TECHNISCHE UNIVERSITÄT DARMSTADT

Metabolic engineering of cannabinoid biosynthesis in tobacco

Vom Fachbereich Biologie der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades

Doctor rerum naturalium

genehmigte Dissertation von

M. Sc. Marcus Frank Geißler

aus Höchst im Odenwald

1. Referent: Prof. Dr. Heribert Warzecha

2. Referent: Prof. Dr. Adam Bertl

Darmstadt 2021

Geißler, Marcus Frank: Metabolic engineering of cannabinoid biosynthesis in tobacco

Darmstadt, Technische Universität Darmstadt

Jahr der Veröffentlichung der Dissertation auf TUprints: 2021

Tag der mündlichen Prüfung: 11.06.2021

Für Jana

Abstract

For millennia, *Cannabis* has been used in various cultural groups because of its versatile applications. Among other things, the plant served as a source of textile fibers, but was also used as a medicine to treat inflammation, cramps or epilepsy. With the discovery of the endocannabinoid system, the plant, which had fallen into disrepute in the mid-20th century because of its excessive use as a recreational drug and is therefore still subject to heavy restrictions in most countries, gained new prestige. Thus, it was found that certain secondary metabolites, known as cannabinoids, could modulate a variety of physiological processes, potentially offering new therapeutic applications. To date, more than 100 of these cannabinoids have been isolated, the best known of which are the psychoactive-acting Δ^9-tetrahydrocannabinol (THC) and the non-psychotropic cannabidiol (CBD). In this context, THC- and CBD-containing products such as Sativex® and Epidyolex®, which are already approved as pharmaceuticals, are finding application in medicine. However, THC is considered a major limitation for clinical use due to its psychoactive character, so non-psychotropic cannabinoids, as well as synthetic non-naturally-occurring cannabinoids that have similar medicinal properties to THC, but do not have the undesirable side effect, will be of key importance in the future. Given that chemical production of synthetic cannabinoids is very costly, heterologous production in host organisms could provide a remedy, potentially leading to industrial-scale production of rare phytocannabinoids or novel synthetic cannabinoid pharmaceuticals not readily offered by cannabis plants.

In this regard, the aim was to establish tobacco as an alternative host organism for the biosynthetic production of cannabinoids. It was shown that it was possible to produce all enzymes involved in cannabinoid biosynthesis in transiently transformed *Nicotiana benthamiana* plants and to detect their activity *in vitro*. Moreover, transient expression of *aae1*, *ols*, and *oac* and supplementation of hexanoic acid *in vivo* resulted in the formation of both, the cannabinoid precursor olivetolic acid (OA) and the new-to-nature C-4 OA glucoside. However, beyond the synthesis of OA, it was not yet possible to reconstruct the biosynthetic pathway *in vivo,* probably due to the lack of sufficient geranyl diphosphate (GPP) supply within tobacco plants. To enable efficient production of cannabinoids in the future, it is therefore essential to eliminate this bottleneck in the biosynthetic pathway.

In addition to the transient approach, stably transformed liquid cell cultures were generated, which expressed the necessary genes for the production of OA or its glucoside. This should, with a view to use in industrial production, enable large-scale cultivation in bioreactors under GMP conditions. In this context, it was possible to integrate the individual genes into *Nicotiana tabacum* by stable

transformation. However, in contrast to transient expression, neither the synthesis of OA nor its glucoside could be detected, most likely due to a mutation in the OLS gene that was used.

The second part of the work dealt with the characterization of the cannabinoid-forming synthases Δ^9-tetrahydrocannabinolic acid synthase (THCAS), cannabichromenic acid synthase (CBCAS) and cannabidiolic acid synthase (CBDAS). It was found that *N*-glycosylation and therefore localization of the proteins *via* the secretory pathway, either into the apoplast or vacuole, is required for the production of the enzymes *in planta*. In addition, *in vitro* experiments with THCAS and CBDAS showed that when organic solvents such as acetonitrile or acetone were added, the product specificity of the enzymes changed from Δ^9-tetrahydrocannabinolic acid (THCA) and cannabidiolic acid (CBDA), respectively, to the synthesis of cannabichromenic acid (CBCA). This in turn indicates that, among other things, also the hydrophobic environment in the glandular trichomes of *Cannabis* could be responsible for the cannabinoid diversity in different *Cannabis* strains.

Since CBCAS has a 93 % amino acid identity to THCAS but does not produce THCA, it stands to reason that small differences in amino acid sequence outside the catalytic center affect cyclization specificities. Therefore, final mutagenesis studies were performed with CBCAS and the goal of producing THCA to gain further insight into the catalytic mechanisms of the synthases. However, none of the so far introduced mutations resulted in the production of the desired cannabinoid, thus further site-directed mutagenesis have to be performed in the future, including the initially neglected amino acids.

Zusammenfassung

Seit Jahrtausenden wird *Cannabis* in diversen Kulturen aufgrund ihrer vielseitigen Anwendungsmöglichkeiten genutzt. So diente die Pflanze unter anderem als Quelle für Textilfasern, fand aber auch als Arzneimittel zur Behandlung von Entzündungen, Krämpfen oder Epilepsie Verwendung. Mit der Entdeckung des Endocannabinoid-Systems erlangte die Pflanze, die Mitte des 20. Jahrhunderts wegen ihres übermäßigen Gebrauchs als Freizeitdroge in Verruf geraten war und bis heute in den meisten Ländern starken Restriktionen unterliegt, neues Ansehen. So fand man heraus, dass bestimmte Sekundärmetabolite, sogenannte Cannabinoide, eine Vielzahl von physiologischen Prozessen modulieren konnten und somit potentiell neue therapeutische Anwendungsmöglichkeiten eröffneten. Bis heute wurden mehr als 100 dieser Cannabinoide isoliert, von welchen die bekanntesten das psychoaktiv-wirkende Δ^9-Tetrahydrocannabinol (THC) und das nicht-psychotrope Cannabidiol (CBD) darstellen. In diesem Zusammenhang finden die bereits als Arzneimittel zugelassenen THC- und CBD-haltigen Produkte wie Sativex® und Epidyolex® Anwendung in der Medizin. Allerdings gilt THC aufgrund seines psychoaktiven Charakters als große Einschränkung für den klinischen Einsatz, weshalb nicht-psychotrope Cannabinoide sowie synthetische, nicht natürlich-vorkommende Cannabinoide, die ähnliche medizinische Eigenschaften wie THC aufweisen, aber nicht die unerwünschte Nebenwirkung haben, in Zukunft von zentraler Bedeutung sein werden. Da die chemische Produktion synthetischer Cannabinoide jedoch sehr kostspielig ist, könnte die heterologe Produktion in Wirtsorganismen Abhilfe schaffen und möglicherweise zu einer Produktion seltener Phytocannabinoide oder neuartiger synthetischer Cannabinoid-Pharmazeutika im industriellen Maßstab führen, die durch den reinen Gebrauch von Cannabispflanzen nicht ohne weiteres zu Verfügung stehen.

Im Hinblick darauf sollte Tabak als potentielle Produktionsplattform für die biosynthetische Herstellung von Cannabinoiden etabliert werden. Dabei konnte gezeigt werden, dass es möglich war alle an der Cannabinoid-Biosynthese beteiligten Enzyme in transient transformierten Tabakpflanzen zu produzieren und deren Aktivität *in vitro* nachzuweisen. Darüber hinaus führte die transiente Expression von *aae1*, *ols* und *oac* und die Supplementierung von Hexansäure *in vivo* sowohl zur Bildung des Cannabinoid-Vorläufermoleküls Olivetolsäure (OA) als auch zur Bildung des neuartigen C-4 OA-Glukosids. Über die Synthese von OA hinaus war es jedoch nicht möglich, den Biosyntheseweg *in vivo* zu rekonstruieren, was wahrscheinlich auf das Fehlen einer ausreichenden Versorgung mit Geranyldiphosphat (GPP) innerhalb der Pflanze zurückzuführen ist. Um zukünftig eine effiziente Produktion von Cannabinoiden zu ermöglichen, ist es daher essentiell diesen Engpass im Biosyntheseweg zu beseitigen.

Zusätzlich zum transienten Ansatz sollten zudem stabil transformierte Flüssigzellkulturen generiert werden, welche die notwendigen Gene für die Produktion von OA oder dessen Glukosid exprimierten. Dies sollte im Hinblick auf einen Einsatz in der industriellen Produktion, eine großtechnische Kultivierung in Bioreaktoren unter GMP-Bedingungen ermöglichen. Im Gegensatz zur transienten Expression konnte jedoch weder die Synthese von OA noch dessen Glukosid nachgewiesen werden, was höchstwahrscheinlich einer Mutation im genutzten OLS-Gen zuzuschreiben ist.

Der zweite Teil der Arbeit beschäftigte sich mit der Charakterisierung der Cannabionid-bildenden Synthasen Δ^9-Tetrahydrocannabinolsäure-Synthase (THCAS), Cannabichromensäure-Synthase (CBCAS) und Cannabidiolsäure-Synthase (CBDAS). Dabei stellte sich heraus, dass für die Produktion der Enzyme *in planta* eine *N*-Glykosylierung und damit verbunden eine Lokalisation der Proteine über den sekretorischen Weg, entweder in den Apoplasten oder die Vakuole, vonnöten ist. Zudem zeigte sich in *in vitro* Versuchen mit THCAS und CBDAS, dass sich unter Zugabe von organischen Lösemitteln wie Acetonitril oder Aceton, die Produktspezifität der Enzyme von Δ^9-Tetrahydrocannabinolsäure (THCA) beziehungsweise Cannabidiolsäure (CBDA) hin zur Synthese von Cannabichromensäure (CBCA) änderte. Dies wiederrum könnte darauf hindeuten, dass unter anderem die hydrophobe Umgebung in den Drüsentrichomen von *Cannabis* für die Cannabinoid-Diversität in verschiedenen *Cannabis* Varietäten verantwortlich ist.

Da die CBCAS eine Aminosäure-Identität von 93 % gegenüber der THCAS aufweist, jedoch kein THCA produziert, liegt es nahe, dass kleine Unterschiede in der Aminosäuresequenz außerhalb des katalytischen Zentrums die Zyklisierungsspezifitäten beeinflussen. Daher wurden abschließend Mutagenese-Studien mit CBCAS und dem Ziel THCA zu produzieren, durchgeführt, um weitere Einblicke in die katalytischen Mechanismen der Synthasen erhalten. Allerdings führte keine der eingeführten Mutationen zur Produktion des gewünschten Cannabinoids, weshalb in Zukunft weitere ortsgerichtete Mutagenesen durchgeführt werden müssen, die auch die zunächst vernachlässigten Aminosäuren mit einschließen.

Table of Contents

List of abbreviations

AAE1	Acetyl activating enzyme 1
AtIPK	*Arabidopsis thaliana* isopentenyl phosphate kinase
CBC	Cannabichromene
CBCA	Cannabichromenic acid
CBCAS	Cannabichromenic acid synthase
CBD	Cannabidiol
CBDA	Cannabidiolic acid
CBDAS	Cannabidiolic acid synthase
CBDV	Cannabidivarin
CBG	Cannabigerol
CBGA	Cannabigerolic acid
CBGAS	Cannabigerolic acid synthase
CBNA	Cannabinolic acid
COVID-19	Coronavirus disease 2019
C-2 CBGA glucoside	C-2 cannabigerolic acid glucoside
C-4 CBGA glucoside	C-4 cannabigerolic acid glucoside
C-2 OA glucoside	C-2 olivetolic acid glucoside
C-4 OA glucoside	C-4 olivetolic acid glucoside
ECS	Endocannabionid system
GB	GoldenBraid
GGPPS	Geranylgeranyl diphosphate synthase
GPP	Geranyl diphosphate
GPPS	Geranyl diphosphate synthase
LimS	Limonene synthase
MEP	2-C-methyl-D-erythritol-4-phosphate
NphB	Soluble aromatic prenyltransferase from *Streptomyces* sp. strain CL190
OA	Olivetolic acid
OAC	Olivetolic acid cyclase
OLS	Olivetol synthase

PNP	Plant natural product
PT4	Prenyltransferase 4 from *Cannabis sativa*
SARS-CoV-2	Severe acute respiratory syndrome coronavirus type 2
THAS1	Tetrahydroalstonine synthase 1
THC	Δ^9-tetrahydrocannabinol
THCA	Δ^9-tetrahydrocannabinolic acid
THCAS	Δ^9-tetrahydrocannabinolic acid synthase
UGT	UDP-glycosyltransferases
2-*O*-GOA	2-*O*-geranyl olivetolic acid

List of figures

List of tables

Chapter 1

Introduction

1.1. *Cannabis* taxonomy

Cannabis is a genus of annual flowering, dioecious plants belonging to the family of *Cannabaceae*. However, regarding the number of different species within the genus and consequently in their nomenclature there is still controversy in the scientific community today [1]. Several suggestions for the number of *Cannabis* species and subspecies are made in literature, mainly referring to geographic, phenotypic, or chemical differences in metabolite composition [2,3]. Generally, three species of *Cannabis* are distinguished, namely *C. sativa*, *C. indica* and *C. ruderalis*, with *C. ruderalis* often referred to as the ancestor of the other species and probably extinct by now [4]. *Cannabis* itself is subdivided into so-called hemp or marihuana. However, from a scientific point of view, these are not subspecies, but are defined only by their cannabinoid composition. While, hemp contains less than 0.3 % of the psychoactive substance Δ^9-tetrahydrocannabinol (THC) and its fibers have been used for thousands of years to make rope, canvas, clothing, paper, shoes and sails, marijuana contains higher amounts of THC and is more commonly used as a recreational drug [5]. Due to the long domestication and breeding history with currently more than 700 strains of *Cannabis*, bearing such colorful names like 'White Widow', 'Blueberry Diesel' or 'Super Lemon Haze', differentiation is often problematic and most authors consider the genus to be a highly variable, monotypic species of *Cannabis sativa* L. [3,6]. In order to avoid misinterpretations in the subsequent texts, *Cannabis*, *Cannabis sativa* and *Cannabis sativa* L. are therefore used synonymously.

1.2. The use of *C. sativa* in medicine: From then to now

For thousands of years, *Cannabis* has been used as a source of textile fibers, oil or in folk medicine for its therapeutic effects. It is assumed to be one of the oldest domesticated crops sharing its origins with the emergence of the first agricultural human societies in Asia as far as 12,000 years ago, from where it eventually began its conquest around the world (Figure 1.2.1) [7–9].

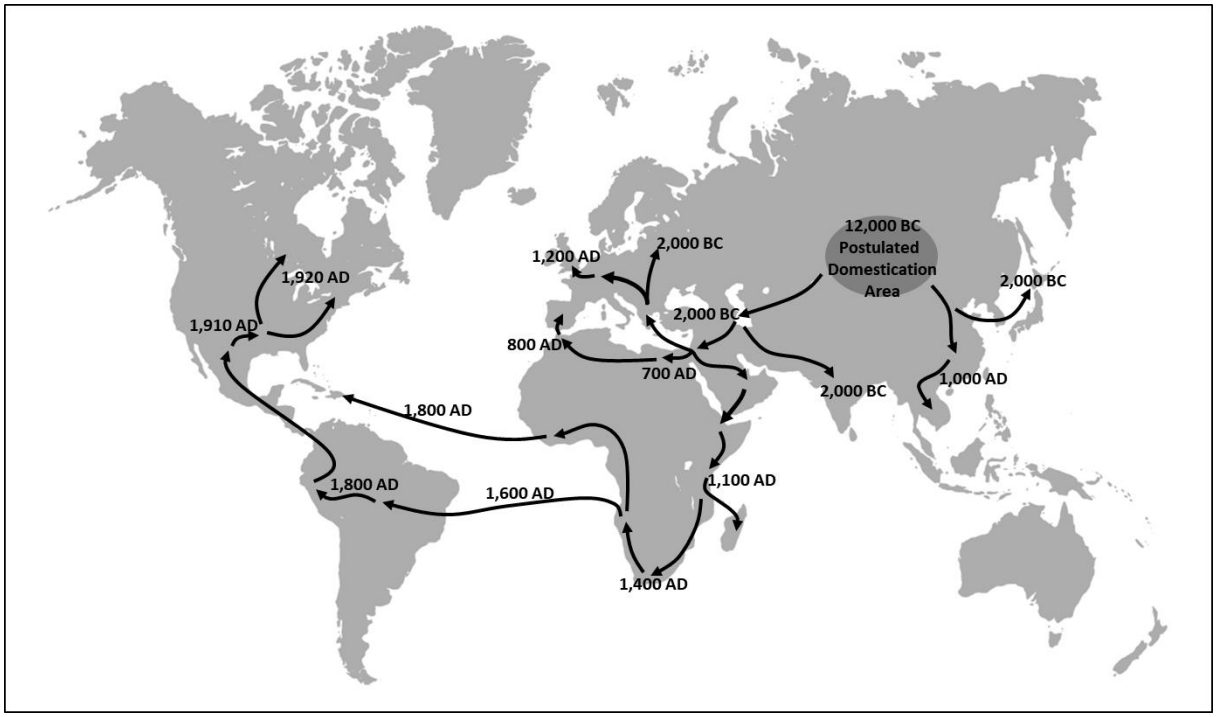


Figure 1.2.1 Historical diffusion of *Cannabis sativa* (modified according to Warf, 2014 [5]). Source world map: Freepik.com.

In Neolithic China hemp was widely used for manufacturing rope, clothing, sails, and bowstrings. The first evidence of *Cannabis* application in medicine was reported in 2,700 B.C., using it as an anesthetic during surgery or the treatment of rheumatic pain, intestinal constipation and malaria [10,11]. Since the Chinese mainly used *Cannabis* seeds for their medicinal purposes, which usually contained hardly any cannabinoids, the health-promoting effect was mainly due to the essential fatty acids and proteins they contained. Today, some of these fatty acids, such as γ-linolenic acid, are known to have therapeutic effects against a variety of inflammatory diseases when administered orally. Regarding the application of the plant as psychoactive drug, however, little has been passed down in texts, as this effect was probably associated with shamanism and ancient texts rarely referred this kind of practice [12,13]. Approximately 2,000 B.C., *Cannabis* found its way to Korea and Japan. In Korea, hemp was most likely cultivated by the Chulmun coastal farmers, while it was used by the Jomon culture in Japan, on the one hand, for their rope-printed pottery and, on the other, by Shinto priests in ceremonial burnings [4,14]. In parallel, *Cannabis* was brought to the South Asian subcontinent and the Middle East between 2,000 and 1,000 B.C., presumably by Aryan invasions [5].

But unlike China, the utilization of *Cannabis* in India was widely disseminated for pharmaceutical application and as a recreational drug, since it was strongly coupled with their religion [12]. Here it was imbibed by the Ayurvedic healing tradition and has been used in a plethora of medical applications among other as an analgesic anticonvulsant, hypnotic, tranquilizing, anesthetic, anti-inflammatory, antibiotic and antispasmodic agent. In this context, the plant was typically mixed with herbs and was smoked as so-called *charas* or consumed in the form of *bhang*, which consisted of *Cannabis* as a mild paste or a mix of tea and milk [5,10,12]. The utilization of *Cannabis* continued to spread throughout the world and was adopted by many different cultures. Thus, the plant eventually found its way to Europe, where in 450 B.C. Herodotus first described that the Scythians burned the seeds of the plant for its euphoric effect. Despite the resounding success of the plant in India, *Cannabis* was initially hardly used for medicinal purposes in Europe, but rather for its fibers. However, this changed with the beginning of the Christian era [10,15]. Approximately in 1,000 A.D., the Persian physician Avicenna described in his medical compendium that *Cannabis* has diuretic, digestive and anti-flatulent properties. It was even reported that the administered resin cured epilepsy in individual cases, but made the patients addicted to it [10,16]. His writings were studied extensively from the thirteenth to the nineteenth centuries and had a lasting influence on Western medicine [17]. However, the first therapeutic uses of *Cannabis* in Western medicine were described in 1839, when the Irish physician William O'Shaughnessy tested the effects of different *Cannabis* forms on animals in order to evaluate the toxicity of the drug. In addition, he reported various successes in the treatment of rheumatism, cramps and especially muscle spasms in tetanus and rabies using *Cannabis* preparations in human trials. O'Shaughnessy's positive results were soon followed by other reports of physicians also achieving success with the use of *Cannabis*, which in turn led to the plants' rapid advance into North America, reaching its spike in the middle of the 1920s [10,15]. With the introduction of vaccines, for example against tetanus, or the development of synthetic painkillers, *Cannabis* became increasingly less important, which, with its amplified use as a recreational drug from the 1930s onwards, led to *Cannabis* being nationally regulated as a drug with the Marihuana Tax Act passed in 1937 and finally to the removal from the American pharmacopoeia in 1941. Beginning in the late 1920s, there was a general trend towards the prohibition of *Cannabis*, not only in the United States but in most Western countries. Consequently, at least in terms of its medicinal value, it became quiet about the plant, until 1960s, when the discovery of *Cannabis*' active ingredients, cannabidiol (CBD) as well as THC, but also cannabigerol (CBG) and cannabidivarin (CBDV) made it possible to isolate the substances and study their effects [10,18–21]. With the discovery of the endogenous G-protein coupled cannabinoid receptors CB_1 and CB_2 in the late 1980s and the subsequent isolation of anandamide and 2-arachidonylglycerol, so-called endocannabinoids, *Cannabis* finally made its comeback and gained renewed scientific interest [22–26]. The newly

acquired knowledge that phytocannabinoids can interact with these receptors and thus modulate a variety of physiological processes provided the opportunity to investigate the therapeutic effects of different phytocannabinoids and synthetic cannabinoids (Figure 1.2.2). For example, CBD has been shown in *in vitro* and animal studies to have anti-anxiety, anti-nausea, anti-arthritic, anti-psychotic, anti-inflammatory, and immunomodulatory properties, whereas THC as well as the synthetic cannabinoid nabilone have been used in the treatment of chemotherapy-induced nausea and vomiting [3,27–30]. With the ongoing COVID-19 pandemic it is even discussed if preparations of *Cannabis* or cannabinoid adjunctive treatments could suppress SARS-CoV-2 caused inflammation and the resulting 'cytokine storm', reducing the severity of COVID-19. However, to be able to say that with certainty, more studies and trials are required [31–33].

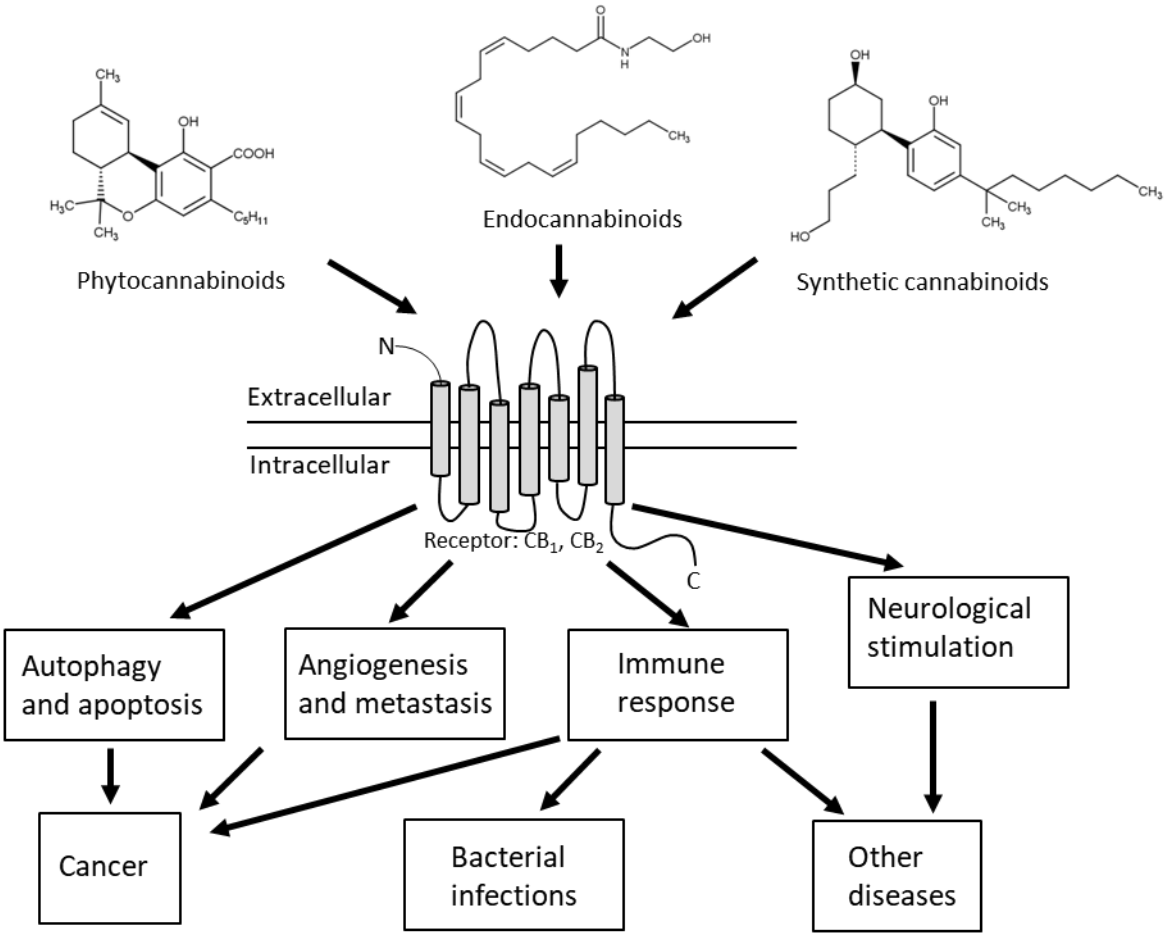


Figure 1.2.2 Involvement of endocannabinoid system in the modulation of different physiological processes (modified according to Śledziński *et al*., 2020 [34]).

1.3. Cannabinoids

Cannabinoids are a group of chemical constituents that were originally found in *C. sativa*. However, with the discovery of the endocannabinoid system (ECS) and the endocannabinoids anandamide and 2-arachidonylglycerol, and the subsequent development of extremely potent synthetic cannabinoid receptor agonists and antagonists, their terminology expanded. Meanwhile, three general types of cannabinoids are distinguished, namely phytocannabinoids, endocannabinoids and synthetic cannabinoids.

Endocannabinoids are a family of naturally occurring *Cannabis*-like biologically active lipids that are produced by vertebrates, regulating the function of multiple organs and tissues. They fulfill the criteria for being classified as neurotransmitters, but in contrast to classical neurotransmitters, they are actually neither synthesized in the cytosol of the neuron nor stored in synaptic vesicles. Specifically, they activate the G-protein coupled receptors CB_1 and CB_2, which are distributed differently throughout the body. While CB_1 receptors are concentrated in the central nervous system, CB_2 is mostly found in immune cells such as lymphocytes [35,36]. When released from postsynaptic neurons into the synaptic cleft, endocannabinoids trigger retrograde (feedback) effects on the presynaptic neuron which in turn causes inhibition of transmitter action at the affected synapse. Mechanisms modulated by the ECS include motor coordination, memory, pain modulation, neuroprotection, appetite, memory and maintenance of homeostasis, among others [36–39]. Endocannabinoids are depicted exemplarily in Figure 1.3.1.

Synthetic cannabinoids represent non-naturally occurring chemical compounds that possess strong binding effects on the endocannabinoid system receptors CB_1 or CB_2, displaying a pharmacological profile similar to the phytocannabinoids THC. There are several classes of synthetic cannabinoids, including the aminoalkylindoles, naphthoylindoles, phenlacetylindoles, cyclohexylphenols, tetramethylcyclopropylindoles, indole and indazole carboxamides, and quinoline esters, with some resembling the structure of the classic cannabinoid THC (CP 55,940), while others are structurally unrelated such as WIN55,212-2 or JWH-018 (Figure 1.3.1) [40].

Phytocannabinoids are defined as group of terpenophenolic plant-derived compounds predominantly produced in *Cannabis,* but also in other plant species including *Lepidium meyenii*, *Piper nigrum*, *Acmella oleracea* or *Radula marginata* that act on the ECS. To date more than 100 different cannabinoids have been reported, which are classified into several subclasses among which the group of cannabigerol type, cannabichromene type, cannabidiol type or tetrahydrocannabinol type cannabinoids contain the most important phytocannabinoids (Figure 1.3.2) [41–49]. In general, cannabinoids are characterized by their bicyclic or tricyclic structure and the majority of them are

derived from the predominantly produced and consequently in *Cannabis* most abundant acidic phytocannabinoids (depending on the *Cannabis* strain), Δ^9-tetrahydrocannabinolic acid-A (THCA), cannabigerolic acid (CBGA), cannabidiolic acid (CBDA), cannabinolic acid (CBNA) and cannabichromenic acid (CBCA) [50]. In a consecutive step, decarboxylation of the parent compound happens non-enzymatically through heat and light exposure as well as auto-oxidation. However, it is important to mention that substances structurally unrelated to conventional phytocannabinoids (THC and CBD), such as β-caryophyllene, have also shown an effect on the ECS by acting as a CB_2 receptor agonist, causing analgesic and anti-inflammatory activity without the undesired psychotropic effects [51].

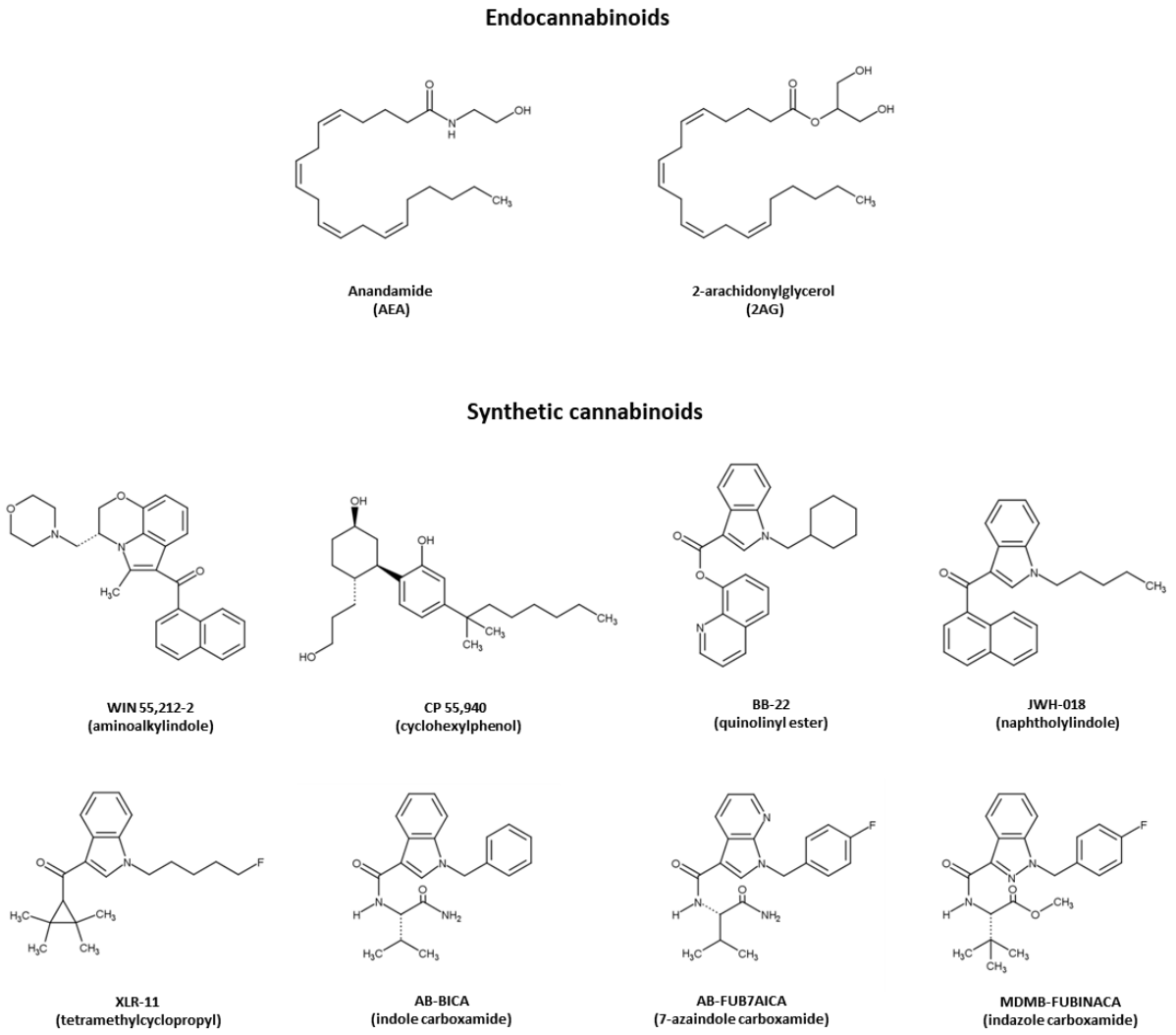


Figure 1.3.1 Structure of endocannabinoids and synthetic cannabinoids (modified according to Walsh and Anderson, 2020 and Chandra *et al.*, 2017 [3,40]).

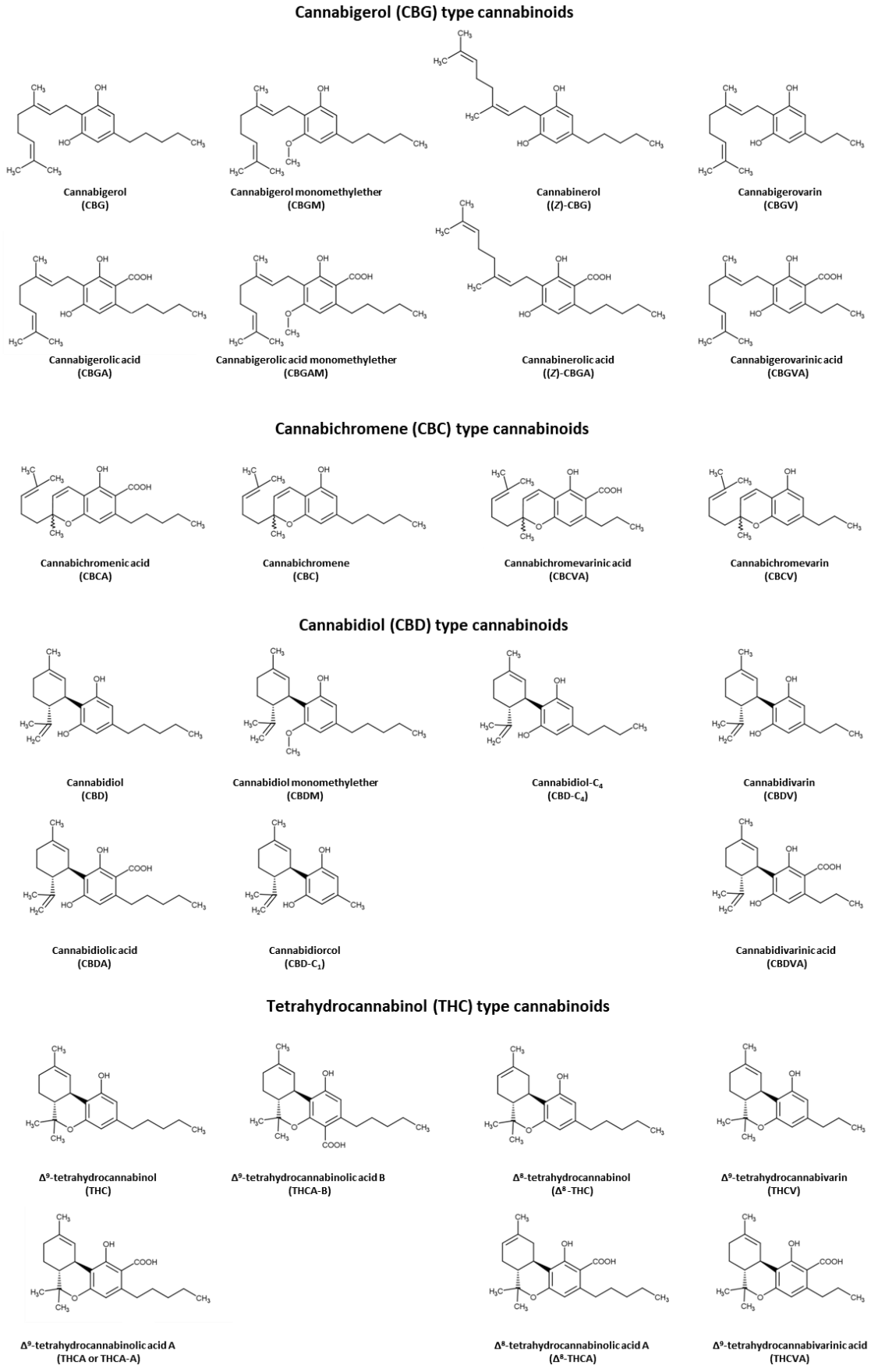


Figure 1.3.2 Overview of the most relevant cannabinoid classes and their structures (modified according to Chandra *et al.*, 2017 [3]).

1.4. Biosynthesis of cannabinoids

In *C. sativa*, cannabinoids are produced and stored in the balloon-shaped secretory cavity of the glandular trichomes. Cannabinoids are found throughout the plant, except in the roots, and at the highest density on the female flowers. Their content varies greatly depending on the plants growing conditions, such as nutrition, humidity, light exposure, as well as harvest time and storage conditions [52–54]. Over the last 20 years, the metabolic route for the production of cannabinoids in *C. sativa* has been elucidated and the enzymes involved identified. In general, the biosynthetic pathway can be divided into three parts: a monoterpene producing part (formation of geranyl diphosphate (GPP) *via* the 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway, localized in plastids), the olivetolic acid (OA) producing part (starting from fatty acid synthesis) and the actual cannabinoid producing part (Figure 1.4.1) [55,56]. The hexanoic acid originating from the fatty acid metabolism is first converted to the active CoA thioester (hexanoyl-CoA) by the acyl activating enzyme 1 (AAE1) [57]. OA is then formed from three molecules of malonyl-CoA and one molecule of hexanoyl-CoA by aldol condensation, catalyzed by the enzymes olivetol synthase (OLS) and olivetol acid cyclase (OAC) [58]. The central precursor for the synthesis of other cannabinoids, cannabigerolic acid (CBGA), is subsequently produced by prenylation of OA by means of the supposedly membrane-bound cannabigerolic acid synthase (CBGAS), also referred as geranyldiphosphate:olivetolate geranyltransferase (GOT), using GPP provided by the MEP pathway [59]. However, to date, there is insufficient scientific knowledge about CBGAS. While there is evidence that CBGAS is an integral membrane protein, such as the recently isolated PT4 or the patented PT1, both deriving from *C. sativa,* Fellermeier and Zenk detected CBGAS activity only in the soluble fraction of *C. sativa* plant crude extract, but not in the particulate fractions obtained by density gradient centrifugation [59–61]. Moreover, besides the actually isolated membrane-bound prenyltransferases, the soluble aromatic prenyltransferase NphB, deriving from the Streptomyces sp. strain CL190, was shown to be a promising candidate for replacing the harder to handle membrane proteins. Accordingly, *Komagataella phaffii* producing NphB was able to convert OA and GPP into CBGA. However, only approximately 15 % of the desired CBGA was formed, whereas 85 % of the enzymes turnover was assigned to the formation of the *O*-prenylated product, 2-*O*-geranyl olivetolic acid (2-*O*-GOA) [62]. In the last consecutive step, the different cannabinoids Δ^9-tetrahydrocannabinolic acid (THCA), cannabidiolic acid (CBDA) and cannabichromenic acid (CBCA) are formed by stereoselective cyclization of the terpene unit of CBGA by the respective enzymes Δ^9-tetrahydrocannabinolic acid synthase (THCAS), cannabidiolic acid synthase (CBDAS) and cannabichromenic acid synthase (CBCAS) [63–65]. These soluble oxidoreductases belong to the berberine-bridge-enzyme (BBE)-like enzyme family (PF 08031) and form a subgroup of the flavine adenine dinucleotide (FAD)-linked oxidase superfamily (SCOPe d.58.32) [66].

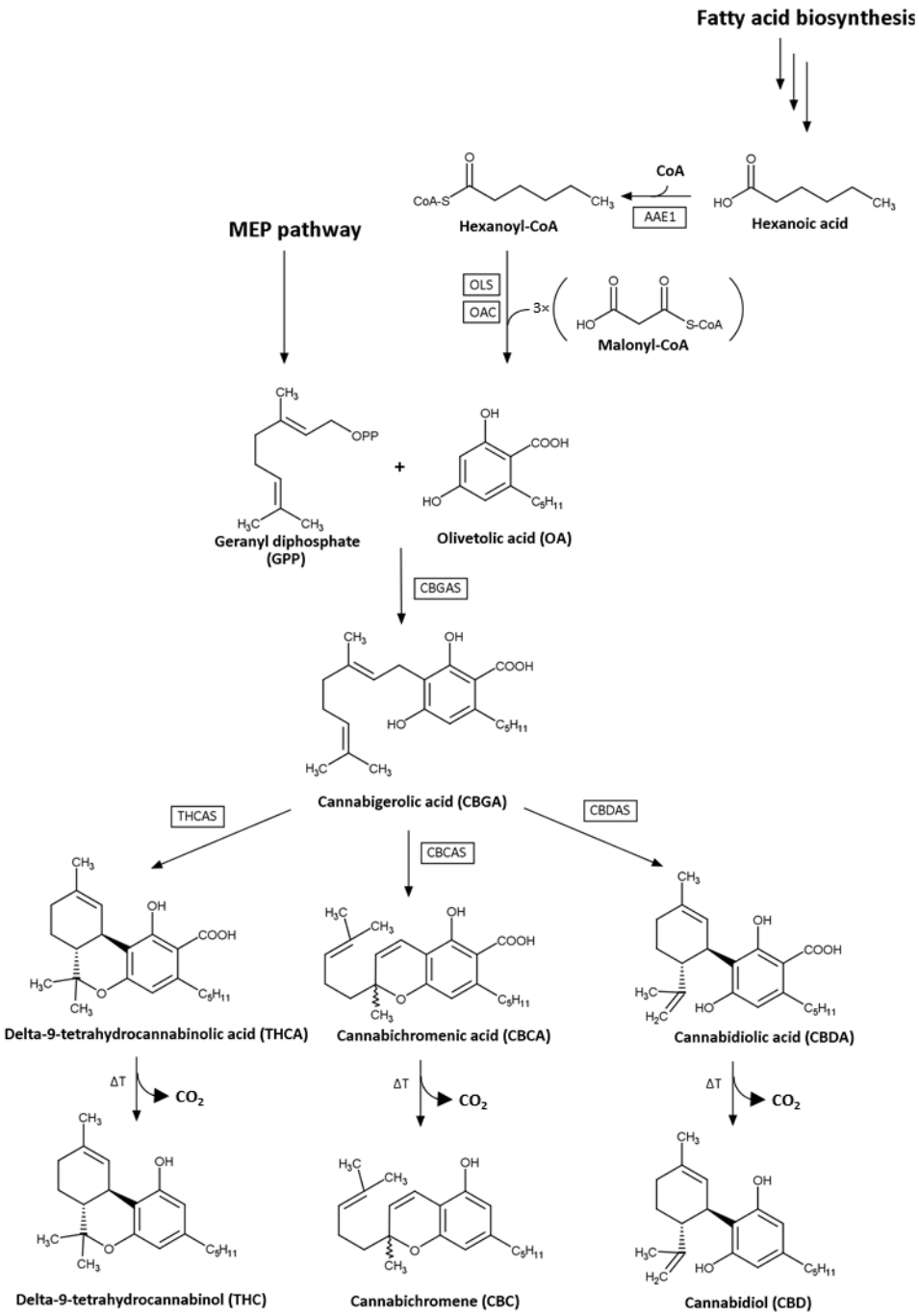


Figure 1.4.1 Schematic depiction of the cannabinoid biosynthesis pathway in *Cannabis sativa* (modified according to Zirpel *et al.* [62]). MEP, 2-C-methyl-D-erythritol-4-phosphate pathway; AAE1, acyl activating enzyme 1; OLS, olivetol synthase; OAC, olivetolic acid cyclase; CBGAS, cannabigerolic acid synthase; THCAS, Δ^9-tetrahydrocannabinolic acid synthase; CBDAS, cannabidiolic acid synthase; CBCAS, cannabichromenic acid synthase.

The structure of THCAS was first elucidated in 2012 by Shoyama *et al.*, showing that the enzyme contains several glycosylation sites, one disulfide bridge and a bi-covalently bound FAD co-factor [67]. Furthermore, THCAS and CBDAS are described to generate hydrogen peroxide as a side-product, when THCA and CBDA are formed [65,68]. Since the THCAS and CBDAS share a high amino

acid sequence similarity of 83 %, and THCAS and the CBCAS even share a similarity of 93 %, it stands to reason that small differences in amino acid sequence determine the cyclization and consequently the product specificity of these enzymes.

Looking at the reaction mechanism in detail, the reaction formally involves hydride abstraction from the benzallylic terpenyl carbon. The formation of the resulting cation allows for the generation of the cyclohexene ring of CBDA and THCA by electrophilic cyclization. Alternatively, the isomerized benzallyl cation can evolve into a quinone methide and generate CBCA by an electrocyclic reaction. While the electrophilic cyclization is enzyme driven and generates chiral products, the electrocyclic reaction is more likely to be spontaneous (Figure 1.4.2) [69].

