and HpaC genes from *E. coli* as well as the PrnF gene from *Pseudomonas protegens* CHA0 were cloned separately into a pET28a(+)-vector sustaining a kanamycin resistance and an N-terminal His$_6$-tag for purification via affinity chromatography (see section 2.4). The constructed plasmids were transformed into *E. coli* DH5α cells for storage and into *E. coli* BL21(DE3) cells for expression.

3.2.2 Substrate spectrum of wild-type 4HPA3H

The substrate requirements of 4HPA3H in *E. coli* as a whole-cell biocatalyst were probed and proved to be well suited for conducting hydroxylation reactions.[70-72] The expression strain BL21(DE3) was applied producing high levels of the recombinant monooxygenase (up to 248 mg l^{-1} of culture) under auto-inducing conditions (see section 2.4.1), but in contrast to previous studies[69-71] the cells were not co-transformed with a plasmid harboring the gene for a flavin reductase component (the endogenous activity proved to be sufficient to enable *in vivo* transformation of substrates, see section 2.5.3). Suspensions of the cells were tested for hydroxylation of a panel of eight phenolic compounds (Fig. 3.1) which belong to different classes of natural products: we included (I) phenolic acids and ketones such as *p*-coumaric acid (**1**), ferulic acid (**3**), rheosmin (**15**) and *p*-hydroxybenzoic acid (**9**), (II) polycyclic compounds such as the hydroxycoumarin umbelliferone (**13**), the stilbene resveratrol (**5**) and the flavonoid naringenin (**7**). Additionally, 2-hydroxycarbazole (**11**) was included, a readily available alkaloid-like substrate resembling carbazoles from plant (*Clauralia* alkaloids)[73] or bacterial origin (carazostatin, carbazomadurins).[74]
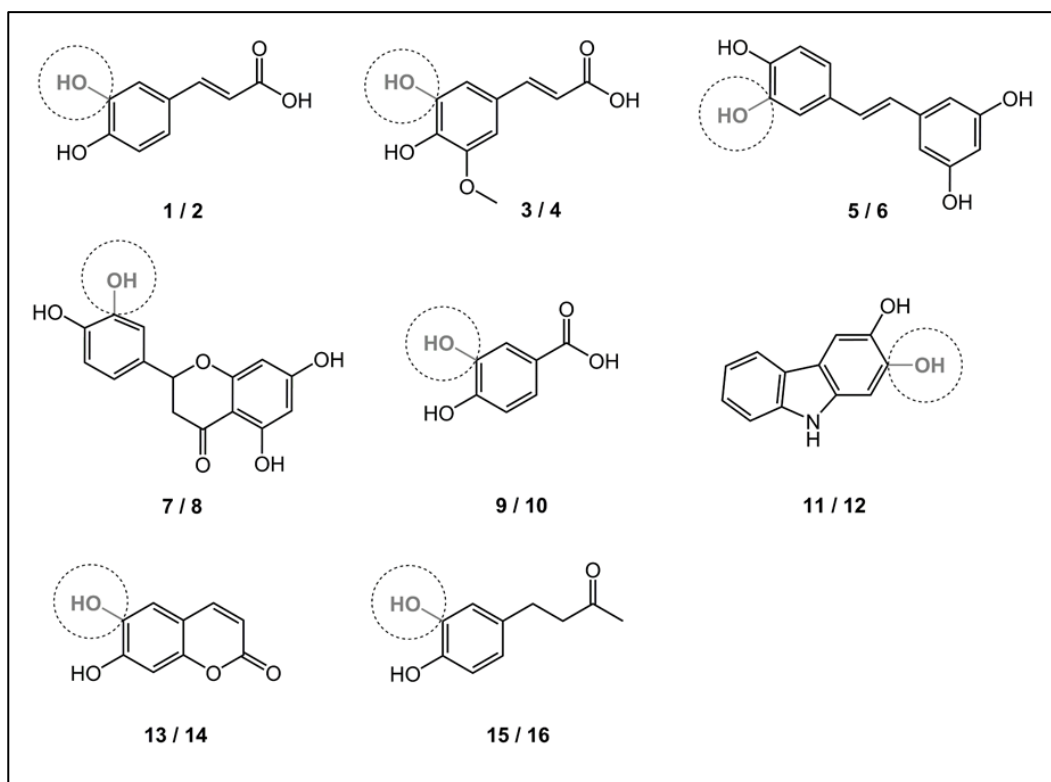
Figure 3.2: Substrates (odd numbers) used in hydroxylation reactions and corresponding catechol products (even numbers). The hydroxyl groups introduced into the monophenols by 4HPA3H-mediated oxidation are depicted in dark grey and marked by circle. **1** *p*-coumaric acid, **2** caffeic acid, **3** ferulic acid, **4** 5-hydroxyferulic acid, **5** resveratrol, **6** piceatannol, **7** naringenin, **8** eriodictyol, **9** *p*-hydroxybenzoic acid, **10** 3,4-dihydroxybenzoic acid, **11** 2-hydroxycarbazole, **12** 2,3-dihydroxycarbazole, **13** umbelliferone, **14** esculetin, **15** rheosmin, **16** 4-(3,4-dihydroxyphenyl)-butan-2-one.

Similar to its known substrate **1**[71], the cells also converted the compounds **13** and **15** with high efficiency (100 %, Fig. 3.2). Remarkably, the biocatalysts were also capable of hydroxylation of **5**, **9** and **11**, although lower yields of catechol products were obtained (78, 39 and 9 %). A potential involvement of other endogenous monooxygenases in these conversions was ruled out by control reactions (i.e. application of cells harboring the empty vector), which showed no conversion. In contrast to the substrates mentioned above, ferulic acid (**3**) and naringenin (**7**) were not accepted, most probably because of their bulky methoxy (**3**) or chromanone substituent (**7**).

We also tested 4HPA3H for conversion of 4-methoxycinnamic acid and 4-aminocinnamic acid. These two analogues of *p*-coumaric acid (**1**) were no substrates for the enzyme (data not shown). Hence, in accordance with a previous study (which showed that 4HPA3H enzymes are unable to hydroxylate cinnamic acid)[69], the presence of a phenolic hydroxyl group in the substrate seems to be mandatory for enzymatic turnover.
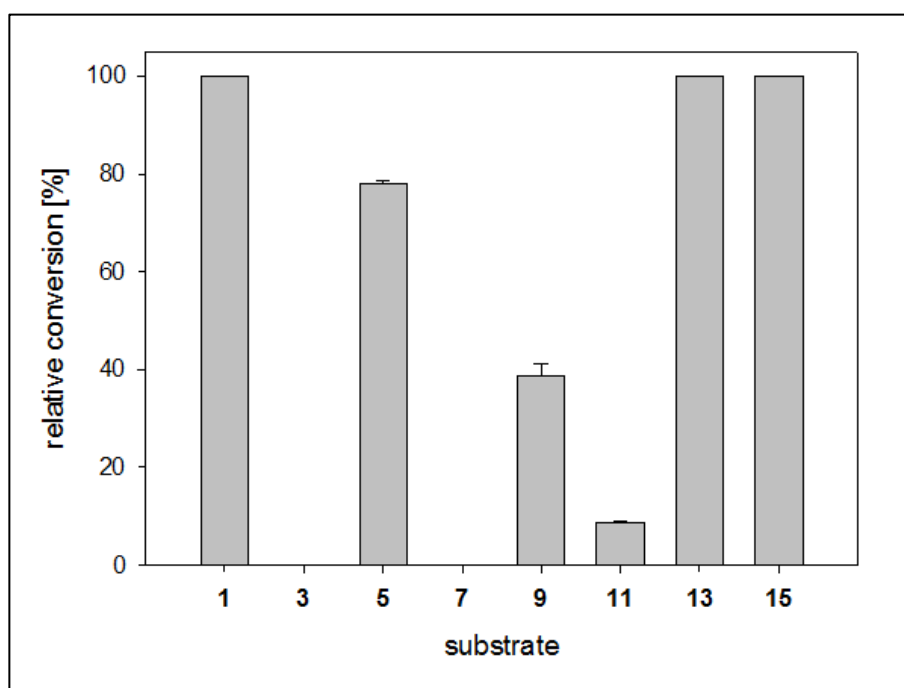
Figure 3.3: Hydroxylation with life whole-cell biocatalysts. Substrates (200 µM) were added to suspensions of bacterial cells expressing 4HPA3H wildtype enzyme. After incubation for 16 hours, catechols and residual substrates were extracted and analyzed by HPLC as described in section 2.5.3.

3.2.3 Optimization of 4HPA3H for hydroxylation of bulky substrates

Unlike the 4HPA3H enzyme used in this study the homologous protein from *Pseudomonas aeruginosa* was reported to readily hydroxylate ferulic acid (**3**).[69] Thus, we speculated whether comparison of both enzymes might yield structural features responsible for this altered substrate scope – which are prerequisites for a rational engineering of 4HPA3H specificity.

In order to gain insight into the architecture of the individual active sites, the three dimensional structures of the proteins were modeled based on crystallographic data of the related enzyme from *Thermus thermophilus* (PDB accession 2YYJ).[72] Both models were aligned to the template, which also contains the substrate 4-hydroxyphenylacetate and FAD (Fig. 3.3, see section 2.7).
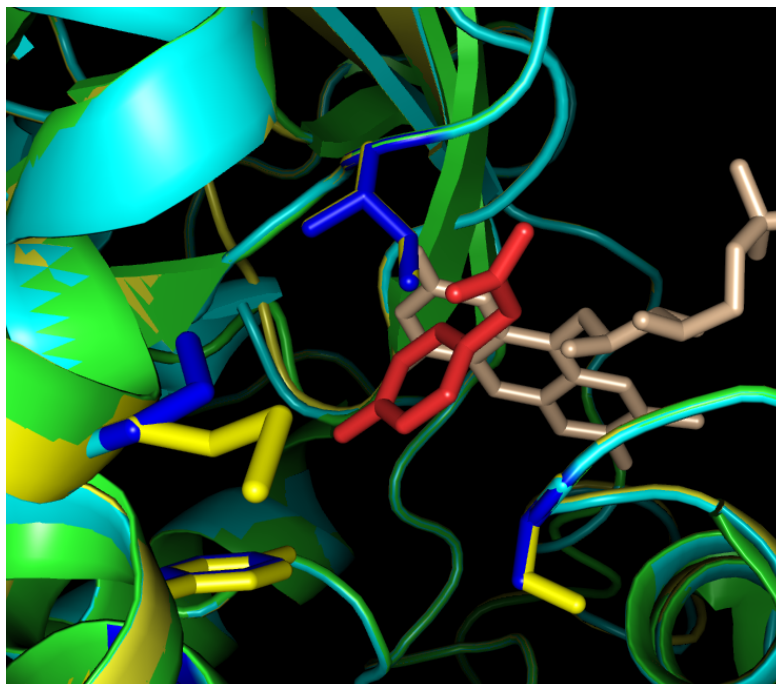
Figure 3.4: Alignment of crystal structure of 4HPA3H from *Thermus thermophilus* (PDB accession 2YYJ) (green) with the models of the enzymes from *E. coli* (pale yellow) and *Pseudomonas aeruginosa* (light blue). The models were generated with the tool Phyre2 as described in the Experimental Section, and aligned by the software PyMol (version 1.2). The amino acid residues in close vicinity to the substrate 4-hydroxyphenylacetate (red) which were subjected to mutagenesis are shown as stick representations in darker coloration (top down, clockwise: I/V157, S/A462, Y/F301, M/P293). Ivory: Riboflavin moiety of FAD.

It was inspected the active sites for residues which are proximate to the aromatic ring of 4-hydroxyphenylacetate, and identified three positions which differ in the enzymes from *E. coli* (M293, Y301, S462) and *Pseudomonas aeruginosa* (P293, F301, A462) only. Remarkably, all these positions are occupied by less voluminous residues in the latter enzyme, resulting in a more spacious active site. Thus, we substituted those residues in the enzyme from *E. coli* to shift the substrate scope in favor of more bulky substrates, yielding the variants M293P, Y301F and S462A. For similar reasons, isoleucine 157, which seems to directly interact with the phenolic moiety of 4-hydroxyphenylacetate (Fig. 3.3), was exchanged against valine. All variants created by site-directed mutagenesis are listed in Table 3.1.

Table 3.2: Overview of all generated single and double variants of the *E. coli* enzyme 4HPA3H.

amino acid residues exchanged in *E. coli* 4HPA3H	
single variant	*double variant*
Y301F	Y301F/S462A
Y301L	Y301L/S462A
Y301I	Y301I/S42A
S462A	Y301F/I157V
M293P	S462A/I157V
I157V	

Similar to the wild-type enzyme, the variants were applied for *in vivo* hydroxylation reactions after appropriate expression was confirmed by SDS-PAGE (see Appendix). Remarkably, the enzymes 4HPA3H-Y301F and -S462A were comparably active towards substrates which were also well accepted by the wild-type enzyme (**1**, **13**, **15**), whereas the variants M293P and, in particular, I157V showed reduced activity (Fig. 3.5). The variants are characterized by an altered substrate specificity: the enzyme Y301F proved to be superior in conversion of the substrates **5**, **9** and **11**. Strikingly, most variants were also capable of oxidation of naringenin (**7**) and, to a minor extend, of ferulic acid (**3**). This effect was even more pronounced in some double variants, which produced up to 26 (I157V/Y301F) and 30 % (Y301F/S462A) of catechol from the bulky substrates 2-hydroxycarbazole (**11**) and naringenin (**7**), respectively (Fig. 3.5).

As especially the substitution Y301F seems to be have the greatest impact on the acceptance of challenging compounds, we decided to introduce a leucine or isoleucine into this position. Both amino acids are similar to phenylalanine with respect to hydrophobicity, but are smaller in size.[75] Strikingly, the preferred substrate of the resulting mutant enzymes was ferulic acid (**3**) (which was converted up to 45 % in particular by 4HPA3H-Y301I), excelling well-known substrates such as *p*-coumaric acid (**1**) (Fig. 3.4). Obviously, subtle changes in size and hydrophobicity of the active site – e.g. the formal replacement of a hydroxyl group by a hydrogen atom in the variants Y301F and S462A only – are sufficient to change the substrate scope of 4HPA3H enzymes.

A similar sensitivity to single amino acid substitutions was observed for the structurally unrelated flavoenzyme *m*-hydroxybenzoate hydroxylase from *Comamonas testosteronei*.[76] Several variants obtained by error-prone polymerase chain reaction were capable of conversion of the phenols lacking the carboxyl group of the natural substrate. This altered substrate preference was

attributed to minor changes in second-shell positions which influence the positioning of active site residues.
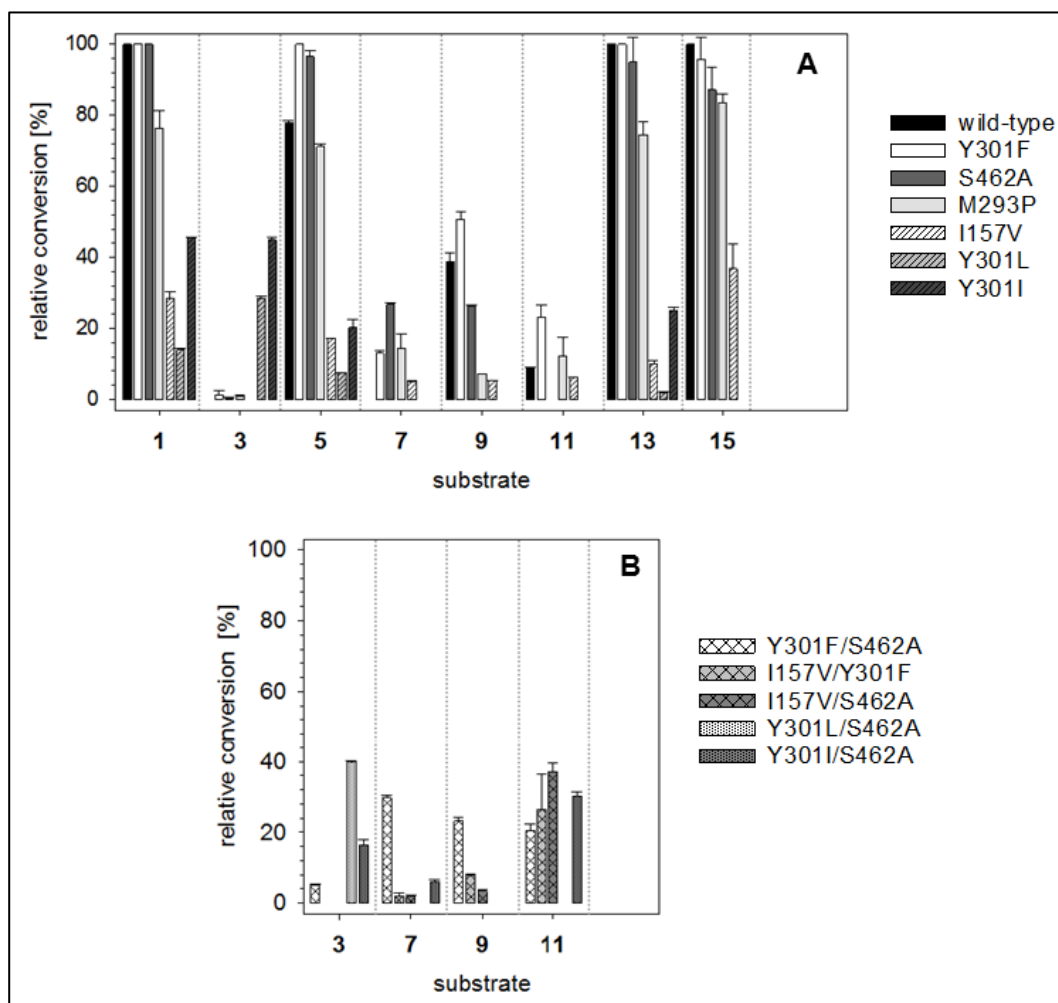
Figure 3.5: Hydroxylation with life whole-cell biocatalysts. Substrates (200 µM) were added to suspensions of bacterial cells expressing 4HPA3H or its single (A) and double variants (B). After incubation for 16 hours, catechols and residual substrates were extracted and analyzed by HPLC as described in section 2.5.3.

3.2.4 Product specificity of 4HPA3H and variants

To gain structural information on the hydroxylation products, whole-cell biotransformations in a large scale (500 mL of cell suspension) were performed (see section 2.5.4.1). Each substrate was converted by the enzyme variant that showed highest conversion in the analytic *in vivo* reactions (see section 3.2.3). Similar to those transformations, the catechol products were extracted from the reaction mixture after 16 hours of incubation. The compounds, which were purified by preparative reserved-phase HPLC, were identified as the catechols shown in Fig. 3.1 by ^{1}H-NMR, MS and MS[2] measurements (see section 2.6.4, 2.6.5 and 2.6.7). Noteworthy, 4HPA3H and its variants demonstrated strict regio- and product-specificity. In contrast to the enzyme from *Pseudomonas aeruginosa*, which performs sequential hydroxylation[69], no further conversion of the catechols into pyrogallol-like polyphenols was observed for both the wild-type enzyme and its variants.

3.2.5 Activity of 4HPA3H and variants

In order to characterize 4HPA3H and its variants with respect to specific activity, the development of a convenient *in vitro* assay was necessary. As the enzymes rely on a reductase component, which provides the cofactor $FADH_2$ (Fig. 3.6), we focused on screening of a suitable reductase enzyme in the first step. The flavin reductase HpaC from *E. coli* – the endogenous partner of 4HPA3H – was not accessible due to formation of mainly insoluble protein under a variety of expression conditions (data not shown). In the search for homologous proteins, a putative flavin:NAD(P)H reductase PrnF from *Pseudomonas protegens* Pf-5 (NCBI accession AAY91318.1) which is a part of the pseudomonad two-component arylamine oxygenase system[77,78] showed to have high sequence identity (99 %) to our present strain *Pseudomonas protegens* CHA0. Recombinant PrnF proved to be well expressed in *E. coli* (see section 2.4.1, in a yield of 175 mg l^{-1} of culture). Similar to literature data[77], purified protein (see Appendix) was highly active in reduction of FAD with NADH (16,660 ± 1,600 U mg^{-1}).

For assaying 4HPA3H in microplate scale (section 2.5.1.1), 4HPA3H was incubated with PrnF, the substrate *p*-coumaric acid (**1**) and the formate dehydrogenase (FDH) from *Candida boidinii*. We introduced this enzyme for regeneration of the costly cofactor NADH:NAD$^+$ (654 € per mg, Sigma Aldrich). During the reaction it is recycled with sodium formate, yielding a sequential three-enzyme cascade for hydride transfer (Fig. 3.6).

HCOOH   NAD   FAD   R-H + O_2

FDH   *PrnF*   *4HPA3H*

CO_2 NADH $FADH_2$ R-OH + H_2O
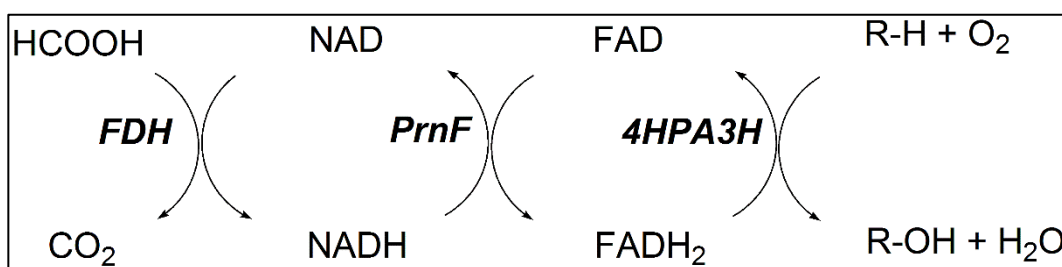
Figure 3.6: Scheme of enzyme cascade for *in vitro* hydroxylation. For recycling of hydride-transferring cofactors, FDH from *Candida boidinii* and the flavin reductase PrnF from *Pseudomonas protegens* were applied.

The conversion of *p*-coumaric acid (**1**) was measured spectrophotometrically by a previously established catechol complexation method, which could be applied for quantitative determination of the product caffeic acid (**2**) (see Appendix). We assayed 4HPA3H under optimized conditions (see next section), yielding a specific activity (6.2 ± 0.5 U mg^{-1}) comparable to this from literature.[71] The two variants Y301F/S462A and Y301I – which were also produced recombinantly (in yields of 295 and 112 mg l^{-1} of culture, respectively) and applied as purified enzymes (see Appendix) – showed reduced activities (1.6 ± 0.3 U mg^{-1} and 1.3 ± 0.6 U mg^{-1}) towards *p*-coumaric acid (**1**). However, similar to the data obtained from whole-cell biotransformations (see section 3.2.3), these variants were beneficial in the conversion of ferulic acid (**3**), which is also reflected by comparably high conversion rates (4HPA3H-Y301F/S462A: 1.06 ± 0.31 U mg^{-1}, 4HPA3H-Y301I: 4.49 ± 0.20 U mg^{-1}).

3.2.6 *In vitro* production of catechols by 4HPA3H/PrnF

In a multitude of biocatalytic conversions, cell-free (*in vitro*) systems proved to be superior to reactions with whole-cell biocatalysts, especially with emphasis on substrate depletion by and toxicity to the used microorganism[79], side reactions[79], and facile product purification.[80] In order to develop an effective and low-cost enzyme system for *in vitro* hydroxylation, we optimized several parameters of the 4HPA3H reaction with respect to maximum reaction rate. By means of our spectrophotometric assay (see section 3.2.4), we screened different buffer

systems and optimized the concentration of the components *p*-coumaric acid (**1**), FAD, NADH and methanol, the incubation temperature and oxygen transfer (i.e. by variation of the shaking frequency) (optimum conditions are given in section 2.5.1.1). In particular, this fine-tuning yielded (I) a reduction in the content of NADH from 200 μM (standard conditions according to reference[71]) to 25 μM without loss of activity (see Fig. 3.6), (II) tolerance of high concentrations of substrate (1 mM) and (III) of the co-solvent methanol (10 %, v/v) (which is beneficial due to enhanced substrate solubility). In contrast, the redox system was susceptible to high concentrations of FAD, which proved to be inhibitory (see Fig. 3.6). In contrast to the cofactors, the accompanying redox enzymes PrnF and FDH were used in excess due to their high specific activity.
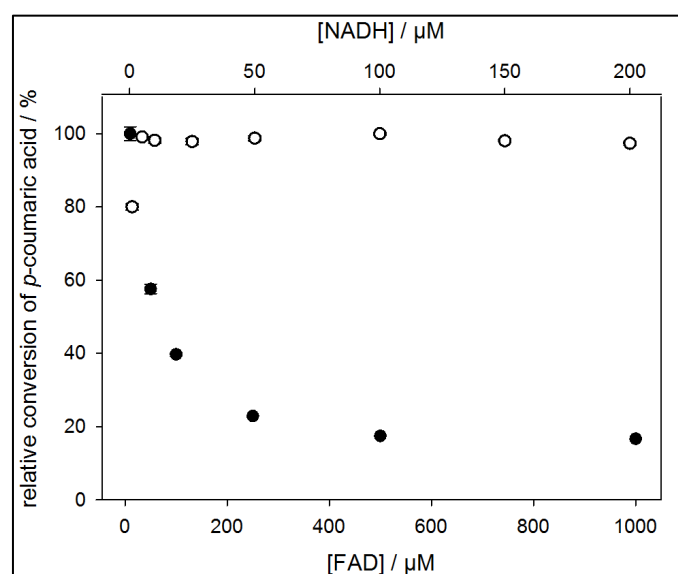
Figure 3.7: Dependence of the *in vitro* conversion of *p*-coumaric acid (1) on the FAD (filled circles) and NADH concentration (open circles). The reactions were performed using the 4HPA3H/PrnF/FDH system as specified in section 2.5.1.1.

Subsequently, we applied the optimized reaction conditions to the *in vitro* oxidation of the substrates shown in Fig. 3.1. In accordance with the whole-cell transformation experiments (Fig. 3.2A), product yields in the conversion by the wild-type 4HPA3H (Fig. 3.7A) ranged from high (**1**, **13**, **15**) or moderate (**5**) to low (**9**, **11**), whereas ferulic acid (**3**) was not oxidized. Consistently, the mutant enzymes Y301F/S462A and Y301I (Fig. 3.7B) proved to be optimal for transformation of the challenging substrates **3** and **7**. However, in contrast to the experiments with life biocatalysts, how a low conversion of naringenin (**7**) even by the wild-type 4HPA3H (25 % yield) was observed (which might be explained by increased substrate availability). The hydroxycarbazole (**11**) was not accepted by the double variant Y301F/S462A, probably because of the high substrate

concentration which seems to reflect strong product inhibition – an effect which has previously been observed for other 4HPA3H enzymes.[69] Noteworthy, HPLC analysis (section 2.6.4) of the reaction products confirmed product specificity. All conversions – except this of naringenin (**7**) by 4HPA3H-Y301I – yielded the previously identified catechols (Fig. 3.1) only.
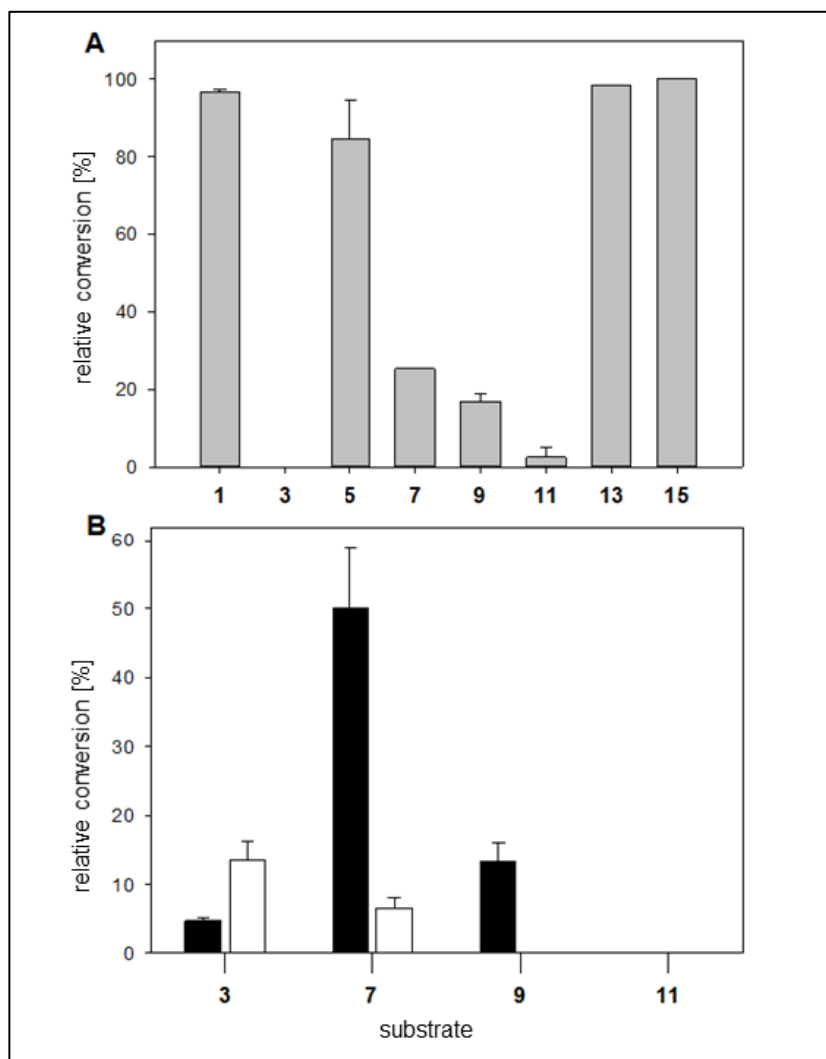
Figure 3.8: *In vitro* hydroxylation by the 4HPA3H/PrnF/FDH system. Purified 4HPA3H (A) or variants (B, Y301I depicted in white and Y301F/S462A depicted in black) were used for conversion of substrates (1 mM) under optimized conditions as specified in section 2.5.1.

3.3 Conclusion

In this chapter the first report on rational design of a flavin-dependent monooxygenase was presented. It was shown that single substitutions, mainly at amino acid position Y301, are sufficient to engineer new selectivities into 4HPA3H, yielding enzymes, which are optimal for conversion of bulky and sterically challenging substrates. This strategy afforded robust biocatalysts, which provide an alternative to existing cytochrome-dependent enzymes, especially due to high yield, strict regio- and product specificity and applicability in *in vitro* and *in vivo* (whole-cell hydroxylation) reaction systems.

In contrast to cytochrome-dependent P450-enzymes, flavin-dependent monooxygenases appear to be relatively abundant in prokaryotic genomes.[40]

4   Chapter II: Enzymatic total synthesis of sinapyl alcohol and related monolignols from *p*-coumaric acid

4.1   Introduction

Lignocellulotic biomass is the most abundant sustainable raw material on earth. Apart from polysaccharides (cellulose, hemicellulose), its main constituent is lignin.[81] While the economic interest concerning the polysaccharides is focused on biofuel production,[81] the lignin building blocks (monolignols) are considered as valuable source for high-value polymeric materials,[82,83] aromatic monomers like vanillin[34] or dimeric lignans with potential application in human disease treatment, e.g. silbyin has hepatoprotective properties.[84] Lignans, an abundant and diverse group of secondary metabolites, have been studied intensively for their diverse structures and variety of biological activities. This structurally diverse organic compounds are characterized by a basic backbone modified in multiple ways and are generally derivatives of phenylpropanoids created naturally via peroxidases and laccases enzymes.[85] Many subgroups of lignans are well known because of their interesting biological, pharmacological or medicinal activities. For the dibenzylbutyro-lactone subgroup, e.g. anti-cancerous, antioxidant, antimicrobial and hepatoprotective biological activities were reported.[85,86]

However, the major amount of the 50 million tons of lignin isolated from pulping processes in 2010 was burned as low-value fuel and only 2 % was used for specialty products.[82] The main barrier for the economic utilization as feedstock for aromatic compounds is the complex and highly crosslinked structure of lignin and its natural linkage to cellulose and hemicellulose. Biocatalytic *in vitro* strategies for the production of pure monolignols would thus constitute interesting alternatives to the extraction from lignin.
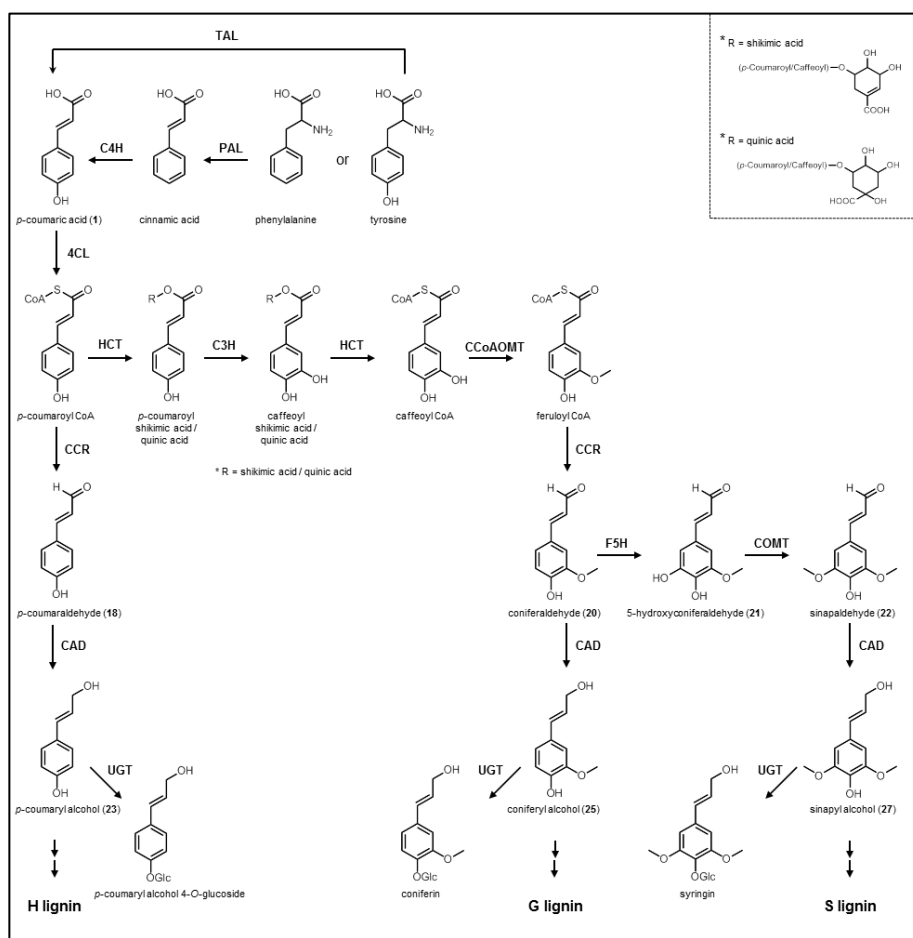
Figure 4.1: Biosynthetic pathway of monolignols in angiosperm. Only the predominant route toward the three main monolignols is shown. TAL: tyrosine ammonia-lyase, PAL: phenylalanine ammonia-lyase, C4H: cinnamate 4-hydroxylase, 4CL: 4-coumarate:CoA ligase, HCT: *p*-hydroxycinnamoyl-CoA:quinate shikimate *p*-hydroxycinnamoyltransferase, C3H: *p*-coumarate 3-hydroxylase, CCoAOMT: caffeoyl-CoA *O*-methyltransferase, CCR: cinnamoyl-CoA reductase, F5H: ferulate 5-hydroxylase, COMT: caffeic acid/5-hydroxyconiferaldehyde *O*-methyltransferase, CAD: cinnamyl alcohol dehydrogenase, UGT: UDP-glucosyltransferase. Figure is modified according to Vanholme et al.[88]

The biosynthesis of monolignols (Fig. 4.1) in plants is well understood and differs depending on plant species as well as time and tissue.[87,88] Monolignol biosynthesis is initiated by the formation of the central intermediate **1** either directly by deamination of L-tyrosine (monocotyledons) or indirectly by deamination of L-phenylalanine and subsequent hydroxylation by cinnamate 4-hydroxylase (C4H, dicotyledons).[87] The monolignols are derived from *p*-coumaric acid (**1**) by modifying both parts of the molecule, the aromatic ring and the aliphatic carboxylate.[87] For the biosynthesis of the simplest monolignol, *p*-coumaryl alcohol (**23**), the carboxylic acid is activated by 4-coumarate-CoA ligase (4CL) and subsequently reduced by cinnamoyl-CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD). The higher oxidized monolignols are

produced by sequential hydroxylation and methylation of the aromatic C3 and C5 positions. Hydroxylated monolignols are produced by sequential oxidation, which is often succeeded by methylation of the hydroxyls at C3 and C5 positions. *p*-Coumarate 3-hydroxylase (C3H) introduces the first hydroxyl group into *p*-coumaric acid (**1**, e.g. in *Populus trichocarpa*) or its shikimic ester (e.g. in *Arabidopsis thaliana*).[89] Contrarily, hydroxylation and methylation at C5 by ferulate 5-hydroxylase (F5H) and caffeic acid/5-hydroxyconiferaldehyde *O*-methyltransferase (COMT) predominantly occur on the aldehyde or alcohol level. This hydroxylation and methylation steps essentially determine the contribution of guaiacyl and syringyl monomers to the gymnosperm and angiosperm pattern.[16] This native biosynthetic route is very inconvenient for a biotechnological process not only due to the different reduction states, at which the aromatic modifications occur. The involvement of instable and membrane-bound P450 enzymes like C3H and F5H makes their application as purified enzymes challenging.[71] Additionally it was shown that interaction of C3H and C4H is essential to exhibit full activity and that these P450s co-localize in multiprotein complexes together with cytochrome P450 reductase (CPR).

