TECHNISCHE UNIVERSITÄT DARMSTADT

Characterization of flavin-dependent tryptophan halogenases and their application in plant metabolic engineering

Vom Fachbereich Biologie der Technischen Universität Darmstadt

zur

Erlangung des akademischen Grades

eines *Doctor rerum naturalium*

genehmigte Dissertation von

Dipl.-Biol. Sabine Fräbel

aus Darmstadt

1. Referent: Prof. Dr. Heribert Warzecha

2. Referent: Prof. Dr. Gerhard Thiel

Tag der Einreichung: 02.11.2015

Tag der mündlichen Prüfung: 18.12.2015

Darmstadt 2016

D 17

I. Table of contents

II. List of abbreviations

General abbreviations

Abbr.	abbreviations
BFP	blue fluorescent protein
CDS	coding sequence
CP	chloroplast targeting
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
ER	endoplasmic reticulum
EV	empty vector
FAD/$FADH_2$	flavin adenine dinucleotide
GB	GoldenBraid
HPLC	high-performance liquid chromatography
LC	liquid chromatography
MS	mass spectrometry
MS/MS	tandem mass spectrometry
m/z	mass-to-charge
n.a.	not available
NADPH	nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
OD	optical density
PCR	polymerase chain reaction
PS	photosystem
pUPD	universal domesticator
RNA	ribonucleic acid
RT	room temperature
UPLC	ultra-performance liquid chromatography
TLC	thin-layer chromatography
TU	transcriptional unit
wt	wild type
YFP	yellow fluorescent protein

Chemical abbreviations

$CaCl_2$	calcium chloride
DTT	dithiothreitol
EtOH	ethanol
HCl	hydrogen chloride
H_2O	water
IAA	indole-3-acetic acid
KBr	potassium bromide
KCl	potassium chloride
KOAc	potassium acetate
MeOH	methanol
MES	morpholinoethanesulfonic acid
MOPS	3-morpholinopropanesulfonic acid
NaCl	sodium chloride
X-Gal	5-bromo-4-chloro-3-indolylb-D-galactopyranoside acid
RbCl	rubidium chloride

Symbols

bp	base pairs
g	gram
g	G-force
h	hour
k	kilo (10^3)
k_{cat}	catalytic rate constant
K_M	Michaelis constant
n	nano (10^{-9})
L	liter
M	molar
m	milli (10^{-3})
min	minute
nt	nucleotide

µ	micro (10^{-6})
R_f	retention factor
rpm	rounds per minute
R_t	retention time
s	second
u	unit
V	volt
v/v	volume per volume
w/v	weight per volume

III. Figures

IV. Abstract

A huge variety of halogenated metabolites found in nature have a profound pharmacological effect or act as antimicrobials like the antibiotics, vancomycin and chloramphenicol, or the antitumor agent, rebeccamycin. Due to the high demand for halogenated compounds, which can be met only partially by chemical synthesis, intense research effort has been undertaken to characterize enzymes catalyzing halogenation reactions in nature and to uncover their reaction mechanisms with the aim to utilize biotechnological production strategies for the retrieval of these high-value compounds. Within the last 15 years, several bacterial flavin-dependent tryptophan halogenases have been characterized in terms of their regiospecific chlorine substitution of arenes. In this regard, halogenation of pharmacologically important secondary metabolites is of special interest, to introduce novel functions into given compounds or enable further modification of the skeleton by substitution. Also, capability of tryptophan-halogenases for application in plant biotechnology has been initially tested. Biosynthesis of chlorinated monoterpene indole alkaloids (MIAs) was previously demonstrated in *Catharanthus roseus* through halogenation of a precursor molecule by two tryptophan halogenases. Based on these findings, catalytic activity of three tryptophan halogenases, namely, RebH wt, RebH Y455W and Stth, was investigated in detail regarding subcellular localization, biosynthesis of valuable fine chemicals and modification of the precursor of all MIAs, strictosidine. In this regard, both the 7-halogenase, RebH wt as well as the 6-halogenase, Stth efficiently catalyzed chlorine substitution of tryptophan and tryptamine in the cytosol and chloroplasts of transiently transformed *Nicotiana benthamiana*. Halogenated products accumulated in high concentrations, up to 6.17 ± 2 ng/mg fresh weight (6-chlorotryptamine). Strikingly, both halogenases were active in chloroplasts without the partner reductase, RebF, whereas no enzymatic activity was observed after translocation to the apoplast. Moreover, tandem halogenation of tryptophan, but not tryptamine, was observed when both enzymes were co-localized in the cytosol or chloroplasts. RebH wt alone also synthesized minor amounts of di-chlorotryptophan. Additionally, both enzymes were shown to efficiently catalyze bromide substitution of tryptophan, resulting in a variety of mono-brominated and di-brominated tryptophan molecules as well as chloro-bromotryptophan. The engineered 7-halogenase, RebH Y455W, reported to predominantly chlorinate tryptamine instead of tryptophan, showed only low catalytic activity *in planta*. This inefficiency could not be compensated by optimization of the metabolic flux through anchoring of the involved enzymes within a protein scaffold. On top of the aforementioned studies on MIA biosynthesis, the halogenases were also incorporated into a newly designed indoxyl biosynthetic pathway to synthesize chlorinated indican derivatives *in planta*.

Introduction of tryptophanase TnaA from *Escherichia coli*, into a previously investigated artificial indican pathway resulted in high yields of 6- and 7-chloroindican. Moreover, subcellular localization of enzymes of this optimized metabolic route was shown to be very flexible and allowed both co-localization and separation of enzymes in the cytosol and chloroplasts. Remarkably, the human CYP450 2A6 mutant L240C/N297Q was also active in chloroplasts, which implies transport of electrons required for substrate oxidation, presumably from the photosystem I, to the cytochrome P450. Further optimization of this production system by introduction of additional enzymes or establishment of stable transgenic tobacco plants and cell cultures might enable efficient and ecological biosynthesis of a huge variety of highly valuable indigoids *in planta*.

To reach the final goal of obtaining halogenated monoterpenoid indole alkaloids in a synthetic pathway, reconstitution and modification of the strictosidine biosynthetic pathway was analyzed *in planta*. In this regard, accumulation of 14 new metabolites was associated with transgene expression; biosynthesis of five of those was enhanced by co-infiltration of the initial precursor, geraniol, whereas four were synthesized exclusively upon geraniol supplementation. However, no actual pathway intermediates could be identified within this group. Even though biosynthesis of precursors was enhanced and constitutive gene expression was facilitated by establishing transgenic tobacco lines, no metabolites of interest were observed. Therefore, the biosynthetic track needs to be further optimized by elimination of potential bottlenecks and replacement of the inefficient RebH Y455W by wild type halogenases. Taken together, the analyzed flavin-dependent tryptophan halogenases represent promising tools for biosynthesis of valuable molecules. Their substantive efficiency enables economical production of halogenated metabolites in high yields *in planta*.

V. Zusammenfassung

Eine Vielzahl natürlich vorkommender halogenierter Substanzen besitzt eine nachgewiesene antimikrobielle oder pharmakologische Wirkung wie beispielsweise die Antibiotika Vancomycin und Chloramphenicol oder das potentielle Chemotherapeutikum Rebeccamycin. Aufgrund der großen Relevanz sowie der aufwendigen und schwierigen chemischen Synthese dieser Substanzen wurde intensiv an der Identifizierung und Charakterisierung halogenierender Enzyme sowie den zugrunde liegenden Reaktionsmechanismen geforscht. Innerhalb der letzten 15 Jahre wurden dabei verschiedene bakterielle, Flavin-abhängige Tryptophan-Halogenasen charakterisiert und hinsichtlich ihrer regio-spezifischen Substitution von Chlorid an aromatische Ringsysteme untersucht. Von großem Interesse ist dabei die Halogenierung sekundärer Pflanzenstoffe, um neue Wirkmechanismen hervorzurufen oder eine weitere Modifizierung des Moleküls durch Substitutionsreaktionen zu vereinfachen. Daher ist die Anwendung dieser Enzyme in der Pflanzenbiotechnologie von großem Interesse. Die erfolgreiche Biosynthese chlorierter monoterpenoider Indolalkaloide (MIAs) durch die Halogenierung von Vorläufermolekülen durch die Tryptophan-Halogenase RebH wurde bereits in *Catharanthus roseus* beschrieben. Basierend auf diesen Erkenntnissen wurden die drei Tryptophan-Halogenasen RebH wt, RebH Y455W und zum ersten Mal Stth hinsichtlich ihrer katalytischen Aktivität, der Synthese hochwertiger Feinchemikalien sowie der Modifizierung von Strictosidin, das Vorläufermolekül aller MIAs, untersucht. In diesem Zusammenhang konnte gezeigt werden, dass sowohl RebH wt als auch Stth die Chlorierung von Tryptophan und Tryptamin mit großer Effizienz im Cytosol sowie in Chloroplasten katalysieren. Die halogenierten Produkte akkumulierten in hohen Konzentrationen mit bis zu 6.17 ± 2 ng/mg Frischgewicht (6-Chlortryptamin) in transient transformierten *Nicotiana benthamiana* Blättern. Bemerkenswerter Weise waren beide Halogenasen auch ohne die Flavinreduktase RebF in Chloroplasten aktiv, während im Apoplasten keine enzymatische Aktivität nachgewiesen werden konnte. Des Weiteren wurde gezeigt, dass die Co-Lokalisation beider Enzyme im Cytosol oder den Chloroplasten eine Doppelhalogenierung von Tryptophan, jedoch nicht von Tryptamin, bewirkt. RebH wt synthetisierte alleine ebenfalls geringer Mengen an Dichlortryptophan. Außerdem waren beide Halogenasen in der Lage, Tryptophan in sehr effizienter Weise zu bromieren, was zur Synthese mono- und dibrominierter Tryptophanderivate sowie Chlor-Bromtryptophan führte. Im Gegensatz dazu wies die Halogenase-mutante, RebH Y455W, welche Tryptamin anstelle von Tryptophan umsetzt, nur eine sehr geringe Aktivität auf. Diese Ineffizienz konnte auch durch die Verbesserung des Metabolitenflusses, nach Verankerung aller relevanten Enzyme an einem Proteingerüst, nicht kompensiert werden. Neben der

katalytischen Charakterisierung, wurden die drei Halogenasen in den ursprünglich von Warzecha *et al.* entwickelten Indoxyl-Biosyntheseweg integriert, um auf diese Weise die Produktion halogenierter Indikanderivate *in planta* zu ermöglichen. Die Synthese dieser wertvollen Feinchemikalien war jedoch weder nach transienter oder stabiler Transformation von Tabak nachweisbar. Die Optimierung des Synthesewegs durch Integration der Tryptophanase TnaA aus *Escherichia coli* resultierte zuletzt in einer hohen Akkumulation von 6- und 7-Chlorindikan. Des Weiteren, ist der optimierte Syntheseweg äußerst flexibel und erlaubt sowohl eine Co-Lokalisation als auch die Aufteilung der Enzyme zwischen dem Cytosol und den Chloroplasten. Interessanter Weise konnte für die die humane Cytochrom P450 Mutante, 2A6 L240C/N297Q, in Chloroplasten ebenfalls eine hohe Aktivität nachgewiesen werden, was auf einen alternativen Elektronentransport, vermutlich vom Photosystem I, hindeutet. Eine weitere Optimierung dieses Produktionssystems durch die Einführung weiterer Enzyme sowie der Etablierung transgener Tabakpflanzen oder Zellkulturen, ermöglicht die effiziente und ökologische Produktion einer Vielzahl halogenierter Indigoide *in planta*.

Zuletzt wurde die Rekonstitution sowie die Modifikation des Strictosidin-Biosynthesewegs zur Produktion halogenierter Derivate dieses MIA-Vorläufermoleküls untersucht. Dabei konnten 14 Metaboliten detektiert werden, deren Akkumulation mit der Expression der Transgene korrelierte; die Biosynthese fünf dieser Moleküle wurde durch die Zugabe des Vorläufers Geraniol verstärkt, während vier Substanzen ausschließlich nach Geraniol-Applikation nachweisbar waren. Allerdings konnte keiner dieser Metaboliten dem Strictosidin-Biosyntheseweg zugeordnet werden. Auch die weitere Optimierung durch eine verbesserte Synthese von Vorläufermolekülen sowie der konstitutiven Genexpression in transgenem Tabak, führte nicht zur Akkumulation der gewünschten Moleküle. Um eine effiziente Produktion halogenierter Strictosidinderivate zu ermöglichen, müssen daher Engpässe im Biosyntheseweg beseitigt und die ineffiziente Mutante RebH Y455W durch Wildtyphalogenasen ersetzt werden. Zusammengefasst ist die Verwendung der hier untersuchten Flavin-abhängigen Tryptophan-Halogenasen zur Biosynthese wertvoller Moleküle in Pflanzen sehr vielversprechend. Ihre hohe Effizienz ermöglicht eine wirtschaftliche und ökologische Produktion wertvoller halogenierter Substanzen mit großen Ausbeuten *in planta*.

1. Introduction

Halogenated metabolites are synthesized in nature by a huge variety of organism, such as bacteria, plants, fungi and mammals [1]. Many of those compounds have biological effects like the antibiotics, vancomycin and chloramphenicol or the antitumor agent, rebeccamycin [2-4]. Due to the pharmacological importance of these halogenated metabolites, intense research effort has been undertaken to characterize reaction mechanisms and halogenating enzymes [5]. Thereby, considerable focus was put on regiospecific halogenation of arenes by flavin-dependent halogenases, which is a challenging task in chemical synthesis [6-9]. Therefore, enzymes catalyzing chlorination and bromination of organic compounds at specific positions represent highly promising tools for efficient biosynthesis of valuable molecules at relatively low costs. Given that many naturally occurring halogenated metabolites have biological functions, chlorination of pharmacologically important molecules, like monoterpenoid indole alkaloids (MIAs), provides the opportunity to introduce novel functions into given metabolites and to simplify their further chemical remodeling in terms of enhanced bioactivity or pharmacokinetics [5, 10]. Such metabolic engineering can be achieved by halogenation of precursor molecules, like tryptophan, by flavin-dependent tryptophan halogenases, as exemplified in *Catharanthus roseus* by Runguphan *et al.* [11]. Modification of specific biosynthetic pathways can be achieved by reconstitution in an alternative host and co-localization of modifying enzymes, such as tryptophan halogenases. The production of plant-derived secondary metabolites in an alternative plant host benefits from cellular compartmentalization of heterologous enzymes and supply of plant-derived precursor molecules [12-14]. However, this complex system also is not free from disadvantages and might lead to inefficient biosynthesis due to loss of intermediates to competing pathways [15-18]. These unintended side reactions can be circumvented by metabolite channeling through co-localization of enzymes in close proximity on a scaffold, as demonstrated by Dueber *et al.* in *Escherichia coli* [19]. Another issue of reconstitution of complex biosynthetic pathways arises from the genetic side. Molecular cloning of several genes by conventional methods is tedious and complex. Each gene needs to be designed and assembled individually to functional transcriptional units and multigene constructs. Therefore, standardized cloning methods are preferable to facilitate easy and efficient reconstitution of entire pathways. In this regard, development of the modular cloning system, GoldenBraid by the group of Diego Orzáez opened up new possibilities in molecular engineering of complex pathways in a timesaving and flexible manner [20, 21].

1.1. Flavin-dependent halogenases

One class of halogenating enzymes found in nature are flavin-dependent halogenases, which originate from bacteria [5]. These enzymes depend on molecular oxygen, chlorine and $FADH_2$, synthesized by partner reductases, as an electron donor [5]. This class can be further subdivided into two major groups. Enzymes of the first group halogenate substrates bound to a thiolated domain of carrier proteins, like the halogenase PltA from *Pseudomonas fluorescens* [22]. Those carrier proteins are involved in the biosynthesis of non-ribosomal peptides (NRPs) or polyketides [23]. Members of the second group are for example the tryptophan halogenases PyrH from *Streptomyces rugosporus* LL-42D005 and RebH from *Lechevalieria aerocolonigenes*, which halogenate small free molecules [24, 25]. Despite single halogenation of tryptophan, tandem chlorination takes place in the biosynthetic pathway of kutznerides, which are non-ribosomal hexadepsipeptides with antifungal properties from *Kutzneria sp.*744 [26]. Here, flavin-dependent halogenases, KtzQ and KtzR catalyze the addition of chlorine at C6 and C7 of tryptophan [26]. Moreover, double halogenation catalyzed by only one enzyme was found in *P. fluorescens*. In this particular case, the flavin-dependent halogenase, PltA catalyzes the addition of chlorine at C4 and C5 positions of the pyrrole ring of pyoluteorin, leading to the biosynthesis of the antifungal agent, pyoluteorin [22].