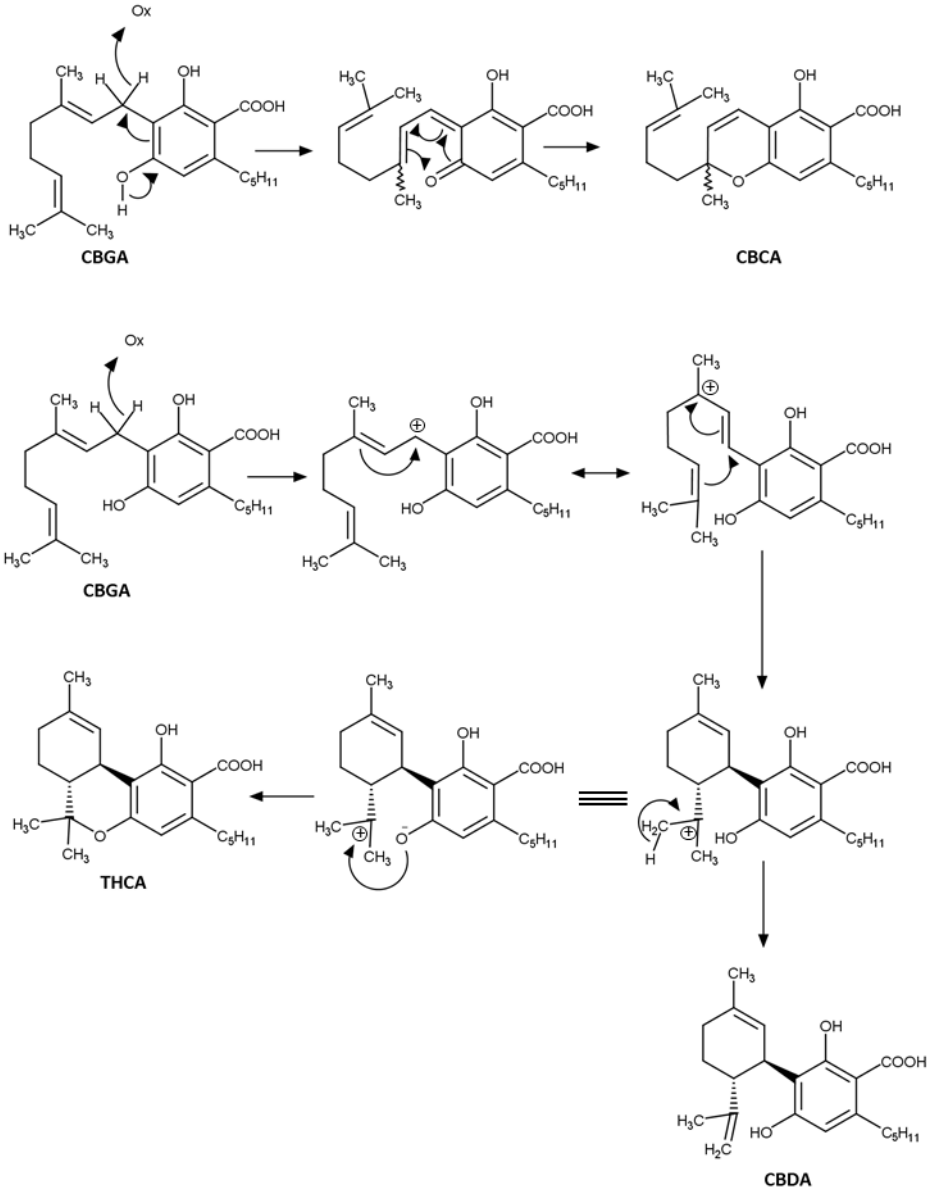


Figure 1.4.2 Biosynthesis of the major phytocannabinoids (modified according to Hanuš *et al.*, 2016 [69]). ; CBGA, cannabigerolic acid; THCA, Δ^9-tetrahydrocannabinolic acid; CBDA, cannabidiolic acid; CBCA, cannabichromenic acid.

1.5. Reconstruction of biosynthetic pathways in heterologous hosts

Engineering of metabolic pathways in alternative hosts holds a promising strategy to improve the production of high valuable secondary metabolites which are often available only in small amounts in their native producers, or to generate new-to-nature compounds optimized for a particular application [70,71]. It is now believed that 25 % of all prescribed medicines are of plant origin and as many as 11 % of the 252 medicines classified by the WHO as basic and essential are derived exclusively from flowering plants [72,73]. As a result, the large-scale production of useful plant natural products (PNPs) received much attention in recent years and their synthesis has been studied in a number of heterologous systems including microorganisms (bacteria or yeast), plant cell/organ cultures, and intact plants [74]. However, the selection of an appropriate expression system is one of the most critical steps in the development process for metabolic pathway reconstruction. Thus, each expression platform offers advantages but also disadvantages. The most commonly used bacterial expression systems produce high levels of the desired recombinant proteins, are easy to maintain, do not require expensive growth media and offer a range of different genetic tools for their molecular modifications. But transcriptional and/or post-translational processing as well as the issue with integrating large-size gene cassettes, make bacterial hosts rather useless for the reconstruction of complex biosynthetic pathways [75,76]. Especially with the view on the production of cannabinoids like THCA, CBDA and CBCA, prokaryotic hosts such as *Escherichia coli* are not suitable, as the FAD-dependent oxygenases THCAS, CBDAS and CBCAS possess several *N*-glycosylation sites which are required for proper folding of the enzymes [67,77,78]. Single-cell eukaryotic microorganisms like yeast combine the low-cost culture media and high protein production levels of prokaryotic hosts with post-transcriptional and post-translational modifications required for eukaryotic proteins. Moreover, similar to bacterial expression systems, the toolbox for genetic manipulation as well as the technologies needed for scale-up production are well developed, and the organisms are Generally Recognized As Safe (GRAS) [76,79]. However, drawbacks include hyperglycosylation of proteins, which can limit their functionality, the use of methanol as an inducer for methylotrophic yeasts such as *Pichia pastoris*, which pose a safety (fire) risk at scale, and occasionally expensive royalties [79,80].

When it is about the production of PNPs, plants offer themselves as a promising heterologous expression platform. Metabolic engineering in plants is particularly warranted as many plant biosynthetic pathways are partitioned in specific subcellular organelles, such as chloroplasts, and require co-enzymes, co-factors, post-translational modifications or regulators and often extensive manipulation of the host organism metabolism if it is not of plant origin [76,81]. In addition, plants offer the possibility of being used as expression hosts either in the form of a whole organism or as a

cell culture, each having its own advantages. While the whole plant requires minimal maintenance and produces higher quantities of target metabolites, the cell cultures yield relatively high biomass and exhibit shorter cultivation times of nine days compared to 6–9 weeks of greenhouse grown plants, as observed in transgenic tobacco plants producing geraniol, thus making cell cultures a more economical option for large scale production [76,81,82]. Moreover, plastid genome engineering has become an emerging tool for biosynthetic pathway reconstruction in plants. It enables the expression of biosynthetic pathway genes as operons, higher yields of produced proteins compared to nucleus transformants, lack of transgene silencing and low-level leakages of transgenes in pollen due to maternal inheritance [83]. Thus, studies have shown that for example transplastomic tobacco plants generated higher quantities of astaxanthin than their corresponding nuclear transformants [84,85]. However, as with the other heterologous expression platforms, there are some drawbacks in plants as well. Among other things, the relatively high cost of engineering, complex transformation protocols as well as the slow growth and reproduction rates [76]. A detailed summary of advantages and disadvantages of several heterologous expression platforms is given in Table 1.5.1.

Table 1.5.1 Advantages and disadvantages of different expression hosts for the production of heterologous proteins and secondary metabolites (adapted from Shanmugaraj *et al*., 2020, Markina *et al.,* 2020, Zhang *et al.,* 2010 and Broadway, 2012 [76,79,80,86]).

Platform	Advantages	Disadvantages	Common species
Bacteria	• Easy to manipulate • Low costs • High expression • Easy to scale up • Short turnaround time • Established regulatory procedures and approval	• Improper protein folding • Lack of post-translational modifications • Endotoxin accumulation	• *Escherichia coli* • *Bacillus subtilis*
Animal cell cultures	• Proper folding and authentic post-translational modifications • Existing regulatory approval • Absence of the cell wall	• High production cost • Expensive media and culture condition requirements • Low growth rate • Potential risk of human virus in culture	• Mammalian cells • Insect cells
Yeast	• Rapid growth and scalable • Easy to manipulate • Low costs • Post-translational modifications of recombinant proteins • Generally Recognized As Safe (GRAS) • High protein production levels	• Difficulty in cell disruption due to the thick and hard cell walls • Hyperglycosylation of proteins • Royalties can be expensive • Use of methanol as inducer is a safety (fire) hazard at scale (methylotrophic yeast)	• *Saccharomyces cerevisiae* • *Pichia pastoris (Komagataella)* • *Candida boidinii* • *Hansenula polymorpha* • *Pichia methanolica* • *Yarrowia lipolytica*
Plants	• Free from pathogen and bacterial toxin contaminants • Economical • Post-translational modification similar to mammalian systems • Host versatility: whole organism or a cell culture • Well suited for heterologous expression of metabolic pathways from other plants • Metabolic pathway can be localized in the chloroplasts	• Regulatory compliance • Limited glycosylation capacity of proteins • Complex transformation protocols • Low growth and reproduction rates	• *Nicotiana benthamiana* • *Nicotiana tabacum* • *Physcomitrella patens*

1.5.1. Tobacco as a platform for heterologous protein and secondary metabolite production

To date, several plant species have been tested for molecular farming purposes, such as the crops alfalfa, lettuce, tomato, potato and soybean [87]. However, tobacco has been the most important expression platform used in green biotechnology. Despite a traditionally negative perception due to its strong association with smoking, tobacco greatly excels other plant species because of its many unique advantages [88]. Since it is a leaf-based expression system, the need for flowering is eliminated and thus the potential for gene escape into the environment through pollen or seed dispersal is significantly reduced. Moreover, the plant is a non-food, non-feed crop, minimizing the regulatory hurdles, since the risk of plant-produced recombinant proteins entering the food chain is not given [89]. Tobacco is readily amenable for genetic modifications either by transient based expression of genes (*via* Agrobacteria or viral induction) or by stable transformation (nuclear or plastid genome) and has therefore become the main plant vehicle for recombinant protein production over the last 30 years. Thus, it is often referred as the 'white mouse' among plants [88]. By now, a variety of therapeutic proteins and vaccines prepared from transgenic tobacco plants are investigated in clinical trials. These include PRX-12 and PRX-102, therapeutic enzymes against Gaucher's disease and Fabry disease, respectively, produced in tobacco BY2 cell cultures, but also ZMapp, aiming against the Ebola virus or antibodies against HIV and influenza, produced in whole tobacco plants [90]. Even a transgenic tobacco derived COVID-19 vaccine named CoVLP from the company Medicago is currently investigated in phase 3 clinical trials [91]. Besides pharmaceuticals, also products ranging from technical enzymes and research reagents to media ingredients and cosmetic products are available [90]. However, tobacco stands out not only as a production chassis for recombinant proteins, but also as a platform for heterologous synthesis of valuable secondary metabolites. Thus, researchers succeeded in reconstructing the biosynthetic pathway of artemisinin combined with the required increase of precursor supply, by introducing 12 transgenes into tobacco plants. The aforementioned PNP is used worldwide to treat infections with multidrug-resistant strains of *Plasmodium falciparum*, the causative pathogen of malaria tropica, and is currently primarily extracted from the leaves and flowers of the annual mugwort (*Artemisia annua*) [92].

1.5.2. Assembly of biosynthetic pathways using the GoldenBraid cloning system

In addition to the lengthy process of elucidating enzymes involved in the metabolic trails of interest, the major technological challenge of synthetic plant biology concerns the construction and transfer of multigene structures into the plant genome [93]. The cloning of several transcriptional units (TUs), each including a promoter, a 5'-untranslated region (5'-UTR), a coding sequence (*cds*), a terminator as well as fusion protein or targeting sequences is often an arduous task, since the introduction of restriction sites, enabling the assembly of DNA parts within a desired vector backbone, requires extensive planning of the experimental procedure and multiple cloning steps. Thus, modular systems, such as the innovative Golden Gate technology, have been developed to simplify cloning and increase its efficiency. The utilization of type IIS restriction enzymes allowed the design of individual tailored four-nucleotide overhangs, as those endonucleases are characterized by the ability to cut the DNA (in case of BsaI and BsmBI) at a specific distance outside their recognition site. While these sites are inversely oriented and flanking the gene of interest, the expression vector contains complementary restriction sites. The recognition sites are then eliminated during the digestion process, leaving the complementary designed four-nucleotide overhangs ready for ligation. Consequently, it was possible to assemble TUs in a one-pot, one-step reaction so that restriction and ligation could be performed together [94]. Based on the Golden Gate cloning system, the modular cloning system (MoClo) was introduced allowing the construction of multigene constructs from a set of standardized genetic level-0 modules, encompassing promoters, 5'-UTRs, targeting sequences, coding sequences (CDS), and terminators. These basic modules are assembled *via* type IIS restriction enzymes into TUs within level-1 vector backbones and further into multigene constructs by fusion of up to six TUs from different level-1 targeting vectors into a level-2 plasmid. However, the utilization of three different IIS-type restriction enzymes as well as seven level-1 and level-2 plasmids and additional end-linker elements made the routine application of this standardized technique rather complex and further development towards simpler applicability was necessary [95]. Concurrently, the GoldenBraid (GB) modular cloning system was developed, which allows the theoretically infinite assembly of multipartite constructs by reducing the destination to two binary plasmids (α- and Ω-level) [96]. In its improved iteration (GB 2.0), the system was made inter-compatible with MoClo and extended by introducing a common entry vector (universal domesticator, pUPD) that allows unlimited extension of the GB library of standardized DNA building blocks. Thus, it is possible to amplify any desired coding sequence with GB primers containing approximately 20 nucleotides of the gene-specific sequence and an additional region containing the pUPD-compatible cleavage site, the relevant recognition site for type IIS endonuclease (BsmBI) and the four-nucleotide fusion sequence that determines the proper integration within the TU. In a one-pot, one-step reaction the CDS amplicon and pUPD are then cut by BsmBI and ligated appropriately using T4 DNA ligase, resulting in

a newly domesticated GB part. In order to ensure a correct final orientation according to their function within the TU, all domesticated parts, including promoters (*Pro*), signal peptides (*sp*), coding sequences, C-terminal fusion protein sequences (*ct*) and terminators (*Ter*) are flanked by a specific four nucleotide 'bar code'. After domestication, the desired GB parts are assembled into TUs within α-level plasmids through BsaI digestion and ligation in a combined reaction step. Subsequently, multigene constructs can be constructed by excising two TUs embedded in two different α-level plasmids using BsmBI and then fusing them into an Ω-level backbone. Due to the unique double loop design of the GoldenBraid vectors, the ligation of two Ω-level multigene expression cassettes within an α-level plasmid enables a theoretically infinite assembly of TUs (Figure 1.5.1). Moreover, the presence of a *lacZ*-cassette in each GB plasmid facilitates blue-white screening of transformed bacterial clones as it is replaced by the desired insert during the digestion and ligation process. Additionally, to counter-select against vectors of the alternative levels, the pUPD, α-level plasmid and Ω-level plasmid are characterized by an ampicillin, kanamycin or spectinomycin resistance, respectively [93]. Since GoldenBraid is intended to serve as a modular assembly system in plant synthetic biology, the GB 1.0 generation vectors originate from the pGreen II binary vector, which multiplies in *Agrobacterium* only in the presence of the pSoup helper plasmid, while in the updated GoldenBraid 2.0 version, the α- and Ω-level destination vectors are derived from the open source pCAMBIA binary vectors, replicating independently in the bacteria and thus simplifying the transformation process [97–99]. Recently the group of Dr. Diego Orzaez introduced the GB 3.0 iteration of the system which extends the α- and Ω- destination plasmid collection with a new set of pCAMBIA-based vectors that exhibit increased *in planta* transformation efficiency, as well as an improved GB webtool interface, enabling integration of *in silico* cloning results and experimental data [100]. Since the one-pot, one-step reaction used remains the same in GB 1.0, GB 2.0 and GB 3.0, the general abbreviation GB is used in all the subsequent texts when referring to GoldenBraid.

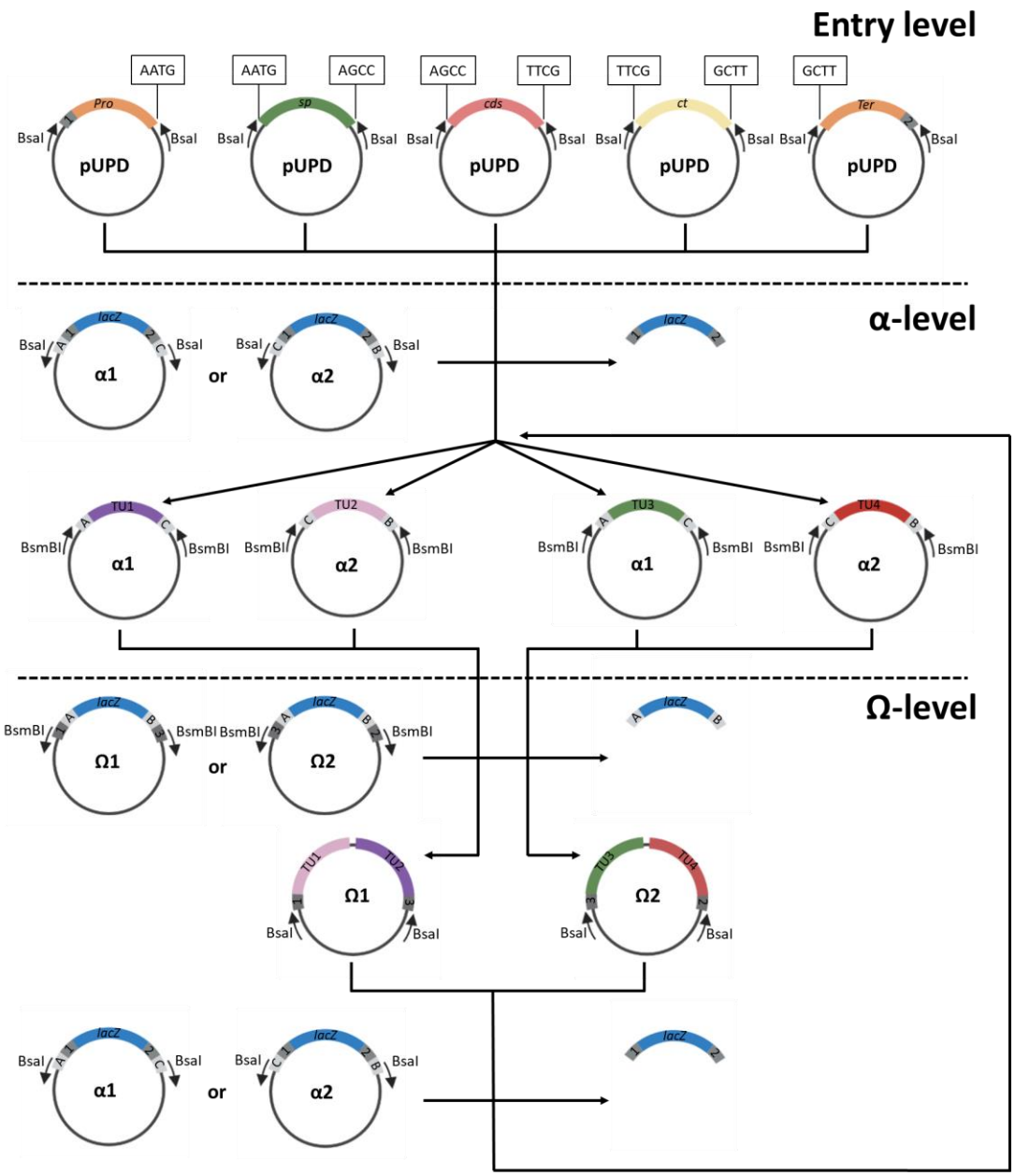


Figure 1.5.1 Schematic illustration of the GoldenBraid (GB) cloning technique (modified according to Sarrion-Perdigones *et al*., 2013 [93]). The domesticated parts, consisting of promoters (*Pro*), signal peptides (*sp*), coding sequences (*cds*), C-terminal fusion protein sequences (*ct*) and terminators (*Ter*), are flanked by specific four-nucleotide 'bar codes' to ensuring the correct final orientation within a transcriptional unit (TU). In a one-pot one-step reaction using *Bsa*I digestion and T4 ligation, assembly of TUs within α-level plasmids is performed. Further, by excising two TUs embedded in two different α-level plasmids using BsmBI and subsequent fusion into Ω-level vector backbones, multigene constructs are created. The GoldenBraid's double-loop design allows further ligation of two Ω-level multigene cassettes within an α-level plasmid, enabling theoretically an infinite assembly of TUs. In addition, each GoldenBraid plasmid contains a *lacZ* expression cassette that is replaced by the corresponding insert during the digestion and ligation process. The additional four-nucleotide fusion sites, designated as A, B, C, 1, 2 and 3 ensure the assembly of TUs as well as multigene constructs in α- and Ω-level destination plasmids.

Chapter 2

Scope of the thesis

Cannabis sativa has recently gained resurgence of interest due to its versatile application possibilities: Among other things, because of its use for the production of bioplastics utilized in automotive sector, but mainly because of its enormous pharmaceutical potential, for example, for the relief of chronic pain or depression [101,102]. Thus, in addition to terpenes and phenolic compounds, the phytochemical portfolio of *C. sativa* also includes the so-called phytocannabinoids, which form a large group of constituents with more than 100 different compounds in ten subclasses [46–49]. However, in most Western countries the use of *C. sativa* as a source for the production of pharmaceutical cannabinoids is restricted, since the plant is treated as an illicit drug due to the main psychoactive ingredient Δ^9-tetrahydrocannabinol (THC) [103]. On this account, there is a necessity to develop alternative strategies, like biotechnological synthesis of defined *Cannabis*-derived compounds, to detach targeted cannabinoid production from a THC producing plant.

Since engineering of cannabinoid biosynthesis is furthermore thought to enable the flux of the pathway towards desired target products, for instance the formation of pharmacologically attractive new-to-nature as well as non-psychotropic cannabinoids, heterologous production has already been investigated mainly in *Saccharomyces cerevisiae* and *Pichia pastoris* [60,62]. Besides yeasts, the utilization of other production chassis such as tobacco could also be a promising alternative, as it is the most commonly used model plant for biotechnological production of proteins and integration of whole biosynthetic pathways by transient and stable transformation is already described [92]. Moreover, the ability of plants to compartmentalize could provide an opportunity to achieve higher titers of the desired substances, as toxic compounds like cannabigerolic acid (CBGA) could be stored in trichomes of tobacco, for example, as it is the case in *C. sativa* [52]. In addition, plant cells possess the capability to carry out various modifications of their produced metabolites, such as oxidation, reduction, hydroxylation, esterification or glycosylation, and thus further increase the product spectrum of potentially therapeutically active cannabinoids [104]. The effects of glycosylation of cannabinoids could also improve their water solubility, leading to possibly less toxicity through storage in the vacuole [105]. Furthermore, plant cell-based production of pharmaceuticals offers superior benefits to public safety, since they reduce the risk of transmitting human pathogens [106]. Consequently, this study deals on the one hand with the evaluation of a plant-based transient and stable expression system for the production of cannabinoids and on the other hand with the characterization and in-depth investigation of late biosynthetic enzymes, to get a better insight into the production of different cannabinoids. The scope of this thesis is further described in detail:

i. The genes encoding proteins involved in cannabinoid biosynthesis have to be introduced into the GoldenBraid cloning system, in order to facilitate the efficient and flexible assembly of the required genetic parts [93]. Moreover, by addition of various targeting sequences, production of the enzymes in different subcellular localizations has to be investigated for optimal integration of the biosynthetic pathway into the host plant, enabling the reconstruction of whole cannabinoid synthesis *via* Agrobacterium mediated transient and stable transformation of *Nicotiana sp.*.

ii. The enzymes Δ^9-tetrahydrocannabinolic acid synthase (THCAS) and cannabidiolic acid synthase (CBDAS) were widely studied in the last years in terms of their reaction mechanism as well as their evolutionary relationship [64,65,67,107–109]. The cannabichromenic acid synthase (CBCAS), on the other hand, received less attention [63,110]. However, since THCAS and CBCAS share a high sequence identity of 93 %, the focus of the work is to examine the CBCAS more closely. Therefore, site-directed mutagenesis of CBCAS will be performed to obtain information which amino acids are responsible for the cyclization specificities that determine the synthesis of the respective product.

Chapter 3

General materials and methods

3.1. Bacterial strains and growth conditions

Escherichia coli TOP10 cells (Thermo Fisher Scientific Inc., Waltham, USA), used for cloning, and *Agrobacterium tumefaciens* strain EHA105 (ICON Genetics, Halle, Germany), applied in transient transformation of plants, were cultivated in liquid lysogeny broth (LB; [111]) under constant agitation (170 rpm) at 37 °C and 28 °C, respectively. For selection, the medium was supplemented with ampicillin (100 µg/mL), kanamycin (50 µg/mL) and spectinomycin (50 µg/mL) and, in case of *A. tumefaciens* strains, additionally with rifampicin (50 µg/mL; EHA105) as well as 100 µM of 3,5-dimethoxy-4-hydroxyacetophenone (acetosyringone; Sigma-Aldrich, St. Louis, MO, USA). For cultivation on LB agar plates, bacterial suspensions were uniformly spread on the solid medium with proper antibiotic supplementation. In case of *E. coli* TOP10 cells, additionally 40 µg/mL of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal; Applichem, Darmstadt, Germany) were added for blue/white selection.

3.2. Molecular cloning methods

3.2.1. Polymerase chain reaction (PCR)

To amplify target sequences from plasmid DNA, polymerase chain reaction was performed. Moreover, genes were extended by addition of 5'- and 3' fusion sites as well as BsmBI recognition sites for integration into the GoldenBraid system. The mixture was set up as follows in a total reaction volume of 50 µL: 7.5 ng – 500 ng of template DNA, 1× Pfu buffer (20 mM Tris (pH 8.8), 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 1 % (v/v) Triton-X 100, 1 mg/mL BSA), 0.2 mM of each dNTP, 0.2 mM of proper forward and reverse primer (see 4.1.3, 5.1.3, 6.1.3 and 7.1.2) as well as 1 µL of Pfu polymerase (in house purified). After the amplification, PCR samples were analyzed by agarose gel electrophoresis (see 3.2.7).

Step	Temperature	Duration	
Initial denaturation	95 °C	5 min	
Denaturation	95 °C	30 sec	35 cycles
Annealing	60 °C	30 sec	
Elongation	72 °C	Variable*	
Final elongation	72 °C	10 min	

*depending on the length of the amplicon

3.2.2. Purification of PCR products

For purification of amplified DNA sequences, the Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, Wisconsin, USA) was used according to the manufacturer's instructions.

3.2.3. Assembly of transcriptional units utilizing the GoldenBraid cloning system

The GB assemblies were performed as described by [93]. In case of new domestications, the 10 µL of the GB reaction mixtures contained 75 ng of PCR products or synthesized DNA, 75 ng of the pUPD vector, 1× T4 DNA ligase buffer (Promega), 1 mM DTT (Carl Roth), 1 µL (3 u/µL) of T4 DNA ligase (Promega) and 1 µL (10 u/µL) of BsmBI (Thermo Fisher Scientific Inc., Waltham, USA). In case of the assembly of transcriptional units, the GB reaction mixture contained 75 ng of each plasmid harboring relevant DNA parts, 75 ng of the appropriate destination vector, 1× T4 DNA ligase buffer, 1 mM DTT, 1 µL (3 u/µL) of T4 DNA ligase and 1 µL (10 u/µL) of BsaI (New England Biolabs, Ipswich, MA, USA) or BsmBI were used for the reaction setup. The GB assemblies were performed in 25- or 50-cycle digestion and ligation reactions for 2 min at 37 °C and 5 min at 16 °C, respectively.

3.2.4. Transformation of competent *E. coli* TOP10 cells

For the amplification of plasmids, stock aliquots of chemically competent *E. coli* TOP10 cells were thawed on ice and 10 µL of the relevant ligation mixture were gently transferred into the 50 µL suspensions. The cells were incubated on ice for 30 min, followed by 1 min heat shock at 42 °C. After that, the bacteria were incubated on ice for another 2 min and 500 µL of liquid LB medium without antibiotics. Subsequently, the cells were regenerated at 37 °C and 750 rpm for 1 h and were plated on LB plates with the required antibiotics (see 3.1). The following incubation took place at 37 °C overnight.

3.2.5. Isolation of plasmid DNA and quantification of nucleic acid content

Plasmid DNA from transformed *E. coli* TOP10 cells was extracted by means of the Plasmid Miniprep Kit I (VWR International GmbH, Darmstadt, DE) according to the manufacturer's protocol. Afterwards, samples were quantified by photometric measurements using the NanoPhotometer® NP80 (Implen GmbH, Munich, Germany).

3.2.6. Restriction analysis

To confirm the correct assembly of the desired DNA fragments in the isolated plasmids, control digestion with restriction enzymes was performed. Therefore, 500 ng of plasmid DNA with 1× Fast digest buffer (Thermo Fisher Scientific Inc.) and 1 µL (3 u/µL) of an appropriate Fast Digest restriction enzyme (Thermo Fisher Scientific Inc.). Restriction was performed for 15 min at 37 °C and analyzed *via* agarose gel electrophoresis (see 3.2.7).

3.2.7. Agarose gel electrophoresis

According to their size, DNA fragments were separated on a 1.0 % (w/v) agarose gel. Therefore, agarose was melted in 1× TAE buffer (40 mM Tris, 20 mM acetic acid and 1 mM EDTA, pH 8.2) and supplemented with 5 µL of the DNA gel stain Nancy-520 (Sigma-Aldrich). DNA samples were mixed with 6× DNA loading dye (Thermo Fisher Scientific Inc.) prior electrophoresis. Separation was then performed in 1× TAE buffer using a voltage of 120 V. Finally, DNA fragments were visualized under UV light by means of the Gel iX Imager (INTAS Science Imaging Instruments GmbH, Göttingen, DE).

3.2.8. Sequencing of DNA samples

Newly domesticated GoldenBraid (GB) constructs within the pUD vectors were sequenced by Eurofins Genomics GmbH using the following primers:

No.	Name	Nucleotide sequence 5'–3'
1	M13 uni (-21)	TGT AAA ACG ACG GCC AGT
2	M13 rev (-29)	CAG GAA ACA GCT ATG ACC

3.2.9. Transformation of competent *A. tumefaciens* EHA105 cells

Chemically competent *A. tumefaciens* cells were transformed by heat shock. Therefore, a 100 µL aliquot was thawed on ice. The bacteria were supplemented with 375 ng of plasmid DNA and the mixture was incubated on ice for 5 min. Subsequently, the samples were frozen in liquid nitrogen for additional 5 min, followed by incubation at 37 °C for 5 min. After the addition of 500 µL of liquid LB medium, the cells were regenerated for 3 h at 28 °C and 750 rpm. Afterwards, the bacteria were transferred on LB plates supplemented with appropriate antibiotics (see 3.1). Finally, the cells were incubated for 48 h at 28 °C.

3.2.10. Colony screen PCR

To check if *A. tumefaciens* have incorporated the desired plasmids after transformation, several colonies were picked and analyzed by colony screen PCR. Therefore, each colony was dissolved in 20 µL of sterile water. 10 µL of this bacterial solution were then mixed with 10 µL of Taq PCR mix resulting in a final reaction composition of 1× Taq buffer, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, 0.1 µM appropriate forward and reverse primer (see 4.1.3, 5.1.3, 6.1.3 and 7.1.2) and 0.5 u of recombinant Taq DNA Polymerase (Thermo Fisher Scientific Inc.). Applied PCR conditions are described below.

Step	**Temperature**	**Duration**	
Initial denaturation	95 °C	5 min	
Denaturation	95 °C	30 sec	35 cycles
Annealing	55 °C	30 sec	
Elongation	72 °C	Variable*	
Final elongation	72 °C	10 min	

*depending on the length of the amplicon

Afterwards, DNA amplicons were analyzed by agarose gel electrophoresis (see 3.2.7) and positive tested clones were subsequently used to inoculate overnight cultures.

3.3. Expression of transgenes in plants

3.3.1. Transient transformation *Nicotiana benthamiana* plants *via A. tumefaciens* mediated gene transfer

Greenhouse-grown *N. benthamiana* plants (cultivated at 21 ± 2 °C under 70 % humidity and 16 h light period) were used for *Agrobacterium*-mediated transient transformation. Overnight cultures of *A. tumefaciens* EHA105 cells harboring the genes of interest were centrifuged at 1,500 ×*g* for 5 min. EHA105 cells carrying pBIN61–P19 were used as negative controls. Subsequently, the obtained cell pellet was resuspended in the infiltration buffer (10 mM MES, 10 mM $MgSO_4$, 100 µM acetosyringone, pH 5.5) to the final OD_{600} of 0.7−1.0. After an incubation period of 2 h at 28 °C, the bacterial suspensions were infiltrated into the leaves of four-week-old plants (on their abaxial side) using a syringe without a needle. When co-infiltrated, the individual *Agrobacterium* suspensions were mixed in equal ratios to keep the concentration of all relevant constructs constant within each experiment. After infiltration, the plants were incubated for five days at 21 ± 2 °C in a climatic chamber under 60 % humidity and 12 h of illumination. Infiltrations were performed in three to five biological replicates.

3.3.2. Fluorescence microscopy

Successful production of recombinant was either ascertained either by fluorescence microscopy (Axioskop 40, Zeiss, Oberkochen, Germany) equipped with a HBO 50 /AC bulb (Osram, Munich, Germany) and filtered by the Zeiss filter set 38, or by the BZ-X710 All-in-One Fluorescence Microscope (KEYENCE, Tokyo, Japan).

3.3.3. Extraction of metabolites

For the extraction of metabolites from plant samples, leaf tissue was frozen in liquid nitrogen and ground with a pestle and a mortar. Subsequently, 150 mg of fresh leaf powder were homogenized by sonication in 200 µL of 80 % (v/v) MeOH for 30 min at RT. Finally, the extracts were purified two times from solid particles by centrifugation at 17,000 ×*g* for 10 min and 4 °C and subjected to HPLC−MS analysis (see 3.5.1).

3.4. Protein biochemical methods

3.4.1. Extraction of total soluble proteins (TSP)

Total soluble protein (TSP) samples were retrieved from previously frozen leaf disks (approximately 100–150 mg) by grinding them in 500 µL of the extraction buffer (0.1 M Tris-HCl buffer pH 8.0, 5 mM EDTA-Na^2, 0.1 M NaCl, 10 mM β-mercaptoethanol) on ice. For PNGase F and PNGase A assays (see 6.1.5.1) the TSP content was quantified by means of the Bradford assay [112]. Therefore, 5 µL of the protein extract was mixed with 1 mL of the Coomassie reagent (100 mg Coomassie Brilliant Blue G250 dissolved in 50 mL of EtOH and 100 mL of H_3PO_4 (85 %), diluted to 1,000 mL with H_2O and filtered) and its optical density was recorded at 595 nm after 2 min of incubation. The standard curve was constructed using BSA at concentrations in the range of 25–200 µg/mL.

3.4.2. Isolation of histidine tagged proteins from plants

After transformation, leaves were cut from plants and homogenized with 2.5 mL Buffer A (50 mM NaH_2PO_4, 10 mM Tris-HCl pH 8.0, 200 mM NaCl) per gram leaf material. Subsequently, 1 mM 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF, Sigma-Aldrich) was added to the buffer. After two times 15 min centrifugation and one centrifugation step of 30 min at 14,600 ×*g* and 4 °C, the supernatant was filtered through a ROTILABO® syringe filter (PVDF, 0.45 µm; Carl Roth). Afterwards, it could be used for protein purification by immobilized metal ion affinity chromatography (IMAC). Protino Columns (Macherey-Nagel GmbH & Co. KG, Düren, Germany) containing 1.5 g Ni-TED Resin (Macherey-Nagel GmbH & Co. KG) were used and gravity-purification was performed. The Column was equilibrated with 12 ml of Buffer A. Afterwards, the crude lysate was applied onto the column. Flow through was collected and stored till analysis at 4 °C. After binding, the resin was washed two times with 12 ml Buffer A. The elution of polyhistidine-tagged proteins took place in three steps using 9 mL Buffer A containing increasing concentrations of imidazole (50 mM, 150 mM and 250 mM). All collected fractions were analyzed for the presence of the proteins of interest by SDS−PAGE and subsequent Coomassie Brilliant Blue staining.

3.4.3. Desalting and concentration of purified proteins

After the protein purification corresponding elution fractions were combined, desalted and concentrated with Vivaspin® Turbo 15 ultrafiltration device (Sartorius AG, Göttingen, Germany). The protein solution was concentrated to a total volume of 2 mL, followed by three times washing with 9 mL of prenyltransferase buffer (50 mM Tris-HCl, 10 mM NaCl, pH 7.5) in case of NphB or

THCAS/CBCAS/CBDAS buffer (100 mM trisodium citrate, pH 5.5) in case of THCAS, CBCAS and CBDAS. After the final concentration of the sample to 500−1,000 µL, protein quantity was determined with 'BCA Protein Assay Kit' (Thermo Fisher Scientific Inc.). The purified proteins were analyzed by SDS−PAGE and visualized by silver nitrate staining. In order to store protein preparations at -20 °C, proteins were diluted in 25 % (v/v) glycerol to a final concentration of 0.5 mg/mL.

3.4.4. Bicinchoninic acid (BCA) assay

Determination of protein concentration was accomplished by means of the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific Inc.). The procedure was performed according to the manufacturer's instructions using bovine serum albumin (BSA) in the range of 0.025-2 mg/mL as a protein standard. Protein dilutions of 1:2, 1:4 and 1:8 in prenyltransferase or THCAS/CBCAS/CBDAS buffer were measured in duplets right after desalting and concentrating (see 3.4.3).

3.4.5. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS−PAGE)

Protein samples were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS−PAGE). For initial concentration of proteins prior entering the resolving gel, a 3.75 % stacking gel was utilized. Dependent on the size of the proteins, a 10 % or 13 % resolving gel was used for separation. The polyacrylamide gels were prepared as follows:

Resolving gel 13 %	**Resolving gel 10 %**	**Stacking gel 3.75 %**
2.17 mL acrylamide solution (30 %)	1.66 mL acrylamide solution (30 %)	0.4 mL acrylamide solution (30 %)
1.25 mL Tris-HCl (1.5 M, pH 8.8)	1.25 mL Tris-HCl (1.5 M, pH 8.8)	0.94 mL Tris-HCl (0.5 M, pH 6.8)
1.47 mL ddH_2O	1.97 mL ddH_2O	1.7 mL ddH_2O
50 µL SDS (10 % (w/v))	50 µL SDS (10 % (w/v))	30 µL SDS (10 % (w/v))
50 µL ammonium persulfate	50 µL ammonium persulfate	30 µL ammonium persulfate
10 µL TEMED	10 µL TEMED	5 µL TEMED

Subsequently, the gel was overlaid with SDS−PAGE running buffer (192 mM glycine, 25 mM Tris, 10 % (w/v) SDS) and loaded with protein samples. Electrophoretic separation was performed for 1 h at 40 V and further 2.5 h at 100 V. Afterwards, the gel was either stained with Coomassie Brilliant Blue (see 3.4.6.1) and silver nitrate (see 3.4.6.2) or was used for further Western blot analysis.

For preparation of protein samples, 25 µL of total soluble protein extract were mixed with an equal volume of 2× SDS-loading dye (100 mM Tris-HCl (pH 8.3), 20 % (v/v) glycerol, 4 % (w/v) SDS, 0.2 % (w/v) bromphenol blue) and boiled at 95 °C for 10 min. In case of purified proteins, 0.5 µg of proteins were dissolved in 25 µL of ddH_2O and prepared as described before. To determine the size of separated proteins, the PageRuler Prestained Protein Ladder (10 to 180 kDa; Thermo Fisher Scientific Inc.) was applied to the polyacrylamide gel.

3.4.6. Visualization of proteins separated *via* SDS−PAGE

3.4.6.1. Coomassie Brilliant Blue staining

The polyacrylamide gels were soaked in Coomassie staining solution (50 % (v/v) MeOH, 10 % (v/v) acetic acid, 0.5 % (w/v) Coomassie Brilliant Blue G 250) for 15 min on an orbital shaker. After visualization the gels were destained for 20 min in destaining solution (30 % (v/v) EtOH, 10 % (v/v) acetic acid), followed by a final destaining step performed overnight.

3.4.6.2. Silver nitrate staining

Compared to classical Coomassie Brillant Blue staining the sensitivity of protein detection on polyacrylamide gels using silver nitrate is typically increased 50−100 times [113], which is why this method is particularly suitable for detecting minor contaminants within a protein preparation.

Firstly, the gel was incubated for 30 min in fixing solution (50 % (v/v) MeOH, 12 % (v/v) acetic acid, 0.02 % (v/v) formaldehyde) followed by two washing steps in 50 % (v/v) EtOH for 10 min. Subsequently, the gel the gel was pre-treated for 1 min in 0.8 mM sodium thiosulfate solution to enhance sensitivity of the coloration. Afterwards, two washing steps with ddH_2O were performed (approximately 20 sec per step) prior adding the coloration solution (15 mM silver nitrate, 0.028 % (v/v) formaldehyde) for 10 min. After two more washing steps in ddH_2O, the gel was incubated in development solution (0.56 M sodium carbonate, 5.65 µM sodium thiosulfate, 0.02 % (v/v) formaldehyde) until the desired signal intensity has been reached. Finally, two washing steps in ddH_2O were performed and the gel was fixed for 5 min in fixation solution (50 % (v/v) MeOH, 12 % (v/v) acetic acid).

3.4.6.3. Western blot analysis

After SDS−PAGE (see 3.4.5), the proteins were transferred to Roti−PVDF membranes (Carl Roth, Karlsruhe, Germany) and subjected to Western blot analysis. For detection, primary his−probe mouse monoclonal IgG antibodies, primary GFP mouse monoclonal IgG antibodies and goat−anti mouse horseradish peroxidase coupled secondary antibodies (Santa Cruz Biotechnology, Heidelberg, Germany) were applied. The proteins were then visualized by chemiluminescence substrate (CheLuminate−HRP PicoDetect kit; Applichem, Darmstadt, Germany) according to the manufacturer's protocol.

3.5. Analytical methods

3.5.1. Reverse−phased liquid chromatography and mass spectrometry (HPLC−MS)

For HPLC−MS analysis, the 1260 Infinity HPLC system (G4225A high performance degasser, G1312B binary pump, G1329B autosampler, G1316C column thermostat, G4212B diode array detector; Agilent, Santa Clara, CA, USA) connected to the 6120 Quadrupole mass spectrometer (Agilent) was used. Five to ten µL of samples were injected onto the analytical LC–MS column (Poroshell 120SB−C18, 3.0×150 mm, 2.7 µm; Agilent). The flow rate was set to 0.5 mL min^{-1} with the mobile phase consisting of ddH_2O/0.1 % formic acid (A) and acetonitrile/0.1 % formic acid (B). The analysis was performed with the following gradient (% B): 0.0–2.0 min, 20; 2.0–8.0 min, 20 to 100; 8.0–12.0 min, 100. After 12 min, the column was re−equilibrated with 20 % B for 5 min. Metabolites were photometrically detected at a wavelength of 260 nm as well as in the negative electrospray ionization (ESI) mode with a full scan MS experiment (m/z 150–800). Further investigation was done in negative selected ion monitoring (SIM) with selected m/z of 223.2, 359.2, 385.2 or 521.2 (detection of olivetolic acid, cannabigerolic acid, olivetolic acid glucoside or cannabigerolic acid glucoside, respectively). Nitrogen was used as a nebulizing (35 psi) and drying gas (350 °C, 12 L min^{-1}). The capillary voltage was set to 3 kV.

Chapter 4

Production of olivetolic acid in transiently transformed *N. benthamiana*

4.1. Materials and methods

4.1.1. Chemicals

Olivetolic acid (OA; Santa Cruz Biotechnology) served as authentic standard as well as substrate in feeding experiments. Hexanoic acid, utilized in feeding experiments, was purchased from Carl Roth (Karlsruhe, Germany).

4.1.2. Plasmids and genetic material

The GoldenBraid (GB) −modified genes encoding for acyl−activating enzyme 1 (AAE1; GenBank accession no: AFD33345.1), olivetol synthase (OLS; GenBank accession no: AB164375.1) and olivetolic acid cyclase (OAC; GenBank accession no: AFN42527.1) originating from *C. sativa* were synthesized by Integrated DNA Technologies (Coralville, IA, USA). Furthermore, the basic GB parts, encompassing the cauliflower mosaic virus (CaMV) 35S promoter (P35S; GB0030), the nopaline synthase terminator (TNos; GB0037), the yellow fluorescent protein (YFP; GB0024), the green fluorescent protein (GFP; GB2237) and the GB destination vectors (pUPD, pDGB2α1, pDGB2α2, pDGB3α1, pDGB3α2, pDGB2Ω1, pDGB2Ω2, pDGB3Ω1 and pDGB3Ω2) were provided by the lab of Dr. Diego Orzáez from The Institute for Plant Molecular and Cellular Biology (IBMCP) of the Polytechnic University of Valencia (UPV) and the Spanish National Research Council (CSIC), Valencia, Spain. Moreover, the pBIN61 plasmid harboring the P19 suppressor of gene silencing was kindly provided by Prof. Sir David Baulcombe from the Cambridge University, UK [114]. Finally, the signaling peptide sequence of calreticulin (er) from *Nicotiana sp*. (GenBank accession no: XM_009806292.1), the chloroplast transit peptide sequence (*cp*), a consensus of dicot sequences, deriving from the pICH20030 plasmid (ICON Genetics, Halle, Germany) – utilized to afford chloroplast localization, as well as a (CaMV) 35S promoter with an incorporated ATG start codon (P35S ATG) and a TNos featuring a 8×his tag (8×his:TNos), were adapted to the GB cloning system in previous projects and were kindly provided by Dr. Sabine Fräbel [115].

4.1.3. Oligonucleotides

All oligonucleotides used in this study were synthesized by Eurofins Genomics GmbH (Ebersberg, Germany). 4-nt overhangs generated by annealing of oligonucleotides for incorporation into the GoldenBraid system are highlighted in bold.

No.	Name	Nucleotide sequence 5'–3'
3	SeqCsAAE1for	TGC CAA TGC ATG TGG ATG CTG TGG
4	pOAC cPCR101	GCA GTG AAG CAT TTG ATT GTA
5	pOAC cPCR201	CGT GGT GTG TAG TCA AAA ATG
6	pOLS cPCR101	TGA GTT TCC CGA CTA CTA CT
7	pOLS cPCR201	CCC ACT AGT AAT TCG AGG TC
8	pAAE1 cPCR101	CTG GTG AAG CAT CTA ATG TAG
9	pAAE1 cPCR201	CAA AGG TGG CAC TCC AAT AG
10	pGFP G101	ATA TCG TCT CAC TCG **AAT G**GT GAG CAA GGG CGA GGA GC
11	pGFP H203	GCG CCG TCT CAC TCG **GGC T**CC GTA CAG CTC GTC CAT GCC GT

4.1.4. Feeding experiments

4.1.4.1. Feeding with olivetolic acid

Wild-type *N. benthamiana* plants were infiltrated with 0.5 mM olivetolic acid (solved in infiltration buffer (10 mM MES, 10 mM $MgSO_4$) and were harvested 1 h, 24 h, 48 h, 72 h, 96 h and 168 h post infiltration. Plants solely infiltrated with infiltration buffer were used as negative control. Finally, metabolites were extracted as described in 3.3.3.