In this Chapter a convenient, artificial pathway for the enzymatic monolignol production starting with *p*-coumaric acid is presented (Fig. 4.2). This pathway comprises the sequential hydroxylation-/*O*-methylation events on the carboxylic acid level (Fig. 4.2A) and the subsequent reduction of all methoxylated derivatives by a uniform enzyme cascade (Fig. 4.2B). All applied enzymes are characterized by high stability and can be purified in high yields by heterologous expression in *E. coli* (see chapter 2.4). Thus, this artificial pathway might serve as platform for the industrial production of all native monolignols in cell-free systems as well as in metabolically engineered producer strains. Compared to a fermentative approach the utilization of purified enzymes provides a more adjustable process environment and does not suffer from yield decrease due to unspecific metabolization of substrates, products or reactive intermediates.
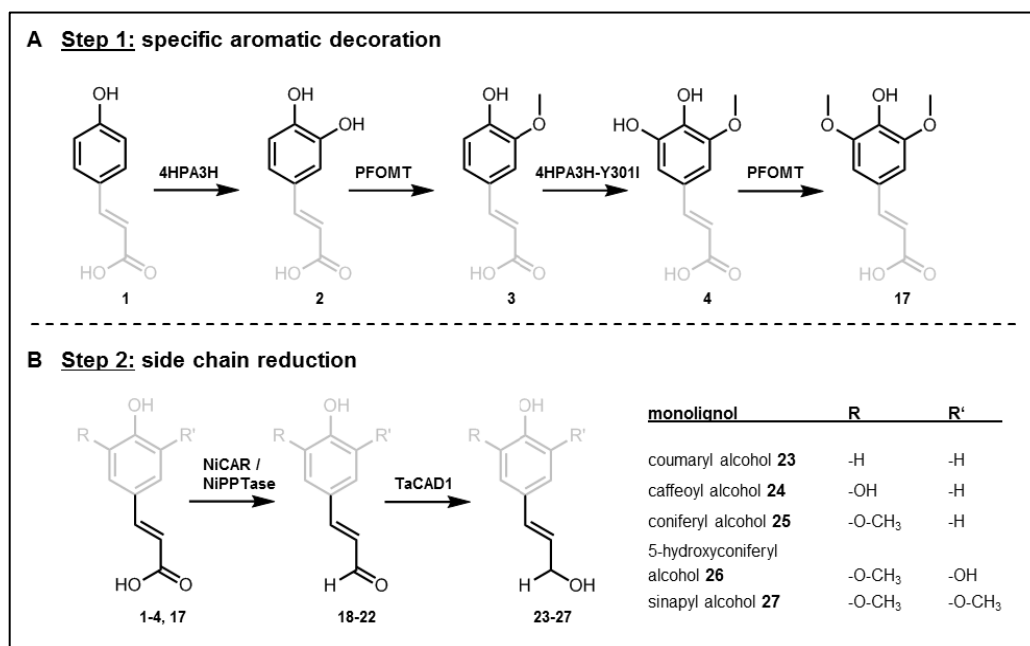
Figure 4.2: Artificial pathway for *in-vitro* biosynthesis of monolignols. A) Specific decoration of the aromatic ring is achieved on the carboxylic acid level. Access to C5-modified compounds is strictly controlled by application of 4HPA3H-Y301I. B) Reduction of all differently decorated cinnamic acid derivatives (**1**-**5**, **17**) into the cinnamyl alcohols (**23**-**27**) is achieved using substrate promiscuous NiCAR and TaCAD1.

4.2 Results

4.2.1 Cloning and transformation of genes

Cloning and transformation of the 4HPA3H and the PrnF genes as well as the site-directed mutagenesis for creating the 4HPA3H variant (Y301I) were already mentioned in chapter I.

The NiCAR and NiPPTase genes from *Nocardia iowensis*, the TaCAD1 gene from *Triticum aestivum* and the ScFDH as well as the G6PDH gene from *S. cerevisiae* pJ69-4A were cloned separately into a pET28a(+)-vector sustaining a kanamycin resistance and an N-terminal His$_6$-tag for purification via affinity chromatography (see section 2.4).

The plasmids of the *O*-methyltransferase from *Mesembryanthemum crystallinum* (PFOMT) and the 4-coumarate-CoA ligase from *Arabidopsis thaliana* (At4CL2) were already available from previous studies.

The synthetic gene of the cinnamoyl-CoA reductase from *Leucaena leucocephala* (LlCCR, constructed by DNA2.0) was amplified from the vector pJ411 and cloned into the pET28a(+)-vector (see above) because of problems during the gene expression of LlCCR under different expression conditions.

The constructed and the even existing plasmids were transformed into *E. coli* DH5α cells for storage and into *E. coli* BL21(DE3) cells for expression.

4.2.2 Expression and purification

After successful gene expression and cell lysis (section 2.4), the proteins were isolated using Immobilized Metal Ion Affinity Chromatography (IMAC). The success of expression, purification and dialysis of the used proteins was visualized by SDS-PAGE (see Appendix). The expression yields differ from adequate (e.g. 4 mg l^{-1} of G6PDH protein) to very good (e.g. 202 mg l^{-1} of PFOMT protein). All other expression yields are listed in section 2.4.1.

4.2.3 Identification of biocatalysts for specific, sequential hydroxylation

Regiospecific hydroxylation is a prerequisite to access differently hydroxylated/methoxylated monolignols. In particular, a biocatalyst with strict specificity for 3'-hydroxylation of *p*-coumaric acid (**1**) and no activity against caffeic (**2**) or ferulic acid (**3**) is mandatory if the mono-substituted derivatives shall be produced. The biosynthesis of caffeic acid (**2**) from *p*-coumaric acid (**1**) in *E. coli* using its endogenous hydroxylase complex was described previously in chapter I. This hydroxylase complex is a two-component flavin dependent monooxygenase system consisting of an oxidase (4-hydroxyphenylacetate 3-hydroxylase, 4HPA3H) catalyzing the aromatic hydroxylation and a reductase (HpaC) that supplies the oxidase with reduced FAD (see chapter I, Fig. 3.4).[90] Besides *p*-coumaric acid, this enzyme was also shown to convert multicyclic phenylpropanoids like naringenin, resveratrol or umbelliferon giving high yields of the respective catechols *in vivo*.[18] Furthermore, a single substitution was sufficient to engineer new functionalities into the monooxygenase. The yielding biocatalyst provides an alternative to existing cytochrome-dependent enzymes especially appropriate to high yield, strict regio- and product-specificity and applicability. Recently, an assay for the *in vitro* conversion of *p*-coumaric acid (**1**) and related aromatic compounds by 4HPA3H was established (see chapter I). Unlike for the *in vivo* conversions, the *E. coli* reductase HpaC had to be replaced by PrnF, the homologous enzyme from *Pseudomonas protegens*, as HpaC was not stable in purified form. Using this approach, the quantitative conversion of 1 mM *p*-coumaric acid (**1**) into caffeic acid (**2**) is achieved without any detectable overoxidation byproduct. Strikingly, also no hydroxylation activity towards ferulic acid (**3**) is observed (Figure 4.3), which identify 4HPA3H as suitable biocatalyst for the production of the monosubstituted *p*-coumaric acid derivatives. In order to

get access to the tri-substituted derivatives, an enzyme with 5'-hydroxylation activity against ferulic acid (**3**) is required. The amino acid residues Y301 and S462 in the *E. coli* enzyme have been identified as key residues determining the substrate specificity for 4HPA3H (see chapter I). Single, conserved substitutions of these residues were sufficient to allow hydroxylation of sterically challenging substrates that were not accepted by 4HPA3H wildtype. Thus, we tested some variants bearing mutations at these positions for 5'-hydroxylation of ferulic acid (**3**, Figure 4.3) in *in vivo* assays and identified the single variant Y301I as superior catalyst for this conversion.

Compared to the wildtype enzyme, this variant shows a reduced activity against *p*-coumaric acid (**1**, wildtype: 6.2 ± 0,5 nmol min^{-1} mg^{-1}, 4HPA3H-Y301I: 1.3 ± 0.6 nmol min^{-1} mg^{-1}). However, a high activity towards ferulic acid (**3**, 4HPA3H-Y301I: 4.49 ± 0.2 nmol min^{-1} mg^{-1}) was observed for the variant, while the wildtype enzyme did not show any conversion of ferulic acid (**3**). Thus, the modified enzyme proved to be key for accessing the desired bi-substituted monolignols.
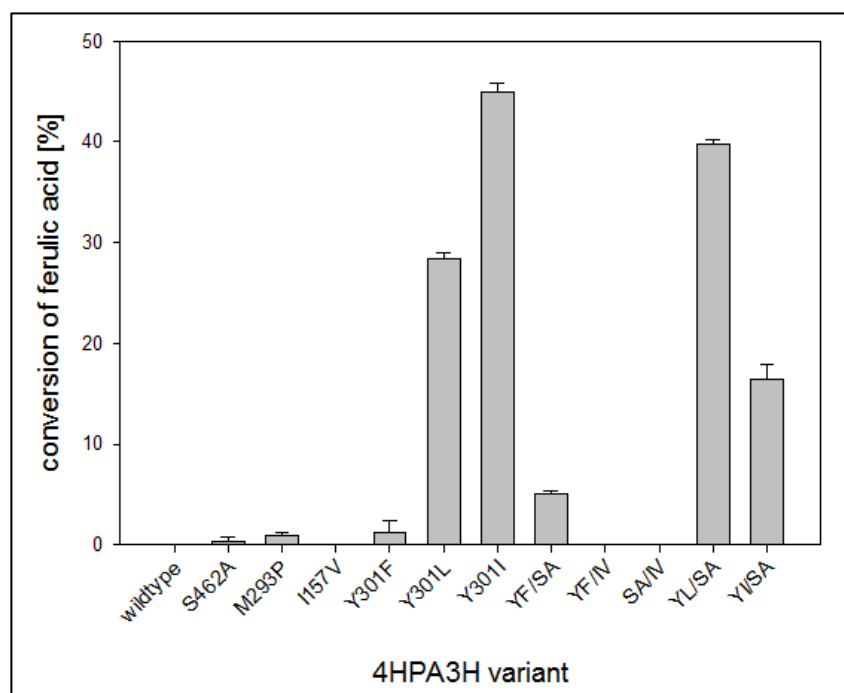
Figure 4.3: Hydroxylation of ferulic acid with 4HPA3H variants in *in vivo* assays. Ferulic acid (**3**, 200 µM) was added to suspensions of bacterial cells expressing 4HPA3H or its variants. After an incubation time of 16 hours, 5-hydroxyferulic acid (**4**) and residual ferulic acid (**3**) were extracted and analyzed by HPLC (see section 2.6.4).

4.2.4 Identification of a biocatalyst for regiospecific *O*-methylation

Both *O*-methylation steps during decoration of the coumaric acid scaffold require the regiospecific modification of a catecholic compound in meta-position. We choose the phenylpropanoid-flavonoid *O*-methyltransferase (PFOMT) from *Mesembryanthemum crystallinum* as potential candidate for these reactions, as this enzyme was shown to catalyze the *ortho*-methylation of different catechols, i.e. preferentially in the *meta*-position of phenylpropanoids (including caffeic acid (**2**) and 5-hydroxy ferulic acid, **4**).[91] In our investigations both **2** and **4** were quantitatively converted into the desired 3'-*O*-methylated product using PFOMT. In both cases, no 4'-*O*-methylation was observed as proved by authentic standards (data not shown).

4.2.5 One-pot process for the enzymatic production of differently decorated cinnamic acids

To verify, that a controlled formation of differently methoxylated monolignol precursors can be accomplished in one-pot multi-enzyme transformations, *p*-coumaric acid (**1**) was incubated with different combinations of the respective catalysts (catalyst set A to D), cosubstrates and accessory enzymes (Table 4.1).

Table 4.1: Different combinations of enzymes (catalyst set A to D) for a one-pot multi-enzyme transformation to obtain different cinnamic acids as monolignol precursor.

	A	B	C	D
4HPA3H	+	+	+	+
PFOMT	–	+	+	+
4HPA3H-Y301I	–	–	+	+
SAM	–	0.2 mM	0.2 mM	1 mM

p-Coumaric acid (**1**) was converted completely into caffeic acid (**2**) in a reaction assay containing only 4HPA3H wildtype (Fig. 4.4 and Fig. 4.5, catalyst set A). The quantitative production of ferulic acid (**3**) was observed when PFOMT was additionally included during the conversion (Fig. 4.4 and Fig. 4.5, catalyst set B). The addition of 4HPA3H-Y301I resulted in the formation of ferulic acid (**3**) and sinapic acid (**17**) (Fig. 4.4. and Fig. 4.5, catalyst set D). The relative distribution of these highly oxidized monolignol precursors can be modulated by adjustment of the cofactor SAM for the *O*-methylation steps: while up to 53 % of 5-hydroxyferulic acid (**4**) and no sinapic acid (**17**), but small amounts of non-converted ferulic acid (**3**) were achieved under low SAM-conditions (Fig. 4.4 and

Fig. 4.5, catalyst set C) the distribution shifts towards the methylated form, sinapic acid (**17**) under SAM-surplus conditions (Fig. 4.4, catalyst set D). Similar as described more in detail in section 5.2.1, we introduced a formate hydrogenase (FDH) from a *Saccharomyces* strain for regeneration of the costly cofactor $NADH:NAD^+$.[c]
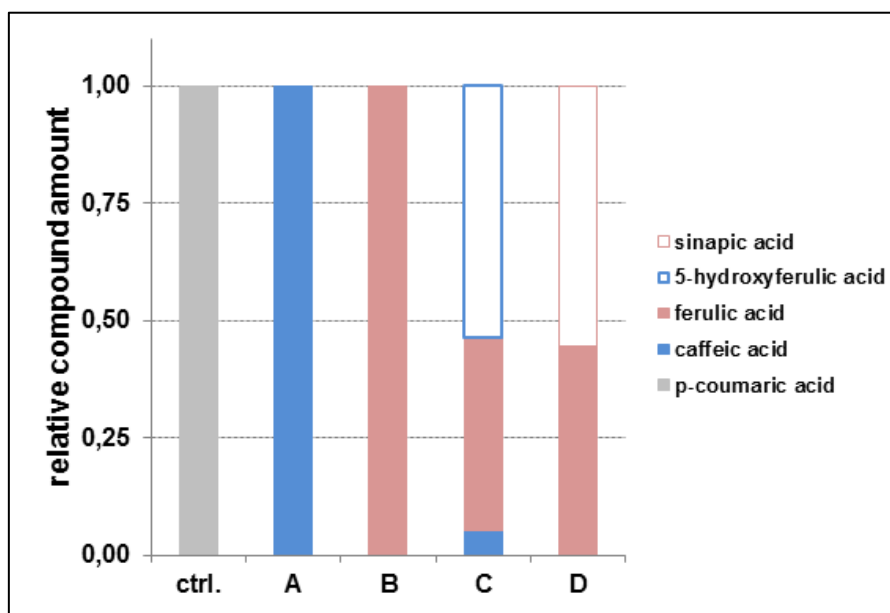
Figure 4.4: Formation of cinnamic acids after conversion of *p*-coumaric acid (1) with different enzyme combinations. For the abbreviations A to D see Table 4.1. In a control reaction (ctrl.) *p*-coumaric acid was incubated in reaction buffer for the total reaction time.

---

[c] Detailed results including error bars are presented in the Appendix.
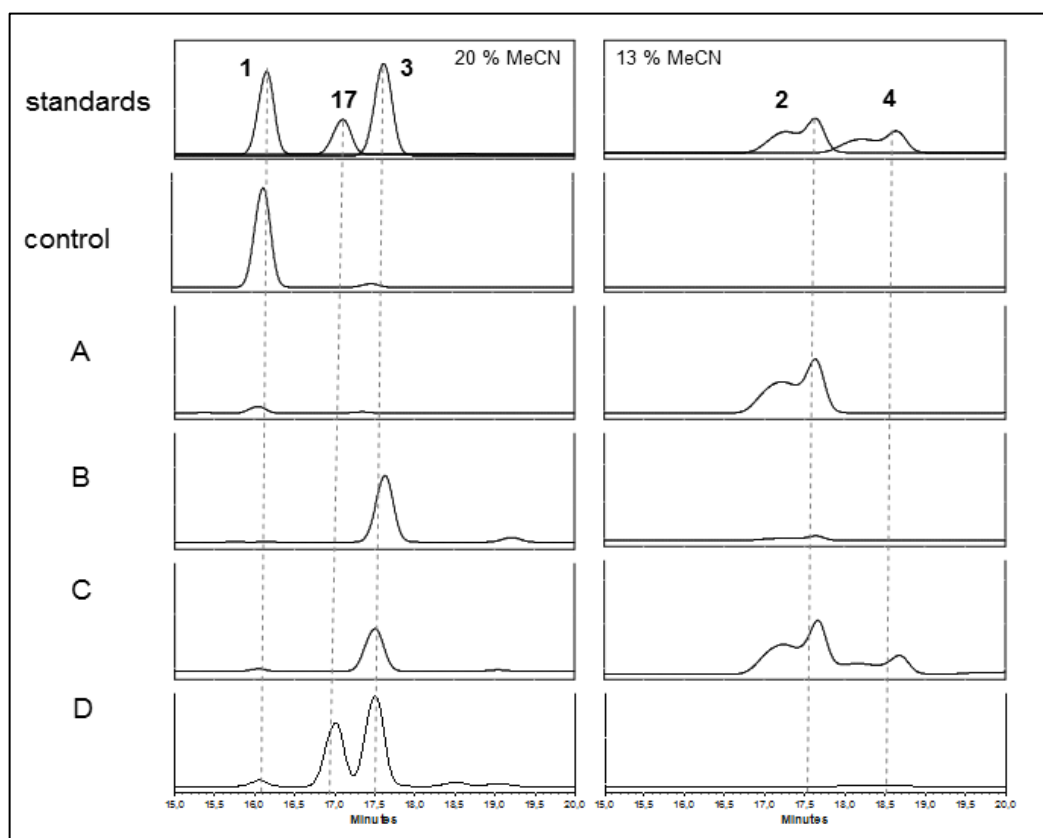
Figure 4.5: HPL chromatograms of decoration reactions. For the abbreviations A to D see Table 4.1. Two different gradients were required for a perfect differentiation of substrates and products: for separation of p-coumaric acid (**1**), sinapic acid (**17**) and ferulic acid (**3**) 20 % (v/v) acetonitrile was added for elution, for separation of caffeic acid (**2**) and 5-hydroxyferulic acid (**4**) 13 % (v/v) acetonitrile was used (see section 2.6.4).

4.2.6 Reduction of cinnamic acids based on an one-enzyme system

Reduction of a carboxylic acid to the respective aldehyde is a thermodynamically unfavorable process and nature evolved two enzymatic strategies for this kind of reaction.[92] In most cases - including monolignol biosynthesis in plants - a two-enzyme process is exploited: the reduction of *p*-coumaric acid derivatives is initiated by the ATP-dependent thiolation of the carboxylic acid precursor yielding an activated CoA-ester. These compounds are then reduced to the aldehydes by a second enzyme releasing CoA.[87] In previous studies, a 4-coumarate CoA-ligase (4CL2) from *Arabidopsis thaliana* was successfully used for the conversion of cinnamic acid derivatives.[93] But 4CL2 was not able to convert sinapic acid (**17**). Furthermore, the chosen cinnamoyl-CoA reductase from *Leucaena leucocephala*, which is well known for the conversion of cinnamoyl-CoA-esters to the related aldehydes, was not accessible due to formation of mainly insoluble protein under a variety of expression conditions (data not shown).

Alternatively, two examples from gram-positive microbes were reported that are able to catalyze this kind of reaction using a single protein called carboxylic acid reductase (CAR). While CAR from *Mycobacterium marinum* is a versatile enzyme for the conversion of aliphatic carboxylic acid,[29] the catalyst from *Nocardia iowensis* was reported to convert aromatic substrates like vanillic acid and ferulic acid.[94,95] Basically, the reaction mechanism of NiCAR resembles that of the two-enzyme system but with two catalytic domains on a single enzyme. The activated thio-ester is formed via a phosphopantetheine post-translational modification and remains covalently bound to the enzyme during the whole reaction cycle (see Fig. 4.6, modified for the reduction of *p*-coumaric acid (**1**)).[96] The reaction cycle was already described by Venkitasubramanian et al. in 2006 for the reduction of benzoic acid to benzaldehyde: After successful phosphopantetheinylation the catalytic cycle starts with the adenylation of benzoate at the N-terminus using ATP and yielding benzoyl-AMP mixed anhydride. Subsequently, the nucleophilic attack of the phosphopantetheine thiol at the carbonyl carbon of benzoyl-AMP releases AMP and the benzoyl thioester phosphopantetheinyl "arm" swings from the adenylating domain to the reduction domain. This leads to the reduction of thioester by NADPH and the release of benzaldehyde, $NADP^+$ and free phosphopantetheinylated CAR, which is ready for another reduction cycle.[96]
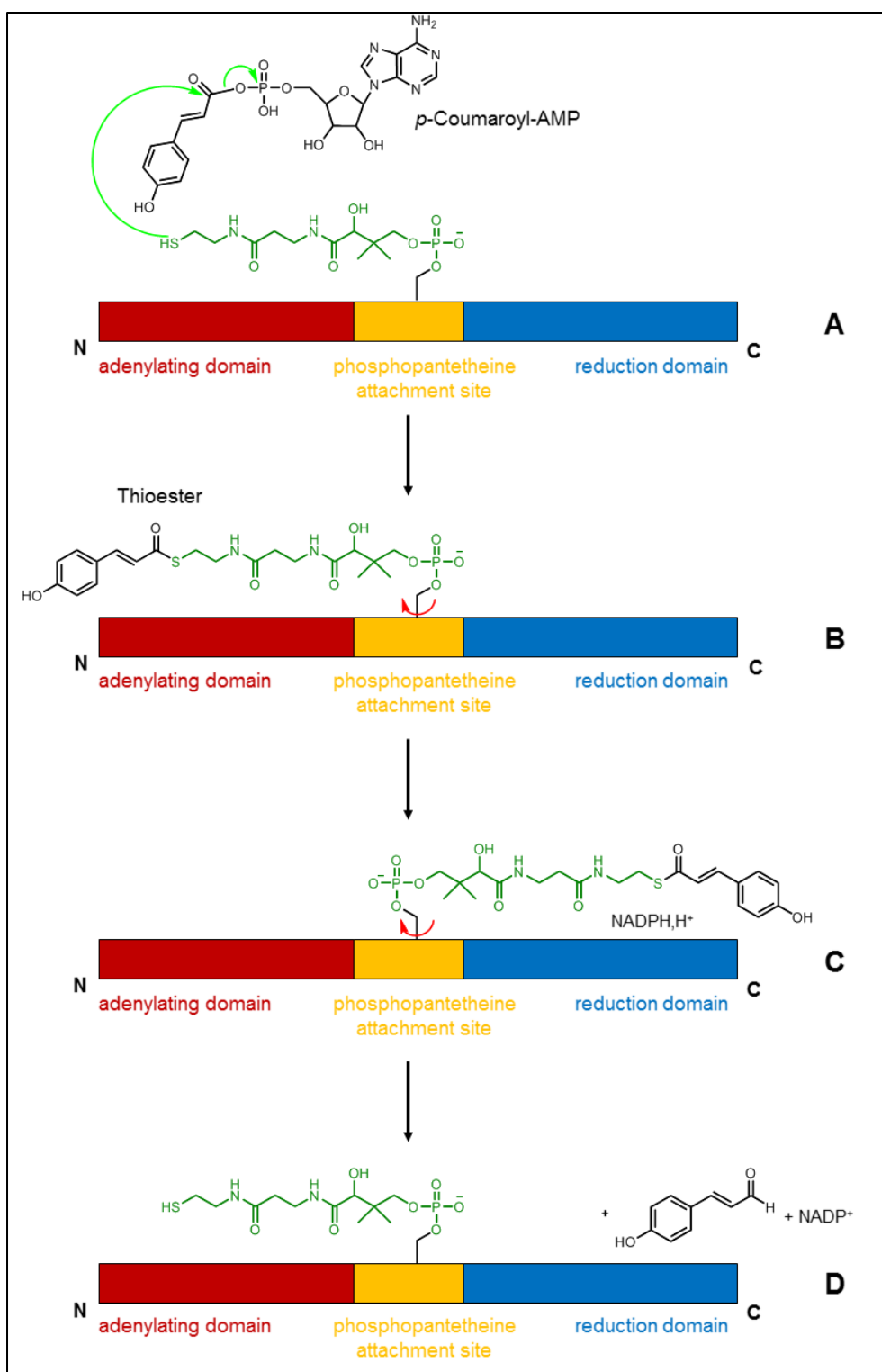
Figure 4.6: Catalytic cycle of phosphopantetheinylated NiCAR mediated-reduction of p-coumaric acid (1) to p-coumaraldehyde (18). A) Adenylation of p-coumarate yielding p-coumaroyl-AMP. Nucleophilic attack of the phosphopantetheine thiol at the carbonyl carbon of p-coumaroyl-AMP releases AMP. B) The p-coumaroyl thioester phosphopantetheinyl "arm" swings from the adenylating domain to the reduction domain. C) Leading to the reduction of thioester by NADPH and D) release p-coumaraldehyde, NADP[+] and free phosphopantetheinylated NiCAR ready for another reduction cycle. Figure is modified for the reduction of p-coumaric acid (1) according to Venkitasubramanian.[96]

We decided to substitute the native enzymes from the plant-derived pathway by NiCAR and tested the conversion of all cinnamic acid derivatives by this enzyme *in vitro*. No activity was observed in a reaction using purified NiCAR, indicating that *E. coli* is not able to phosphopantetheinylate this enzyme. Preincubation of purified NiCAR with the related phosphopantetheine transferase (NiPPTase) from *Nocardia iowensis* and coenzyme A yielded active enzyme and the formation of all expected aldehydes (**18**-**22**, Figure 4.8), which could be detected by LC-MS (see Appendix). For regeneration of the very costly cofactor NADPH (1,454 € per gram, Sigma Aldrich) we introduced a glucose-6-phosphate dehydrogenase (G6PDH, see Fig. 4.7) from *S. cerevisiae* (section 3.2.4). We tested, that G6PDH could be used as regeneration system for the formation of cinnamaldehydes (**18**-**22**) as well as cinnamyl alcohols (**23**-**27**).
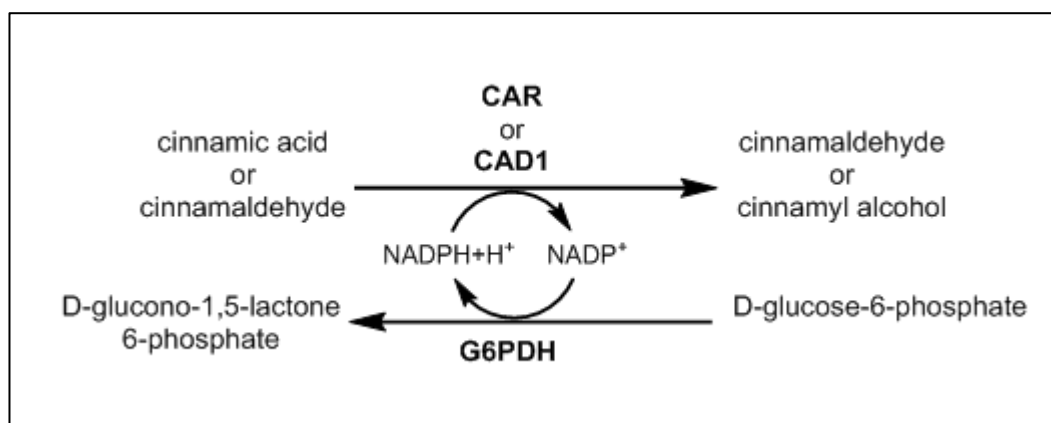
Figure 4.7: Scheme of recycling enzyme cascades used in *in vitro* transformations. For recycling during the reduction reactions, glucose-6-phosphate dehydrogenase (G6PDH) from *S. cerevisiae* was applied for recycling of NADPH.
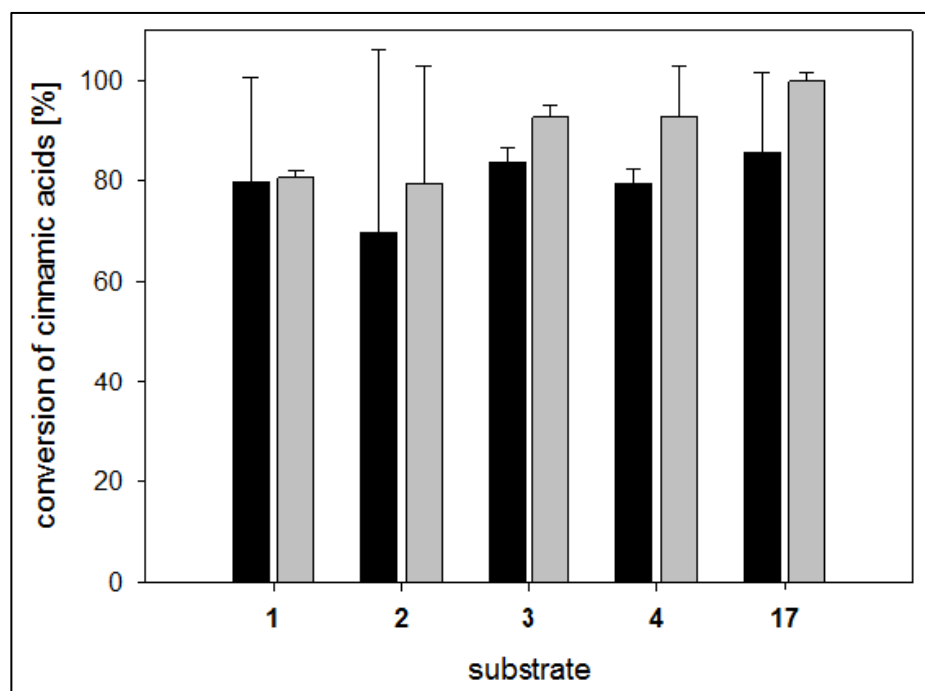
Figure 4.8: Conversion of cinnamic acid derivatives with two different enzyme systems. *p*-Coumaric acid (**1**), caffeic acid (**2**), ferulic acid (**3**), 5-hydroxyferulic acid (**4**) and sinapic acid (**17**) were incubated with NiCAR/NiPPTase (depicted in black) and NiCAR/NiPPTase/TaCAD1 (depicted in black) yielded in good conversions to the corresponding cinnamaldehydes (**18**-**22**) or cinnamyl alcohols (**23**-**27**).

The phosphopantetheinylated amino acid residue S689 could be experimentally verified by peptide analysis in cooperation with Petra Majovsky and Dr. Wolfgang Hoehenwarter (IPB, Halle (Saale)). After in-solution digestion with a combination of two endoproteinases (see section 2.6.3), the purified and desalted peptides were measured by LC-MS (see Appendix). By means of this method, phosphopantetheinylation of NiCAR could be detected. Non-phosphopantetheinylated NiCAR was only present in a marginal amount (data not shown). Furthermore, it could be detected that the phosphopantetheinyl group is bound in the analyzed sequence at serine 6 (i.e. S689) with a probability of at least 96.62 % (Table 4.2).

Table 4.2: Probability of site analysis in analyzed sequence.

Score	Mr (calc)	Delta	Sequence of peptide	Site Analysis
44.8	2213.9916	-0.0006	DLGGDSLSALSFSNLLHE	Phosphopantetheine at S6 96.62 %
30.1	2213.9916	-0.0006	DLGGDSLSALSFSNLLHE	Phosphopantetheine at S8 3.26 %
15.7	2213.9916	-0.0006	DLGGDSLSALSFSNLLHE	Phosphopantetheine at S11 0.12 %

Since the quantification of the aldehydes proved to be difficult (due to their high reactivity and low stability as well as the lack of available synthetic standards) we coupled the NiCAR-dependent aldehyde formation to the subsequent reduction into the corresponding alcohols (Fig. 4.2B). After addition of cinnamyl-alcohol dehydrogenase from *Triticum aestivum* (TaCAD1), which is able to accept cinnamaldehydes with differing substitution patterns,[97,98] the respective alcohols were produced very efficiently (Fig. 4.8). An accumulation of the aldehyde intermediates was not detectable under these conditions. Coupling of TaCAD1 with the NiCAR/NiPPTase system seems to result in higher conversion of all cinnamic acid derivatives.

As expected, the energetically challenging reduction of the carboxylic acids by NiCAR constitutes the rate limiting step for this two-step conversion. Quantification was possible for a few cinnamaldehydes and cinnamyl alcohols of which commercial product standards were available (Tab. 4.3).