1.1.1. Flavin-dependent tryptophan halogenases investigated in the presented studies

Two bacterial halogenases were characterized in more detail in terms of their catalytic properties and functional application in transiently transformed *Nicotiana benthamiana*: RebH, a tryptophan 7-halogenase found in *L. aerocolonigenes*, and Stth, a tryptophan 6-halogenase originating from the soil bacterium *Streptomyces toxytricini*. Moreover, a RebH mutant, generated by amino acid exchange, was selected for metabolic engineering of two biosynthetic pathways in transgenic *Nicotiana tabacum* plants.

The tryptophan halogenase, RebH and its partner flavin-reductase, RebF were first characterized by Yeh *et al.* in 2005 [25]. RebH has a molecular weight of 60,292 g/mol and forms homodimers comprising 530 amino acids per chain [4, 27]. The validated efficiency of halogenation catalyzed by this enzyme revealed a K_m value of 2.0 µM and a turnover number k_{cat} of 1.4 min^{-1} for tryptophan [25]. Besides substitution of chlorine to the heteroaromatic ring, RebH also brominates tryptophan, whereas substitution of iodide and fluoride is not catalyzed. Furthermore, the partner reductase has a molecular

weight of 18,434 g/mol and catalyzes the reduction of FAD to $FADH_2$. Its kinetic characterization revealed a k_{cat} of 108 min^{-1} for NADH oxidation as well as a K_m value of 0.7 µM for FAD, whereas NADPH was not oxidized [25]. Subsequent work by Payne *et al.* in 2013 shed light on the regiospecific halogenation of tryptophan at the C7 position [6]. Additionally, a broad variety of potential RebH substrates were discovered, including molecules like tryptamine, substituted naphthalenes and tryptoline. In contrast to the natural substrate, halogenation of these molecules is not limited to the C7 position of arenes, but it does not necessarily take place at the most preferable position for electrophilic substitution [6]. This regioselective halogenation is a unique feature of RebH compared to other characterized halogenases. For example, the tryptophan 7-halogenase, PrnA from *P. fluorescens* converts tryptamine as well as a variety of substituted indole derivatives, but halogenation of these molecules occurs at the electronically most preferred position for electrophilic substitution [7]. One reason for the improved regioselectivity of RebH in contrast to PrnA is, most probably, the exchange of one amino acid within the active site of the enzyme facilitating the formation of a hydrogen bond between tryptophan and Asn467 [6]. In PrnA the amino acid residue of Leu456 is located at the same position and does not stabilize the substrate in this way [6].

Glenn *et al.* characterized different mutants of RebH after replacement of distinct amino acids within the active site of the halogenase [28]. One of those, RebH Y455W, was found to halogenate tryptamine instead of the natural substrate, tryptophan. Due to this aspect, the mutated RebH constitutes an interesting tool for modification of tryptamine-derived secondary metabolites. In fact, introduction of RebH Y455W into *Catharanthus roseus* hairy roots resulted in the formation of a chlorinated monoterpene indole alkaloid [28]. In contrast to wild type halogenases, RebH Y455W does not reduce the tryptophan pool of the cells or contribute to accumulation of 7-chloroptryptophan [11].

The second flavin-dependent halogenase characterized within the presented study is the tryptophan 6-halogenase, Stth from *S. toxytricini*. This enzyme was first described by Zeng *et al.* in 2011 [29]. The encoding gene is located within an NRP synthetase gene cluster adjacent to the putative flavin reductase gene, *stfre*. Amino acid sequence alignments of Stth revealed 76 % identity to the tryptophan 6-halogenase, KtzR from *Kutzneria sp.* 744. Moreover, Stth comprises 523 amino acids and has a molecular weight of 58,715 g/mol. In contrast to RebH, Stth features a three times higher K_m value of

21 ± 3 µM in terms of tryptophan conversion [29]. Thus, the validated turnover number k_{cat} of 1.53 ± 0.12 min^{-1} was similar to the one of RebH [29].

1.1.2. Reaction mechanism of flavin-dependent tryptophan halogenases

Crystal structures of tryptophan 7-halogenases, PrnA from *P. fluorescens* and RebH from *L. aerocolonigenes* shed light on possible mechanisms of the regiospecific arene halogenation [27, 30]. $FADH_2$, synthesized by NAD-dependent flavin reductases, like RebF from *L. aerocolonigenes*, plays an essential role in this process. The molecule freely diffuses into the N-terminal FAD module of flavin-dependent halogenases, where it is located close to the characteristic signature motifs of these enzymes, GxGxxG and WxWxIP [31]. The crystal structure of PnrA revealed a 10 Å tunnel separating the flavin and tryptophan, thus making a direct interaction of both molecules unlikely [30]. In fact, $FADH_2$ reacts with molecular oxygen to form a FAD(C4a)-OOH intermediate [32]. Subsequently, the flavin peroxide is hydroxylated by a nucleophilic attack of a chlorine ion, which is stabilized in the active site by the amide nitrogen of Thr348 and Gly349 in PnrA [30]. It has been controversially discussed if the resulting hypochlorous acid (HOCl) performs a direct nucleophilic attack on the heteroaromatic ring, which might be disadvantageous for regioselective substitution [30, 33, 34]. Moreover, within the active sites of both halogenases, Lys79 is located closely to the flavin molecule and is positioned only 4.1 Å above C7 of tryptophan [30, 33]. Therefore, formation of chloramine and subsequent regioselective halogenation through a Lys-εNH-Cl intermediate seems the most likely mechanism for electrophilic aromatic substitution of arenes at C7 position [33]. In addition, residues of Glu357 (RebH) or Glu346 (PrnA), are perfectly positioned for deprotonation of the halogenated tryptophan leading to the final product [8]. The proposed reaction mechanism of flavin-dependent tryptophan halogenases is illustrated in fig. 1.1.

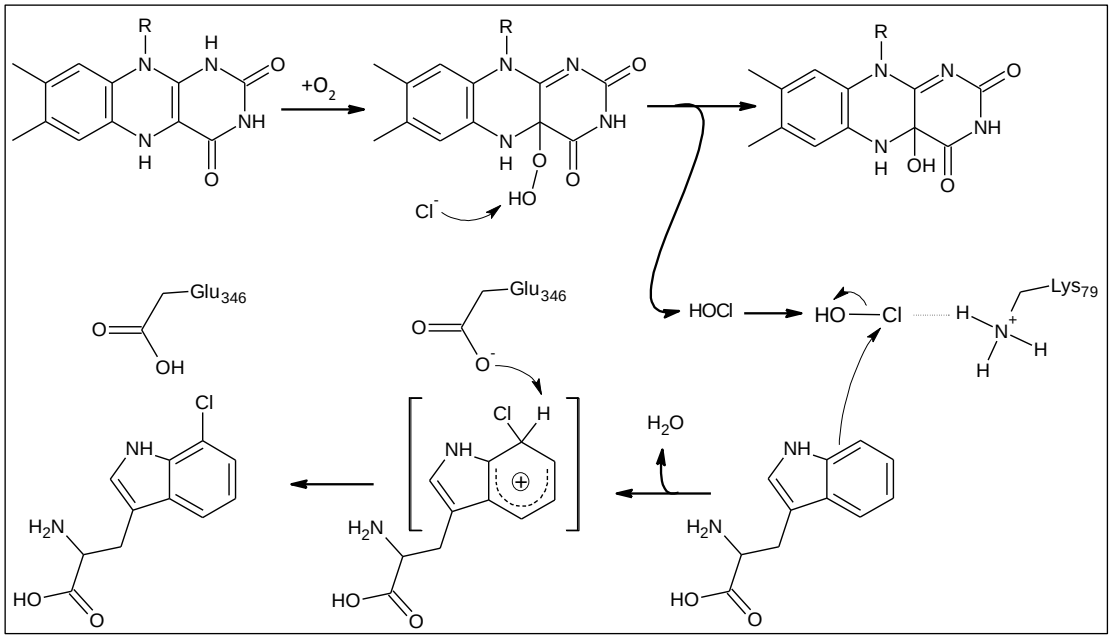

Figure 1.1: Proposed reaction mechanism of tryptophan 7-halogenases (modified based on [30, 32, 33]). $FADH_2$ and molecular oxygen build a FAD(C4a)-OOH intermediate, which is attacked by the chloride ion leading to the formation of hypochlorous acid (HOCl). HOCl reacts with the amine of Lys79 to form a Lys-εNH-Cl intermediate. Finally, aromatic substitution of chlorine at C7 of tryptophan takes place. Following deprotonation, the final product 7-chlorotryptophan is released from the enzyme.

1.2. Molecular engineering of biosynthetic pathways in transgenic plants

Plants produce a huge variety of valuable metabolites, which are for example used as pharmaceuticals like the chemotherapy drug vinblastine from *Catharantus roseus*, the narcotic morphine from *Papaver somniferum* or the antimitotic agent taxol from *Taxus brevifolia* [35-37]. Many of those molecules are secondary metabolites, synthesized to protect the plant against herbivore attacks, UV radiation, etc. or to attract pollinators [38]. The concentration of secondary metabolites within the plant is usually relatively low and, consequently, a considerable amount of tissue is required for extraction of valuable molecules, leading to high production costs. However, isolation of these compounds from plant tissue might be more economical compared to chemical synthesis of molecules with complex structures. Therefore, enzymatic biosynthesis could be a promising alternative to the tedious extraction process or chemical synthesis, leading to regio- and stereospecific modification of metabolites. However, for reconstitution of an entire pathway, all enzymes contributing to the biosynthesis of the desired molecule need to be known. This is often not the case for long and complex biosynthetic pathways of secondary

metabolites. Application of alternative enzymes catalyzing the same reaction can be a key step for successful production of valuable compounds in a heterologous host [14]. Moreover, different organisms might be suitable as heterologous hosts. Due to fast growth, high yields and relatively low costs, bacterial cell cultures are commonly used as expression systems [39]. However, prokaryotic cells can be disadvantageous for production of enzymes that require special compartmentalization. This refers especially to cytochromes P450 often involved in biosynthesis of secondary metabolites [40]. Eukaryotic P450s are localized mainly in the membrane of the endoplasmic reticulum (ER), together with NADPH-cytochrome P450 oxidoreductases that transfer required electrons to the P450s [41]. As a result, biosynthesis of heterologous P450s in a prokaryotic host can be challenging, since co-localization of an NADPH-P450 oxidoreductase as well as truncation or modification of the N-terminus of the P450s might be necessary to preserve enzymatic activity or increase stability [42-44]. Therefore, an eukaryotic expression system can be more suitable for biosynthesis of plant-derived metabolites. Yeast represents a popular eukaryotic host for heterologous production of enzymes and secondary metabolites in synthetic biology [45]. In contrast to bacterial expression systems, the biosynthesis of recombinant enzymes in yeast benefits from glycosylation and cellular compartmentalization in the ER membrane or mitochondria [45, 46]. Moreover, in plants recombinant proteins can be targeted to chloroplasts or the apoplastic space, leading to high accumulation of the desired enzymes [47, 48]. Finally, plants are particularly suitable for production of pharmaceuticals in contrast to mammalian cell cultures, as they harbor neither animal pathogens nor viral or ontogenetic DNA, which is an important safety issue [49, 50].

Biosynthesis of secondary metabolites by recombinant enzymes in plants can be achieved through stable or transient transformation of the host organism. The transient transformation of tobacco is a well-established method facilitating the production of recombinant enzymes within a few days *in planta* [51]. For this purpose, transgenes are integrated into the genome of mature leaves through *Agrobacterium*-mediated gene transfer [52]. Therefore, DNA sequences are flanked by 25 bp imperfect repeats (right and left border) and transported by involvement of Vir proteins into the nucleus, where they are integrated by illegitimate recombination during DNA repair processes [52]. To facilitate constitutive production of heterologous enzymes and desired metabolites, generation of transgenic plants is an efficient approach. Two well-established techniques for nuclear transformation are biolistic bombardment and the *Agrobacterium*-mediated gene transfer [52, 53]. Thereby, transgenes are

translocated into the nucleus and randomly integrated into the plant genome by illegitimate recombination [52, 53]. A special application of biolistic bombardment is the chloroplast transformation. Hereby, transgenes are flanked by homologous sequences and integrated into the chloroplast genome by homologous recombination [54]. Due to the high quantity of chloroplasts within one cell (up to 120 in *Arabidopsis thaliana* mesophyll cells [55]), recombinant enzymes accumulate in high amounts within the plant. Moreover, multiple copies of the chloroplast genome are present in one organelle, each comprising the transgene, which facilitates massive accumulation of recombinant proteins in homoplasmic plants [47]. Additionally, the chloroplast genome is maternally inherited, which eliminates the risk of transgene propagation through pollen [56]. Due to the tedious and complex regeneration process of homoplasmic plants, biosynthetic pathways analyzed in course of the undertaken studies were reconstituted by transient and stable nuclear transformation of *Nicotiana sp.*

1.2.1. Biosynthesis of indigoid molecules

The blue pigment indigo and its violet derivative tyrian purple (6,6'-dibromoindigo) have been used as natural dyes for thousands of years, representing wealth and power [57, 58]. In contrast to other natural dyes, these molecules are extremely stable, which contributes to their popularity and continuous use over several centuries [30]. The indigo precursors, isatan B and indican, are synthesized from indole in indigoferous plants, such as *Indigofera tinctoria*, *Baphicacanthus cusia* and *Polygonum tinctorium* [59, 60]. The physiological function of these secondary metabolites as well as their biosynthesis are still not completely understood. In case of leaf damage, precursor molecules are enzymatically hydrolyzed and the reactive aglycone, indoxyl, dimerizes in the presence of molecular oxygen to form the blue pigment, indigo [58]. For the cloth dying process, the hydrophobic molecule is reduced to its water-soluble *leuco* form featuring a yellowish color [57]. The characteristic dark blue coloration develops during the subsequent drying procedure [57]. The first chemical synthesis of indigo was established by Adolf von Baeyer in 1878; it was further optimized for industrial use and finally commercialized by BASF in 1897 [43]. Indican can be synthesized in four steps from indoxylic acid methylester and acetobromoglucose, as described by Robertson *et al.* [61]. Even though chemical synthesis of indigo is prioritized over extraction of plant-derived indican for industrial production [57, 58], the method requires high temperatures, protection of the sugar moiety and purification from byproducts [61].

Another naturally occurring indigoid, 6,6'-dibromoindigo also known as tyrian purple, is found in some species of sea snails of Thaisidae and Muricidae families, like *Nucella lapillus* [62]. This pigment was used for thousands of years to dye cloth to dark violet [43]. Due to its rare presence in molluscs and the tremendous extraction effort, it was a symbol for wealth and power [42]. In contrast to the plant-derived blue indigo, tyrian purple is synthesized from tyrindoxyl sulfate, which was found in *Dicathais orbita* [63]. It is assumed that this metabolite is activated by an arylsulphatase and molecular oxygen resulting in formation of the reactive intermediates tyrindolinone, tyriverdin and 6-bromoisantin [62, 64]. Finally, photolytic cleavage of tyriverdin yields the purple pigment, 6,6'-dibromoindigo [65, 66]. It was postulated that the initial substitution of bromide within precursor molecules occurs through tryptophan bromination by a bromoperoxidase, which was detected in hyprobranchial extracts of *Hexaplex trunculus* [67]. Moreover, co-localization of the bromoperoxidase, tryptophan and tyrindoxyl sulfate was demonstrated histologically in hypobranchial glands, gonoducts and eggs of *Dicathais orbita* [64]. Even though many intermediates have been suggested to contribute to tyrian purple biosynthesis in marine molluscs, final evidence of this proposed reaction mechanisms needs to be provided [62, 64]. Moreover, the initial biosynthesis steps including the brominated substrate leading to the tyrindoxyl sulfate precursor are not yet uncovered [62, 64].