4.1.4.2. Feeding with hexanoic acid

Transiently transformed *N. benthamiana* plants were fed with 4 mM of hexanoic acid (solved in infiltration buffer (10 mM MES, 10 mM $MgSO_4$) four days after initial infiltration with agrobacteria harboring *aae1*, *ols* and *oac*. Transiently transformed plants solely infiltrated with infiltration buffer were used as negative control. The plants were harvested after 24 h of incubation and metabolites were extracted as described in 3.3.3.

4.2. Results and discussion

4.2.1. Molecular cloning of genes involved in the biosynthesis of olivetolic acid

In order to take the first step towards heterologous production of cannabinoids in tobacco, it was necessary to establish plants that can provide precursors for the actual cannabinoid production. While plants are already able to synthesize geranyl pyrophosphate (GPP) for monoterpene production *via* the plastidic 2-C-methyl-D-erythritol-4-phosphate (MEP) and small amounts of hexanoic acid have been found in tobacco, essential for the biosynthesis of olivetolic acid (OA), the initial focus was on the integration of pathway genes up to OA production [56,57,116]. Therefore, the coding sequences for acyl activating enzyme 1 (AAE1), olivetol synthase (OLS) and olivetolic acid cyclase (OAC) were procured from Integrated DNA Technologies harboring the appropriate fusion sites (AGCC and GCAG) and BsmBI recognition sites for incorporation into the GoldenBraid (GB; see 3.2.3) cloning system. For the domestication of *aae1*, *ols* and *oac*, the obtained gene sequences were ligated (*via* the BsmBI-generated cleavage sites) into the universal domesticator (see 3.2.3) and transformed into chemically competent *Escherichia coli* TOP10 cells (see 3.2.4), followed by isolation of the amplified plasmids and verification by sequencing (see 3.2.8; oligonucleotides 1-3).

The pUDs harboring the verified coding sequences (CDS) were then utilized to generate different transcriptional units (TUs). Therefore, the CDS were ligated to the cauliflower mosaic virus 35S promoter (P35S) fused with the 5-UTR (omega leader) of tobacco mosaic virus and the nopaline synthase gene terminator from *Agrobacterium tumefaciens* featuring a 8×his tag (8×his:TNos). In addition to the importance of general availability of precursor molecules in the chosen host organism, the choice of the localization of the biosynthetic pathway to be introduced also plays an important role regarding to the precursor supply, since spatial distribution of metabolic pathways in eukaryotes is much more complex than for example in bacterial hosts [117]. Consequently, to investigate the production of the enzymes in different subcellular localizations, the 5´-end of *aae1*, *ols* or *oac* were fused either with the chloroplast transit peptide sequence (cp) or the ER/apoplast signal peptide sequence (er). For cytosolic production, the 5´-end of the coding sequences was directly fused to the ATG start codon following the 5´-UTR regulatory sequences (P35S ATG). For easy visualization of cellular localization by UV microscopy, also AAE1, OLS and OAC variants with a C–terminally fused YFP were cloned for all subcellular localizations. Furthermore, a variant of OLS with an N–terminally fused GFP was constructed. For this purpose, fusion overhangs were incorporated into GFP by PCR (see 3.2.1) utilizing the appropriate primer pair (see 4.1.3; oligonucleotides 10 and 11), followed by purification the amplified product (see 3.2.2), domestication and cloning with the P35S ATG, *ols* and 8×his:TNos into an α-level plasmid (see 3.2.3). Subsequently, competent *E. coli* TOP10 cells were transformed with the GB reaction mixtures *via* the heat shock

method and uniformly spread on solid LB medium with proper antibiotics (according to the selection marker genes of the particular destination vector) as well as X-Gal for blue-white screening (see 3.1). After cultivation of bacterial transformants, plasmid DNA was extracted (see 3.2.5) and all assemblies were confirmed by restriction analysis with appropriate enzymes (see 3.2.6) on an agarose gel (see 3.2.7). For transient transformation of *Nicotiana benthamiana* plants, the desired constructs were introduced into chemically competent *A. tumefaciens* EHA105 cells (see 3.2.9). Afterwards, the bacteria were transferred on LB agar plates supplemented with relevant antibiotics followed by incubation for 48 h at 28 °C. The gene transfer was confirmed by colony PCR (see 3.2.10) using the matching primer pairs (see 4.1.3; oligonucleotides 4-9). Finally, glycerol stocks of overnight cultures were prepared and stored at -80 °C. Figure 4.2.1 illustrates the expression constructs generated for this part of the study.

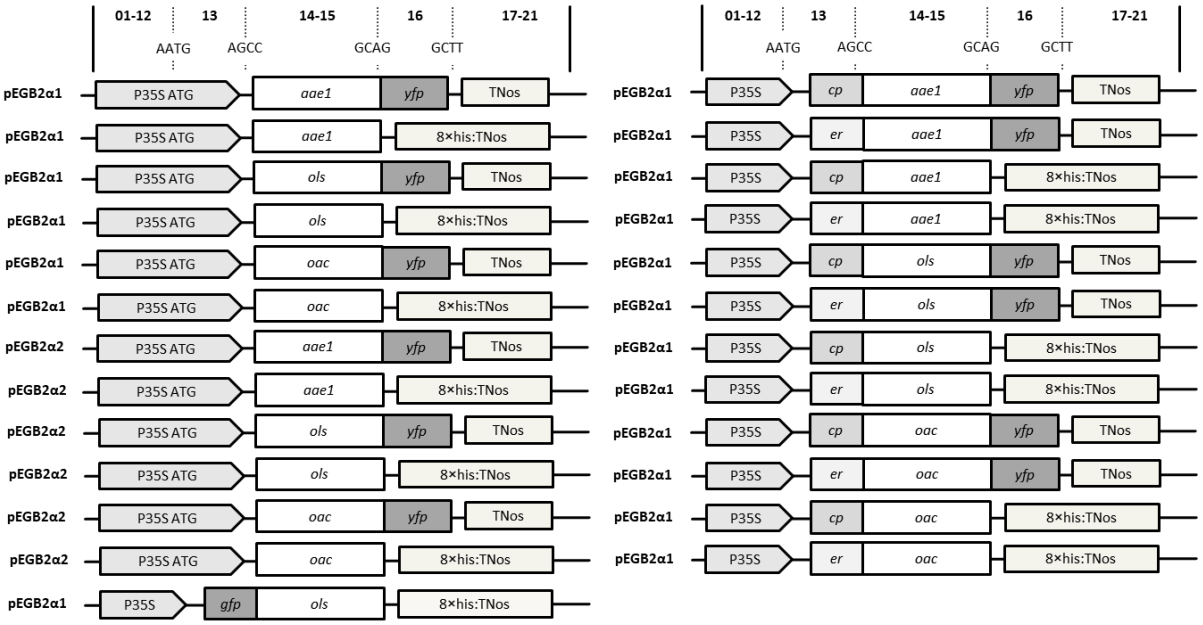


Figure 4.2.1 (A) Schematic representation of the generated AAE1, OLS and OAC constructs utilizing the GoldenBraid cloning technique. The capital letters show the four−nucleotide overhangs ensuring correct final orientation within the transcriptional unit (TU), while the numbers above the scheme represent standard GoldenBraid classes within the TU structure [93]. P35S, cauliflower mosaic virus (CaMV) 35S promoter; P35S ATG, cauliflower mosaic virus (CaMV) 35S promoter with an integrated start codon ensuring cytosolic localization; TNos, nopaline synthase terminator; 8×his:TNos, nopaline synthase terminator comprising an 8×his-tag; cp, chloroplast targeting sequence; er, ER/apoplast signal peptide. Boxes are not drawn to scale.

4.2.2. Expression of *aae1*, *ols* and *oac* in *N. benthamiana* plants

Following the successful cloning of the different transcriptional units, it was necessary to evaluate whether and in which different compartments of the plant cells the proteins could be produced. Therefore, *N. benthamiana* plants were transiently transformed with Agrobacteria harboring the various expression constructs and the gene coding for the P19 protein to avoid post-transcriptional gene silencing [114] (see 3.3.1). After five days post inoculation the plants were harvested and either analyzed by fluorescence microscopy or utilized for extraction of total soluble proteins (see 3.4.1), followed by Western blot detection (see 3.4.6.3). In case of cytosolic localized proteins, fluorescence signals could be observed for each of the three enzymes (Figure 4.2.2 A). Moreover, the production of proteins could be additionally confirmed by Western blot analysis through comparison of the protein sizes detected with their calculated ones (OAC:YFP – 40 kDA; AAE1-YFP – 95 kDa; GFP:OLS:8×his – 70 kDa; OAC:8×his – 14 kDa; Figure 4.2.2 B). For proteins targeted to the chloroplasts, it was possible to detect cp:AAE1:YFP, cp:OAC:YFP and cp:OLS:YFP either protein biochemically by Western blot analysis or by fluorescence microscopy, whereas in case of the apoplast localized AAE1, OLS and OAC only the OAC:YFP variant was detected by Western blot analysis and microscopically (Figure 4.2.2 B; Figure S4.3.1). However, the absence of protein bands in the Western blot detection does not necessarily mean that no protein is present, but that other factors, such as too low a concentration of applied protein or masking of the antigen by the nonfat dry milk used, may also play a role [118,119]. Nevertheless, since GPP as a substrate plays a crucial role for the synthesis of CBGA, localization of OA biosynthesis in the apoplast would not be the best choice anyway, since, as already mentioned, the production of GPP takes place *via* the MEP pathway localized in chloroplasts [56]. Therefore, targeting AAE1, OLS and OAC to the chloroplasts would be the most favorable. However, it has also been reported, that GPP could be transported across the envelope membrane of chloroplasts, enabling metabolic cross talk between cytosol and plastids, which would also make the production of OA in the cytosol an alternative worth mentioning [120].

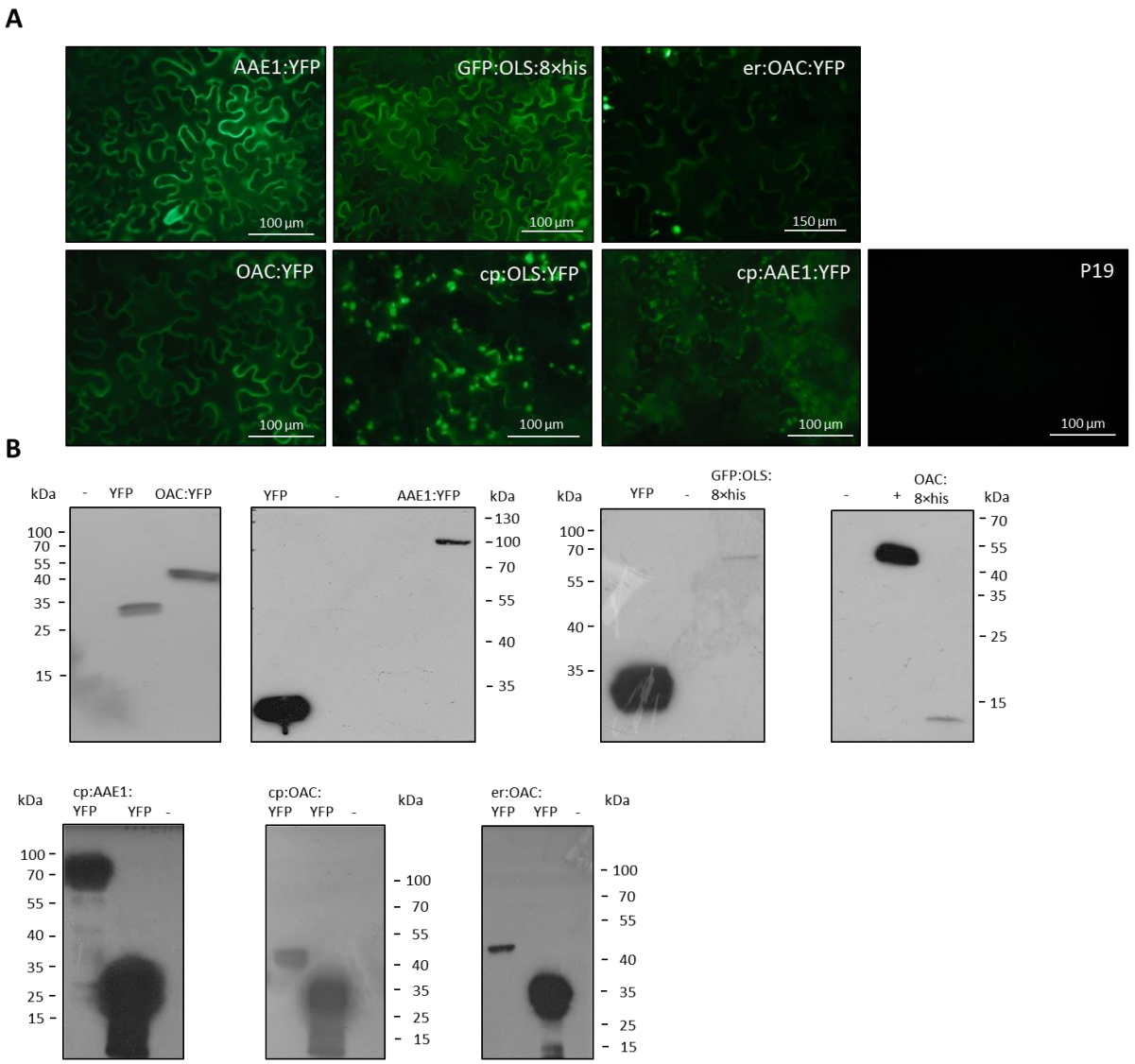


Figure 4.2.2 (A) Fluorescence microscopy of different AAE1, OLS and OAC variants. *N. benthamiana* plants expressing *p19* served as a negative control. **(B)** Western blot analysis of recombinant AAE1, OLS and OAC. Heterologous produced proteins were either detected using his-probe mouse monoclonal IgG antibodies (OAC:8×his) or GFP primary antibodies in case of C−terminal fused YFP and N−terminal GFP fused proteins (OAC:YFP; AAE1:YFP; GFP:OLS:8×his; cp:AAE1:YFP; cp:OAC:YFP and er:OAC:YFP). A recombinant esterase (OeEst228; [121]), featured with a C−terminal 6×his tag fusion (+) as well as recombinant produced YFP served as positive controls. TSP extracts of plants producing solely P19 were used as negative controls (-). cp, chloroplast targeting sequence; er, ER/apoplast signal peptide.

4.2.3. Feeding of *N. benthamiana* plants with the cannabinoid biosynthetic pathway intermediate olivetolic acid

After the production of the enzymes involved in OA biosynthesis could be largely verified for localization in the chloroplast or cytoplasm, it was necessary to verify their activity. However, before plants were transiently transformed, preliminary experiments had to be carried out. Therefore, it was necessary to determine if OA is toxic to the tobacco cells, as it has been shown for cannabigerolic acid (CBGA) and Δ^9-tetrahydrocannabinolic acid (THCA) fed to *C. sativa* cell suspension cultures, and whether the extraction method used was suitable to isolate the OA injected or produced in plant cells [109]. Therefore, *N. benthamiana* plants were initially fed with 0.5 mM of OA (see 4.1.4.1) and incubated for 24 h. Metabolites were extracted (see 3.3.3) and subjected to HPLC−MS analysis (see 3.5.1). The isolated substances were screened in the negative electrospray ionization (ESI) mode with a full scan MS experiment (m/z 150–800) and at a wavelength of 260 nm. Surprisingly, not only OA could be detected, but also another substance which eluted approximately 2 min earlier, exhibiting an m/z value of 385.2 (Figure 4.2.3 A and B). To further investigate the newly formed compound and the detectability of both substances over several days, *N. benthamiana* plants were again infiltrated with 0.5 mM of OA, followed by harvesting of plants after different time intervals ranging from 2 h to 168 h (see 4.1.4.1). After extraction, metabolites were subjected to HPLC−MS analysis and detection was performed in negative selected ion monitoring (SIM) with selected m/z of 223.2 and 385.2. Screening with the selected m/z of 385.2 revealed two substances eluting at almost the same retention time (Figure 4.2.3 C). As it is known that plants are able to glycosylate secondary metabolites under the action of UDP-glycosyltransferases (UGTs) as a mechanism for plant stress management, enabling for example the storage of toxic compounds in the vacuole, it could be assumed that both newly formed substances represent the glycosylated forms of OA [105]. Moreover, the metabolites were detectable even after 168 h of incubation, which could suggest the storage in the vacuole like described in case of glucosylated isoflavones in soybeans [122] (Figure 4.2.3 C). In the process of glycosylation, OA could potentially be glycosylated either at the hydroxy group at C-2, or at C-4, which would be consistent with two peaks being detected (Figure 4.2.3 C). In particular, since a reversed-phase HPLC column was used, in which hydrophilic substances elute earlier, the assumption of glycosylation would be supported [123]. Furthermore, the detected m/z value of 385.2 would correspond to the mass of OA containing a glucose moiety. However, due to the lack of authentic standards and to ultimately confirm the hypothesis, the substances need to be isolated and analyzed using other methods, such as nuclear magnetic resonance (NMR) spectroscopy. Nevertheless, for the further course of the project it was assumed that the extracted substances are the glucosylated forms of OA, since the results were in

consensus with a recent publication from Gülck *et al*. showing that chemically synthesized C-4 OA glucoside resembled the metabolite eluting first, followed by the C-2 OA glucoside [124].

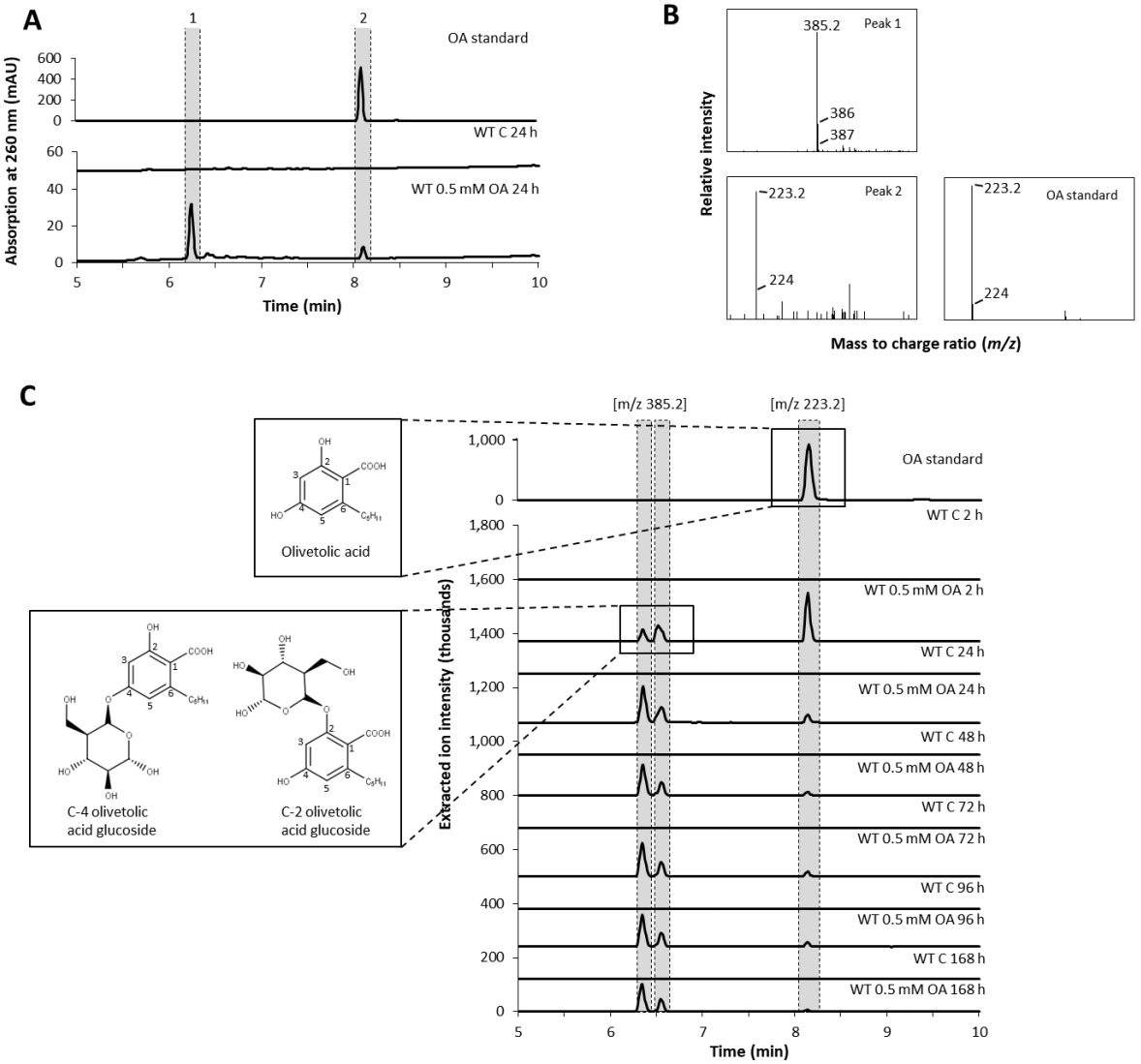


Figure 4.2.3 (A) HPLC–MS analysis of initial studies on extractability of olivetolic acid (OA) infiltrated in leaves of *N. benthamiana* wild-type plants (WT). After 24 h of incubation a new substance was detectable (1) in addition to the originally injected OA (2). As a negative control (WT C) wild-type plants were solely infiltrated with infiltration buffer. **(B)** Reference fragmentation of metabolites detected in A. Peak 1 resembles OA with an m/z of 223.2 $[M-H]^-$, while the new substance (2) had an m/z of 385.2 $[M-H]^-$. **(C)** HPLC–MS analysis of traceability studies of OA and the newly produced compound in infiltrated *N. benthamiana* WT plants. In negative selected ion monitoring, two peaks with an m/z of 385.2 were detected. The first produced metabolite could represent C-4 OA glucoside, while the second one could represent C-2 OA glucoside [124]. Even after 168 h of incubation, OA glucosides were detectable.

4.2.4. Production of glycosylated OA in transiently transformed *N. benthamiana* plants

Next, metabolic engineering of OA biosynthesis should be achieved by introduction of the pathway genes into the plant host. Proteins previously detected by either Western blot or fluorescence microscopy were selected for reconstruction (Figure 4.2.2). To simplify the transformation process, multigene constructs were assembled for production of the desired proteins. Therefore, the AAE1:YFP TU was fused to the GFP:OLS:8×his TU into the pDGB3Ω1-level plasmid (TOA 1), whereas the OAC:8×his TU was cloned together with the P19 TU (kindly provided by Dr. Iryna Gerasymenko) into the pDGB3Ω2-level plasmid (TOA 2) (Figure 4.2.4 A). Unfortunately, due to time constraints, it was not possible to clone the TUs containing the sequences for the chloroplast localized proteins into multigene constructs as well. Hence, for transient transformation of *N. benthamiana* plants, Agrobacteria harboring the constructs TOA 1 and TOA 2, ensuring cytosolic localization of proteins, were co-infiltrated, while in case of the chloroplast localized ones, Agrobacteria, containing the single TUs (cp:AAE1:YFP, cp:OLS:YFP and cp:OAC:YFP), were co-injected. Four days post infiltration the plants were then supplemented with 4 mM hexanoic acid and harvested after another 24 h of incubation (see 4.1.4.2). Transiently transformed plants solely infiltrated with infiltration buffer were used as negative control. As an additional control, plants expressing *p19* (in case of cytosolic localized enzymes) and plants infiltrated with *A. tumefaciens* containing the pDGB2α2 empty vector (in case of proteins targeted to the chloroplasts) were used. The extracted metabolites were administered to HPLC−MS and analyzed in terms of production of olivetolic acid and its glucosylated forms (see 3.5.1). Localization of enzymes in the cytosol resulted clearly in the synthesis of C-4 OA glucoside, when supplemented with hexanoic acid. This would imply that the endogenous UGTs in *N. benthamiana* exhibit higher regio-selectivity to the C-4 positioned hydroxyl group of OA when it is produced in a stepwise enzymatic conversion and is not directly available in large amounts (Figure 4.2.3 B). Besides the C-4 glucosylated OA, also small amounts of OA could be observed when proteins were localized in the cytoplasm (Figure 4.2.4 B). Further analysis of proteins targeted to the chloroplasts revealed also the formation of the C-4 OA glucoside, but in considerably lower amounts as detected for the glucoside produced by enzymes localized in the cytosol (peak area of 1700 ± 32 ion counts s^{-1} compared to 24854 ± 1125 ion counts s^{-1}, respectively) (Figure 4.2.4 C). However, due to the deviant experimental design, it is difficult to compare the results with each other. Thus, absence of P19 in the transformation set up with the chloroplast localized enzymes could have resulted in increased gene silencing, which in turn could have had negative effects on the synthesis of the desired substances [114]. Nevertheless, general feasibility of OA reconstitution in the cytosol or the chloroplasts was proven. Finally, the C-4 OA glucoside was even detected in plants transformed with the pathway genes and without supplementation of hexanoic acid (Figure 4.2.4 C). This was exclusively observed for the chloroplast localized pathway, as the peaks in Figure 4.2.4 B

could be considered as background noise due to their general appearance in all of the controls. These findings could be explained by the fact that fatty acids are synthesized in plastids and that tobacco was already shown to produce hexanoic acid in small amounts [116,125].

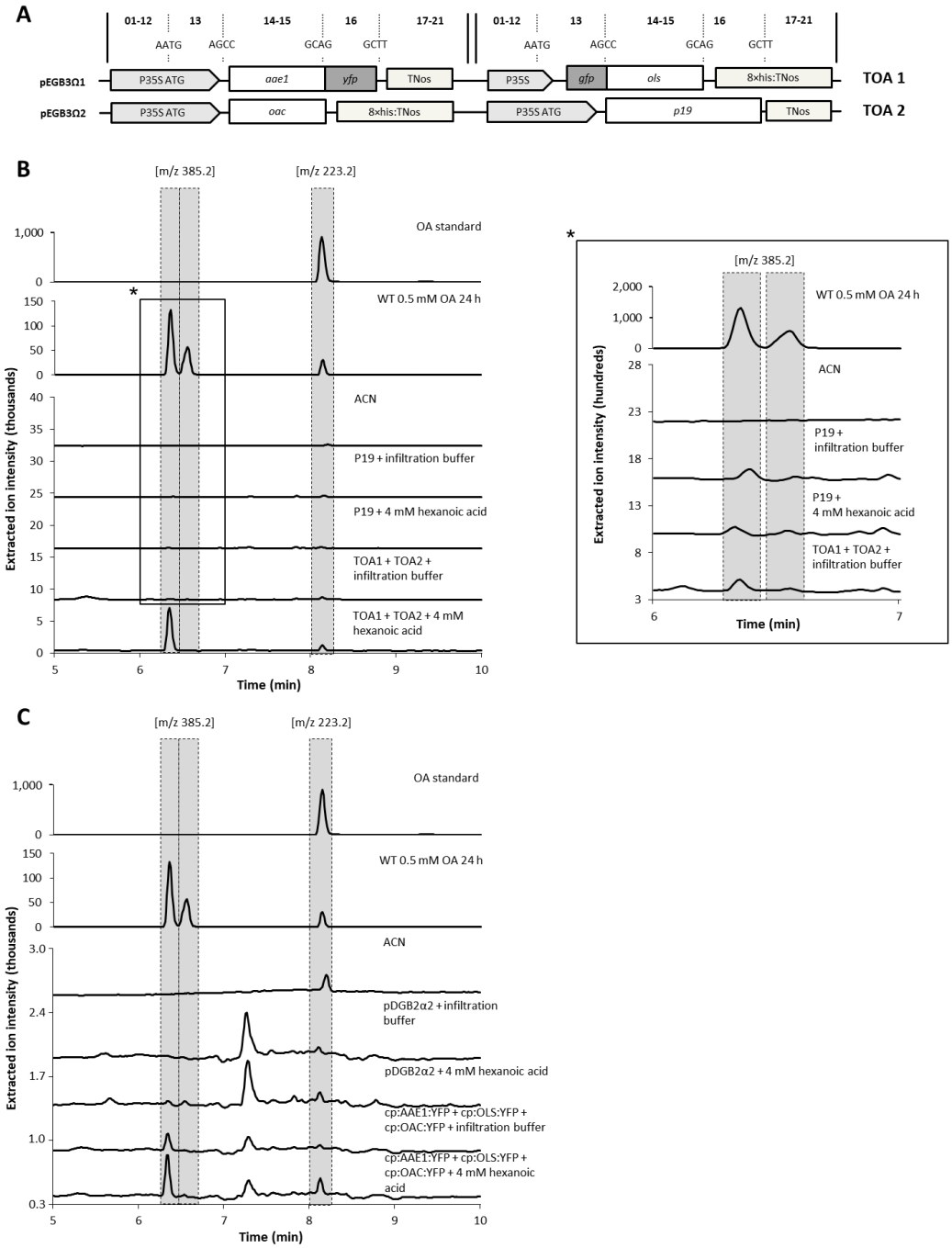


Figure 4.2.4 (A) Multigene expression cassettes generated for the production of olivetolic acid (OA). Abbreviation of each construct is listed on the right. Boxes are not drawn to scale. **(B** and **C)** HPLC–MS analysis revealed production of OA glucoside in *N. benthamiana* plants expressing the pathway genes localized in the cytosol **(B)** or in chloroplasts **(C)** with additional supplementation of 4 mM hexanoic acid. Due to the lack of authentic standards, wild-type *N. benthamiana* plants infiltrated with OA served as a positive control. Plants infiltrated with infiltration buffer as well as plants injected with Agrobacteria harboring P19 or the pDGB2α2 empty vector and supplemented with hexanoic acid were used as negative controls.

In summary, the initial results are promising regarding integration of the entire cannabinoid biosynthesis into the favored host. As cannabinoids in general are strongly hydrophobic, their usage in drug formulation is rather complicated which in turn also leads to a limitation of their delivery options [126]. However, as a positive side-effect, it was observed that the precursor olivetolic acid produced *in planta* was glucosylated, paving the way for future new to nature cannabinoid-based pharmaceuticals, as glycosylation of desired substances would enhance the solubility and therefore the bioavailability of the potential drugs [127,128]. Besides the improved bioavailability, the glycosylated drugs could also act as prodrugs, which is for example described for senna glycosides. When administered orally, sugars were cleaved in the colon by β-glucosidases secreted by the intestinal microbiome [129]. Moreover, since cannabinoids are toxic to plant cells, causing mitochondrial permeability transitions and DNA degradation, glycosylation of cannabinoids *in planta* could be beneficial as they could be converted into non-toxic metabolites as described in the case of steroidal glycoalkaloids (SGAs) and xenobiotics [130–132]. Accordingly, the attempt to achieve higher yields in tobacco does not necessarily have to be coupled to the storage of cannabinoids in the secretory cavity the glandular trichomes as it is the case in *C. sativa* [54,133].

4.3. Supporting information

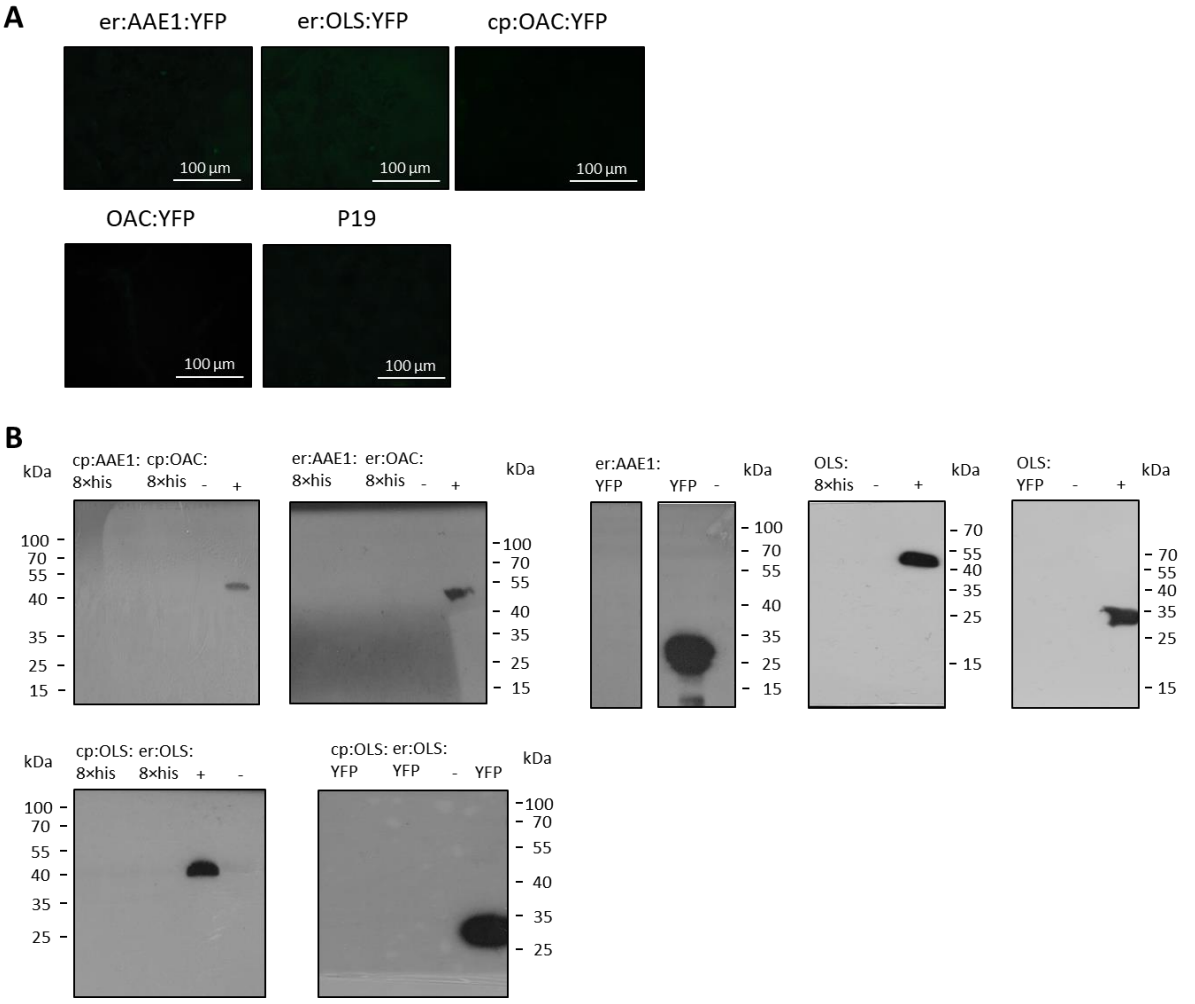


Figure S4.3.1 (A) Fluorescence microscopy of er:AAE1:YFP, er:OLS:YFP, cp:OAC:YFP and cp:OLS:YFP. Plants expressing *p19* served as a negative control. **(B)** Western blot analysis of total soluble protein extracts from *N. benthamiana* transformed with Agrobacteria harboring different AAE1, OLS and OAC constructs. His-probe mouse monoclonal IgG antibodies or GFP primary antibodies in case of C−terminal fused YFP and N−terminal GFP fused proteins were used for detection. A recombinant esterase (OeEst228; [121]), featured with a C−terminal 6×his tag fusion (+) as well as recombinant produced YFP served as positive controls. TSP extracts of plants producing solely P19 were used as negative controls (-).

Chapter 5

Biosynthesis of cannabigerolic acid

5.1. Materials and methods

5.1.1. Chemicals

Cannabigerolic acid (CBGA) used as authentic standard was purchased from Sigma-Aldrich, while 2-*O*-geranyl olivetolic acid (2-*O*-GOA) was kindly provided by the lab of Prof. Oliver Kayser from the Department of Biochemical and Chemical Engineering at the Technical University of Dortmund (Dortmund, Germany). Geranyl diphosphate (GPP) and olivetolic acid (OA), serving as substrates for prenyltransferase assays, were obtained from Axon Medchem (Groningen, The Netherlands) and Santa Cruz Biotechnology, respectively.

5.1.2. Plasmids and genetic material

The soluble aromatic prenyltransferase NphB from *Streptomyces* sp. strain CL190 (GenBank accession no: AB187169.1) was synthesized by Integrated DNA Technologies (Coralville, IA, USA), while the pICH11599 harboring the coding sequence for the membrane bound *Cannabis sativa* prenyltransferase 4 (PT4; GenBank accession no: BK010648.1) was kindly provided by the lab of Prof. Oliver Kayser. The basic GB parts, GB destination vectors, the pBIN61-P19 plasmid and signal peptide sequences were obtained as described in 4.1.2.

5.1.3. Oligonucleotides

All oligonucleotides used in this study were synthesized by Eurofins Genomics GmbH (Ebersberg, Germany). 4-nt overhangs generated by annealing of oligonucleotides for incorporation into the GoldenBraid system are highlighted in bold. BsmBI recognition sites are underlined.

No.	Name	Nucleotide sequence 5'–3'
12	pCsPT4 H101	ATA TCG TCT CAC TCG **AGC C**GG ACT CTC ATT AGT TTG TAC C
13	pCsPT4 M201	GCG CCG TCT CAC TCG **CTG C**AA TGA AAA CAT ATA CTA AAT ATT C
14	pCsPT4 cPCR101	GGT TCT CCT CAT CAT GAA TCT G
15	pCsPT4 cPCR201	CAA GAG CTG ATG TGG TTG CAG
16	pNphB cPCR101	CTG AGC ACG TTC CAG GAC ACG CTC G
17	pNphB cPCR201	CGA GTC GAA CGC CTT CAG CAG TCC G
18	pNphB Q295L fw	CAC CGA TGT CCT GCG CGG ACT GC
19	pNphB Q295L rev	ATG TGG TAG TAC GCG CCC AG

5.1.4. Feeding with olivetolic acid

Transiently transformed *N. benthamiana* plants were infiltrated with 0.5 mM olivetolic acid (solved in infiltration buffer (10 mM MES, 10 mM $MgSO_4$)) and harvested 24 h post infiltration. Plants solely infiltrated with infiltration buffer were used as negative control. Finally, metabolites were extracted as described in 3.3.3.

5.1.5. Molecular cloning methods

5.1.5.1. Site-directed mutagenesis (SDM)

For introduction of mutation Q295L into the NphB gene, site-directed mutagenesis (SDM) was performed by means of the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) was performed according to the manufacturer's protocol.

5.1.6. Protein biochemical methods

5.1.6.1. Microsomal preparation

Microsomal preparation of transiently transformed *N. benthamiana* was carried out according to Diesperger *et al.* (1974) with few modifications. All steps were performed at 4 °C. Firstly, the material of six plants each was snap-frozen in liquid nitrogen and homogenized in just as much solubilization buffer (100 mM Tris-HCl; 20 % (v/v) saccharose, 10 mM KCl, 20 mM β-mercaptoethanol, 50 mM ascorbic acid, pH 8.5) for about five minutes with a d ULTRA-TURRAX T 18 basic (IKA-Werke, Staufen in Breisgau, Germany). The homogenate was stirred for 15 min and then extruded through one layer of miracloth (Merck Millipore, Burlington, MA, USA). The filtrate was then centrifuged (30 min, 13,000 ×*g*, 4 °C). Subsequently, 1 M of $MgCl_2$ solution was added stepwise to the supernatant till the final concentration of 50 mM was reached. After 60 minutes of stirring, the samples were centrifuged for 1 hour at 38,000 ×*g* at 4 °C. The debris was resuspended in 1.5 mL prenyltransferase buffer (50 mM Tris-HCl, 10 mM NaCl, pH 7.5) and homogenized in a glass homogenizer. Samples were then subjected to Western blot analysis and activity assays.

5.1.6.2. Isolation of chloroplasts

Chloroplasts were isolated based on the protocol of Kley *et al*., 2010 [134]. For the isolation 2× chloroplast isolation buffer (300 mM sorbitol, 1 mM $MgCl_2$, 50 mM HEPES/KOH (pH 7.5), 5 mM EGTA, 5 mM EDTA, 10 mM $NaHCO_3$ and freshly added 0.5 mM DTT) was mixed with 100 % Percoll (GE Healthcare, Buckinghamshire, UK) to yield a 50 % Percoll solution. Prior isolation, the solution was cooled to 4 °C. 3 g of transiently transformed *N. benthamiana* leaf material were harvested and homogenized in 23 mL of 1× chloroplast isolation buffer for 5 s using an ULTRA-TURRAX T 18 basic (IKA-Werke). Subsequently, the homogenate was filtered through miracloth (Merck Millipore) and mixed with the 50 % Percoll solution. After centrifugation (10 min, 2,000 ×*g*, 4 °C), the upper layer, containing broken chloroplasts, was discarded. The pellet was washed with 14 mL of 1× isolation buffer, followed by an additional centrifugation step for five minutes (2,000 ×*g*, 4 °C). Finally, the pellet containing intact chloroplasts was resolved in 200 µL of prenyltransferase buffer (50 mM Tris-HCl, 10 mM NaCl, pH 7.5) and used for prenyltransferase activity assays or microscopic examinations.

5.1.6.3. Prenyltransferase activity assays with plant crude extract

100 mg of frozen leaf material were homogenized with 100 µL of prenyltransferase assay buffer (50 mM Tris-HCl, 10 mM NaCl, pH 7.5) and centrifuged at 17,000 ×*g* for 30 min at RT. For activity assays, either 87 µL of the TSP prepared in this way, 87 µL of the previously isolated chloroplasts or 87 µL of microsomal preparations were mixed with 10 µL of (10 mM) GPP, 1 µL (50 mM) of olivetolic acid and 2 µL (500 mM) of $MgCl_2$. Afterwards, the mixtures were incubated at 30 °C and 750 rpm for 24 h. The assays were quenched by adding 275 µL of ice-cold acetonitrile, followed by incubation on ice for 30 min. Finally, the supernatants were purified two times from solid particles by centrifugation (17,000 ×*g*, 30 min, 4 °C) and subjected to HPLC−MS analysis (see 3.5.1). The amount of formed product was quantified by peak integration and subsequent conversion from area counts into nM by means of serial standard dilutions (CBGA 1.6–800 nM). Assays were performed in four replicates.

5.1.6.4. Prenyltransferase activity assays with purified proteins

Reaction mixtures contained 10 µg of NphB:8×his or 10 µg of NphB(Q295L):8×his as well as 1 mM of GPP, 0.5 mM of OA and 10 mM of $MgCl_2$ in a total volume of 100 µL buffered in prenyltransferase assay buffer (50 mM Tris-HCl, 10 mM NaCl, pH 7.5). The reaction took place for 24 h at 30 °C and 750 rpm. To terminate the reactions, 275 µL of ice-cold acetonitrile was added, followed by incubation on ice for 30 min. Finally, the supernatants were purified two times from solid particles by centrifugation (17,000 ×*g*, 30 min, 4 °C). Assays were performed in four replicates and analyzed by HPLC−MS (see 3.5.1).

5.1.7. Analytical methods

5.1.7.1. Confocal laser scanning microscopy

For verification of heterologous produced CsPT4, isolated chloroplasts were analyzed by confocal laser scanning microscopy utilizing a Leica TCS SP5 II spectral confocal laser scanning microscope (Leica Microsystems, Wetzlar, Germany). Samples were excited with an argon laser (488 nm; 20 % intensity). Emission spectral ranges were set to 500-540 nm for GFP signal detection and 660-700 nm for chlorophyll signal detection using a 100× oil objective (HCX PL APO CS 100.0 × 1.44 OIL UV, Leica Microsystems) [135,136].

5.2. Results and discussion

5.2.1. Molecular cloning of aromatic prenyltransferases

As a result of the successful generation of olivetolic acid (OA) in transiently transformed *N. benthamiana* described in Chapter 4, it had to be further investigated, if the subsequent step in the biosynthesis of cannabinoids through prenylation of OA, forming cannabigerolic acid (CBGA), could be achieved by integrating aromatic prenyltransferases into the transformation setup. Therefore, the membrane-bound aromatic prenyltransferase 4 (PT4) from *C. sativa*, as well as the soluble aromatic prenyltransferase NphB originating from *Streptomyces* sp. strain CL190 were integrated into the GoldenBraid (GB) system [60,137]. For domestication of PT4, the pICH11599 vector harboring the PT4 coding sequence was used as a template to incorporate appropriate fusion overhangs by PCR (see 3.2.1 and 5.1.3; oligonucleotides 12 and 13). Subsequently, the PCR product was purified (see 3.2.2) and ligated into the universal domesticator *via* a GB reaction (see 3.2.3). In case of the soluble aromatic prenyltransferase NphB, the GB–modified sequence was synthesized by Integrated DNA Technologies harboring the appropriate fusion sites (AGCC and GCAG) and BsmBI recognition sites for ligation into the universal domesticator. The resulting GB reaction mixtures were transformed into chemically competent *Escherichia coli* TOP10 (see 3.2.4), followed by isolation of plasmid DNA (see 3.2.5), restriction analysis (see 3.2.6) and verification by sequencing (see 3.2.8; oligonucleotides 1 and 2).

Additionally, the mutation Q295L was introduced which was shown to increase CBGA formation, since NphB is described to geranylate OA preferentially at the 2-*O* position producing 2-*O*-geranyl olivetolic acid (2-*O*-GOA) [62,138]. The domesticated NphB served as a template for site-directed mutagenesis of NphB(Q295L) (see 5.1.5.1) utilizing the appropriate primer pair (see 5.1.3; oligonucleotides 18 and 19). Subsequently, the cloned vector was amplified in *E. coli* TOP10 and verified by sequencing (see 3.2.8). As already described in chapter 4, the production of enzymes in different subcellular localizations had to be investigated. Therefore, for assembly of transcriptional units in α-level plasmids, the coding sequence (CDS) of NphB and NphB(Q295L) were either fused with the chloroplast transit peptide sequence (cp) or the ER/apoplast signal peptide sequence (er), while in case for production in the cytoplasm, the CDS was directly fused to the P35S ATG. For the PT4 on the other hand, only the sequence with the native plastidal targeting was utilized and ligated to the 35S promoter and the nopaline synthase terminator (TNos). To visualize the proteins during fluorescence microscopy, *yfp* or *gfp* were C–terminally fused to the enzymes. The assembled GB constructs were then transformed into chemically competent *A. tumefaciens* EHA105 cells (see 3.2.9) and positive clones were confirmed by colony PCR (see 3.2.10) using the relevant primer pairs (see 5.1.3; oligonucleotides 14-17). Finally, glycerol stocks of overnight cultures were prepared and stored

at -80 °C. Figure 5.2.1 illustrates the expression constructs generated specifically for this part of the study.