Table 4.3: Total yield of the expected aldehydes and alcohols. Conversions with NiCAR/NiPPTase or NiCAR/NiPPTase/TaCAD1 resulted in good yields for all products. Quantification was possible only where commercial product standards were available.

substrate	c(aldehyde) [µM]	c(alcohol) [µM]
1	unquantifiable	838 ± 31
2	unquantifiable	878 ± 131
3	541 ± 29	771 ± 30
4	unquantifiable	unquantifiable
17	546 ± 24	805 ± 8

Interestingly, the conversion of *p*-coumaric acid (**1**), caffeic acid (**2**), ferulic acid (**3**) and sinapic acid (**17**) with NiCAR/NiPPTase/TaCAD1 resulted in the expected cinnamyl alcohols only (see in the Appendix), whereas the conversion of 5-hydroxyferulic acid (**4**) showed a mixture of the substrate **4**, aldehyde **21**, alcohol **26** and a further product in the HPL chromatogram (Fig. 4.9).
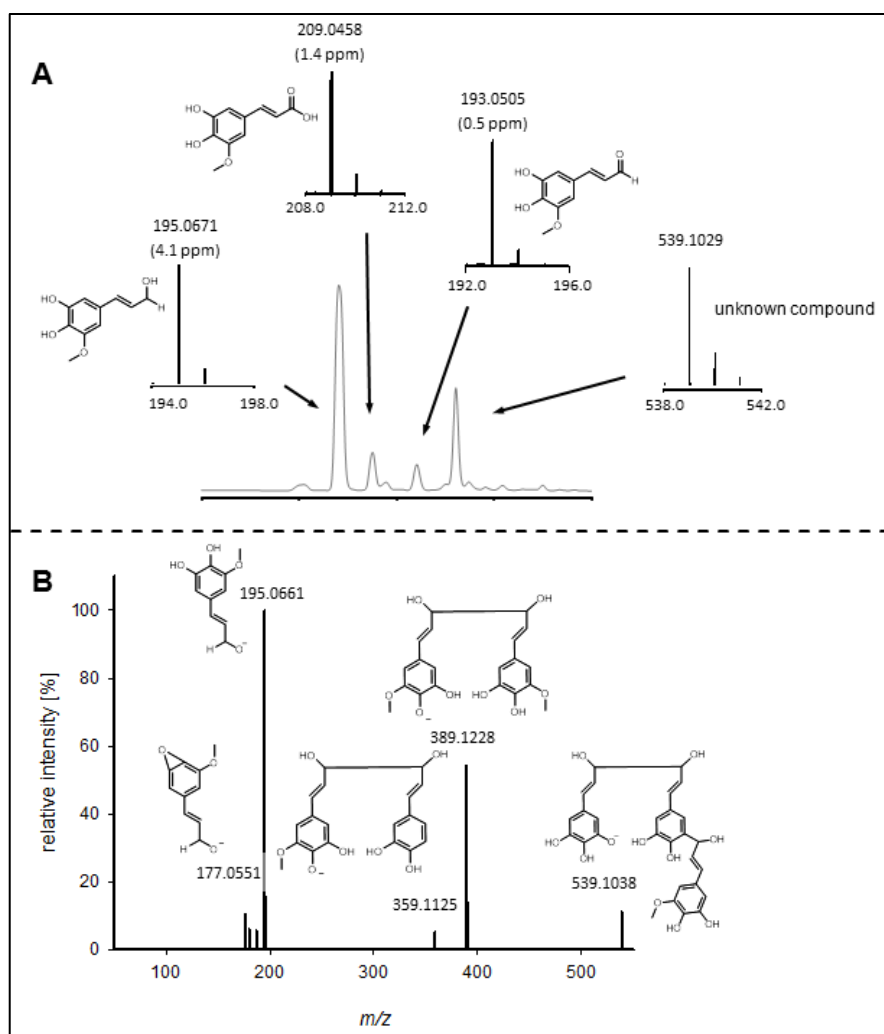
Figure 4.9: Extracted HPL chromatogram and MS analysis of the products of the NiCAR/NiPPTase/TaCAD1 assay using 5-hydroxyferulic acid (4) as substrate. A) For all peaks (substrate, product and side products) full MS data are shown. B) Fragmentation pattern of the unknown compound and the proposed structures for each fragment are presented.

Mass spectrometric analysis of the unknown peak leads to the assumption that the side product might be an oligomer (trimer, Fig. 4.9) of the actually expected product 5-hydroxyconiferyl alcohol (**26**).[d]

4.2.7 Spectroscopic analysis of sinapic acid derivatives

After success in converting the substrate we analyzed the conversion of sinapic acid (**17**) - the most decorated cinnamic acid derivative - to the respective aldehyde (**22**) and alcohol (**27**). The compounds were produced in preparative scale for NMR studies (section 2.5.4.2, figures and complete data are listed in the Appendix).

[d] Structure analysis was performed by Dr. Andrej Frolov, Dept. NWC, IPB Halle.

Sinapaldehyde (**22**) and sinapyl alcohol (**27**) could be successfully identified by ^{1}H NMR. For both compounds the configuration of the two protons attached to the cinnamyl double bond is distinctive. While for an *E*-configuration a coupling constant between 15 and 16 Hz is expected, for a *Z*-configuration the coupling constant is lower (approx. 12 Hz). Both compounds could be obtained in the desired *E*-configuration (**22** and **27**: ^{3}J(H, H) = 15.8 Hz). In addition, both compounds showed an intense singlet multiplicity for the protons of both methoxy groups at the aromatic ring (**22**: δ = 3.95 ppm; **27**: δ = 3.91 ppm). For a differentiation of both compounds, the chemical shift and the multiplicity of the respective functional group has to be observed. For compound **22** the proton of the cinnamaldehyde group revealed a doublet (^{3}J(H, H) = 7.7 Hz) downfield at δ = 9.66 ppm. Instead the protons of the of the cinnamyl alcohol group of compound **27** revealed a doublet (^{3}J(H, H) = 5.9 Hz) upfield at δ = 4.31 ppm. Both sinapic acid derivatives, sinapaldehyde (**22**) and sinapyl alcohol (**27**), exhibit a variation in the fragmentation pattern (Fig. 4.10).
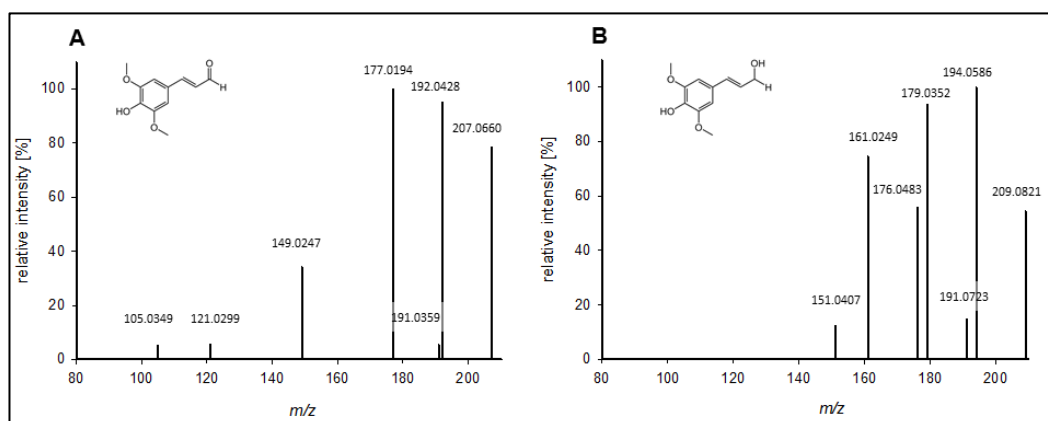
Figure 4.10: MS/MS fragmentation pattern of the products of the A) NiCAR/NiPPTase assay and the B) NiCAR/NiPPTase/TaCAD1 assay using sinapic acid (17) as a substrate. A) sinapaldehyde (**22**) and B) sinapyl alcohol (**27**).

4.3　Conclusion and discussion

The chimeric enzymatic pathway for the cell-free production of specific monolignols from *p*-coumaric acid (**1**) presented in this work is divided into the aromatic decoration at the carboxylic acid level and the subsequent reduction of the decorated acids into the corresponding alcohols. The crucial step for accessing multiply methoxylated monolignol derivatives is the utilization of an engineered 4HPA3H variant (Y301I) that catalyzes 5'-hydroxylation of ferulic acid (**3**) and of a regiospecific OMT with relaxed substrate but strict 3'-regiospecificity.

Concerning the reduction of the carboxylic acid side chain, it was shown that the disadvantageous plant-derived two-enzyme system can be substituted by the bacterial one-component carboxylic acid reductase from *Nocardia iowensis*.

Metabolic engineering of microorganisms for the synthesis of plant natural products is a fast growing field[4,99] and the reaction cascade presented here might be transferred into microbial host organisms for the fermentative production of monolignols known from plant metabolism. Previous examples for such a strategy include the biosynthesis of 12.1 mg l^{-1} caffeic acid (**2**) from endogenous tyrosine in *E. coli*, which was achieved by overexpression of a dual-enzyme cascade comprising tyrosine ammonia lyase (TAL) from *Rhodobacter capsulatus* and 4HPA3H.[18] Later the authors could increase the titer of caffeic acid (**2**) to 766 mg l^{-1} by engineering a tyrosine over-producing strain.[100] However, even this advanced system did not produce substituted cinnamates such as sinapic acid (**17**). Another group was able to produce 22 mg l^{-1} *p*-coumaryl alcohol (**23**) in *E. coli* from tyrosine by overexpression of an artificial phenylpropanoid pathway comprising TAL, 4CL, CCR and CAD.[101] The yield of *p*-coumaryl alcohol was increased to 50 mg l^{-1} by combinatorial operon optimization for the expression of the involved enzymes.[102] However, the applied 4CL/CCR system for reduction of the carboxylic acid into the alcohol failed in the conversion of highly oxidized species such as sinapic acid because of the absence of a specific oxidase. In contrast, the NiCAR/NiPPTase-based system presented in this study can be combined with the 4HPA3H-Y301I / PFOMT reaction for specific aromatic methoxylation, allowing for the formation of all monolignols commonly found in plants from their acid precursors. Apart from the alcohols, the corresponding aldehydes are also interesting products as their higher intrinsic reactivities might be favorable for subsequent polymerization to biogenic macromolecules or in condensation reactions. However, this high reactivity constitutes a bottleneck for their microbial production, as host enzymes like alcohol dehydrogenases or aldo-keto reductases rapidly convert aldehydes either to the corresponding alcohols[103] or oxidation products (see chapter I). Most interestingly, Kunjapur et al. recently engineered an *E. coli* strain for the accumulation of aromatic aldehydes (e.g. vanillin which could be highly overproduced compared to control strains[104]) and ketones (e.g. phenylacetylcarbinol[105,106]). Due to its high productivity and the ability to tolerate toxic levels of carbonyl compounds, a combination of such a strain with our enzymatic platform for specific decoration and reduction of cinnamic acids might provide a suitable background for both, the production of

monomeric monolignol compounds (including the cinnamaldehydes) as well as of oxidative coupling products (lignans, lignin) by fermentation.

5   Chapter III: A cell-free multienzyme cascade for synthesis of phenylpropanoids

5.1 Introduction

Phenylpropanoids, a structurally diverse class of phenolic natural products, are involved in a number of physiological processes in plants, such as lignification or development of floral and fruit pigmentation (see section 1).[16,3] These ubiquitous metabolites are part of the human diet, and increasing evidence indicates that their consumption is beneficial to health (e.g. due to antioxidant and anti-inflammatory properties).[3] Special phenylpropanoids, e.g. vanillin[107] or eriodictyol[108], are also important aroma or taste-modifing compounds, which contribute to the scent and flavor of foods (Fig. 5.1[e], left).

Vanillin (4-hydroxy-3-methoxybenzaldehyde, Fig. 5.1) is one of the most used flavoring compounds in the world. It is a white crystalline powder with an intensely sweet taste and a typical vanilla-like odor.[34] In addition, vanillin has antimicrobial and antioxidant properties and is an important precursor for the production of drugs, e.g. dopamine or papaverine. There is increasing demand for phenolic natural products in the food and pharmaceutical industry.
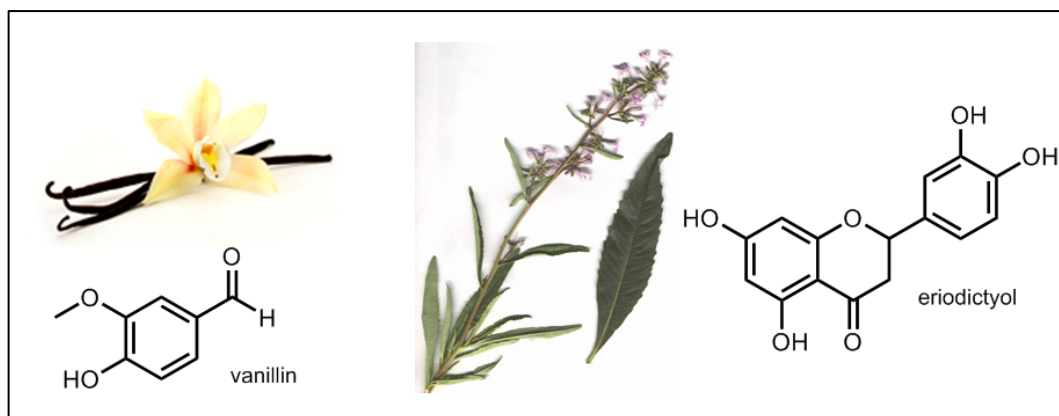
Figure 5.1: The phenylpropanoids vanillin (28) from *Vanilla* pods (left) and eriodictyol (8) from *Eriodictyon californicum* (right).

---

[e] source *Vanilla* pods: Symrise AG
source *Eriodictyon californicum:* http://www.fs.fed.us/wildflowers/plant-of-the-week/eriodictyon_sp.shtml

The bitter taste of some foodstuff is tolerated in only a few cases, e.g. coffee or tea. In most cases the compounds, which are responsible for the bitter, astringent taste, are masked among others by e.g. sodium chloride or sugar. There are many substances described as bitter, e.g. phenolic acid derivatives, flavonoids and terpenes[108], but many of them provide health benefits.[109] As a consequence of masking bitter compounds there is an increasing intake of salt and sugar which can leads to higher blood pressure or diabetes.[108,109] A lot of different methods were developed to avoid these limitations and to find healthier alternatives, e.g. use of sweeteners, addition of proteins (milk in coffee/tea) or nucleotides such as AMP, which were supposed as acting directly on the bitter receptors expressed on taste cells.[108] In the late 19th century, extracts of *Eriodictyon californicum* (Fig. 5.1, right, common name *Herba Santa*) were proposed as bitter masking agents. Ley et al. indicated in 2005 that eriodictyol (**8**, Fig. 5.1) and the homoeriodictyol sodium salt isolated from *Herba Santa* showed the most remarkable masking effects against the bitter compound caffeine.[108]

But the availability from natural sources of these compounds is often limited. Apart from chemosynthesis, strategies for a sustainable production of these ingredients in recombinant microorganisms (which harbor the corresponding biosynthetic pathways) have been developed in recent years.[4,36,110]

In particular, at the beginning our research was focused on biosynthesis of vanillin, the main aroma constituent of *Vanilla* pods.[107,111] The demand of the worldwide market cannot be met by extracted natural vanillin. Pathway engineering of genetically modified yeasts and bacteria afforded the generation of vanillin from phenylpropanoid precursors (e.g. ferulic acid[111]) or even from metabolic intermediates, resulting in *de novo* synthesis from glucose feedstock.[7] Based on recombinant yeast, a fermentative production system for this "natural" aroma is already in the process of commercialization.[f]

With a market price of 13,860 € per gram (Sigma-Aldrich), eriodictyol (**8**) is a quite expensive compound. In contrast, one gram of naringenin (**7**) is available for 23.60 € (Sigma-Aldrich) and can be used as cheap precursor. This offers the opportunity for enzymatic hydroxylation of naringenin (**7**) to obtain the expensive bitter-masking eriodictyol (**8**).

In contrast to the numerous examples for *in vivo* synthesis of phenylpropanoids, reconstituted pathways comprising cascades of isolated enzymes have not been applied for this purpose yet. There are several reasons preventing the use of

---

[f] 'A sustainable production route', can be found under http:/www.evolva.com/products/vanillin, 2014.

enzymes – especially those from the plant producers – in cell-free systems, such as low activity / stability under *in vitro* conditions, and generation of side products from intermediates because of lacking compartmentalization.[7] In addition, the pathways to most phenylpropanoids include modification steps (e.g. hydroxylation and methylation), which are mandatory for biological activities[110] but require costly cofactors such as hydride donors and the methyl group donor *S*-adenosyl-L-methionine (SAM).

5.2 Results and discussion

5.2.1 Cloning and transformation of genes

Cloning of 4HPA3H, PrnF, PFOMT, At4CL2, FDH, G6PDH as well as the site-directed mutagenesis for creating the 4HPA3H variant (Y301F/S462A) were already mentioned in the previous two chapters.

The LeOPR1 gene from *Lycopersicon esculentum* was cloned into a pET28a(+)-vector sustaining a kanamycin resistance and an N-terminal His$_6$-tag for purification via affinity chromatography (see section 2.4).

Gene cloning and transformation of SAHN, SAHH, FCoAHL and the site-directed mutagenesis for creating the SAMS variant (I317V) were already implemented by Dr. Martin Dippe.

The constructed and the even existing plasmids were transformed into *E. coli* DH5α cells for storage and into *E. coli* BL21(DE3) cells for expression.

5.2.2 Expression and purification

After successful gene expression and cell lysis (section 2.4), the proteins were isolated using Immobilized Metal Ion Affinity Chromatography (IMAC). The success of expression, purification and dialysis of the used proteins was visualized by SDS-PAGE (see Appendix). The yields differ from good (e.g. 56 mg l^{-1} of LeOPR1 protein) to very good (e.g. 295 mg l^{-1} of 4HPA3H-Y301F/S462A protein) expression yields. All other expression yields are listed in section 2.4.1.

Overexpression, purification and characterization of additional enzymes used in module M and module L was done by Dr. Martin Dippe, IPB Halle (Saale). The expression conditions, expression yields and specific activities are listed in the Appendix.

5.2.3 Cell-free enzymatic synthesis of vanillin

For the development of the model *in vitro* system, we focused on the modification of the phenylpropanoid parent compound *p*-coumaric acid (**1**) and the synthesis of vanillin. Besides its commercial relevance, this flavor compound is toxic to microbial production systems even at millimolar concentrations, a restriction that is insignificant in cell-free production. In the initial step of vanillin biosynthesis, substance **1** is converted to caffeic acid (**2**) by hydroxylation, which is subsequently transformed into ferulic acid (**3**) by selective methylation[112] (structures are depicted in Figure 5.2A). In *Vanilla* plants, this intermediate is hydrolyzed to vanillin by a bi-functional lyase enzyme.[107] For *in vitro* reconstitution of this pathway, we first selected suitable enzymes based on three criteria: (I) sufficient yield in recombinant expression, (II) high activity and (III) suitable substrate scope and regio- / product specificity. In the following sections, detailed information on the individual performance of the recombinant enzymes used for the reaction steps is summarized:

Hydroxylation of *p*-coumaric acid: As only a partial sequence of the endogenous monooxygenase from *Vanilla planifolia*[113] is available, we substituted this enzyme by 4-hydroxyphenylacetate hydroxylase (4HPA3H) from *E. coli*. This flavin-dependent enzyme has been used for conversion of *p*-coumaric acid (**1**) to caffeic acid (**2**) by whole-cell biocatalysts[71] (see chapter I). It is part of a two-enzyme system for hydroxylation of monophenols, including a reductase component, which is necessary for cofactor supply. For this purpose, the reductase (PrnF) homologue from *Pseudomonas protegens* was applied, which was capable of reducing flavin adenine dinucleotide (FAD) to $FADH_2$ with high specific activity. The PrnF reaction itself is consuming NADH.[77] In order to use this costly hydride donor in non-stoichiometric amounts, we employed the well-known formate dehydrogenase from yeast[114] to recycle formed NAD^+ with inexpensive formate. The resulting multi-enzyme cascade was tested for conversion of three substrates from the vanillin pathway in an *in vitro* assay. It was functional in hydroxylation of *p*-coumaric acid (**1**, with a specific activity of 6.2 ± 0.5 nmol min^{-1} mg^{-1}), but showed considerable oxidative degradation of vanillin into unidentified side products (data not shown)[g], suggesting to avoid these oxidative conditions in the final steps to this aldehyde.

---

[g] The conversion of vanillin and *p*-hydroxybenzaldehyde with 4HPA3H was tested in first experiments of chapter I. Both substrates showed a confusing spectra in HPLC, so that they might be metabolized.
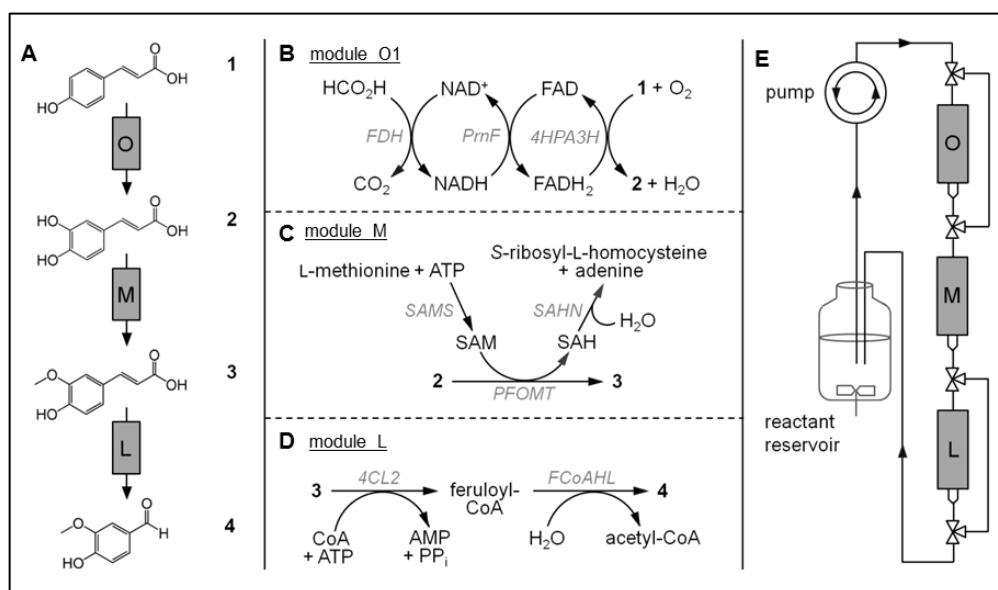
Figure 5.2: Modular multi-enzyme system for the transformation of phenylpropanoids. Left: A) Scheme of the reconstructed pathway with modules, phenylpropanoid intermediates and products (**1** *p*-coumaric acid, **2** caffeic acid, **3** ferulic acid, **28** vanillin). Middle: A detailed view of the enzymatic steps which are combined in the individual cartridges is depicted, including B) hydride transfer in aromatic hydroxylation (module O1), C) cofactor (SAM) generation and by-product (SAH) removal in O-methylation (module M), and D) the formation of vanillin (**28**) by hydration/retro-aldol type lysis of a CoA ester (module L). Individual enzymes are shown in italics in grey, AMP: adenosine-5´-monophosphate, PPi: inorganic diphosphate. Right: E) experimental setup of sequentially arranged cartridges (modules). Bypasses allow excluding individual modules to either generate different products or reaction orders, or to avoid interference of early intermediates with late biocatalytic steps or vice versa (see text). Figure was created by Dr. Martin Dippe.

Methylation of caffeic acid (**2**): For methyl transfer, the phenylpropanoid-flavonoid *O*-methyltransferase (PFOMT) from the ice plant *Mesembryanthemum crystallinum*, which was produced in high yield in heterologous overexpression, was used. We confirmed its high specificity for the meta-position in catechols which was reported previously.[91] In methylation of caffeic acid (**2**) it produced ferulic acid (**3**) as the sole product (see section 4.2.2).

In order to enable cofactor (SAM) generation and to increase methylation yield, we introduced two additional enzymes: First, we applied a variant of the SAM synthase (SAMS) from *Bacillus subtilis* for production of SAM from the comparably less expensive substrates L-methionine and ATP. This variant, which differs from the wild-type by exchange of a single amino acid residue (I317V), is superior to the parental enzyme in *in vitro* synthesis of SAM due to higher activity and decreased product inhibition.[49] Furthermore, we tested the ability of two hydrolases to remove *S*-adenosyl-L-homocysteine (SAH), the inhibitory by-

product of most methyltransferase reactions,[115,116] from the reaction mixture. Both SAH hydrolase (SAHH) from *Corynebacterium efficiens*[117] and SAH nucleosidase (SAHN) from *E. coli*[118] were functional in cleavage of SAH (yielding L-homocysteine or *S*-ribosyl-L-homocysteine, respectively). However, SAHH proved to be inapplicable because of very low activity.

Lyase reaction: For the conversion of ferulic acid (**3**) into vanillin (**28**), we used a combination of 4-coumarate-coenzyme A (CoA) ligase 2 (4CL2) from *Arabidopsis thaliana*[119] and feruloyl-CoA hydratase/lyase (FCoAHL) from *Pseudomonas fluorescens*.[120] A similar two-component enzyme system has been proposed to catalyze this reaction in the native biosynthetic pathway of *Vanilla* orchids, before the bi-functional CoA-independent lyase mentioned above was discovered recently.[107] In our reconstituted reaction sequence, 4CL2 is applied to generate feruloyl-CoA from ferulic acid, CoA and one equivalent of ATP. This activation is followed by hydrolysis of the CoA ester to vanillin by FCoAHL in a retro-aldol type reaction as the final reaction step.

A combination of the steps, even if the individual performance of single enzymes is well established, or cascade reaction modules for cofactor (re)generation are known, is usually not readily functional without causing severe interdependency problems. True orthogonality of enzymatic reactions is often claimed but in reality commonly not given. Substrates, intermediates and product(s) can be reactive with more than one enzyme employed in cascades or in an organism.

 A multitude of adverse complexity-related problems can arise from, e.g., different pH and temperature optima of the enzymes, salt or O_2 sensitivity, insufficient flux, or reactivity / cross-inhibition of intermediates or products with any other component. Most organisms solve these problems either by compartmentalization or assembly lines (clusters or metabolons) that reduce external interference. The problem increases if the individually ideal enzymes for an artificially assembled pathway come from different organisms.[7]

For example, from the enzymes used in our route towards vanillin, 4CL2 shows a relaxed substrate specificity and converts a broad range of cinnamic acid derivatives.[119] Indeed, it was less active on ferulic acid (**3**, 43 ± 1 nmol min^{-1} mg^{-1}) than on its precursors *p*-coumaric acid (**1**) and caffeic acid (**2**, 783 ± 19 and 554 ± 5 nmol min^{-1} mg^{-1}). Similarly, FCoAHL was able to hydrolyze the CoA derivatives of all three acids with different efficiency (specific activities on CoA esters of **1**, **2** and **3**: 7700 ± 220, 4930 ± 630 and 4500 ± 200 nmol min^{-1} mg^{-1}), leading to the

formation of the corresponding benzaldehydes. In terms of pathway engineering, the promiscuity of both 4CL2 and FCoAHL is a drawback: In a non-compartmented system containing all eight enzymes for synthesis of vanillin (**28**), the starting material **1** is prone to conversion into *p*-hydroxybenzaldehyde.

Due to this potential interference of enzyme reactions, we decided to combine the individual enzymes involved in one of the three basic steps of the reaction sequence to separate subunits (modules). Enzyme portions leading to balanced activities were mixed and applied to commercial columns for affinity immobilization, which contain matrix-bound Co^{2+} ions. Thus, the enzymes were randomly immobilized via the N-terminal hexahistidine tag. These matrix-supported biocatalyst assemblies ("modules") separately harbor the sub-cascades for certain standard reaction steps in phenylpropanoid biosynthesis, including tributaries, for the oxidation (module O1 an O2), methylation (module M) or lyase reaction (module L). Thus they combine the advantages of a modular setup with those of immobilized enzymes. The sequential assembly of modules enables a tight control of the conversion, accompanied by repression of side reactions. In addition, the "scaffolding" provided by immobilization confers optimal stoichiometries in multi-enzyme systems.[121] It also proved to stimulate substrate channeling which leads to enhanced reaction rates[33] and facilitates product purification and analysis (as solutions do not contain any biocatalysts).[80]

In a first experiment, module O1 was used for the generation of caffeic acid. The substrate **1** (200 µM) and all cosubstrates (FAD, NADH, sodium formate) were dissolved in a buffer containing $MgCl_2$ and KCl, which was tested for compatibility with all used enzymes (Mg^{2+} and K^+ ions are essential for activity of PFOMT, SAMS and 4CL2). The reactant solution was applied to the module, which contained 20 mU of the oxidase 4HPA3H and an excess of the cofactor regeneration enzymes PrnF and FDH, by a circulating pump. As reflected by the time course of the conversion, repetitive application is necessary for high conversion: Transformation of 76 % of the substrate was achieved after approximately 120 passages (Fig. 5.3A). In principal, an increased productivity (i.e. identical conversion with reduced passage numbers) might also be achieved by other technical improvements (such as O_2 enrichment, slower pumping, altered design or parallelization of modules) or higher catalyst loading.

Next, we coupled the hydroxylation to the methylation reaction. Therefore, module M (containing PFOMT, SAMS and SAHN) was integrated, and L-methionine and ATP – the substrates of SAMS – were added to the reaction to

enable internal supply of costly SAM. Thus, caffeic acid (**2**) from module O was completely converted into ferulic acid (**3**, Fig. 5.3B). The high efficiency of the methylation cascade was assured by excess application of PFOMT and SAMS (70 mU each), and by removal of the by-product SAH by SAHN, which results in generation of *S*-ribosyl-L-homocysteine and adenine (detected by HPLC [data not shown] and UHPLC/ESI-MS [see Appendix]).
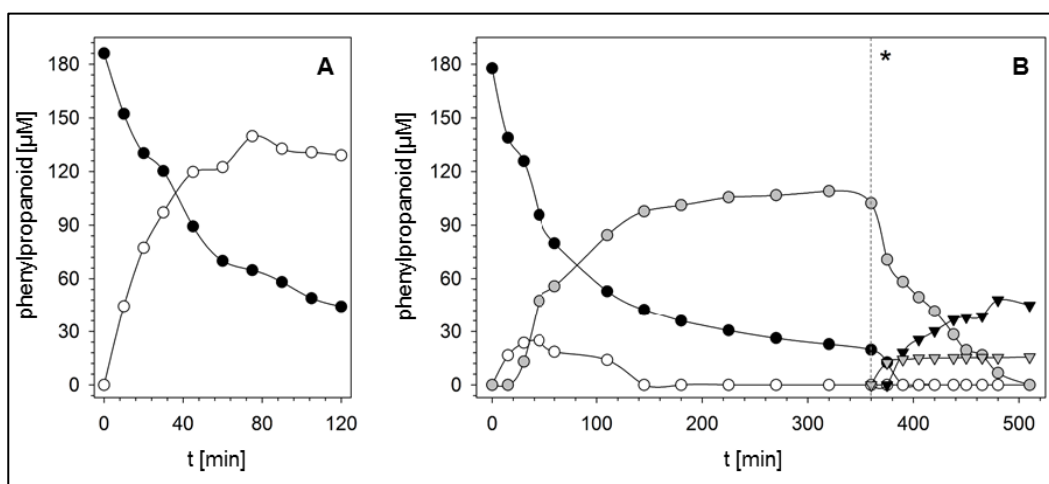
Figure 5.3: Time course of phenylpropanoid transformations. A) *Ortho*-hydroxylation of *p*-coumaric acid by module O1. B) Conversion of *p*-coumaric acid (**1**) to vanillin (**28**) by modules O1, M and L. Subsequent to the oxidation and methylation steps (catalyzed by modules O1 and M, respectively), formation of vanillin was initiated by activating module L and bypassing module O (see asterisk* and Fig. 5.2). The transformations were performed as described in the section 2.8.2. Samples were taken from the reactant reservoir and analyzed by HPLC for (●) *p*-coumaric acid (**1**), (○) caffeic acid (**2**), (●, filled grey) ferulic acid (**3**), (▼) vanillin (**28**) and (▼) *p*-hydroxybenzaldehyde (by-product from **1**). Figure was created by Dr. Martin Dippe.

In order to convert ferulic acid (**3**) to vanillin (**28**), module L (harboring 4CL2 and FCoAHL) was applied. Ideally, for this step module O is removed or shortcut by a valve system to avoid oxidative degradation of produced vanillin (see above). The reaction was started by addition of CoA, and after further incubation the production of vanillin by the 4CL2/FCoAHL system was detected.

Interestingly, residual *p*-coumaric acid (**1**) was transformed into *p*-hydroxybenzaldehyde after integration of module L (Figure 5.3B). This side reaction catalyzed by 4CL2 and FCoAHL proceeded with higher reaction rate than the competing formation of vanillin (**28**). This fact clearly supports a modular (compartmented) set-up of the reaction cascade to vanillin, which is crucial for a directed synthesis of the compound from **1** as well as for high product yield.

5.2.4 Cell-free enzymatic hydroxylation of naringenin (7)

In cooperation with Symrise AG (Holzminden, Germany) we tested the modular system for a further conversion: the hydroxylation of the bitter-tasting compound naringenin (**7**) into the bitter-masking compound eriodictyol (**8**) using the 4HPA3H variant Y301/S462A which showed the highest conversions for the chosen substrate **7** in previous studies (see chapter I).

The system is the same as described for module O1 in section 5.2.1. The reactant solution was applied to the module, which contained 200 µM substrate **7**, 35 mU of the oxidase variant 4HPA3H-Y301F/S462A and an excess of the cofactor regeneration enzymes PrnF and FDH, by a circulating pump. After reaction time of 240 min 67.7 µM eriodictyol (**8**, Fig. 5.4) and 33.3 µM of non-transformed naringenin (**7**) could be determined. In other words, in less than 4 hours reaction time 19.5 mg l^{-1} of the product **8** (more than 33 % conversion of **7**) were obtained. However, these are absolute values since the effective values are higher.
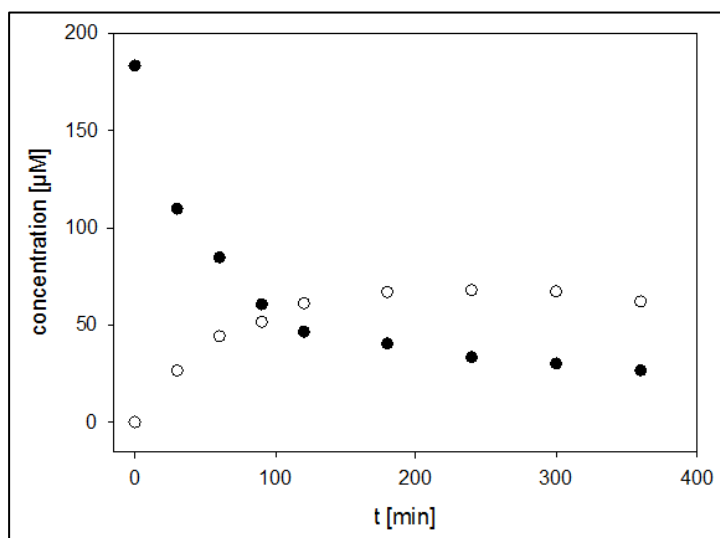
Figure 5.4: Time course of eriodictyol formation. Conversion of naringenin (**7**, ●) to eriodictyol (**8**, ○) by module O2.

The decrease of eriodictyol (**8**) decreases after 240 min does not mean that no further product **8** is or was formed. Optically recognizable with rising reaction time is an intensive purple color of the column, which contains matrix-bound Co^{2+} ions. Due to its catechol-like structure, eriodicytol (**8**) likely possesses as all catechols a high affinity to bind to the column. As a result of this the column was washed with elution buffer (see section 2.4.2) after finishing reaction time and the eluate was measured by HPLC. In this solution 29.1 µM eriodictyol (**8**) and 13.6 µM non-converted substrate **7** were found. Two further washing steps (water and ethanol) did not yield further product. Nevertheless, the column still retains the purple color after many washing steps.