Besides chemical synthesis, formation of indigo was also described for a prokaryotic system by Gillam *et al.* in *E. coli* cell cultures producing human cytochrome P450s 2A6 and 2E1 [75]. Both P450s oxidized indole, which derived from tryptophan degradation catalyzed by *E. coli* enzymes [75]. Thus synthesized, indoxyl dimerized to indigo and resulted in blue coloration of bacterial cell cultures [75]. Furthermore, plant-based biosynthesis of indoxyl was established by Warzecha *et al.* in *N. tabacum* cell cultures and plants by co-expression of two genes, *bx1* and *2A6*, encoding for an indole synthase and a human cytochrome P450, respectively [14] (fig. 1.2). In maize BX1 catalyzes the first enzymatic step of the biosynthesis of the fungicide and insecticide 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA), and represents a homologue of the α-subunit of tryptophan synthase (TS) [68]. TS forms hetero-tetramers of two α- and two β-subunits and converts indole-3-glycerol phosphate to produce tryptophan [69]. Thereby, indole is formed as an intermediate by the α-subunit and is immediately converted to tryptophan by the β-subunit [69]. Due to this metabolic channeling, indole is never released from the enzyme [69]. In contrast, BX1 functions independently and synthesizes indole from

indole-3-glycerolphosphate in an efficient manner, exhibiting a K_m value of 0.013 mM and k_{cat} of 2.8 s^{-1} [68]. In transgenic tobacco, indole is subsequently oxidized by the human cytochrome P450 2A6 to form the instable indoxyl [14]. In humans, 2A6 is mainly localized in liver and kidney cells and metabolizes several pharmaceuticals, such as the anesthetics methoxyflurane and halothane [70, 71]. In addition, it predominantly contributes to oxidation of S-(-)-nicotine resulting in the formation of S-(-)-nicotine $\Delta^{1'\text{-}5'}$-iminium ion, which is subsequently converted to cotinine [72]. At the same time, nicotine has been shown to efficiently inhibit 2A6 (K_i value of 4.4 ± 0.6 µM) [73]. However, enzymatic activity of the recombinant 2A6 in transgenic *N. tabacum* was sufficient for indole oxidation and indican accumulation [14]. Moreover, modification of 2A6 substrate specificity was achieved by exchange of two amino acids, leading to a double mutant 2A6 L240C/N297Q [74]. It was then demonstrated that the altered cytochrome P450 converted a diverse variety of substituted indole derivatives including chloro- and bromoindoles [75]. In a plant-based indoxyl production system, this unstable metabolite is immediately glycosylated by plant glycosyltransferases and, presumably, stored in the vacuole [14, 76]. Therefore, blue pigmentation of transgenic tobacco resulting from indoxyl dimerization was not described for intact cells [14].

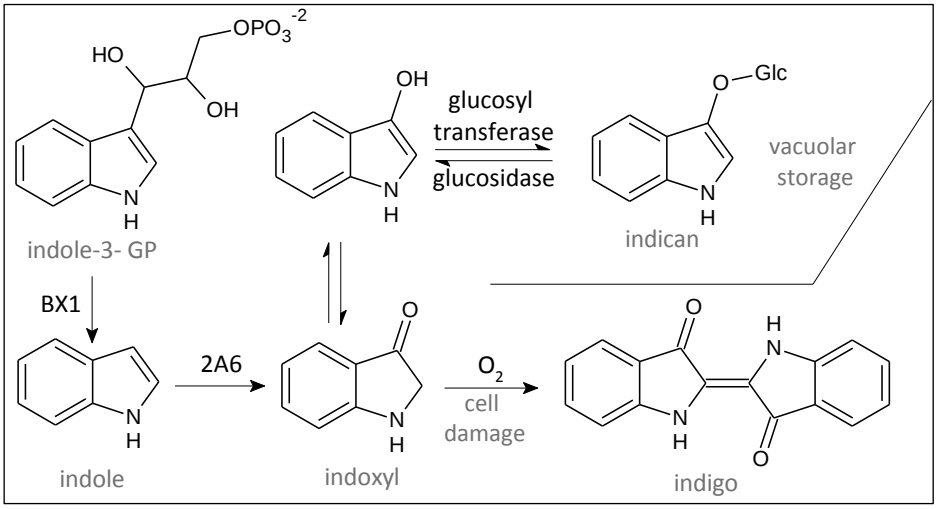

Figure 1.2: Biosynthesis of indoxyl in transgenic tobacco (modified based on [14]). Indole-3-glycerol phosphate (indole-3-GP) is converted by the indole synthase BX1 originating from maize. Subsequently, indole is oxidized by the human cytochrome P450 2A6. Glucosylation by endogenous tobacco glucosyltransferases leads to the biosynthesis of indican, which is stored in the vacuole. After cell disruption indican is deglucosylated and the resulting indoxyl dimerizes, in the presence of oxygen to the blue pigment, indigo.

Further modification of this pathway, resulting in generation of halogenated indole and indoxyl, can be achieved through introduction of tryptophan halogenases. Thereby, a broad variety of regioselectively chlorinated and brominated indigo derivatives could be synthesized using different halogenases *in planta*. This pathway modification requires the exchange of the recombinant 2A6 to the optimized 2A6 L240C/N297Q mutant, which was shown to oxidize halogenated indole derivatives [74, 75].

Halogenated indigoids find application in molecular biology and biochemical assays [77-79]. The most prominent example of those fine chemicals is the widely used 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside acid also known as X-Gal [80]. This indoxyl derivative is substituted with both chlorine and bromide and is covalently bound to galactose [80]. It is applied as substrate for the bacterial β-galactosidase, encoded by the reporter gene, *lacZ,* which is used for histological assays, blue-white screening of transformed bacteria, etc. [77, 80, 81]. After hydrolyzation of galactose by the β-galactosidase, the unstable aglycone, 5-bromo-4-chloro-indoxyl dimerizes in presence of molecular oxygen to a blue indigo derivative [82]. The actual market value of X-Gal is about 422 €/g (Sigma-Aldrich, USA) [83]. This relatively high price is further surpassed by the costs of rare indigoid derivatives like 6-chloroindican (Salmon Glu), with its current sales price of 153 € per 25 mg (Sigma-Aldrich, USA) [84].

1.2.2. Strictosidine biosynthetic pathway

The monoterpenoid indole alkaloid (MIA) precursor, strictosidine is synthesized in different plant species of Apocynaceae and Rubiaceae in diverse tissues, for example in leaves of *C. roseus* or roots of *Rauvolfia serpentina* [85]. Strictosidine is the precursor of more than 3000 known MIAs, some of which are used for medical treatment, e.g. the chemotherapeutics, vinblastine and vincristine from *C. roseus* or the antiarrhythmic agent, ajmaline from *R. serpentina* [85-87]. These secondary metabolites accumulate in minor amounts in the host plants (0.0002 % fresh weight, vincristine and vinblastine), which consequently results in immense costs of their isolation and purification [12]. Given that chemical synthesis of these complex molecules is challenging or even impossible, enzymatic biosynthesis seems to be a promising technique for industrial production. Therefore, all enzymatic steps need to be identified to reconstitute the entire pathway in transgenic plants. Even though enzymes contributing to biosynthesis of pharmaceutically important MIA metabolites like ajmaline and vinblastine are not completely discovered yet, all enzymes involved in the biosynthesis of their precursor, strictosidine have been identified within the last 26 years [12]. The metabolic route in question includes eleven enzymes of

which four are cytochrome P450s (fig. 1.3) [12]. Moreover, the pathway is localized in two different cell types and is further subdivided on a subcellular level [12].

The initial step of strictosidine biosynthesis is performed by geraniol synthase (GES) that converts geranyl diphosphate to the monoterpene, geraniol [88]. Even though formation of geraniol was known long before to be the rate-limiting step in indole alkaloid accumulation in hairy roots and cell cultures [89, 90], identification of the gene sequence as well as characterization of GES from *C. roseus* was not described until 2012 [88]. Biosynthesis of geraniol is localized in chloroplasts of internal phloem-associated parenchyma (IPAP) cells in *C. roseus* leaves [88]. Additionally, precursors of geranyl diphosphate are synthesized by co-localized enzymes of the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway [91]. Subsequently, the monoterpene is oxidized to 8-hydroxygeraniol and further to 8-oxogeraniol by geraniol 8-oxidase (G8O) [92]. This cytochrome P450 (CYP76B6) was initially designated as geraniol 10-hydroxylase (G10H) [93]. Additionally, 8-hydroxygeraniol is converted by 8-hydroxy-geraniol oxidoreductase, the next enzyme in the strictosidine biosynthetic pathway [12, 94]. It catalyzes several reversible steps to synthesize 8-oxogeranial from 8-hydroxygeraniol, forming 8-oxogeraniol or 8-hydroxygeranial as intermediates [12, 94]. Moreover, 8-HGO efficiently converts primary alcohols, such as geraniol, nerol and farnesol [12]. Krithika *et al.* recently characterized 10-hydroxygeraniol dehydrogenase from *C. roseus* (Cr10HGO), catalyzing the same reversible hydroxylation of 8-hydroxy-geraniol to 8-oxogeraniol and further, to 8-oxogeranial [95]. This enzyme shares only low sequence identity with 8-HGO though catalyzing the same reaction [12]. Furthermore, primary alcohols, such as geraniol and nerol were shown to be poor substrates for Cr10HGO [95]. Subsequently, 8-oxogeranial is cyclized by iridoid synthase (IS), probably either by Diels-Alder cycloaddition or by Michael addition [96]. This cyclization results in the formation of the open dialdehydes, *cis*- iridodial and *trans*-iridodial, as well as the bicyclic *cis*-*trans*-nepetalactol, which freely convert from the open dialdehyde to the bicyclic conversion and *vice versa* [96]. This manner of iridoid biosynthesis differs from all known reaction mechanisms of terpene cyclases that initiate cyclization of propenyl diphosphates, like geranyl diphosphate, through a cationic intermediate [97, 98]. C*is*-*trans* nepetalactol is further oxidized to 7-deoxyloganetic acid by iridoid oxidase (IO), a cytochrome P450 (CYP76A26) [12, 99]. Subsequently, 7-deoxyloganetic acid is glycosylated by the glucosyltransferase, 7-DLGT (CrUGT8), which is also localized in IPAP cells of *C. roseus* leaves [12, 100]. Thus synthesized 7-deoxy-loganic acid is oxidized to loganic acid by the cytochrome P450, 7-DLH

(CYP72A224) [12, 101]. Subsequent methylation of loganic acid by methyltransferase, LAMT takes place in the epidermal cells of *C. roseus* and results in loganin biosynthesis [102]. This change of enzyme localization implies the transport of loganic acid, synthesized in IPAP cells, to the epidermis, where the last steps of strictosidine biosynthesis take place [102]. However, the transport mechanisms are not completely understood and transporters of the iridoid intermediates have not yet been characterized. Subsequently, oxidation of loganin is catalyzed by secologanin synthase (SLS), the last cytochrome P450 of the pathway (CYP72A1) [103]. The final step of strictosidine biosynthesis, the stereospecific condensation of secologanin and tryptamine to 3α-(*S*)-strictosidine, was discovered more than two decades ago [104, 105]. It is catalyzed by strictosidine synthase (STR) in a Pictet-Spengler–type reaction, by condensation of the secoiridoid glucoside and the amine function of tryptamine. The latter is synthesized in the cytosol of epidermal cells through decarboxylation of tryptophan by tryptophan decarboxylase (TDC) [106]. Moreover, STR is localized in the vacuole, which implies the transport of tryptamine and secologanin over the tonoplast membrane by an unknown mechanism [107, 108]. De Waal *et al.* identified six isoforms of STR encoded by a single-copy gene in *C. roseus* [109]. In addition, STR possesses an N-terminal vacuolar signal peptide and is glycosylated [109]. However, both factors are not essential for enzyme activity [109]. Finally, determination of *R. serpentina* STR crystal structure revealed a six-bladed β-propeller fold [110].

Regulation of strictosidine biosynthesis on a transcriptional level is not completely understood. The transcription factor ORCA3, first described in *C. roseus* cell suspension cultures, increases gene expression of *tdc*, *str* and the cytochrome P450-reductase gene, *cpr* [111]. Accumulation of this transcription factor itself is inducible by the plant stress phytohormone, methyl jasmonate, which is synthesized as a response to herbivore attacks and leads to accumulation of alkaloids in *C. roseus* [112-114].

Stepwise transient expression of the strictosidine biosynthetic pathway genes resulted in the complete conversion from geraniol to strictosidine [12]. In this regard, biosynthesis of 7-deoxyloganic acid from geranyl diphosphate was proved in the first experimental step through expression of the initial pathway genes, from *gpps* to *7-dlgt* [12]. In the second approach, strictosidine biosynthesis was analyzed after supplementation of iridodial, iridotrial or 7-deoxyloganic acid to tobacco leaves transformed with the last seven genes, starting from *io* up to *str* [12]. Moreover, the production of strictosidine in

Saccharomyces cerevisiae was accomplished by Brown *et al.* in 2015 [115]. However, the heterologous pathway reconstitution resulted in poor yield of the final product depending on co-expression of seven additional genes and three deletions of endogenous genes [115]. Interestingly, strictosidine synthase, localized within the vacuoles of plant cells, was active in the yeast cytosol, offering scope for metabolic engineering of this pathway by alternative enzyme localization [115]. Given that biosynthesis of strictosidine is inefficient in other eukaryotic hosts, whereas transient transformation of *N. benthamiana* led to accumulation of strictosidine and pathway intermediates, a plant-based expression system seems to be the most promising model for reconstitution and modification of the pathway of interest [12, 115].

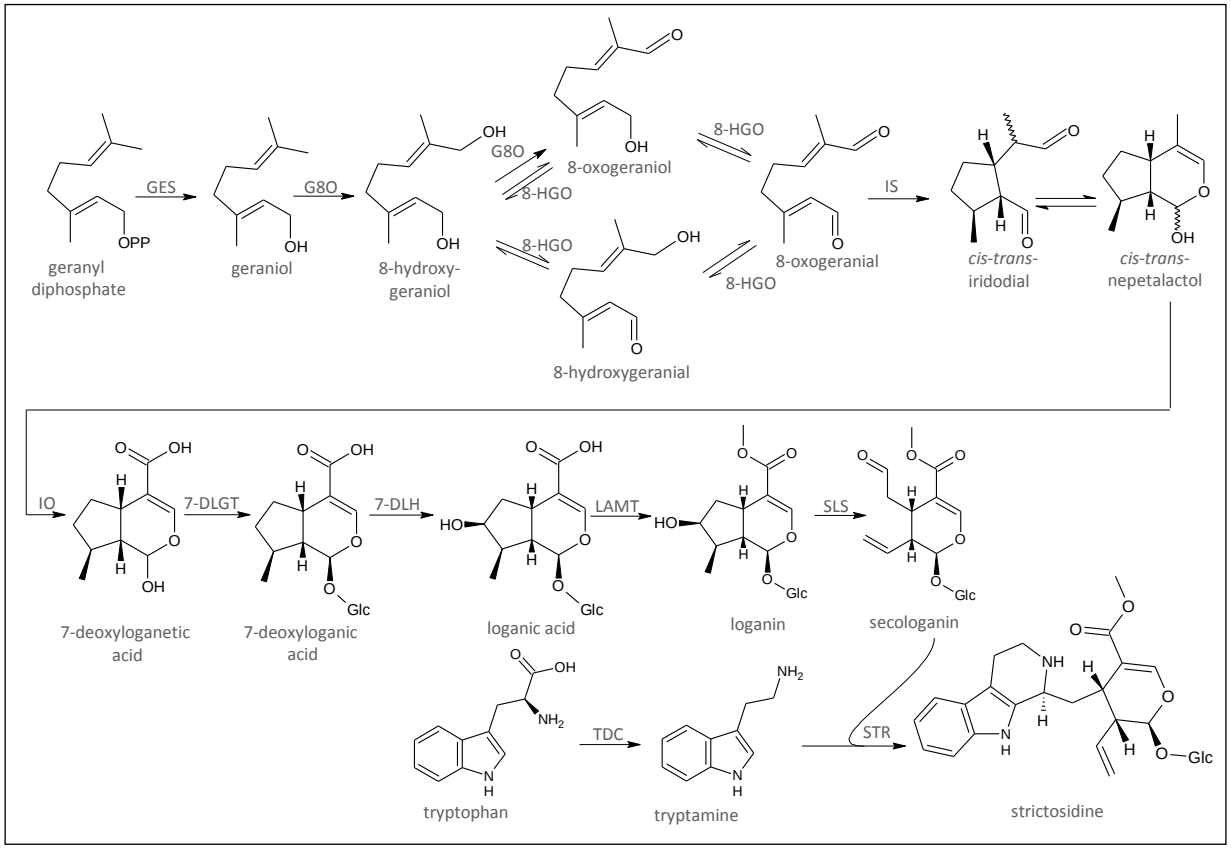

Figure 1.3: The proposed strictosidine biosynthetic pathway in *C. roseus* (modified based on [12]). Strictosidine biosynthesis is catalyzed by eleven enzymes in different cell types of *C. roseus* leaves. The initial steps, from geraniol to loganic acid, are localized in internal phloem-associated parenchyma cells (IPAP), whereas the latter reactions take place in epidermal cells. GES: geraniol synthase, G8O: geraniol 8-oxidase, 8-HGO: 8-hydroxygeraniol oxidoreductase, IS: iridoid synthase, IO: iridoid oxidase, 7-DLGT: 7-deoxyloganetic acid glucosyltransferase, 7-DLH: 7-deoxyloganic acid hydroxylase, LAMT: loganic acid O-methyltransferase, SLS: secologanin synthase, TDC: tryptophan decarboxylase, STR: strictosidine synthase