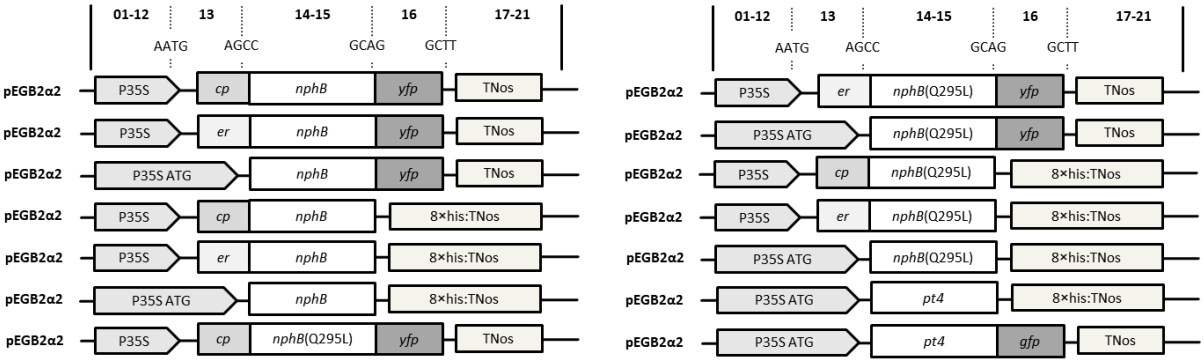


Figure 5.2.1 Cloning of different prenyltransferase expression cassettes using the GoldenBraid cloning system. The capital letters show the four-nucleotide overhangs ensuring correct final orientation within the transcriptional unit (TU), while the numbers above the scheme represent standard GoldenBraid classes within the TU structure [93]. P35S, cauliflower mosaic virus (CaMV) 35S promoter; P35S ATG, cauliflower mosaic virus (CaMV) 35S promoter with an integrated start codon ensuring cytosolic localization; TNos, nopaline synthase terminator; 8×his:TNos, nopaline synthase terminator comprising an 8×his-tag; cp, chloroplast targeting sequence; er, ER/apoplast signal peptide. Boxes are not drawn to scale.

A detailed description of all chronological steps performed during assembly of GB constructs and transformation of *A. tumefaciens* used for transformation of tobacco plants is given in chapter 4.2.1.

5.2.2. Heterologous production of PT4 and NphB

N. benthamiana plants were infiltrated with Agrobacteria harboring the relevant prenyltransferase genes. Transformation with the pBIN61 P19 plasmid was utilized as a negative control. After five days of incubation, the plants were harvested and proteins were extracted. In case of NphB:8×his and NphB(Q295L):8×his total soluble proteins (TSP) were isolated (see 3.4.1), while in case for the transmembrane protein PT4:8×his, microsomes were prepared (see 5.1.6.1). The protein and vesicle-like fragment preparations were then subjected to Western blot analysis and visualized using his-probe mouse monoclonal IgG antibodies (see 3.4.6.3). As shown in Figure 5.2.2, only the different variants of NphB:8×his and NphB(Q295L):8×his could be detected with a calculated mass of approximately 35 kDa, whereas the production of PT4 could not be ascertained. Nevertheless, as described in Chapter 4, the successful detection of heterologous produced proteins by Western blot relies on many different parameters and does not necessarily mean that the desired protein is not synthesized. Particularly, detection of transmembrane proteins in general is more challenging due to difficulties in terms of their solubilization [118,119,139].

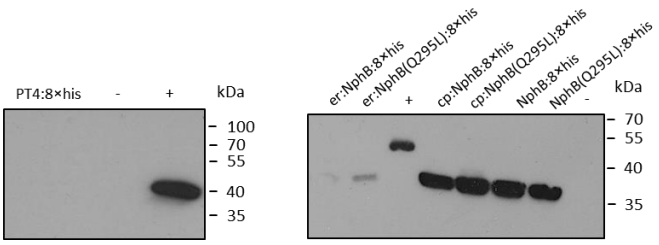


Figure 5.2.2 Western blot analysis of heterologous produced prenyltransferases. +, recombinant esterase (OeEst228; [121]), featured with a C–terminal 6×his tag fusion. -, TSP extracts of plants producing solely P19; cp, chloroplast targeting sequence; er, ER/apoplast signal peptide.

Therefore, *N. benthamiana* plants were co-transformed with Agrobacteria harboring the PT4:GFP and P19 for visualization by fluorescence microscopy. Plants solely infiltrated with P19 served as a negative control. The plants were harvested five days post inoculation and chloroplasts were isolated (see 5.1.6.2). Following that, the PT4:GFP isolates were investigated by confocal laser scanning microscopy (CLSM; see 5.1.7.1). Since chloroplasts emit an auto fluorescence signal with a peak of approximately 680 nm caused by the pigment chlorophyll, the emission was observed on one hand in the spectral range of 660-700 nm, to ensure that the observed structures were the desired organelles, and on the other hand in the spectral range of 500-540 nm for the detection of GFP [135,136]. The subsequent analysis of the merged images of the auto fluorescence of the chloroplasts and the fluorescence of PT4:GFP clearly revealed the expression of *pt4:gfp* and the expected localization in the chloroplasts, when compared to the P19 control (Figure 5.2.3).

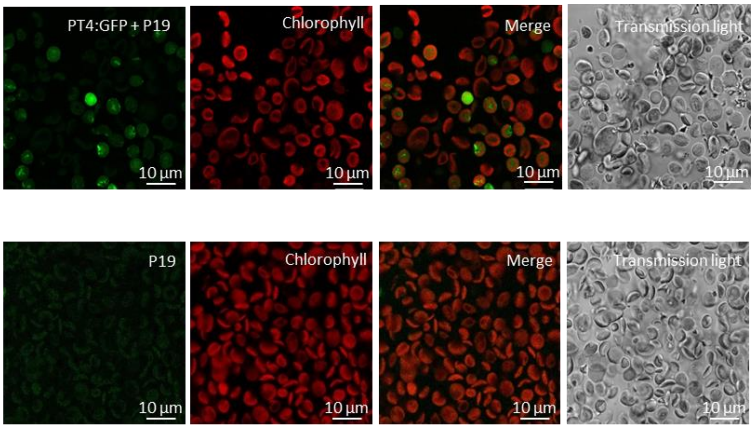


Figure 5.2.3 Confocal laser scanning microscopy of isolated chloroplasts using a Leica TCS SP5 II spectral confocal laser scanning microscope. *N. benthamiana* plants were expressing *pt4:gfp* and *p19*, or only *p19* (negative control).

Nevertheless, the detection of GFP in the desired compartment does not automatically imply that a functional prenyltransferase was produced, albeit fusion partners are sometimes known to stabilize the membrane proteins, as it is the case for G protein-coupled receptors [140]. Consequently, enzyme activity was measured by subjecting the isolated chloroplasts or microsomal preparations to prenyltransferase activity assays (see 5.1.6.3). After 24 h of incubation the assays were quenched and analyzed by HPLC−MS in negative selected ion monitoring (SIM) with selected m/z of 359.2. Surprisingly, prenylation of olivetolic acid yielding in CBGA was identified in the reaction mixture with chloroplasts or microsomes containing PT4:GFP as well as in the P19 control mixtures when compared to the authentic standard (Figure 5.2.4 A). However, when considering cannabis research from recent years, it was observed, that cannabinoids or cannabimimetic compounds are not exclusively produced in *C. sativa*, but also in other plant species including *Lepidium meyenii*, *Piper nigrum*, *Acmella oleracea* or *Radula marginata,* whereby in the latter, even the upstream genes responsible for the production of the central precursor CBGA have been identified [41–45]. In order to investigate if the production of CBGA in *N. benthamiana* might be linked to endogenous prenyltransferases, crude plant extracts from wild-type plants were used and subjected to prenyltransferase activity assays (see 5.1.6.3). Crude plant extracts from wild-type (WT) *N. benthamiana* plants, in which either one or both substrates (OA and geranyl diphosphate (GPP)) were deficient, served as negative controls. Furthermore, prenyltransferase assay buffer supplemented with the substrates was used as an additional control. After the analysis of the outcomes by HPLC−MS, CBGA was identified in assays containing both substrates. However, this was also the case for the samples which contained only buffer instead of the plants proteins (Figure 5.2.4 B). Actually, it has been described that prenylation *via* friedel-crafts alkylation does not to occur spontaneously but requires an enzymatic catalyst or a Lewis acid such as aluminum chloride or ferric chloride [141]. However, since the reaction occurred without the intentional addition of protein and without the presence of a Lewis acid, it is possible that either one of the ingredients used was contaminated with proteins or that prenylation occurs spontaneously by a yet unknown chemical mechanism.

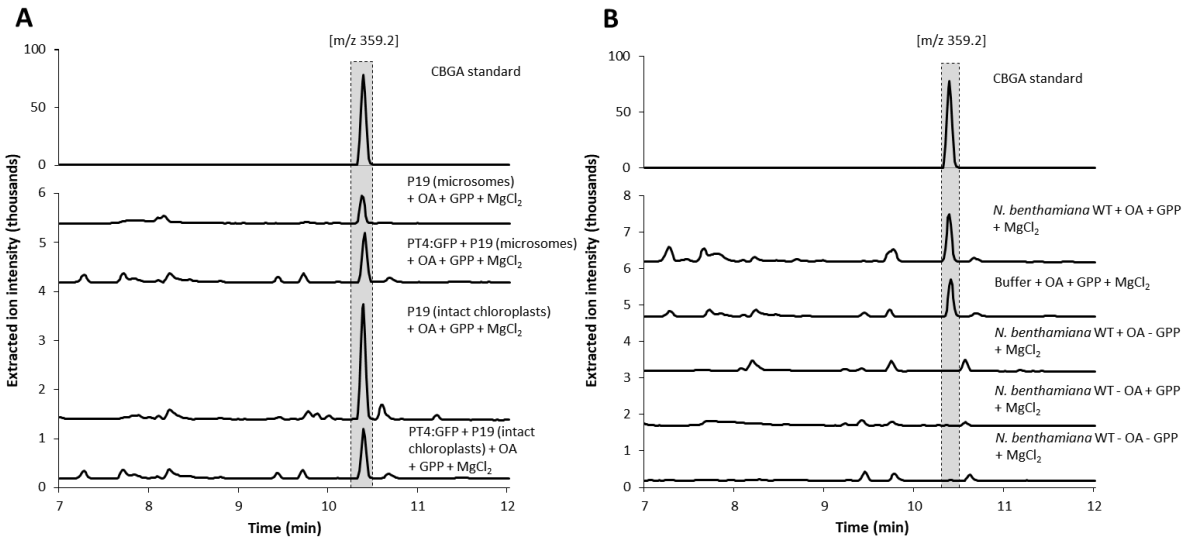


Figure 5.2.4 (A) HPLC–MS analysis of *in vitro* prenyltransferase assays of heterologous produced PT4 utilizing isolated chloroplasts or microsomal preparations. *N. benthamiana* plants solely producing P19 were used as a control. Surprisingly, prenylation of olivetolic acid (OA) yielding in cannabigerolic acid was observed in the reaction mixture with PT4 as well as in the control mixture. **(B)** HPLC–MS analysis of *in vitro* prenyltransferase assays with total soluble proteins of *N. benthamiana* wild-type leaves. Leaf extracts supplemented with only one of the two substrates or none of the substrates were used as controls. The control, which contained no protein but both substrates, showed a turnover to CBGA as well.

Although the ability of PT4 to produce CBGA has been confirmed by Luo *et al.* when expressed in yeast, no activity could be detected in *N. benthamiana* in *in vitro* assays in the course of this thesis [60]. However, this was most likely due to a handling problem with membrane-bound proteins in *in vitro* assays, as a recent study has shown that the heterologous expression of *pt4* in *N. benthamiana* plants, followed by supplementation with OA and GPP, resulted in the synthesis of CBGA and glycosylated CBGA *in vivo* [124]. In fact, the characterization of transmembrane proteins is generally challenging as the reconstitution for example is strongly reliant on the careful selection of the detergent species used for solubilisation [142]. Therefore, for ongoing studies on cannabinoid pathway integration, an alternative protein was sought that is easier to handle in *in vitro* studies. Hence, focus was set on the soluble aromatic prenyltransferase NphB.

Since the expression of *nphB* and *nphB*(Q295L) was already proven (Figure 5.2.2 and Figure S5.3.1), the next step was the affirmation of their activity. Therefore, 20 *N. benthamiana* plants were transiently co-transformed with Agrobacteria containing either NphB:8×his and P19 or NphB(Q295L):8×his and P19, respectively. After five days of incubation, plants transformed with the same construct were pooled and the recombinant proteins were purified by immobilized metal ion affinity chromatography (IMAC) utilizing a column matrix pre-charged with Ni^{2+} ions (see 3.4.2). The resulting elution fractions were subjected to SDS−PAGE (see 3.4.5) and visualized by Coomassie Brilliant Blue staining (see 3.4.6.1). Both purification approaches showed, that the desired proteins featuring the calculated size of approximately 35 kDa were present in elution fractions 1 and 2 (data not shown). Therefore, both fractions were combined, desalted by replacing the purification buffer containing imidazole with prenyltransferase buffer (50 mM Tris-HCl, 10 mM NaCl, pH 7.5), and concentrated. Following that, samples were quantified by BCA assay and diluted in 25 % (v/v) glycerol to a final concentration of 0.5 mg/mL (see 3.4.3). Finally, quality of the purified proteins was analyzed by SDS−PAGE and staining with silver nitrate (see 3.4.5 and 3.4.6.2). As shown in Figure 5.2.5 C no noteworthy contamination with other proteins transferred during the purification process was observed and only the expected bands with the size of approximately 35 kDa, representing the desired prenyltransferases, were detected.

To verify their catalytic activity, the proteins were then applied in prenyltransferase activity assays (see 5.1.6.4). Therefore, 10 µg of NphB:8×his or 10 µg of NphB(Q295L):8×his as well as 1 mM of GPP, 0.5 mM of OA and 10 mM of $MgCl_2$ were mixed in a total volume of 100 µL buffered in prenyltransferase assay buffer (50 mM Tris-HCl, 10 mM NaCl, pH 7.5). After 24 h of incubation at 30 °C, reactions were stopped and subsequently measured by HPLC−MS at a wavelength of 260 nm as well as in the negative electrospray ionization (ESI) mode with a full scan MS experiment (m/z 150–800) (see 3.5.1). The analysis revealed that both heterologous produced enzymes were active when compared to the negative control and authentic standards. Moreover, as described in literature, NphB predominantly geranylated OA at the 2-*O* position forming 2-*O*-geranyl olivetolic acid (2-*O*-GOA), whereas the introduced mutation Q295L resulted in an increased CBGA production and almost no more 2-*O*-GOA formation (Figure 5.2.5 A and B) [62,138]. Since studies with yeast had already shown that 2-*O*-GOA did not serve as a substrate for Δ^9-tetrahydrocannabinolic acid synthase (THCAS), there was no further need for the NphB wild-type enzyme [62]. Thus, NphB(Q295L) was utilized in all subsequent experiments.

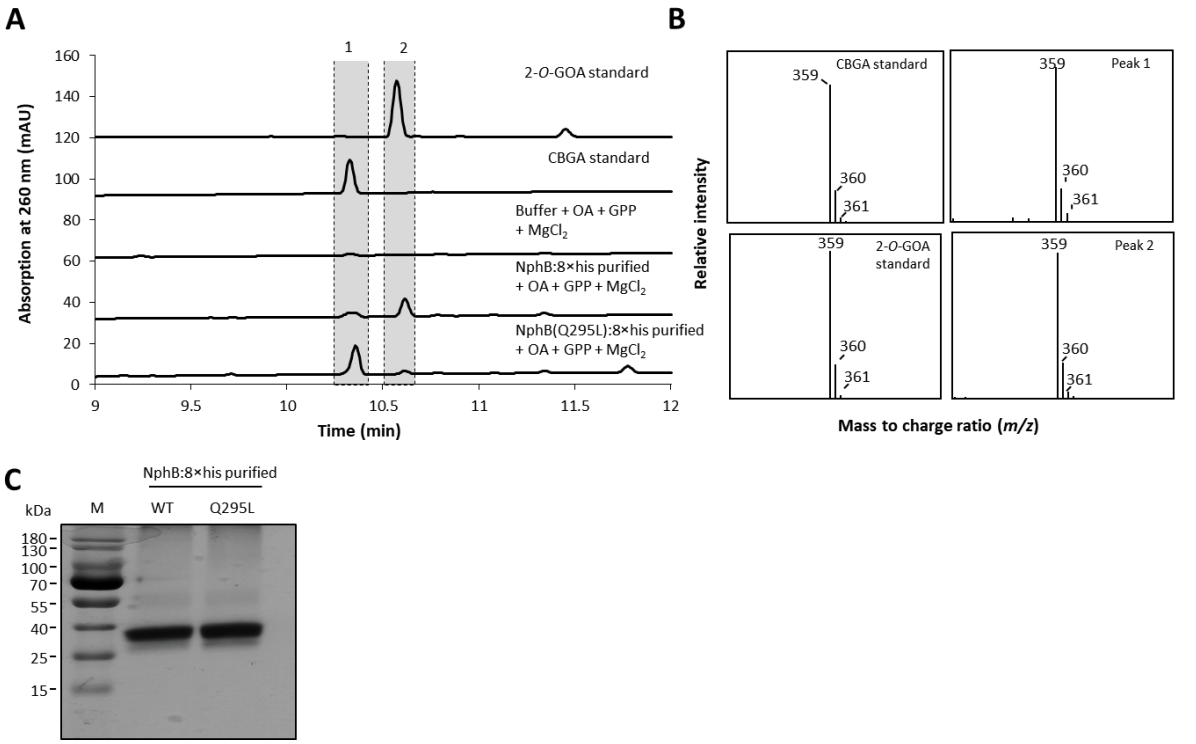


Figure 5.2.5 (A) HPLC–MS analysis of assays for prenyltransferase activity of NphB and NphB(Q295L) towards production of cannabigerolic acid. The wild-type enzyme mainly produces 2-*O*-geranyl olivetolic acid (2-*O*-GOA), while the mutant mainly synthesizes the desired pathway intermediate cannabigerolic acid (CBGA). **(B)** Mass spectrum of the detected substances (Peaks 1 and 2) corresponding to authentic standards. **(C)** SDS–PAGE of purified NphB:8×his wild-type (WT) and NphB(Q295L):8×his visualized by silver nitrate staining. Numbers to the left indicate molecular mass of marker proteins (M).

5.2.3. Investigation of cannabigerolic acid production *in vivo* after expression of *nphB*(Q295L)

The proof of activity of *in planta* produced aromatic prenyltransferase NphB(Q295L) was the next crucial step towards cannabinoid biosynthesis in the new plant host. However, for whole pathway integration, it was necessary to investigate whether CBGA could be produced *in vivo* when *nphB*(Q295L) was expressed. Consequently, plants were co-infiltrated with *A. tumefaciens* EHA105 cells carrying either *nphB*(Q295L):8×his and *p19* or *cp:nphB*(Q295L):8×his and *p19*, respectively. Plants solely expressing *p19* were used as controls. Although it was possible to produce OA and C-4 OA glucoside by introducing the necessary genes, the quantities of the detected metabolites were rather small, representing a potential bottleneck in the synthesis of CBGA (Figure 4.2.4 B and C). Therefore, in order to exclude this from the outset, OA was supplemented by injecting it into the leaves on the fourth day after infiltration with the desired prenyltransferase constructs. Transiently transformed plants injected with infiltration buffer instead of OA served as a negative control (see 5.1.4). After one day of further incubation, *N. benthamiana* plants were harvested and metabolites were extracted (see 3.3.3). Unfortunately, no CBGA was detected by HPLC–MS analysis

in plants producing the chloroplast targeted NphB mutant or the cytosolic localized one, when compared to the authentic standard (Figure 5.2.6 A). However, recent studies have shown that endogenous glycosyltransferases were not only capable of glucosylating OA but also CBGA when infiltrated into *N. benthamiana* plants resulting mostly in the production of C-4 CBGA glucoside (molecular mass of 522.2 g/mol) [124]. Therefore, the plant extracts were further screened for the production of glucosylated CBGA. But even here, no newly formed metabolite could be detected (Figure 5.2.6 B). The absence of CBGA or its glucoside could be caused by several reasons: on one hand, the infiltrated OA could have been completely converted to OA glucoside, despite still being detectable in previous experiments and that the introduced prenyltransferase did not accept the glucoside as a substrate (Figure 4.2.3 C). On the other hand, low availability of GPP could have been the bottleneck in this approach. In order to get to the bottom of the points mentioned, the plant extracts were examined for the presence of OA and OA glucoside. The HPLC–MS analysis revealed that the plants expressing the prenyltransferase genes still contained OA, which was comparable to the amounts found in the control plants (Figure 5.2.6 C). In view of this, even if OA glucoside was not accepted as a substrate, the amount of OA should have been sufficient to form CBGA. Furthermore, since the injection of OA together with GPP in *N. benthamiana* plants producing PT4 was shown to lead to the synthesis of CBGA glucoside, it can be excluded that the localization of NphB(Q295L) in chloroplasts poses a problem regarding the import of OA into the organelles [124]. Therefore, it is most likely that the availability of GPP is the limiting factor, thus *in planta* GPP pool needs to be increased. (Over)expression of *Arabidopsis thaliana* isopentenyl phosphate kinase gene (AtIPK) in tobacco for example was shown to increase monoterpene formation derived from the plastidic MEP pathway, by enhancing the production of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), the building blocks of terpenoids. Moreover, AtIPK also supported monoterpene production *via* the cytosolic mevalonate pathway (MVA) without participation of MEP pathway as shown in experiments where the key enzyme 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR) of MEP pathway was inhibited. Thus implementation of AtIPK into the transformation setup could not only enable CBGA production with the help of chloroplast-localized NphB(Q295L), but also whole CBGA biosynthesis could be translocated to the cytosol [143,144].

Additionally, integration of geranyl diphosphate synthases (GPPS), catalyzing the condensation of DMAPP and IPP to form GPP, could further enhance the GPP biosynthesis [145,146]. In this regard, the cytosolically localized *Lithospermum erythrorhizon* GPPS (LeGPPS) or the small subunit of the heterotetrameric GPPS derived from *Antirrhinum majus* (AmGPPS:SSU) might be promising candidates [147]. In case of AmGPPS:SSU, it has already been demonstrated that the introduction in stable transformed tobacco resulted in exclusive formation of GPP by binding to endogenous

geranylgeranyl diphosphate synthases (GGPPS), forming a functional GPPS. Unfortunately, the decrease of di- and tetraterpene levels led to limited chlorophyll, carotenoid and gibberellin amounts which in turn caused light-sensitivity and growth-restriction in transgenic plants [148]. Therefore, application of AmGPPS:SSU should be chosen wisely and rather be limited to the transient transformation approaches.

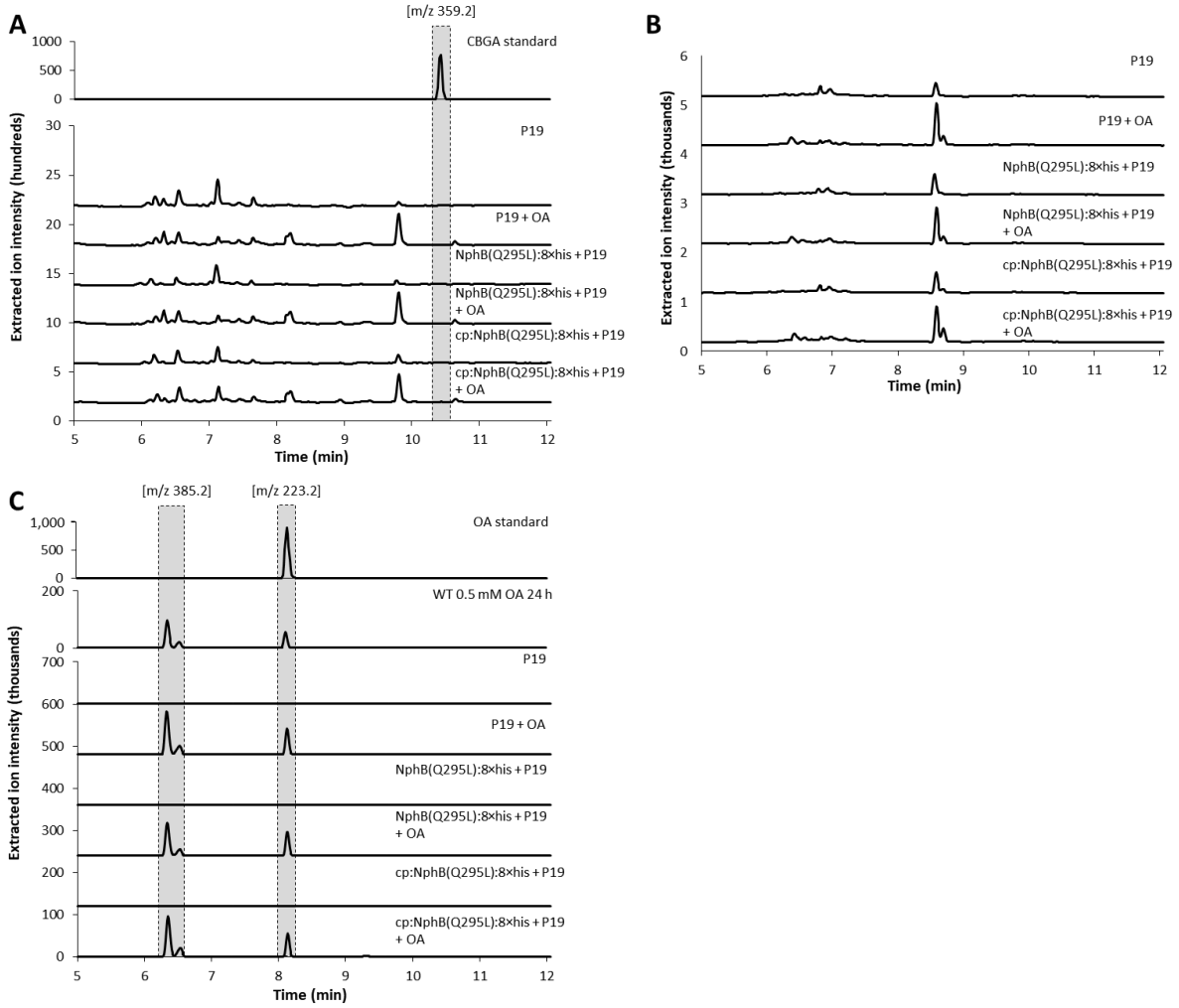


Figure 5.2.6 Studies on *in vivo* production of cannabigerolic acid and cannabigerolic acid glucosides. Metabolites were extracted from *N. benthamiana* plants infiltrated with olivetolic acid (OA) and expressing *nphB*(Q295L):8×his, *cp:nphB*(Q295L):8×his and *p19*. The extracts were analyzed by HPLC−MS in negative selected ion monitoring (SIM) with selected m/z of 359.2 (**A**; detection of cannabigerolic acid (CBGA)), in negative SIM with selected m/z of 521.2 (**B**; detection of CBGA glucoside) and in negative SIM with selected m/z of 223.2 and 385.2 (**C**; detection of OA and OA glucoside, respectively). No newly formed metabolites associated with heterologously produced enzymes could be detected in CBGA and CBGA glucoside analyses (**B** and **C**), while the amount of OA and OA glucoside remained the same compared to the P19 control **(A).** Due to the lack of authentic standards, wild-type *N. benthamiana* plants infiltrated with OA served as a positive control.

However, in order to integrate the entire cannabinoid biosynthetic pathway in future attempts, it was important to clarify, particularly with regard to the production of glycosylated cannabinoids, whether NphB(Q295L) accepts glucosylated OA as a substrate at all. As mentioned before, no authentic standards of the desired glycosylated substances were available and therefore 100 mg of the plant material previously infiltrated with OA and the prenyltransferase constructs were homogenized in prenyltransferase assay buffer and subjected to activity assays (see 5.1.6.3). The OA and OA glucosides in the extract should then serve as substrate for the heterologously produced prenyltransferases. Additionally, 1 mM of GPP and 10 mM of $MgCl_2$ were added into the reaction mixture, while plant extracts without supplementation of GPP served as negative controls. After 24 h of incubation the assays were then terminated and analyzed by HPLC−MS for the production of CBGA and CBGA glucoside.

CBGA was detected in samples containing the desired prenyltransferases and the supplemented GPP, indicating that the OA contained in the extract was sufficient for the production of CBGA (Figure 5.2.7 A). Subsequently, the reaction mixtures were also screened in the negative SIM with selected m/z of 521.2 for the C-4 and C-2 CBGA glucoside but no newly formed metabolite could be observed (Figure 5.2.7 B). Interestingly, after screening the mixtures for the presence of OA and OA glucoside, only OA was detected, although it was shown that the extracts from plants infiltrated with OA contained both, OA and its glucosides when prenyltransferase assay buffer was used as extraction reagent (Figure 5.2.7 C). Considering that the extracts contained all sorts of soluble tobacco proteins, the absence of OA glucoside could be explained by endogenous β-glucosidases cleaving the sugar moiety of OA glucosides. These proteins are essential for plants as they play a role in lignin formation, in phytohormone regulation or as a defense mechanism by activating phytochemical responses against protruding herbivores and pathogens [149–155]. Moreover, it is assumed that glucosidases in dicotyledonous plants like tobacco mostly accumulate in the apoplast or the cell wall and therefore the disruption of cells during the homogenization process could have led to the deglucosylation of OA glucosides in the following assays [156]. This, in turn, would imply that without reaction mixtures containing the purified enzymes and substrates, no final clarification about substrate acceptance could be given.

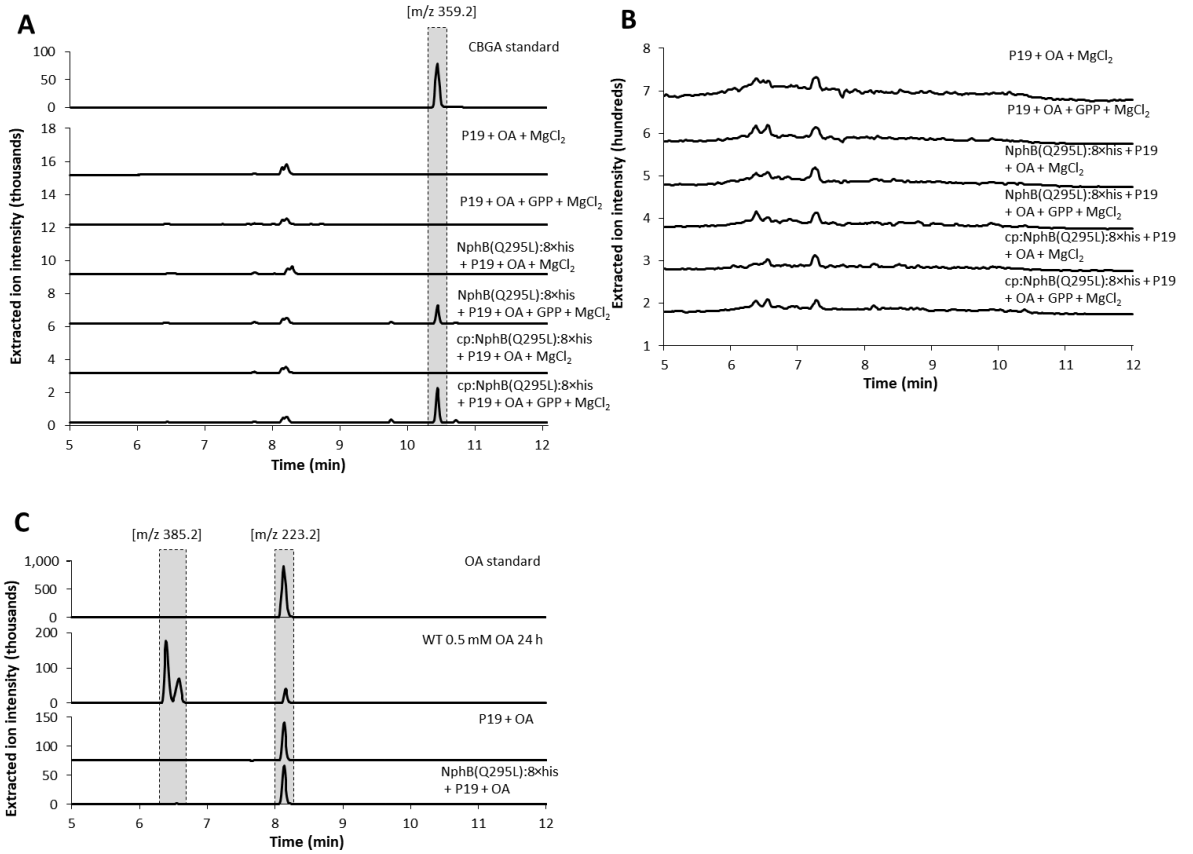


Figure 5.2.7 Studies on *in vitro* production of cannabigerolic acid and cannabigerolic acid glucosides. Metabolites and total soluble proteins were extracted from *N. benthamiana* plants infiltrated with olivetolic acid (OA) and expressing *nphB*(Q295L):8×his*, cp:nphB*(Q295L):8×his and *p19*. The extracts were subjected to prenyltransferase assays and addtionally supplemented with GPP. Subsequently, the assays were analyzed by HPLC−MS in negative selected ion monitoring (SIM) with selected m/z of 359.2 (**A**; detection of cannabigerolic acid (CBGA)), in negative selected ion monitoring (SIM) with selected m/z of 521.2 (**B**; detection of CBGA glucoside) and in negative selected ion monitoring (SIM) with selected m/z of 223.2 and 385.2 (**C**; detection of OA and OA glucoside, respectively). CBGA was formed in conjunction with heterologously produced enzymes, whereas no additional peak could be detected in CBGA glucoside analyses (**A** and **B**). **(C)** Screening of assays for OA and OA glucoside revealed, that just OA could be identified. Since lack of authentic standards, wild-type *N. benthamiana* plants infiltrated with OA served as a positive control.

5.3. Supporting information

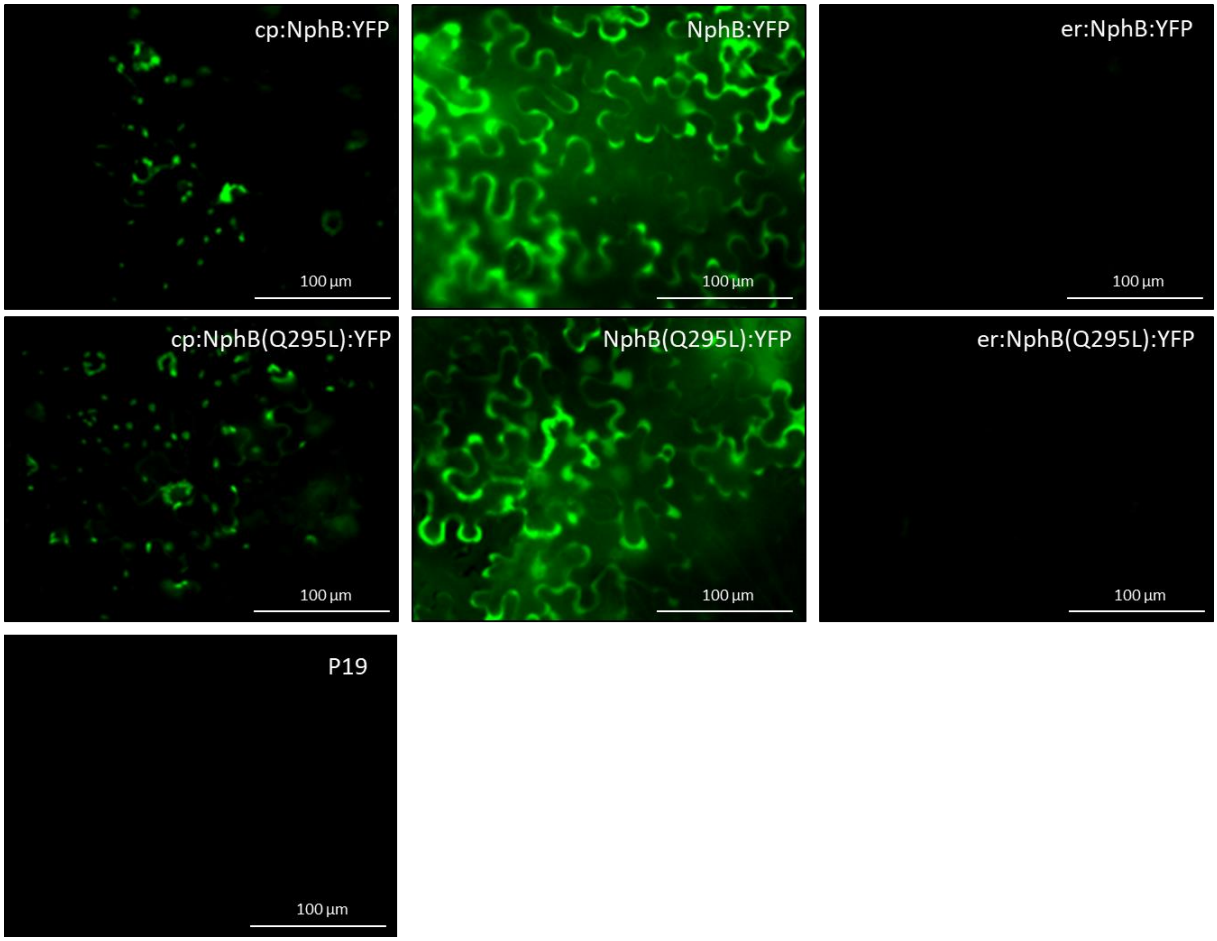


Figure S5.3.1 Fluorescence microscopy of NphB and NphB(Q295L) localized in different compartments of *N. benthamiana* cells. Plants expressing *p19* served as a negative control. cp, chloroplast targeting sequence; er, ER/apoplast signal peptide.

Chapter 6

Characterization of late biosynthetic enzymes in *planta*

6.1. Materials and methods

6.1.1. Chemicals

Cannabigerolic acid (CBGA) as well as cannabichromenic acid (CBCA) and cannabidiolic acid (CBDA) were purchased from Sigma-Aldrich, while Δ^9-tetrahydrocannabinolic acid (THCA) was procured from THC Pharm (Frankfurt am Main, Germany). Myrcene, linalool, limonene and α-pinene were kindly provided by Prof. Jürg Gertsch from the Institute of Biochemistry and Molecular Medicine at the University of Bern (Bern, Switzerland).

6.1.2. Plasmids and genetic material

The GoldenBraid (GB) -modified cannabidiolic acid synthase gene (*cbdas;* GenBank accession no: AB292682) and cannabichromenic acid synthase gene (*cbcas*; sequence as published in WO 2015/196275 A1) as well as *cbcas*(Mut1) and *cbcas*(Mut2) without the first 84 bp (native signal peptide) were purchased from GenScript (Piscataway Township, NJ, USA). The Δ^9-tetrahydrocannabinolic acid synthase (THCAS; GenBank accession no: AB057805) from *C. sativa* was adapted to the GB system in course of previous research projects and the transcriptional units (TU) pEGB2α1 er:THCAS:6×his, pEGB2α1 cp:THCAS:6×his and pEGB2α1 er:THCAS:YFP was kindly provided by Dr. Jascha Volk [157]. The vacuolar signal sequence (*erV*), stemming from the endogenous targeting sequence of the strictosidine synthase gene from *Rauvolfia serpentina* (GenBank accession no: X62334.1) was adapted to the GB cloning system in previous projects and was kindly provided by Dr. Sabine Fräbel [115]. The basic GB parts and the GB destination vectors, the pBIN61-P19 plasmid and signal peptide sequences were provided as described in 4.1.2.

6.1.3. Oligonucleotides

All oligonucleotides used in this study were synthesized by Eurofins Genomics GmbH (Ebersberg, Germany). 4-nt overhangs generated by annealing of oligonucleotides for incorporation into the GoldenBraid system are highlighted in bold. BsmBI recognition sites are underlined.

No.	Name	Nucleotide sequence 5'–3'
20	pTHCAS H101	GCG C<u>CG TCT C</u>AC TCG **AGC C**AA TCC TCG AGA AAA CT
21	pTHCAS M201	GCG C<u>CG TCT C</u>AC TCG **CTG C**AT GAT GAT GCG GTG GA
22	SeqTHCAS/CBDAS/CBCASfor	GCG TAG GTG GAC ACT TTA GTG GAG G
23	pCBCAS/pTHCAS cPCR 102	AAC TGC GTG GGT TGA AGC CGG AGC T
24	pCBCAS/pTHCAS cPCR 202	GCC CAT GTA TCT CCA TGT TC
25	pCBDAS cPCR102	AAC TGC ATG GGT TGA AGC CGG AGC T
26	pCBDAS cPCR202	GTA GAC TTT GGG ACA GCA ACC AG
27	pTHCAS Y175F fw	TCC TGG TGG GTT TTG CCC TAC TG
28	pTHCAS Y175F rev	AAA CTA AGA TTC TCA TTC TTC TCA TTG ATC C
29	pTHCAS/CBCAS Y484F fw	GTA TCT CAA TTT TAG GGA CCT TG
30	pTHCAS/CBCAS Y484F rev	GCC AAT CTT GGA TTT TGG
31	pTHCAS Y354F fw	AAC CAT CTT CTT CAG TGG TGT TG
32	pTHCAS Y354F rev	GTA TCA ATC CAG CTA AAT TC
33	pCBCAS E40K fw	ATG CTT CTC GAA ATA TAT TCC TAA C
34	pCBCAS E40K rev	TTA AGG AAG TTT TCT TGA GG
35	pCBCAS Y41H fw	CTT CTC GGA ACA TAT TCC TAA CAA TC
36	pCBCAS Y41H rev	ATA TAT TCC GAG AAG CAT TTA AG
37	pCBCAS P46V fw	TCC TAA CAA TGT AGC AAA TCC AAA ATT C
38	pCBCAS P46V rev	ATA TAT TCC GAG AAG CAT TTA AG
39	pCBCAS V63I fw	GTA TAT GTC TAT CCT GAA TTC GAC
40	pCBCAS V63I rev	GTG AGT TGT GCT GGT TAA
41	pCBCAS T74I fw	TCT TAG ATT CAT CTC TGA TAC AAC
42	pCBCAS T74I rev	TTT TGT ATT GTC GAA TTC AG
43	pCBCAS V90N fw	TCC TTC AAA TAA CTC CCA TAT CCA G
44	pCBCAS V90N rev	GTG ACA ATA ACG AGT GGT TTT G

45	pCBCAS I129V fw	CCC ATT TGC TGT AGT AGA CTT GAG
46	pCBCAS I129V rev	ACT TGA GAT ATG TAG GAC AAA C
47	pCBCAS M288V fw	CAA AGA TTT AGT GCT CAC GAC TC
48	pCBCAS M288V rev	TCA TAC TTG TAA GCA ATA TTT TG
49	pCBCAS R296K fw	TTC AGA ACT AAG AAT ATT ACA GAT AAT C
50	pCBCAS R296K rev	GTG AGT CGT GAG CAT TAA ATC
51	pCBCAS L318H fw	TTC CAT TTT TCA TGG TGG AGT GG
52	pCBCAS L318H rev	GAG AAG TAA CCA TGT ACT G
53	pCBCAS L345F fw	TTG CAA AGA ATT CAG CTG GAT TG
54	pCBCAS L345F rev	TCA GTT TTT TTA ATA CCC AAC
55	pCBCAS Y360F fw	TGT TGT AAA TTT CAA CAC TGC TAA TTT TAA AAA G
56	pCBCAS Y360F rev	CCA CTG TAG AAG ATG GTT G
57	pCBCAS D424E fw	TGG TAT AAT GGA GGA GAT TTC AGA ATC
58	pCBCAS D424E rev	CCG TAA GGG TAC AAC ACA TAC
59	pCBCAS P494H fw	AAA AAC TAA TCA TGA GAG TCC TAA TAA TTA C
60	pCBCAS P494H rev	CCT AAA TCA AGG TCC CTA TAA TTG
61	pCBCAS E495A fw	GAA AAA CTA ATC CTG CGA GTC CTA ATA ATT AC
62	pCBCAS E495A rev	CTA AAT CAA GGT CCC TAT AAT TGA GA
63	pCBCAS E40K Y41H fw	ATG CTT CTC GAA ACA TAT TCC TAA CAA TC
64	pCBCAS P494H E495A fw	AAA AAC TAA TCA TGC GAG TCC TAA TAA TTA C

6.1.4. Molecular cloning methods

6.1.4.1. Site-directed mutagenesis (SDM)

For introduction of single and double mutations into the THCAS and CBCAS genes, site-directed mutagenesis (SDM) was performed by means of the Q5 Site-Directed Mutagenesis Kit (New England Biolabs, Ipswich, MA, USA), by utilizing Phusion DNA polymerase (Thermo Fisher Scientific Inc.) or by using Q5 Hot Start High Fidelity DNA Polymerase (New England Biolabs). While PCR protocol for the Q5 Site-Directed Mutagenesis Kit was conducted according to the manufacturer's protocol, the reaction mixtures, as well as the cycler programs for Phusion polymerase and Q5 Hot Start High Fidelity DNA Polymerase were set up as follows.

Phusion DNA Polymerase reaction mix (total 50 µL)	
Phusion High-Fidelity 5× Buffer	1×
Plasmid DNA	12.5 ng
Primer fw (see 6.1.3)	0.5 µM
Primer rev (see 6.1.3)	0.5 µM
dNTP's	0.4 mM
Phusion DNA Polymerase	2 u

Phusion DNA Polymerase cycler protocol			
Step	**Temperature**	**Duration**	
Initial denaturation	98 °C	2 min	
Denaturation	98 °C	15 sec	30 cycles
Annealing	60 °C	30 sec	
Elongation	72 °C	3 min	
Final elongation	72 °C	5 min	

Q5 Hot Start High Fidelity DNA Polymerase reaction mix (total 50 µL)	
Q5® Hot Start High-Fidelity 2× Master mix	1×
Plasmid DNA	12.5 ng
Primer fw (see 6.1.3)	0.5 µM
Primer rev (see 6.1.3)	0.5 µM

Q5 Hot Start High Fidelity DNA Polymerase cycler protocol			
Step	**Temperature**	**Duration**	
Initial denaturation	98 °C	30 sec	
Denaturation	98 °C	10 sec	25 cycles
Annealing	60 °C	30 sec	
Elongation	72 °C	1:30 min	
Final elongation	72 °C	2 min	

6.1.4.2. Kinase, ligase and DpnI (KLD) treatment

After amplification the SDM products were circularized using the Q5® Site-Directed Mutagenesis Kit (New England Biolabs) according to the manufacturer's instructions or by performing all enzymatic reactions separately in the order as listed in the following protocol. The Polynucleotide Kinase Kit by Thermo Fisher Scientific Inc. was used for the phosphorylation reaction. T4 DNA ligase and T4 DNA ligase buffer were purchased from Promega, while DpnI fast digest as well as fast digest buffer were procured from Thermo Fisher Scientific Inc.