For our cooperation partner, a possible discrimination of (+)- or (-)-naringenin during the conversion was of interest. For identification we used chiral reversed-phase high-performance liquid chromatography. A kinetic assay was performed and samples were taken at defined time points. No note worthy discrimination of naringenin enantiomers during the conversion and no selective formation of eriodictyol enantiomers during the formation could be observed (Fig. 5.5).
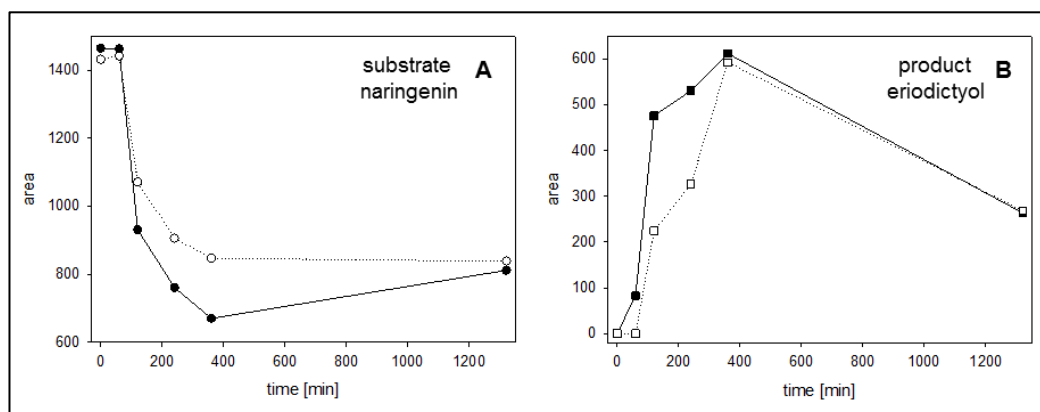
Figure 5.5: Area under the curve of the enantiomers naringenin (7, left) and eriodictyol (8, right) during the reaction. A) Area under the curve of (+)-naringenin (depicted with closed circles) to (-)-naringenin (depicted with open circles). B) Area under the curve of (+)-eriodictyol (depicted with closed squares) to (-)-eriodictyol (depicted with open squares).

Although the conversion of (+)-naringenin is faster at the beginning of the reaction, it leveled out to a balanced ratio over time (Fig. 5.6). This shows that the oxidase variant 4HPA3H-Y301F/S462A did not significantly prefer one enantiomeric substrate or form one favored enantiomeric product. However, a slight discrimination in relative rates is present and might be worked out if desired by directed evolution.
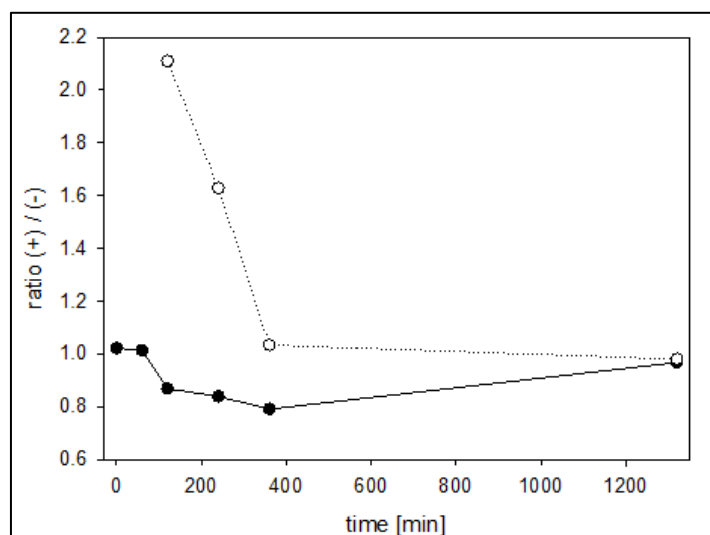
Figure 5.6: Ratio between the enantiomers of naringenin (7, ●) and eriodictyol (8, ○) during the reaction.

Nonetheless, a solid-phase-based flow method could be established for the hydroxylation of naringenin (**7**) with the 4HPA3H variant Y301F/S462A. During the reaction time, no formation of any byproducts could be detected. The enzyme catalysts were stable for several days (data not shown). Recommendable is the utilization of another (covalent) immobilization system, which is not based on metal-chelate columns, because of binding not only proteins but also products. A covalent, metal-free immobilization could be an option.

In cooperation with Symrise AG, the results of this study have been applied for patent (Application No./Patent No. 15188136.4 - 1501).[122]

5.3　Conclusion and Outlook

5.3.1　Conclusion

In conclusion, it could be shown that a cascade of eight immobilized biocatalysts – including enzymes for efficient cofactor supply and by-product removal – is suitable for the synthesis of the model phenylpropanoid vanillin. In principle, this kind of modular system is also applicable for the production of other types of phenylpropanoids, e.g. flavonoids as eriodictyol. It is advantageous if starting material, intermediates or products display high toxicity or lack orthogonality to basic metabolism in whole-cell biocatalytic systems. It can be easily extended for combinatorial syntheses of (non-natural) derivatives, e.g. by addition of further modules for modification reactions such as prenyl or glycosyl transfer.

5.3.2 Reductive Production of ketone flavors

As mentioned before the system is extendable for the production of many other flavors and fragrances. For construction of another module we chose two ketones as potential products, which are famous in the flavor industry: rheosmin (**15**, commonly called raspberry ketone) occurs in a variety of fruits (e.g. cranberries and blackberries) and is the primary flavor compound in raspberries[123]. Zingerone (**31**, also called vanillylacetone) was identified as a pungent principle of ginger[124] (Fig. 5.7A[h]).
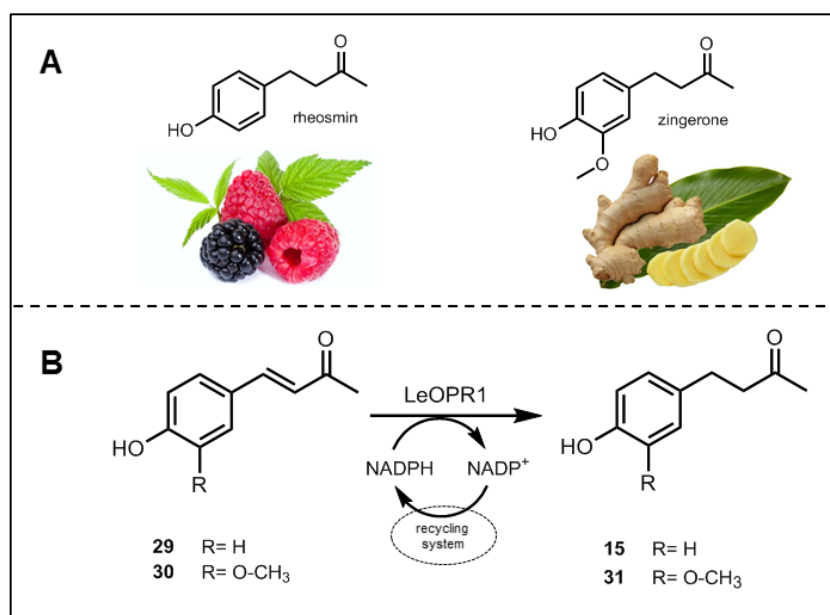
Figure 5.7: Converted ketones produced by the reductive modular system. A) Rheosmin (**15**), characteristic aroma compound of raspberry or blackberry (left), zingerone (**31**), a pungent principle of ginger (right). B) Scheme of reaction catalyzed by 12-oxophytodienoate reductase from *Lycopersicon esculentum* (LeOPR1).

One step during the biosynthetic formation of the ketones **15** and **31** is the reduction of the C=C bonds within the cinnamic acid scaffold. Chemically the catalytic hydrogenation of the α,β-unsaturated ketones (**29** and **30**) can be performed by using a combination of palladium and Raney nickel. However, it can lead to significant over-hydrogenation resulting in alcohols as byproducts, which complicates cleanup and isolation of the pure ketone products. As an option 0.5 % rhodium on alumina can be used as catalyst which is more expensive than palladium but it allows no overhydrogenation such as ketone reduction.[125]

But for flavor and fragrance industry there has been legislative discrimination between chemically identical food additives from synthetic and natural sources. Therefore, it has become advantageous to produce food additives using isolated enzymes or whole-cell biocatalysts.[126] One well-known enzyme family for the reduction of many commercially useful substrates is the old yellow enzyme (OYE) family. The OYE family is a large group of flavin-dependent redox biocatalysts which are often used in industrial applications in the reduction of activated alkenes.[126] OYE was one of the first enzymes isolated from brewer`s bottom yeast and was the first enzyme shown to possess a flavin cofactor.[127] OYE enzymes are outstanding catalysts because of their broad specificity, stability and availability in large quantities.[126] OYE-like enzymes have been identified in yeasts, bacteria, plants, as well as parasitic eukaryotes and are able to reduce a wide variety of substrates (e.g. conjugated enals, enones, α,β-unsaturated carboxylic acids, imides and nitroalkenes).[128]

The OYE2 from *S. cerevisiae* is able to convert substituted and nonsubstituted α,β-unsaturated ketones and was pre-existing in our working group.[i] Physiologically, the enzyme is part of a detoxification system for a broad spectrum of electrophilic compounds, e.g. acrolein.[129] We tested purified OYE2 enzyme for a possible conversion of the α,β-unsaturated ketones (**29** and **30**) to our desired products rheosmin (**15**) and zingerone (**31**). However, no conversion of the unsaturated ketones to the saturated ketones could be observed (data not shown).

In the search for another enzyme, 12-oxophytodienoate reductase (OPR, isoenzyme OPR1) from *Lycopersicum esculentum* (tomato) was chosen as an alternative enzyme because of its remarkably broad substrate spectrum.[128,130] It was termed OPR because it catalyzes the NADPH-dependent reduction of 12-oxophytodienoate in higher plants (e.g. *Arabidopsis thaliana*) to yield 3-oxo-2(2'(Z)-pentenyl)-cyclopentane-1-octanoic acid. The OPR-catalyzed reaction is part of the biosynthetic pathway to jasmonic acid, a compound with multiple hormonal activities in higher plants.[127,131] However, the function of the isoenzyme OPR1 in the pathway is unclear.[132] This flavin-dependent enzyme belongs to the 'old yellow enzyme family' as well and is able to catalyze the stereoselective reduction of activated alkenes (e.g. α-methylmaleic acid, ketoisophorone[130,133], α,β-unsaturated dicarboxylic acids and their corresponding diesters[128]). The internal flavin cofactor ($FMNH_2$) is reduced at the expense of a nicotinamide

---

[i] OYE2 from *S. cerevisiae* pJ69-4A was cloned, overexpressed and purified by Dr. Martin Dippe, Dept. NWC, IPB Halle

cofactor and transfers a hydride to the β-carbon atom of the used substrate while an essential tyrosine residue (conserved along the OYE family) adds a proton to the α-carbon from the opposite side. As a corollary of the stereochemistry of this mechanism, the addition of H_2 proceeds in a strict *trans*-specific fashion.[128,130,133,134]

After RNA isolation from tomato plant (*Lycopersicum esculentum*) the specific OPR1 gene was cloned, overexpressed and purified according to the experimental section. The activity assay was performed in cuvettes at a wavelength of 410 nm, because the unsaturated substrates (**29** and **30**) show a good absorbance at this wavelength while the saturated derivatives (**15** and **31**) did not shown any absorbance (in the Appendix).

First tests with the purified enzyme show, that in *in vitro* studies the reaction depends on proper recycling of the cofactor NADPH. If the reaction is performed without regeneration system the reaction stagnates after a certain time and no further conversion of the α,β-unsaturated substrates can be observed (see Fig. 5.8A, exemplary for the conversion of **29**). However, the addition of a regeneration system (glucose-6-phosphate dehydrogenase and glucose-6-phosphate, see section 4.2.4) leads to continuous consumption of the substrate and formation of the saturated product (see Fig. 5.8B). For further tests with OPR1 the use of the regeneration system was indispensable.
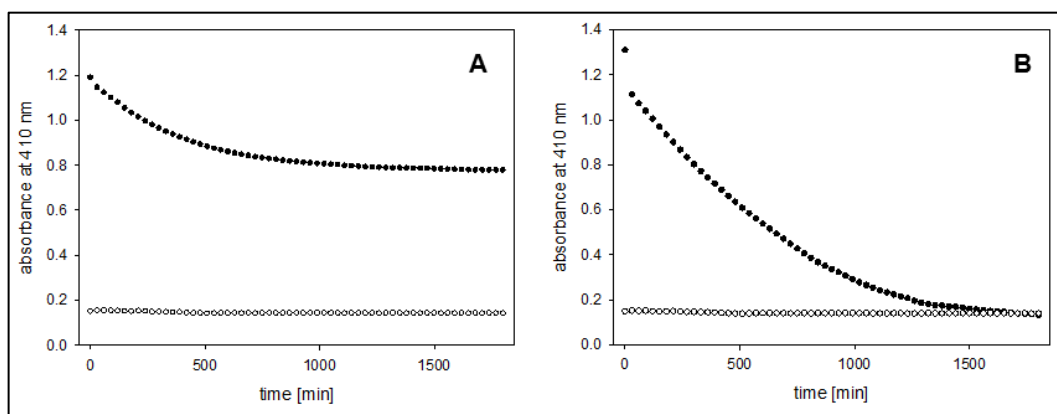
Figure 5.8: Time course of the LeOPR1 reaction with or without the usage of a cofactor regeneration system. A) Conversion of 100 μM substrate **29** (●) or **30** (○) with 0.8 mg ml^{-1} LeOPR1 and 200 μM NADPH. B) Conversion of 100 μM substrate **29** (●) or **30** (○) with 0.8 mg ml^{-1} LeOPR1 and 200 μM NADPH in the presence of 5 mM G6P and 0.5 μg ml^{-1} G6PDH.

The general setup of the used system is the same as described for module O1 in section 5.2.1. The reactant solution was applied to the module (called module R), which contained 200 µM substrate (**29** or **30**), 35 mU LeOPR1 and an excess of the cofactor regeneration enzyme G6PDH, by a circulating pump.
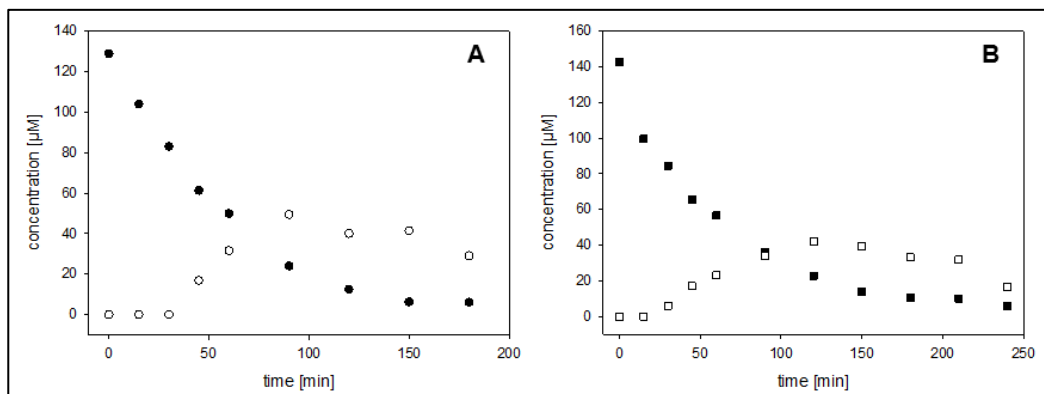
Figure 5.9: Time course of reductive formation of saturated ketones. A) Conversion of dehydrozingerone (**30**, ●) to zingerone (**31**, ○) by module R. B) Conversion of dehydrorheosmin (**29**, ■) to rheosmin (**15**, □) by module R.

After a reaction time of 90 minutes, 49.4 µM zingerone (**31**) and 23.99 µM non-transformed precursor (**30**) were found (Fig. 5.9A). In the late phase of the reaction (180 min) the amount of **30** further decreased, and a continuing formation of product was observed (**31**, Fig. 5.9A). The formation of rheosmin (**15**) was performed similar to this of zingerone. After reaction time of 120 min 42.1 µM rheosmin (**15**) and 22.8 µM precursor (**29**) could be determined (Fig. 25B). Towards 240 min only 5.8 µM of **29** could be determined, but also here a decrease of product (**15**) is present (Fig. 5.9B).

5.3.3 Outlook

This chapter illustrates that the chosen immobilization system can also be used as proof-of-concept for the conversion of different substance classes by the usage of different enzymes types but not for further conversions to produce higher product yields. Depending on application area, different immobilization techniques could be used. Potentially, enzymes can be immobilized by entrapment in a polymer network such as organic polymers or more stable inorganic silica matrices.[135,136] The advantages of the immobilization method are that it can be used universally for any enzyme and has mild procedure conditions. But there are disadvantages as well, e.g. there is often a large diffusional barrier, loss of enzyme activity because of a leakage or denaturation as a matter of free radicals.[137] To improve some of this disadvantages,

Wachtmeister et al. used encapsulated cells in a whole-cell catalytic "teabag" in combination with a micro-aqueous reaction system that allows high substrate loadings and resulting in high product concentrations.[138] The encapsulation of lyophilized biocatalyst into a polyvinylidene fluoride (PVDF) western blotting membrane resulted in catalytic "teabags", which can be easily handled, stored, removed and recycled. With this "teabags" enzymes can be reutilized and the system can even be applied for enzyme cascades. Furthermore, they described the applicability on a larger scale by using a SpinChem reactor. They demonstrated that, despite the presence of a diffusion limitation by the "teabag" membrane, thresholds of industrial feasibility can still be exceeded.[139]

In general there is no standard immobilization technique for each enzyme or enzyme class. For application methods the enzyme of choice should be tested with different immobilization techniques.

Technical improvements, like in line analytics of conversion that is controlling a valve system, better immobilization and carrier systems, and inclusion of further reactions for cofactor supply / substitution (e.g. enzymatic regeneration of ATP from polyphosphates[140] as starting material and substitution of CoA by less expensive thiol mimics[141]) will allow more advanced modular cell-free multistep biosynthesis in the future.

6   Final summary and outlook

6.1   Final summary

This work presents a variety of starting points for enzymatic synthesis and modifications of different central metabolites of the biosynthesis of ligin (e.g. cinnamaldehydes and cinnamyl alcohols) and flavors (e.g. vanillin and rheosmin) using soluble microbial and plant-derived enzymes with sufficient stability and their possible applications.

In the first chapter, a rational design of a flavin-dependent monooxygenase from *E coli* was presented. It is known that the *ortho*-hydroxylation (mostly P450 reactions) of hydroxycinnamic acids is a crucial step for the biosynthesis of e.g. coumarins in plants.[17] The monooxygenase 4HPA3H is able to hydroxylate cinnamic acid (**1**), umbelliferon (**13**) and rheosmin (**15**) to their related catechols completely. Further, it was shown that single substitutions, in this case mainly at position Y301, are sufficient to engineer new functionalities into 4HPA3H. The yielding enzymes are optimal for conversion of sterically challenging substrates (Fig. 6.1), e.g. conversion of the flavonoid naringenin (**7**) to eriodictyol (**8**). The variant Y301F/S462A provides an alternative for the conversion of naringenin (**7**) by existing membrane-bound cytochrome-dependent enzymes (e.g. the P450 enzymes BM3[93]). All generated variants showed applicability in different reaction systems due to high product yield and strict regio- and product specificity. In addition, an optimized *in vitro* assay was established for the conversion and further for the production of catechols with the 4HPA3H or its variants. As already mentioned, a steady focus shall be on the avoidance of expensive cofactors. Within this assay, a regeneration system for NADH (by the usage of FDH from *Candida boidinii*, later the gene was cloned from *Saccharomyces cerevisiae*) was integrated, which leads to the possibility of implementation in additional studies. As an alternative for the endogenous reductase partner of 4HPA3H, which was not accessible due to formation of many insoluble protein, the flavin:NAD(P)H reductase PrnF from *Pseudomonas protegens* was used for regeneration of FADH$_2$ within the *in vitro* formations.
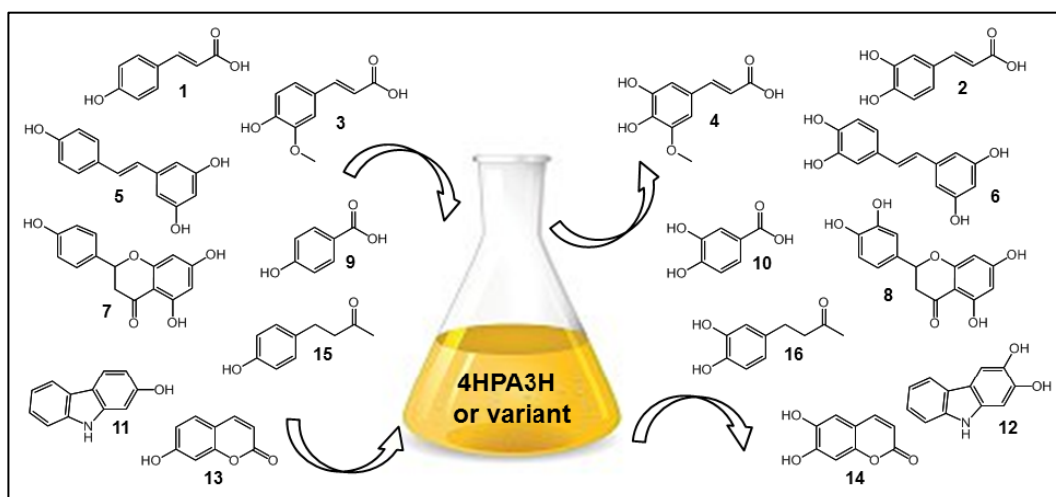
Figure 6.1: Substrates (odd numbers) used in hydroxylation reactions and the corresponding catechol products (even numbers) with the monooxygenase 4HPA3H or its variants.

The results of the first chapter were one essential step for further studies presented in the second chapter. Since biocatalytic *in vitro* strategies for the production of pure monolignols would thus constitute interesting alternatives to the extraction from lignin, the main focus was the enzymatic pathway for cell-free production of specific monolignols starting from *p*-coumaric acid (**1**).

One initial step for accessing the highly methoxylated derivatives is the hydroxylation of ferulic acid (**3**). Whereas the 4HPA3H wildtype enzyme was not able to convert ferulic acid (**3**) to 5-hydroxyferulic acid (**4**), the engineered variant 4HPA3H-Y301I presented in chapter I showed good results (upon 45 %).

The second reaction step – the regiospecific methylation of the oxidized scaffold – was carried out by an *O*-methyltrasnferase (PFOMT from *Mesembryanthemum crystallinum*) with relaxed substrate- but strict 3'-regiospecificity. PFOMT catalyzes the conversion of caffeic acid (**2**) to ferulic acid (**3**) as well as the conversion of 5-hydroxyferulic acid (**4**) to sinapic acid (**17**). Consequently, it was possible to transform all required cinnamic acid derivatives by different combinations of all enzymes. However, the variant 4HPA3H-Y301I remains the limiting factor for a full conversion to sinapic acid despite good conversion results.

Moreover, the plant-derived two-enzyme system (4CL2 and CCR) for reduction of the carboxylate to the alcohol could be replaced by a bacterial enzyme with two domains (NiCAR), which carries out the same reaction. Indeed, the NiCAR enzyme has to be activated by phosphopantetheinylation of the coenzyme NiPPTase, but purified NiCAR could be obtained in good yields, in contrast to CCR from *Leucaena leucocephala*, which was not accessible due to formation of

much insoluble protein. The activated enzyme NiCAR catalyzes the conversion of all tested cinnamic acids (**1**-**4** and **17**) to the corresponding cinnamaldehydes (**18**-**22**). Furthermore, by pairing NiCAR/NiPPTase with the cinnamyl alcohol dehydrogenase TaCAD1 from *Triticum aestivum* the instable cinnamaldehydes could be converted into more stable cinnamyl alcohols.

Despite the oxidation-sensitive of the cinnamaldehydes it is possible to prepare them by the established *in vitro* assay, whereas in *in vivo* assays the aldehydes might be toxic for *E. coli* cells or could be metabolized.

Likewise as in the first chapter, regeneration systems of cofactors for *in vitro* conversions are very important to achieve a robust conversion in the reduction reaction. By the use of G6PDH from *S. cerevisiae*, oxidized NADP$^+$ could be regenerated to the reduced NADPH in the system.

Finnigan et al. tested five different carboxylic acid reductases from different strains in kinetic studies with various substrates. These enzymes showed that they have a broad but similar substrate specificity. In case of cinnamic acid ((*E*)-3-phenylprop-2-enoic acid) the carboxylic acid reductase from *Mycobacterium phlei* (MpCAR) revealed a higher k_{cat}/K_m level (MpCAR: 240 min^{-1} mM^{-1}, NiCAR: 170 min^{-1} mM^{-1}) than our used enzyme from *Nocardia iowensis*.[142] For further studies these enzyme should be tested with the cinnamic acid derivatives as well.

In the third chapter, a cell-free multienzyme cascade for the synthesis of different phenylpropanoids was presented. The use of immobilized enzymes simplifies the purification of the reaction mixture hugely and leads to higher passes. Moreover, enzyme-loaded cartridges can be removed flexible. Within this research topic investigated enzymes of previous studies could be used to yield good conversions. For the production of vanillin (**28**) e.g. the monooxygenase 4HPA3H and the *O*-methyltransferase PFOMT could be applied (Fig. 6.2A). Similarly, the 4HPA3H variant, which revealed the best results for the conversion of naringenin (**7**) in chapter I (4HPA3H-Y301F/S462A), was used in combination with the reductase (PrnF) for the hydroxylation of the flavonoid **7** to eriodictyol (**8**) in one cascade (Fig. 6.2B). A third application of the cell-free multienzyme cascade was used for the hydrogenation of two α,β-unsaturated ketones (**29** and **30**). The well-known enzyme OYE2 from *S. cerevisiae* was tested for the conversion of the ketones **29** and **30**, but no transformation to the saturated products rheosmin (**15**) and zingerone (**31**) could be observed. Therefore, we tested the alternative enzyme OPR1 from *Lycopersicum esculentum*. With the use of this enzyme good conversions could be obtained in *in vitro* assays as well as the cascade reactions (Fig. 6.2C).
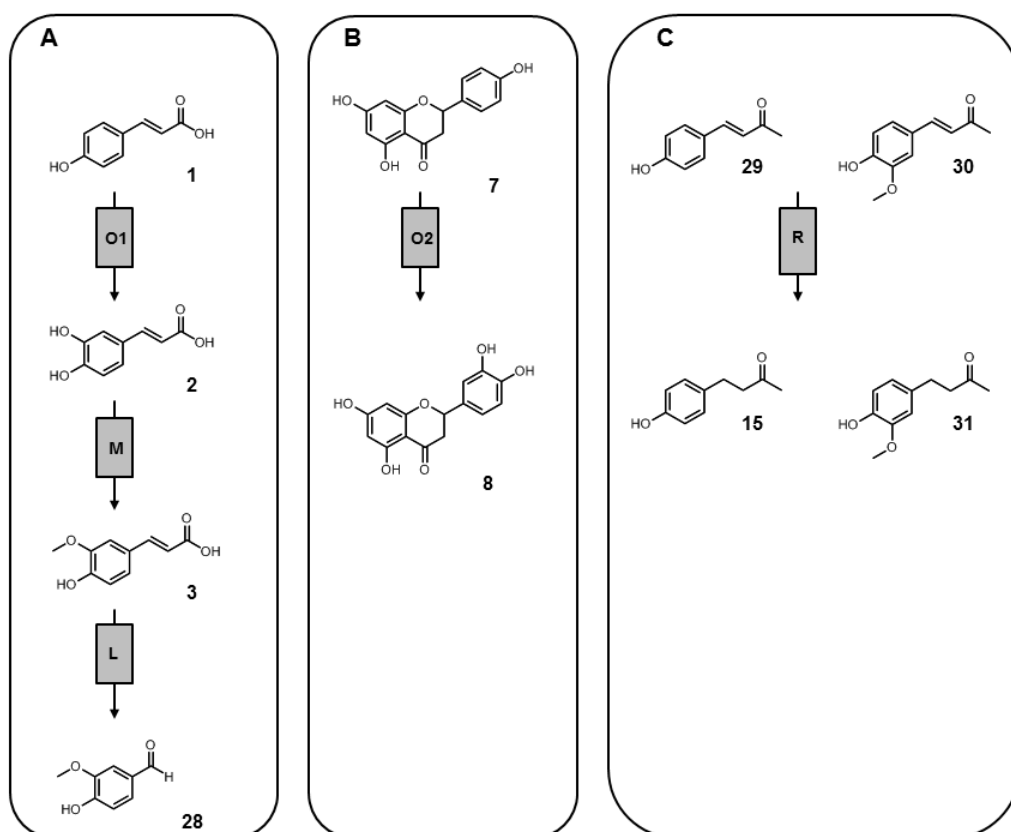
Figure 6.2: Multienzyme cascade for synthesis of different phenylpropanoids. Conversion of A) *p*-coumaric acid (**1**) to vanillin (**28**) by modules O1, M and L; B) naringenin (**7**) to eriodictyol (**8**) by modul O2; C) dehydrorheosmin (**29**) or dehydrozingerone (**30**) to rheosmin (**15**) or zingerone (**31**) by module R.

Additionally, regeneration systems for required cofactors, e.g. NADH, NADPH and $FADH_2$ could be integrated. However, the regeneration of some important cofactors is missing, e.g. ATP or coenzyme A.

ATP is often referred to as the "molecular unit of currency" of intracellular energy transfer. This small molecule transports chemical energy within cells for metabolism. Since stoichiometric amounts of ATP for transformations are necessary, a regeneration system will be an economical opportunity. Kameda et al. submitted in 2001 an ATP regeneration system using polyphosphate-AMP phosphotransferase (PAP) and polyphosphate kinase (PKK). PAP catalyzes the phospho-conversion of adenosine monophosphate (AMP) to adenosine diphosphate (ADP), and PPK catalyzes ATP formation from ADP by using inorganic polyphosphate as cheap phosphate donor (43 € per kg, Sigma Aldrich).[140] Further, the regeneration of AMP to ATP is possible by a combined action of two groups of PPK2 enzymes from *Pseudomonas aeruginosa*.[143]

Likewise, a regeneration of the expensive coenzyme A (trilithium salt, 2010 € per g, Sigma Aldrich) should be integrated. Satoh et al. established a novel enzyme-

catalyzed poly-3-hydroxybutyrate synthesis system, which is capable for recycling of CoA. In this system, the CoA molecules that were released during condensation and polymerization reactions, were reused successfully for the synthesis of acetyl-CoA.[144] One further opportunity could be to develop a novel route that does not require any coenzyme. For example, Fururya et al. developed an artificial pathway for the production of vanillin that consists of a coenzyme-independent decarboxylase and a coenzyme-independent oxygenase.[111]

6.2 Outlook

As extensively presented in the third chapter, the cascade system can be used for the successful formation of selected phenylpropanoids. Because of this the cell-free multienzyme cascade might be useful for improving results of chapter II as well. Based on this method, combinations of different modules would result by targeted control in a variety of cinnamic acids, cinnamaldehydes and cinnamyl alcohols. In Fig. 6.1 an idea of an automatized process for special production of central metabolites is presented. This process could be operated as required including regeneration of diverse cofactors.
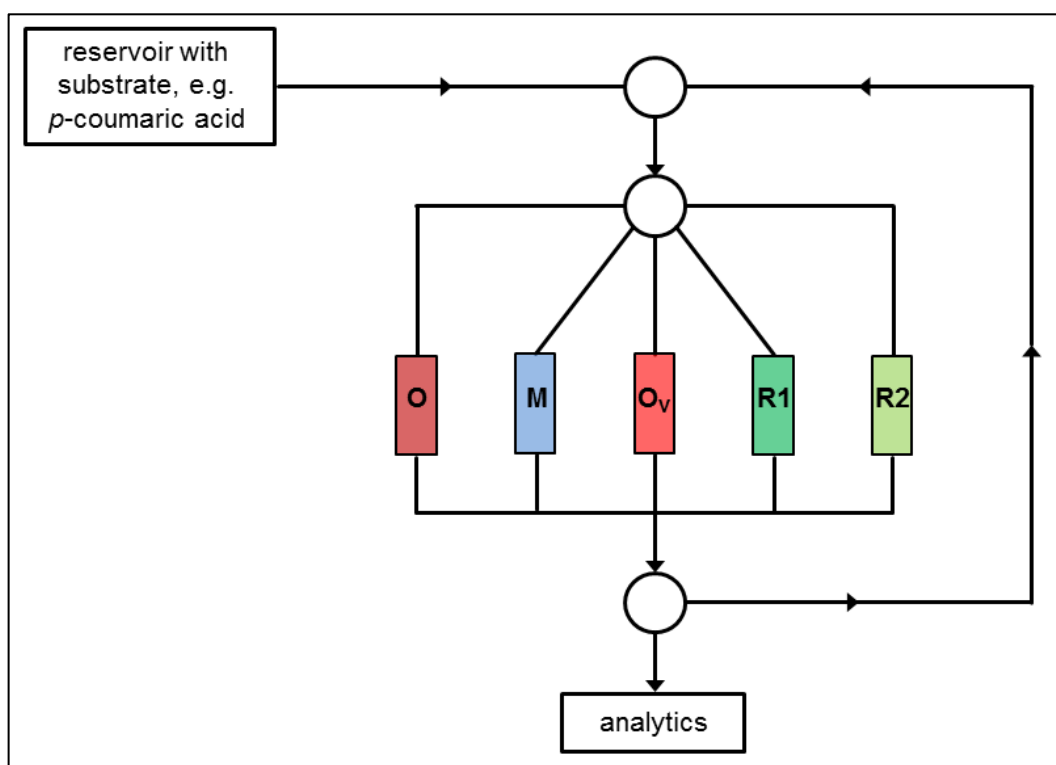
Figure 6.3: Idea of an automatized process for special production of central metabolites, e.g. cinnamic acids, cinnamaldehydes or cinnamyl alcohols. Special modules including regeneration systems for cofactors, O: monooxygenase 4HPA3H, M: methyltransferase PFOMT, O_v: 4HPA3H variant (Y301I), R1: oxidoreductase 1 NiCAR with NiPPTase, R2: oxidoreductase 2 TaCAD1.