Following biosynthesis in the vacuole, strictosidine is transported to the nucleus and deglucosylated by strictosidine β-D-glucosidase (SGD) [107]. The resulting reactive aglycone is the progenitor of more than 3000 known MIAs [86, 87, 116]. Stavrinides *et al.* recently characterized tetrahydroalstonine synthase (THAS) converting a stereoisomer of the strictosidine aglycone to tetrahydroalstonine, a precursor of ajmalicine in *C. roseus* [86]. In addition, bimolecular fluorescence complementation assays suggested a C-terminal interaction of SGD and THAS within the nucleus, leading to metabolic channeling of the unstable aglycone [86]. Besides the precursor function of strictosidine for monoterpene indole alkaloid biosynthesis, it was proposed that its reactive aglycone might protect the plant against herbivore attacks. In this regard, Guirimand *et al.* demonstrated protein precipitation after deglucosylation of strictosidine and formation of an unstable dialdehyde [107]. However, final evidence for insecticidal effect through protein precipitation in insects was not provided in this study [107].

Through further engineering of strictosidine synthase, several mutants characterized by an expanded substrate scope were developed [117]. Replacement of Val214 by Met resulted in conversion of 5-, and 6-chlorotryptamine, while the substrates were not processed by the wild type enzyme. However, the catalytic capacity of the mutant for tryptamine conversion was reduced 7-fold compared to STR [117]. Nevertheless, this strictosidine synthase mutant was used by Runguphan *et al.* to integrate halogen atoms into MIAs of *C. roseus* [11]. Co-localization of STRV214M with either tryptophan 5-halogenase, PyrH or tryptophan 7-halogenase, RebH in transgenic *C. roseus* hairy roots resulted in the formation of halogenated indole alkaloids in yields up to 26 µg per gram of fresh weight of plant tissue [11].

1.3. Modular assembly of biosynthetic pathways through application of the GoldenBraid cloning system

Construction of biosynthetic pathways in alternative hosts is a promising strategy for production of rare natural metabolites of great value [12, 118]. Moreover, final product yields can be increased by elimination of enzymatic bottlenecks or constitutive gene expression [17]. However, reconstitution of entire biosynthetic pathways is a challenging task on a molecular level. In this regard, conventional 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 targeting or fusion protein sequences is a tedious process. One of the limitations of the traditional approach is the introduction of restriction sites,

which are necessary to enable fusion of DNA parts into a cloning vector and remain in the final DNA construct after ligation. Therefore, introduction of additional parts needs to be accomplished using other restriction enzymes. This issue can be circumvented by application of particular type IIS restriction enzymes, that cut DNA outside of their recognition sequence leaving 4-nt overhangs that can be designed individually [119]. This principle was first fully applied by the group of Dr. Sylvestre Marillonnet from ICON Genetics in 2008 who developed the innovative Golden Gate cloning technology [119]. This technique is based on *Bsa*I recognition sites, which are inversely oriented, flanking genes of interest as well as the entry sites of the vectors, and are eliminated upon ligation of DNA parts. Moreover, in the initially proposed protocol, digestion and ligation were performed in one reaction step at 50°C and resulted in a successful assembly of the desired plasmids after no more than 5 min. That way, up to three inserts originating from different entry vectors could be ligated to one expression vector. However, Golden Gate cloning was adapted for introduction of single gene sequences into expression vectors, whereas assembly of several genes as transcriptional units in one plasmid was not possible [119]. Three years later, the same group introduced an optimized version of Golden Gate cloning called Modular Cloning system (MoClo), which allows assembly of several TUs in a modular and automated manner [120]. Therefore, all parts, including promoters, 5'-UTRs, targeting sequences, CDSs and terminators, are initially ligated into level-0 destination plasmids using a type IIS restriction enzyme. Each part is flanked by different 4 bp fusion sites with regard to their final position in the transcriptional unit. In a consecutive step, these single parts are assembled to TUs into level-1 destination vectors. Further, assembly of multigene constructs is achieved by fusion of up to six TUs from different level-1 plasmids into one level-2 destination vector. To make this system as flexible as possible, seven level-1 and level-2 plasmids were designed for directional cloning of various numbers of parts and constructs. Therefore, MoClo was a first step towards modular and standardized cloning, but the use of 14 different destination vectors and three different type IIS restriction enzymes, made routine application of this method rather complex [120]. Therefore, an improved modular cloning system called GoldenBraid (GB) was developed by the group of Dr. Diego Orzáez in 2011 [20]. The main benefit of this technique is the introduction of a double loop cloning design that allows easy and (theoretically) infinite assembly of multiple TUs. Destination vectors were reduced to two levels, α and Ω, each including two different kinds of plasmids. For assembly into a destination vector, inserts of two plasmids from the same level can be combined in that of the other. For example, two transcriptional units ligated into α1 and α2 plasmids, respectively, are assembled to a multigene construct within the Ω1 destination vector. Subsequently, two multigene

constructs each ligated into an Ω1 or Ω2 plasmid are assembled in the α1 destination vector. This infinite cloning procedure is based on the use of two different type IIS restriction enzymes, *Bsm*BI and *Bsa*I. For counter-selection against plasmids of the alternative level, destination vectors confer different antibiotic resistances. Furthermore, assembly of GoldenBraid constructs is achieved by alternating restriction and ligation in a one-pot-one-step reaction. Each GoldenBraid plasmid contains a *lacZ*-cassette for blue-white screening of transformed bacterial clones. Moreover, the GoldenBraid system is compatible with the MoClo technology (on level-0), which facilitates a shift to the new system. Each part used for assembly of TUs like promoters, CDSs, signal peptides, fusion protein sequences and terminators, is first ligated into an entry plasmid (pE). To ensure stereospecific cloning, each part is flanked by specific 4 bp fusion sites according to its function within the TU. This modular and standardized system facilitates the exchange, reuse and continuous extension of domesticated parts by each laboratory using the method. Moreover, the GoldenBraid system is designed for plant synthetic biology. Therefore, α and Ω destination vectors are based on Ti-plasmids used for *A. tumefaciens*-mediated gene transfer. In this regard, GB1 destination vectors originate from the pGreen II binary plasmid, which can replicate in *Agrobacterium* only in the presence of the pSoup plasmid, providing the required replication functions *in trans* [20, 121]. A new, updated version of the GB plasmids was established by Sarrion-Perdigones and Vazquez-Vilar *et al.* in 2013 by introduction of the GoldenBraid 2.0 cloning system [21]. The new α and Ω destination vectors originate from the open source pCAMBIA binary vectors [122, 123]. These plasmids are double the size of pGreen-based vectors, but replicate independently in *Agrobacterium* and therefore, simplify the transformation process [122, 123]. In addition, a new entry plasmid, the universal domesticator, pUPD, was designed for incorporation of parts into the GB 2.0 system [21]. Finally, the GB grammar was extended for a broad variety of TUs. New fusion sites were designated for incorporation of signal peptide sequences (sp), regulatory elements, like 5'- and 3'-UTRs and fusion protein sequences (e.g., genes encoding for fluorescent proteins) [21]. The basic principles of the GB 2.0 cloning system are illustrated in fig. 1.4.

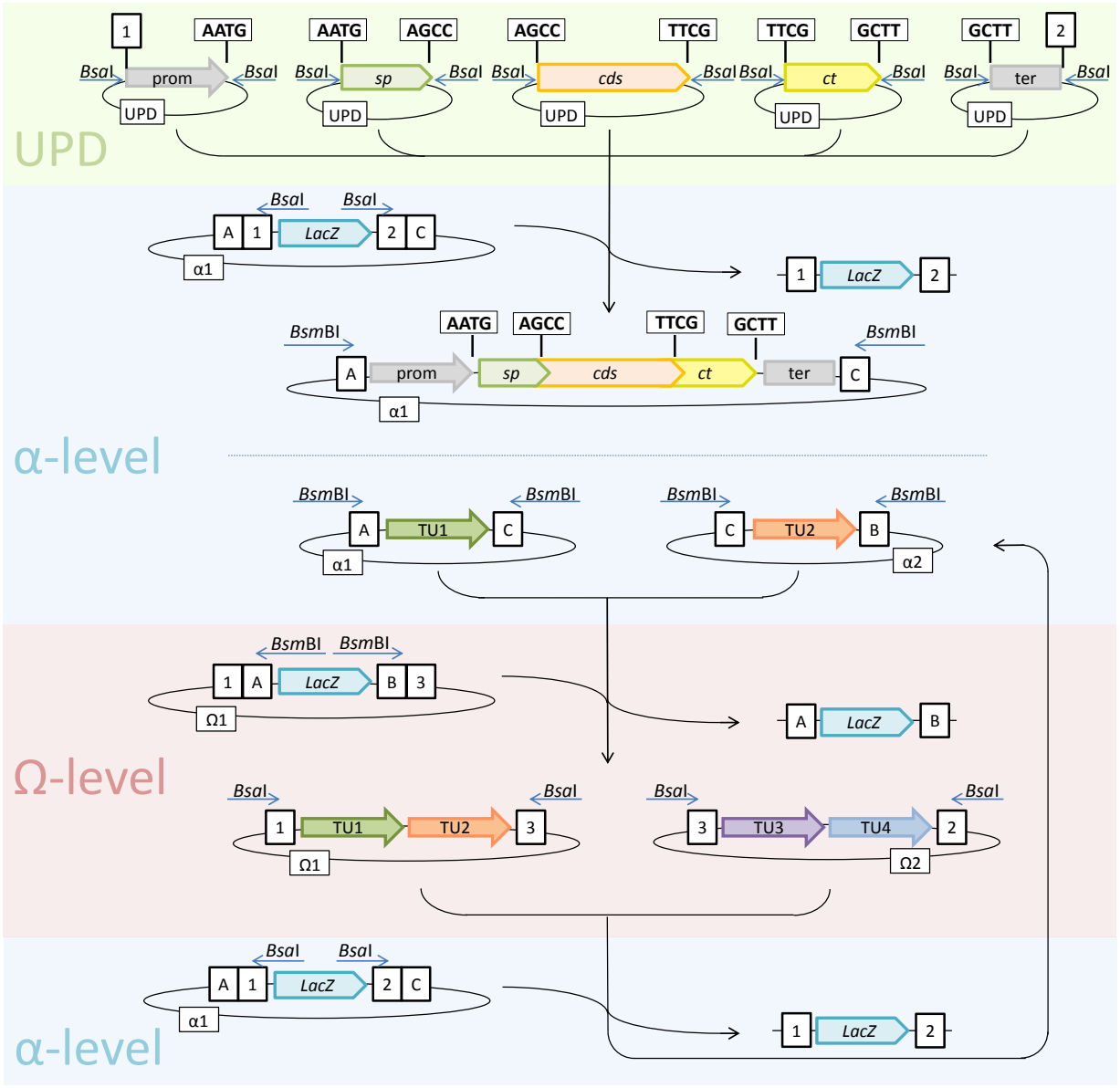

Figure 1.4: Basic principles of the GoldenBraid 2.0 cloning system (modified based on [21]). Each part, e.g., promoters (prom), signal sequences (sp), coding sequences (CDS), C-terminal fusion protein sequences (ct) and terminators (ter), is ligated into the universal domesticator. To enable a correct final orientation, each part is flanked by specific 4 bp fusion sites. Single parts can be assembled to transcriptional units (TU) within α level plasmids by alternating *Bsa*I digestion and ligation in a one-pot-one-step reaction. Further, two TUs of two different α level plasmids are cut by *Bsm*BI and fused to a multigene construct in an Ω level plasmid. Subsequently, multigene constructs of two different Ω level plasmids can be ligated into an α level plasmid. Consequently, the double loop design of GoldenBraid 2.0 facilitates infinite assembly of TUs. All α and Ω level entry plasmids include a *lacZ* cassette which is replaced by the insert. Additionally, 4 bp fusion sites, designated as A, B, C, 1, 2 and 3, are used for assembly of TUs and multigene constructs in α and Ω level plasmids.

1.4. Metabolite channeling by anchoring recombinant enzymes within a protein scaffold

Biosynthesis of metabolites in a heterologous host can be a challenging task due to flux imbalances by virtue of competing pathways or accumulation of toxic intermediates [16, 115, 119, 124]. Besides regulation of gene expression by specific promoter strength or enzyme optimization in terms of turnover number through directed evolution [123, 124], metabolite channeling provides another promising tool for the improvement of metabolite flux, as demonstrated by Dueber *et al.* [19, 125, 126]. This method is based on fusion of recombinant enzymes to ligand peptides [19]. Further, the corresponding ligand binding domains are embedded within a protein scaffold, which enables the attachment of recombinant enzymes and, thereby, improves the conversion of synthesized intermediates. The first binding domain adapted for scaffold assembly was generated from a GTP binding domain from the actin polymerization switch,N-WASP (GBD). Additionally, the Src homology 3 domain (SH3) from the adaptor protein CRK as well as the PSD95/DlgA/Zo-1 domain (PDZ) from the adaptor protein syntrophin were fused to build the final scaffold. The principle of metabolic flux improvement was investigated in *E. coli* by co-localization of three enzymes of the mevalonate pathway. Several protein scaffolds differing in the final numbers of binding domains were analyzed. The most efficient combination resulted in a remarkable 77-fold increase in mevalonate production compared to the biosynthesis by non-scaffolded enzymes. However, biosynthetic rate of the final product depended on the number of each binding domain molecules as well as levels of recombinant enzymes and the protein scaffold. The factors might differ for each biosynthetic pathway and need to be optimized individually [19].

1.5. Aim of the study

A huge variety of halogenated metabolites, whether found in nature or chemically synthesized, find application as pharmaceuticals, like the antibiotic vancomycin, or as fine chemicals such as, X-Gal used in molecular biology or histochemistry [2, 77, 79]. Due to their great value, much research effort has been undertaken in characterization of enzymes that catalyze halogen substitution of arenes in nature [5]. A representative class of those enzymes are flavin-dependent tryptophan halogenases from bacteria that catalyze chlorination and bromination of tryptophan in a highly regiospecific manner [6, 26, 31]. Such precise halogenation is hardly achievable by chemical synthesis, since electrophilic aromatic substitution

occurs at the most preferable position for electrophilic substitution, and therefore, products usually need to be purified from mixtures [9]. Therefore, regiospecific halogenation of aromatic molecules catalyzed by flavin-dependent halogenases is a promising tool for inexpensive and efficient biosynthesis of valuable products. Moreover, modification of pharmacologically important molecules, such as monoterpenoid indole alkaloids, by chlorination or bromination might introduce novel functions and enables easy modification of the MIA skeletons by substitution in terms of enhanced bioactivity or pharmacokinetics [5, 10]. Such metabolic engineering can be achieved by modification of precursor molecules, like tryptophan, by flavin-dependent tryptophan halogenases [11]. Due to the enormous potential of flavin-dependent tryptophan halogenases for modification of plant-derived pharmaceuticals as well as biosynthesis of expensive fine chemicals, their application in plant synthetic biology will be investigated in detail:

- Catalytic activity of two flavin-dependent tryptophan halogenases, namely RebH wt and Stth, will be analyzed in different cell compartments of transiently transformed *N. benthamiana* to evaluate their application in a plant-based expression system. Moreover, substrate scope of both halogenases will be investigated in detail. Additionally, the halogenase mutant, RebH Y455W established by Glenn *et al*., halogenating tryptamine instead of tryptophan, will be characterized *in planta* [28].