Reaction	**5'-phosphorylation (total 10 µL)**	**Ligation (total 10 µL)**	**DpnI treatment (total 20 µL)**
Mixture	20 ng PCR product	5 µL phosphorylated DNA	10 µL ligation mixture
	1× Reaction buffer A	1× T4 DNA ligase buffer	1× Fast digest buffer
	1 mM ATP	1 u T4 DNA ligase	2 u DpnI fast digest
	1 u T4 polynucletoidkinase		
Incubation	37 °C, 20 min	16 °C, overnight	37 °C, 2 h
Inactivation	75 °C, 10 min	75 °C, 10 min	75 °C, 10 min

6.1.5. Protein biochemical methods

6.1.5.1. PNGase F and PNGase A assays

The deglycosylation reaction of THCAS was performed with PNGase F or PNGase A (New England Biolabs), according to the manufacturer's instructions. Therefore, 500 u of PNGase F or 5 u of PNGase A were mixed with 20 or 5 µg TSP extracts, respectively. The resulting sample was then separated on 10 % SDS–PAGE and subjected to Western blot analysis (see 3.4.6.3).

6.1.5.2. THCAS/CBCAS/CBDAS activity assays with plant crude extract

150 mg of frozen plant material was homogenized in 500 µL of THCAS/CBDAS/CBCAS reaction buffer (100 mM trisodium citrate, pH 5.5) at RT and centrifuged at 17,000 ×*g* for 15 min. Thereupon, *in vitro* activity assays were performed with the addition of different amounts of substrate or solvent into the mixture. The reaction mixtures contained 92.8 µL of the extracted total soluble proteins as well as 0.05−0.2 mM of CBGA and ACN in the range of 1.8−7.2 % (v/v) in a total volume of 100 µL. The reaction took place for 2 h at 37 °C. To terminate the reactions, 275 µL of ice-cold acetonitrile were added, followed by incubation on ice for 30 min. Finally, the supernatants were purified two times from solid particles by centrifugation (17,000 ×*g*, 30 min, 4 °C). THCA, CBCA and CBDA were then analyzed by using HPLC−MS (see 3.5.1). Assays were performed in four biological replicates.

6.1.5.3. THCAS/CBCAS/CBDAS activity assays with purified proteins

Reaction mixtures contained 1.25 µg of er:THCAS:6×his, 2.5 µg of er:CBDAS:8×his or 2.5 µg of er:CBCAS:8×his as well as 0.05 mM of CBGA and an appropriate solvent (acetonitrile (ACN), acetone, dichloromethane (DCM), isopropanol, EtOH and DMSO) or monoterpene (myrcene, linalool, α-pinene and limonene) in the range of 1.8−7.2 % (v/v) in a total volume of 100 µL buffered in THCAS/CBDAS/CBCAS reaction buffer (100 mM trisodium citrate, pH 5.5). The reaction took place for 2 h at 37 °C. To terminate the reactions, 275 µL of ice-cold acetonitrile was added, followed by incubation on ice for 30 min. Finally, the supernatants were purified two times from solid particles by centrifugation (17,000 ×*g*, 30 min, 4 °C). Assays were performed in four replicates and analyzed by HPLC−MS (see 3.5.1).

6.1.5.4. Quantification of enzyme activity

In general, the catalytic activity of enzymes can be described in katal (kat), where one katal refers to an enzyme catalyzing one mole of substrate per second. However, since THCAS and CBDAS produce two different products and the specific activity towards the production of each cannabinoid should be determined, substrate conversion was equated with product formation based on the fact that one molecule of CBGA is converted to one molecule of THCA, CBCA or CBDA [64,65,77]. Therefore, the amount of formed product was quantified by peak integration at the wavelength of 260 nm and subsequent conversion from area counts into katal, calculated by means of serial standard dilutions (THCA 800 nM–0.05 mM; CBDA 800 nM–0.05 mM; CBCA 6.0 nM–0.05 mM). The specific enzyme activities were calculated with:

$$kat = \frac{mol}{s}$$

and expressed in kat per gram fresh weight (kat g_{FW}^{-1}) in case of using proteins from crude plant extracts or in kat per mg protein (kat mg^{-1}) in case of utilizing purified proteins. For mutagenesis studies, the activity of each mutant was normalized on the average activity of their corresponding wild-type enzyme.

6.1.5.5. Peroxide assays

To investigate the oxygen dependency in the production of CBCA, activity assays with purified proteins were performed as described in 6.1.5.3. One half of the assay volume (A) was used for peroxidase assays without terminating the reaction, whereas the other half (B) was terminated after a total assay duration of 2.5 h (directly after measurement of H_2O_2 concentration of A) and subjected to HPLC–MS analysis. The peroxidase assays were performed with Pierce Quantitative Peroxidase: Aqueous Compatible Formulation Kit (Thermo Fisher Scientific Inc.) according to the manufacturer's instructions. The concentration of hydrogen peroxide was determined by means of a serial standard solution in the range of 1−1000 µM. Reaction mixtures without the substrate CBGA served as a negative control. Finally, the amount of produced CBCA and hydrogen peroxide were calculated and compared to each other by using the equation:

$$n_x = \frac{m_x}{M_x}$$

6.1.6. Software based analysis

6.1.6.1. CBGA docking simulation

The available crystal structure of THCA synthase (pdb entry 3vte.1.A) does not include its substrate CBGA. In order to investigate the reaction mechanism a static docking using AutoDock Vina was performed: The structure of CBGA was obtained from the ZINC database 3 and converted into the PDBQT format required by AutoDock Vina using Pybel 4 [158–160]. The structure of the THCA synthase was converted into PDBQT using the tools provided by AutoDock Vina. The substrate was docked in a 10 Å vicinity of the FAD *N8* atom, as this atom is probably essential for the reaction mechanism. The docking process resulted in nine models of the CBGA conformation. The model with the highest score was coincidentally the conformation with the highest similarity to the proposed binding mechanism [67]. The hydrogen atoms were added to the selected docked CBGA structure *via* a custom Python script using the bioinformatics library Biotite [161]. Finally, the protein model was visualized by means of the UCSF Chimera software [162].

6.1.6.2. Homology modeling of CBCAS

Modelling of the three dimensional structure of CBCAS was performed using the fully automated protein homology modelling server SWISS-MODEL [163]. Therefore, the CBCAS amino acid sequence was submitted to the server and aligned with the sequence of THCAS (pdb entry 3vte.1.A). Based on the available crystal structure of the reference protein, comparative modelling was performed using the implemented engine ProMod3. Qualitative Model Energy ANalysis (QMEAN) was represented with a score of -1.13, expressing the major geometrical aspects of protein structures, while the Global Model Quality Estimation (GMQE) had a score of 0.94, describing the expected accuracy of a model built with that alignment and template [163,164]. The protein model was visualized by means of the UCSF Chimera software [162].

6.2. Results and discussion

6.2.1. Molecular cloning of THCAS, CBDAS and CBCAS

Although complete cannabinoid biosynthesis *in vivo* without addressing the assumed bottleneck in geranyl diphosphate (GPP) availability was not feasible, it was nevertheless considered a milestone as it was possible to produce all the enzymes involved in the cannabinoid biosynthetic pathway heterologously in *N. benthamiana* and to demonstrate their activity. Furthermore, it was even succeeded to form olivetolic acid (OA) and its glucosides *in vivo* (Figure 4.2.4). To finally round off the whole story of the producibility of all enzymes involved in the synthesis and to further investigate the biosynthesis of the individual cannabinoids, Δ^9-tetrahydrocannabinolic acid synthase (THCAS), cannabidiolic acid synthase (CBDAS) and cannabichromenic acid synthase (CBCAS) had to be introduced in the GoldenBraid (GB) cloning system. Even though it has been reported that THCAS is secreted into the extracellular space of the *Cannabis* glandular trichomes and thus being processed *via* the secretory pathway, different subcellular localizations have been included in the engineering setup here as well, to definitively rule out any alternative localization possibilities [109]. In order to domesticate the cannabidiolic acid synthase (CBDAS) and the cannabichromenic acid synthase (CBCAS) genes, GB–modified sequences without the first 84 bp (signal peptide) were synthesized by Integrated DNA Technologies harboring the appropriate fusion sites (AGCC and GCAG) and BsmBI recognition sites. Since in the case of Δ^9-tetrahydrocannabinolic acid synthase (THCAS) transcriptional units (TUs) for the localization of the enzyme to the apoplast and in to chloroplasts had already been generated in course of previous studies (pEGB2α1 er:THCAS:6×his, pEGB2α1 cp:THCAS:6×his and pEGB2α1 er:THCAS:YFP; [157]), pEGB2α1 er:THCAS:6×his was used as a template to incorporate fusion overhangs (AGCC and GCAG) by PCR for cloning of cytosolic and vacuolar localized THCAS (see 3.2.1 and 6.1.3; oligonucleotides 20 and 21). Subsequently, the PCR product was purified (see 3.2.2) and ligated into the universal domesticator *via* a GB reaction (see 3.2.3). The resulting GB reaction mixtures were transformed into chemically competent *E. coli* TOP10 (see 3.2.4), followed by isolation of plasmid DNA (see 3.2.5), restriction analysis (see 3.2.6) and verification by sequencing (see 3.2.8; oligonucleotides 1, 2 and 22).

Again, various TUs were generated for the localization of the enzymes in different compartments of the plant cells. Therefore, for assembly of transcriptional units in α-level plasmids, the coding sequence (CDS) of CBDAS and CBCAS were either fused with the chloroplast transit peptide sequence (cp) or the ER/apoplast signal peptide sequence (er), while in case for production in the cytoplasm, the CDS was directly fused to the P35S ATG. Since for THCAS, the apoplast and chloroplast localized versions were already cloned as mentioned above, only THCAS constructs containing a vacuolar signal sequence (erV) as well as the cytosolic localized THCAS were assembled during the current

project. Unfortunately, due to time constraints only the THCAS was fused with the erV targeting sequence. Furthermore, for visualization of proteins during fluorescence microscopy, *yfp* was C−terminally fused to the enzymes. The assembled GB constructs were then transformed into chemically competent *A. tumefaciens* EHA105 cells (see 3.2.9) and positive clones were confirmed by colony PCR (see 3.2.10) using the relevant primer pairs (see 6.1.3; oligonucleotides 23−26). Finally, glycerol stocks of overnight cultures were prepared and stored at -80 °C. Figure 6.2.1 illustrates the expression constructs generated particularly for this part of the project.

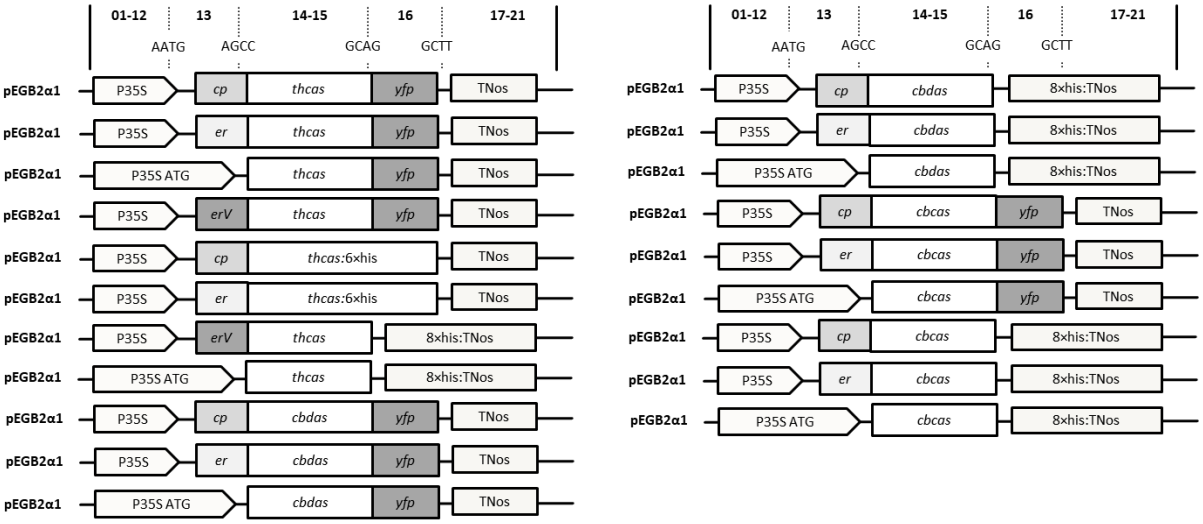


Figure 6.2.1 Cloning of different THCAS, CBDAS and CBCAS expression cassettes using the GoldenBraid cloning system. The capital letters show the four-nucleotide overhangs ensuring correct final orientation within the transcriptional unit (TU), while the numbers above the scheme represent standard GoldenBraid classes within the TU structure [93]. P35S, cauliflower mosaic virus (CaMV) 35S promoter; P35S ATG, cauliflower mosaic virus (CaMV) 35S promoter with an integrated start codon ensuring cytosolic localization; TNos, nopaline synthase terminator; 8×his:TNos, nopaline synthase terminator comprising an 8×his-tag; cp, chloroplast targeting sequence; er, ER/apoplast signal peptide; erV, vacuolar signal sequence. Boxes are not drawn to scale.

A detailed description of all chronological steps performed during assembly of GB constructs and transformation of *A. tumefaciens* used for transformation of tobacco plants is given in chapter 4.2.1.

6.2.2. Heterologous production of late biosynthetic enzymes in *N. benthamiana* plants

6.2.2.1. Production of Δ^9-tetrahydrocannabinolic acid synthase (THCAS)

After introduction of THCAS, CBDAS and CBCAS into the GB cloning system, *N. benthamiana* plants were transiently co-transformed with Agrobacteria harboring the different THCAS constructs and P19. Five to seven days post inoculation plants were harvested and protein accumulation in various compartments was examined by fluorescence microscopy and Western blot analysis (see 3.3.2 and 3.4.6.3). Surprisingly, in addition to the fluorescence of the proteins processed *via* the secretory pathway and located in the apoplast or the so-called bulbs, representing spherical structures often observed within the lumen of vacuoles, protein accumulation was also detected in the chloroplasts (Figure 6.2.2 A) [165–167]. In parallel, using total soluble proteins extracted from plants producing THCAS with C−terminally fused his-tag, neither chloroplast targeting, nor cytosolic localization yielded in detectable protein. Only enzymes from extracts containing proteins with apoplast or vacuolar targeting could be detected with a clearly visible band at around 80 kDa (Figure 6.2.2 B). To examine the detected fluorescence in chloroplasts more deeply, Western blot analysis was also performed with crude plant extracts containing the cp:THCAS:YFP fusion. When visualized, a band was observed with a size of approximately 30 kDa, resembling the size of native YFP (Figure 6.2.2 B). The rationale behind could be that post translational modifications like glycosylation obtained in the ER might contribute to the correct folding of the enzyme as shown with THCAS produced in transgenic *Komagataella phaffii* [68]. Therefore, upon transport into chloroplasts and after cleavage of the signal peptide, the pre-protein could lack the ability to fold correctly resulting in a degradation or cleavage of the misfolded part [168]. To ultimately confirm that the heterologous plant system provides glycosylated THCAS, plant crude extracts containing er:THCAS:6×his were subjected to PNGase F treatment, catalyzing the cleavage of asparagine-linked high mannose, complex and hybrid oligosaccharides unless containing an α(1,3)-linked core fucose, or PNGase A treatment, cleaving also *N*-linked glycans with α(1,3)-linked core fucose residues (see 6.1.5.1) [169,170]. Subsequent Western blot analysis indicated that PNGase F treated proteins did not show any alteration in their protein size. In contrast, PNGase A treatment resulted in slightly smaller size of the THCAS, indicating that the protein was glycosylated when transported through the secretory system and that the sugar antenna probably contained a α(1,3)-linked core fucose residue, preventing action of PNGase F on the molecule (Figure 6.2.2 C). However, the PNGase A treated THCAS showed only a small reduction in size and appeared larger (approximately 70 kDa) than the calculated value of 59 kDa. At this point it cannot be excluded that the THCAS contains other modifications not effected by PNGase A treatment [77].

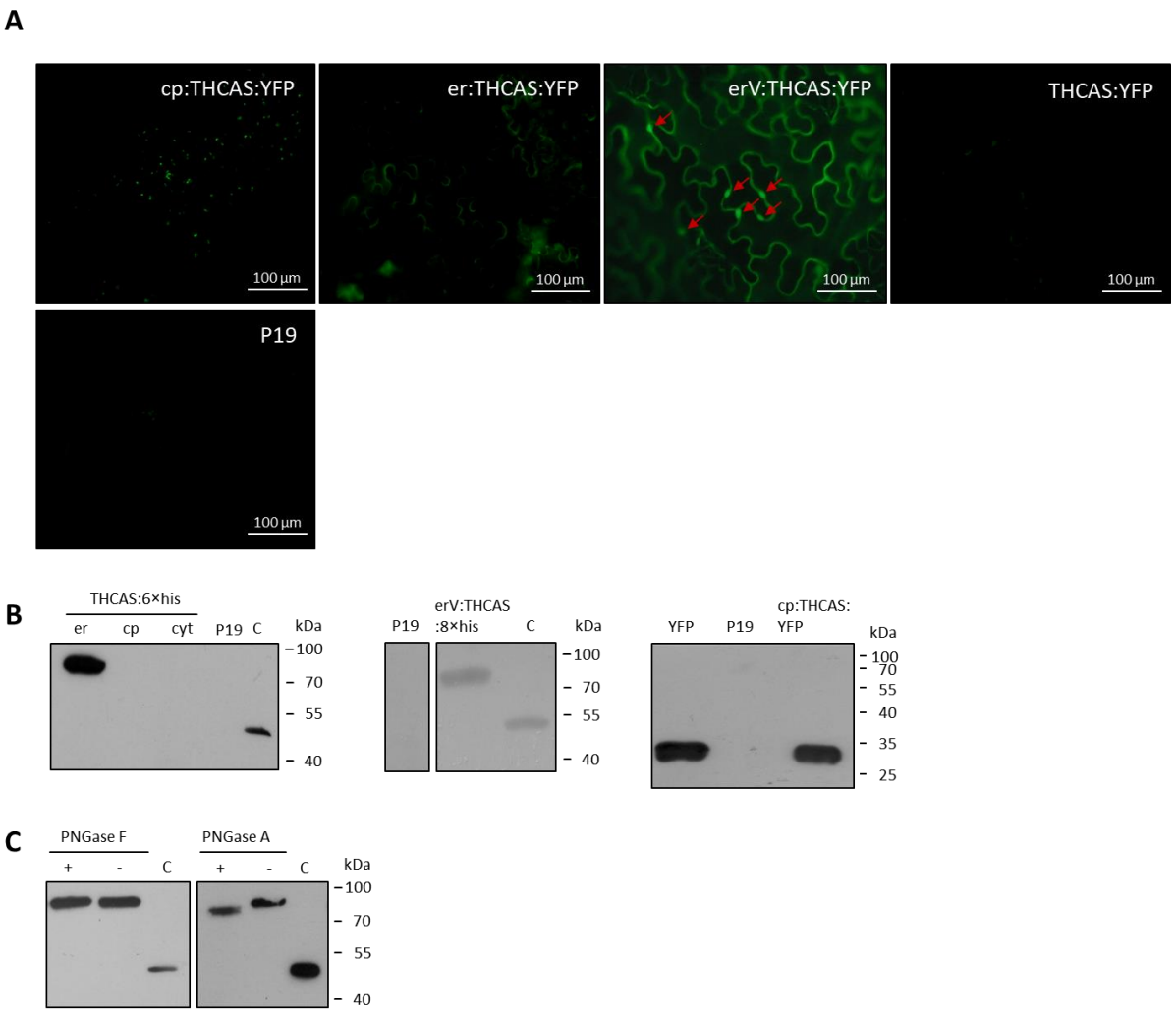


Figure 6.2.2 (A) Localization of Δ^9-tetrahydrocannabinolic acid synthase (THCAS) fused with YFP and targeted to the chloroplast (cp), apoplast (er), vacuole (erV) or without a targeting signal. Plants solely producing P19 served as a negative control. Red arrows could indicate so-called bulbs, representing spherical structures often observed within the lumen of vacuoles [165–167]. **(B)** Western blot analysis of recombinant THCAS. Heterologous produced proteins were either detected using his-probe mouse monoclonal IgG antibodies or GFP primary antibodies in case of C–terminal fused YFP proteins. A recombinant esterase (OeEst228; [121]), featured with a C–terminal 6×his tag fusion (C) as well as recombinant produced YFP served as positive controls. -, TSP extracts of plants producing solely P19. A band of around 80 kDa representing THCAS could only be detected after ER/apoplast or vacuolar targeting of the protein. **(C)** Western blot analysis of PNGase F and PNGase A treated recombinant er:THCAS:6×his. The deglycosylation reaction of er:THCAS:6×his with PNGase A resulted in a slightly lower molecular weight compared to PNGase F treated proteins as well as the negative controls without any treatment. cyt, cytosolic localization; +, treatment with PNGase; -, reaction mixture without PNGase.

6.2.2.2. Enzymatic activity of heterologous produced THCAS

To verify the enzymatic activity of THCAS produced in *N. benthamiana* plants, crude protein extracts containing the THCAS with C−terminally fused his-tag were applied to activity assays (see 6.1.5.2), followed by HPLC−MS analysis (see 3.5.1). Through comparison with authentic standards it could be observed, that only the *in planta* produced THCA synthase which was localized to the apoplast or the vacuole was active (Figure 6.2.3 and Figure S6.3.1 A). This would also be consistent with the obtained results from Western blot analysis and the statement that glycosylation obtained in the endoplasmic reticulum (ER), was essential for the correct folding of the enzyme (Figure 6.2.2 B and C) [64]. Furthermore, THCAS was not only able to produce THCA but also CBCA (Figure 6.2.3).

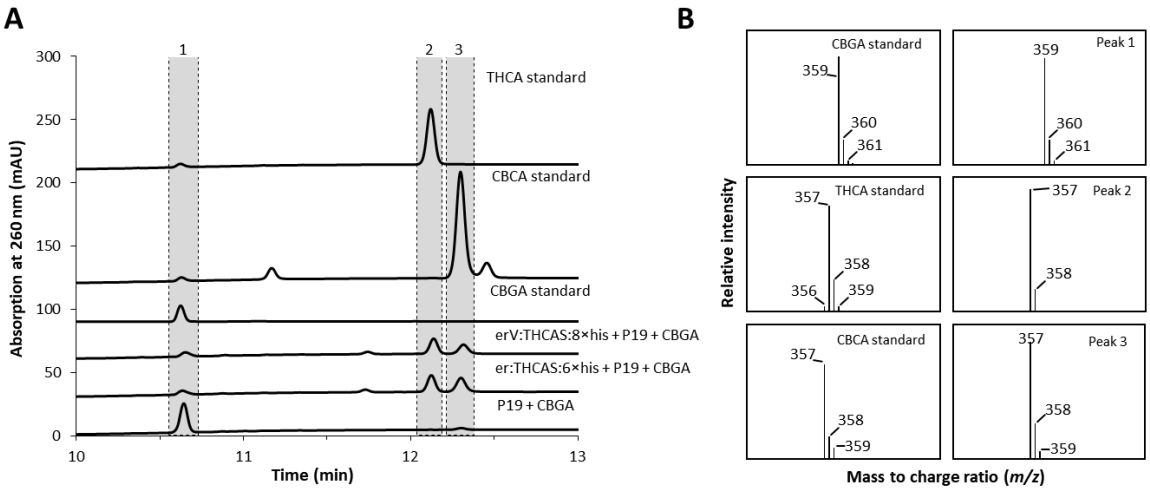


Figure 6.2.3 (A) Production of Δ^9-tetrahydrocannabinolic acid (THCA) and cannabichromenic acid (CBCA) with extracts of *N. benthamiana* plants transiently transformed with er:THCAS:6×his, erV:THCAS:8×his and P19. **(B)** Mass spectrum of the synthesized THCA and CBCA (Peaks 2 and 3, respectively) corresponding to authentic standards.

It was recently suggested that the *thcas* sequence originates from the *cbdas* sequence as the result of a gene duplication event [108]. Moreover, since the CBDAS protein sequence shares a high homology to the published THCAS and CBCAS sequences (83 % and 82 %, respectively) and even sharing the same putative catalytic base (Tyr485 in case of THCAS and CBCAS as well as Tyr484 in case of CBDAS), it was assumed that all known cannabinoid forming oxidoreductases might be derived from a common ancestor (Figure 6.2.4) [65,67]. Thus, it is not surprising that the enzymes exhibit overlapping product specificity. Hence, THCAS does convert CBGA to different products in a plant matrix, suggesting that enzyme promiscuity additionally contributes to the diversity of cannabinoids in *Cannabis* [77].

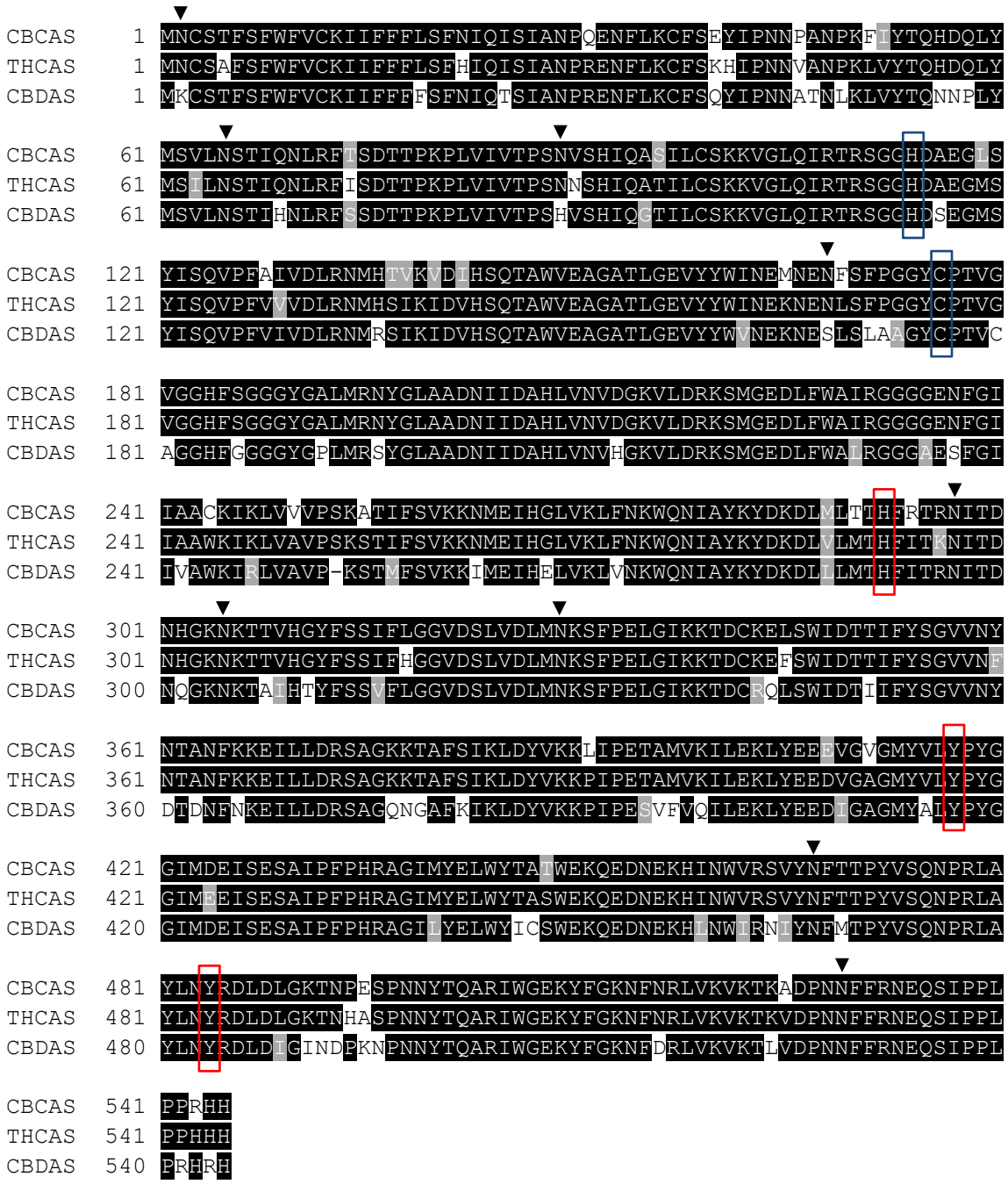


Figure 6.2.4 Multiple sequence alignment of CBCAS (sequence as published in WO 2015/196275 A1), THCAS (GenBank accession no: AB057805) and CBDAS (GenBank accession no: AB292682) using BoxShade software. Amino acids identical are represented in white letters on black background, while amino acids shaded in grey have similar characteristics/properties to those in the same position. FAD binding positions are boxed in blue, while the putative catalytic amino acids are boxed in red. Putative *N*-glycosylation sites are marked with black arrows.

6.2.2.3. Production of cannabidiolic acid synthase (CBDAS) and cannabichromenic acid synthase (CBCAS)

Besides THCAS, also CBDAS and CBCAS were heterologously produced in transiently transformed *N. benthamiana* plants. Since both enzymes share, as mentioned before, a high homology to THCAS, containing the same putative *N*-glycosylation sites, it was assumed that also CBDAS and CBCAS had to be glycosylated in the ER in order to fold properly (Figure 6.3.4). Therefore, Western blot and fluorescence analysis for particularly these two enzymes were skipped and the crude plant extracts containing CBDAS or CBCAS with C−terminally fused his-tag were directly subjected to activity assays (see 6.1.5.2). As expected, HPLC−MS analysis revealed, that also here, only the ER/apoplast localized proteins were active. Moreover, CBDAS also showed an overlapping activity as it was able to produce CBDA as well as CBCA. Only CBCAS was capable of forming just one of the products, namely CBCA (Figure 6.2.5 and Figure S6.3.1 B and C).

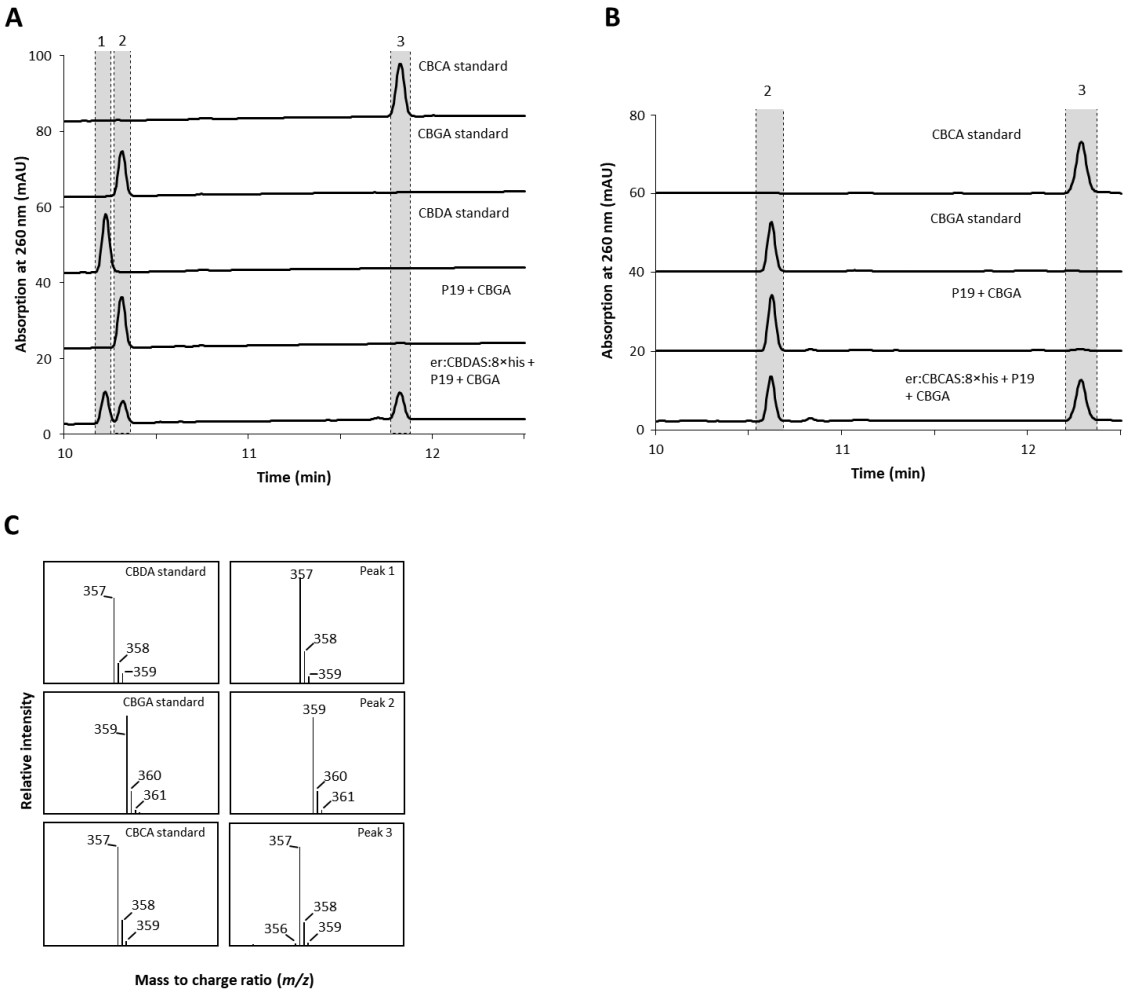
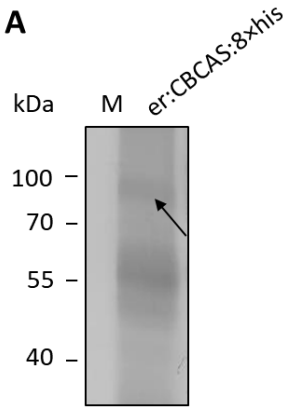
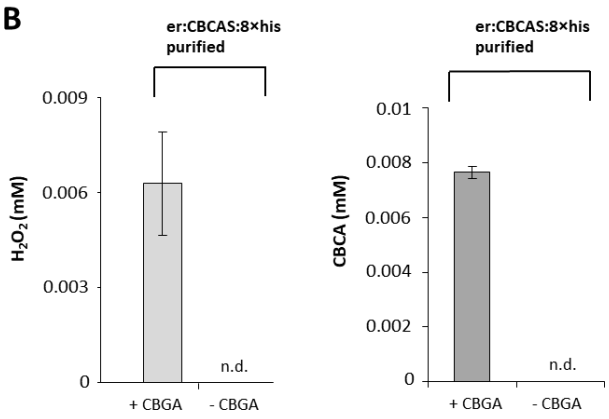


Figure 6.2.5 Production of cannabidiolic acid (CBDA) and cannabichromenic acid (CBCA) with plant extracts of *N. benthamiana* plants transiently transformed with er:CBDAS:8×his and P19 **(A)** or er:CBCAS:8×his and P19 **(B)**. CBDAS showed an overlapping activity as it was able to produce CBDA as well as CBCA **(C)** Mass spectrum of the synthesized CBDA and CBCA (Peaks 1 and 3, respectively) corresponding to authentic standards.

Morimoto *et al*. claimed that the CBCA synthase does not require molecular oxygen and cofactors and is therefore not related to THCAS or CBDAS [63]. However, the production of CBCA by both CBDAS and THCAS, which require molecular oxygen for FAD regeneration, suggests that a possible CBCAS is indeed related to the other synthases. To ultimately confirm the oxygen dependency, CBCAS was heterologous produced in *N. benthamiana* (see 3.3.1) and purified *via* immobilized metal ion affinity chromatography (IMAC) (see 3.4.2). The elution fractions were then separated on SDS−PAGE (see 3.4.5) and visualized by Coomassie Brilliant Blue staining (see 3.4.6.1). The desired proteins with the calculated size of approximately 80 kDa were present in elution fractions 1 and 2 (data not shown). Hence, both fractions were combined and replaced with THCAS/CBDAS/CBCAS buffer (100 mM trisodium citrate, pH 5.5), followed by concentration of the proteins. Afterwards, samples were quantified by BCA assay and diluted in 25 % (v/v) glycerol to a final concentration of 0.5 mg/mL (see 3.4.3). Finally, quality of the purified proteins was analyzed by loading 0.5 µg of each protein onto SDS−PAGE, followed by staining with silver nitrate (see 3.4.5 and 3.4.6.2). As shown in Figure 6.2.6 A a weak band with the size of approximately 80 kDa, representing the desired protein, was detected.

2.5 µg of the purified protein were then utilized in activity assays (see 6.1.5.3) as well as in hydrogen peroxide assays (see 6.1.5.5). Metabolite concentrations were then calculated by means of serial standard dilutions (CBCA 6.0 nM–0.05 mM and H_2O_2 1−1000 µM). Reaction mixtures without the substrate CBGA served as a negative control. When comparing the amount of CBCA and H_2O_2 produced, it was observed that approximately equal amounts of the two metabolites were formed (Figure 6.2.6 B), confirming that also CBCAS is oxygen dependent, most likely utilizing FAD as a co-factor.





% ACN (v/v)	Amount of CBCA (pmol)	Amount of H_2O_2 (pmol)
1.8	0.382 ± 0.011	0.317 ± 0.082

Figure 6.2.6 (A) SDS−PAGE of purified er:CBCAS:8×his visualized by silver nitrate staining. Arrows indicate the desired proteins with the expected size of approximately 80 kDA. Numbers to the left indicate molecular mass of marker proteins (M). **(B)** Measurement of hydrogen peroxide and CBCA concentrations after CBCA synthase activity assays. Based on the calculated amounts of hdyrogen peroxide and cannabigerolic acid produced, it was observed that hydrogen peroxide release appears to be proportional to CBCA formation. The assays containing no CBGA as substrate served as a negative control. N.d. , not detected.

In summary, regarding the reconstitution of the biosynthetic pathway, this work showed that all enzymes involved in cannabinoid biosynthesis could be manufactured in *N. benthamiana* and their activity could be demonstrated. For future engineering efforts in *N. benthamiana*, however, bottleneck alleviation (see Chapter 5), maximizing pathway flux through metabolic channeling, as well as the optimal supply of co-factors would be expedient [171]. Moreover, although it was shown in recent studies that injection of CBGA into leaves of *N. benthamiana* expressing *thcas* containing its native signal peptide (leading to secretion of the protein into the apoplast) resulted in production of THCA, it was not possible to conclude from these experiments how endogenously produced CBGA would behave and whether it would also be secreted from the cell [124]. Consequently, localization of late biosynthetic enzymes within the cell, like in the vacuole (Figure 6.2.2), has to be strongly considered as it is already the case in engineered yeast [60].

6.2.3. In-depth investigation of late biosynthetic enzymes

6.2.3.1. Examination of the enzymes product specificity

Surprisingly, during initial experiments, a change in product specificity from THCA to CBCA was observed upon addition of different amounts of substrate during assays with crude plant extracts containing er:THCAS:6×his (chromatograms not shown). To further investigate this occurrence, the activity of er:THCAS:6×his towards production of THCA and CBCA was quantified and expressed in kat per gram fresh weight (see 6.1.5.4). Thus, increasing the concentration of CBGA from 0.05 mM to 0.2 mM in the assay mixtures resulted not only in increased production of THCA (2.5 ± 0.24 fkat g_{FW}^{-1} to 4 ± 0.25 fkat g_{FW}^{-1}), but also resulted in almost twice as much CBCA (7.5 ± 0.34 fkat g_{FW}^{-1}) being formed compared to THCA (4 ± 0.25 fkat g_{FW}^{-1}) (Figure 6.2.7).

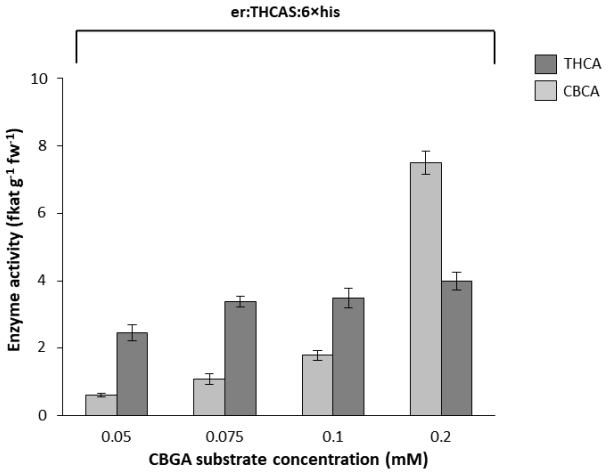


Figure 6.2.7 Quantification of specific enzyme activity in *in vitro* assays using crude extracts of transiently transformed *N. benthamiana* plants containing er:THCAS:6×his and different CBGA amounts.

However, since the substrate was dissolved in acetonitrile (ACN), it was necessary to clarify whether the product change is related to the increased amount of substrate available or dependent on the solvent. Therefore, different amounts of ACN in the range of 1.8−7.2 % (v/v) were added to the assay mixtures while maintaining constant amounts of substrate (see 6.1.5.2). This clearly resulted in a product switch that was caused by the modified solvent concentration and not the concentration of substrate (Figure 6.2.8 A). Zirpel *et al.* already demonstrated that the product specificity of THCAS and CBDAS strongly depends on the pH of their environment. When the assay mixtures were shifted from acidic (pH 4.5) to more basic conditions (pH 7.5) THCAS and CBDAS changed their product formation from THCA and CBDA, respectively, to CBCA [172]. To investigate if the addition of ACN had an impact on the pH in the reaction mixtures leading to the switch of product specificity, assay buffer (pH 5.5), buffer containing 1.8 % (v/v) ACN or buffer containing 7.2 % (v/v) ACN were dribbled

onto pH test strips. When compared to the control (assay buffer with pH 5.5), no change of pH has been observed, when ACN was added to the buffer (Figure 6.2.8 B). Thus, a product change caused by an altered pH value could be excluded. To further examine whether CBDAS and CBCAS also change their product specificity when supplemented with ACN, *N. benthamiana* plant crude extracts containing er:CBDAS:8×his or er:CBCAS:8×his were subjected to activity assays harboring different amounts of ACN and 0.05 mM of CBGA. HPLC−MS analysis and subsequent quantification of enzyme activity showed that when the solvent concentration was increased, the formation of CBCA catalyzed by CBDAS was at least maintained at the same level, while CBDA production could no longer be detected (solvent concentration 7.2 % (v/v)). In case of CBCAS, on the other hand, the addition of solvents solely led to the abolishment of the enzymes activity (Figure 6.2.8 C).

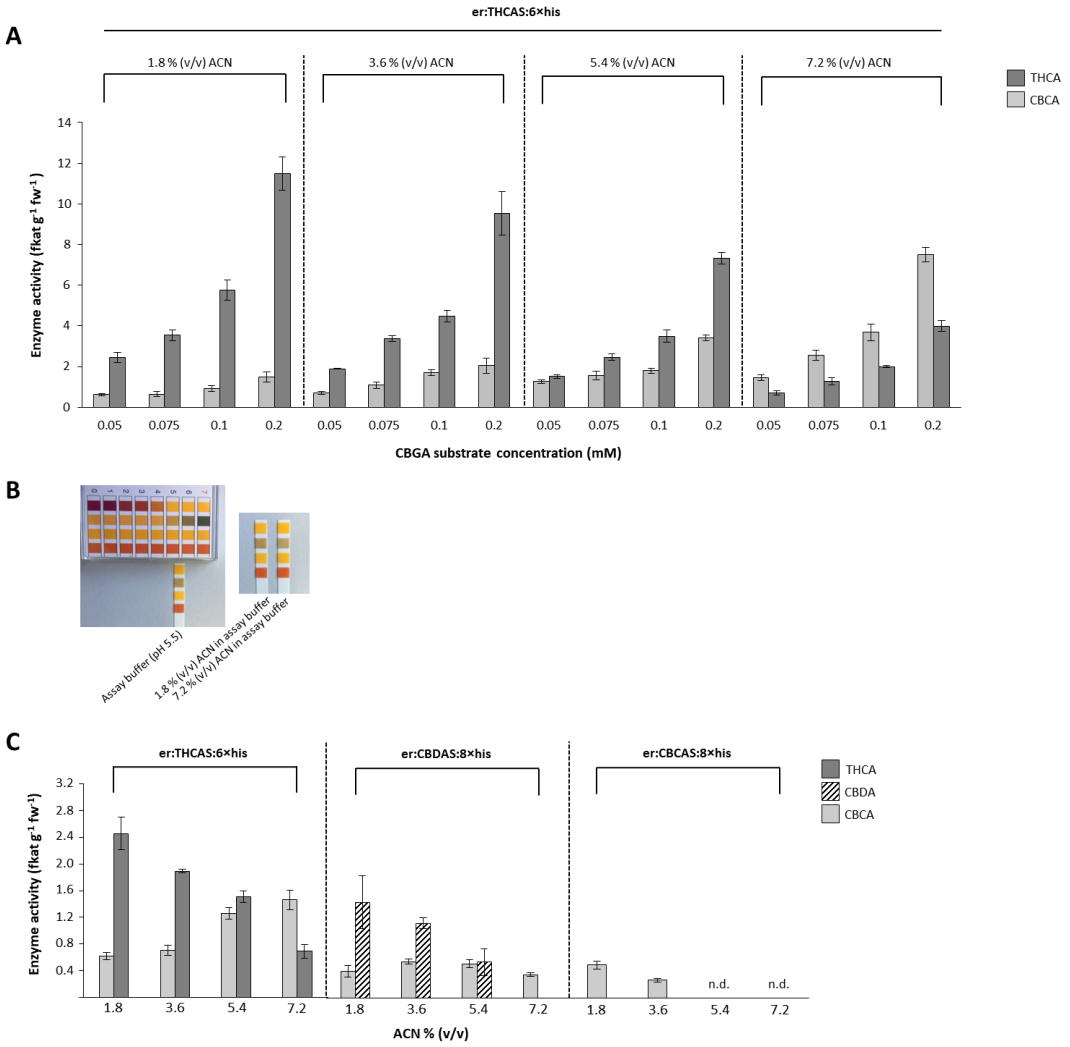


Figure 6.2.8 (A) Quantification of specific enzyme activity in *in vitro* assays using plant crude extracts containing er:THCAS:6×his and various amounts of substrate as well as acetonitrile (ACN). **(B)** Determination of the pH value of assay buffer containing ACN. **(C)** Quantification of specific enzyme activity in *in vitro* assays using plant crude extracts containing er:THCAS:6×his, er:CBDAS:8×his or er:CBCAS:8×his and various concentrations of ACN. In case of er:THCAS:6×his and er:CBDAS:8×his, a product switch towards CBCA due to the increased ACN addition, could be observed. N.d. , not detected.