One further possibility for effective applications in industry could be the use of some methods published in the field of synthetic biology. Here, e.g. Marillonnet and coworkers established a fast track assembly cloning method called 'Golden Gate Cloning'. With this robust, simple and efficient method it is possible to assemble at least nine separate DNA fragments into one acceptor vector.[145] The presented protocol by Engler et al. is based on the use of type IIs restriction enzymes (e.g. BsaI or BpiI).[51,145] It is performed in one step and one tube by simply subjecting a mix of 10 undigested input plasmids (nine insert plasmids and one acceptor vector) to a restriction-ligation and transformation of the resulting mix into competent cells (Fig. 6.2).[51] The only requirement is to construct a set of compatible restriction sites in the entry clones and the acceptor vector.[115] This allows generating libraries of recombinant genes by combining several fragments in one reaction.[51,146]
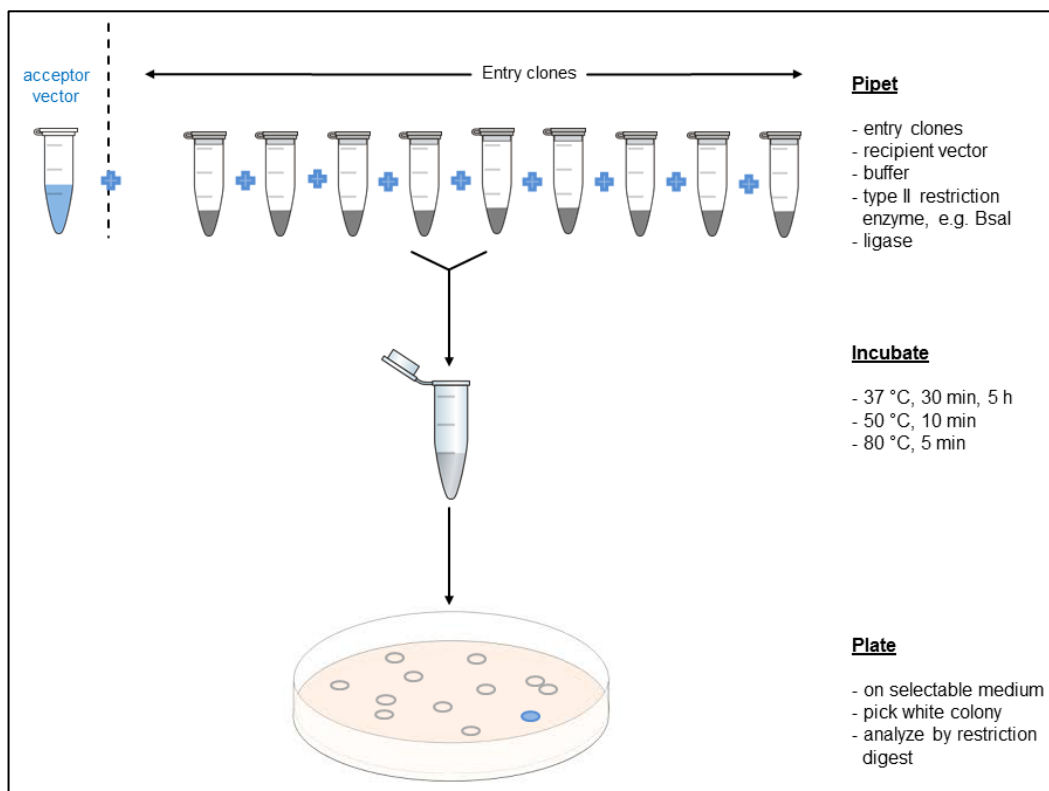
Figure 6.4: Schematic overview of Golden Gate Cloning system. In principle entry clones and acceptor clones were pipetted, incubated and plated on a selectable medium.

'Golden Gate Cloning' cannot only be used for *in vivo* reactions. It can be used for *in vitro* conversions as well, because it is possible to introduce fragments, which allow a subsequent purification or additional applications of the generated module (e.g. by His_6-tag or maltose-binding protein).

Based on this method, it is conceivable to create a changeable set of enzymes from the monolignol pathway. For the production of differently decorated cinnamyl alcohols, e.g *p*-coumaryl (**23**) or sinapyl alcohol (**27**), many immediate enzyme combinations could be designed. Thereby, it might be possible to design a specialized production system for the central metabolites of the biosynthesis of lignin or for flavors. It will allow a high variability for getting access to phenylpropanoids. More or less this method will be the 'synthetic biology'-answer to the automatized process mentioned before.

7   Appendix

7.1   Appendix A – Nucleotide and protein sequences

4HPA3H – nucleotide sequence

ATGAAACCAGAAGATTTCCGCGCCAGTACCCAACGTCCTTTCACCGGGGAAGAGTATCTGAAAAGCCTGCA
GGATGGTCGCGAGATCTATATCTATGGCGAGCGAGTGAAAGACGTCACCACTCATCCGGCATTTCGTAATG
CGGCAGCGTCTGTTGCCCAGCTGTACGACGCACTGCACAAACCGGAGATGCAGGACTCTCTGTGTTGGAA
CACCGACACCGGCAGCGGCGGCTATACCCATAAAATTCTTCCGCGTGGCGAAAAGTGCCGACGACCTGCG
CCAGCAACGCGACGCCATCGCTGAGTGGTCACGCCTGAGCTATGGCTGGATGGGCCGTACCCCAGACTA
CAAAGCCGCTTTCGGTTGCGCACTGGGCGCGAATCCGGGCTTTTACGGTCAGTTCGAGCAGAACGCCCGT
AACTGGTACACCCGTATTCAGGAAACTGGCCTCTACTTTAACCACGCGATTGTTAACCCACCGATCGATCG
TCATTTGCCGACCGATAAAGTGAAAGACGTTTACATCAAGCTGGAAAAAGAGACTGACGCCGGGATTATCG
TCAGCGGTGCGAAAGTGGTTGCCACCAACTCGGCGCTGACTCACTACAACATGATTGGCTTCGGCTCGGC
ACAAGTGATGGGCGAAAACCCGGACTTCGCACTGATGTTCGTTGCGCCAATGGATGCCGATGGCGTGAAA
TTAATCTCCCGCGCCTCTTATGAGATGGTCGCGGGTGCTACCGGCTCGCCATACGACTACCCGCTCTCCA
GCCGCTTCGATGAGAACGATGCGATTCTGGTGATGGATAACGTGCTGATTCCATGGGAAAACGTGCTGATC
TACCGCGATTTTGATCGCTGCCGTCGCTGGACGATGGAAGGCGGTTTTGCCCGTATGTATCCGCTGCAAG
CCTGTGTGCGCCTGGCAGTGAAATTAGACTTCATTACGGCACTGCTGAAAAAATCACTCGAATGTACCGGC
ACCCTGGAGTTCCGTGGTGTGCAGGCCGATCTCGGTGAAGTGGTAGCGTAGCGCAACACCTTCTGGGCAT
TGAGTGACTCGATGTGTTCAGAAGCAACGCCCGTGGGTCAACGGGGCTTATTTACCGGATCATGCCGCACT
GCAAACCTATCGCGTACTGGCACCAATGGCCTACGCGAAGATCAAAAACATTATCGAACGCAACGTTACCA
GTGGCCTGATCTATCTCCCTTCCAGTGCCCGTGACCTGAATAATCCGCAGATCGACCAGTATCTGGCGAAG
TATGTGCGCGGTTCGAACGGTATGGATCACGTCCAGCGCATCAAGATCCTCAAACTGATGTGGGATGCTAT
TGGCAGCGAATTTGGTGGTCGTCACGAACTGTATGAAATCAACTACTCCGGTAGCCAGGATGAGATTCGCC
TGCAGTGTCTGCGCCAGGCACAAAACTCCGGCAATATGGACAAGATGATGGCGATGGTTGATCGCTGCCT
GTCGGAATACGACCAGGACGGCTGGACTGTGCCGCACCTGCACAACAACGACGATATCAACATGCTGGAT
AAGCTGCTGAAATAA

4HPA3H – protein sequence

M K P E D F R A S T Q R P F T G E E Y L K S L Q D G R E I Y I Y G E R V K D V T T H P A F R N A A A S V
A Q L Y D A L H K P E M Q D S L C W N T D T G S G G Y T H K F F R V A K S A D D L R Q Q R D A I A E W
S R L S Y G W M G R T P D Y K A A F G C A L G A N P G F Y G Q F E Q N A R N W Y T R I Q E T G L Y F N
H A I V N P P I D R H L P T D K V K D V Y I K L E K E T D A G I I V S G A K V V A T N S A L T H Y N M I G F
G S A Q V M G E N P D F A L M F V A P M D A D G V K L I S R A S Y E M V A G A T G S P Y D Y P L S S R
F D E N D A I L V M D N V L I P W E N V L I Y R D F D R C R R W T M E G G F A R M Y P L Q A C V R L A V
K L D F I T A L L K K S L E C T G T L E F R G V Q A D L G E V V A W R N T F W A L S D S M C S E A T P W
V N G A Y L P D H A A L Q T Y R V L A P M A Y A K I K N I I E R N V T S G L I Y L P S S A R D L N N P Q I D
Q Y L A K Y V R G S N G M D H V Q R I K I L K L M W D A I G S E F G G R H E L Y E I N Y S G S Q D E I R L
Q C L R Q A Q N S G N M D K M M A M V D R C L S E Y D Q D G W T V P H L H N N D D I N M L D K L L K
Stop

HpaC – nucleotide sequence

ATGCAATTAGATGAACAACGCCTGCGCTTTCGTGACGCGATGGCCAGCCTGTCGGCAGCGGTAAATATTAT
CACCACCGAGGGCGACGCCGGACAATGCGGGATTACGGCAACGGCCGTCTGCTCGGTCACGGATACACC
ACCGTCGCTGATGGTGTGCATTAACGCCAACAGTGCGATGAACCCGGTTTTTCAGGGCAACGGCAAGTTG
TGCGTCAACGTCCTCAACCATGAGCAGGAACTGATGGCACGCCACTTCGCGGGCATGACAGGCATGGCGA
TGGAAGAGCGTTTTAGCCTCTCATGCTGGCAAAAAGGTCCGCTGGCGCAGCCGGTGCTAAAAGGTTCGCT
GGCCAGTCTTGAAGGTGAGATCCGCGATGTGCAGGCAATTGGCACACATCTGGTGTATCTGGTGGAGATT
AAAAACATCATCCTCAGTGCAGAAGGTCATGGACTTATCTACTTTAAACGCCGTTTCCATCCGGTGATGCTG
GAAATGGAAGCTGCGATTTAA

HpaC – protein sequence

M Q L D E Q R L R F R D A M A S L S A A V N I I T T E G D A G Q C G I T A T A V C S V T D T P P S L M V C
I N A N S A M N P V F Q G N G K L C V N V L N H E Q E L M A R H F A G M T G M A M E E R F S L S C W Q
K G P L A Q P V L K G S L A S L E G E I R D V Q A I G T H L V Y L V E I K N I I L S A E G H G L I Y F K R R F
H P V M L E M E A A I Stop

PrnF – nucleotide sequence

ATGAATGCTGCCACCGAAACCAAAGTTCACGATTTGCTCGATGCCGAGGGCCGCGATGTCCGCGATGCCC
GTGAACTGCGCAACGTGCTGGGGCAGTTTGCTACCGGTGTGACCGTGATCACCACCCGCACCGCAGACG
GCCGCAACGTCGGTGTGACGGCCAACTCGTTCTCCTCACTGTCGCTGTCACCGGCGCTGGTGCTCTGGAG
CCTGGCACGCACGGCACCGAGCCTGAAGGCCTTTTGCTCGGCGAGCCACTTCGCCATCAACGTACTGGG
CGCGCACCAGCACCACCTGTCGGAGCAGTTCGCACGGGCAGCAGCCGACAAGTTCGCCGGTGTAGCTCA
TTCCTATGGCAAGGCGGGAGCCCCGGTGCTGGACGATGTGGTGGCAGTGCTGGTGTGCCGCAACGTCAC
CCAGTACGAGGGCGGTGACCACCTGATCTTCATTGGCGAGATCGAACAATACCGCTACAGCGGCGCAGAA
CCGCTGGTCTTCCATGCAGGCCAGTACCGGGGGCTAGGGAGCAATAGAGCAGAAAGCGTCCTCAAGCAC
GAATAG

PrnF – protein sequence

M N A A T E T K V H D L L D A E G R D V R D A R E L R N V L G Q F A T G V T V I T T R T A D G R N V G V
T A N S F S S L S L S P A L V L W S L A R T A P S L K A F C S A S H F A I N V L G A H Q H H L S E Q F A R
A A A D K F A G V A H S Y G K A G A P V L D D V V A V L V C R N V T Q Y E G G D H L I F I G E I E Q Y R Y
S G A E P L V F H A G Q Y R G L G S N R A E S V L K H E Stop

ScFDH – nucleotide sequence

ATGTCGAAGGGAAAGGTTTTGCTGGTTCTTTACGAAGGTGGTAAGCATGCTGAAGAGCAGGAAAAGTTATT
GGGGTGTATTGAAAATGAACTTGGTATCAGAAATTTCATTGAAGAACAGGGATACGAGTTGGTTACTACCAT
TGACAAGGACCCTGAGCCAACCTCAACGGTAGACAGGGAGTTGAAAGACGCTGAAATTGTCATTACTACGC
CCTTTTTCCCCGCCTACATCTCGAGAAACAGGATTGCAGAAGCTCCTAACCTGAAGCTCTGTGTAACCGCT
GGCGTCGGTTCAGACCATGTCGATTTAGAAGCTGCAAATGAACGGAAAATCACGGTCACCGAAGTTACTGG
TTCTAACGTCGTTTCTGTCGCAGAGCACGTTATGGCCACAATTTTGGTTTTGATAAGAAACTATAATGGTGG
TCATCAACAAGCAATTAATGGTGAGTGGGATATTGCCGGCGTGGCTAAAAATGAGTATGATCTGGAAGACA
AAATAATTTCAACGGTAGGTGCCGGTAGAATTGGATATAGGGTTCTGGAAAGATTGGTCGCATTTAATCCGA
AGAAGTTACTGTACTACGACTACCAGGAACTACCTGCGGAAGCAATCAATAGATTGAACGAGGCCAGCAAG
CTTTTCAATGGCAGAGGTGATATTGTTCAGAGAGTAGAGAAATTGGAGGATATGGTTGCTCAGTCAGATGTT
GTTACCATCAACTGTCCATTGCACAAGGACTCAAGGGGTTTATTCAATAAAAAGCTTATTTCCCACATGAAA
GATGGTGCATACTTGGTGAATACCGCTAGAGGTGCTATTTGTGTCGCAGAAGATGTTGCCGAGGCAGTCAA
GTCTGGTAAATTGGCTGGCTATGGTGGTGATGTCTGGGATAAGCAACCAGCACCAAAAGACCATCCCTGG
AGGACTATGGACAATAAGGACCACGTGGGAAACGCAATGACTGTTCATATCAGTGGCACATCTCTGGATGC
TCAAAAGAGGTACGCTCAGGGAGTAAAGAACATCCTAAATAGTTACTTTTCCAAAAAGTTTGATTACCGTCC
ACAGGATATTATTGTGCAGAATGGTTCTTATGCCACCAGAGCTTATGGACAGAAGAAATAA

ScFDH – protein sequence

M S K G K V L L V L Y E G G K H A E E Q E K L L G C I E N E L G I R N F I E E Q G Y E L V T T I D K D P E
P T S T V D R E L K D A E I V I T T P F F P A Y I S R N R I A E A P N L K L C V T A G V G S D H V D L E A A
N E R K I T V T E V T G S N V V S V A E H V M A T I L V L I R N Y N G G H Q Q A I N G E W D I A G V A K N
E Y D L E D K I I S T V G A G R I G Y R V L E R L V A F N P K K L L Y Y D Y Q E L P A E A I N R L N E A S K
L F N G R G D I V Q R V E K L E D M V A Q S D V V T I N C P L H K D S R G L F N K K L I S H M K D G A Y L
V N T A R G A I C V A E D V A E A V K S G K L A G Y G G D V W D K Q P A P K D H P W R T M D N K D H V
G N A M T V H I S G T S L D A Q K R Y A Q G V K N I L N S Y F S K K F D Y R P Q D I I V Q N G S Y A T R A
Y G Q K K Stop

PFOMT – nucleotide sequence

ATGGATTTTGCTGTGATGAAGCAGGTCAAAAATACAGGATTGCTACAGAGTGAGGAGTTATGCCAGTATATT
CTCCGAACTAGTGTCTATCCGCGAGAAGCAGGGTTCCTCAAGGAACTCAGGGAAGCCAATGAAAGTCACC
CAGACTCTTATATGTCGACTTCACCACTTGCTGGACAATTGATGTCATTCGTTCTAAAATTAGTGAATGCAAA
GAAGACTATTGAAGTTGGAGTCTTTACAGGATACTCCCTCTTACTCACTGCTCTTTCAATTCCTGATGATGG
AAAGATTACGGCAATTGATTTCGACAGAGAGGCATATGAGATTGGCTTGCCATTTATCAGAAAAGCTGGTGT
GGAGCACAAAATCAACTTCATTGAATCGGATGCTATGCTAGCTCTTGACAATCTTCTGCAAGGACAAGAGA
GCGAGGGGAGTTACGACTTTGGCTTTGTTGATGCGGACAAACCTAACTACATCAAGTACCATGAGAGGTTG
ATGAAACTAGTCAAGGTGGGTGGCATAGTCGCTTATGACAACACATTATGGGGTGGAACTGTAGCCCAGCC
TGAATCCGAAGTACCAGATTTCATGAAGGAAAACAGAGAAGCTGTTATTGAACTCAACAAGTTGCTTGCTGC
TGATCCTCGTATCGAGATTGTACATCTTCCTTTGGGTGATGGTATCACTTTCTGCAGGCGTCTTTATTGA

PFOMT – protein sequence

M D F A V M K Q V K N T G L L Q S E E L C Q Y I L R T S V Y P R E A G F L K E L R E A N E S H P D S Y M
S T S P L A G Q L M S F V L K L V N A K K T I E V G V F T G Y S L L L L T A L S I P D D G K I T A I D F D R E
A Y E I G L P F I R K A G V E H K I N F I E S D A M L A L D N L L Q G G Q E S E G S Y D F G F V D A D K P N
Y I K Y H E R L M K L V K V G G I V A Y D N T L W G G T V A Q P E S E V P D F M K E N R E A V I E L N K L
L A A D P R I E I V H L P L G D G I T F C R R L Y Stop

LICCR- nucleotide sequence

ATGCCAGCTGCCGCCCCTGCCCCAACCGCCGCTAACACGACTTCTTCTGGTTCTGGTCAAACCGTCTGCG
TTACCGGTGCGGGTGGCTTTATTGCCAGCTGGATCGTCAAGTTGCTGCTGGAGCGTGGTTACACCGTTCG
TGGTACCGTTCGCAATCCGGACGATAGCAAGAATAGCCACCTGAAAGAGCTGGAAGGCGCGGAAGAACGT
CTGACGCTGCATAAGGTTGATTTGCTGGACCTGGAGAGCGTCAAGGCAGTTATCAATGGTTGCGACGGCA
TTATCCACACCGCTAGCCCGGTGACCGACAACCCGGAAGAAATGGTCGAACCGGCGGTGAACGGCGCGA
AAAATGTGATTATTGCAGCGGCAGAGGCGAAAGTCCGTCGTGTGGTTTTCACGAGCAGCATCGGTGCGGT
TTACATGGACCCGAGCCGCAATATTGACGAAGTGGTGGATGAGAGCTGTTGGTCGAACCTGGAGTACTGC
AAAAACACGAAAAACTGGTACTGTTATGGTAAAGCAGTCGCGGAGCAAGCCGCGTGGGACGAAGCCAAGG
CGCGTGGTGTGGATCTGGTGGTTGTGAACCCGGTTCTCGTCCTGGGTCCGCTGCTGCAGTCCACCATGAA
TGCAAGCACCATCCACATCCTGAAATATCTGACCGGCTCCGCAAAGACTTATGCGAACGCAACGCAAGCAT
ACGTGCATGTTAAAGATGTCGCGTTGGCTCACGTTCTGGTCTACGAGATTCCGAGCGCGTCCGGTCGCTAT
CTGTGCAGCGAGAGCTCACTGCATCGTGGCGAACTGGTTGAAATCTTGGCAAAGTACTTCCCGGAGTACC
CGATTCCGACCAAGTGCAGCGATGAGAAGAACCCTCGTGCGAAGGCGTATACCTTTTCGAATAAACGCCT
GAAAGATCTGGGTCTGGAGTTCACCCCGGTGCACCAGTGTTTGTATGACACGGTGAAAAGCCTGCAGGAT
AAGGGCCACTTACCGCTGCCGACGAAGTAA

LICCR – protein sequence

M P A A A P A P T A A N T T S S G S G Q T V C V T G A G G F I A S W I V K L L L E R G Y T V R G T V R N
P D D S K N S H L K E L E G A E E R L T L H K V D L L D L E S V K A V I N G C D G I I H T A S P V T D N P
E E M V E P A V N G A K N V I I A A A E A K V R R V V F T S S I G A V Y M D P S R N I D E V V D E S C W S
N L E Y C K N T K N W Y C Y G K A V A E Q A A W D E A K A R G V D L V V V N P V L V L G P L L Q S T M
N A S T I H I L K Y L T G S A K T Y A N A T Q A Y V H V K D V A L A H V L V Y E I P S A S G R Y L C S E S S
L H R G E L V E I L A K Y F P E Y P I P T K C S D E K N P R A K A Y T F S N K R L K D L G L E F T P V H Q
C L Y D T V K S L Q D K G H L P L P T K Stop

NiCAR – nucleotide sequence

ATGGCAGTGGATTCACCGGATGAGCGGCTACAGCGCCGCATTGCACAGTTGTTTGCAGAAGATGAGCAGG
TCAAGGCCGCACGTCCGCTCGAAGCGGTGAGCGCGGCGGTGAGCGCGCCCGGTATGCGGCTGGCGCAG
ATCGCCGCCACTGTTATGGCGGGTTACGCCGACCGCCCGGCCGCCGGGCAGCGTGCGTTCGAACTGAAC
ACCGACGACGCGACGGGCCGCACCTCGCTGCGGTTACTTCCCCGATTCAAGACCATCACCTATCGCGAAC
TGTGGCAGCGAGTCGGCGAGGTTGCCGCGGCCTGGCATCATGATCCCGAGAACCCCTTGCGCGCAGGTG
ATTTCGTCGCCCTGCTCGGCTTCACCAGCATCGACTACGCCACCCTCGACCTGGCCGATATCCACCTCGG
CGCGGTTACCGTGCCGTTGCAGGCCAGCGCGGCGGTGTCCCAGCTGATCGCTATCCTCACCGAGACTTC
GCCGCGGCTGCTCGCCTCGACCCCGGAGCACCTCGATGCGGCGGTCGAGTGCCTACTCGCGGGCACCA
CACCGGAACGACTGGTGGTCTTCGACTACCACCCCGAGGACGACGACCAGCGTGCGGCCTTCGAATCCG
CCCGCCGCCGCCTTGCCGACGCGGGCAGCTCGGTGATCGTCGAAACGCTCGATGCCGTGCGTGCCCGG
GGCCGCGACTTACCGGCCGCGCCACTGTTCGTTCCCGACACCGACGACGACCCGCTGGCCCTGCTGATC
TACACCTCCGGCAGCACCGGAACGCCGAAGGGCGCGATGTACACCAATCGGTTGGCCGCCACGATGTGG
CAGGGGAACTCGATGCTGCAGGGGAACTCGCAACGGGTCGGGATCAATCTCAACTACATGCCGATGAGCC
ACATCGCCGGTCGCATATCGCTGTTCGGCGTGCTCGCTCGCGGTGGCACCGCATACTTCGCGGCCAAGA
GCGACATGTCGACACTGTTCGAAGACATCGGCTTGGTACGTCCCACCGAGATCTTCTTCGTCCCGCGCGT
GTGCGACATGGTCTTCCAGCGCTATCAGAGCGAGCTGGACCGGCGCTCGGTGGCGGGCGCCGACCTGGA
CACGCTCGATCGGGAAGTGAAAGCCGACCTCCGGCAGAACTACCTCGGTGGGCGCTTCCTGGTGGCGGT
CGTCGGCAGCGCGCCGCTGGCCGCGGAGATGAAGACGTTCATGGAGTCCGTCCTCGATCTGCCACTGCA
CGACGGGTACGGGTCGACCGAGGCGGGCGCAAGCGTGCTGCTCGACAACCAGATCCAGCGGCCGCCGG
TGCTCGATTACAAGCTCGTCGACGTGCCCGAACTGGGTTACTTCCGCACCGACCGGCCGCATCCGCGCG
GTGAGCTGTTGTTGAAGGCGGAGACCACGATTCCGGGCTACTACAAGCGGCCCGAGGTCACCGCGGAGA
TCTTCGACGAGGACGGCTTCTACAAGACCGGCGATATCGTGGCCGAGCTCGAGCACGATCGGCTGGTCTA
TGTCGACCGTCGCAACAATGTGCTCAAACTGTCGCAGGGCGAGTTCGTGACCGTCGCCCATCTCGAGGCC
GTGTTCGCCAGCAGCCCGCTGATCCGGCAGATCTTCATCTACGGCAGCAGCGAACGTTCCTATCTGCTCG
CGGTGATCGTCCCCACCGACGACGCGCTGCGCGGCCGCGACACCGCCACCTTGAAATCGGCACTGGCCG
AATCGATTCAGCGCATCGCCAAGGACGCGAACCTGCAGCCCTACGAGATTCCGCGCGATTTCCTGATCGA
GACCGAGCCGTTCACCATCGCCAACGGACTGCTCTCCGGCATCGCGAAGCTGCTGCGCCCCAATCTGAAG
GAACGCTACGGCGCTCAGCTGGAGCAGATGTACACCGATCTCGCGACAGGCCAGGCCGATGAGCTGCTC
GCCCTGCGCCGCGAAGCCGCCGACCTGCCGGTGCTCGAAACCGTCAGCCGGGCAGCGAAAGCGATGCT

CGGCGTCGCCTCCGCCGATATGCGTCCCGACGCGCACTTCACCGACCTGGGCGGCGATTCCCTTTCCGC
GCTGTCGTTCTCGAACCTGCTGCACGAGATCTTCGGGGTCGAGGTGCCGGTGGGTGTCGTCGTCAGCCC
GGCGAACGAGCTGCGCGATCGGCGAATTACATTGAGGCGGAACGCAACTCGGGCGCGAAGCGTCCCAC
CTTCACCTCGGTGCACGGCGGCGGTTCCGAGATCCGCGCCGCCGATCTGACCCTCGACAAGTTCATCGAT
GCCCGCACCCTGGCCGCCGCCGACAGCATTCCGCACGCGCCGGTGCCAGCGCAGACGGTGCTGCTGAC
CGGCGCGAACGGCTACCTCGGCCGGTTCCTGTGCCTGGAATGGCTGGAGCGGCTGGACAAGACGGGTG
GCACGCTGATCTGCGTCGTGCGCGGTAGTGACGCGGCCGCGGCCCGTAAACGGCTGGACTCGGCGTTCG
ACAGCGGCGATCCCGGCCTGCTCGAGCACTACCAGCAACTGGCCGCACGGACCCTGGAAGTCCTCGCCG
GTGATATCGGCGACCCGAATCTCGGTCTGGACGACGCGACTTGGCAGCGGTTGGCCGAAACCGTCGACC
TGATCGTCCATCCCGCCGCGTTGGTCAACCACGTCCTTCCCTACACCCAGCTGTTCGGCCCCAATGTCGT
CGGCACCGCCGAAATCGTCCGGTTGGCGATCACGGCGCGGCGCAAGCCGGTCACCTACCTGTCGACCGT
CGGAGTGGCCGACCAGGTCGACCCGGCGGAGTATCAGGAGGACAGCGACGTCCGCGAGATGAGCGCGG
TGCGCGTCGTGCGCGAGAGTTACGCCAACGGCTACGGCAACAGCAAGTGGGCGGGGGAGGTCCTGCTG
CGCGAAGCACACGATCTGTGTGGCTTGCCGGTCGCGGTGTTCCGTTCGGACATGATCCTGGCGCACAGC
CGGTACGCGGGTCAGCTCAACGTCCAGGACGTGTTCACCCGGCTGATCCTCAGCCTGGTCGCCACCGGC
ATCGCGCCGTACTCGTTCTACCGAACCGACGCGGACGGCAACCGGCAGCGGGCCCACTACGACGGTCTG
CCCGCCGATTTCACGGCGGCGGCCGATCACCGCGCTCGGCATCCAAGCCACCGAAGGCTTCCGGACCTAC
GACGTGCTCAATCCGTACGACGATGGCATCTCCCTCGATGAATTCGTCGACTGGCTCGTCGAATCCGGCC
ACCCGATCCAGCGCATCACCGACTACAGCGACTGGTTCCACCGTTTCGAGACGGCGATCCGCGCGCTGC
CGGAAAAGCAACGCCAGGCCTCGGTGCTGCCGTTGCTGGACGCCTACCGCAACCCCTGCCCGGCGGTCC
GCGGCGCGATACTCCCGGCCAAGGAGTTCCAAGCGGCGGTGCAAACAGCCAAAATCGGTCCGGAACAGG
ACATCCCGCATTTGTCCGCGCCACTGATCGATAAGTACGTCAGCGATCTGGAACTGCTTCAGCTGCTCTGA

NiCAR – protein sequence

M A V D S P D E R L Q R R I A Q L F A E D E Q V K A A R P L E A V S A A V S A P G M R L A Q I A A T V M
A G Y A D R P A A G Q R A F E L N T D D A T G R T S L R L L P R F K T I T Y R E L W Q R V G E V A A A W
H H D P E N P L R A G D F V A L L G F T S I D Y A T L D L A D I H L G A V T V P L Q A S A A V S Q L I A I L
T E T S P R L L A S T P E H L D A A V E C L L A G T T P E R L V V F D Y H P E D D D Q R A A F E S A R R
R L A D A G S S V I V E T L D A V R A R G R D L P A A P L F V P D T D D D P L A L L I Y T S G S T G T P K
G A M Y T N R L A A T M W Q G N S M L Q G N S Q R V G I N L N Y M P M S H I A G R I S L F G V L A R G
G T A Y F A A K S D M S T L F E D I G L V R P T E I F F V P R V C D M V F Q R Y Q S E L D R R S V A G A
D L D T L D R E V K A D L R Q N Y L G G R F L V A V V G S A P L A A E M K T F M E S V L D L P L H D G Y
G S T E A G A S V L L D N Q I Q R P P V L D Y K L V D V P E L G Y F R T D R P H P R G E L L L K A E T T I
P G Y Y K R P E V T A E I F D E D G F Y K T G D I V A E L E H D R L V Y V D R R N N V L K L S Q G E F V T
V A H L E A V F A S S P L I R Q I F I Y G S S E R S Y L L A V I V P T D D A L R G R D T A T L K S A L A E S I
Q R I A K D A N L Q P Y E I P R D F L I E T E P F T I A N G L L S G I A K L L R P N L K E R Y G A Q L E Q M
Y T D L A T G Q A D E L L A L R R E A A D L P V L E T V S R A A K A M L G V A S A D M R P D A H F T D L
G G D S L S A L S F S N L L H E I F G V E V P V G V V V S P A N E L R D L A N Y I E A E R N S G A K R P T
F T S V H G G G S E I R A A D L T L D K F I D A R T L A A A D S I P H A P V P A Q T V L L T G A N G Y L G
R F L C L E W L E R L D K T G G T L I C V V R G S D A A A A R K R L D S A F D S G D P G L L E H Y Q Q L
A A R T L E V L A G D I G D P N L G L D D A T W Q R L A E T V D L I V H P A A L V N H V L P Y T Q L F G P
N V V G T A E I V R L A I T A R R K P V T Y L S T V G V A D Q V D P A E Y Q E D S D V R E M S A V R V V
R E S Y A N G Y G N S K W A G E V L L R E A H D L C G L P V A V F R S D M I L A H S R Y A G Q L N V Q
D V F T R L I L S L V A T G I A P Y S F Y R T D A D G N R Q R A H Y D G L P A D F T A A A I T A L G I Q A T
E G F R T Y D V L N P Y D D G I S L D E F V D W L V E S G H P I Q R I T D Y S D W F H R F E T A I R A L P
E K Q R Q A S V L P L L D A Y R N P C P A V R G A I L P A K E F Q A A V Q T A K I G P E Q D I P H L S A P
L I D K Y V S D L E L L Q L L Stop

NiPPTase – nucleotide sequence

ATGATCGAGACAATTTTGCCTGCTGGTGTCGAGTCGGCTGAGCTGCTGGAGTATCCGGAGGACCTGAAGG
CGCATCCGGCGGAGGAGCATCTCATCGCGAAGTCGGTGGAGAAGCGGCGCCGGGACTTCATCGGGGCC
AGGCATTGTGCCCGGCTGGCGCTGGCTGAGCTCGGCGAGCCGCCGGTGGCGATCGGCAAAGGGGAGCG
GGGTGCGCCGATCTGGCCGCGCGGCGTCGTCGGCAGCCTCACCCATTGCGACGGATATCGGGCCGCGG
CGGTGGCGCACAAGATGCGCTTCCGTTCGATCGGCATCGATGCCGAGCCGCACGCGACGCTGCCCGAAG
GCGTGCTGGATTCGGTCAGCCTGCCGCCGGAGCGGGAGTGGTTGAAGACCACCGATTCCGCACTGCACC
TGGACCGTTTACTGTTCTGCGCCAAGGAAGCCACCTACAAGGCGTGGTGGCCGCTGACCGCGCGCTGGC
TCGGCTTCGAGGAAGCGCACATCACCTTCGAGATCGAAGACGGCTCCGCCGATTCCGGCAACGGCACCTT
TCACAGCGAGCTGCTGGTGCCGGGACAGACGAATGACGGTGGGACGCCGCTGCTTTCGTTCGACGGCCG
GTGGCTGATCGCCGACGGGTTCATCCTCACCGCGATCGCGTACGCCTGA

NiPPTase – protein sequence

M I E T I L P A G V E S A E L L E Y P E D L K A H P A E E H L I A K S V E K R R R D F I G A R H C A R L A L
A E L G E P P V A I G K G E R G A P I W P R G V V G S L T H C D G Y R A A A V A H K M R F R S I G I D A
E P H A T L P E G V L D S V S L P P E R E W L K T T D S A L H L D R L L F C A K E A T Y K A W W P L T A
R W L G F E E A H I T F E I E D G S A D S G N G T F H S E L L V P G Q T N D G G T P L L S F D G R W L I
A D G F I L T A I A Y A Stop

TaCAD1 – nucleotide sequence

ATGGGCAGCGTCGACGCCTCCGAGACGACGGTCACCGGGTGGGCCGCCAGGGACGCCACCGGCCACCT
CTCCCCCTACACATACACCCTCAGGAAAACAGGCCCTGAAGACGTGGTGTTGAAGGTAAAATACTGCGGC
ATCTGCCACACAGACGTCCACCAGGTTAAGAACGACCTCGGCGCTTCCAAGTACCCCATGGTCCCAGGGC
ATGAGGTGGTTGGCGAGGTGGTGGAGGTCGGGCCGGAGGTGAGCAAGTTCCGCGCCGGCGACGTGGTG
GGCGTCGGGGTGATCGTGGGGTGCTGCCGCGACTGCCGGCCGTGCAAGGCCAACGTCGAGCAGTACTG
CAACAAGAAGATCTGGTCGTACAACGACGTCTACACCGACGGCAAGCCGACGCAGGGCGGCTTCGCCTC
CGCCATGGTCGTCGATCAAAAGTTCGTGGTGAAGATCCCCGCCGGGCTGGCGCCGGAGCAGGCGGCGCC
GCTGCTGTGCGCGGGCGTGACAGTGTACAGCCCGCTGAAGCACTTCGGGCTCATGACCCCCGGCCTCCG
CGGCGGCATCCTGGGCCTGGGAGGCGTGGGCCACATGGGCGTGAAGGTCGCCAAGTCCATGGGCCACC
ACGTGACGGTGATCAGCTCGTCCAACAAGAAGCGGGCCGAGGCCATGGACGACCTGGGCGCCGACGCCT
ACCTCGTCAGCTCCGACGCCGACCAGATGGCCGCCGCCGCCGACTCGCTGGACTACATCATCGACACCG
TGCCGGTCAAGCACCCGCTGGAGCCCTACCTGGCGCTGCTCAAGATGGACGGCAAGCTGGTGCTGATGG
GCGTGATCGCGGAGCCGCTCAGCTTCGTGTCCCCCATGGTGATGCTGGGGAGGAAGACCATCACGGGGA
GCTTCATCGGGAGCATGGACGAGACGGAGGAGGTGCTCCAGTTCTGCGTCGACAAGGGGCTCACCTCGC
AGATCGAGGTCGTCAAGATGGACTACGTTAACCAGGCGTTCGAGAGGCTCGAGCGCAACGACGTGCGCTA
CCGCTTCGTCGTCGACGTCGGCGGGAGCAACATCGAGGACGCCGCCTGA