- Following characterization of RebH wt and Stth in a plant host, their application for molecular engineering of a biosynthetic pathway will be evaluated. In this regard, reconstitution of the indoxyl biosynthetic pathway, established by Warzecha *et al.* [14], and co-localization of flavin-dependent tryptophan halogenases RebH wt, Stth and RebH Y455W could probably lead to formation of halogenated indican derivatives. The potential of economic and ecological biosynthesis *in planta* will be evaluated in stable and transiently transformed *Nicotiana sp.*

- Furthermore, modification of the strictosidine biosynthetic pathway will be undertaken through co-expression of all pathway genes as well as *rebH Y455W* and *rebF* in an alternative host plant. This molecular engineering endeavor will be attempted by stable and transient transformation of *Nicotiana sp.*, potentially leading to biosynthesis of halogenated strictosidine *in planta*.

- Moreover, optimization of metabolic flux by anchoring of recombinant enzymes within a protein scaffold will be investigated *in planta*. Since this method was proven to increase final product yields of a plant-derived pathway in *E. coli*, linkage of particular enzymes to a protein scaffold will be explored in terms of optimization of the metabolic flux within the analyzed biosynthetic pathways [19].

- Transformation of diverse halogenase gene constructs as well as construction of two biosynthetic pathways requires an enormous cloning effort. To facilitate the efficient and flexible assembly, all required genetic parts will be integrated to the GoldenBraid 2.0 cloning system [21]. Additionally, the library of GB standard parts will be extended to meet the requirements of the innovative modular cloning approaches.

2. Materials and methods

2.1. Materials

2.1.1. Devices

ACQUITY UPLC System	Waters (Milford, USA)
ALPHA 1-4 LDplus freez-dryer	MartinChrist (Osterode am Harz, DE)
Analytical balance ATILON ATL-84-I AccuLab	Sartorius (Göttingen, DE)
Autoclave Systec V-65	Systec (Wettenberg, DE)
Axioskop 40	Zeiss (Oberkochen, DE)
Benchtop Orbital Shaker MaxQ 4450	Thermo Fisher Scientific (Waltham, USA)
Centrifuge 5804 R	Eppendorf (Hamburg, DE)
Centrifuge Micro Star 17R	VWR (Radnor, USA)
Digital pH meter Typ 647-1	Knick (Berlin, DE)
Electrophoresis Power Supply PAC 300	BioRad (Hercules, USA)
Gel System PerfectBlue Midi S	VWR (Radnor, USA)
Gene gun Biolistic PDS-1000/He System	BioRad (Hercules, USA)
Incubator B 5050 E	Heraeus (Hanau, DE)
Infinity 1260 HPLC System	Agilent (Santa Clara, USA)
Incubation Shaker	Ecotron Infors (Bottmingen, CH)
Intas UV-System	Intas (Göttingen, DE)
Leica TCS SP5 II spectral confocal laser scanning microscope	Leica Microsystems (Wetzlar,DE)
Magnetic Stirrers RCT Basic	IKA (Staufen, DE)
Mastercycler epgradient S	Eppendorf (Hamburg, DE)
Microwave R-93ST-AA	Sharp (Osaka, J)
Mixer Mill MM 400	Retsch (Haan, DE)
NanoDrop ND - 1000 Spectrophotometer	VWR (Radnor, USA)
Pipetts 10 µL, 100 µL, 200 µL, 1000 µL	Eppendorf (Hamburg, DE)
Regrigerated Vapor Trap RVT100	Savant Instruments, Inc. (Holbrook, USA)
Sonorex ultrasonic bath	Bandelin (DE)
Spectrafuge™ Mini Laboratory Centrifuge	Labnet International, Inc. (Edison, USA)
SpeedVac Plus SC110A	Savant Instruments, Inc. (Holbrook, USA)

SYNAPT G2-Si High Definition Mass Spectrometry	Waters (Milford, USA)
Thermal Cycler T100	BioRad (Hercules, USA)
ThermoMixer Comfort	Eppendorf (Hamburg, DE)
Vortex-Genie 2	Scientific Industries (Bohemia, USA)
Water bath	Köttermann (Uetze/ Hänigsen, DE)

2.1.2. Chemicals, reagents, consumables

2.1.2.1. Chemicals

Acetic acid	Applichem (Darmstadt, DE)
Acetonitrile, HPLC grade	VWR (Radnor, USA)
Agar-agar	BaccaraRose (Alpen, DE)
Agarose	Applichem (Darmstadt, DE)
Benzylaminopurine	DUCHEFA Biochemie B.V. (Haarlem, NL)
5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside acid (X-Gal)	Applichem (Darmstadt, DE)
Calcium chloride	Applichem (Darmstadt, DE)
Cetyltrimethylammonium Bromid (CTAB)	Applichem (Darmstadt, DE)
3,5-Dimethoxy-4-hydroxyacetophenon (Acetosyringon)	Sigma-Aldrich (St. Louis, USA)
EDTA	Applichem (Darmstadt, DE)
Formic acid, HPLC grade	Carl Roth (Karlsruhe, DE)
Glycerol	Applichem (Darmstadt, DE)
Indole-3-acetic acid	Merck (Darmstadt, DE)
Manganese chloride	Applichem (Darmstadt,DE)
2-Mercaptoethanol	Applichem (Darmstadt, DE)
Methanol, HPLC grade	VWR (Radnor, USA)
$MgSO_4$	Applichem (Darmstadt, DE)
MOPS	Applichem (Darmstadt, DE)
Murashige-Skoog salts	DUCHEFA Biochemie B.V. (Haarlem, NL)
Myo-Inositol	Applichem (Darmstadt, DE)
Nancy-520	Sigma-Aldrich (St. Louis, USA)
1-Naphthaleneacetic acid	Sigma-Aldrich (St. Louis, USA)
Nicotinic acid	Applichem (Darmstadt, DE)

Potassium acetate	Applichem (Darmstadt, DE)
Phyto-agar	DUCHEFA Biochemie B.V. (Haarlem, NL)
Pyridoxine	Sigma-Aldrich (St. Louis, USA)
Rubidium chloride	Applichem (Darmstadt, DE)
Saccharose	Südzucker (Mannheim, DE)
Sodium chloride	Applichem (Darmstadt, DE)
Thiamine	Sigma-Aldrich (St. Louis, USA)
Tris	Applichem (Darmstadt, DE)
Trypton	Applichem (Darmstadt, DE)
Yeast extract	Applichem (Darmstadt, DE)

2.1.2.2. Analytical chemicals used as reference compounds in liquid chromatography

Indican, biochemica	AppliChem (Darmstadt, DE)
5-Chloroindole, 98 %	Sigma-Aldrich (St. Louis, USA)
6-Chloroindole, 99 %	Alfa Aesar (Ward Hill, USA)
7-Chloroindole, 98 %	Alfa Aesar (Ward Hill, USA)
Indole	Merck (Darmstadt, DE)
Indole-3-acetic acid–aspartate	Sigma-Aldrich (St. Louis, USA)
Loganin, for HPLC	Carl Roth (Karlsruhe, DE)
Secologanin	Sigma-Aldrich (St. Louis, USA)
Tryptamine	AppliChem (Darmstadt, DE)
Tryptophan	Serva (Heidelberg, DE)

Strictosidine and 5-chlorotryptamine standard metabolites were kindly provided by Prof. Dr. Joachim Stöckigt from Mainz.

11-Hydroxyiridodial glycoside, iridodial glycoside, iridotrial glycoside, loganic acid, loganin and secologanin were kindly provided by Dr. Heiko Rischer from the Technical Research Center of Finland (VTT) for the LC-MS analyses of metabolites of strictosidine biosynthetic pathway (see 2.5.5.7).

2.1.2.3. Antibiotics and herbicides

Transformed bacteria were incubated in LB growth medium supplemented with antibiotics according to the selection marker genes harbored by the applied plasmids [127]. Transgenic *N. tabacum* plants were cultivated in RMOP and MS growth media supplemented with phosphinothricin and/or additional antibiotics [128]. Final concentrations of the applied antibiotics and phosphinothricin are specified in tab. 2.1.

Tab 2.1: Manufacturers and working concentrations of antibiotics and phosphinothricin used for cultivation of bacterial cultures und transgenic tobacco plants.

Name	manufacturer	working concentration [mg/L]
Ampicillin	AppliChem (Darmstadt, DE)	100
Gentamicin	AppliChem (Darmstadt, DE)	15
Kanamycin	AppliChem (Darmstadt, DE)	50
Rifampicin	AppliChem (Darmstadt, DE)	50
Spectinomycin	Duchefa (Haarlem, NL)	50
Cefotaxime	Duchefa (Haarlem, NL)	300
Phosphinothricin	Duchefa (Haarlem, NL)	5

2.1.2.4. Enzymes

*Bsa*I	New England Biolabs GmbH (Ipswich, USA)
*Bsm*BI	Thermo Fisher Scientific Inc. (Waltham, USA)
Fast Digest restriction enzymes	Thermo Fisher Scientific Inc. (Waltham, USA)
T4 DNA ligase	Promega (Madison, USA)
T4 polynucleotide kinase	Thermo Fisher Scientific Inc. (Waltham, USA)

2.1.2.5. Kits

E.Z.N.A Plasmid DNA Mini Kit I	Omega Bio-Tek (Norcross, USA)
PureLink® HiPure Plasmid Midiprep Kit	Thermo Fisher Scientific Inc. (Waltham, USA)
PureLink® Quick Gel Extraction Kit	Thermo Fisher Scientific Inc. (Waltham, USA)

2.1.2.6. Consumables

Gold microcarriers (0.6 µm)	Biorad (Hercules, USA)
Macrocarriers	Biorad (Hercules, USA)
Rupture disks (1100 psi)	Biorad (Hercules, USA)
Bd plastipak tuberculin syringe, 1 mL	Becton Dickinson GmbH (Madrid, ES)
GeneRuler 100 bp DNA ladder	Thermo Fisher Scientific Inc. (Waltham, USA)
GeneRuler 1 kb DNA ladder	Thermo Fisher Scientific Inc. (Waltham, USA)
Microscope cover glasses 0.13 – 0.16 mm	VWR International (Radnor, USA)
Microscope slides	VWR International (Radnor, USA)

2.1.3. Culture media

Culture media were sterilized by autoclaving at 121°C for 20 min. Subsequently, the media were supplemented with appropriate antibiotics.

<u>Bacterial growth media</u>

Luria-Bertani medium [127]:	10 g/L	trypton
	5 g/L	yeast extract
	10 g/L	NaCl
	pH 7.0	
	addition of 15 g/L agar for solid media.	

<u>Plant media</u>

Murashige-Skoog medium [128]:	4.3 g/L	Murashige-Skoog salts
	30 g/L	saccharose
	0.1 g/L	myo-inositol
	1 mL/L	B5 vitamins
	7 g/L	phyto-agar
	pH 5.7	

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	
B5-vitamins:	0.1 g/ mL	myo-inositol
	0.01 g/mL	thiamine
	1 g/mL	nicotinic acid
	1 g/mL	pyridoxine

2.1.4. Bacterial strains

E. coli cells of the following genotypes were used for plasmid amplification:

TOP10 (Thermo Fisher Scientific Inc., Waltham, USA): F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G

DH5α (New England Biolabs, Ipswich, USA): F^- Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rk^-, mk^+) *pho*A *sup*E44 λ^- *thi*-1 *gyr*A96 *rel*A1

Agrobacterium tumefaciens

Agrobacterium tumefaciens strain EHA105 (ICON Genetics, Halle, DE) carrying the pTiBo542DT-DNA helper plasmid was used for the transient transformation of *N. benthamiana* plants [129]. Cells were cultivated in LB medium supplemented with rifampicin and a second antibiotic with regards to the resistance gene of the utilized GoldenBraid plasmid (tab. 2.8). Cells of *A. tumefaciens* strain GV3101 (DSMZ Braunschweig, DE) carrying the pMP90 helper plasmid were used for transient transformation of *N. benthamiana* and stable transformation of *N. tabacum* plants. Cells were cultivated in LB medium supplemented with rifampicin, gentamycin and a third antibiotic with regards to the resistance gene of the applied GoldenBraid plasmid (tab. 2.8).

2.1.5. Plants

Nicotiana tabacum L. cv. Petit Havana wild type plants as well as transgenic *N. tabacum* BX1#12 plants generated by Warzecha *et al.* [14] were used for stable transformation. *N. tabacum* BX1#12 plants carried the *bx1* gene encoding for indole synthase from maize as well as the *bar* resistance gene for selection on phosphinothricin. *N. tabacum* plants were cultivated in a sterile environment at 26°C and under continuous illumination.

N. benthamiana plants were used for transient transformation *via Agrobacterium*-mediated gene transfer. The plants were cultivated in the Otto-Stocker greenhouse (Technische Universität Darmstadt, DE) at 23°C, 60 % humidity and with 16 h of illumination.

N. tabacum L. cv. Petit Havana wt, as well as F1 generation of *N. tabacum* BX1#12, BX1/2A6#2-13 and rB11 plants were used for infiltration of indole derivatives. BX1/2A6 plants generated by Warzecha *et al.* [14] carried the previously described *bx1* gene as well as the *2A6* gene encoding for a human cytochrome P450. Generation of the transgenic tobacco line rB11 was part of the presented research project and is described in more detail in chapter 3.2.6. The plants were cultivated in the Otto-Stocker greenhouse (Technische Universität Darmstadt, DE) at 23°C, 60 % humidity and with 16 h of illumination

2.1.6. Plasmids and genetic material

For assembly of GoldenBraid (GB) transcriptional units, standard parts were provided as domesticated sequences from the group of Dr. Diego Orzáez from the Spanish National Research Council (CSIC) in Valencia, Spain. α and Ω level destination vectors based on pGreen (pDGB1) or pCAMBIA (pDGB2) plasmids were also provided by the aforementioned research group, as was the α2 plasmid, pHUGE for the assembly of extensive multigene constructs. All plasmids used in course of the undertaken studies are listed, according to official designations and GB ID numbers, in tab. 2.2.

Tab. 2.2: Overview of all domesticated sequences and destination vectors provided by the lab of Dr. Diego Orzáez. Official designations and ID_numbers are specified.

designation	GoldenBraid ID_number	designation	GoldenBraid ID_number
pP35S	GB0030	pDGB1_omega1	pDGB1_omega1
pTnos	GB0037	pDGB1_omega1R	pDGB1_omega1R
pPUbq3	GB0272	pDGB1_omega2	pDGB1_omega2
pTAct2	GB0210	pDGB1_omega2R	pDGB1_omega2R
pYFP_CT	GB0024	pDGB2_alpha1	pDGB2alpha1
pbfp_CT	GB0034	pDGB2_alpha1R	pDGB2_alpha1R
pEGB Tnos:NptII:Pnos	GB0226	pDGB2_alpha2	pDGB2_alpha2
BASTA (*bar*)	GB0023	pDGB2_alpha2R	pDGB2_alpha2R
pPnos	GB0072	pDGB2_omega1	pDGB2_omega1
pDGB1_alpha1	pDGB1_alpha1	pDGB2_omega1R	pDGB2_omega1R
pDGB1_alpha1R	pDGB1_alpha1R	pDGB2_omega2	pDGB2_omega2
pDGB1_alpha2	pDGB1_alpha2	pDGB2_omega2R	pDGB2_omega2R
pDGB1_alpha2R	pDGB1_alpha2R	pUPD	pUPD

Besides genes of interest, targeting sequences and special promoter and terminator varieties were introduced to the GB system. Targeting sequences were extended with appropriate GB fusion sites to facilitate assembly with promoter and coding sequences. The vacuolar signal sequence, *erV* was obtained from the endogenous targeting sequence of strictosidine synthase gene, *str* from *Rauvolfia serpentina* [130]. Moreover, the apoplast targeting sequence, *er* was amplified from the pICH17620 plasmid (ICON Genetics, Halle, Saale, DE). This sequence originates from the endogenous targeting of calreticulin from *Nicotiana sp.* Furthermore, the chloroplast targeting sequence, *cp*, originating from the small subunit Rubisco gene, was amplified from the pICH20030 plasmid (ICON Genetics, Halle, Saale). In addition, the benzoxazinless1 gene, *bx1* from *Zea mays* (GeneBank NM_001111749.1) was integrated into the GoldenBraid cloning system omitting the endogenous chloroplast targeting sequence. Molecular cloning of all targeting sequences described above as well as *bx1* was performed by Dr. Agata Staniek.