Since the assays were performed with extracts containing all kinds of soluble proteins as well as water soluble metabolites, it was further investigated that the product specificity upon addition of solvents was not influenced by other tobacco proteins still present in the crude extract leading for example to protein-protein interactions [173]. Therefore, also er:THCAS:6×his and er:CBDAS:8×his were heterologous produced in *N. benthamiana* (see 3.3.1) and purified as described before for CBCAS *via* immobilized metal ion affinity chromatography (IMAC) (see 3.4.2). Finally, quality of the purified proteins was analyzed by loading 0.5 µg of each protein onto SDS−PAGE, followed by staining with silver nitrate (see 3.4.5 and 3.4.6.2). As shown in Figure 6.2.9 the expected bands with the size of approximately 80 kDa, representing the desired proteins, were detected in both cases.

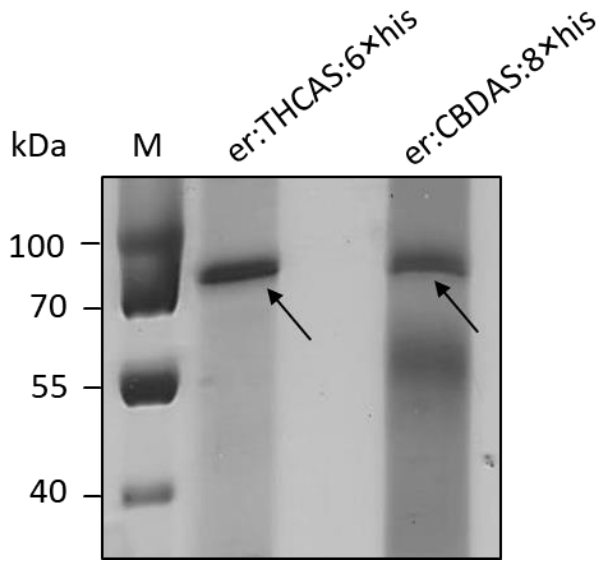
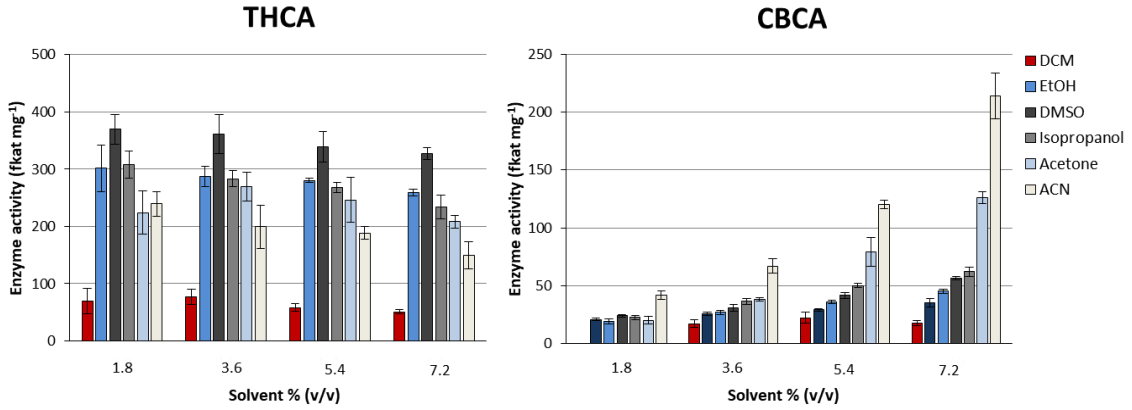


Figure 6.2.9 SDS−PAGE of purified er:THCAS:6×his and er:CBDAS:8×his visualized by silver nitrate staining. Arrows indicate the desired proteins with the expected size of approximately 80 kDA. Numbers to the left indicate molecular mass of marker proteins (M).

The purified proteins were then used for activity assays with addition of different organic co-solvents. The reaction mixtures contained 1.25 µg of er:THCAS:6×his, 2.5 µg of er:CBDAS:8×his or 2.5 µg of er:CBCAS:8×his as well as 0.05 mM of CBGA and an appropriate solvent in the range of 1.8−7.2 % (v/v) in a total volume of 100 µL buffered in THCAS/CBDAS/CBCAS reaction buffer (100 mM trisodium citrate, pH 5.5). After incubation for 2 h at 37 °C, the reactions were terminated with 275 µL of ice-cold acetonitrile (see 6.1.5.3). Subsequently, THCA, CBDA and CBCA formation was analyzed by HPLC−MS (see 3.5.1). The addition of different co-solvents clearly indicated their impact on the product specificity, at least for THCAS and CBDAS and most strikingly when ACN was used (Figure 6.2.10, Figure 6.2.11 and Figure 6.2.12). In literature, the reasons for alterations in the relative activity of enzymes, caused by the addition of organic co-solvents, are widely discussed.

Some of the most important are among others: (i) That organic co-solvents have an influence on the general properties of the buffer, like its polarity or the hydrophobicity, thus leading to a change in the solvation of the ground and/or transition states of the enzymatic reaction which in turn alters electrostatic, hydrophobic or other interactions of the enzyme with its substrate. (ii) That organic co-solvents cause a change in the protein conformation and its flexibility [174–177]. Therefore, it could be assumed that a conformational change in THCAS and CBDAS promoted a product switch to CBCA. Solely in case of dichloromethane (DCM) added into the reaction mixtures, enzyme activities were quiet low and no overall switch of product specificity could be observed (Figure 6.2.10, Figure 6.2.11 and Figure 6.2.12). This could be explained by the fact that DCM is insoluble in water and thus forming an aqueous-organic biphasic system, which in turn may not have had any effect on the conformation of the proteins and their product specificity. Moreover, the lower activity could be a consequence of the poorer accessibility of the proteins to the substrate CBGA dissolved in DCM.

A

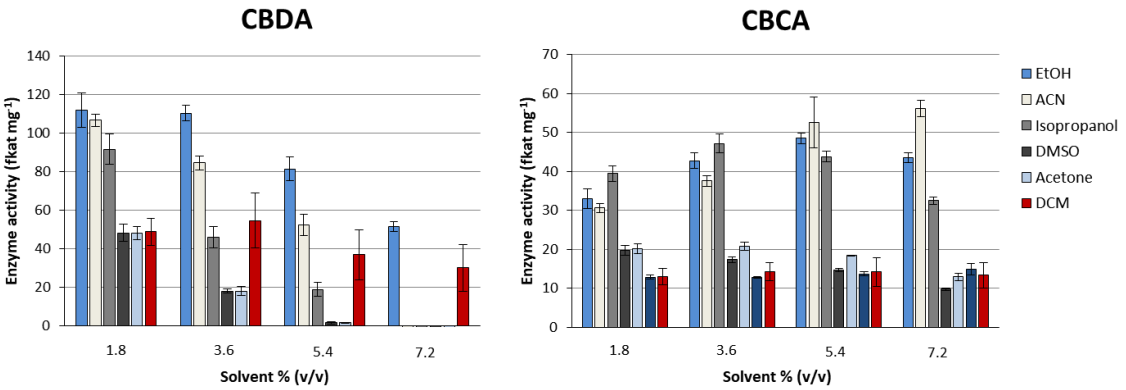


B

er:THCAS:6×his						
Solvent	**ACN**	**Acetone**	**DCM**	**Isopropanol**	**EtOH**	**DMSO**
	THCA : CBCA	THCA : CBCA	THCA : CBCA	THCA : CBCA	THCA : CBCA	THCA : CBCA
1.8 % (v/v)	5.7 : 1	11 : 1	only THCA	13.7 : 1	15.8 : 1	15.4 : 1
3.6 % (v/v)	3 : 1	7 : 1	4.5 : 1	7.8 : 1	10.7 : 1	11.6 : 1
5.4 % (v/v)	1.6 : 1	3.1 : 1	2.6 : 1	5.4 : 1	7.8 : 1	8.2 : 1
7.2 % (v/v)	1 : 1.5	1.6 : 1	2.9 : 1	3.8 : 1	5.7 : 1	5.8 : 1

Figure 6.2.10 (A) Detection of enzyme activity of purified er:THCAS:6×his produced in transiently transformed *N. benthamiana* plants and with the addition of different solvents into the reaction mixture**.** *In vitro* assays of purified er:THCAS:6×his revealed a change in product specificity from THCA to CBCA after the addition of different solvents. The effect was most striking by supplementing higher amounts of acetone and acetonitrile (ACN) into the assay mixtures. **(B)** Alteration of the ratios of THCA to CBCA produced by er:THCAS:6×his with non-aqueous solvents applied to the assay mixtures.

A



B

er:CBDAS:8×his

Solvent	ACN	Acetone	DCM	Isopropanol	EtOH	DMSO
	CBDA : CBCA	CBDA : CBCA	CBDA : CBCA	CBDA : CBCA	CBDA : CBCA	CBDA : CBCA
1.8 % (v/v)	3.5 : 1	2.7 : 1	3.8 : 1	2.3 : 1	3.4 : 1	2.4 : 1
3.6 % (v/v)	2.3 : 1	1.3 : 1	3.9 : 1	1 : 1	2.6 : 1	1 : 1
5.4 % (v/v)	1 : 1	1 : 6.5	2.6 : 1	1 : 2.3	1.7 : 1	1 : 9
7.2 % (v/v)	Only CBCA	Only CBCA	2.3 : 1	Only CBCA	1.2 : 1	Only CBCA

Figure 6.2.11 (A) Detection of enzyme activity of purified er:CBDAS:8×his produced in transiently transformed *N. benthamiana* plants and with the addition of different solvents into the reaction mixture**.** *In vitro* assays of purified er:CBDAS:8×his revealed a change in product specificity from CBDA to CBCA. The effect was most striking by adding higher amounts of EtOH, isopropanol and acetonitrile (ACN) into the assay mixtures. **(B)** Alteration of the ratios of CBDA to CBCA produced by er:CBDAS:8×his with non-aqueous solvents applied to the assay mixtures.

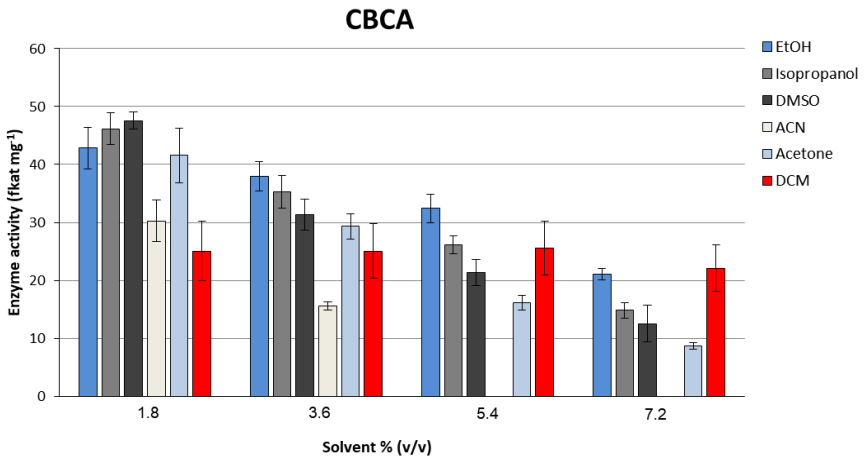


Figure 6.2.12 Detection of enzyme activity of purified er:CBCAS:8×his produced in transiently transformed *N. benthamiana* plants and with the addition of different solvents into the reaction mixture**.** *In vitro* assays of purified er:CBCAS:8×his revealed only a decrease of CBCA production.

However, organic solvents in aqueous buffers do not reflect the actual conditions of a biological system. In general, THCAS, CBDAS and CBCAS as well as metabolites like cannabinoids or terpenes are described to accumulate in the secretory cavity of *Cannabis* glandular trichomes [54,178]. Therefore, it was necessary to put the findings in a physiological context and to mimic the hydrophobic environment in glandular trichomes more accurate. Thus CBGA was dissolved in distillates of some of the main monoterpenoids (myrcene, α-pinene, limonene and linalool) present in *C. sativa* and used *in vitro* assays with purified er:THCAS:6×his (see 6.1.5.3) [49]. Subsequently, production of THCA and CBCA was analyzed by HPLC−MS (see 3.5.1). Again, the formation of a biphasic system was observed as previously in assays with DCM and no overall switch of product specificity was recognizable. The increase of monoterpene concentrations in the *in vitro* assays resulted only in a lowering of enzyme activity regarding THCA and CBCA production. When comparing the individual monoterpenes with each other, it was observed that upon the addition of limonene, no activity was detectable above a concentration of 5.4 % (v/v). The strongest effect on the enzyme activity, however, had the addition of the monoterpene linalool, for which no conversion from CBGA to THCA or CBCA could be detected even at the lowest amounts (Figure 6.2.13). However, inhibition of enzymes by terpenes is no unusual observation. In incubation experiments with boreal forest soil, the addition of monoterpenoids (α-pinene, carene and myrcene) caused a small decrease in β-glucosidase and chitinase activities. This was even more striking in *in vitro* assays and was explained by the ability of terpenes to bind to enzymes causing complex formation and precipitation of the proteins [179–183].

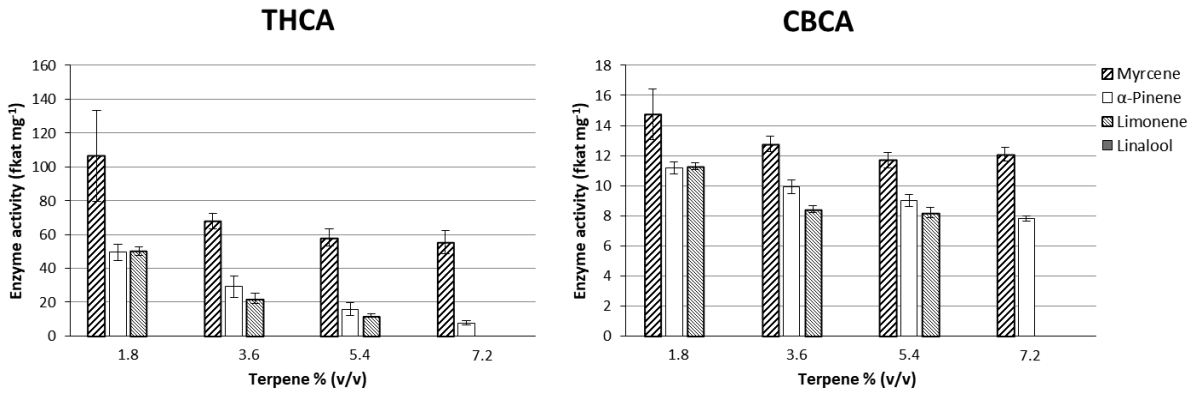


Figure 6.2.13 Enzyme activity of purified er:THCAS:6×his produced in transiently transformed *N. benthamiana* plants and with the addition of various terpenes into the reaction mixture. The increase of terpene content led to a decrease of CBCA and THCA.

Although the desired product switch caused by the THCAS in a solvent-containing lipophilic environment could not be reproduced, the apparent inhibitory property of the added monoterpenes may indicate that they could contribute to some extent to the amount of cannabinoids produced in trichomes of *C. sativa*. However, it must also be taken into account that monoterpenes represent only a part of the constituents in trichomes. Analysis of the exudates of *C. sativa* capitate-stalked glandular trichome storage cavities showed that next to aliphatic hydrocarbons and terpenoids other amphiphilic and hydrophilic compounds like fatty acid amines, fatty alcohols, carboxylic acids, amino acids, sugars and polyols were identified [184]. Therefore, it has to be considered that the heterogeneity in trichome composition is potentially influencing or interrupting the reactions of terpenes with THCAS. However, in order to address the problem with the more accurate representation of the trichome environment, future experiments could incorporate natural deep eutectic solvents (NADES) into the assay mixture. NADES comprise natural compounds, such as organic acids, amino acids and different sugar species and would partially explain how nature formulates lipophilic ingredients without assistance of organic solvents [185]. Thus *in vitro* assays containing for example NADES combined with water insoluble natural products like terpenes and the substrate CBGA could be performed.

Nevertheless, the obtained results could be the first indication, that not only the intracellular pH or the gene expression levels of the different synthases are crucial for the decision which cannabinoids are produced, but that also the general hydrophobic environment of trichomes as well as the interplay of various compounds with the proteins could contribute to the cannabinoid diversity in different *Cannabis* strains [172,186].

6.2.3.2. Docking of CBGA into the putative active site of THCAS

Since the crystal structure of THCAS as well as the amino acids involved in the catalytic activity are known, it was considered to conduct molecular dynamics (MD) simulations of THCAS binding to its substrate CBGA either in presence of water or acetonitrile, to further investigate if the product switch is caused by a potential conformational change of the enzyme upon addition of solvents [67]. However, first of all docking of CBGA into THCAS had to be performed, as no tertiary structure of the enzyme in complex with its substrate was available. For docking simulation, the crystal structure of THCA synthase (pdb entry 3vte.1.A) as well as the structure of CBGA obtained from the ZINC database 3 was used. Afterwards, the protein model was visualized by means of the UCSF Chimera software (see 6.1.6.1). The initial docking simulation of THCAS in aqueous solution revealed that Tyr484, assumed to be involved in the catalytic activity of THCAS by deprotonation of the hydroxyl group at the O6' position of CBGA, seems to be constantly far away from the O6' (average distance of 7.7 Å) [67]. Moreover, other atoms of CBGA (C10-tail) would sterically prevent the deprotonation at the O6' atom (Figure 6.2.14 upper left picture). However, the amino acids Tyr175 and Tyr354, which are adjacent and exposed to the active site cleft, also did not appear to be in a better steric position (Figure 6.2.14 upper right picture and lower picture).

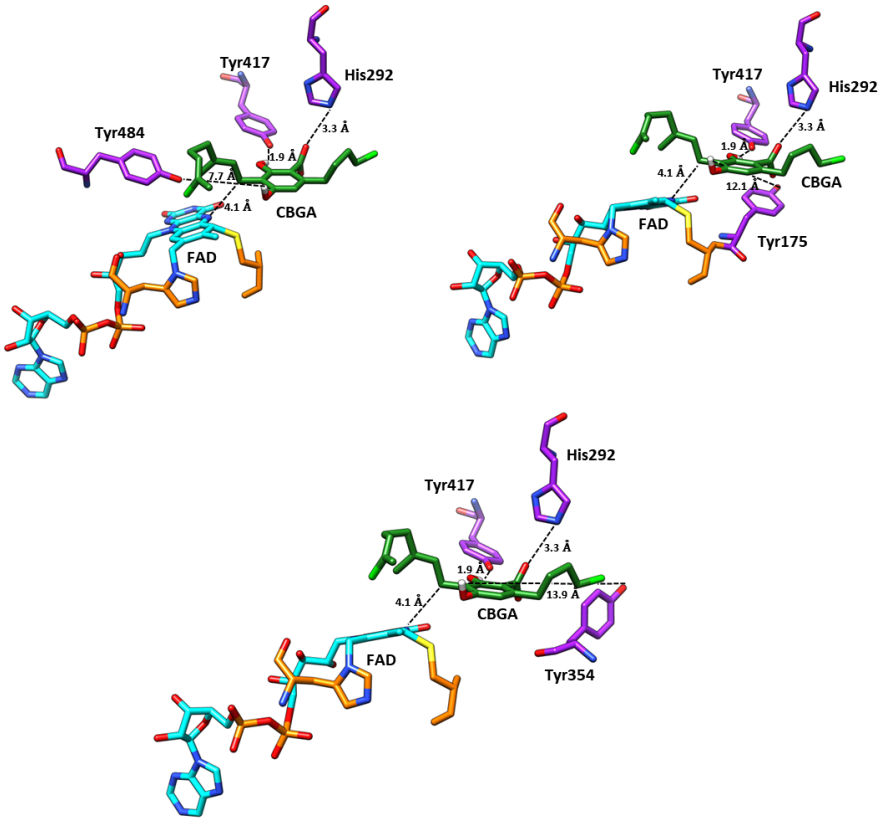


Figure 6.2.14 Substrate docking of CBGA into the putative catalytic site of THCAS WT. FAD is colored in cyan, while CBGA and putative catalyctic amino acids are colored in green and purple, respectively. The image was generated using UCSF chimera [162].

Consequently, based on the modeling results, two assumptions could be made: (i) The obtained crystal structure without the bound substrate is far off from an active conformation, meaning that without co-crystallization of THCAS and its substrate CBGA, further MD simulations could not deliver conclusive results regarding the conformational change of the protein when solvent would be added. (ii) Mutagenesis studies have shown that exchange of tyrosine at position 484 to phenylalanine abolished the activity of the mutant enzyme. However, in the case of amino acids Tyr354 and Tyr175, it was suspected only on the basis of structural data that they do not have a direct role in the catalytic process, but perhaps in substrate binding [67]. Therefore, it could also be possible that either Tyr354 or Tyr175 represents the catalytic base and that Tyr484, contrary to what has been suggested, may not be directly involved in the reaction mechanism but, for example, is structure-determining for the enzyme.

Although the first assumption could not be further verified during the time frame of this thesis, mutagenesis studies of THCAS were performed in order to investigate the second assumption.

6.2.3.3. Single site-directed mutagenesis of THCAS

For generation of different THCAS mutants, the already domesticated THCAS (see 6.2.1) served as a template for site-directed mutagenesis (see 6.1.4.1) utilizing the appropriate primer pairs (see 6.1.3; oligonucleotides 27−32). After gel-electrophoretic analysis (see 3.2.7) of the PCR products and KLD-treatment (see 6.1.4.2) the cloned vector was amplified in *E. coli* TOP10 and verified by sequencing (see 3.2.8). Subsequently, TUs for expression of mutated *thcas in planta* were assembled by fusing the CDS with the 35S promoter, the ER targeting peptide and the nopaline synthase terminator containing a 8×his-tag (8×his:TNos) into α-level plasmids using the GB cloning system (see 3.2.3). The assembled GB constructs were then transformed into chemically competent *A. tumefaciens* EHA105 cells (see 3.2.9) and positive clones were confirmed by colony PCR (see 3.2.10) using the relevant primer pairs (see 6.1.3; oligonucleotides 23-26). The generated expression constructs are depicted in Figure 6.2.15 A. Following cloning, *A. tumefaciens* EHA105 harboring the desired constructs were co-infiltrated with *A. tumefaciens* EHA105 cells carrying P19 into *N. benthamiana* plants. Plants co-infiltrated with Agrobacteria harboring er:THCAS:6×his and Agrobacteria containing P19 served as a positive control, while plants solely expressing *p19* served as a negative control. Five days post inoculation, TSP were extracted (see 3.4.1) and subjected to activity assays containing 0.05 mM of CBGA and 1.8 (v/v) ACN (see 6.1.5.2), followed by monitoring the conversion of CBGA to THCA and CBCA *via* HPLC−MS analysis (see 3.5.1). The *in vitro* experiments of the different mutants unveiled that THCAS(Y354F) and THCAS(Y175F) had no significant impact on the production of THCA compared to the wild type enzyme. In contrast, the mutation of THCAS(Y484F) led to an almost complete loss of activity (see Figure 6.2.15 B and C). However, it is important to emphasize that not purified proteins were used in the assays, but plant extracts containing the desired protein. Therefore, no defined amount of the actual protein was applied, but only a defined amount of plant material. Accordingly, parameters such as the influence of a mutation on the accumulation of the protein could not be considered in the analysis, which means that without additional evidence, such as a Western blot analysis, differences in activity compared to the WT enzyme cannot ultimately be attributed to the inactivity of the proteins. Therefore, to exclude, that the loss of THCAS(Y484F) activity was due to the absence of the protein in the extract, additional Western blot analysis was performed with THCAS(Y484F) and the WT enzyme (see 3.4.6.3). But as depicted in Figure 6.2.15 D, the mutein was detected, indicating that the loss of activity was no result of degraded proteins (Figure 6.2.15 D). Therefore, the abolishment of catalytic activity of THCAS(Y484F) confirmed the results of Shoyama *et al*. that Tyr484 is the amino acid most likely involved in the catalytic activity [67].

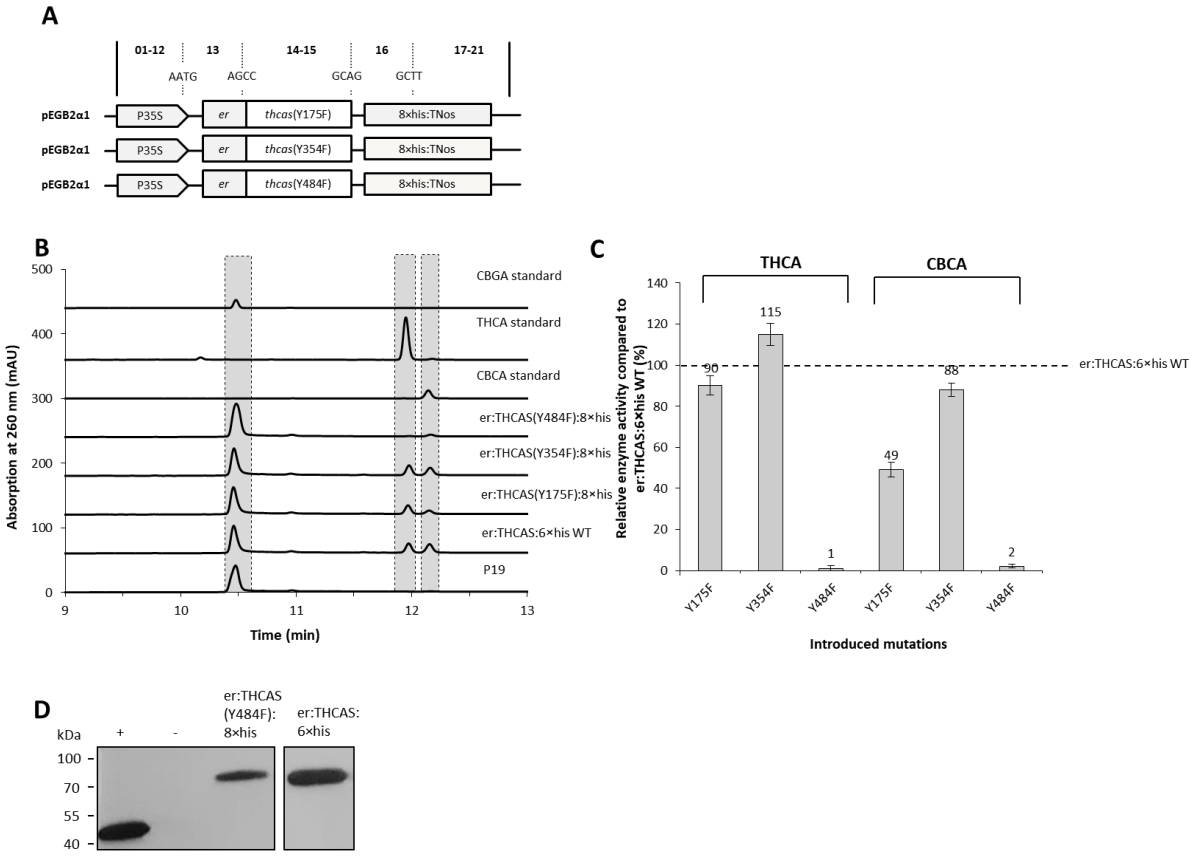


Figure 6.2.15 (A) Expression cassettes harboring different THCAS mutants assembled by utilizing the GoldenBraid cloning system. The capital letters show the four-nucleotide overhangs ensuring correct final orientation within the transcriptional unit (TU), while the numbers above the scheme represent standard GoldenBraid classes within the TU structure [93]. Boxes are not drawn to scale. **(B)** The different mutants were produced in transiently transformed *N. benthamiana* and total soluble protein extracts were used for enzyme activity assays. In case of the er:THCAS(Y484F):8×his mutant, no activity towards THCA and CBCA production was observed. **(C)** Relative enzyme activity of mutated THCAS compared to er:THCAS:6×his WT. **(D)** Western blot analysis of heterologous produced er:THCAS(Y484F):8×his and er:THCAS:6×his. +, recombinant esterase (OeEst228; [121]), featured with a C–terminal 6×his tag fusion. -, TSP extracts of plants producing solely P19; er, ER/apoplast signal peptide.

6.2.4. Mutagenesis studies on CBCAS

Since THCAS has been discussed in detail, attention should once again be paid to the CBCAS. As already mentioned before, THCAS and CBCAS share a high homology in their amino acid sequence (Figure 6.2.4; 93 %). Although the two enzymes do not even differ with regard to the catalytic amino acids described for THCAS, CBCAS, as shown in previous experiments, produces exclusively CBCA and thus does not share the properties of THCAS, which forms both THCA and CBCA (Figure 6.2.3 and Figure 6.2.5 B and C). Therefore, it stands to reason that small differences in amino acid sequence outside the catalytic center could affect the cyclization specificities of the aforementioned proteins. That amino acids do not have to be directly involved in the catalytic mechanism to affect product formation was already described, e.g. for tetrahydroalstonine synthase 1 (THAS1). The introduction of mutation E59A resulted here in a slight change of the product ratio from tetrahydroalstonine to mayumbine [187]. Consequently, to better understand the cyclization reaction forming cannabinoids and to determine the amino acids responsible for the regulation of product specificity with regard to THCA production, single site-directed mutagenesis of CBCAS was considered. Overall, the amino acid sequence of THCAS and CBCAS differs in 37 of the 545 amino acids (Figure 6.3.16). To slightly narrow down the mutations to be performed, amino acids were selected on the following basis: As already shown in previous experiments, CBDAS was only able to produce CBDA and CBCA, but no THCA formation was detected (Figure 6.2.5 A). Therefore, amino acids of CBDAS that are identical to amino acids of THCAS but different from those of CBCAS were initially assumed to play no role in the regulation of THCA production. Consequently, 15 amino acids were selected that were either different for all three proteins or identical to those of CBDAS but different to the amino acids of THCAS. Moreover, Tyr484 was selected for site-directed mutagenesis in order to confirm that also this amino acid is responsible for the enzymes activity (Figure 6.2.16).

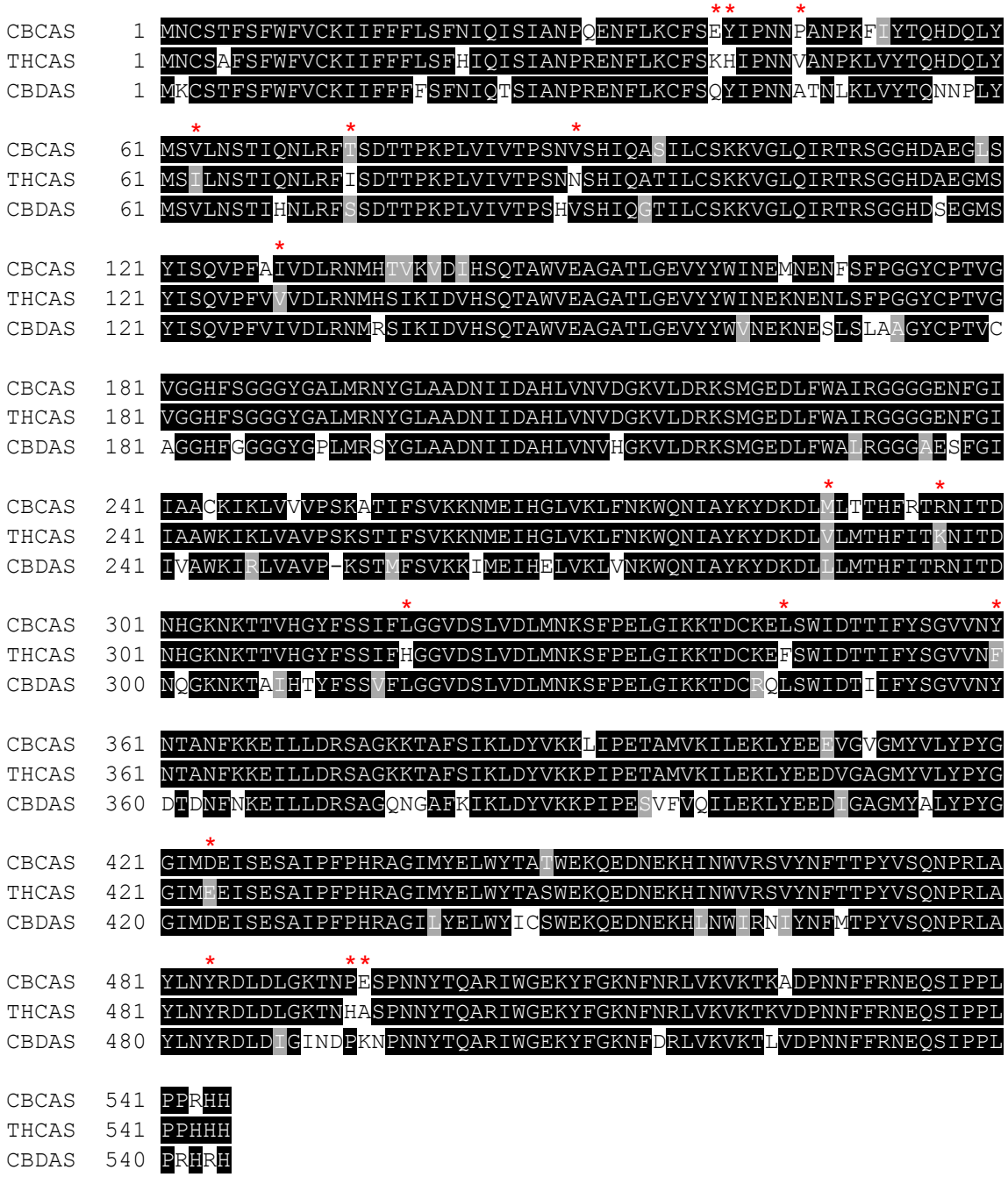


Figure 6.2.16 Multiple sequence alignment of CBCAS (sequence as published in WO 2015/196275 A1), THCAS (GenBank accession no: AB057805) and CBDAS (GenBank accession no: AB292682) using BoxShade software. Amino acids identical are represented in white letters on black background, while amino acids shaded in grey have similar characteristics/properties to those in the same position. Amino acids considered for site-directed mutagenesis are marked with red asterisks.

6.2.4.1. Single site-directed mutagenesis of CBCAS

The different CBCAS mutants were generated by utilizing the already domesticated CBCAS (see 6.2.1) as a template. Site-directed mutagenesis (see 6.1.4.1) was performed with appropriate primer pairs (see 6.1.3; oligonucleotides 29, 30 and 33−62), followed by KLD-treatment (see 6.1.4.2). Finally, the cloned vector was amplified in *E. coli* TOP10 and verified by sequencing (see 3.2.8). Subsequently, each mutated *cbcas* was fused with the 35S promoter (P35S), the ER/apoplast signal peptide (er) and the nopaline synthase gene terminator containing a 8×his-tag (8×his:TNos) and ligated into an α-level plasmid using the GB cloning system (see 3.2.3). The assembled GB constructs were transformed into chemically competent *A. tumefaciens* EHA105 cells (see 3.2.9) and positive clones were confirmed by colony PCR (see 3.2.10) using the relevant primer pairs (see 6.1.3; oligonucleotides 23-26). The generated expression constructs are depicted in Figure 6.2.16 A. *N. benthamiana* plants were co-infiltrated with Agrobacteria harboring the corresponding mutein genes as well as with Agrobacteria containing P19. As a control, wild-type er:CBCAS:8×his and plants solely producing P19 were used. After five days of incubation, *N. benthamiana* plants were harvested and total soluble proteins were extracted (see 3.4.1). The obtained extract was then subjected to activity assays containing 1.8 % (v/v) ACN and 0.05 mM of CBGA (see 6.1.5.2). HPLC−MS analysis (see 3.5.1) revealed that none of the mutated proteins was able to produce the desired THCA (Figure 6.2.16 B). As expected, substitution of Tyr484 resulted in a completely abolished activity towards CBCA as already seen in case of THCAS. However, some of the other single-mutations resulted in a lower activity towards formation of CBCA, most strikingly observed for the muteins T74I, V90N, M288V, R296K and Y360F (Figure 6.2.16 C). To investigate whether the decrease in activity could have been due to the reduced amount of protein in the extract caused by the mutation, it was also reasonable to evaluate protein production by Western blot analysis (see 3.4.6.3). But in the case of CBCAS, it was not even possible to detect the wild-type protein (Figure 6.2.16 D). Consequently, it was not feasible to draw concrete conclusions if the effect of these single amino acid mutations on the reduced turnover of CBCA is caused by low expression rates or reduced catalytic activity as a result of conformational changes in the proteins structure. Nevertheless, the focus was still on the production of THCA.

In studies with monoterpene cyclases/synthases (mTC/S), the introduction of more than one mutation was needed to change the product specificity of limonene synthase (LimS). Substitution of three amino acids in close proximity to each other (S454G, C457V, M458I) resulted in the increased production of bi-cyclic products like β-pinene compared to its natural formed product limonene [188]. Since the single-mutagenesis did not lead to the preferred THCA formation, it was therefore considered to generate double mutants in the next step.

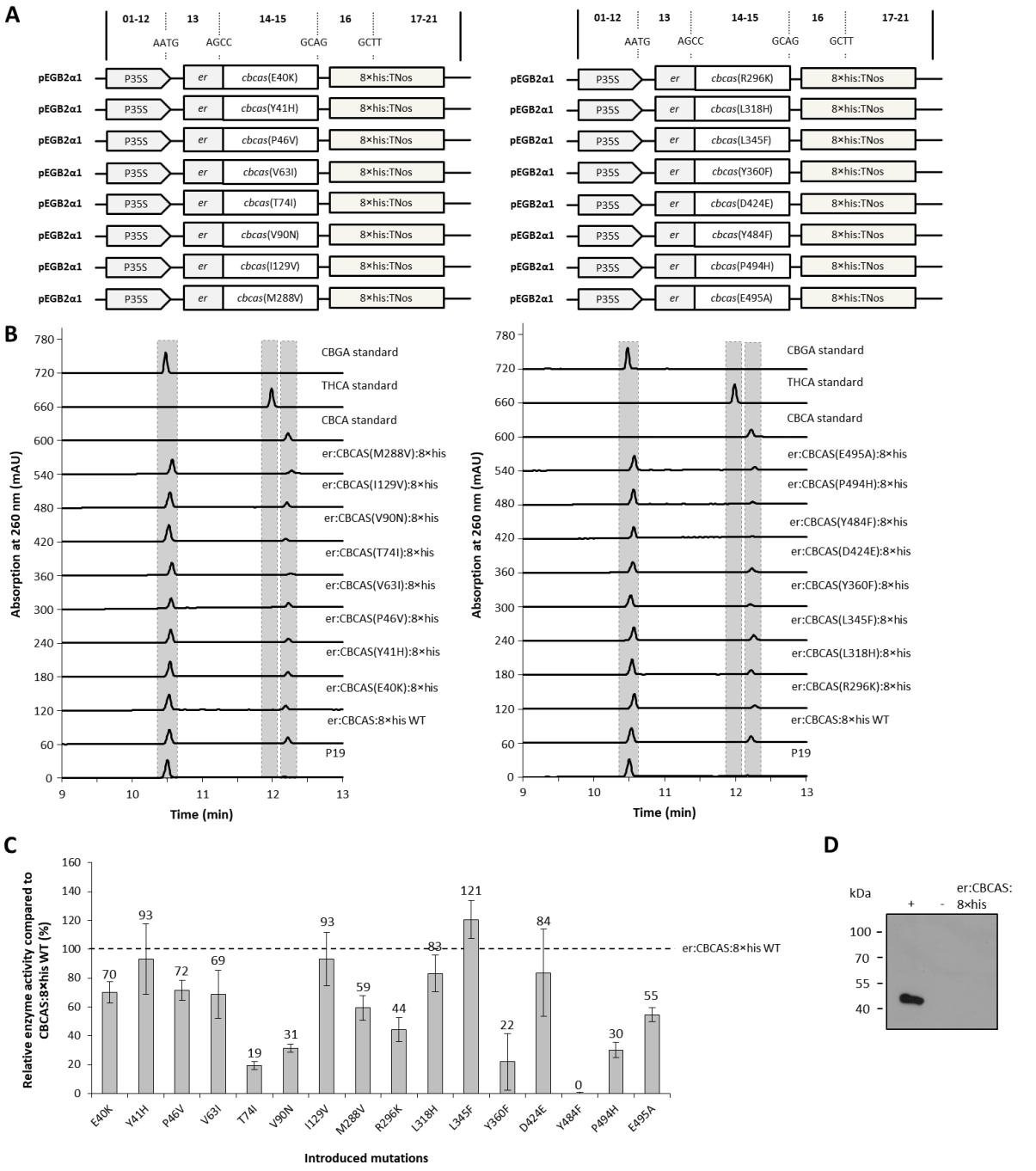
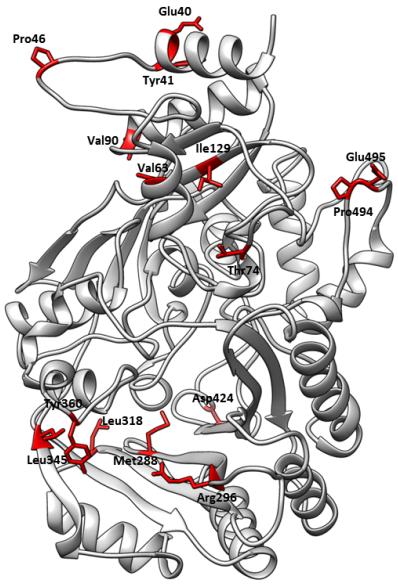


Figure 6.2.16 (A) Cloning of expression cassettes harboring different CBCAS single mutants utilizing the GoldenBraid cloning system. The capital letters show the four-nucleotide overhangs ensuring correct final orientation within the transcriptional unit (TU), while the numbers above the scheme represent standard GoldenBraid classes within the TU structure [93]. Boxes are not drawn to scale. **(B)** HPLC–MS analysis of activity assays from total soluble proteins extracted from transiently transformed *N. benthamiana* plants. None of the single mutations let to the production of THCA. **(C)** Relative enzyme activity of mutated CBCAS compared to er:CBCAS:8×his wild-type (WT). The production of CBCA was most strikingly reduced for the mutants Y484F, Y360F, T74I, V90N, R296K and M288V. **(D)** Western blot analysis of heterologous produced er:CBCAS:8×his. +, recombinant esterase (OeEst228; [121]), featured with a C–terminal 6×his tag fusion. -, TSP extracts of plants producing solely P19; er, ER/apoplast signal peptide.

6.2.4.2. Introduction of double mutations into CBCAS

For selection of double amino acid substitutions, modeling of the three dimensional structure of CBCAS was performed using the fully automated protein homology modeling server SWISS-MODEL and the already deposited structure of THCAS (pdb entry 3vte.1.A; see 6.1.6.2, followed by visualization of the protein by means of the UCSF Chimera software (Figure 6.2.17 A) [163]. Most of the physical interactions between amino acids are caused by direct contact between their residues or the hydration shells. In particular, potential of van der Waals interactions is strongest in the range 2.5–5.0 Å and decreases rapidly with a spatial distance longer than 5.0−7.0 Å, while the limit for the formation of hydrogen bonds is even only about 4.2 Å. Only coulomb interactions have a longer range of approximately 10–15 Å, which implies that everything beyond this distance could be ignored. Consequently, 15 Å was set as the maximum spatial distance criterion for the selection of double mutants to be generated [189–191]. This resulted in ten potential candidates for double amino acid substitutions with a spatial distribution ranging from 6.2–14.3 Å. Unfortunately, it was possible to generate only four of the ten double mutants in the period of this work (Figure 6.2.17 B). The mutations were introduced by site-directed mutagenesis (see 6.1.4.1) using pUDs containing CBCAS single mutants as a template and the appropriate primer pairs (see 6.1.3; oligonucleotides 34, 47, 48, 53, 54, 63 and 64), followed by assembling into an α-level plasmid and transformation into Agrobacteria as already described previously for the generated single mutants (see 6.2.4.1). The utilized expression constructs are depicted in Figure 6.2.17 C. After transformation of *N. benthamiana* plants (see 3.3.1), total soluble proteins were extracted and subjected to activity assays (see 6.1.5.2). But also here, HPLC−MS analysis (see 3.5.1) revealed that no THCA was produced. Only activity towards CBCA formation was detected, which, as in the case of the single mutants, was also lower compared to the wild-type enzyme (Figure 6.2.17 D and E).

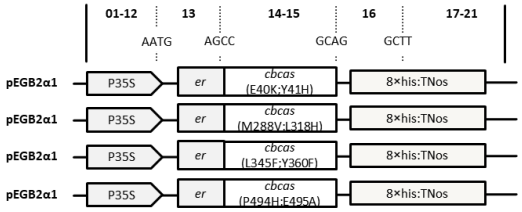
A



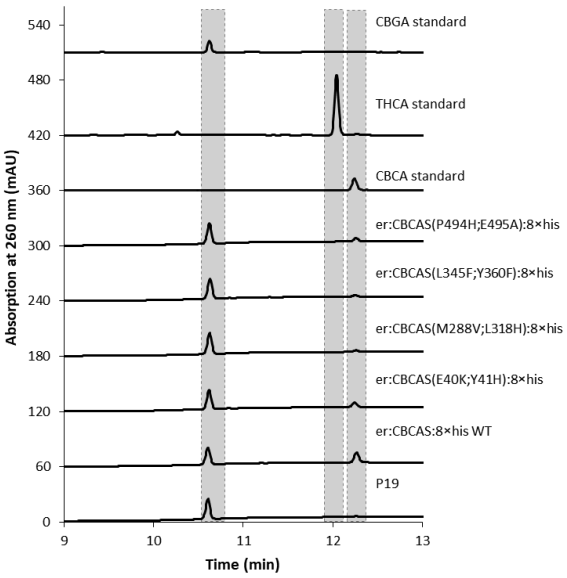
B

Amino acids	Distance in Å
Glu40 and Tyr41	**6.2**
Val63 and Ile129	6.2
Pro494 and Glu495	**6.8**
Met288 and Leu318	**7.9**
Arg296 and Tyr360	9.6
Met288 and Asp424	11.6
Leu345 and Tyr360	**11.9**
Thr74 and Ile129	11.9
Leu318 and Leu345	12.4
Val63 and Thr74	14.3

C



D



E

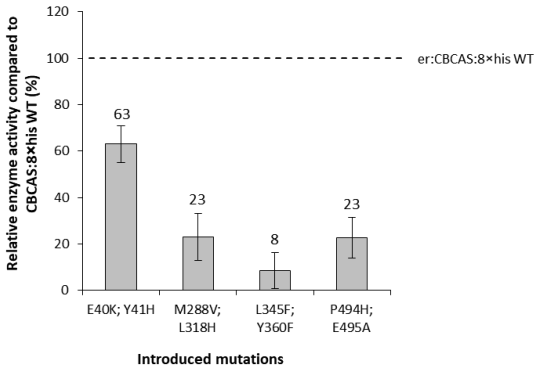


Figure 6.2.17 (A) Predicted 3D structure of CBCAS. Amino acids that are candidates for double mutations are highlighted in red. CBCAS image was generated using UCSF chimera and pdb entry 3vte.1.A [162]. **(B)** Spatial distance between different amino acids. Amino acids that were mutated during this work are written in bold. **(C)** Cloned expression cassettes harboring different CBCAS double mutants. Boxes are not drawn to scale. **(D)** HPLC–MS analysis of activity assays from total soluble proteins extracted from transiently transformed *N. benthamiana* plants. No production of THCA could be observed. **(E)** Relative enzyme activity of mutated CBCAS compared to er:CBCAS:8×his WT. All mutants exhibited a lower enzyme activity compared to the wild-type enzyme, most strikingly for the mutants L345F;Y360F and P494H;E495A.