TaCAD1 – protein sequence

M G S V D A S E T T V T G W A A R D A T G H L S P Y T Y T L R K T G P E D V V L K V K Y C G I C H T D V
H Q V K N D L G A S K Y P M V P G H E V V G E V V E V G P E V S K F R A G D V V G V G V I V G C C R D
C R P C K A N V E Q Y C N K K I W S Y N D V Y T D G K P T Q G G F A S A M V V D Q K F V V K I P A G L A
P E Q A A P L L C A G V T V Y S P L K H F G L M T P G L R G G I L G L G G V G H M G V K V A K S M G H
H V T V I S S S N K K R A E A M D D L G A D A Y L V S S D A D Q M A A A A D S L D Y I I D T V P V K H P L
E P Y L A L L K M D G K L V L M G V I A E P L S F V S P M V M L G R K T I T G S F I G S M D E T E E V L Q
F C V D K G L T S Q I E V V K M D Y V N Q A F E R L E R N D V R Y R F V V D V G G S N I E D A A STOP

G6PDH – nucleotide sequence

ATGAGTGAAGGCCCCGTCAAATTCGAAAAAAATACCGTCATATCTGTCTTTGGTGCGTCAGGTGATCTGGC
AAAGAAGAAGACTTTTCCCGCCTTATTTGGGCTTTTCAGAGAAGGTTACCTTGATCCATCTACCAAGATCTT
CGGTTATGCCCGGTCCAAATTGTCCATGGAGGAGGACCTGAAGTCCCGTGTCCTACCCCACTTGAAAAAAC
CTCACGGTGAAGCCGATGACTCTAAGGTCGAACAGTTCTTCAAGATGGTCAGCTACATTTCGGGAAATTAC
GACACAGATGAAGGCTTCGACGAATTAAGAACGCAGATCGAGAAATTCGAGAAAAGTGCCAACGTCGATGT
CCCACACCGTCTCTTCTATCTGGCCTTGCCGCCAAGCGTTTTTTTGACGGTGGCCAAGCAGATCAAGAGTC
GTGTGTACGCAGAGAATGGCATCACCCGTGTAATCGTAGAGAAACCTTTCGGCCACGACCTGGCCTCTGC
CAGGGAGCTGCAAAAAAACCTGGGGCCCCTCTTTAAAGAAGAAGAGTTGTACAGAATTGACCATTACTTGG
GTAAAGAGTTGGTCAAGAATCTTTTAGTCTTGAGGTTCGGTAACCAGTTTTTGAATGCCTCGTGGAATAGAG
ACAACATTCAAAGCGTTCAGATTTCGTTTAAAGAGAGGTTCGGCACCGAAGGCCGTGGCGGCTATTTCGAC
TCTATAGGCATAATCAGAGACGTGATGCAGAACCATCTGTTACAAATCATGACTCTCTTGACTATGGAAAGA
CCGGTGTCTTTTGACCCGGAATCTATTCGTGACGAAAAGGTTAAGGTTCTAAAGGCCGTGGCCCCCATCGA
CACGGACGACGTCCTCTTGGGCCAGTACGGTAAATCTGAGGACGGGTCTAAGCCCGCCTACGTGGATGAT
GACACTGTAGACAAGGACTCTAAATGTGTCACTTTTGCAGCAATGACTTTCAACATCGAAAACGAGCGTTG
GGAGGGCGTCCCCATCATGATGCGTGCCGGTAAGGCTTTGAATGAGTCCAAGGTGGAGATCAGACTGCAG
TACAAAGCGGTCGCATCGGGTGTCTTCAAAGACATTCCAAATAACGAACTGGTCATCAGAGTGCAGCCCGA
TGCCGCTGTGTACCTAAAGTTTAATGCTAAGACCCCTGGTCTGTCAAATGCTACCCAAGTCACAGATCTGAA
TCTAACTTACGCAAGCAGGTACCAAGACTTTTGGATTCCAGAGGCTTACGAGGTGTTGATAAGAGACGCCC
TACTGGGTGACCATTCCAACTTTGTCAGAGATGACGAATTGGATATCAGTTGGGGCATATTCACCCCATTAC
TGAAGCACATAGAGCGTCCGGACGGTCCAACACCGGAAATTTACCCCTACGGATCAAGAGGTCCAAAGGG
ATTGAAGGAATATATGCAAAAACACAAGTATGTTATGCCCGAAAAGCACCCTTACGCTTGGCCCGTGACTAA
GCCAGAAGATACGAAGGATAATTAG

G6PDH – protein sequence

M S E G P V K F E K N T V I S V F G A S G D L A K K K T F P A L F G L F R E G Y L D P S T K I F G Y A R S
K L S M E E D L K S R V L P H L K K P H G E A D D S K V E Q F F K M V S Y I S G N Y D T D E G F D E L R
T Q I E K F E K S A N V D V P H R L F Y L A L P P S V F L T V A K Q I K S R V Y A E N G I T R V I V E K P F
G H D L A S A R E L Q K N L G P L F K E E E L Y R I D H Y L G K E L V K N L L V L R F G N Q F L N A S W
N R D N I Q S V Q I S F K E R F G T E G R G G Y F D S I G I I R D V M Q N H L L Q I M T L L T M E R P V S
F D P E S I R D E K V K V L K A V A P I D T D D V L L G Q Y G K S E D G S K P A Y V D D D T V D K D S K
C V T F A A M T F N I E N E R W E G V P I M M R A G K A L N E S K V E I R L Q Y K A V A S G V F K D I P N
N E L V I R V Q P D A A V Y L K F N A K T P G L S N A T Q V T D L N L T Y A S R Y Q D F W I P E A Y E V L
I R D A L L G D H S N F V R D D E L D I S W G I F T P L L K H I E R P D G P T P E I Y P Y G S R G P K G L K
E Y M Q K H K Y V M P E K H P Y A W P V T K P E D T K D N Stop

LeOPR1 – nucleotide sequence

ATGGAAAATAAAGTCGTTGAAGAGAAACAAGTAGACAAGATCCCTCTAATGAGCCCTTGTAAAATGGGAAA
GTTTGAGTTATGTCATAGAGTTGTATTGGCACCATTAACAAGGCAAAGATCTTATGGTTATATTCCTCAACCA
CATGCTATACTTCATTACTCACAAAGAAGTACAAATGGTGGCCTTCTAATAGGAGAGGCCACAGTAATATCT
GAGACTGGCATAGGGTACAAAGATGTACCTGGTATATGGACAAAAGAGCAAGTGGAGGCTTGGAAACCAA
TTGTAGATGCAGTTCATGCTAAAGGAGGAATCTTCTTTTGCCAAATTTGGCATGTTGGTAGAGTTTCCAACA
AAGATTTTCAGCCCAATGGAGAGGATCCTATCTCCTGCACAGACAGAGGACTAACACCTCAAATTCGTTCC
AATGGCATAGATATTGCACACTTTACACGACCTAGACGGTTGACAACAGATGAAATTCCTCAAATTGTTAAC
GAATTTCGAGTTGCTGCTAGAAACGCAATTGAAGCTGGATTTGATGGGGTTGAGATCCACGGAGCTCATGG
CTATCTAATTGATCAGTTTATGAAAGATCAAGTTAACGATCGAAGTGATAAATATGGAGGGTCTTTAGAGAAT
CGTTGTAGATTTGCACTTGAAATAGTGGAAGCAGTTGCAAATGAGATTGGATCTGACCGAGTTGGTATAAG
GATATCCCCATTTGCGCATTATAATGAAGCAGGGGACACGAACCCGACTGCTTTGGGACTTTACATGGTGG
AATCGTTGAACAAGTATGATCTCGCGTATTGCCATGTGGTTGAGCCTAGGATGAAAACAGCTTGGGAAAAA
ATTGAATGTACTGAAAGCCTTGTACCGATGAGGAAGGCATATAAAGGTACTTTTATAGTAGCTGGTGGTTAC
GATAGAGAAGATGGAAACAGAGCTTTGATTGAAGATCGAGCTGATCTTGTTGCGTATGGACGTTTATTCATA
TCTAATCCAGATTTACCAAAGCGATTTGAGCTAAATGCTCCTCTTAACAAGTATAACAGAGACACATTTTATA
CTTCTGATCCAATTGTTGGCTATACTGATTATCCATTTCTAGAAACCATGACATGA

LeOPR1 – protein sequence

M E N K V V E E K Q V D K I P L M S P C K M G K F E L C H R V V L A P L T R Q R S Y G Y I P Q P H A I L H
Y S Q R S T N G G L L I G E A T V I S E T G I G Y K D V P G I W T K E Q V E A W K P I V D A V H A K G G I
F F C Q I W H V G R V S N K D F Q P N G E D P I S C T D R G L T P Q I R S N G I D I A H F T R P R R L T T
D E I P Q I V N E F R V A A R N A I E A G F D G V E I H G A H G Y L I D Q F M K D Q V N D R S D K Y G G S
L E N R C R F A L E I V E A V A N E I G S D R V G I R I S P F A H Y N E A G D T N P T A L G L Y M V E S L
N K Y D L A Y C H V V E P R M K T A W E K I E C T E S L V P M R K A Y K G T F I V A G G Y D R E D G N R
A L I E D R A D L V A Y G R L F I S N P D L P K R F E L N A P L N K Y N R D T F Y T S D P I V G Y T D Y P F
L E T M T Stop

7.2 Appendix B – MS and NMR data

NMR and MS data for compounds produced by whole-cell bioconversion (chapter I, section 3.2.3)

Caffeic acid (2): ^{1}H NMR : δ = 7.53 (d, ^3J(H, H) = 15.79 Hz, 1H; CH), δ = 7.04 (s, 1H; CH), δ = 6.93 (d, ^3J(H, H) = 7.89 Hz, 1H; CH), δ = 6.78 (d, ^3J(H, H) = 7.89 Hz, 1H; CH), δ = 6.21 (d, ^3J(H, H) = 15.79 Hz, 1H; CH) ppm; ESI-FTMS: m/z: 179.0351 [M-H]$^-$; (-)-ESI-MS^2 (25 eV): m/z (%): 179.0352 (32) [M-H]$^-$, 135.0454 (100) [M-H-CO$_2$]$^-$.

5-Hydroxyferulic acid (4): ^{1}H NMR: δ = 7.52 (d, ^3J(H, H) = 15.79 Hz, 1H; CH), δ = 7.26 (s, 1H; CH), δ = 7.24 (s, 1H; CH), δ = 6.26 (d, ^3J(H, H) = 15.78 Hz, 1H; CH), δ = 3.29 (s, 3H; CH$_3$) ppm; ESI-FTMS: m/z: 209.0457 [M-H]$^-$; (-)-ESI-MS^2 (25 eV): m/z (%): 209.0455 (46) [M-H]$^-$, 194.0221 (100) [M-H-CH$_3$]$^-$, 165.0558 (44) [M-H-CO$_2$]$^-$, 150.0324 (14) [M-H-C$_2$H$_3$O$_2$]$^-$.

Piceatannol (6): ^{1}H NMR: δ = 6.97 (s, 1H; CH), δ = 6.89 (d, ^3J(H, H) = 16.22 Hz, 1H; CH), δ = 6.82 (d, ^3J(H, H) = 7.89 Hz, 1H; CH), δ = 6.73 (d, ^3J(H, H) = 7.9 Hz, 1H; CH), δ = 6.74 (d, ^3J(H, H) = 16.22 Hz, 1H; CH), δ = 6.43 (2 x s, 2H; 2 x CH), δ = 6.16 (s, 1H; CH) ppm; ESI-FTMS: m/z: 243.0664 [M-H]$^-$; (-)-ESI-MS^2 (35 eV): m/z (%): 243.0665 (4) [M-H]$^-$, 225.0559 (100) [M-H-H$_2$O]$^-$, 215.0715 (14) [M-H-CO]$^-$, 201.0559 (64) [M-H-C$_2$H$_2$O]$^-$, 181.0661 (8) [M-H-CH$_2$O$_3$]$^-$, 175.0767 (36) [M-H-C$_3$O$_2$]$^-$, 159.0454 (15) [M-H-C$_4$H$_4$O$_2$]$^-$.

Eriodictyol (8): ^{1}H NMR: δ = 7.31 (d, ^3J(H, H) = 8.77 Hz, 1H; CH), δ = 6.91 (s, 1H; CH), δ = 6.81 (d, ^3J(H, H) = 8.77 Hz, 1H; CH), δ = 6.89 (s, 1H;CH), δ = 6.88 (s, 1H; CH), δ = 5.34 (dd, ^3J(H, H) = 13.0, ^2J(H, H) = 3.07 Hz, 1H, CH), δ = 3.11 (dd, ^3J(H, H) = 13.0, ^3J(H, H) = 17.1, 1H; CH), δ = 2.69 (dd, ^3J(H, H) = 17.1, ^2J(H, H) = 3.0 Hz, 1H; CH) ppm; ESI-FTMS: m/z: 287.0565 [M-H]$^-$; (-)-ESI-MS^2 (25 eV): m/z (%): 287.0565 (6) [M-H]$^-$, 181.0145 (4) [M-H-C$_7$H$_6$O]$^-$, 151.0040 (100) [M-H-C$_8$H$_8$O$_2$]$^-$, 125.0247 (4) [M-H-C$_9$H$_6$O$_3$]$^-$.

3,4-Dihydroxybenzoic acid (10): ^{1}H NMR: δ = 7.41 (s, 1H; CH), δ = 7.03 (d, ^3J(H, H) = 8.34 Hz, 1H; CH), δ = 6.70 (d, ^3J(H, H) = 8.33 Hz, 1H; CH); ESI-FTMS: m/z: 153.0195 [M-H]$^-$; (-)-ESI-MS^2 (25 eV): m/z (%): 153.0196 (6) [M-H]$^-$, 109.0298 (100) [M-H-CO$_2$]$^-$.

2,3-Dihydroxycarbazole (12): ^{1}H NMR: δ = 7.80 (d, ^3J(H, H) = 7.46 Hz, 1H; CH), δ = 7.39 (s, 1H; CH), δ = 7.30 (d, ^3J(H, H) = 7.89 Hz, 1H; CH), δ = 7.19 (dd, ^3J(H, H) = 7.01 Hz, ^3J(H, H) = 7.89 Hz, 1H; CH), δ = 7.02 (dd, ^3J(H, H) = 7.01 Hz, ^3J(H, H) = 7.45 Hz, 1H; CH) ppm; ESI-FTMS: *m/z*: 198.0560 [M-H]$^-$; (-)-ESI-MS^2 (35 eV): *m/z* (%): 198.0562 (2) [M-H]$^-$, 180.0456 (100) [M-H-H_2O]$^-$, 170.0613 (40) [M-H-CO]$^-$, 154.0664 (18) [M-H-CO_2]$^-$, 130.0664 (4) [M-H-C_3O_2]$^-$.

Esculetin (14): ^{1}H NMR: δ = 7.77 (d, ^3J(H, H) = 9.21 Hz, 1H; CH), δ = 6.94 (s, 1H; CH), δ = 6.76 (s, 1H; CH), δ = 6.18 (d, ^3J(H, H) = 9.21 Hz, 1H; CH) ppm; ESI-FTMS: *m/z*: 177.0191 [M-H]$^-$; (-)-ESI-MS^2 (25 eV): *m/z* (%): 177.0193 (14) [M-H]$^-$, 149.0244 (7) [M-H-CO]$^-$, 133.0296 (100) [M-H-CO_2]$^-$, 105.0347 (8) [M-H-C_2O_3]$^-$, 89.0398 (3) [M-H-C_2O_4]$^-$.

4-(3,4-Dihydroxyphenyl)-butan-2-one (16): ^{1}H NMR: δ = 6.65 (d, ^3J(H, H) = 7.89 Hz, 1H; CH), δ = 6.61 (s, 1H; CH), δ = 6.48 (d, ^3J(H, H) = 7.9 Hz, 1H; CH), δ = 3.31 (t, ^3J(H, H) = 6.58 Hz, 2H; CH_2), δ = 2.71 (t, ^3J(H, H) = 6.58 Hz, 2H; CH_2), δ = 2.10 (s, 3H, CH_3) ppm; ESI-FTMS: *m/z*: 179.0714 [M-H]$^-$; (-)-ESI-MS^2 (25 eV): *m/z* (%): 179.0715 (42) [M-H]$^-$, 121.0297 (100) [M-H-C_3H_6O]$^-$, 109.0297 (14) [M-H-C_4H_6O]$^-$.

MS data from cinnamaldehydes and cinnamyl alcohols (chapter II, section 4.2.4)

Table B 1: Key ions in the positive ion Qq-TOF mass spectra of the cinnamaldehydes and cinnamyl alcohols.

| compound	scan mode exact mass [*m/z*]	*m/z* (relative intensity), collisions energy = 25 eV
p-coumaraldehyde (**18**)	148.05	148.0051 (9 %), 147.5565 (11 %), 147.0452 (100 %), 129.0345 (48 %), 119.0500 (68 %), 117.0348 (11 %), 93.0348 (10 %)
caffeyl aldehyde (**19**)	164.05	164.0498 (10 %), 163.0397 (100 %), 162.0325 (6 %), 145.0295 (24 %), 135.0446 (76 %), 134.0370 (45 %), 121.0294 (34 %), 119.0503 (10 %), 117.0347 (6 %), 107.0499 (38 %)
coniferaldehyde (**20**)	178.06	177.0561 (87 %), 163.0399 (14 %), 162.0329 (100 %), 161.0256 (8 %), 134.0380 (42 %)
5-hydroxyconiferaldehyde (**21**)	194.06	194.0539 (24 %), 193.0512 (100 %)
sinapaldehyde (**22**)	208.07	207.0660 (78 %), 192.0428 (95 %), 191.0359 (5 %), 177.0194 (100 %), 149.0247 (34 %), 121.0299 (6 %), 105.0349 (5 %)
p-coumaryl alcohol (**23**)	150.07	149.0620 (27 %), 131.0509 (100 %), 130.0432 (95 %), 103.0561 (37 %), 102.0485 (8 %)
caffeoyl alcohol (**24**)	166.06	165.0568 (87 %), 147.0459 (100 %), 146.0390 (27 %), 130.0435 (20 %), 129.0353 (97 %), 119.0514 (19 %), 101.0402 (68 %)
coniferyl alcohol (**25**)	180.08	179.0724 (40 %), 164.0484 (81 %), 161.0618 (34 %), 147.0450 (6 %), 146.0377 (100 %), 145.0306 (6 %)
5-hydroxyconiferyl alcohol (**26**)	196.07	195.0669 (59 %), 180.0431 (100 %), 179.0361 (9 %), 177.0568 (7 %), 163.0409 (8 %), 162.0328 (65 %), 151.0410 (11 %)
sinapyl alcohol (**27**)	210.09	209.0821 (54 %), 194.0586 (100 %), 191.0723 (15 %), 179.0352 (94 %), 176.0483 (56 %), 161.0249 (75 %), 151.0407 (12 %)
unknown compound (see section 4.2.7)		collisions energy = 10 eV 539.1038 (11 %), 390.126411 (14 %), 389.1228 (54 %), 359.1125 (5 %), 196.0690 (15 %), 195.0661 (100 %), 187.0966 (5 %), 180.0423 (6 %), 177.0551 (10 %)

^{1}H-NMR data and spectra of sinapaldehyde (22) and sinapyl alcohol (27, chapter II, section 4.2.4)

**Sinapaldehyde (22)** ^{1}H NMR (CDCl$_3$): δ = 9.66 (d, ^3J(H, H) = 7.7 Hz, 1H; CH), δ = 7.39 (d, ^3J(H, H) = 15.8 Hz, 1H; CH), δ = 6.82 (s, 2H; CH), δ = 6.61 (dd, ^3J(H, H) = 7.7 Hz, ^3J(H, H) = 15.8 Hz, 1H; CH), δ = 3.95 (s, 6H; CH$_3$) ppm.
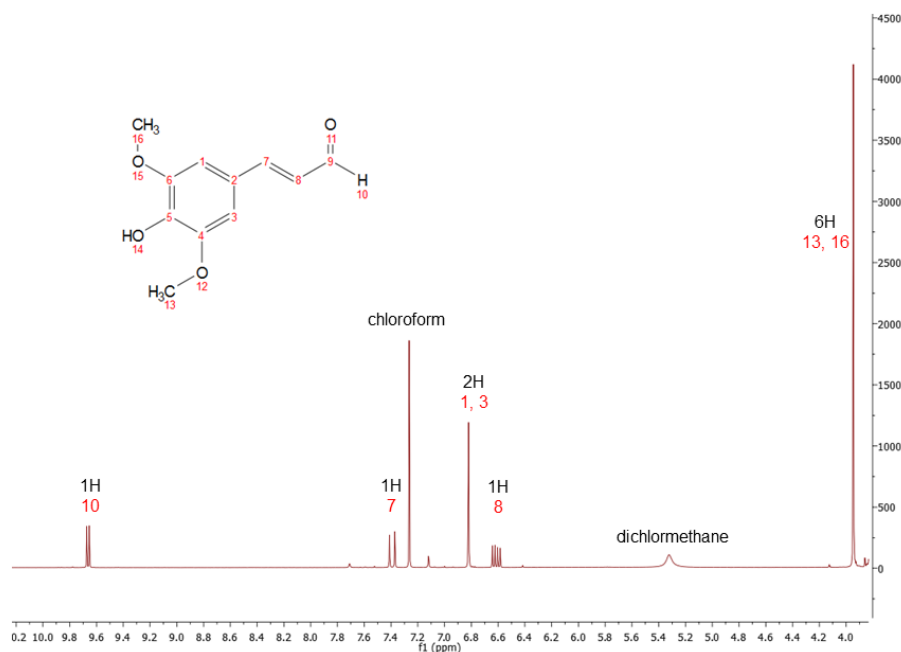
Sinapyl alcohol (27) ^{1}H NMR (CDCl$_3$): δ = 6.64 (s, 2H; CH), δ = 6.53 (d, ^3J(H, H) = 15.8 Hz, 1H; CH); δ = 6.24 (dt, ^3J(H, H) = 5.9 Hz, ^3J(H, H) = 15.8 Hz, 1H; CH), δ = 5.72 (br. s, 1H, OH), δ = 4.31 (d, ^3J(H, H) = 5.9 Hz, 2H; CH), δ = 3.91 (s, 6H; CH$_3$) ppm.
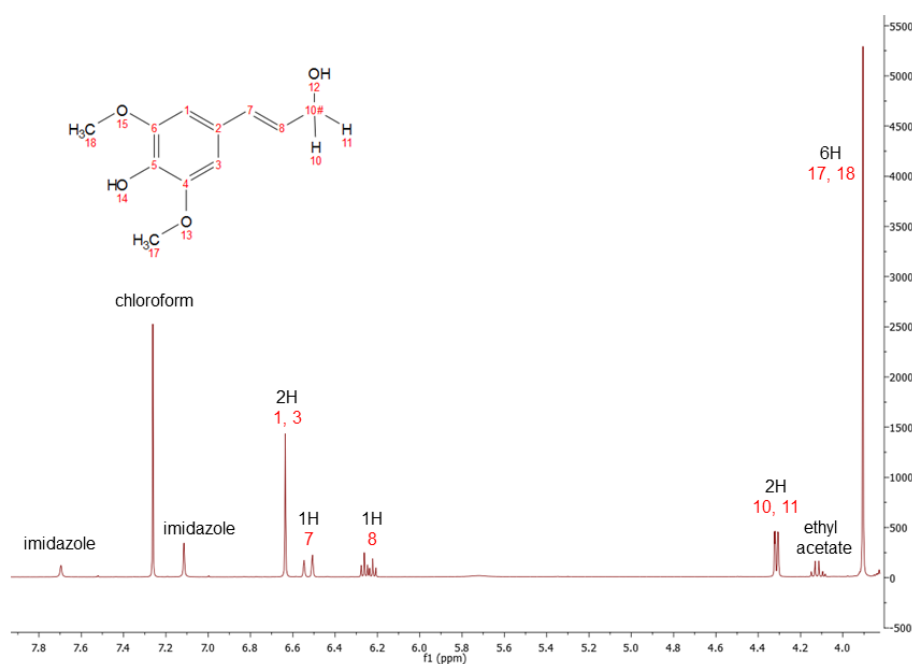
Figure B 1: LC-MS spectrum of the phosphopantetheinylated serine 689 in NiCAR. The purified and desalted peptides were measured after in-solution digestion with combination of two endoproteinases (AspN and Glu-C, chapter II, section 4.2.6).
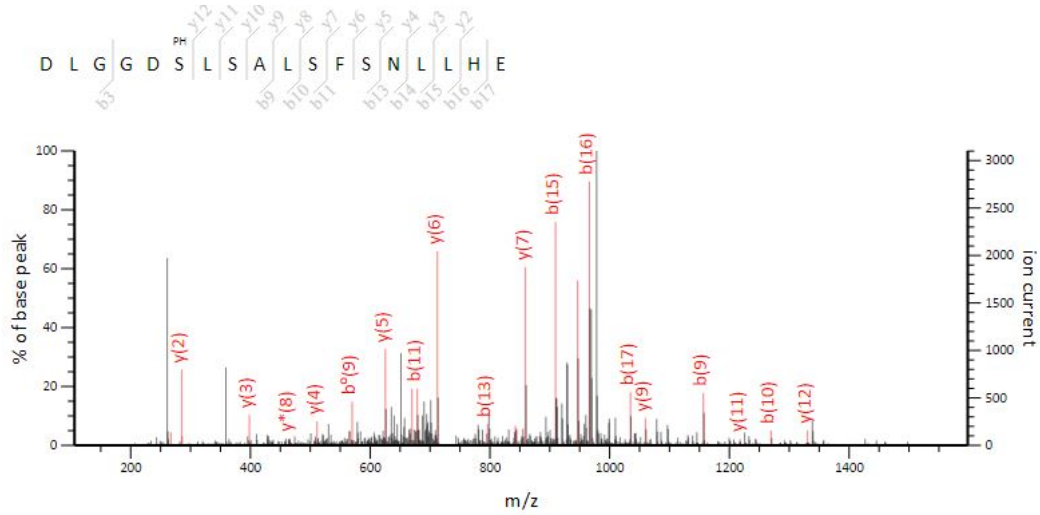
MS data from multi-enzyme cascade (chapter III, section 5.2.1)

Caffeic acid (2): m/z 179.0351 ([M-H]$^-$, calc. for C$_9$H$_7$O$_4^-$ 179.0350); negative ion ESI-MS^2 (25 % normalized collision energy) [m/z, rel. int. (%)]: 179.0352 ([M-H]$^-$, 32), 135.0454 ([M-H-CO$_2$]$^-$, 100, calc. for C$_8$H$_7$O$_2^-$ 134.0452).

Ferulic acid (3): m/z 193.0507 ([M-H]$^-$, calc. for C$_{10}$H$_9$O$_4^-$ 193.0506); negative ion ESI-MS^2 (40 % normalized collision energy) [m/z, rel. int. (%)]: 178.0272 ([M-H-CH$_3$]$^-$, 64, calc. for C$_9$H$_6$O$_4^-$ 178.0272), 149.0609 ([M-H-CO$_2$]$^-$, 100, calc. for C$_9$H$_9$O$_2^-$ 149.0608), 134.0374 ([M-H-CH$_3$-CO$_2$]$^-$, 35, calc. for C$_8$H$_6$O$_2^-$ 134.0373). Rt 11.7 min.

Vanillin (28): m/z 151.0401 ([M-H]$^-$, calc. for C$_8$H$_7$O$_3^-$ 151.0401); negative ion ESI-MS^2 (40 % normalized collision energy) [m/z, rel. int. (%)]: 136.0167 ([M-H-CH$_3$]$^-$, 100, calc. for C$_7$H$_4$O$_3^-$ 136.0166). Rt 10.4 min.

***p*-Hydroxybenzaldehyde:** m/z 121.0296 ([M-H]$^-$, calc. for C$_7$H$_5$O$_2^-$ 121.0295); negative ion ESI-HCD (200 % normalized collision energy) [m/z, rel. int. (%)]: 121.0296 ([M-H]$^-$, 100), 92.0269 ([M-H-CHO]$^-$, 61, calc. for C$_6$H$_4$O$^-$ 92.0268). Rt 8.5 min.

Adenine: m/z 134.0474 ([M-H]$^-$, calc. for C$_5$H$_4$N$_5^-$ 134.0472). Rt 0.9 min.

7.3 Appendix C – Gel electrophoresis after PCR

Figure C 1: Agarose gel electrophoresis of PCR products stained with ethidium bromide and photographed under UV light. A) M: Gene Ruler^{TM} 1 kb DNA Ladder, lane 1: 4HPA3H (1563 bp), lane2: HpaC (513 bp), B) lane 1: Prnf (561 bp). PCR was performed as described in section 2.2.1.
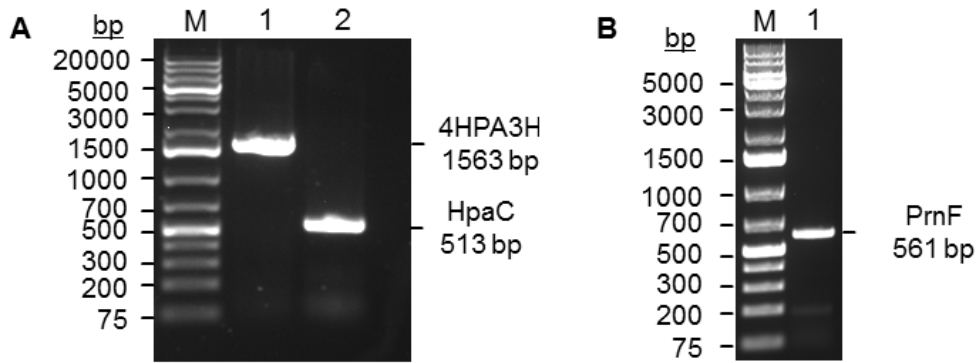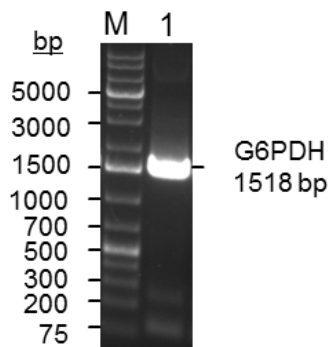

Figure C 2: Agarose gel electrophoresis of PCR products stained with ethidium bromide and photographed under UV light. M: Gene RulerTM 1 kb DNA Ladder, lane 1: ScFDH (1131 bp). PCR was performed as described in section 2.2.1.
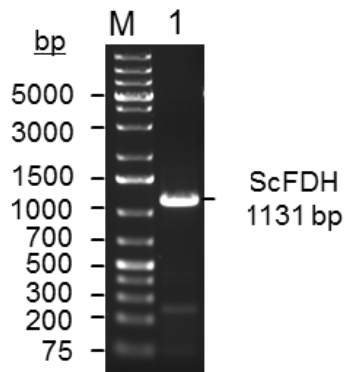
Figure C 3: Agarose gel electrophoresis of PCR products stained with ethidium bromide and photographed under UV light. M: Gene RulerTM 1 kb DNA Ladder, lane 1: G6PDH (1518 bp). PCR was performed as described in section 2.2.1.
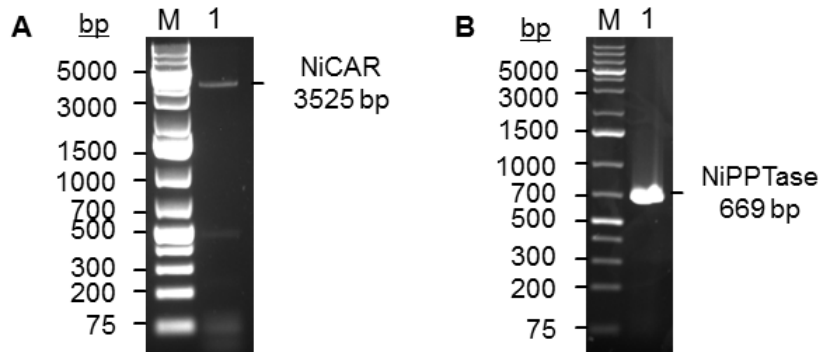
Figure C 4: Agarose gel electrophoresis of PCR products stained with ethidium bromide and photographed under UV light. A) M: Gene Ruler™ 1 kb DNA Ladder, lane 1: NiCAR (3525 bp), B) lane 1: NiPPTase (669 bp). PCR was performed as described in section 2.2.1.
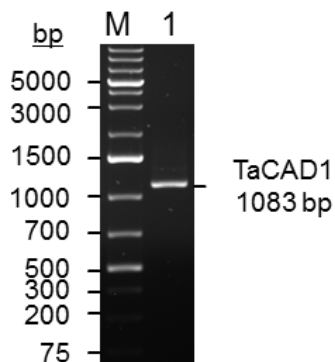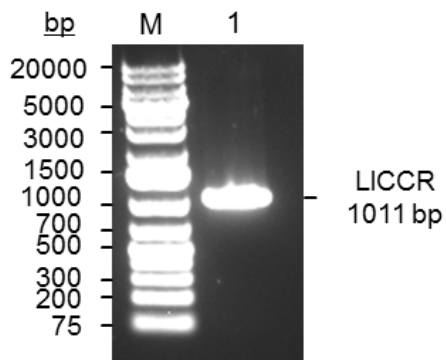

Figure C 5: Agarose gel electrophoresis of PCR products stained with ethidium bromide and photographed under UV light. M: Gene RulerTM 1 kb DNA Ladder, lane 1: TaCAD1 (1083 bp). PCR was performed as described in section 2.2.1.

Figure C 6: Agarose gel electrophoresis of PCR products stained with ethidium bromide and photographed under UV light. M: Gene RulerTM 1 kb DNA Ladder, lane 1: LlCCR (1011 bp). PCR was performed as described in section 2.2.1.
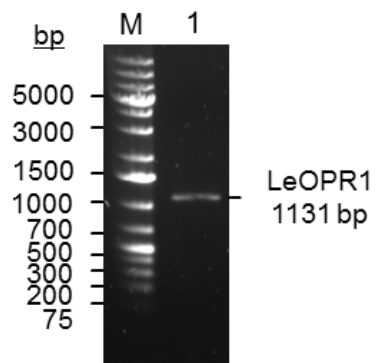
Figure C 7: Agarose gel electrophoresis of PCR products stained with ethidium bromide and photographed under UV light. M: Gene RulerTM 1 kb DNA Ladder, lane 1: LeOPR1 (11131 bp). PCR was performed as described in section 2.2.1.