In addition, the pool of standard parts was extended to simplify assembly of transcriptional units. To facilitate cytosolic enzyme localization, the cauliflower mosaic virus (CaMV) 35S and the nopaline synthase promoter sequences were extended to incorporate the *AGCC* GB fusion site as well as an *ATG*

start codon. P35S_ATG was generated by PCR amplification using primers listed in tab. 2.5. PNos_ATG was synthesized by Integrated DNA Technologies (Coralville, USA). Furthermore, two additional Nos terminators were generated. The TNos_his was modified by replacement of the original 5' GB fusion site with *GCAG* as well as addition of eight histidine codons and a stop codon. Primers used for modification and amplification of TNos_his are listed in tab. 2.5. A second alternative Nos terminator with the 5'-GB fusion site *GCAG* and a stop codon was synthesized by Integrated DNA Technologies (Coralville, USA).

DNA sequences of *rebH wt* (GeneBank BAC15758) and *rebH Y455W* were kindly provided by Prof. Sarah E. O'Connor from the John Innes Centre in Norwich, UK. Furthermore, DNA sequences of strictosidine biosynthetic pathway genes: *ges* (GeneBank JN882024.1), *g8o* (GeneBank KF561461.1), *8hgo* (GeneBank KF302069.1), *is* (Genebank JX974564.1), *io* (Genebank KF302066.1), *7-dlgt* (Genebank KF415118.1), *7-dlh* (GeneBank KF302067.1), *lamt* (GeneBank EU057974.1), *sls* (GeneBank L10081.1), *str* (GeneBank X61932.1) and *tdc* (GeneBank X67662.1) from *Catharanthus roseus* were kindly provided by Prof. Dr. Johan Memelink from the Leiden University, NL.

In addition, some gene sequences were synthesized commercially for integration into the GoldenBraid cloning system. This applies to *stth* (GeneBank HQ844046.1) and *rebF* (GeneBank BAC15756) synthesized by Thermo Fisher Scientific Inc. (Waltham, USA). Furthermore, geranyl diphosphate synthase (GPPS) gene, *Pagpps* from *Picea abies* (GenBank GQ369788.1), the coding sequence of the small subunit of GPPS, *Amgpps* from *Antirrhinum majus* (GenBank AY534686.1) as well as the 1-deoxy-D-xylulose synthase gene, *dxs2* from *Solanum lycopersicum* (GenBank AY687353.1) were synthesized by Integrated DNA Technologies (Coralville, USA), as was the protein scaffold sequence $G_1S_1P_1$ and ligand sequences GBD, SH3 and PDZ [19].

Finally, the pBIN61-P19 plasmid for suppression of gene silencing was kindly provided by Prof. Sir David Baulcombe from the Cambridge University, UK. The P19 protein binds to mircoRNA, normally recruited by the RNA-induced silencing complex (RISC) to degrade complimentary mRNA [131]. The p19 gene originating from the Tomato bushy stunt virus was regulated by the CaMV 35S promoter and terminator.

2.1.7. Primers

All primers and oligonucleotides used in this study were synthesized by Eurofins Genomics (Ebersberg, DE). For integration of different genes into the GB cloning system, CDSs were extended with appropriate fusion sites and *Bsm*BI recognition sites *via* PCR. Furthermore, all internal *Bsm*BI and *Bsa*I recognition sites were eliminated by introduction of silent mutations. All primers used for domestication of *rebH* are listed in tab. 2.3. In addition, integration of all genes of strictosidine biosynthetic pathway in the GB cloning system was achieved using primers listed in tab. 2.4. Moreover, the DNA sequence encoding for the protein scaffold was amplified *via* PCR using primers listed in tab. 2.5. Finally, primers used for screening of transformed *A. tumefaciens* colonies and transgenic *N. tabacum* plants by PCR are listed in tab. 2.6.

Tab. 2.3: Primers used for PCR amplification of *rebH* DNA fragments for domestication through GoldenBraid cloning. *BsmBI* recognition sites are underlined and GB fusion sites are highlighted in bold.

name	primer sequence 5'→3'	GB fusion site
prebH H101d	AGCGTCTCACTCG**AGCC**TCCGGCAAGATTGACAAGATCCTC	AGCC
prebH G101d	AGCGTCTCACTCG**AATG**TCCGGCAAGATTGACAAGATCCTC	AATG
prebH 201	TGCGTCTCTGAGTCTCCGGGTCGAGGTGCCACAT	---
prebH 301	GACGTCTCGACTCAGCCCCTCAACAGGAT	---
prebH 401	TCCGTCTCCATTGTCTCGATCTCGCGGTTG	---
prebH 501	AGCGTCTCACAATGTTCGACGACACGCGCGACTT	---
prebH 601	AGCGTCTCCAGTGTCTCGAGCAGGTTCCGCTGC	---
prebH 701	TGCGTCTCACACTGCCGAGCCTCCACGAGTT	---
prebH M801d	CGCGTCTCACTCG**CTGC**GCGGCCGTGCTGTTGCCT	GCAG
prebH N802d	CGCGTCTCACTCG**AAGC**TTAGCGGCCGTGCTGTTGCCT	GCTT
prebH M803	CGCGTCTCA**CTGC**GCGGCCGTGCTGTTGCCTCAG	GCAG

Tab. 2.4: Primers used for PCR amplification of the genes of the strictosidine biosynthetic pathway for domestication through GoldenBraid cloning. *BsmBI* recognition sites are underlined and GB fusion sites are highlighted in bold.

name	primer sequence 5'→3'	GB fusion site	gene
pGS H101	GCGCCGTCTCACTCG**AGCC**TCTTCATCATCATCATCATCTTCATCC	AGCC	*ges*
pGS M201	GCGCCGTCTCACTCG**CTGC**AAAACAAGGTGTAAAAAACAAAGCTTTTAC	CAGC	*ges*
pG10H101	GCGCCGTCTCGCTCG**AATG**GATTACCTTACCATAATATTAAC	AATG	*g8o*
pG10H201	CTTTCGTCTCCGTAGACTTCTCCATCGTGAGGCTAC	---	*g8o*
pG10H102	GATGCGTCTCTCTACGAAAAGTTTTGAATTCTAATATA	---	*g8o*
pG10H202	GCGCCGTCTCGCTCG**CTGC**AAGGGTGCTTGGTACAGCAC	GCAG	*g8o*
p10HGO101	GCGCCGTCTCGCTCG**AATG**ACCAAGACCAATTCCCC	AATG	*8-hgo*
p10HGO201	GCGCCGTCTCGCTCG**CTGC**GAACTTGATAACAACTTTGACACA	GCAG	*8-hgo*
pMTC101	GCGCCGTCTCGCTCG**AATG**AGTTGGTGGTGGAAGAG	AATG	*is*
pMTC201	GCGCCGTCTCGCTCG**CTGC**AGGAATAAACCTATAATCCCTCAT	GCAG	*is*
pCYP76A26101	GCGCCGTCTCGCTCG**AATG**GCGACCATCACTTTCGA	AATG	*io*
pCYP76A26201	GCGCCGTCTCGCTGTCTCTTCGACCATTCTG	---	*io*
pCYP76A26102	GCGCCGTCTCGACAGACATGGAGAATTTGCCGTATTTGCAAGCAGTTGT AAAGGAGACACTTAGATTG	---	*io*
pCYP76A26202	GCGCCGTCTCGCTCG**CTGC**GATATGAACTCTCTTCTTAGGG	GCAG	*io*
pDLGT101	GCGCCGTCTCGCTCG**AATG**GGTTCTCAAGAAACAAATTTG	AATG	*7-dlgt*
pDLGT201	GCGCCGTCTCGTGATCTCTTGTTACAACTGTAATAC	---	*7-dlgt*
pDLGT102	GCGCCGTCTCGATCAACTCCTGGAATTCTGGT	---	*7-dlgt*
pDLGT202	GCGCCGTCTCGCTCG**CTGC**AATAATCAGTGATTTTATGTAATCAACG	GCAG	*7-dlgt*
p7DLH101	GCGCCGTCTCGCTCG**AATG**GAATTGAACTTCAAATCAATTATTTT	AATG	*7-dlh*
p7DLH201	GCGCCGTCTCGTTGCCATTTTGGCTTCCAATAG	---	*7-dlh*
p7DLH102	GCGCCGTCTCGGCAATGTCTCTAATCTTGAGGCGTTTTTCT	---	*7-dlh*
p7DLH202	GCGCCGTCTCGCTCG**CTGC**GAGTTTGTGCAGAATCAAATGAG	GCAG	*7-dlh*
pLAMT101	GCGCCGTCTCGCTCG**AATG**GTTGCCACAATTGATTCC	AATG	*lamt*
pLAMT201	ACCACGTCTCAGTCCTCAACTGAGGGATGATAT	---	*lamt*
pLAMT102	CCCTCGTCTCAGGACTTGGAAATGGTGATAGA	---	*lamt*
pLAMT202	GCGCCGTCTCGCTCG**CTGC**ATTTCCCTTGCGTTTCAAGAC	GCAG	*lamt*
pSLS101	GCGCCGTCTCGCTCG**AATG**GAGATGGATATGGATACC	AATG	*sls*
pSLS201	GCGCCGTCTCGCTCG**CTGC**GCTCTCAAGCTTCTTGTAGAT	GCAG	*sls*
TDC GB H101d	ATCGTCTCACTCG**AGCC**GGCAGCATTGATTCAACAA	AGCC	*tdc*
TDC GB N201d	TACGTCTCACTCG**AAGC**TCAAGCTTCTTTGAGCAAATCATC	GCTT	*tdc*

TDC GB M201d	TACGTCTCACTCG**CTGC**AGCTTCTTTGAGCAAATCATC	GCAG	*tdc*
GB2str1FWs	GGCGTCTCGCTCG**AGCC**AAAGAGATTTTGATTGAGG	AGCC	*str*
GB2str1RVs	GGCGTCTCGCTCG**AAGC**TTAATGACTTGAAACAAAAGAATTTCC	GCTT	*str*
GB2str1-fRV	GGCGTCTCGCTCG**CTGC**ATGACTTGAAACAAAAGAATTTCC	GCAG	*str*
pStrvm 201	ATCGTCTCTCATTTCTGCCCCACCTGGAAC	---	*strV214M*
pStrvm 301	TGCGTCTCAAATGAGTGCAGATAGCTCCTTTGTT	---	*strV214M*

Tab. 2.5: Primers used for amplification of all other DNA sequences by PCR for domestication through GoldenBraid cloning. *BsmBI* recognition sites are underlined and GB fusion sites are highlighted in bold.

name	primer sequence 5'→3'	GB fusion site
scaffold for	TTAGCCGTCTCACTCG**AGCC**GGATCTACCAAAGCAGATATTGGC	AGCC
scaffold rv new	ATGCCGTCTCACTCG**AAGC**TTAACCTTTGAAATACGGGCTAACTTC	GCTT
P35SATG fw	CGTACGTCTCTCTCG**GGAG**ACTAGAGCCAAGCTGATCTCCTTTGCACTAGAG CCAAGC	GGAG
P35SATG rv	ACTACGTCTCGCTCG**GGCT**GAGGAAGCCATTTCGACTAGAATAGTAAATTGT AATGTTGTTTGTTG	AGCC
TNos GBM101	AGATCGTCTCACTCG**GCAG**GGCATCATCACCATCATCACCATCACTAAGGAA TGGATCTTCGATCCCG	GCAG
TNos GBB201	AGATCGTCTCACTCG**AGCG**CGAGTCGGTCCCATTATTGAAGCATTTATCA	CGCT

Tab. 2.6: Primers used for DNA screening of *A. tumefaciens* colonies and transgenic *N. tabacum* plants.

name	primer sequence 5'→3'	used for
p35S-Cf3	CCACGTCTTCAAAGCAAGTGG	*A. tumefaciens* screening
p35S-Cr4	TCCTCTCCAAATGAAATGAACTTCC	
pG10H 301	GTGGAAGAGATTTCGAGCTGATTCC	*N. tabacum* screening
pG10H 401	AAGGGTGCTTGGTACAGCACGC	
p10HGO301	TCATTCCTCTTCTGTGTGGTAGAACTG	*N. tabacum* screening
p10HGO 401	GAACTTGATAACAACTTTGACACAATCAGG	
pMTC 301	TGGCTGAGTTAATGAAGGATAAAGATCAAG	*N. tabacum* screening
pMTC 401	AGGAATAAACCTATAATCCCTCATCTTATCAATAC	

2.2. Molecular cloning methods

2.2.1. Polymerase chain reaction

Target DNA sequences were amplified using gene-specific primers and *Pfu* polymerase in course of polymerase chain reaction (PCR) [132]. Furthermore, gene sequences were extended by addition of 5'- and 3'-fusion sites compatible with the GB cloning strategy as well as *Bsm*BI recognition sites. The applied PCR program was adjusted for the addition of long overhangs by means of primer annealing at low annealing temperature in five initial amplification cycles. Thus, in the initial phase of PCR, only sequence-specific parts of the primers bound to the complimentary DNA sequences. Subsequent 25 cycles were performed with an increased annealing temperature to ensure highly specific amplification. At this stage, entire primer sequences, including additional overhangs, associate with the template. The PCR mixture was composed of 50-500 ng template DNA, 1x *Pfu*-buffer[1], 0.2 mM of each dNTP, 0.2 mM of each forward and reverse primer as well as 2 µL of (in-house) purified *Pfu*-polymerase. The total reaction volume was 100 µL.

PCR-program:

step	temperature [°C]	duration [min]	
initial denaturation	94	5	
denaturation	94	0.5	
annealing	50-55	0.5	5 cycles
elongation	72	0.5-2	
intermediate elongation	72	10	
denaturation	94	0.5	
annealing	60-67	0.5	25 cycles
elongation	72	0.5-2	
final elongation	72	10	

[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

2.2.2. Agarose gel electrophoresis

DNA fragments generated by PCR or plasmid digestion were separated according to their sizes by electrophoresis in 1-2 % (w/v) agarose gels. Therefore, agarose was melted in 1x TAE buffer[2] and supplemented with 5 µL of the fluorescent stain, Nancy-520 (Sigma-Aldrich, USA) for visualization of DNA fragments under UV-light. DNA samples were mixed with a 6x loading dye[3] prior to electrophoresis. Separation of DNA fragments was performed in 1x TAE buffer at 120 V for 45 - 60 min and fragment patterns were documented at 312 nm with an Intas UV-System (Intas Science Imaging Instruments GmbH, DE).

[2]1x TAE (Tris-acetate): 40 mM Tris, 1.4 % (v/v) acetic acid, 1 mM EDTA, pH 8.2

[3]6x loading dye: 60 % glycerol, 10 mM Tris-HCl, 60 mM EDTA, 0.03 % bromophenol blue, pH 7.6

2.2.3. DNA extraction from agarose gels

For purification of amplified DNA sequences, the target bands were cut out from gels following electrophoresis. Subsequently, the DNA fragments of interest were purified using the PureLink® Quick Gel Extraction Kit (Thermo Fisher Scientific, USA) according to the manufacture's recommendations.

2.2.4. Quantification of DNA

DNA concentration was determined by photometric measurements using the NanoDrop ND-1000-spectrometer (VWR, USA). 1 µL of DNA solution was applied to the instrument and measurements were performed according to the manufacture's recommendations.