6.2.4.3. Generation of CBCAS/THCAS chimeras

In a final attempt to narrow down the amino acids that might be responsible for the cyclization specificity, two chimeras, each containing half of the mutations selected at the beginning, were ordered from GenScript (Figure 6.2.18 A and B). The already GB-modified synthetic genes, harboring the appropriate fusion sites (AGCC and GCAG) and BsmBI recognition sites, were introduced into the GB cloning system (see 3.2.3) and transformed into chemically competent *E. coli* TOP10 cells (see 3.2.4), followed by isolation of the amplified plasmids and verification by sequencing (see 3.2.8; oligonucleotides 1-3). Afterwards, each of the domesticated mutants was ligated to the cauliflower mosaic virus 35S promoter (P35S), the ER/apoplast signal peptide sequence (er) as well as the 8×his:TNos into an α-level plasmid. Finally, the plasmids containing the expression constructs were transformed in *A. tumefaciens* EHA105 (see 3.2.9) and used for transient expression in *N. benthamiana* plants (see 3.3.1). Unfortunately, here too, no THCA production and an almost complete abolished activity towards CBCA formation was observed after assaying the produced chimeras (Figure 6.2.18 C and D).

Recently it was shown that substitution of A414V in CBDAS resulted in 3-fold higher catalytic activity for the production of CBDA and a 19-fold increase in THCA formation, suggesting that this amino acid regulates both, the enzymes activity and specificity [172]. However, although CBCAS exhibits valine at the same position, this amino acid does not appear to be crucial for the specificity in CBCAS (Figure 6.3.16). Therefore, it is recommended that amino acids which were initially excluded, in order to limit the amount of mutants to be generated, should also be included in future mutagenesis studies. However, what can be said with certainty is that there does not seem to be the "one" amino acid that is responsible for the promiscuity of the enzymes, but it rather depends on their conformation. Nevertheless, to gain additional insight into the mechanistic differences between CBCAS and THCAS, it should be considered to crystallize both, CBCAS and THCAS, together with their substrate CBGA for x-ray analysis and subsequent comparison of both structures and their catalytic centers.

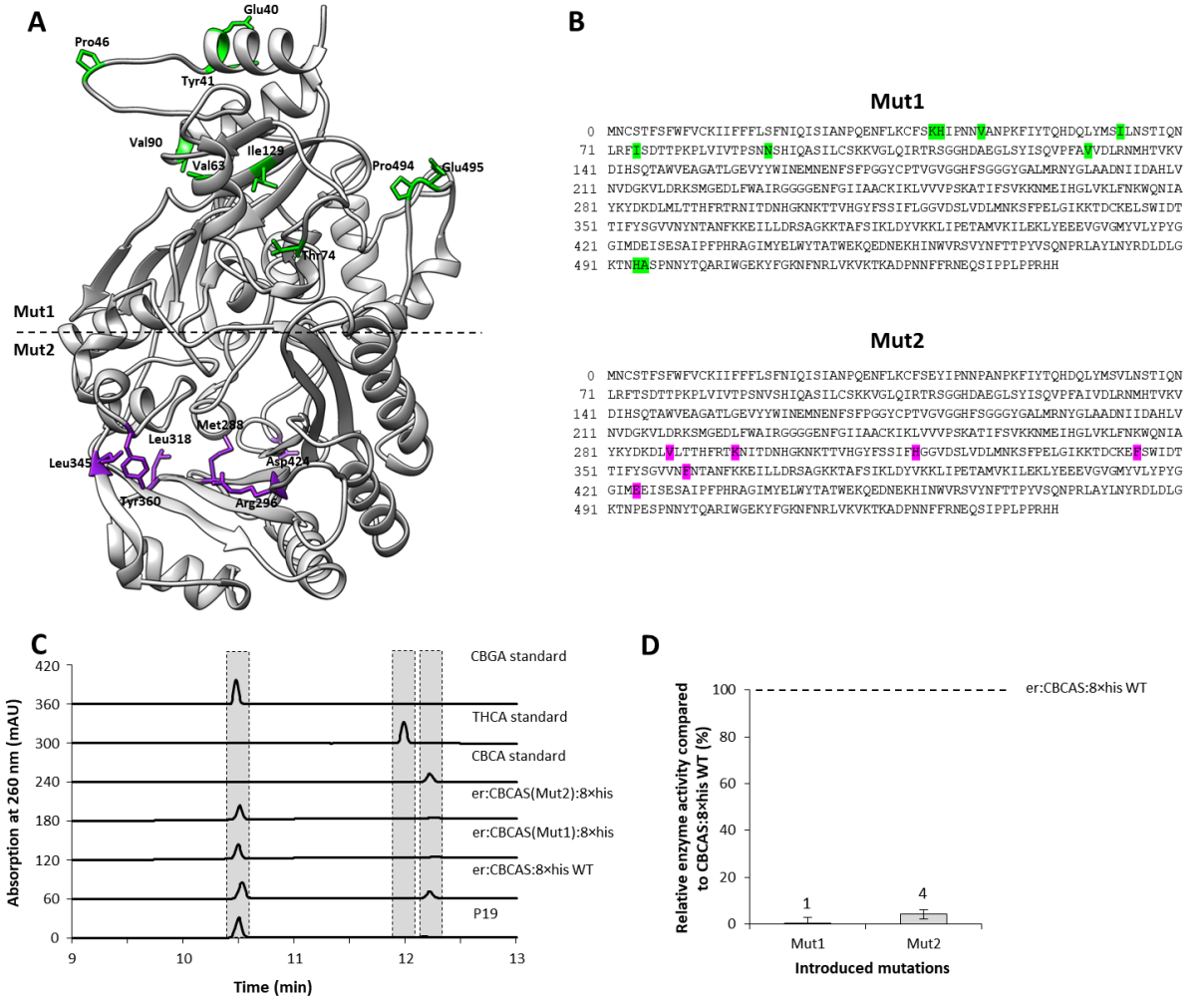


Figure 6.2.18 (A) Predicted 3D structure of CBCAS. Amino acids that are selected for mutation in order to generate CBCAS(Mut1) and CBCAS(Mut2) are highlighted in green and purple, respectively. CBCAS image was generated using UCSF chimera and pdb entry 3vte.1.A [162]. **(B)** Amino acid sequence of CBCAS(Mut1) and CBCAS(Mut2). Mutations of CBCAS(Mut1) and CBCAS(Mut2) are colored in green and purple, respectively. **(C)** HPLC−MS analysis of activity assays from total soluble proteins extracted from transiently transformed *N. benthamiana* plants. No production of THCA could be observed. **(D)** Relative enzyme activity of mutated CBCAS compared to er:CBCAS:8×his WT. The activity of both mutated enzymes was decreased significantly compared to the wild-type enzyme.

6.3. Supporting information

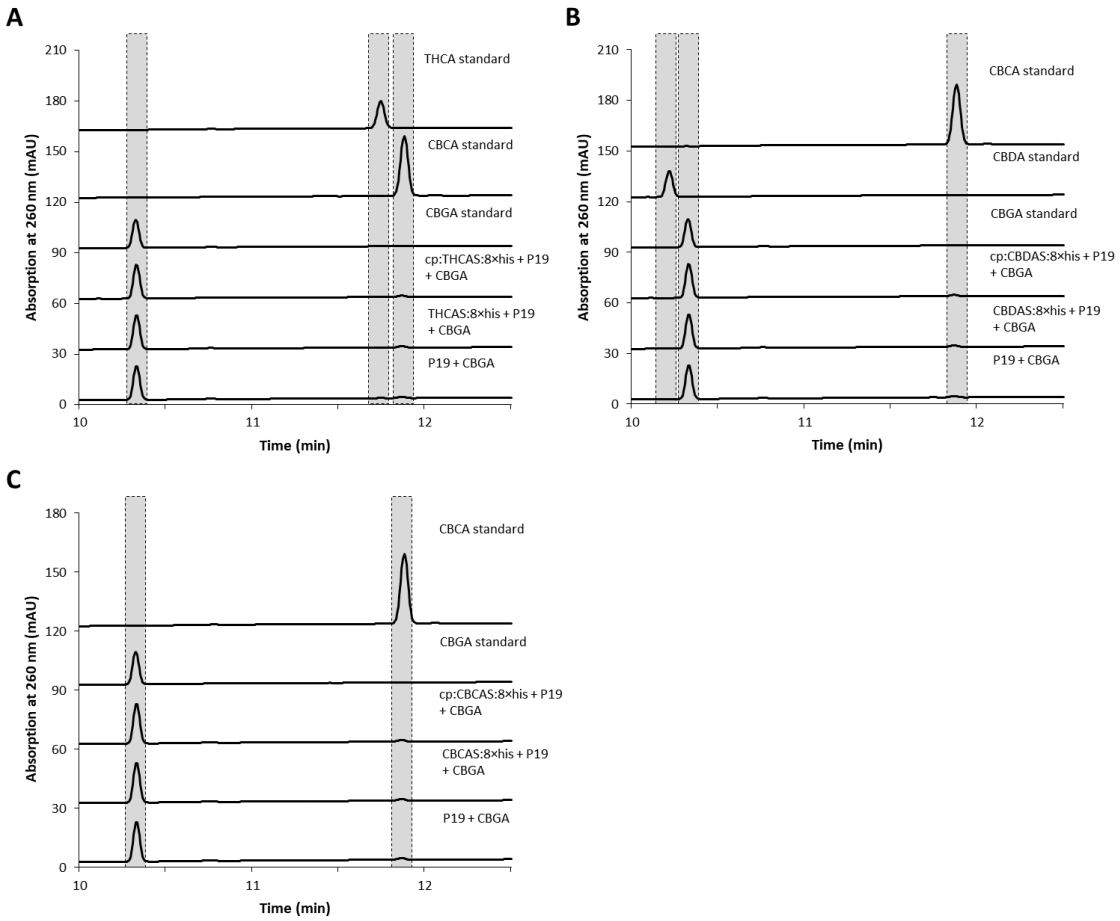


Figure S6.3.1 HPLC–MS analysis of *N. benthamiana* plants producing cytosolic and chloroplast localized THCAS:6×his, CBDAS:8×his or CBCAS:8×his. Plants solely producing P19 served as a negative control.

Chapter 7

Generation of stable transformed tobacco plants for the production of cannabinoids

7.1. Materials and methods

7.1.1. Plasmids and genetic material

The resistance marker gene *nptII* encoding for a neomycin phosphotransferase II (GB0226) was provided by the lab of Dr. Diego Orzáez. The basic GB parts and the GB destination vectors, as well as signal peptide sequences and olivetolic acid pathway genes were provided as described in 4.1.2.

7.1.2. Oligonucleotides

All oligonucleotides used in this study were synthesized by Eurofins Genomics GmbH (Ebersberg, Germany). 4-nt overhangs generated by annealing of oligonucleotides for incorporation into the GoldenBraid system are highlighted in bold. BsmBI recognition sites are underlined.

No.	Name	Nucleotide sequence 5'–3'
65	pOAC cPCR101	GCA GTG AAG CAT TTG ATT GTA
66	pOAC cPCR201	CGT GGT GTG TAG TCA AAA ATG
67	pOLS cPCR101	TGA GTT TCC CGA CTA CTA CT
68	pOLS cPCR201	CCC ACT AGT AAT TCG AGG TC
69	pAAE1 cPCR101	CTG GTG AAG CAT CTA ATG TAG
70	pAAE1 cPCR201	CAA AGG TGG CAC TCC AAT AG
71	pNptII H101	ATA TCG TCT CAC TCG **AGC C**AT TGA ACA AGA TGG ATT GCA
72	pNptII N201	GCG CCG TCT CAC TCG **AAG C**TC AGA AGA ACT CGT CAA GAA
73	pGFP H101	ATA TCG TCT CAC TCG **AGC C**GT GAG CAA GGG CGA GGA GC
74	pGFP M201	GCG CCG TCT CAC TCG **CTG C**CT AGT ACA GCT CGT CCA TGC
75	pGFP cPCR101	CGT AAA CGG CCA CAA GTT CAG C
76	pGFP cPCR201	TAC AGC TCG TCC ATG CCG TGA GT

7.1.3. Generation of transgenic plants

7.1.3.1. Sterilization of tobacco seeds

Nicotiana tabacum L. cv. Petit Havana wild-type plants were used for establishing of stable transformed tobacco lines. In this regard, their seeds were sterilized in 1 mL of 6 % (v/v) sodium hypochlorite for 20 min. Afterwards, the seeds were washed five times in 1 mL of sterile water and placed on MS medium. Finally, the plants were cultivated under sterile conditions at 26 °C and under 12 h of illumination.

7.1.3.2. Stable transformation of *N. tabacum* L. cv. Petit Havana plants

For stable transformation of tobacco plants, *A. tumefaciens* EHA105 cells, harboring the desired genetic constructs, were inoculated into 15 mL of liquid LB medium supplemented with antibiotics and incubated overnight at 28 °C. Afterwards, the culture was centrifuged for 4 min at 4,000 ×*g* and the pellet was resuspended in 25 mL of liquid LB medium without antibiotics. Three leaves of three to four weeks old *N. tabacum* L. cv. Petit Havana plants, cultivated under sterile conditions, were cut into 4×4 mm pieces and incubated in the bacterial culture for 10 min under continuously mixing. Subsequently, the leaf pieces were transferred onto solid RMOP medium (4.3 g/L Murashige-Skoog salts, 30 g/L saccharose, 0.1 g/L myo-inositol, 1 mg/L 6-benzylaminopurine, 0.1 mg/L 1-naphthalene acetic acid, 4.3 g/L phyto-agar, pH 5.8) without antibiotics and incubated for 24 h in the dark. For the selection of successfully transformed tobacco cells, the leaf parts were transferred onto solid RMOP medium supplemented with kanamycin (100 µg/mL) as well as cefotaxime (300 µg/mL), to get rid of the agrobacteria. The leaf parts were then further cultivated at 26 °C under 12 h of illumination. Following the selection process of callus tissue, emerging shoots were transferred onto solid Murashige and Skoog (MS; [192]) medium containing the appropriate antibiotics to facilitate the recovery of transgenic tobacco lines. After root development, plants were cultivated in potting soil in the greenhouse.

In case of establishing liquid cell cultures, callus tissue was inoculated into 120 mL of liquid Linsmaier and Skoog (LS; [193]) medium and supplemented with kanamycin (100 µg/mL) and cefotaxime (300 µg/mL).

7.1.3.3. Isolation of genomic DNA from plant tissue

To verify the successful generation of transgenic plants, genomic DNA was extracted and screened for the integration of transgenes utilizing PCR. Firstly, 750 µL of extraction buffer (1 % sarcosyl, 0.8 M NaCl, 22 mM EDTA, 0.22 M Tris-HCl, 0.8 % CTAB, 0.14 M mannitol; supplementation with 1.4 µL/mL of β-mercaptoethanol prior usage) and 750 µL of chloroform were mixed with 100 mg of ground and frozen plant material, followed by an incubation step of 45 min at 65 °C. Subsequently, the samples were centrifuged at 7,500 ×*g* for 10 min at 4 °C and the upper aqueous phase was transferred into a new tube and mixed with an equal volume of isopropanol. The DNA was precipitated by centrifugation at maximum speed for 10 min at 4 °C and the obtained pellet was washed twice in 70 % ice-cold EtOH. Finally, the genomic DNA was resuspended in 50 µL of H_2O and stored at 4 °C until further analysis.

7.1.3.4. Feeding of transgenic *N. tabacum* plants with hexanoic acid

Four week old transgenic *N. tabacum* plants were infiltrated with 4 mM of hexanoic acid (solved in infiltration buffer (10 mM MES, 10 mM $MgSO_4$). The plants were harvested after 24 h of incubation and metabolites were extracted as described in 3.3.3.

7.1.3.5. Feeding of liquid cell and preparation for HPLC–MS analysis

To test the possible production of olivetolic acid, liquid cell cultures (120 mL batches) putatively expressing the pathway genes (*aae1*, *ols* and *oac*) were fed with 20 mM of hexanoic acid or 0.2 mM of olivetolic acid. As a negative control, callus cultures expressing *gfp* were or *N. tabacum* wild-type were used. After 24 h of incubation, cells were harvested and prepared for HPLC–MS analysis. Therefore, the media of liquid cell cultures were removed and the remaining cells were lyophilized at -80 °C in a vacuum of 0.25 mbar for 24-48 h using the ALPHA 1-4 LDplus freeze-dryer (MartinChrist, Osterode am Harz, Germany). Afterwards, 20 mg of freeze dried cells were homogenized by sonication in 200 µL of 80 % (v/v) MeOH for 30 min at RT. Finally, the extracts were purified two times from solid particles by centrifugation at 17,000 ×*g* for 10 min and 4 °C and subjected to HPLC–MS analysis (see 3.5.1).

7.2. Results and discussion

7.2.1. Reconstitution of olivetolic acid (OA) biosynthesis in stable transformed *Nicotiana tabacum*

In parallel with the transient transformation experiments in *N. benthamiana*, stably transformed *N. tabacum* lines producing cannabinoids were of major interest. Therefore, genes required for the production of OA had to be integrated into the genome of *N. tabacum* plants. For the selection process of transgenic *N. tabacum* plants, the resistance marker *nptII* encoding neomycin phosphotransferase II was selected [194]. For domestication of NptII and GFP, the pEGB vector harboring the NptII coding sequence (CDS) as well as the pUPD containing the GFP CDS was used as a template to incorporate appropriate fusion overhangs by PCR (see 3.2.1 and 7.1.2; oligonucleotides 71−74). Subsequently, the PCR products were purified (see 3.2.2) and ligated into the universal domesticator *via* a GB reaction (see 3.2.3). Afterwards, each CDS was fused to the P35S ATG and the nopaline synthase terminator (TNos) and assembled into an α-level plasmid. For the biosynthesis of OA by stable transformation, already assembled TUs of the pathway genes, aae1:yfp, gfp:ols:8×his and oac:8×his (see 4.2.1) were assembled together with the NptII TU into an α-level plasmid (SOA; Figure 7.2.1 A). As a control, the cloned GFP TU was fused to the NptII TU and ligated into the pDGB3Ω1-level plasmid (SGFP; Figure 7.2.1 A).

A detailed description of all chronological steps performed during assembly of GB constructs and transformation of *A. tumefaciens* used for transformation of tobacco plants is given in chapter 4.2.1.

After assembly of expression cassettes, stable transgenic plants were generated by transformation of plants, cultivated under sterile conditions (see 7.1.3.1), with *A. tumefaciens* EHA105 cells, harboring the desired genetic constructs (see 7.1.3.2). Four transgenic lines for each construct (SOA and SGFP) were successfully regenerated from callus cultures. To verify integration of the desired genes, genomic DNA was extracted from leaf tissue of cultivated transgenic plants (see 7.1.3.3) and screened for integration of the corresponding DNA sequences by PCR (see 3.2.1) using gene specific primers (see 7.1.2; oligonucleotides 65−70, 75 and 76). In this regard transgenes were detected in all lines harboring the SOA construct as well as in #29 of SGFP line. Since SGFP served as a control, only line #29 was mentioned in all experiments (Figure 7.2.1 B). Surprisingly, only the aae1 and oac transgene was detectable in all of the four SOA lines. However, it is not unusual that low purity of extracted genomic DNA causes a false negative result due to inhibition of polymerase-based amplification. Therefore, it was assumed that all transgenes were integrated, as it is not likely that *ols*, which is flanked by *aae1* and *oac*, was cut from the T-DNA in all lines. In order to investigate if also the desired proteins were produced, total soluble proteins were extracted (see 3.4.1) and loaded onto SDS−PAGE (see 3.4.6), followed by Western blot analysis (see 3.4.6.3). But only in SOA line #25 a

faint band mirroring the size of AAE1:YFP with approximately 95 kDA was detected, when GFP primary antibodies were used. In contrast, the control SGFP#29 showed a clear production of GFP (see Figure 7.2.1 C). However, since protein accumulation in stable transgenic plants is generally lower than in transiently transformed ones and as the data collected in previous transient experiments already resulted in weak protein detection (see Chapter 4.2.2), it was not surprising [106]. Nevertheless, production of OA in transgenic plants was further analyzed.

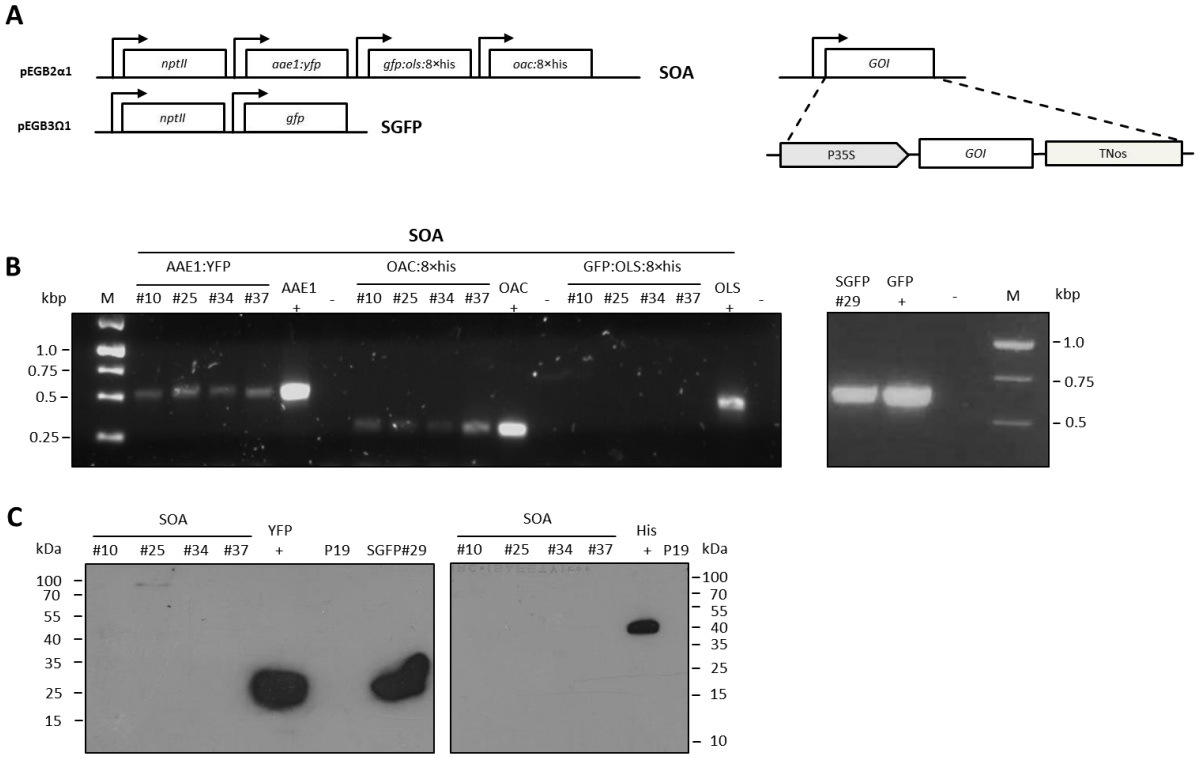


Figure 7.2.1 (A) Schematic representation of the generated multigene GoldenBraid DNA constructs used for stable transformation of *N. tabacum* plants. Each transcriptional unit (TU) consisted of the gene of interest (GOI), the cauliflower mosaic virus (CaMV) 35S promoter (P35S) and the nopaline synthase terminator (TNos). The arrows represent the direction of transcription. Boxes are not drawn to scale. Abbreviations are specified on the right. **(B)** PCR amplification of *aae1*, *oac*, *ols* and *gfp* from genomic DNA isolated from transgenic *N. tabacum* lines transformed either with construct SOA or SGFP. +, 15 ng of plasmid DNA (pUD harboring either *aae1*, *oac*, *ols* or *gfp*); -, PCR reaction mixture containing genomic DNA of wild-type *N. tabacum* plants; M, GeneRuler 1 kb DNA ladder. **(C)** Western blot analysis of total soluble proteins (TSP) extracted from transgenic *N. tabacum* lines. Heterologous produced proteins were either detected using GFP primary antibodies (left blot) or his-probe mouse monoclonal IgG antibodies (right blot). A recombinant esterase (OeEst228; [121]), featured with a C–terminal 6×his tag fusion (His +) as well as recombinant produced YFP (YFP +) served as positive controls. TSP extracts of plants producing solely P19 were used as negative controls (P19).

7.2.2. Analysis of transgenic plants and liquid cell cultures

To verify the biosynthesis of OA, four week old transgenic *N. tabacum* plants were infiltrated with 4 mM of hexanoic acid and incubated for 24 h (see 7.1.3.4). After the extraction of metabolites (see 3.3.3), samples were subjected to HPLC−MS analysis (see 3.5.1). However, when screened for the specific masses of OA glucosides or OA, no newly formed metabolites, which could be assigned to the desired substances, were detected (Figure 7.2.2).

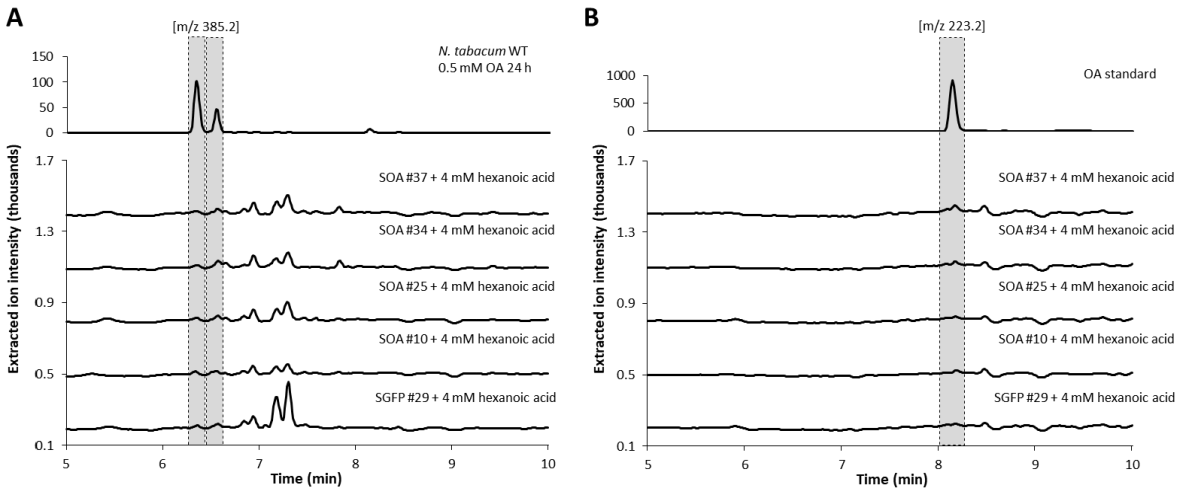


Figure 7.2.2 HPLC−MS analysis of metabolites extracted from transgenic plants harboring the desired genetic constructs. Prior extraction the plants were supplemented with 4 mM of hexanoic acid. The extracts were analyzed in negative selected ion monitoring (SIM) with selected m/z of 385.2 (**A**; detection of olivetolic acid (OA) glucoside) and in negative SIM with selected m/z of 223.2 (**B**; detection of OA). No newly formed metabolites associated with heterologously produced enzymes could be detected in analyses. Due to the lack of authentic standards, wild-type *N. tabacum* plants supplemented with olivetolic acid served as a positive control for OA glucoside measurements. Plants harboring the construct SGFP served as a negative control.

In parallel liquid cell cultures were generated from SOA lines #10, #24, #35 and SGFP line #29. Only SOA line #37 showed growth issues and was therefore excluded from further experiments. To examine if OA was also glucosylated in undifferentiated callus cells, wild-type (WT) *N. tabacum* liquid cell cultures were initially fed with 0.2 mM of OA and incubated for 24 h. Afterwards, the cells were harvested and prepared for HPLC−MS analysis (see 7.1.3.5). The resulting mass chromatograms were then compared to chromatograms obtained from WT *N. tabacum* cells without any supplementation. Here as well, when screened in negative selected ion monitoring with a with selected (SIM) m/z of 223.2 (detection of OA) and in negative SIM with selected m/z of 385.2 (detection of OA glucoside), OA as well as C-2 and C-4 OA glucoside was detected, showing that the undifferentiated cells were converting OA and that the extraction method was suitable to extract the desired metabolites (Figure 7.2.3 A). Thereupon, transgenic cell cultures (SOA and SGFP) were supplemented with 20 mM of hexanoic acid and incubated for 24 h. After isolation of metabolites (see 7.1.3.5), the extracts were

analyzed by HPLC−MS (see 3.5.1). Unfortunetaly, also here, neither OA nor OA glucosides were detected in extracts of transgenic cells. Although there was a peak in samples screened for the specific mass of OA, the substance was also present in control samples that contained the solvent ACN. Therefore, it is more likely that this was a contamination of the column rather than the desired cannabinoid precursor (Figure 7.2.3 B and C).

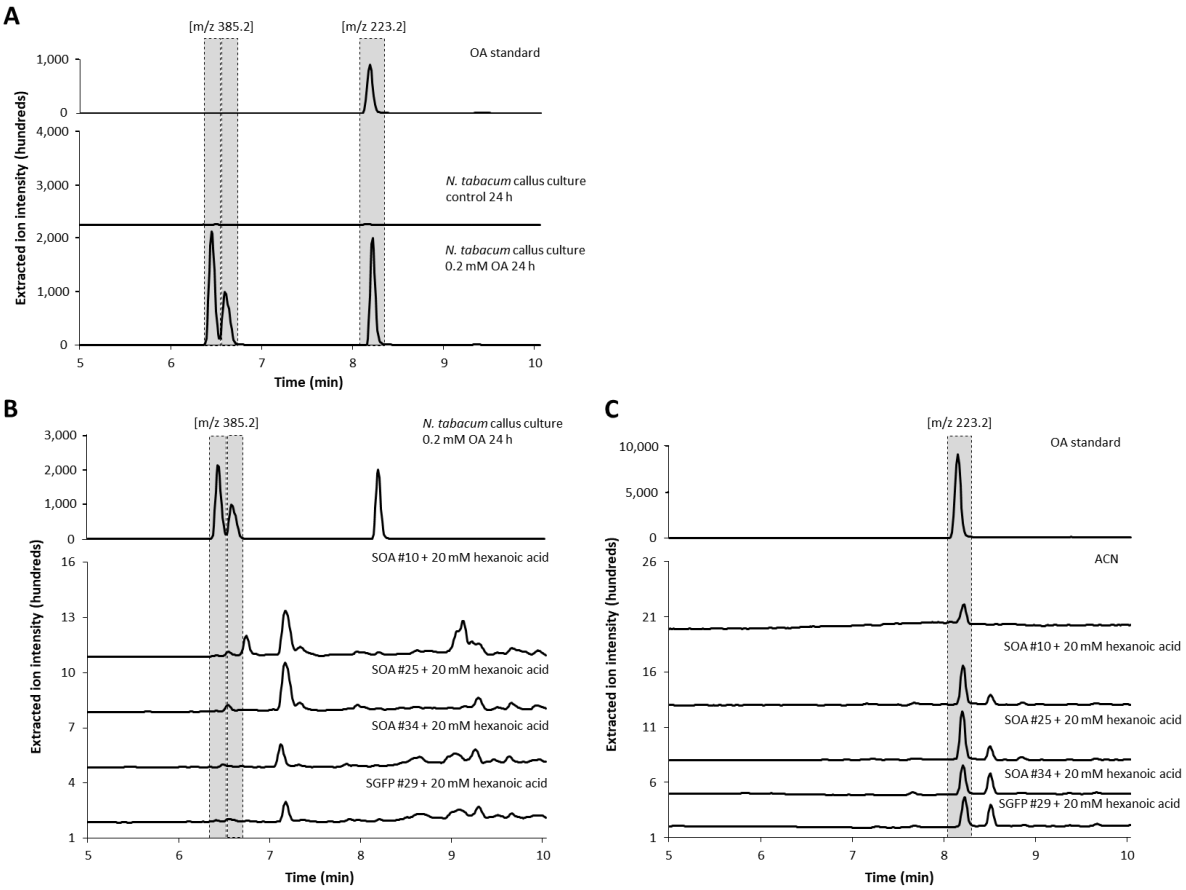


Figure 7.2.3 (A) Feeding of *N. tabacum* callus culture suspension with olivetolic acid (OA). OA was absorbed and part of the substance was glycosylated as already observed for infiltrated *N. benthamiana* plants. *N. tabacum* callus culture suspension without supplementation served as a negative control. **(B** and **C)** HPLC−MS analysis of metabolites extracted from freeze-dried calli harboring the desired genetic constructs. Prior extraction the cells were fed with 20 mM hexanoic acid. The extracts were analyzed in negative selected ion monitoring (SIM) with selected m/z of 385.2 (**B**; detection of OA glucoside) and in negative SIM with selected m/z of 223.2 (**C**; detection of OA). No newly formed metabolites associated with introuced genes could be detected in analyses**.** Since lack of authentic standards, wild-type *N. tabacum* liquid cell culture was supplemented with olivetolic acid served as a positive control for OA glucoside measurements. Cells harboring the construct SGFP served as a negative control.

To narrow down the bottleneck, for example, transgene expression could be further investigated on a transcriptional level by performing Northern blot analysis or quantitative real-time PCR (qPCR). However, these experiments require extensive preparation and especially qPCR is highly error-prone if the method is not well established and validated [195]. In order to rule out all other eventualities

before carrying out these rather time-consuming experiments, the SOA construct used for the transformation of plants was sequenced once again. Surprisingly, the sequencing revealed an insertion of 63 base pairs (bp) in the ols gene which resulted in a frame shift and in a premature stop codon (Figure 7.2.4). Consequently, it was assumed that the protein truncated by 105 amino acids (280 instead of 385) was not functional and therefore the OLS was most likely the bottleneck in this approach.

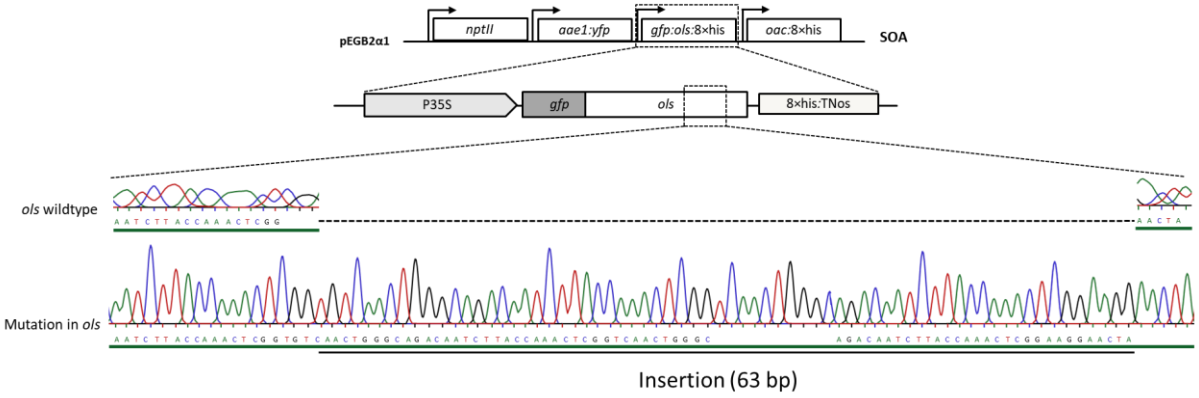


Figure 7.2.4 Sequencing of construct SOA, used for stable transformation of *N. tabacum* plants. The sequencing revealed an insertion of 63 base pairs (bp) in the ols gene which resulted in a frame shift and in a premature stop codon. Boxes are not drawn to scale.

However, when it comes to the generation of stable transgenic plants in general, the path to the desired product is often a very error-prone process. Thus one of the major hurdles in the generation of transgenic lines concerns the expression of transgenes that are particularly susceptible to silencing mediated by small RNAs (sRNAs) compared to endogenous ones [196]. To counteract the gene silencing, different genetic elements, such as promoters and introns, have been demonstrated to have an impact on the stability of transgenes [197–200]. But also the choice of terminators is a key factor that has a major effect on transgene expression levels. The utilization of the *Arabidopsis thaliana* heat shock protein terminator (tHSP) for example in combination with the P35S showed less gene silencing than combination of commonly used P35S and TNos regulatory elements. This observation was assigned to less read-through transcription and therefore less mRNAs lacking poly(A) tails [196]. Besides the choice of regulatory elements, also directions of transgenes should be considered during plasmid design when assembling two or more expression cassettes as the intentional introduction of repetitive sequences into a transgenic locus has in some cases been associated with undesirable negative effects on the expression of transgenes and their stability [201–204]. Consequently, since the construct has to be cloned again harboring the functional ols gene, the general design of the here used multigene expression construct could be reconsidered.

Conclusions

Cannabis research has experienced great resurgence during the last decade, especially in the field of biosynthetic production of cannabinoids in microbial hosts, predominantly in *Saccharomyces cerevisiae* or *Komagataella phaffii*, but also in *Kluyveromyces marxianus* and *Yarrowia lipolytica* [60,62,172,205–207]. However, the utilization of tobacco as production chassis described in the present work depicts a promising basis for the future research regarding cannabinoid production in this heterologous plant host. Initially, it was observed that injection of OA into leaves of tobacco led to the production of the novel-to-nature compounds C-4 and C-2 OA glucosides, which could in turn offer new opportunities regarding glucosylated cannabinoid-based pharmaceuticals [124,127,128].

Furthermore, transient expression of *aae1*, *ols* and *oac* and supplementation of hexanoic acid yielded in the *in vivo* formation of the cannabinoid precursor olivetolic acid (when cytosolic localized proteins were used; Figure 4.2.4 B) and C-4 OA glucoside (with cytosolic and chloroplast localized proteins; Figure 4.2.4 B and C). However, it needs to be clarified whether the downstream proteins such as NphB(Q295L) or THCAS and CBDAS are able to accept the glucosylated substrates and converting them into the valuable glucosylated cannabinoids. At least for the formation of THCA or CBCA, the acceptance of C-4 cannabigerolic acid glucoside as a substrate would be unlikely, since the ring closure happens at the C-4 localized hydroxyl group [63,67]. Consequently, it would be necessary to investigate if maximization of the pathway flux by metabolic channeling trough scaffold proteins, regulating spatial and temporal organization of desired molecules, would curtail the formation of C-2 OA glucoside and deliver more acceptable precursors for cannabinoid biosynthesis [208,209]. Nevertheless, even in the case that the substrates could not be processed, co-expression of β-glucosidase genes like *bgl1* from *Aspergillus niger* could be considered [210].

Although there are still questions pending regarding the glucosylated cannabinoids, the issue that should be definitively addressed in the next step is the increase of GPP supply. The introduction of NphB(Q295L) localized either in the cytosol or the chloroplast and supplemented with sufficient amounts of OA did not result in the formation of CBGA, indicating that GPP is the bottleneck in this approach (Figure 5.2.6 and Figure 5.2.7). In this regard, it would be advantageous to boost the general GPP pool by incorporation of *A. thaliana* isopentenyl phosphate kinase (AtIPK) and different GPP synthases such as the *L. erythrorhizon* GPPS (LeGPPS) or the small subunit of the heterotetrameric GPPS derived from *A. majus* (AmGPPS:SSU) into the transformation approach [143,147,148].

In parallel to the transient approach, it was considered to establish stable transgenic plants, more precisely liquid cell cultures. These were thought to simplify the feeding of precursors and, especially

with regard to the utilization in industrial production, could enable large-scale cultivation in bioreactors under GMP conditions. However, the obtained transgenic lines expressing *aae1*, *ols* and *oac* did not produce olivetolic acid or any of its glucosides (Figure 7.2.2 and Figure 7.2.3). This was most likely due to an insertion of 63 base pairs (bp) in the ols gene which resulted in a frame shift and in a premature stop codon. Consequently, generation of transgenic lines expressing the correct *ols* together with *aae1* and *oac* could make OA production still feasible.

In summary, although it has not yet been possible to synthesize cannabinoids such as CBGA, THCA, or CBDA *in vivo*, at least the intermediate OA has been formed. In addition, NphB(Q295L) and the late biosynthetic enzymes THCAS, CBDAS and CBCAS were successfully produced in the heterologous plant host and their activity was confirmed in *in vitro* assays. However, in order to establish tobacco as a competitive production host in the future, further extensive engineering efforts are required.

Besides reconstruction of the cannabinoid biosynthetic pathway, the late biosynthetic enzymes were studied in more detail. It was shown that THCAS and CBDAS exhibited product promiscuity when co-solvent systems were used, indicating that not only the intracellular pH or the gene expression levels of the different synthases, but also the hydrophobic environment and the various compounds in *Cannabis* glandular trichomes should be considered to be responsible for the cannabinoid diversity in different *Cannabis* strains [172,186]. Nevertheless, further experiments comprising for example with assays containing NADES have to be performed in order to confirm this hypothesis.

Finally, mutagenesis studies with CBCAS were performed with the aim of producing THCA and in order to get more insights into the catalytic mechanisms of the late biosynthetic enzymes. Unfortunately, none of the introduced mutations resulted in the production of the desired cannabinoid. Therefore, further site-directed mutagenesis, which includes the initially omitted amino acids, as well as co-crystallization of CBCAS and THCAS, together with their substrate CBGA for x-ray analysis and subsequent comparison of both structures and their catalytic centers, needs to be performed.

References

1. **Russo EB**. 2007. History of Cannabis and Its Preparations in Saga, Science, and Sobriquet. *Chem Biodivers* **4**(8): 1614–48.
2. **Hillig KW**. 2004. A chemotaxonomic analysis of terpenoid variation in *Cannabis*. *Biochem Syst Ecol* **32**(10): 875–91.
3. **Chandra S, ElSohly MA, Lata H**. 2017. *Cannabis sativa L*: Botany and biotechnology. Springer International Publishing.
4. **Clarke RC, Merlin M**. 2016. *Cannabis*: Evolution and ethnobotany. University of California Press.
5. **Warf B**. 2014. HIGH POINTS: AN HISTORICAL GEOGRAPHY OF CANNABIS. *Geogr Rev* **104**(4): 414–38.
6. **Gloss D**. 2015. An Overview of Products and Bias in Research. *Neurotherapeutics* **12**(4): 731–4.
7. **Clarke RC**. 2007. Traditional *Cannabis* Cultivation in Darchula District, Nepal—Seed, Resin and Textiles. *J Ind Hemp* **12**(2): 19–42.
8. **Russo EB, Jiang H-E, Li X, Sutton A**, *et al*. 2008. Phytochemical and genetic analyses of ancient cannabis from Central Asia. *J Exp Bot* **59**(15): 4171–82.
9. **Abel E**. 1980. Marihuana: The First Twelve Thousand Years. Springer Science and Business Media.
10. **Zuardi AW**. 2006. History of *cannabis* as a medicine: a review. *Rev Bras Psiquiatr* **28**(2): 153–7.
11. **Merlin MD**. 1973, 1972. Man and Marijuana: Some Aspects of Their Relationship. A.S. Barnes.
12. **Touw M**. 1981. The Religious and Medicinal Uses of *Cannabis* in China, India and Tibet. *J Psychoact Drugs* **13**(1): 23–34.
13. **Kapoor R, Huang Y-S**. 2006. Gamma Linolenic Acid: An Antiinflammatory Omega-6 Fatty Acid. *Curr Pharm Biotechnol* **7**(6): 531–4.
14. **Nelson SM**. 2007. THE ARCHAEOLOGY OF KOREA. Cambridge University Press.
15. **Mikuriya TH**. 1969. Marijuana In Medicine: Past, Present and Future. *Calif Med* **110**(1): 34–40.
16. **Lozano I**. 2001. The Therapeutic Use of *Cannabis sativa* (L.) in Arabic Medicine. *J Cannabis Ther* **1**(1): 63–70.
17. **Mahdizadeh S, Khaleghi Ghadiri M, Gorji A**. 2015. Avicenna's Canon of Medicine: a review of analgesics and anti-inflammatory substances. *Avicenna J Phytomed* **5**(3): 182–202.