7.4 Appendix D – SDS-PAGE gels of purified enzymes

Figure D 1: SDS-PAGE of cells expressing 4HPA3H (lane 1), the variants Y301F (lane 2), S462A (lane 3), M293P (lane 4), I157V (lane 5), Y301L (lane 6), Y301I (lane 7), Y301F/S462A (lane 8), Y301L/S462A (lane 9), Y301I/S462A (lane 10), Y301F/I157V (lane 11) and I157V/S462A (lane 12), or harboring the empty vector pET-28a(+) (lane 13). Lysates from equal amounts of cells were applied to polyacrylamide gels with 10 % (w/v) crosslinking. M: molecular markers.
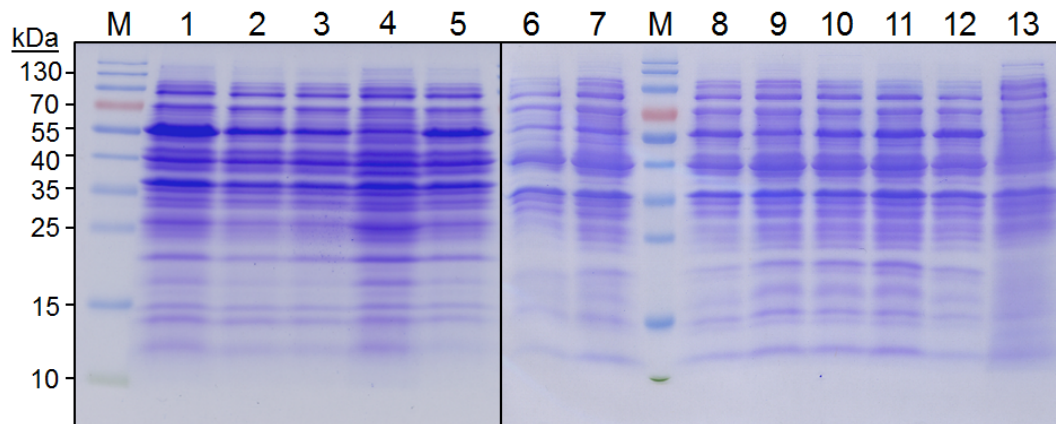
Figure D 2: SDS-PAGE of purified enzymes used in *in vitro* hydroxylation (section 3.2.4). Equal amounts of protein (5 µg) were used for separation in polyacrylamide gels with (10 % (w/v) crosslinking as specified in section 2.6.2. Lane 1: wild-type 4HPA3H, lane 2: 4HPA3H-Y301I, lane 3: 4HPA3H-Y301F/S462A, lane 4: PrnF, M: molecular marker.
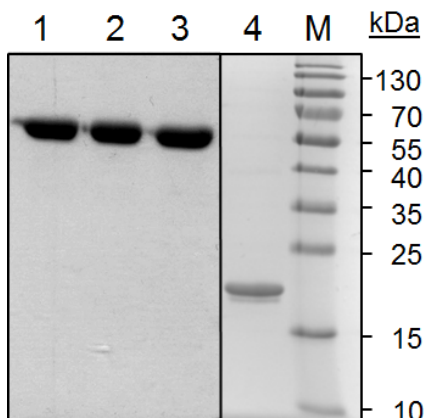
Figure D 3: SDS-PAGE of all purified enzyme used in chapter II. Purified protein (2 µg) was applied to gels with 14 % (w/v) crosslinking. M: molecular marker, lane 1: 4HPA3H, lane 2: PrnF, lane 3: FHD, lane 4: PFOMT, lane 5: 4HPA3H-Y301I, lane 6: NiCAR, lane 7: NiPPT, lane 8: TaCAD1, lane 9: G6PDH.
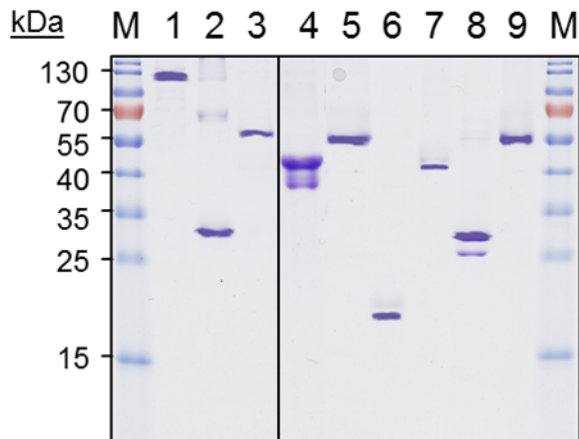
Figure D 4: SDS-PAGE of the enzymes used in section 5.2.1. Purified protein (5 µg) was applied to gels with 10 % (w/v) crosslinking. Lanes show the electropherograms of the molecular markers (M, apparent molecular masses indicated at left side), 4HPA3H (lane 1), PrnF (lane 2), ScFDH (lane 3), SAMS-I317V (lane 4), PFOMT (lane 5), SAHH (lane 6), SAHN (lane 7), 4CL2 (lane 8) and FCoAHL (lane 9).
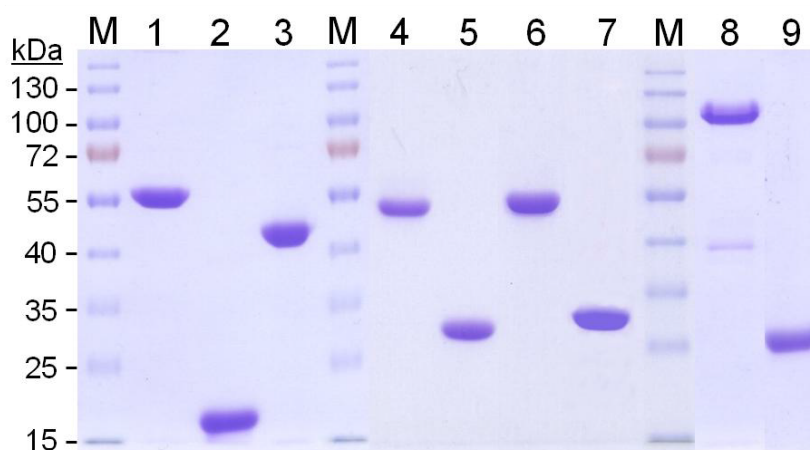
Figure D 5: SDS-PAGE of enzymes used in section 5.2.2. Purified protein (5 µg) was applied to gels with 14 % (w/v) crosslinking. M: molecular marker, lane 1: 4HPA3H-Y301F/S462A, lane 2: PrnF, lane 3: ScFDH.
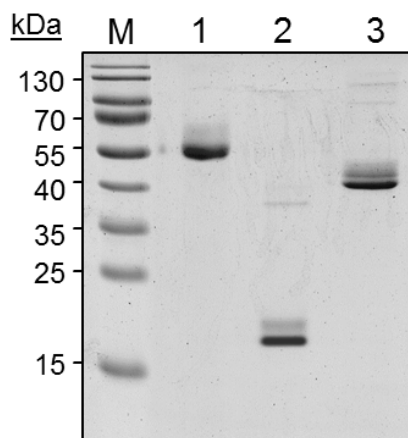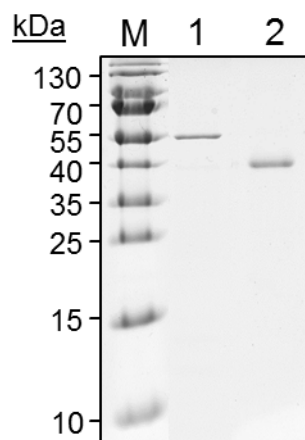

Figure D 6: SDS-PAGE of enzymes used in section 5.3. Purified protein (2 µg) was applied to gels with 14 % (w/v) crosslinking. M: molecular marker, lane 1: G6PDH, lane 2: LeOPR1.

7.5 Appendix E – Activity assays

Figure E 1: Catechol complex formation assay for determination of 4HPA3H activity (chapter I). (A) Standard curve for the determination of caffeic acid (**2**). Mixtures of caffeic acid (**2**) and *p*-coumaric acid (**1**) were incubated with $FeCl_3$ and analyzed at 595 nm as described in section 2.5.1.1. (B) Time course of the oxidation of *p*-coumaric acid (**1**) by wild-type 4HPA3H. The enzymatic reaction was performed under conditions given in chapter I, and product caffeic acid (**2**) was detected as Fe^{3+} complex as described in section 2.5.1.1. Data are the means ± standard deviations obtained from three measurements.
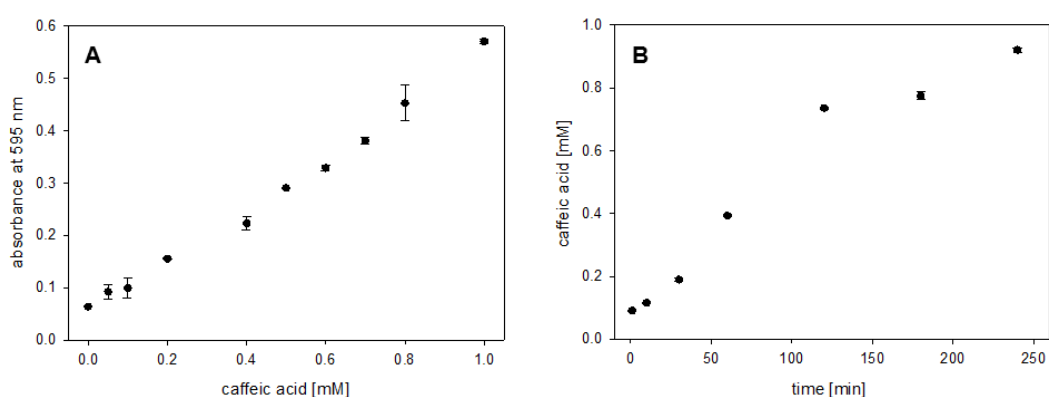
Figure E 2: The absorption spectrums of substrates and products for the LeOPR1-reaction at different wavelengths (chapter III). A) Absorbance of dehydrozingerone (**30**, depicted in red) and zingerone (**31**, depicted in yellow) at wavelengths between 300 – 550 nm. B) Absorbance of dehydrorheosmin (**29**, depicted in red) and rheosmin (**15**, depicted in yellow) at wavelengths between 300 – 550 nm.
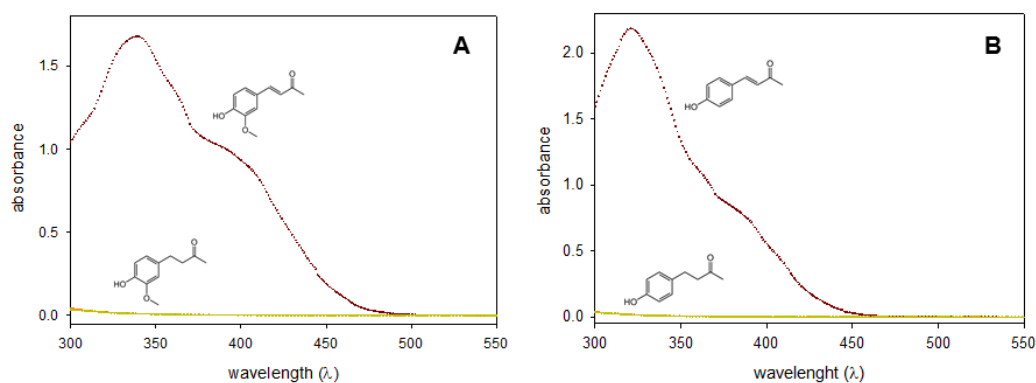
Figure E 3: Formation of cinnamic acids after conversion of *p*-coumaric acid (1) with different enzyme combinations (chapter II). For the abbreviations A to D see Table 4.1 in section 4.2.3.
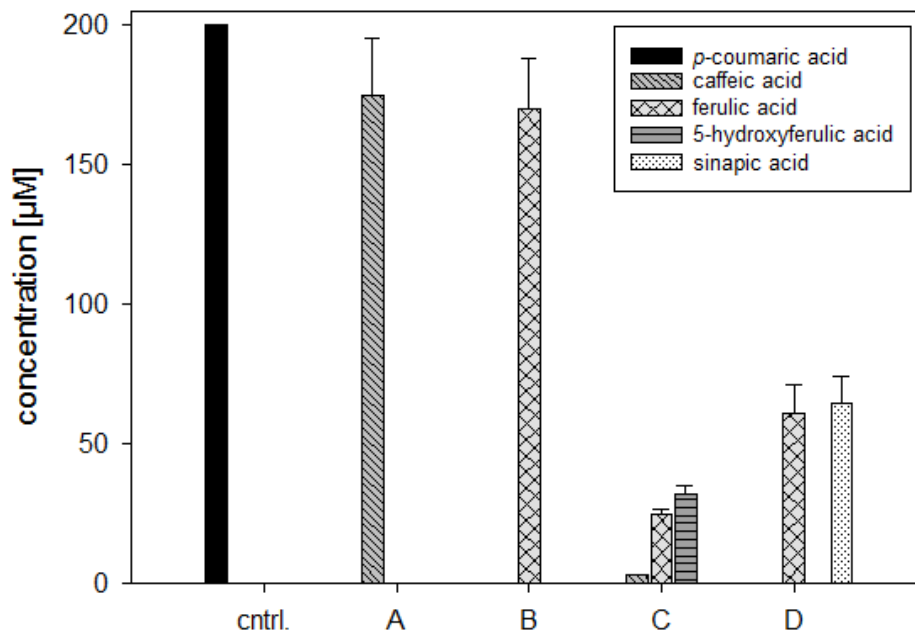
7.6   Appendix F – HPL chromatograms

Figure F 1: HPL chromatogram of (+)- and (-)-naringenin separated by chiral reversed-phase high-performance liquid chromatography (chapter III, section 5.2.4). A) HPLC-CD chromatogram of separated (+)- and (-)-naringenin. B) UV-Vis trace of separated (+)-and (-)-naringenin at a wavelength of 280 nm.
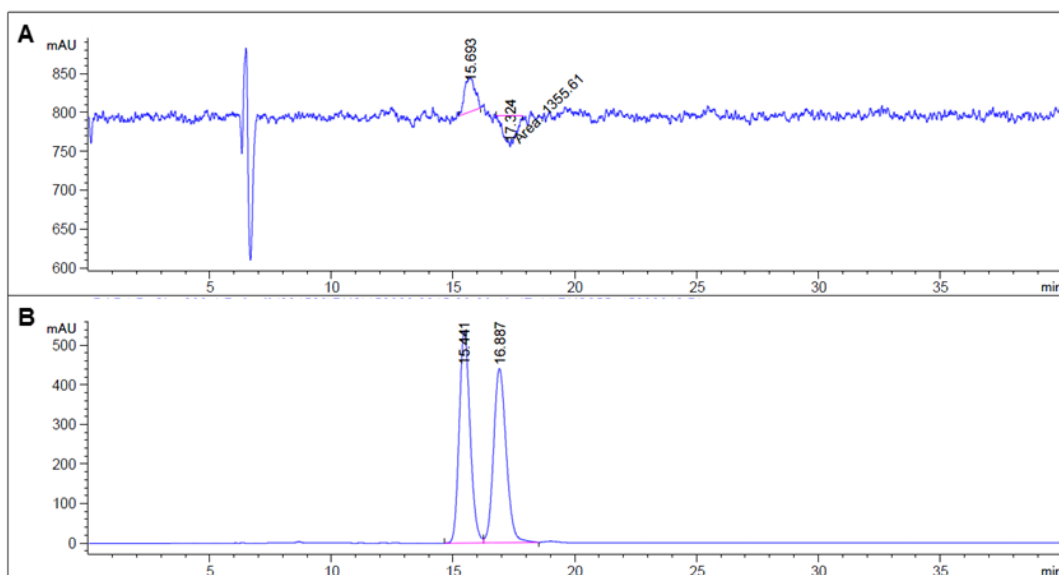
Figure F 2: HPL chromatogram of (+)- and (-)-eriodictyol separated by chiral reversed-phase high-performance liquid chromatography (chapter III, section 5.2.4). A) HPLC-CD chromatogram of separated (+)- and (-)-eriodictyol. B) UV-Vis trace of separated (+)- and (-)-eriodictyol at wavelength of 280 nm.
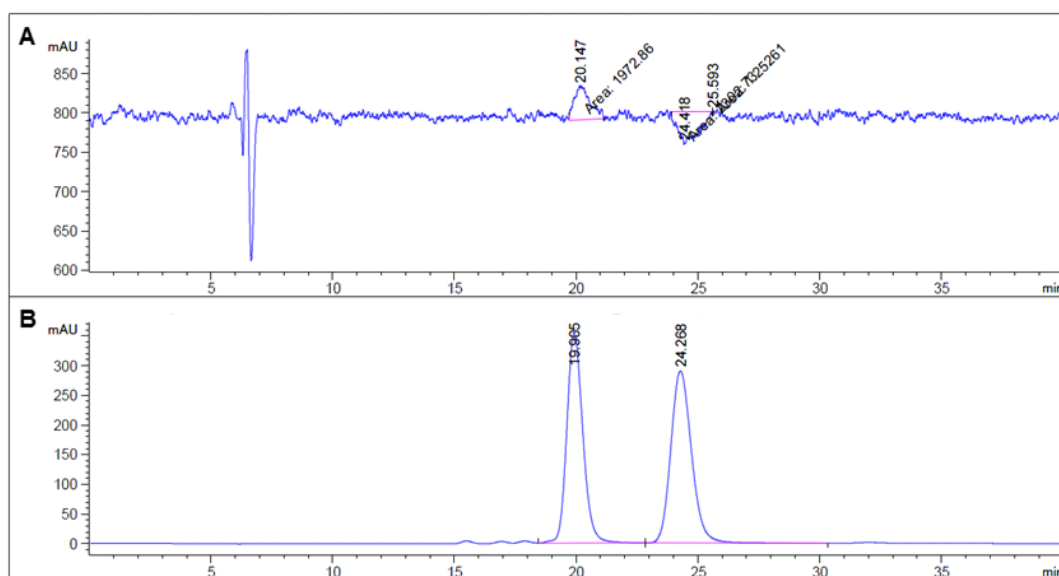
Figure F 3: HPL chromatograms of reduction reactions (chapter II, section 4.2.6). Conversion of cinnamic acid derivatives with NiCAR/NiPPTase to the corresponding cinnamaldehyde or with NiCAR/PPTase + TaCAD1 to the corresponding cinnamyl alcohol: A) ferulic acid (**3**) to coniferaldehyde (**20**) or coniferyl alcohol (**25**); B) sinapic acid (**17**) to sinapaldehyde and sinapyl alcohol (**27**); C) p-coumaric acid (**1**) to p-coumaraldehyde or p-coumaryl alcohol (**23**); D) caffeic acid (**2**) to caffeyl aldehyde or caffeoyl alcohol (**24**); E) 5-hydroxyferulic acid (**4**) to 5-hydroxyconiferaldehyde or 5-hydroxyconiferyl alcohol.
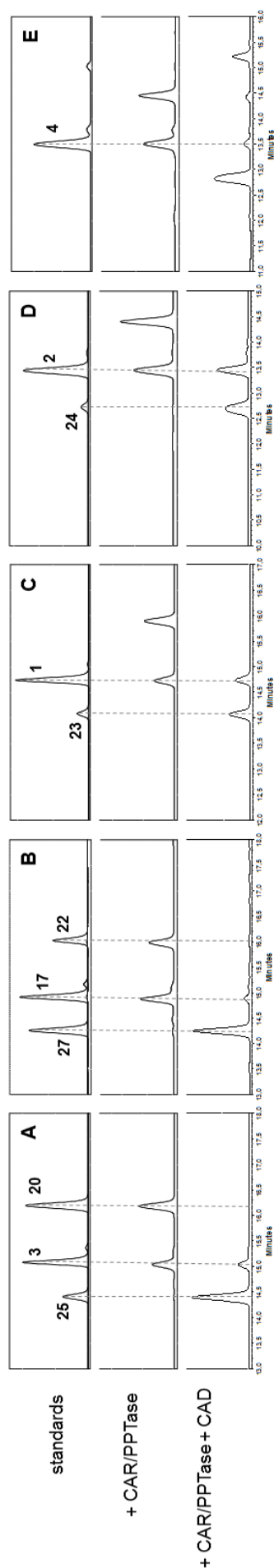
7.7 Appendix G – Calibrations

Calibrations for specific aromatic decoration (chapter II, section 4.2.3)

Figure G 1: Calibration curve for *p*-coumaric acid (**1**) in the concentration range from 70 to 280 µM for tests during specific aromatic decoration assay**.**

calibration *p*-coumaric acid [70 - 280 µM]
$y = 67612 \cdot x - 1857355$
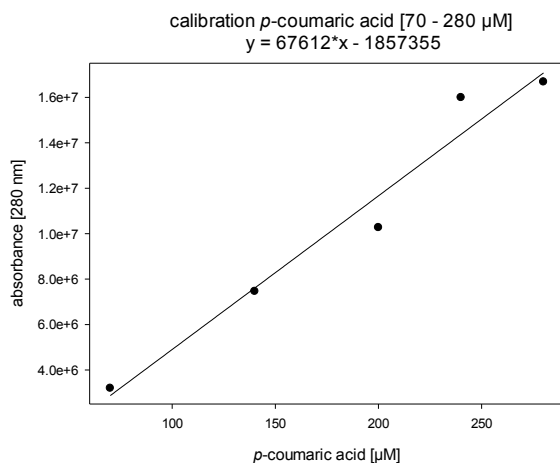
Figure G 2: Calibration curve for caffeic acid (**2**) in the concentration range from 70 to 280 µM for tests during specific aromatic decoration assay.

calibration caffeic acid [70 - 280 µM]
$y = 27601 \cdot x + 615187$
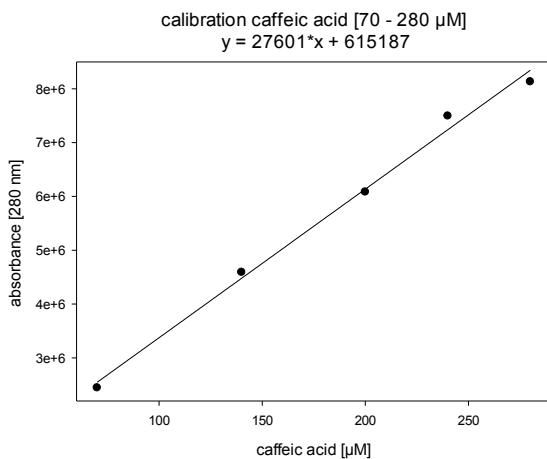
Figure G 3: Calibration curve for ferulic acid (**3**) in the concentration range from 70 to 280 µM for tests during specific aromatic decoration assay.

calibration ferulic acid [70 - 280 µM]
y = 31858*x - 745560
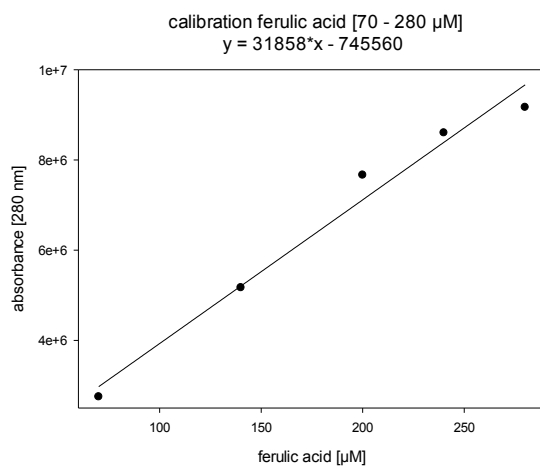
Figure G 4: Calibration curve for 5-hydroxyferulic acid (**4**) in the concentration range from 70 to 280 µM for tests during specific aromatic decoration assay.

calibration 5-hydroxyferulic acid [70 -280 µM]
y = 20564*x - 345569

Figure G 5: Calibration curve for sinapic acid (**17**) in the concentration range from 70 to 280 µM for tests during specific aromatic decoration assay.

calibration sinapic acid [70 - 280 µM]
y = 19549*x + 802892
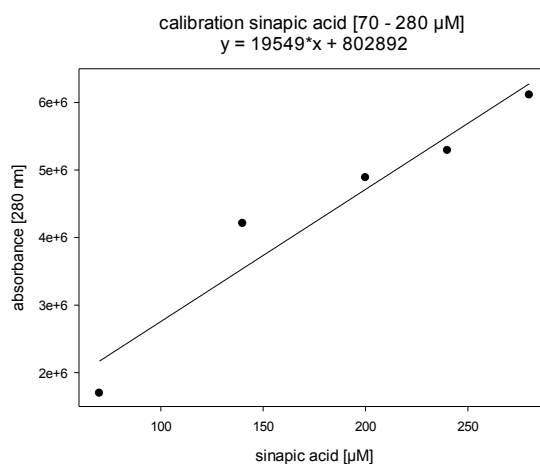
Calibrations for side chain reduction (chapter II, section 4.2.4)

Figure G 6: Calibration curve for coniferaldehyde (**20**) in the concentration range from 70 to 280 µM for tests during side chain reduction assay.
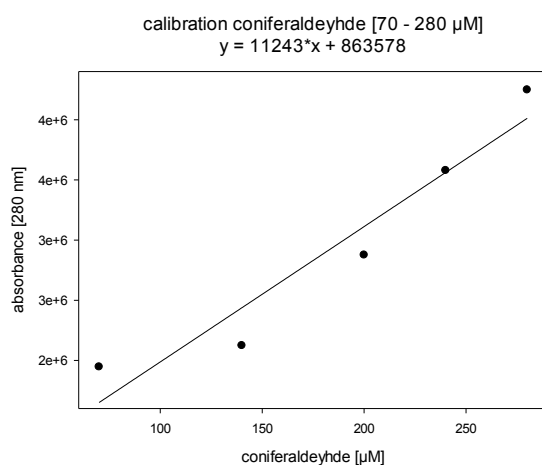
Figure G 7: Calibration curve for sinapaldehyde (**22**) in the concentration range from 70 to 280 µM for tests during side chain reduction assay.
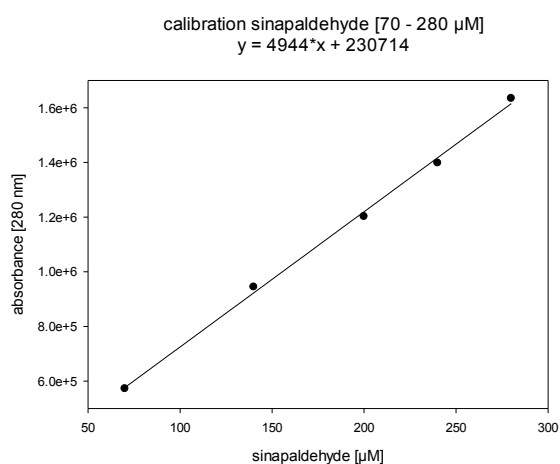
Figure G 8: Calibration curve for *p*-coumaryl alcohol (**23**) in the concentration range from 70 to 280 µM for tests during side chain reduction assay.
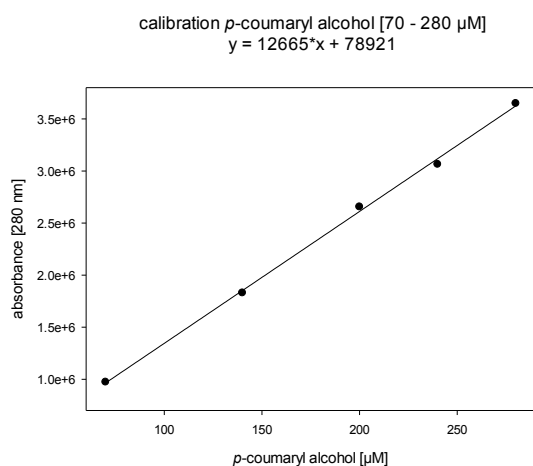
Figure G 9: Calibration curve for caffeoyl alcohol (**24**) in the concentration range from 70 to 280 µM for tests during side chain reduction assay.

calibration caffeoyl alcohol [70 - 280 µM]
y = 6708*x + 417584
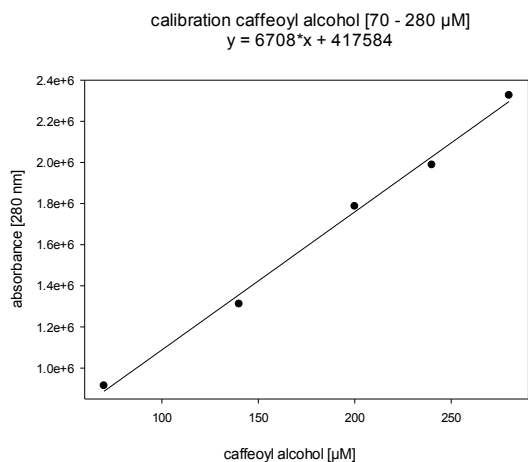
Figure G 10: Calibration curve for coniferyl alcohol (**25**) in the concentration range from 70 to 280 µM for tests during side chain reduction assay.

calibration coniferyl alcohol [70 - 280 µM]
y = 18621*x + 101830
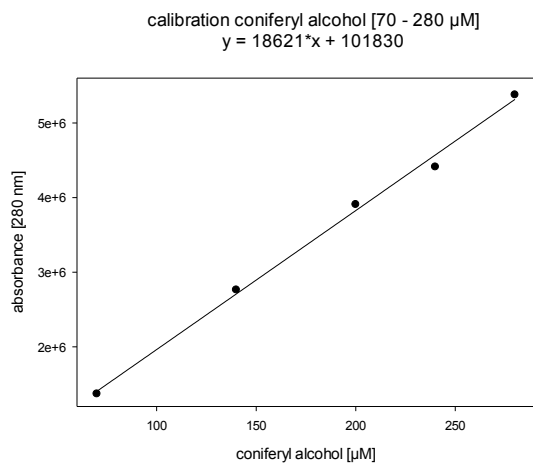
Figure G 11: Calibration curve for sinapyl alcohol (**26**) in the concentration range from 70 to 280 µM for tests during side chain reduction assay.

calibration sinapyl alcohol [70 - 280 µM]
y = 20019*x + 185100
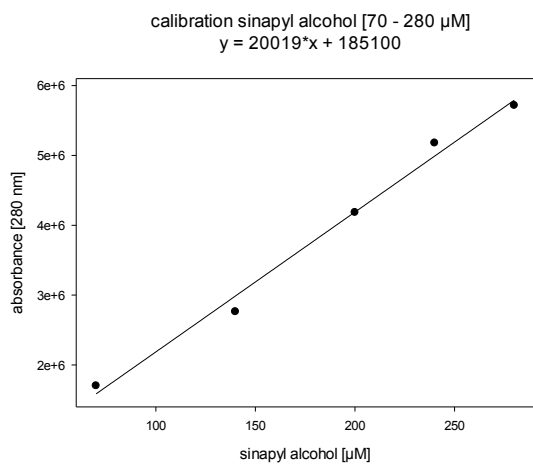
Calibrations for multi-enzyme cascade (chapter III)

Figure G 12: Calibration curve for *p*-coumaric acid (**1**) in the concentration range from 70 to 280 μM for tests during the multi-enzyme cascade (section 5.2.3).
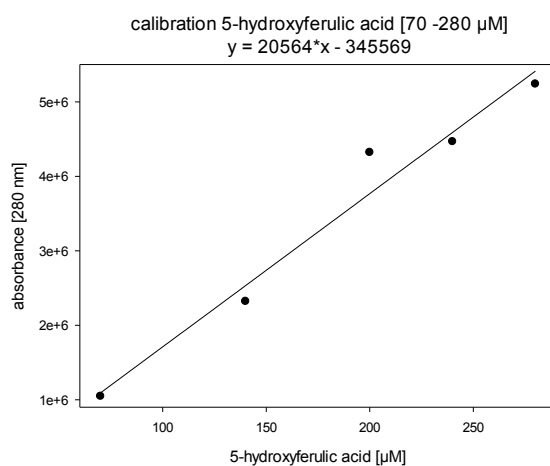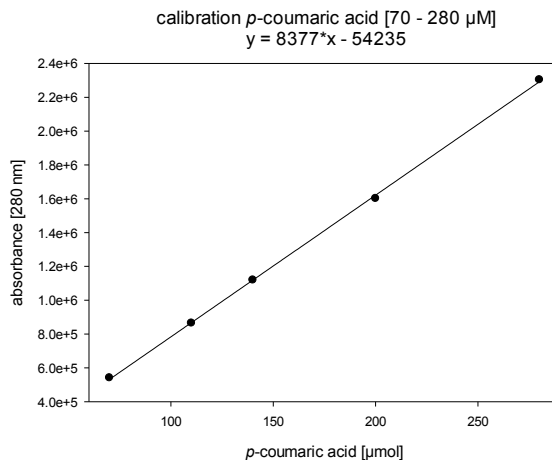

Figure G 13: Calibration curve for caffeic acid (**2**) in the concentration range from 70 to 280 μM for tests during the multi-enzyme cascade (section 5.2.3).
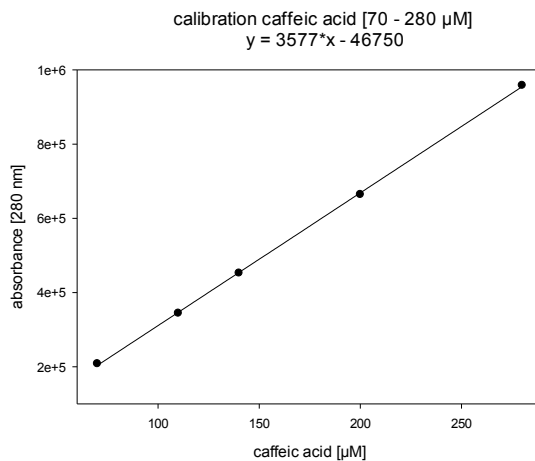

Figure G 14: Calibration curve for ferulic acid (**3**) in the concentration range from 70 to 280 μM for tests during the multi-enzyme cascade (section 5.2.3).
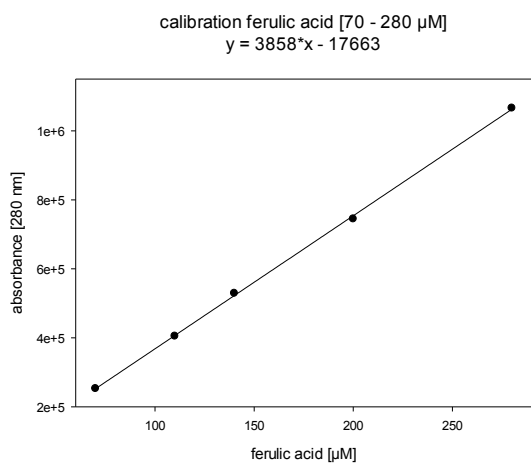
Figure G 15: Calibration curve for vanillin (**28**) in the concentration range from 70 to 280 µM for tests during the multi-enzyme cascade (section 5.2.3).
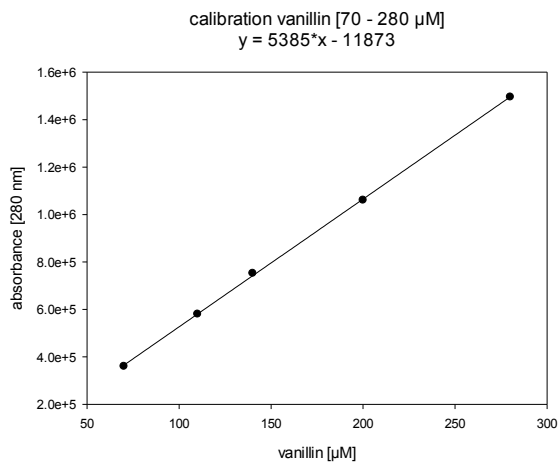
Figure G 16: Calibration curve for *p*-hydroxybenzaldehyde in the concentration range from 70 to 280 µM for tests during the multi-enzyme cascade (section 5.2.3).
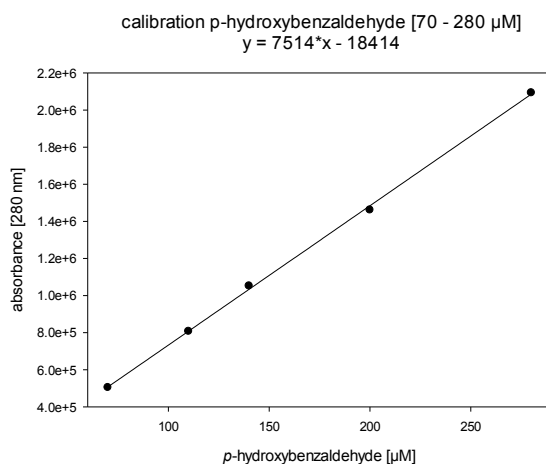
Figure G 17: Calibration curve for adenine in the concentration range from 70 to 280 µM for tests during the multi-enzyme cascade (section 5.2.3).
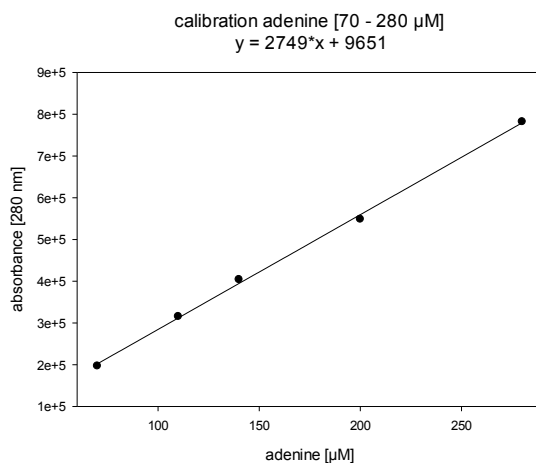
Figure G 18: Calibration curve for naringenin (**7**) in the concentration range from 70 to 280 µM for tests during the multi-enzyme cascade (section 5.2.4).