2.2.5. Annealing and phosphorylation of oligonucleotides

For the addition of an octuplet histidine-tag (His-tag) to coding sequences, two oligonucleotides were annealed, forming appropriate 4-nt fusion sites, prior to the GB reaction (tab 2.7). Ligation of the annealed His-tag oligonucleotides to coding sequences was performed either during their domestication or TU assembly. For the annealing reaction, 40 pmol of each oligonucleotide were mixed with 2 µL of 10x ligase buffer (Promega, USA) and filled up with water to a final volume of 20 µL. The mixture was incubated for 5 min at 95°C and cooled down at room temperature (RT) for at least 2 h. Subsequently, 2 µL of the annealed oligonucleotide sample were mixed into a 20 µL phosphorylation reaction mixture

consisting of 1x PNK A buffer, 1 mM ATP and 10 u of T4 polynucleotide kinase. The reaction was incubated for 20 min at 37°C followed by enzyme inactivation at 75°C for 10 min. Contrary to the regular protocol, 3 µL of the phosphorylated oligonucleotide preparation were applied to the GoldenBraid reaction mixture. Domestication of endogenous 2A6 targeting sequence was achieved according to the same protocol by annealing of oligonucleotides listed in tab. 2.7. In contrast to the His-tag fusion, the targeting sequence was domesticated and verified by sequencing.

Tab. 2.7: Oligonucleotides used for the fusion of 8x His-tag or 2A6 targeting sequence to genes during GB reaction. GoldenBraid fusion sites are highlighted in bold; 4-nt overhangs generated by annealing of oligonucleotides are underlined.

Name	**Sequence 5'→3'**	**GB overhang 5'→3'**	**application**
His-tag MN30	**GCAG**GGCATCATCACCATCATCACCATCACTAAGCTT	GCAG	domestication
His-tag MN40	CTCG**AAGC**TTAGTGATGGTGATGATGGTGATGATGCC	GCTT	
His-tag MN50	**GCAG**GGCATCATCACCATCATCACCATCACTAA	GCAG	TU assembly
His-tag MN60	**AAGC**TTAGTGATGGTGATGATGGTGATGATGCC	GCTT	
2A6 targ10	CTCG**AATG**CTCGCTTCTGGTATGTTATTAGTGGCTTTAC TGGTTTGCTTGACTGTCATGGTACTAGCC	AATG	domestication of 2A6 targeting
2A6 targ20	CTCG**GGCT**AGTACCATGACAGTCAAGCAAACCAGTAAA GCCACTAATAACATACCAGAAGCGAGCATT	AGCC	

2.2.6. Digestion and ligation of plasmid DNA *via* GoldenBraid 2.0 cloning

The GoldenBraid 2.0 cloning technique facilitates fusion of DNA fragments in a one-step/one-pot reaction by alternating restriction and ligation of plasmid DNA [21]. The method was applied for generation of all plasmids used for transient and stable transformation of tobacco. 10 µL of each GB reaction mixture contained of 75 ng of each expression plasmid harboring DNA parts, 75 ng of destination vector, 1x ligase buffer, 1 mM DTT, 3 u T4-ligase and 10 u *Bsa*I (for cloning into α level plasmids) or 10 u *Bsm*BI (for cloning into the universal domesticator and Ω level plasmids). For the domestication of single parts, the reaction was set up as described above. Instead of expression and destination plasmids, 75 ng of purified PCR product (see 2.2.3) as well as 75 ng of the universal

domesticator plasmid were applied to the mixture. The GoldenBraid assembly was performed in 25 cycles, including 2 min digestion at 37°C followed by a ligation step at 16°C for 5 min. Finally, the complete reaction mixture was applied for transformation of competent *E. coli* cells (see 2.2.8.). Selection marker genes of all GB destination vectors are specified in tab. 2.8.

Tab. 2.8: Selection marker genes of GoldenBraid destination vectors. Bacteria transformed with GoldenBraid constructs were cultivated on growth media supplemented with the appropriate antibiotic.

GoldenBraid plasmid	selection marker gene	antibiotic used for selection
universal domesticator	*bla*	ampicillin
α level	*nptI*	kanamycin
Ω level	*aadA*	spectinomycin

2.2.7. Preparation of competent cells

2.2.7.1. Preparation of chemically competent *E. coli* cells

E. coli cells were transferred from glycerol stocks into 3 mL LB medium without antibiotics and incubated overnight at 37°C and 140 rpm. Moreover, 100 mL LB medium were inoculated with 1 mL of overnight culture and incubated at 37°C and 140 rpm until cells reached an OD_{600} of 0.5. Next, the culture was centrifuged at 3000x *g* for 10 min at 4°C. The pellet was resuspended in 30 mL of ice-cold TfB1 buffer[4] and incubated for 90 min on ice. Subsequently, the cell suspension was centrifuged at 3000x *g* for 10 min at 4°C and gently resuspended in 4 mL of ice-cold TfB2 buffer[5]. Finally, *E. coli* cells were frozen, in 50 µL aliquots, in liquid nitrogen and stored at -80°C.

[4]TfB1 buffer: 100 mM RbCl, 50 mM $MnCl_2$, 30 mM KOAc, 10 mM $CaCl_2$, 15 % (v/v) glycerol, pH 5.8

[5]TfB2 buffer: 10 mM RbCl, 10 mM MOPS, 75 mM $CaCl_2$, 15 % (v/v) glycerol, pH 6.8

2.2.7.2. Preparation of chemically competent *A. tumefaciens* cells

A. tumefaciens cells were transferred from glycerol stocks into 3 mL of LB medium supplemented with rifampicin (EHA105) or rifampicin/gentamycin (GV3101) and incubated overnight at 28°C and 140 rpm. 1 mL of the overnight culture was transferred into 100 mL of LB medium supplemented with the appropriate antibiotics and further incubated at 28°C and 140 rpm until cells reached an OD_{600} of 0.5. Subsequently, the bacteria culture was centrifuged at 4°C for 20 min at 1400x *g* and the cell pellet was resuspended in 10 mL of TE buffer[6] and, once again, centrifuged at 4°C for 10 min and 1400x *g*. Finally, 5 mL of LB medium supplemented with proper antibiotics were used to resuspend the cells and 100 µL aliquots were stored at -80°C.

[6]TE buffer: 10 mM Tris, 1 mM EDTA, pH 8.0

2.2.8. Transformation of competent cells

2.2.8.1. Transformation of competent *E. coli* cells

For the transformation of chemically competent *E. coli* DH5α or TOP10 cells *via* heat shock, a 50 µL aliquot was thawed on ice. Subsequently, 10 µL of the GB reaction (see 2.2.6) or 500 ng of plasmid DNA were added to the cells. The mixture was incubated for 20 min on ice and heat shock was performed for 1 min at 42°C. After incubation for additional 2 min on ice, 500 µL of LB growth medium were added to the *E. coli* cells. Subsequently, the cells were incubated at 37°C and 750 rpm for 1 h and plated out on LB medium supplemented with antibiotics for selection of positive clones. For blue/white screening of cells transformed with GB plasmids, LB plates were supplemented with 40 µg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside acid (X-Gal). Finally, *E. coli* cells were incubated at 37°C overnight.

2.2.8.2. Transformation of competent *A. tumefaciens* cells

For the transformation of chemically competent *A. tumefaciens* cells by heat shock, a 100 µL aliquot was thawed on ice and supplemented with 300 ng - 500 ng of plasmid DNA. The mixture was incubated on ice for 5 min followed by freezing in liquid nitrogen for additional 5 min. Finally, the cells were incubated at 37°C for 5 min. Following the addition of 500 µL of LB liquid medium, the cell suspension was incubated for 4h at 28°C and 750 rpm. Subsequently, the bacteria were transferred on LB plates

supplemented with rifampicin (EHA105) or rifampicin and gentamycin (GV3101) as well as another antibiotic corresponding to the selection marker gene of the applied plasmid (tab 2.8). Finally, the cells were incubated for 48 h at 28°C.

2.2.9. Colony screening of transformed bacteria

For the identification of successfully transformed bacterial cells, colonies were picked from plates and diluted in 20 µL water. 10 µL of cell suspension were added to 10 µL of PCR reaction mixture. The final PCR reaction sample was composed of 1x *Taq*-buffer[7], 0.2 mM $MgCl_2$, 0.2 mM of each dNTP, 0.1 mM of each primer and 0.2 µL of (in-house) purified *Taq*-polymerase. PCR products were separated by agarose gel electrophoresis and the fragment sizes were documented under UV light. Colonies presenting the genes of interest were transferred into 3 mL of LB medium supplemented with appropriate antibiotics (tab 2.8) and incubated overnight at either 37°C (*E. coli*) or 28°C (*A. tumefaciens*) and 140 rpm.

PCR program:

step	temperature [°C]	duration [min]	
initial denaturation	94	5 min	
denaturation	94	0.5	
annealing	55-60	0.5	30 cycles
elongation	72	0.5-2	
final elongation	72	10	

[7]*Taq*-buffer : 10 mM Tris (pH 8.8), 50 mM KCl, 0.08 % (v/v) Triton-X 100

2.2.10. Plasmid DNA preparation

Plasmid DNA preparation was performed with the E.Z.N.A Plasmid DNA Mini Kit I (Omega Bio-Tek, USA) according to the manufacturer's recommendations.

2.2.11. Sequencing of DNA fragments

For verification of either DNA sequences in the universal domesticator or of coding sequences (CDSs) within transcriptional units, 750-1500 ng plasmid DNA were sequenced by Eurofins Genomics using the following primers:

Name	Sequence 5‘ → 3‘	application
M13 uni (-21)	TGTAAAACGACGGCCAGT	Sequencing of domesticated DNA sequences
M13 rev (-29)	CAGGAAACAGCTATGACC	
P35S 3'sense	TACAACAATTACCACAACAAAACAAACAAC	Sequencing of CDSs within transcriptional units
TNos 5'antisense	CTTTATTGCCAAATGTTTGAACGATC	

2.2.12. Control digestion of plasmid DNA

To confirm the correct assembly of GB parts into transcriptional units and multigene constructs, plasmid DNA was digested with restriction enzymes and analyzed by agarose gel electrophoresis (see 2.2.2). Therefore, 200-300 ng of plasmid DNA were mixed with 1x Fast Digest buffer and 3-5 u of an appropriate Fast Digest restriction enzyme. The reaction was incubated for 15 min at 37°C.

2.2.13. Preparation of glycerol stocks

For long-term storage of bacterial cells, 500 µL of relevant cell cultures were briefly mixed with 300 µL of 50 % (v/v) glycerol. Subsequently, the cells were frozen in liquid nitrogen and stored at -80°C.

2.3. Software-based analyses

2.3.1. Prediction of signal peptides by means of TargetP

Signal peptides were identified using the online tool, TargetP 1.1 Server [133]. Cleavage site prediction in plants was calculated in the default mode for all recombinant enzymes.

2.3.2. Ligand docking simulation by means of AutoDock 4.2

Software-based calculation of potential ligand-receptor interactions is a timesaving method, requiring reasonable effort, for evaluation of binding and localization of new substrates within the active sites of given enzymes based on their crystal structures. This principle was applied for docking of indole and indican molecules into the active site of RebH wt using the open source docking software, AutoDock 4.2 [134]. The crystal structure of the RebH-tryptophan complex was obtained from the Protein Data Bank

entry 2E4G. Moreover, 3D structures of the analyzed ligands were obtained as XML files from the PubChem database of the National Center for Biotechnology Information (NCBI). For AutoDock calculations, XML files were transformed into the mol2 format using the open source software, OpenBabel [135]. Additionally, PDBQT files for RebH wt and the ligands were generated by computing Gasteiger charges and assigning AD4 type atoms. Furthermore, the ligands were modified by adding hydrogen atoms. Moreover, ligand torsion angles were kept flexible, whereas amino acid residues of the enzyme were set rigid. To keep the computation time and effort minimal, affinity and electrostatic grids were calculated for a defined grid box within the active site of the enzyme. The box was centered according to the position of the tryptophan ligand (x=3.720; y=34.077; z=-9.128) with 40 numbers of points in each dimension. Each ligand was docked into the active site of RebH using the Lamarckian genetic algorithm with standard parameters in 100 docking runs.

2.4. Generation and characterization of transgenic plants

2.4.1. Sterilization of tobacco seeds

N. tabacum L. cv. Petit Havana wt and transgenic BX1#12 plants were transformed for establishing of transgenic tobacco lines. For this purpose, their seeds were sterilized in 1 mL of 6 % (v/v) sodium hypochlorite for 20 min. Subsequently, the seeds were washed five times in 1 mL of sterile water and placed on MS medium (wt) or on MS medium supplemented with phosphinothricin (BX1#12). Finally, the plants were cultivated in a sterile environment at 26°C and under continuous illumination.

2.4.2. Stable transformation of *N. tabacum*

The establishing of transgenic tobacco lines by stable transformation with recombinant genes provides the opportunity of long-term biosynthesis and accumulation of desired metabolites *in planta*. However, regeneration of transformed tobacco plants is a time-consuming process, in contrast to transient approaches. Nuclear transformation of *N. tabacum* plants was achieved using two different methods: first, biolistic bombardment with DNA-coated gold particles was performed by means of a gene gun (see 2.4.2.1). Additionally, tobacco was transformed *via Agrobacterium*-mediated gene-transfer (see 2.4.2.2). For the selection of successfully transformed cells, tobacco leaves were cultivated on RMOP medium supplemented with proper antibiotics or herbicides. Following the selection process of

calli, emerging shoots were transferred onto MS selection medium to facilitate the recovery of transgenic tobacco lines. After root development, plants were cultivated in potting soil in the greenhouse.

2.4.2.1. Biolistic bombardment

The biolistic bombardment was performed according to Sanford *et al.* [53]. The target DNA was precipitated on gold particles used as carries. The particles were then shot by virtue of helium pressure onto plant leaves in a vacuum chamber. The DNA carried into the cell nucleus was thus randomly integrated into the plant genome by illegitimate recombination during DNA repair processes.

Particle preparation

Preparation of gold particles was performed on ice and all centrifugation steps were done at 4°C. EtOH and H_2O were cooled on ice prior to use. 2 mg of 0.6 µm gold particles were resuspended in 100 µL of EtOH and mixed for 1 min. The mixture was centrifuged for 1 s at 4000x *g* and the supernatant was discarded. Subsequently, the particles were resuspended in 1 mL of H_2O and centrifuged for 1 s at 4000x *g*. The supernatant was discarded and the particles were resuspended in 250 µL of H_2O. 25 µg of plasmid DNA were applied to the particles and mixed immediately. For co-transformation of two plasmids, 12.5 µg of each plasmid DNA were added to the particles. Further, 250 µL of $CaCl_2$ (2.5 M) and 50 µL of spermidine (0.1 M) were added to the particles and mixed immediately after each step. Subsequently, the gold particles were incubated on ice for 10 min and briefly mixed every minute. The suspension was centrifuged for 1 s at 3500x *g* and gently resuspended in 600 µL of EtOH. After centrifugation for 1 s at 3500x *g*, the gold particles were washed for a second time in EtOH and finally resuspended in 65 µL EtOH.

Biolistic bombardment

Biolistic bombardment was performed using the Gene Gun Biolistic PDS-1000/He System (Biorad, USA). All biolistic equipment was sterilized in isopropanol or autoclaved prior to use. Leaves of three to four weeks old *N. tabacum* L. cv. Petit Havana plants grown in sterile environment were used for the transformation process. For each shot, one leaf was placed abaxial side up on a sterile filter paper on an RMOP medium plate. The bombardment was performed in a vacuum of 27 inch Hg using a 1100 psi rapture disc and 7 µL of the gold particle suspension. The macrocarrier with applied gold particles was

placed in the top slot of the instrument, the tobacco leaf in the third one. After bombardment, leaves were incubated for 24 h at 26°C in the dark. For the selection of transgenic cells, the leaves were cut in 4x4 mm pieces and transferred onto RMOP medium supplemented with appropriate antibiotics, abaxial side down. Finally, the leaves were cultivated at 26°C under continuous illumination.