18. **Russo E, Guy GW**. 2006. A tale of two cannabinoids: The therapeutic rationale for combining tetrahydrocannabinol and cannabidiol. *Med Hypotheses* **66**(2): 234–46.
19. **Mechoulam R, Shvo Y**. 1963. Hashish—I. *Tetrahedron* **19**(12): 2073–8.
20. **Gaoni Y, Mechoulam R**. 1964. Isolation, Structure, and Partial Synthesis of an Active Constituent of Hashish. *J Am Chem Soc* **86**(8): 1646–7.
21. **Vollner L, Bieniek D, Korte F**. 1969. Haschisch XX. Canabidivarin, ein neuer Haschisch-Inhaltsstoff. *Tetrahedron Lett*(3): 145–7.
22. **Devane WA, Dysarz FA, Johnson MR, Melvin LS**, *et al*. 1988. Determination and characterization of a cannabinoid receptor in rat brain. *Mol Pharmacol* **34**(5): 605–13.
23. **Devane WA, Hanus L, Breuer A, Pertwee RG**, *et al*. 1992. Isolation and structure of a brain constituent that binds to the cannabinoid receptor. *Science* **258**(5090): 1946–9.
24. **Mechoulam R, Ben-Shabat S, Hanus L, Ligumsky M**, *et al*. 1995. Identification of an endogenous 2-monoglyceride, present in canine gut, that binds to cannabinoid receptors. *Biochem Pharmacol* **50**(1): 83–90.
25. **Matsuda LA, Lolait SJ, Brownstein MJ, Young AC**, *et al*. 1990. Structure of a cannabinoid receptor and functional expression of the cloned cDNA. *Nature* **346**(6284): 561–4.
26. **Munro S, Thomas KL, Abu-Shaar M**. 1993. Molecular characterization of a peripheral receptor for cannabinoids. *Nature* **365**(6441): 61–5.
27. **Burstein S**. 2015. Cannabidiol (CBD) and its analogs: a review of their effects on inflammation. *Bioorg Med Chem* **23**(7): 1377–85.
28. **Carlini EA**. 2004. The good and the bad effects of (-) *trans*-delta-9-tetrahydrocannabinol (Delta 9-THC) on humans. *Toxicon* **44**(4): 461–7.
29. **Pertwee RG**. 2006. Cannabinoid pharmacology: the first 66 years. *Br J Pharmacol* **147** (1): 163-71.
30. **Badowski ME**. 2017. A review of oral cannabinoids and medical marijuana for the treatment of chemotherapy-induced nausea and vomiting: a focus on pharmacokinetic variability and pharmacodynamics. *Cancer Chemother Pharmacol* **80**(3): 441–9.
31. **Rossi F, Tortora C, Argenziano M, Di Paola A**, *et al*. 2020. Cannabinoid Receptor Type 2: A Possible Target in SARS-CoV-2 (CoV-19) Infection? *Int J Mol Sci* **21**(11): 3809.
32. **Onaivi ES, Sharma V**. 2020. Cannabis for COVID-19: can cannabinoids quell the cytokine storm? *Future Sci OA* **6**(8): FSO625.
33. **Byrareddy SN, Mohan M**. 2020. SARS-CoV2 induced respiratory distress: Can cannabinoids be added to anti-viral therapies to reduce lung inflammation? *Brain Behav Immun* **87**: 120–1.
34. **Śledziński P, Nowak-Terpiłowska A, Zeyland J**. 2020. Cannabinoids in Medicine: Cancer, Immunity, and Microbial Diseases. *Int J Mol Sci* **22**(1).

35. **Piomelli D**. 2013. Endocannabinoids; *Encyclopedia of Biological Chemistry*. Elsevier. p 194–196.

36. **Fisar Z**. 2009. Phytocannabinoids and endocannabinoids. *Curr Drug Abuse Rev* **2**(1): 51–75.

37. **Pacher P, Bátkai S, Kunos G**. 2006. The endocannabinoid system as an emerging target of pharmacotherapy. *Pharmacol Rev* **58**(3): 389–462.

38. **Tasker JG**. 2006. Rapid glucocorticoid actions in the hypothalamus as a mechanism of homeostatic integration. *Obesity* **14**: 259-265.

39. **Marsicano G, Lutz B**. 2006. Neuromodulatory functions of the endocannabinoid system. *J Endocrinol Invest* **29** (3): 27–46.

40. **Walsh KB, Andersen HK**. 2020. Molecular Pharmacology of Synthetic Cannabinoids: Delineating CB1 Receptor-Mediated Cell Signaling. *Int J Mol Sci* **21**(17): 6115.

41. **Hajdu Z, Nicolussi S, Rau M, Lorántfy L**, *et al*. 2014. Identification of endocannabinoid system-modulating *N*-alkylamides from *Heliopsis helianthoides var. scabra* and *Lepidium meyenii*. *J Nat Prod* **77**(7): 1663–9.

42. **Reynoso-Moreno I, Najar-Guerrero I, Escareño N, Flores-Soto ME**, *et al*. 2017. An Endocannabinoid Uptake Inhibitor from Black Pepper Exerts Pronounced Anti-Inflammatory Effects in Mice. *J Agric Food Chem* **65**(43): 9435–42.

43. **Dallazen JL, Maria-Ferreira D, Da Luz BB, Nascimento AM**, *et al*. 2020. Pharmacological potential of alkylamides from *Acmella oleracea* flowers and synthetic isobutylalkyl amide to treat inflammatory pain. *Inflammopharmacology* **28**(1): 175–86.

44. **Hussain T, Espley RV, Gertsch J, Whare T**, *et al*. 2019. Demystifying the liverwort *Radula marginata*, a critical review on its taxonomy, genetics, cannabinoid phytochemistry and pharmacology. *Phytochem Rev* **18**(3): 953–65.

45. **Hussain T, Plunkett B, Ejaz M, Espley RV**, *et al*. 2018. Identification of Putative Precursor Genes for the Biosynthesis of Cannabinoid-Like Compound in *Radula marginata*. *Front Plant Sci* **9**: 537.

46. **Brenneisen R**. 2007. Chemistry and Analysis of Phytocannabinoids and Other Cannabis Constituents. In Elsohly MA ed; *Marijuana and the Cannabinoids*. Humana Press. 17–49.

47. **Elsohly MA, Slade D**. 2005. Chemical constituents of marijuana: the complex mixture of natural cannabinoids. *Life Sci* **78**(5): 539–48.

48. **Radwan MM, Elsohly MA, Slade D, Ahmed SA**, *et al*. 2009. Biologically active cannabinoids from high-potency *Cannabis sativa*. *J Nat Prod* **72**(5): 906–11.

49. **Fischedick JT, Hazekamp A, Erkelens T, Choi YH**, *et al*. 2010. Metabolic fingerprinting of *Cannabis sativa* L., cannabinoids and terpenoids for chemotaxonomic and drug standardization purposes. *Phytochemistry* **71**(17-18): 2058–73.

50. **Degenhardt F, Stehle F, Kayser O**. 2017. The Biosynthesis of Cannabinoids; *Handbook of Cannabis and Related Pathologies*. Elsevier. 13–23.

51. **Marcu JP**. 2016. An Overview of Major and Minor Phytocannabinoids; *Neuropathology of Drug Addictions and Substance Misuse*. Elsevier. 672–678.

52. **Happyana N, Agnolet S, Muntendam R, van Dam A**, *et al*. 2013. Analysis of cannabinoids in laser-microdissected trichomes of medicinal *Cannabis sativa* using LCMS and cryogenic NMR. *Phytochemistry* **87**: 51–9.

53. **Elhendawy MA, Wanas AS, Radwan MM, Azzaz NA**, *et al*. 2019. Chemical and Biological Studies of *Cannabis sativa* Roots. *Med Cannabis Cannabinoids* **1**(2): 104–11.

54. **Andre CM, Hausman J-F, Guerriero G**. 2016. *Cannabis sativa*: The Plant of the Thousand and One Molecules. *Front Plant Sci* **7**: 19.

55. **Schachtsiek J, Warzecha H, Kayser O, Stehle F**. 2018. Current Perspectives on Biotechnological Cannabinoid Production in Plants. *Planta Med* **84**(4): 214–20.

56. **Eisenreich W, Bacher A, Arigoni D, Rohdich F**. 2004. Biosynthesis of isoprenoids via the non-mevalonate pathway. *Cell Mol Life Sci* **61**(12): 1401–26.

57. **Stout JM, Boubakir Z, Ambrose SJ, Purves RW**, *et al*. 2012. The hexanoyl-CoA precursor for cannabinoid biosynthesis is formed by an acyl-activating enzyme in *Cannabis sativa* trichomes. *Plant J* **71**(3): 353–65.

58. **Gagne SJ, Stout JM, Liu E, Boubakir Z**, *et al*. 2012. Identification of olivetolic acid cyclase from *Cannabis sativa* reveals a unique catalytic route to plant polyketides. *Proc Natl Acad Sci U S A* **109**(31): 12811–6.

59. **Fellermeier M, Zenk MH**. 1998. Prenylation of olivetolate by a hemp transferase yields cannabigerolic acid, the precursor of tetrahydrocannabinol. *FEBS Lett* **427**(2): 283–5.

60. **Luo X, Reiter MA, d'Espaux L, Wong J**, *et al*. 2019. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* **567**(7746): 123–6.

61. **Page JE, Boubakir Z**. 2010. AROMATIC PRENYLTRANSFERASE FROM CANNABIS. US20090272057P;US20090272117P **C12N15/54;C07K14/415;C12N1/38;C12N9/99;C12P1/00**(WO2011017798 (A1)).

62. **Zirpel B, Degenhardt F, Martin C, Kayser O**, *et al*. 2017. Engineering yeasts as platform organisms for cannabinoid biosynthesis. *J Biotechnol* **259**: 204–12.

63. **Morimoto S, Komatsu K, Taura F, Shoyama Y**. 1998. Purification and characterization of cannabichromenic acid synthase from *Cannabis sativa*. *Phytochemistry* **49**(6): 1525–9.

64. **Sirikantaramas S, Morimoto S, Shoyama Y, Ishikawa Y**, *et al*. 2004. The gene controlling marijuana psychoactivity: molecular cloning and heterologous expression of Delta1-tetrahydrocannabinolic acid synthase from *Cannabis sativa* L. *J Biol Chem* **279**(38): 39767–74.

65. **Taura F, Sirikantaramas S, Shoyama Y, Yoshikai K**, *et al.* 2007. Cannabidiolic-acid synthase, the chemotype-determining enzyme in the fiber-type *Cannabis sativa*. *FEBS Lett* **581**(16): 2929–34.

66. **Daniel B, Konrad B, Toplak M, Lahham M**, *et al*. 2017. The family of berberine bridge enzyme-like enzymes: A treasure-trove of oxidative reactions. *Arch Biochem Biophys* **632**: 88–103.

67. **Shoyama Y, Tamada T, Kurihara K, Takeuchi A**, *et al*. 2012. Structure and Function of Δ1-Tetrahydrocannabinolic Acid (THCA) Synthase, the Enzyme Controlling the Psychoactivity of *Cannabis sativa*. *J Mol Biol* **423**(1): 96–105.

68. **Taura F, Dono E, Sirikantaramas S, Yoshimura K**, *et al*. 2007. Production of Δ^1-tetrahydrocannabinolic acid by the biosynthetic enzyme secreted from transgenic *Pichia pastoris*. *Biochem Biophys Res Commun* **361**(3): 675–80.

69. **Hanuš LO, Meyer SM, Muñoz E, Taglialatela-Scafati O**, *et al.* 2016. Phytocannabinoids: a unified critical inventory. *Nat Prod Rep* **33**(12): 1357–92.

70. **Luo Y, Li B-Z, Liu D, Zhang L**, *et al*. 2015. Engineered biosynthesis of natural products in heterologous hosts. *Chem Soc Rev* **44**(15): 5265–90.

71. **Calgaro-Kozina A, Vuu KM, Keasling JD, Loqué D**, *et al*. 2020. Engineering Plant Synthetic Pathways for the Biosynthesis of Novel Antifungals. *ACS Cent Sci* **6**(8): 1394–400.

72. **Veeresham C**. 2012. Natural products derived from plants as a source of drugs. *J Adv Pharm Technol Res* **3**(4): 200–1.

73. **Yun U-W, Yan Z, Amir R, Hong S**, *et al.* 2012. Plant natural products: history, limitations and the potential of cambial meristematic cells. *Biotechnol Genet Eng Rev* **28**: 47–59.

74. **Miralpeix B, Rischer H, Häkkinen ST, Ritala A**, *et al*. 2013. Metabolic engineering of plant secondary products: which way forward? *Curr Pharm Des* **19**(31): 5622–39.

75. **Zhang H, Boghigian BA, Armando J, Pfeifer BA**. 2011. Methods and options for the heterologous production of complex natural products. *Nat Prod Rep* **28**(1): 125–51.

76. **Markina NM, Kotlobay AA, Tsarkova AS**. 2020. Heterologous Metabolic Pathways: Strategies for Optimal Expression in Eukaryotic Hosts. *Acta Naturae* **12**(2): 28–39.

77. **Geissler M, Volk J, Stehle F, Kayser O**, *et al*. 2018. Subcellular localization defines modification and production of Δ^9-tetrahydrocannabinolic acid synthase in transiently transformed *Nicotiana benthamiana*. *Biotechnol Lett* **40**(6): 981–7.

78. **Carvalho Â, Hansen EH, Kayser O, Carlsen S**, *et al*. 2017. Designing microorganisms for heterologous biosynthesis of cannabinoids. *FEMS Yeast Res* **17**(4).

79. **Zhang N, An Z**. 2010. Heterologous Protein Expression in Yeasts and Filamentous Fungi. In Baltz RH, Davies JE, Demain AL, Bull AT *et al*. eds; *Manual of Industrial Microbiology and Biotechnology*. ASM Press. 145–156.

80. **Broadway N**. 2012. Recombinant Protein Expression: Vector-Host Systems. *Mater Methods* **2**.

81. **Ikram NKBK, Zhan X, Pan X-W, King BC**, *et al*. 2015. Stable heterologous expression of biologically active terpenoids in green plant cells. *Front Plant Sci* **6**: 129.

82. **Vasilev N, Schmitz C, Dong L, Ritala A**, *et al*. 2014. Comparison of plant-based expression platforms for the heterologous production of geraniol. *Plant Cell Tiss Organ Cult* **117**: 373–80.

83. **Bharadwaj RKB, Kumar SR, Sathishkumar R**. 2019. Green Biotechnology: A Brief Update on Plastid Genome Engineering. In Sathishkumar R, Kumar SR, Hema J, Baskar V eds; *Advances in Plant Transgenics: Methods and Applications*. Springer Singapore. 79–100.

84. **Hasunuma T, Miyazawa S-I, Yoshimura S, Shinzaki Y**, *et al*. 2008. Biosynthesis of astaxanthin in tobacco leaves by transplastomic engineering. *Plant J* **55**(5): 857–68.

85. **Ralley L, Enfissi EMA, Misawa N, Schuch W**, *et al*. 2004. Metabolic engineering of ketocarotenoid formation in higher plants. *Plant J* **39**(4): 477–86.

86. **Shanmugaraj B, Malla A, Phoolcharoen W**. 2020. Emergence of Novel Coronavirus 2019-nCoV: Need for Rapid Vaccine and Biologics Development. *Pathogens* **9**(2).

87. **Makhzoum A, Benyammi R, Moustafa K, Trémouillaux-Guiller J**. 2014. Recent advances on host plants and expression cassettes' structure and function in plant molecular pharming. *BioDrugs* **28**(2): 145–59.

88. **Tremblay R, Wang D, Jevnikar AM, Ma S**. 2010. Tobacco, a highly efficient green bioreactor for production of therapeutic proteins. *Biotechnol Adv* **28**(2): 214–21.

89. **Conley AJ, Zhu H, Le LC, Jevnikar AM**, *et al*. 2011. Recombinant protein production in a variety of *Nicotiana* hosts: a comparative analysis. *Plant Biotechnol J* **9**(4): 434–44.

90. **Park KY, Wi SJ**. 2016. Potential of plants to produce recombinant protein products. *J Plant Biol* **59**(6): 559–68.

91. **Medicago**. Study of a Recombinant Coronavirus-Like Particle COVID-19 Vaccine in Adults: NCT04636697, CP-PRO-CoVLP-021. Available at: https://ClinicalTrials.gov/show/NCT04636697.

92. **Malhotra K, Subramaniyan M, Rawat K, Kalamuddin M**, *et al*. 2016. Compartmentalized Metabolic Engineering for Artemisinin Biosynthesis and Effective Malaria Treatment by Oral Delivery of Plant Cells. *Mol Plant* **9**(11): 1464–77.

93. **Sarrion-Perdigones A, Vazquez-Vilar M, Palací J, Castelijns B**, *et al*. 2013. GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. *Plant Physiol* **162**(3): 1618–31.

94. **Engler C, Kandzia R, Marillonnet S**. 2008. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**(11): e3647.

95. **Weber E, Engler C, Gruetzner R, Werner S**, *et al*. 2011. A Modular Cloning System for Standardized Assembly of Multigene Constructs. *PLoS One* **6**(2): e16765.

96. **Sarrion-Perdigones A, Falconi EE, Zandalinas SI, Juárez P**, *et al*. 2011. GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. *PLoS One* **6**(7): e21622.

97. **Hellens RP, Edwards EA, Leyland NR, Bean S**, *et al*. 2000. pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. *Plant Mol Biol* **42**(6): 819–32.

98. **Chi-Ham CL, Boettiger S, Figueroa-Balderas R, Bird S**, *et al*. 2012. An intellectual property sharing initiative in agricultural biotechnology: development of broadly accessible technologies for plant transformation. *Plant Biotechnol J* **10**(5): 501–10.

99. **Komori T, Imayama T, Kato N, Ishida Y**, *et al*. 2007. Current status of binary vectors and superbinary vectors. *Plant Physiol* **145**(4): 1155–60.

100. **Vazquez-Vilar M, Quijano-Rubio A, Fernandez-Del-Carmen A, Sarrion-Perdigones A**, *et al*. 2017. GB3.0: a platform for plant bio-design that connects functional DNA elements with associated biological data. *Nucleic Acids Res* **45**(4): 2196–209.

101. **Marsh G**. 2003. Next step for automotive materials. *Mater Today* **6**(4): 36–43.

102. **Luque JS, Okere AN, Reyes-Ortiz CA, Williams PM**. 2021. Mixed methods study of the potential therapeutic benefits from medical cannabis for patients in florida. *Complement Ther Med*: 102669.

103. **Aguilar S, Gutiérrez V, Sánchez S, Nougier M ed**. 2018. Medicinal cannabis policies and practices around the world. *International Drug Policy Consortium*.

104. **Morreel K, Saeys Y, Dima O, Lu F**, *et al*. 2014. Systematic structural characterization of metabolites in *Arabidopsis* via candidate substrate-product pair networks. *Plant Cell* **26**(3): 929–45.

105. **Le Roy J, Huss B, Creach A, Hawkins S**, *et al*. 2016. Glycosylation Is a Major Regulator of Phenylpropanoid Availability and Biological Activity in Plants. *Front Plant Sci* **7**: 735.

106. **Chen Q, Lai H, Hurtado J, Stahnke J**, *et al*. 2013. Agroinfiltration as an Effective and Scalable Strategy of Gene Delivery for Production of Pharmaceutical Proteins. *Adv Tech Biol Med* **1**(1): 103.

107. **Allen L, Fu J, Pritchard J, Wellendorf R**, *et al*. 2016. Complex Variability within the THCA and CBDA Synthase Genes in *Cannabis* Species. *J Forensic Investigation* **4**(1).

108. **Onofri C, Meijer EPM** de, **Mandolino G**. 2015. Sequence heterogeneity of cannabidiolic- and tetrahydrocannabinolic acid-synthase in *Cannabis sativa* L. and its relationship with chemical phenotype. *Phytochemistry* **116**: 57–68.

109. **Sirikantaramas S, Taura F, Tanaka Y, Ishikawa Y**, *et al*. 2005. Tetrahydrocannabinolic acid synthase, the enzyme controlling marijuana psychoactivity, is secreted into the storage cavity of the glandular trichomes. *Plant Cell Physiol* **46**(9): 1578–82.

110. **Page JE, Stout JM**. 2015. CANNABICHROMENIC ACID SYNTHASE FROM CANNABIS SATIVA. US201462018128P **C12N15/53;C07K16/40;C12N1/19;C12N1/21;C12N5/10;C12N9/02;C12N15/113;C12N15/63; C12P17/06**(WO2015196275 (A1)).

111. **Bertani G**. 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. *J Bacteriol* **62**(3): 293–300.

112. **Bradford MM**. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248–54.

113. **Simpson RJ**. 2007. Staining proteins in gels with silver nitrate. *CSH Protoc* **2007**: pdb.prot4727.

114. **Lakatos L, Szittya G, Silhavy D, Burgyán J**. 2004. Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. *EMBO J* **23**(4): 876–84.

115. **Fräbel S**. 2016. Characterization of flavin-dependent tryptophan halogenases and their application in plant metabolic engineering. Dissertation.

116. **Sun S-H, Xie J-P, Xie F-W, Zong Y-L**. 2008. Determination of volatile organic acids in oriental tobacco by needle-based derivatization headspace liquid-phase microextraction coupled to gas chromatography/mass spectrometry. *J Chrom A* **1179**(2): 89–95.

117. **Wahrheit J, Nicolae A, Heinzle E**. 2011. Eukaryotic metabolism: measuring compartment fluxes. *Biotechnol J* **6**(9): 1071–85.

118. **Mahmood T, Yang P-C**. 2012. Western blot: technique, theory, and trouble shooting. *N Am J Med Sci* **4**(9): 429–34.

119. **Pillai-Kastoori L, Schutz-Geschwender AR, Harford JA**. 2020. A systematic approach to quantitative Western blot analysis. *Anal Biochem* **593**: 113608.

120. **Bick JA, Lange BM**. 2003. Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. *Arch Biochem Biophys* **415**(2): 146–54.

121. **Volk J, Sarafeddinov A, Unver T, Marx S**, *et al*. 2019. Two novel methylesterases from *Olea europaea* contribute to the catabolism of oleoside-type secoiridoid esters. *Planta* **250**(6): 2083–97.

122. **de Brito Francisco R**, **Martinoia E**. 2018. The Vacuolar Transportome of Plant Specialized Metabolites. *Plant Cell Physiol* **59**(7): 1326–36.

123. **Rivas F, Parra A, Martinez A, Garcia-Granados A**. 2013. Enzymatic glycosylation of terpenoids. *Phytochem Rev* **12**(2): 327–39.

124. **Gülck T, Booth JK, Carvalho Â, Khakimov B**, *et al*. 2020. Synthetic Biology of Cannabinoids and Cannabinoid Glucosides in *Nicotiana benthamiana* and *Saccharomyces cerevisiae*. *J Nat Prod* **83**(10): 2877–93.

125. **Li N, Xu C, Li-Beisson Y, Philippar K**. 2016. Fatty Acid and Lipid Transport in Plant Cells. *Trends Plant Sci* **21**(2): 145–58.

126. **Huestis MA**. 2007. Human cannabinoid pharmacokinetics. *Chem Biodivers* **4**(8): 1770–804.

127. **Thorson J, Hosted Jr. T, Jiang J, Biggins J**, *et al*. 2001. Natures Carbohydrate Chemists The Enzymatic Glycosylation of Bioactive Bacterial Metabolites. *Curr Org Chem* **5**(2): 139–67.

128. **Hardman JM, Brooke RT, Zipp BJ**. 2017. Cannabinoid glycosides: *In vitro* production of a new class of cannabinoids with improved physicochemical properties. BioRxiv doi: https://doi.org/10.1101/104349

129. **Hardcastle JD, Wilkins JL**. 1970. The action of sennosides and related compounds on human colon and rectum. *Gut* **11**(12): 1038–42.

130. **Shoyama Y, Sugawa C, Tanaka H, Morimoto S**. 2008. Cannabinoids act as necrosis-inducing factors in *Cannabis sativa*. *Plant Signal Behav* **3**(12): 1111–2.

131. **Morimoto S, Tanaka Y, Sasaki K, Tanaka H**, *et al*. 2007. Identification and characterization of cannabinoids that induce cell death through mitochondrial permeability transition in *Cannabis* leaf cells. *J Biol Chem* **282**(28): 20739–51.

132. **Wang S, Alseekh S, Fernie AR, Luo J**. 2019. The Structure and Function of Major Plant Metabolite Modifications. *Mol Plant* **12**(7): 899–919.

133. **Livingston SJ, Quilichini TD, Booth JK, Wong DCJ**, *et al*. 2020. Cannabis glandular trichomes alter morphology and metabolite content during flower maturation. *Plant J* **101**(1): 37–56.

134. **Kley J, Heil M, Muck A, Svatos A**, *et al.* 2010. Isolating intact chloroplasts from small *Arabidopsis* samples for proteomic studies. *Anal Biochem* **398**(2): 198–202.

135. **Kodama Y**. 2016. Time Gating of Chloroplast Autofluorescence Allows Clearer Fluorescence Imaging *In Planta*. *PLoS One* **11**(3): e0152484.

136. **Chalfie M, Tu Y, Euskirchen G, Ward WW**, *et al*. 1994. Green fluorescent protein as a marker for gene expression. *Science* **263**(5148): 802–5.

137. **Kuzuyama T, Noel JP, Richard SB**. 2005. Structural basis for the promiscuous biosynthetic prenylation of aromatic natural products. *Nature* **435**(7044): 983–7.

138. **Kayser O, Stehle FO**. 2019. BIOTECHNOLOGICAL PRODUCTION OF CANNABINOIDS. DE201810117233 **C12N9/10;C12N15/81;C12P5/00;C12P7/22**(WO2020016287 (A1)).

139. **Tsuji Y**. 2020. Transmembrane protein western blotting: Impact of sample preparation on detection of SLC11A2 (DMT1) and SLC40A1 (ferroportin). *PLoS One* **15**(7): e0235563.

140. **Chun E, Thompson AA, Liu W, Roth CB**, *et al*. 2012. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure* **20**(6): 967–76.

141. **Schultz EE, Braffman NR, Luescher MU, Hager HH**, *et al*. 2019. Biocatalytic Friedel-Crafts Alkylation Using a Promiscuous Biosynthetic Enzyme. *Angew Chem Int Ed* **58**(10): 3151–5.

142. **Seddon AM, Curnow P, Booth PJ**. 2004. Membrane proteins, lipids and detergents: not just a soap opera. *Biochim Biophys Acta* **1666**(1-2): 105–17.

143. **Henry LK, Gutensohn M, Thomas ST, Noel JP**, *et al*. 2015. Orthologs of the archaeal isopentenyl phosphate kinase regulate terpenoid production in plants. *Proc Natl Acad Sci U S A* **112**(32): 10050–5.

144. **Vranová E, Coman D, Gruissem W**. 2013. Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu Rev Plant Biol* **64**: 665–700.

145. **Burke CC, Wildung MR, Croteau R**. 1999. Geranyl diphosphate synthase: cloning, expression, and characterization of this prenyltransferase as a heterodimer. *Proc Natl Acad Sci U S A* **96**(23): 13062–7.

146. **Wang K, Ohnuma S**. 1999. Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution. *Trends Biochem Sci* **24**(11): 445–51.

147. **Ueoka H, Sasaki K, Miyawaki T, Ichino T**, *et al*. 2020. A Cytosol-Localized Geranyl Diphosphate Synthase from Lithospermum erythrorhizon and Its Molecular Evolution. *Plant Physiol* **182**(4): 1933–45.

148. **Orlova I, Nagegowda DA, Kish CM, Gutensohn M**, *et al.* 2009. The Small Subunit of Snapdragon Geranyl Diphosphate Synthase Modifies the Chain Length Specificity of Tobacco Geranylgeranyl Diphosphate Synthase in Planta. *Plant Cell* **21**(12): 4002–17.

149. **Dharmawardhana DP, Ellis BE, Carlson JE**. 1992. Characterization of vascular lignification in *Arabidopsis thaliana*. *Can J Bot* **70**(11): 2238–44.

150. **Dharmawardhana DP, Ellis BE, Carlson JE**. 1995. A β-Glucosidase from Lodgepole Pine Xylem Specific for the Lignin Precursor Coniferin. *Plant Physiol* **107**(2): 331–9.

151. **Seshadri S, Akiyama T, Opassiri R, Kuaprasert B**, *et al*. 2009. Structural and Enzymatic Characterization of Os3BGlu6, a Rice β-Glucosidase Hydrolyzing Hydrophobic Glycosides and (1→3)- and (1→2)-Linked Disaccharides. *Plant Physiol* **151**(1): 47–58.

152. **Wang P, Liu H, Hua H, Wang L**, *et al*. 2011. A vacuole localized β-glucosidase contributes to drought tolerance in *Arabidopsis*. *Chin Sci Bull* **56**(33): 3538–46.

153. **Pankoke H, Buschmann T, Müller C**. 2013. Role of plant β-glucosidases in the dual defense system of iridoid glycosides and their hydrolyzing enzymes in *Plantago lanceolata* and *Plantago major*. *Phytochemistry* **94**: 99–107.

154. **Sampedro J, Valdivia ER, Fraga P, Iglesias N**, *et al*. 2017. Soluble and Membrane-Bound β-Glucosidases Are Involved in Trimming the Xyloglucan Backbone. *Plant Physiol* **173**(2): 1017–30.

155. **Morant AV, Jørgensen K, Jørgensen C, Paquette SM**, *et al*. 2008. β-Glucosidases as detonators of plant chemical defense. *Phytochemistry* **69**(9): 1795–813.

156. **Holden DW, Rohringer R**. 1985. Peroxidases and glycosidases in intercellular fluids from noninoculated and rust-affected wheat leaves isozyme assay on nitrocellulose blots from two-dimensional gels. *Plant Physiol* **79**(3): 820–4.

157. **Volk J**. 2014. Investigating the Recombinant Production of THC in Heterologous Plant Models. Forschungspraktikum Masterarbeit.

158. **Trott O, Olson AJ**. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**(2): 455–61.

159. **Sterling T, Irwin JJ**. 2015. ZINC 15--Ligand Discovery for Everyone. *J Chem Inf Model* **55**(11): 2324–37.

160. **O'Boyle NM, Morley C, Hutchison GR**. 2008. Pybel: a Python wrapper for the OpenBabel cheminformatics toolkit. *Chem Cent J* **2**: 5.

161. **Kunzmann P, Hamacher K**. 2018. Biotite: a unifying open source computational biology framework in Python. *BMC Bioinformatics* **19**(1): 346.

162. **Pettersen EF, Goddard TD, Huang CC, Couch GS**, *et al*. 2004. UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem* **25**(13): 1605–12.

163. **Waterhouse A, Bertoni M, Bienert S, Studer G**, *et al*. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**(W1): 296–303.

164. **Benkert P, Tosatto SCE, Schomburg D**. 2008. QMEAN: A comprehensive scoring function for model quality assessment. *Proteins* **71**(1): 261–77.

165. **Hunter PR, Craddock CP, Di Benedetto S, Roberts LM**, *et al*. 2007. Fluorescent Reporter Proteins for the Tonoplast and the Vacuolar Lumen Identify a Single Vacuolar Compartment in Arabidopsis Cells. *Plant Physiol* **145**(4): 1371–82.

166. **Saito C, Ueda T, Abe H, Wada Y**, *et al*. 2002. A complex and mobile structure forms a distinct subregion within the continuous vacuolar membrane in young cotyledons of *Arabidopsis*. *Plant J* **29**(3): 245–55.

167. **Segami S, Makino S, Miyake A, Asaoka M**, *et al*. 2014. Dynamics of Vacuoles and H^+-Pyrophosphatase Visualized by Monomeric Green Fluorescent Protein in *Arabidopsis*: Artifactual Bulbs and Native Intravacuolar Spherical Structures. *Plant Cell* **26**(8): 3416–34.

168. **Lee DW, Jung C, Hwang I**. 2013. Cytosolic events involved in chloroplast protein targeting. *Biochim Biophys Acta* **1833**(2): 245–52.

169. **Tarentino AL, Trimble RB, Plummer TH**. 1989. Enzymatic approaches for studying the structure, synthesis, and processing of glycoproteins. *Methods Cell Biol* **32**: 111–39.

170. **Tretter V, Altmann F, März L**. 1991. Peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F cannot release glycans with fucose attached α1→3 to the asparagine-linked *N*-acetylglucosamine residue. *Eur J Biochem* **199**(3): 647–52.

171. **Barone RP, Knittel DK, Ooka JK, Porter LN**, *et al*. 2020. The production of plant natural products beneficial to humanity by metabolic engineering. *Curr Plant Biol* **24**: 100121.

172. **Zirpel B, Kayser O, Stehle F**. 2018. Elucidation of structure-function relationship of THCA and CBDA synthase from *Cannabis sativa* L. *J Biotechnol* **284**: 17–26.

173. **Kuo Y-M, Henry RA, Andrews AJ**. 2016. Measuring specificity in multi-substrate/product systems as a tool to investigate selectivity *in vivo*. *Biochim Biophys Acta* **1864**(1): 70–6.

174. **Mozhaev VV, Khmelnitsky YL, Sergeeva MV, Belova AB**, *et al*. 1989. Catalytic activity and denaturation of enzymes in water/organic cosolvent mixtures. α-Chymotrypsin and laccase in mixed water/alcohol, water/glycol and water/formamide solvents. *Eur J Biochem* **184**(3): 597–602.

175. **Stepankova V, Damborsky J, Chaloupkova R**. 2013. Organic co-solvents affect activity, stability and enantioselectivity of haloalkane dehalogenases. *Biotechnol J* **8**(6): 719–29.

176. **Tanford C**. 1968. Protein denaturation. *Adv Protein Chem* **23**: 121–282.

177. **Timasheff SN**. 1970. Protein-solvent interactions and protein conformation. *Acc Chem Res* **3**(2): 62–8.

178. **Ross SA, ElSohly MA**. 1996. The volatile oil composition of fresh and air-dried buds of *Cannabis sativa*. *J Nat Prod* **59**(1): 49–51.

179. **Adamczyk S, Adamczyk B, Kitunen V, Smolander A**. 2015. Monoterpenes and higher terpenes may inhibit enzyme activities in boreal forest soil. *Soil Biol Biochem* **87**: 59–66.

180. **Adamczyk S, Kiikkilä O, Kitunen V, Smolander A**. 2013. Potential response of soil processes to diterpenes, triterpenes and tannins: Nitrification, growth of microorganisms and precipitation of proteins. *Appl Soil Ecol* **67**: 47–52.

181. **Adamczyk B, Kitunen V, Smolander A**. 2009. Polyphenol oxidase, tannase and proteolytic activity in relation to tannin concentration in the soil organic horizon under silver birch and Norway spruce. *Soil Biol Biochem* **41**(10): 2085–93.

182. **Triebwasser DJ, Tharayil N, Preston CM, Gerard PD**. 2012. The susceptibility of soil enzymes to inhibition by leaf litter tannins is dependent on the tannin chemistry, enzyme class and vegetation history. *New Phytol* **196**(4): 1122–32.

183. **Schimel JP, Cates RG, Ruess R**. 1998. The role of balsam poplar secondary chemicals in controlling soil nutrient dynamics through succession in the Alaskan taiga. *Biogeochemistry* **42**(1/2): 221–34.

184. **Rodziewicz P, Loroch S, Marczak Ł, Sickmann A**, *et al.* 2019. Cannabinoid synthases and osmoprotective metabolites accumulate in the exudates of *Cannabis sativa* L. glandular trichomes. *Plant Sci* **284**: 108–16.

185. **Liu Y, Friesen JB, McAlpine JB, Lankin DC**, *et al.* 2018. Natural Deep Eutectic Solvents: Properties, Applications, and Perspectives. *J Nat Prod* **81**(3): 679–90.

186. **Weiblen GD, Wenger JP, Craft KJ, Elsohly MA**, *et al*. 2015. Gene duplication and divergence affecting drug content in *Cannabis sativa*. *New Phytol* **208**(4): 1241–50.

187. **Stavrinides A, Tatsis EC, Caputi L, Foureau E**, *et al*. 2016. Structural investigation of heteroyohimbine alkaloid synthesis reveals active site elements that control stereoselectivity. *Nat Commun* **7**: 12116.

188. **Leferink NGH, Ranaghan KE, Karuppiah V, Currin A**, *et al*. 2019. Experiment and Simulation Reveal How Mutations in Functional Plasticity Regions Guide Plant Monoterpene Synthase Product Outcome. *ACS Catal* **8**(5): 3780–91.

189. **Mitternacht S, Berezovsky IN**. 2011. On the Importance of Amino Acid Sequence and Spatial Proximity of Interacting Residues for Protein Folding. *J Biomol Struct Dyn* **28**(4): 607–9.

190. **Stickle DF, Presta LG, Dill KA, Rose GD**. 1992. Hydrogen bonding in globular proteins. *J Mol Biol* **226**(4): 1143–59.

191. **Nemethy G, Pottle MS, Scheraga HA**. 1983. Energy parameters in polypeptides. 9. Updating of geometrical parameters, nonbonded interactions, and hydrogen bond interactions for the naturally occurring amino acids. *J Phys Chem* **87**(11): 1883–7.

192. **Murashige T, Skoog F**. 1962. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiol Plant* **15**(3): 473–97.

193. **Linsmaier EM, Skoog F**. 1965. Organic Growth Factor Requirements of Tobacco Tissue Cultures. *Physiol Plant* **18**(1): 100–27.

194. **Bevan MW, Flavell RB, Chilton MD**. 1983. A chimaeric antibiotic resistance gene as a selectable marker for plant cell transformation. *Nature* **304**: 184–7.

195. **Bustin SA**. 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* **25**(2): 169–93.

196. **F de Felippes F, McHale M, Doran RL, Roden S**, *et al*. 2020. The key role of terminators on the expression and post-transcriptional gene silencing of transgenes. *Plant J* **104**(1): 96–112.

197. **Que Q, Wang HY, English JJ, Jorgensen RA**. 1997. The Frequency and Degree of Cosuppression by Sense Chalcone Synthase Transgenes Are Dependent on Transgene Promoter Strength and Are Reduced by Premature Nonsense Codons in the Transgene Coding Sequence. *Plant Cell* **9**(8): 1357–68.

198. **Bleys A, Vermeersch L, van Houdt H, Depicker A**. 2006. The frequency and efficiency of endogene suppression by transitive silencing signals is influenced by the length of sequence homology. *Plant Physiol* **142**(2): 788–96.

199. **Christie M, Croft LJ, Carroll BJ**. 2011. Intron splicing suppresses RNA silencing in Arabidopsis. *Plant J* **68**(1): 159–67.

200. **Dadami E, Moser M, Zwiebel M, Krczal G**, *et al*. 2013. An endogene-resembling transgene delays the onset of silencing and limits siRNA accumulation. *FEBS Lett* **587**(6): 706–10.

201. **Peremarti A, Twyman RM, Gómez-Galera S, Naqvi S**, *et al*. 2010. Promoter diversity in multigene transformation. *Plant Mol Biol* **73**(4-5): 363–78.

202. **Mette MF, van der Winden J, Matzke MA, Matzke AJ**. 1999. Production of aberrant promoter transcripts contributes to methylation and silencing of unlinked homologous promoters *in trans*. *EMBO J* **18**(1): 241–8.

203. **Mette MF, Aufsatz W, van der Winden J, Matzke MA**, *et al.* 2000. Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J* **19**(19): 5194–201.

204. **Mourrain P, van Blokland R, Kooter JM, Vaucheret H**. 2007. A single transgene locus triggers both transcriptional and post-transcriptional silencing through double-stranded RNA production. *Planta* **225**(2): 365–79.

205. **Zirpel B, Degenhardt F, Zammarelli C, Wibberg D**, *et al*. 2018. Optimization of Δ^9-tetrahydrocannabinolic acid synthase production in *Komagataella phaffii* via post-translational bottleneck identification. *J Biotechnol* **272-273**: 40–7.

206. **Mikheev M, Gao D**. 2018. PRODUCTION OF CANNABINOIDS IN YEAST. US201762531827P **C12P17/06;C12P7/42**(WO2019014490 (A1)).

207. **Poulos JL, Farnia AN**. 2014. Production of cannabinoids in yeast. US201462024099P **C12P7/42;C12N15/81**(WO2016010827A1).

208. **Good MC, Zalatan JG, Lim WA**. 2011. Scaffold proteins: hubs for controlling the flow of cellular information. *Science* **332**(6030): 680–6.

209. **Shih M-L, Morgan JA**. 2020. Metabolic flux analysis of secondary metabolism in plants. *Metab Eng Commun* **10**: e00123.

210. **Espinoza-Sánchez EA, Torres-Castillo JA, Rascón-Cruz Q, Zavala-García F**, *et al*. 2016. Production and characterization of fungal β-glucosidase and bacterial cellulases by tobacco chloroplast transformation. *Plant Biotechnol Rep* **10**(2): 61–73.

Declaration of own work

Experiments, data analysis, data representation and writing of the thesis were exclusively performed by me with exception of:

Chapter 4

- Cloning of α-level AAE1, OLS and OAC transcriptional units, for localization of proteins in the chloroplast or the apoplast was performed by Christoph Neubauer (Master Thesis). Western blot analysis of the produced proteins was performed by him as well. All experiments were performed under my supervision.

Chapter 5

- Cloning of α-level NphB and NphB(Q295L) transcriptional units was performed by Felix Hamburger (Bachelor Thesis) and Karen Schlechte (Bachelor Thesis), respectively, under my supervision.
- Microsomal and chloroplast preparations of PT4 as well as subsequent Western blot analysis and activity assays were performed by Lisa Niederreiter (Master Thesis) under my supervision.
- Confocal laser scanning microscopy was performed with help of Dr. Dominique Tandl.

Chapter 6

- Cloning of α-level THCAS and CBCAS transcriptional units was performed by Felix Hamburger (Bachelor Thesis) and Christoph Neubauer (Bachelor Thesis), respectively, under my supervision.
- Docking of CBGA into the putative active site of THCAS was performed by Patrick Kunzmann. Visualization was done by me.
- Introduction of different mutations into CBCAS and THCAS, followed by transient transformation of *N. benthamiana* plants and activity assays was performed with the support of my students Milana Kremenovic (Master research internship), Yannik Käseberg (Bachelor Thesis) and Franziska Heid (Bachelor Thesis).
- Some parts of Chapter 6 were already published in a similar way in Geissler *et al.*, 2018.

Danksagung

An erster Stelle möchte ich mich recht herzlich bei Herrn Prof. Dr. Heribert Warzecha bedanken, welcher mir nicht nur das Anfertigen meiner Dissertation in seiner Arbeitsgruppe ermöglichte, sondern mir stets mit Rat zur Seite stand. Vielen Dank für deine Unterstützung und das Vertrauen, welches du mir während dieser Zeit entgegengebracht hast.

Ganz herzlich möchte ich mich auch bei Herrn Prof. Dr. Bertl für die Übernahme des Zweitgutachtens bedanken. Mein herzlicher Dank gilt außerdem der Arbeitsgruppe von Herrn Prof. Dr. Kayser für deren Zusammenarbeit während dieses Cannabis-Projektes.

Ein riesiges Dankeschön geht an die gesamte Arbeitsgruppe Warzecha für das fantastische Arbeitsklima, die tollen Gespräche, die Hilfestellungen und Ideenaustausche und natürlich für die einzigartigen AG Events. Die Zeit mit euch war wundervoll und unvergesslich! Dabei möchte ich vor allem noch einmal explizit unsere „Labormama“ Simone Bartl-Zimmermann, aber auch Dr. Jascha Volk, Dr. Sabine Fräbel und Elisabeth Haumann hervorheben. Ihr seid mir in meiner gesamten Zeit in der Arbeitsgruppe, die nicht erst mit der Promotion, sondern bereits mit der Bachelorarbeit im Jahr 2014 begann, zu wirklich guten Freunden geworden.

Zudem möchte ich mich bei Dr. Iryna Gerasymenko für die vielen hilfreichen Ratschläge und Ideen, aber auch für das Korrekturlesen der Arbeit bedanken.

Des Weiteren möchte ich mich bei Patrick Kunzmann für die durchgeführte Substratdocking-Simulation und Dr. Dominique Tandl für die Hilfe am CLSM bedanken.

Mein ganz besonderer Dank gilt ebenfalls meinen Masteranden, Bacheloranden und Praktikanten Felix Hamburger, Karen Schlechte, Christoph Neubauer, Theresa Ernst, Louis Sollinger, Yannik Käseberg, Franziska Heid, Lisa Niederreiter, Phillip Dahlhaus, Andrea Nigl und Milana Kremenovic, welche mich Tatkräftig im Labor unterstützten.

Außerdem danke ich Birgit und Renate für die zuverlässige und gute Pflege aller Pflanzen, die ich im Zuge meiner Promotion benötigte.

Auch möchte ich mich bei meinen Eltern und Großeltern, sowie bei meiner Schwester bedanken, die mich während meiner gesamten akademischen Laufbahn stets uneingeschränkt unterstützt haben. Abschließend gilt jedoch mein größter Dank, dir, Jana. Dafür, dass du mir während den stressigen und auch anstrengenden Phasen meiner Doktorandenzeit stets mit deiner liebevollen Unterstützung zur Seite standst und maßgeblich dazu beitrugst, dass ich es überhaupt schaffte die Motivation und die Energie für den eingeschlagenen Weg aufzubringen. Danke für Alles!

Curriculum Vitae

Personal information

Name	Marcus Frank Geißler
Address	Berbigstraße 49, 64739 Höchst im Odenwald, Germany
Date of Birth	22/01/1991
Place of Birth	Groß-Umstadt
Nationality	German

Academic career

06/2017 –06/2021	**PhD Student** at the Department of Biology Technische Universität Darmstadt, Darmstadt (Germany) Plant Biotechnology and Metabolic Engineering **Title:** 'Metabolic engineering of cannabinoid biosynthesis in tobacco'
04/2015 – 05/2017	**Master of Science**, Technische Biologie Technische Universität Darmstadt, Darmstadt (Germany) Plant Biotechnology and Metabolic Engineering **Title:** 'Metabolic engineering of the seco-iridoid pathway in *Nepeta cataria* L.'
10/2011 – 04/2015	**Bachelor of Science**, Biology Technische Universität Darmstadt, Darmstadt (Germany) Plant Biotechnology and Metabolic Engineering **Title:** 'Molecular cloning and characterization of reductases involved in alkaloid biosynthesis in *Rauvolfia serpentina*'
07/2001 – 07/2010	**General University Entrance Qualification** Ernst-Göbel-Schule, Höchst im Odenwald (Germany)

Conference contributions

Poster

- 2nd German Conference on Synthetic Biology (GASB), Berlin (Germany), September 2018: **'Biotechnological production of cannabinoids: Characterization of late biosynthetic enzymes *in planta.'***

Publications

11/2020 European Patent Application EP 20 20 7220.3

M. Sc. Marcus Geissler, M. Sc. Lisa Niederreiter, Dr. Heribert Warzecha; Production of cannabinoids/phytocannabinoids with a plant extract.

06/2018 **Geissler M.**, Volk J., Stehle F., Kayser O., Warzecha H.; Subcellular localization defines modification and production of Δ^9-tetrahydrocannabinolic acid synthase in transiently transformed *Nicotiana benthamiana*, Biotechnology Letters (2018) 40: 981.

03/2016 **Geissler M.**, Burghard M., Volk J., Staniek A., Warzecha H.; A novel cinnamyl alcohol dehydrogenase (CAD)-like reductase contributes to the structural diversity of monoterpenoid indole alkaloids in *Rauvolfia*, Planta (2016) 243: 813.

Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit entsprechend den Regeln guter wissenschaftlicher Praxis selbstständig und ohne unzulässige Hilfe Dritter angefertigt habe.

Sämtliche aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie sämtliche von Anderen direkt oder indirekt übernommenen Daten, Techniken und Materialien sind als solche kenntlich gemacht. Die Arbeit wurde bisher bei keiner anderen Hochschule zu Prüfungszwecken eingereicht.

Darmstadt, den 08.07.2021