calibration naringenin [70 - 280 µM]
y = 47034*x - 301975
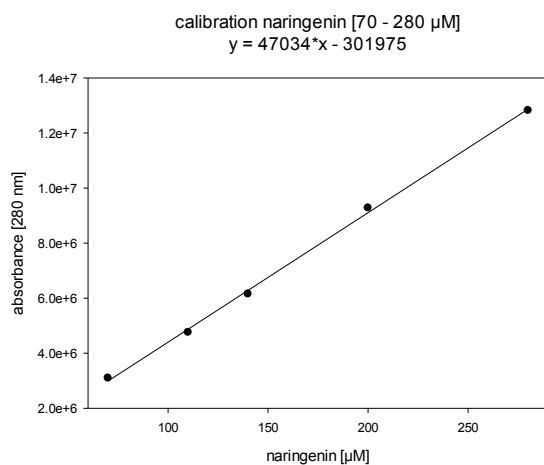
Figure G 19: Calibration curve for eriodictyol (**8**) in the concentration range from 70 to 280 µM for tests during the multi-enzyme cascade (see section 5.2.4).

calibration eriodictyol [70 - 280 µM]
y = 49683*x - 114360
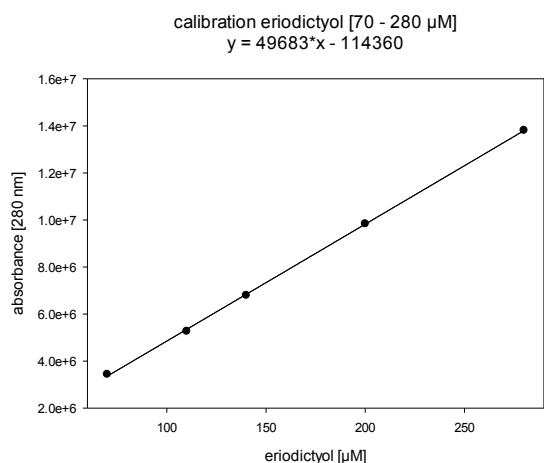
Figure G 20: Calibration curve for dehydrorheosmin (**29**) in the concentration range from 70 to 280 µM for tests during the multi-enzyme cascade (section 5.3.2).

calibration 4-(4-hydroxyphenyl)-3-buten-2-one [70 - 280 µM]
y = 34592*x + 96953
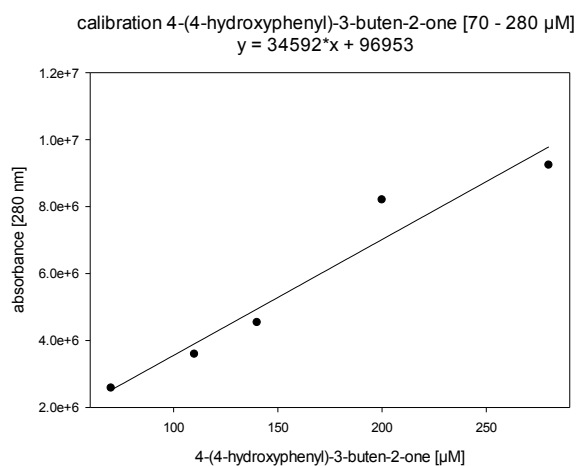
Figure G 21: Calibration curve for rheosmin (**15**) in the concentration range from 70 to 280 µM for tests during the multi-enzyme cascade (see section 5.3.2).

calibration rheosmin [70 - 280 µM]
$$y = 11638*x + 101597$$
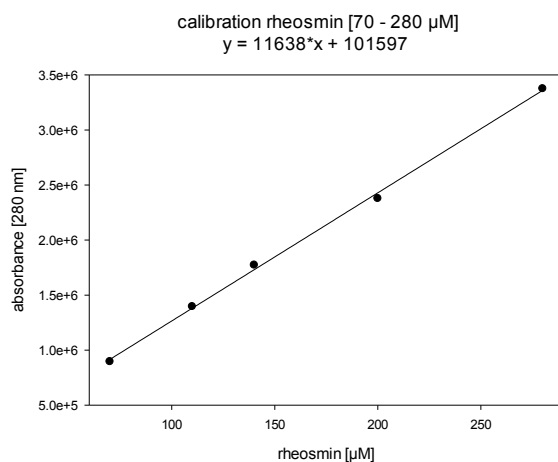
Figure G 22: Calibration curve for 4-(4-hydroxy-3-methoxyphenyl)-3-buten2-one (**30**) in the concentration range from 70 to 280 µM for tests during the multi-enzyme cascade (see section 5.3.2).

calibration 4-(4-hydroxy-3-methoxyphenyl)-3-buten-2-one [70 - 280 µM]
$$y = 17578*x + 2995$$
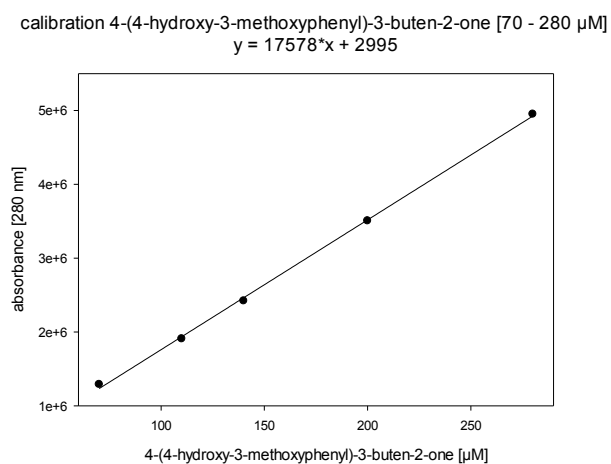
Figure G 23: Calibration curve for zingerone (**31**) in the concentration range from 70 to 280 µM for tests during the multi-enzyme cascade (see section 5.3.2).

calibration zingerone [70 - 280 µM]
$$y = 10031*x + 430319$$
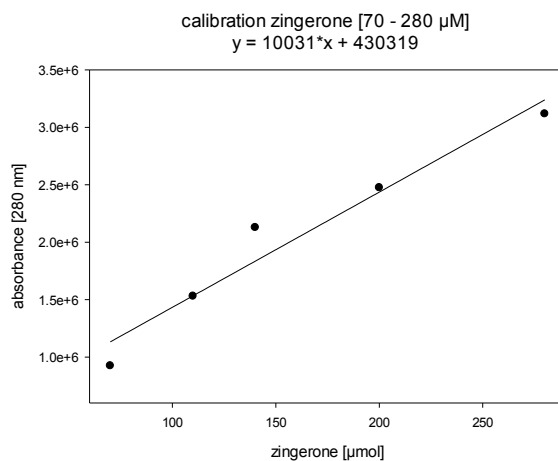
7.8 Appendix H – Production of recombinant enzymes for multi-enzyme cascade (chapter III, section 5.2.3)

variant (NCBI accession)	expression conditions	expression yield [mg l^{-1}]	specific activity [mU mg^{-1}]
SAMS-I317V (NP_390933.1 (wild-type))	22 °C, LB media, 50 µM IPTG, 2 h	102	126.4 ± 2.7
SAHN (NP_414701.1)	37 °C, LB media, 1 mM IPTG, 4 h	21	5.4 ± 0.1
SAHH (EEW50513.1)	37 °C, LB media, 1 mM IPTG, 4 h	70	1.6 ± 0.1
At4CL2 (Q9S725.2)	25 °C, overnight, auto-induced media	35	43.1 ± 0.6
FCoAHL (AAZ23790.1)	25 °C, overnight, auto-induced media	202	2170 ± 250

8   References

(1)   Olivoto, T.; Nardino, M.; Carvalho, I. R.; Follmann, D. N.; Szareski, V. J.; Ferrari, M.; Junior De Pelegrin, A.; Queiróz De Souza, V. *African J. Agric. Rsearch* **2017**, *12* (2), 71–84.

(2)   Croteau, R.; Kutchan, T. M.; Lewis, N. G. *Biochem. Mol. Biol. plants* **2000**, *24*, 1250–1319.

(3)   Korkina, L. G. *Cell. Mol. Biol.* **2007**, *53* (1), 15–25.

(4)   Marienhagen, J.; Bott, M. *J. Biotechnol.* **2013**, *163*, 166–178.

(5)   Verpoorte, R.; Contin, A.; Memelink, J. *Phytochem. Rev.* **2002**, *1*, 13–25.

(6)   Ncube, B.; Van Staden, J. *Molecules* **2015**, *20*, 12698–12731.

(7)   Staniek, A.; Bouwmeester, H.; Fraser, P. D.; Kayser, O.; Martens, S.; Tissier, A.; van der Krol, S.; Wessjohann, L.; Warzecha, H. *Biotechnol. J.* **2014**, *9* (3), 326–336.

(8)   Li, J. W.-H.; Vederas, J. C. *Science (80-.).* **2009**, *325*, 161–165.

(9)   *Biotechnology: Secondary Metabolites*, 2nd ed.; Ramawat, K. G., Merillon, J. M., Eds.; CRC Press, **2007**.

(10)   Mabry, T. J.; Ulubelen, A. *J. Agric. Food Chem* **1980**, *28* (2), 188–196.

(11)   Douglas, C. J. *Trends Plant Sci.* **1996**, *1* (6), 171–178.

(12)   Trantas, E. A.; Koffas, M. A. G.; Xu, P.; Ververidis, F.; Paolocci, F.; Casati, P.; Marienhagen, J. *Front. Plant Sci.* **2015**, *6*, 7.

(13)   Hannemann, F.; Bichet, A.; Ewen, K. M.; Bernhardt, R. *Biochim. Biophys. Acta 1770* **2006**, 330–344.

(14)   Neufeld, K.; Marienhagen, J.; Schwaneberg, U.; Pietruszka, J. *Green Chem.* **2013**, *15*, 2408–2421.

(15)   Ferrer, J. L.; Austin, M. B.; Stewart, C.; Noel, J. P. *Plant Physiol. Biochem.* **2008**, *46*, 356–370.

(16)   Vogt, T. *Mol. Plant* **2010**, *3* (1), 2–20.

(17)   Bourgaud, F.; Hehn, A.; Larbat, R.; Doerper, S.; Gontier, E.; Kellner, S.; Matern, U. *Phytochemistry Reviews*. **2006**.

(18)   Lin, Y.; Yan, Y. *Biotechnol. Bioeng* **2014**, *111*, 1895–1899.

(19)   Lin, Y.; Sun, X.; Yuan, Q.; Yan, Y. *Metab. Eng.* **2013**, *18*, 69–77.

(20)   Dixon, R. A.; Achnine, L.; Kota, P.; Liu, C. J.; Reddy, M. S. S.; Wang, L. *Mol. Plant Pathol.* **2002**, *3* (5), 371–390.

(21)   https://www.epa.gov/greenchemistry.

(22)   Sheldon, R. A. *Green Chem.* **2016**, *18*, 3180–3183.

(23)   Hatti-Kaul, R.; Törnvall, U.; Gustafsson, L.; Börjesson, P. *Trends Biotechnol.* **2007**, *25* (3), 119–124.

(24)   https://echa.europa.eu/de/regulations/reach.

(25)   Dunn, P. J. *Chem. Soc. Rev.* **2012**, *41* (41), 1452–1461.

(26)   Anastas, P. T.; Warner John. C., *Green Chemistry: Theory and Practice*; Oxford University Press: New York, **1998**; pp 29–56.

(27)   Illanes, A. In *Enzyme Biocatalysis: Principles and Applications*; Springer, **2008**.

(28)   Cioc, R. C.; Ruijter, E.; Orru, R. V. A. *Green Chem.* **2014**, *16*, 2958–2975.

(29)   Kalim Akhtar, M.; Turner, N. J.; Jones, P. R.; Haselkorn, R. *PNAS* **2013**, *110* (1), 87–92.

(30)   Faber, K. In *Biotransformations in Organic Chemistry*; Springer-Verlag: Heidelberg, **2011**; pp 3–9.

(31)   Schmid, A.; Dordick, J. S.; Hauer, B.; Klener, A.; Wubbolts, M.; Witholt, B. *Nature* **2001**, *409*, 258–268.

(32)   Gosset, G. *Curr. Opin. Biotechnol.* **2009**, *20*, 651–658.

(33)   Zhang, Y. H. P. *Biotechnol. Adv.* **2011**, *29*, 715–725.

(34)   Kaur, B.; Chakraborty, D. *Appl. Biochem. Biotechnol.* **2013**, *169*, 1353–1372.

(35)   Nielsen, J.; Keasling, J. D. *Cell* **2016**, *164* (6), 1185–1197.

(36)   Wessjohann, L. A.; Keim, J.; Weigel, B.; Dippe, M. *Curr. Opin. Chem. Biol.* **2013**, *17* (2), 229–235.

(37)   Chen, X.; Zhou, L.; Tian, K.; Kumar, A.; Singh, S.; Prior, B. A.; Wang, Z. *Biotechnol. Adv.* **2013**, *31* (8), 1200–1223.

(38)   Woolston, B. M.; Edgar, S.; Stephanopoulos, G. *Annu. Rev. Chem. Biomol. Eng.* **2013**, *4*, 259–288.

(39)   Wang, S.; Zhang, S.; Xiao, A.; Rasmussen, M.; Skidmore, C.; Zhan, J. *Metab. Eng.* **2015**, *29*, 153–159.

(40)   Torres Pazmino, D. E.; Winkler, M.; Glieder, A.; Fraaije, M. W. *J. Biotechnol.* **2010**, *146*, 9–24.

(41)   Ajikumar, P. K.; Xiao, W.-H.; Tyo, K. E. J.; Wang, Y.; Simeon, F.; Leonard, E.; Mucha, O.; Phon, T. H.; Pfeifer, B.; Stephanopoulos, G. *Science (80-.).* **2010**, *330*, 70–74.

(42)   Xu, P.; Li, L.; Zhang, F.; Stephanopoulos, G.; Koffas, M. *PNAS* **2014**, *111* (31), 11299–11304.

(43)   Ishii, T.; Araki, M. *Plant Cell Rep.* **2016**, *35* (7), 1507–1518.

(44)   Kwon, S. J.; Mora-Pale, M.; Lee, M. Y.; Dordick, J. S. *Curr. Opin. Chem. Biol.* **2012**, *16* (1–2), 186–195.

(45)   Mora-Pale, M.; Sanchez-Rodriguez, S. P.; Linhatrdt, R. J.; Dordick, J. S.; Koffas, M. A. G. *Plant Sci.* **2013**, *210*, 10–24.

(46) Van Oort, M. *Enzymes in food technology*, 2nd ed.; Whitehurst, R. J., Van Oort, M., Eds.; John Wiley & Sons, **2010**.

(47) Bolivar, J. M.; Eisl, I.; Nidetzky, B. *Catal. Today* **2016**.

(48) Jones, J. A.; Koffas, M. A. G. In *Synthetic Biology and Metabolic Engineering in Plants and Microbes*; Sarah E. O´Connor, Ed.; Academic Press, **2016**; pp 179–193.

(49) Dippe, M.; Brandt, W.; Rost, H.; Porzel, A.; Schmidt, J.; Wessjohann, L. A. *Chem. Commun.* **2015**, No. 51, 2637–2640.

(50) Studier, F. W. *Protein Expr Purif* **2005**, *41* (1), 207–234.

(51) Engler, C.; Gruetzner, R.; Kandzia, R.; Marillonnet, S. *PLoS One* **2009**, *4* (5), e5553.

(52) Inoue, H.; Nojima, H.; Okayama, H. *Gene* **1990**, *96* (1), 23–28.

(53) AAT Bioquest® Product Technical Information Sheet. *Photometric Detection of Nicotinamide Adenine Dinucleotides*; **2012**.

(54) Laemmli, U. K. *Nature* **1970**, *227*, 680–685.

(55) Majovsky, P.; Naumann, C.; Lee, C. W.; Lassowskat, I.; Trujillo, M.; Dissmeyer, N.; Hoehenwarter, W. *J. Proteome Res.* **2014**, *13* (4246–4258).

(56) Kelley, L. A.; Sternberg, M. J. E. *Nat. Protoc.* **2009**, *4* (3), 363–371.

(57) Ullrich, R.; Hofrichter, M. *Cell. Mol. Life Sci.* **2007**, *64*, 271–293.

(58) Yadav, G. D.; Asthana, N. S. *Appl. Catal. A Gen.* **2003**, *244* (2), 341–357.

(59) Woodley, J. M. *Trends Biotechnol.* **2008**, *26* (6), 321–327.

(60) Nolan, L. C.; O'Connor, K. E. *Biotechnol. Lett.* **2008**, *30*, 1879–1891.

(61) Urlacher, V. B.; Girhard, M. *Trends Biotechnol.* **2012**, *30* (1), 26–36.

(62) Fasan, R. *ACS Catal.* **2012**, *2*, 647–666.

(63) Bogazkaya, A. M.; Von Bühler, C. J.; Kriening, S.; Busch, A.; Seifert, A.; Pleiss, J.; Laschat, S.; Urlacher, V. B. *Beilstein J. Org. Chem.* **2014**, *10*, 1347–1353.

(64) Sucharitakul, J.; Tinikul, R.; Chaiyen, P. *Arch. Biochem. Biophys.* **2014**, *555–556*, 33–46.

(65) Huijbers, M. M. E.; Montersino, S.; Westphal, A. H.; Tischler, D.; Van Berkel, W. J. H. *Arch. Biochem. Biophys.* **2014**, *544*, 2–17.

(66) Ceccoli, R. D.; Bianchi, D. A.; Rial, D. V. *Front. Microbiol.* **2014**, *5*, 25.

(67) van Berkel, W. J. H.; Kamerbeek, N. M.; Fraaije, M. W. *J. Biotechnol.* **2006**, *124*, 670–689.

(68) Galán, B.; Díaz, E.; Prieto, M. A.; García, J. L. *J. Bacteriol.* **2000**, *182* (3), 627–636.

(69) Furuya, T.; Kino, K. *Appl. Microbiol. Biotechnol.* **2014**, *98* (3), 1145–1154.

(70) Coulombel, L.; Nolan, L. C.; Nikodinovic, J.; Doyle, E. M.; O'Connor, K. E. *Appl. Microbiol. Biotechnol.* **2011**.

(71) Lin, Y.; Yan, Y. *Microb. Cell Fact.* **2012**, *11*, 42.

(72) Kim, S. H.; Hisano, T.; Takeda, K.; Iwasaki, W.; Ebihara, A.; Miki, K. *J. Biol. Chem.* **2007**, *282* (45), 33107–33117.

(73) Songsiang, U.; Thongthoom, T.; Boonyarat, C.; Yenjai, C. *J. Nat. Prod.* **2011**, *74*, 208–211.

(74) Hieda, Y.; Anraku, M.; Choshi, T.; Tomida, H.; Fujioka, H.; Hatae, N.; Hori, O.; Hirose, J.; Hibino, S. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 3530–3533.

(75) Voet, D.; Voet, J. G. *Biochem.*, 3rd ed.; Wiley-VCH, **2004**.

(76) Chang, H. K.; Zylstra, G. J. *Biochem. Biophys. Res. Commun.* **2008**, *371*, 149–153.

(77) Lee, J. K.; Zhao, H. *J. Bacteriol.* **2007**, *189* (23), 8556–8563.

(78) Tiwari, M. K.; Singh, R. K.; Lee, J. K.; Zhao, H. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 1344–1347.

(79) You, C.; Percival Zhang, Y. H. *Adv. Biochem. Eng. Biotechnol.* **2013**, *131*, 89–119.

(80) Homaei, A. A.; Sariri, R.; Vianello, F.; Stevanato, R. *J. Chem. Biol.* **2013**, *6*, 185–205.

(81) Weng, J. K.; Li, X.; Bonawitz, N. D.; Chapple, C. *Curr. Opin. Biotechnol.* **2008**, *19*, 166–172.

(82) Upton, B. M.; Kasko, A. M. *Chem. Rev.* **2016**, *116*, 2275–2306.

(83) Llevot, A.; Grau, E.; Carlotti, S.; Grelier, S.; Cramail, H. *Macromol. Rapid Commun.* **2016**, *37* (1), 9–28.

(84) Pilkington, L. I.; Barker, D. *Nat. Prod. Rep.* **2015**, *32*, 1359–1528.

(85) Fazary, A. E.; Alfaifi, M. Y.; Saleh, K. A.; Alshehri, M. A.; Eldin, S.; Elbehairi, I. *Nat. Prod. Chem. Res.* **2016**, *4* (4), 1000226.

(86) Mao, J.; Yu, N.-J.; Yang, Y.; Zhao, Y.-M. *J. Int. Pharm. Res.* **2014**, *41*, 275–281.

(87) Boerjan, W.; Ralph, J.; Baucher, M. *Annu. Rev. Plant Biol* **2003**, *54*, 519–546.

(88) Vanholme, R.; Morreel, K.; Ralph, J.; Boerjan, W. *Curr. Opin. Plant Biol.* **2008**, *11*, 278–285.

(89) Schoch, G. A.; Morant, M.; Abdulrazzak, N.; Asnaghi, C.; Goepfert, S.; Petersen, M.; Ullmann, P.; Werck-Reichhart, D. *Environ. Chem. Lett.* **2006**, *4*, 127–136.

(90) Xun, L.; Sandvik, E. R. *Appl. Environ. Microbiol.* **2000**, *66* (2), 481–486.

(91) Ibdah, M.; Zhang, X. H.; Schmidt, J.; Vogt, T. *J. Biol. Chem.* **2003**, *278* (45), 43961–43972.

(92) Napora-Wijata, K.; Strohmeier, G. A.; Winkler, M. *Biotechnol. J.* **2014**, *9*, 822–843.

(93) Bauer, A.-K. Biocatalytic synthesis of taste-modifying flavonoids, ed. L. Wessjohann, Ph.D. thesis, Martin Luther University Halle-Wittenberg, **2016**.

(94) Li, T.; Rosazza, J. P. N. *J. Bacteriol.* **1997**, *179* (11), 3482–3487.

(95) He, A.; Li, T.; Daniels, L.; Fotheringham, I.; Rosazza, J. P. N. *Appl. Environ. Microbiol.* **2004**, *70* (3), 1874–1881.

(96) Venkitasubramanian, P.; Daniels, L.; Rosazza, J. P. N. *J. Biol. Chem.* **2007**, *282* (1), 478–485.

(97) Ma, Q. H. *J. Exp. Bot.* **2010**, *61* (10), 2735–2744.

(98) Mitchell, H. J.; Hall, S. A.; Stratford, R.; Hall, J. L.; Barber, M. S. *Planta* **1999**, *208*, 31–37.

(99) Sun, X.; Shen, X.; Jain, R.; Lin, Y.; Wang, J.; Sun, J.; Wang, J.; Yan, Y.; Yuan, Q. *Chem. Soc. Rev.* **2015**, *44* (44), 3760–3785.

(100) Huang, Q.; Lin, Y.; Yan, Y. *Biotechnol. Bioeng.* **2013**, *110* (12), 3188–3196.

(101) Jansen, F.; Gillessen, B.; Mueller, F.; Commandeur, U.; Fischer, R.; Kreuzaler, F. *Biotechnol. Appl. Biochem.* **2014**, *61* (6), 646–654.

(102) Van Summeren-Wesenhagen, P. V; Voges, R.; Dennig, A.; Sokolowsky, S.; Noack, S.; Schwaneberg, U.; Marienhagen, J. *Microb. Cell Fact.* **2015**, *14*, 79.

(103) Rodriguez, G. M.; Atsumi, S. *Metab. Eng.* **2014**, *25*, 227–237.

(104) Kunjapur, A. M.; Tarasova, Y.; Prather, K. L. J. *J. Am. Chem. Soc.* **2014**, *136*, 11644–11654.

(105) Meyer, D.; Neumann, P.; Parthier, C.; Friedemann, R.; Nemeria, N.; Jordan, F.; Tittmann, K. *Biochemistry* **2010**, *49*, 8197–8212.

(106) Meyer, D.; Walter, L.; Kolter, G.; Pohl, M.; Müller, M.; Tittmann, K. *J. Am. Chem. Soc.* **2011**, *133*, 3609–3616.

(107) Gallage, N. J.; Hansen, E. H.; Kannangara, R.; Olsen, C. E.; Motawia, M. S.; Jørgensen, K.; Holme, I.; Hebelstrup, K.; Grisoni, M.; Lindberg Møller, B. *Nat. Commun.* **2014**, *5*, 1–14.

(108) Ley, J. P.; Krammer, G.; Reinders, G.; Gatfield, I. L.; Bertram, H. J. *J. Agric. Food Chem.* **2005**, *53*, 6061–6066.

(109) Taylor, A.; Hort, J. *Modifying flavour in food*; Woodhead Publishing Limited, **2007**.

(110) Gruber-Khadjawi, M.; Dippe, M.; Tengg, M.; Wessjohann, L. A. In *Cascade biocatalysis: Integrating stereoselective and environmentally friendly*

reactions, Vol. 1; Riva, S., Fessner, W.-D., Eds.; Wiley VCH: Weinheim, **2014**; pp 393–426.

(111) Furuya, T.; Miura, M.; Kino, K. *Chembiochem* **2014**, *15*, 2248–2254.

(112) Neglshi, O.; Sugiura, K.; Negishi, Y. *J. Agric. Food Chem.* **2009**, *57*, 9956–9961.

(113) Fock-Bastide, I.; Palama, T. L.; Bory, S.; Lecolier, A.; Noirot, M.; Joet, T. *Plant Physiol. Biochem.* **2014**, *74*, 304–314.

(114) Serov, A. E.; Popova, A. S.; Fedorchuk, V. V; Tishkov, V. I. *Biochem. J.* **2002**, *367* (3), 841–847.

(115) Pihlavisto, P.; Reenila, I. *J. Chromatogr.* **2002**, *781*, 359–372.

(116) Wei, H.; Zhang, R.; Wang, C.; Zheng, H.; Li, A.; Chou, K. C.; Wei, D. Q. *J. Theor. Biol.* **2007**, *244*, 692–702.

(117) Lozada-Ramírez, J. D.; Martínez-Martínez, I.; García-Carmona, F.; Sánchez-Ferrer, A. *Biotechnol. Prog.* **2008**, *24*, 120–127.

(118) Cornell, K. A.; Swarts, W. E.; Barry, R. D.; Riscoe, M. K. *Biochem. Biophys. Res. Commun.* **1996**, *228*, 724–732.

(119) 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. *Phytochemistry* **2005**, *66*, 2072–2091.

(120) Mitra, A.; Kitamura, Y.; Gasson, M. J.; Narbad, A.; Parr, A. J.; Payne, J.; Rhodes, M. J. C.; Sewter, C.; Walton, N. J. *Arch. Biochem. Biophys.* **1999**, *365*, 10–16.

(121) Lee, H.; DeLoache, W. C.; Dueber, J. E. *Metab. Eng.* **2012**, *14*, 242–251.

(122) Wessjohann, L. A.; Riemer-Köhler, S.; Dippe, M.; Geißler, T.; Geißler, K.; Ley, J. Biotechnological Methods for Providing 3,4-Dihydroxyphenyl Compounds and Methylated Variants Thereof. EP 15188136.4-1501, **2015**.

(123) Deifel, A. *Zeitschrift für Leb. und Forsch.* **1989**, *188* (4), 330–332.

(124) Monge, P.; Scheline, R.; Solheim, E. *Xenobiotica* **1976**, *6* (7), 411–423.

(125) Smith, L. R. *Chem. Educ.* **1996**, *1* (3), 1–18.

(126) Toogood, H. S.; Gardiner, J. M.; Scrutton, N. S. *ChemCatChem* **2010**, *2*, 892–914.

(127) Straßner, J.; Fürholz, A.; Macheroux, P.; Amrhein, N.; Schaller, A. *J. Biol. Chem.* **1999**, *274* (49), 35067–35073.

(128) Stueckler, C.; Hall, M.; Ehammer, H.; Pointner, E.; Kroutil, W.; Macheroux, P.; Faber, K. *Org. Lett.* **2007**, *9* (26), 5409–5411.

(129) Trotter, E. W.; Collinson, E. J.; Dawes, I. W.; Grant, C. M. *Appl. Environ. Microbiol.* **2006**, *72* (7), 4885–4892.

(130) Hall, M.; Stueckler, C.; Ehammer, H.; Pointner, E.; Oberdorfer, G.; Gruber,

K.; Hauer, B.; Stuermer, R.; Kroutil, W.; Macheroux, P.; Faber, K. *Adv. Synth. Catal.* **2008**, *350*, 411–418.

(131) Wasternack, C.; Parthier, B. *Trends Plant Sci.* **1997**, *2* (8), 302–307.

(132) Strassner, J.; Schaller, F.; Frick, U. B.; Howe, G. A.; Weiler, E. W.; Amrhein, N.; Macheroux, P.; Schaller, A. *Plant J.* **2002**, *32*, 585–601.

(133) Hall, M.; Stueckler, C.; Kroutil, W.; Macheroux, P.; Faber, K. *Angew. Chemie - Int. Ed.* **2007**, *119*, 4008–4011.

(134) Shimoda, K.; Ito, D. I.; Izumi, S.; Hirata, T. *J. Chem. Soc. Perkin Trans.* **1996**, *1* (4), 355–358.

(135) Sheldon, R. A.; Van Pelt, S. *Chem. Soc. Rev. Chem. Soc. Rev* **2013**, *42* (42), 6223–6235.

(136) Faber, K. In *Biotransformations in Organic Chemistry*; Springer-Verlag: Heidelberg, **2011**; pp 356–367.

(137) Scouten, W. H.; Luong, J. H. T.; Stephen Brown, R. *Trends Biotechnol.* **1995**, *13*, 178–185.

(138) Wachtmeister, J.; Jakoblinnert, A.; Kulig, J.; Offermann, H.; Rother, D. *ChemCatChem* **2014**, *6*, 1051–1058.

(139) Wachtmeister, J.; Mennicken, P.; Hunold, A.; Rother, D. *ChemCatChem* **2016**, *8*, 607–614.

(140) Kameda, A.; Shiba, T.; Kawazoe, Y.; Satoh, Y.; Ihara, Y.; Munekata, M.; Ishige, K.; Noguchi, T. *J. Biosci. Bioeng.* **2001**, *91* (6), 557–563.

(141) Arora, P.; Vats, A.; Saxena, P.; Mohanty, D.; Gokhale, R. S. *J. Am. Chem. Soc.* **2005**, *127*, 9388–9389.

(142) Finnigan, W.; Thomas, A.; Cromar, H.; Gough, B.; Snajdrova, R.; Adams, J. P.; Littlechild, J. A.; Harmer, N. J. *ChemCatChem* **2017**, *9* (6), 1005–1017.

(143) Nocek, B.; Kochinyan, S.; Proudfoot, M.; Brown, G.; Evdokimova, E.; Osipiuk, J.; Edwards, A. M.; Savchenko, A.; Joachimiak, A.; Yakunin, A. F.; Haselkorn, R. *PNAS* **2008**, *105* (46), 17730–17735.

(144) Satoh, Y.; Tajima, K.; Tannai, H.; Munekata, M. *J. Biosci. Bioeng.* **2003**, *95* (4), 335–341.

(145) Engler, C.; Kandzia, R.; Marillonnet, S. *PLoS One* **2008**, *3* (11), e3647.

(146) Werner, S.; Engler, C.; Weber, E.; Gruetzner, R.; Marillonnet, S. *PLoS One* **2012**, *6* (2), e16765.

Presentations and Paper manuscripts

Paper manuscripts

2016   S. Herrmann, M. Dippe, M. Meyer, P. Majovsky, M. Pietzsch, L. A. Wessjohann: Enzymatic synthesis of sinapyl alcohol and related monolignols from *p*-coumaric acid, Biotechnology and Bioengineering, in preparation.

2015   M. Dippe, S Riemer-Köhler, B. Weigel, A.-K. Bauer, A. Laub, J. Schmidt, L. A. Wessjohann: A modular cell-free multi-enzyme cascade for vanillin synthesis and phenylpropanoid modification, Angewandte Chemie, under revision.

2014   S. Herrmann, M. Dippe, M. Pietzsch, L. A. Wessjohann: Rational engineering of the substrate spectrum of a FAD-dependent monooxygenase from *Escherichia coli*, ChemCatChem, in preparation.

Presentations (Poster)

M. Dippe, S. Riemer-Köhler, B. Weigel, A.-K. Bauer, A. Laub, L. A. Wessjohann: *A modular cell-free multi-enzyme cascade for vanillin synthesis*, 27. Irseer Naturstofftage, 25.-27.02.2015, Irsee

M. Dippe, S. Riemer-Köhler, B. Weigel, A.-K. Bauer, A. Laub, L. A. Wessjohann: *A modular cartridge system for flavor production*, 11th Plant Science Student Conference, 02.-05.06.2015, Halle (Saale)

S. Riemer-Köhler, D. Meyer, M. Pietzsch, L. A. Wessjohann, *Biocatalytic route for cell-free synthesis of sinapyl alcohol and related monolignols*, 8th International Congress on Biocataysis, 28.08.-01.09.2016, Hamburg-Harburg

Curriculum vitae

Personal details

Name	Dipl.-LMChem. Susann Herrmann, nee Riemer-Köhler
Date of Birth	December 20th, 1988
Place of Birth	Cottbus
Nationality	German

Education

07/2013 – 07/2017	Ph.D. student Leibniz Institute of Plant Biochemistry, Halle (Saale), Germany, Department Bioorganic Chemistry Thesis title: Enzymatic Synthesis of Plant Phenylpropanoids Supervisor: Prof. Dr. Ludger A. Wessjohann
10/2008 – 04/2013	Food chemistry studies at the Martin Luther University Halle-Wittenberg, final grade 1.5
06/2008	Abitur, final grade 2.1 (Ludwig-Leichhardt-Gymnasium, Cottbus)
2001-2008	Secondary school, (Ludwig-Leichhardt-Gymnasium, Cottbus)

Postgraduate Training

07/2014	GMP training course (SFi Halle)
11/2013	Workshop about protein electrophoresis (SERVA)

Additional Work

07/2011 – 06/2012	Student assistant at the Leibniz Institute of Plant Biochemistry in Halle (Saale)/Germany, Department Bioorganic Chemistry

Halle (Saale),

Susann Herrmann

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe. Andere, als die von mir angegebene Quellen und Hilfsmittel, habe ich nicht benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen habe ich als solche kenntlich gemacht. Weiterhin wurde diese Arbeit an keiner weiteren wissenschaftlichen Einrichtung zur Erlangung des akademischen Grades eingereicht.

Halle (Saale),

Susann Herrmann