2.4.2.2. *A. tumefaciens*-mediated gene transfer

A. tumefaciens bacteria are used in plant biotechnology as a tool for nuclear transformation of plants. The desired recombinant genes are flanked by 25 bp imperfect repeats (right and left border) and transported through involvement of Vir proteins into the plant nucleus [52]. Finally, the genes of interest are integrated randomly into the plant genome through illegitimate recombination during DNA repair processes.

For stable transformation of tobacco plants, the desired *A. tumefaciens* GV3101 cells were inoculated into 15 mL of LB medium supplemented with antibiotics and incubated overnight at 28°C. The culture was centrifuged for 4 min at 4000x *g* and the pellet was resuspended in 25 mL of LB medium without antibiotics. Three leaves of three to four weeks old *N. tabacum* L. cv. Petit Havana plants, cultivated in a sterile environment, were cut into 4x4 mm pieces and incubated in the bacterial culture for 10 min. For an optimal result, the leaves were continuously mixed during the entire incubation period. Subsequently, the leaf pieces were transferred onto RMOP medium without antibiotics and incubated for 24 h in the dark. For the selection of successfully transformed tobacco cells, the leaf parts were transferred onto ROMP medium supplemented with cefotaxime and appropriate antibiotics for selection of transformed cells. The leaf pieces were further cultivated at 26°C under continuous illumination.

2.4.3. Transient transformation of tobacco *via A. tumefaciens* infiltration

Transient biosynthesis of recombinant enzymes in *N. benthamiana* was achieved by infiltration of transgenic *A. tumefaciens* EHA105 or GV3101 cells. Therefore, 10 mL of an overnight culture, carrying the desired transgenes, were centrifuged at 8000x *g* for 5 min and the cell pellet was resuspended in the ICON buffer[8] to a final OD_{600} of 1. Subsequently, the bacterial suspension was infiltrated into leaves of three to four weeks old tobacco plants using a syringe without a needle. For the combination of different *Agrobacterium* cultures, the amount of each construct was kept constant within one experiment by adding the corresponding amount of cell suspension carrying the pDGB2α2 plasmid. To increase

transformation efficiency, the plants were incubated overnight at RT in the dark. Final incubation was performed at 26°C and 60 % humidity with 12 h of illumination in a climate chamber. Successful biosynthesis of recombinant enzymes was verified three to seven days after transformation by observing YFP or BFP fluorescence (see 2.5.1). As a variation of this protocol, genes of the indoxyl biosynthetic pathway were co-expressed with the *p19* gene encoding for a gene silencing repressor, by co-infiltration of *A. tumefaciens* cells carrying the pBIN61-P19 plasmid (see 3.2.3).

[8]ICON buffer: 10 mM $MgSO_4$, 10 mM MES, 100 µM 3,5-Dimethoxy-4-hydroxyacetophenon (Acetosyringon), pH 5.5

2.4.4. Extraction of genomic DNA from plant tissue

For the identification of transgenic plants, genomic DNA was extracted and screened for integration of transgene by PCR. Therefore, plant leaves were frozen in liquid nitrogen and ground with a mortar and a pestle. 750 µL of extraction buffer[9] and 750 µL of chloroform were applied to 100 mg of the plant powder and the mixture was incubated for 45 min at 65°C. Subsequently, the samples were centrifuged at 7500x *g* for 10 min at 4°C. The upper aqueous phase was transferred into a fresh 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 corresponding 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.

[9]Extraction buffer: 1 % sarcosyl, 0.8 M NaCl, 22 mM EDTA, 0.22 M Tris-HCl, 0.8 % CTAB, 0.14 M mannitol
The extraction buffer was supplemented with 1.4 µL/mL of β-mercaptoethanol prior to use.

2.4.5. Grafting of transgenic tobacco plants

For the removal of nicotine from transgenic *N. tabacum* plants, tobacco scions were grafted onto *Solanum lycopersicum* var. TA234 rootstocks. In this regard, three to four weeks old tobacco and tomato plants with identical stem diameters were selected for the grafting process. Scion and rootstock plants were cut at matching stem diameters at the same time to guarantee identical cut angles. Except for two leaves, all remaining leaves of the tobacco scion as well as all leaves of the tomato rootstock were removed from the plants to avoid fatal loss of water. Finally, scions were attached by clips to the rootstocks and the grafted plants were kept at 26°C in a darkened chamber at high humidity to increase

the grafting efficiency. After one week, the plants were cultivated in the greenhouse at 26°C, 60 % humidity and with 12 h illumination.

2.4.6. Infiltration of metabolites into tobacco leaves

For the enhancement of metabolic flux and for activity assays of transiently synthesized enzymes, metabolites were infiltrated into tobacco leaves. The desired chemicals were dissolved in a H_2O:MeOH solution (19:1 v/v) to a final concentration of 400 µM and infiltrated into tobacco leaves using a syringe without a needle. Control infiltrations were performed with the H_2O:MeOH solution solely. Plant samples were harvested two days after metabolite infiltration. In case of transient expression of biosynthetic pathway genes in *N. benthamiana*, metabolite feeding was performed four days after the initial *A. tumefaciens* infiltration.

2.4.7. Lyophilization of plant material

Leaves of *N. tabacum* PaGPPS-IS lines were freeze-dried to avoid any impact on LC-MS analyses by an increased water content in plants featuring the altered phenotype. Therefore, 2-5 g of fresh leaf material were frozen in liquid nitrogen and ground with a mortar and a pestle. The leaf powder was lyophilized at -80°C in a vacuum of 0.25 mbar for 24-48 h using the ALPHA 1-4 LDplus freeze-dryer (MartinChrist, DE). Finally, the samples were stored at -20°C.

2.5. Analytical methods

2.5.1. Fluorescent microscopy

Successful transformation of infiltrated tobacco plants was verified by virtue of fluorescence of the yellow fluorescent protein (YFP) fused to the recombinant enzymes using the Axioskop 40 microscope (Zeiss, DE). The plant cells were illuminated with UV-light generated by the HBO 50 /AC bulb (Osram, DE) and filtered by the Zeiss filter set 38. The samples were analyzed at 200x magnification using a 20-x objective (Achroplan 20x/0.45, Zeiss, DE).

2.5.2. Confocal laser scanning microscopy (CLSM)

For verification of correct localization of recombinant enzymes, transiently transformed *N. benthamiana* leaves were analyzed by confocal laser scanning microscopy using the Leica TCS SP5 II spectral confocal laser scanning microscope (Leica Microsystems, DE). YFP emission was measured in the range of 510-600 nm after excitation at 514 nm using a 20x oil objective (HCX PL APO CS 20.0 X 0.70 IMM UV, Leica Microsystems, DE).

2.5.3. Thin-layer chromatography

The technique of thin-layer chromatography (TLC) enables separation of metabolites in complex mixtures [136]. An advantage of this analytical method is, in contrast to column chromatography, the possibility of visualization of metabolite coloration. Furthermore, additional chemical treatments like deglycosylation, can be easily implemented. The separation of metabolites is based on different interactions of molecules with the solid and mobile phase. A typical polar solid phase consists of a silica gel bound either to a glass, aluminum or plastic plate. The liquid phase, or eluent, is composed of organic solvents and, depending on the chemical properties of the separated metabolites, water in different ratios. For the separation process, samples are dissolved in organic solvents and applied to the plate. While the eluent is drawn up to the upper edge of the plate by virtue of capillary action, metabolites are separated by virtue of divergent interaction strengths with the organic eluent and the polar silica gel. These interactions can be characterized by the retention factor (R_f) value, which describes the ratio of migration distance of single metabolites to the migration distance of the eluent front.

Sample preparation

For investigation of the biosynthesis of indole, indican and their halogenated derivatives by thin-layer chromatography, 200 mg leaf tissue were frozen in liquid nitrogen and ground using a pestle and a mortar. The samples were supplemented with 1 mL of acetone and 5 µL of acetic acid and mixed for 15 min at RT. Following centrifugation at full speed for 10 min, the supernatant was concentrated at 42°C using the Speed Vac Plus SC110A instrument (Savant Instruments Inc., USA). Finally, the dried extracts were resuspended in 50 mL of MeOH and stored at -20°C.

Sample separation by thin-layer chromatography

Plant extracts were separated on silica aluminum TLC sheets pre-coated with a fluorescent indicator UV_{254}. 20 µL of plant extracts or 1-5 µg of standard metabolites were applied in 1 cm lines to the plates. For separation of highly lipophilic metabolites, like indole and its derivatives, the mobile phase was composed of chloroform and ethyl acetate in ratio of 4:1 (v/v). More hydrophilic molecules, such as indican and its derivatives, were separated using a mobile phase consisting of chloroform, methanol and water in a ratio of 70:28:2 (v/v/v). After metabolite separation, the thin-layer plates were digitally documented under UV light. Molecules masking the UV fluorescent dye on the plates were detected at 254 nm, whereas the inherent fluorescent of metabolites was documented at 366 nm. For visualization of indican on the plates, indoxyl dimerization was initiated by deglucosylation of indican. Therefore, the plates were sprayed with a 5 % HCl solution (v/v in EtOH). After a 15 min incubation at 100°C, the plates were photographed for documentation and R_f values of metabolites of interest were calculated according to formula (1).

(1) $$R_f = \frac{distance\ metabolite\ [mm]}{distance\ solvent\ front\ [mm]}$$

2.5.4. Preparation of plant extracts for liquid chromatography

2.5.4.1. Methanol extraction

For the extraction of metabolites from plant samples, leaf tissue was frozen in liquid nitrogen and ground with a pestle and a mortar. Subsequently, 100 mg of fresh leaf powder or 10 mg of lyophilized leaf tissue 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 full speed for 10 min at 4°C and stored at -20°C.

2.5.4.2. Acidic methanol extraction

The extraction of metabolites deriving from the strictosidine biosynthetic pathway was performed according to [12] in the research group of Dr. Heiko Rischer at the Technical Research Center of Finland. For the purification of molecules, 200 mg of leaf tissue were frozen in liquid nitrogen and ground using a pestle and a mortar. Subsequently, the samples were homogenized in 600 µL of extraction solution[10] for

1 min with three grinding balls using the Mixer Mill 400 (Retsch, DE). The samples were purified two times by centrifugation at full speed for 10 min at 4°C. Finally, the samples were stored at -20°C.

[10]Extraction solution: 0.133 % formic acid (v/v) in MeOH

2.5.5. Reversed-phase liquid chromatography and mass spectrometry

Liquid chromatography enables the separation of complex chemical mixtures of almost all kinds based on diverse interactions of single metabolites with the stationary phase [137]. In case of reversed phase chromatography, separation columns are packed with porous silica particles with attached carbon chains, creating a stationary phase with a substantial nonpolar surface. The mobile phase consists of an aqueous and an organic solvent. In high-performance (HP) and ultra-performance liquid chromatography (UPLC), the mobile phase is transported by high-capacity pumps enabling a time-saving and highly efficient sample separation at high pressure [137]. Finally, metabolites eluting from the stationary phase are detected by UV-absorption measurement or mass spectrometry.

In mass spectrometry, metabolites are separated according to their mass-to-charge ratio (m/z) [138]. This method is commonly used downstream of gas and liquid chromatography applications in an on-line process. Metabolites eluting from LC columns can be vaporized and ionized by electrospray ionization (ESI). In case of quadrupole mass analyzers, the resulting ions are subsequently accelerated and separated in an oscillating electric field *in vacuo* according to their mass-to-charge ratio, determined by deflection of their trajectory. Subsequently, the ions can be detected by an electron multiplier, which, at the same time, increases their electric signal. Alternatively, m/z values of metabolites can be determined according to their velocity in the electric field using a time-of-flight instruments (TOF). Here, ions can be detected using an analog-to-digital converter (ADC) that translates ion signals into digital values.

Tandem mass spectrometry (MS/MS) constitutes a special application of the technique. It allows further characterization of the structure and functional groups of single metabolites by fragmentation using triple quadrupole mass spectrometers [139]. Initially, molecules are separated like in a conventional MS approach according to their mass-to-charge ratios by the first quadrupole. Subsequently, metabolites with distinct m/z values are selected for fragmentation by collision with inert gas particles in the collision

cell. Finally, the resulting fragments are further separated according to their m/z values by another quadrupole [139].

2.5.5.1. Characterization of flavin-dependent tryptophan halogenases by high-performance liquid chromatography

For the separation of halogenated tryptamine and tryptophan derivatives, plant extracts were analyzed by reversed-phase HPLC by means of the 1260 Infinity instrument (Agilent, USA). Metabolites were separated on the C18 ZORBAX 300SB column (Agilent, USA) using a nonlinear mobile phase gradient (tab. 2.9). Subsequently, the eluting molecules were detected by a diode array detector (DAD) measuring UV-absorption spectra instead of distinct wavelengths. The following parameters and instrument settings were used for the analyses:

HPLC parameters

column:	ZORBAX 300SB, C18, 4.6x150 mm, 5 µm
aqueous phase:	0.1 % (v/v) formic acid in water
organic phase:	acetonitrile
flow rate:	1 mL/min
column temperature:	30°C
injection volume:	10 µL of plant extracts; 1 µL of reference metabolites (0.1 mg/mL)
spectrum range:	200-400 nm
signal detection:	238 nm, 280 nm

Table 2.9: Composition of aqueous and organic phase for sample separation on the C18 ZORBAX 300SB column used for the characterization of flavin-dependent tryptophan halogenases.

time [min]	0.1 % formic acid	acetonitrile	
0.0 – 4.0	90 %	10 %	isocratic
4.0 – 9.0	70 %	30 %	gradient
9.0 – 12.0	70 %	30 %	isocratic
12.0 – 13.0	0 %	100 %	gradient
13.0 – 16.0	0 %	100 %	isocratic
16.0 – 17.0	90 %	10 %	gradient
17.0 – 20.0	90 %	10 %	isocratic

2.5.5.2. Analyses of indole and indican derivatives by high-performance liquid chromatography

Analyses of indole, indican and their derivatives in plant extracts were performed by reversed-phase HPLC by means of the 1260 Infinity instrument (Agilent, USA). Metabolites were separated on the C14 ZORBAX Bonus RP column (Agilent, USA) using a nonlinear mobile-phase gradient (tab. 2.10) and were subsequently detected by a diode array detector (DAD). The following parameters and instrument settings were used for the analyses:

HPLC parameters

column:	ZORBAX Bonus-RP, C14, 4.6x150 mm, 5 µm
aqueous phase:	0.1 % (v/v) formic acid in water
organic phase:	methanol
flow rate:	1 mL/min
column temperature:	30°C
injection volume:	10 µL of plant extracts; 1 µL of reference metabolites (0.1 mg/mL)
spectrum range:	200-400 nm
signal detection:	280 nm

Tab. 2.10: Composition of aqueous and organic phase for sample separation on the Bonus RP C14 column for the analyses of indole and indican derivatives.

time [min]	0.1 % formic acid	methanol	
0.0 – 3.0	70 %	30 %	isocratic
3.0 – 5.0	30 %	70 %	gradient
5.0 – 10.0	30 %	70 %	isocratic
10.0 – 11.0	0 %	100 %	gradient
11.0 – 15.0	0 %	100 %	isocratic
15.0 – 16.0	70 %	30 %	gradient
16.0 – 20.0	70 %	30 %	isocratic

2.5.5.3. Repetition of HPLC runs on a UPLC-MS system for the characterization of flavin-dependent halogenases

For the analyses of metabolites synthesized by Stth and RebH wt, the samples were separated on the C18 ZORBAX 300SB HPLC column coupled to the Acquity UPLC system (Waters, USA) using the HPLC gradient (see 2.5.5.1). Eluting metabolites were first detected with a photodiode array (PDA) detector (Waters, USA) and subsequently analyzed by mass spectrometry. Retention time values of the investigated metabolites were compared to those obtained in previous HPLC runs and m/z values were analyzed by the Micromass Quattro Premier triple quadrupole mass spectrometer (Waters, USA). MS/MS measurements were performed for several metabolites using the same LC conditions. All analyses were performed in cooperation with Dr. Markus Krischke at the Institute for Pharmaceutical Biology of the Julius Maximilian University of Würzburg. The following parameters and instrument settings were used for LC-MS analyses:

UPLC parameters

column:	ZORBAX 300SB, C18, 4.6x150 mm, 5 µm
flow rate:	1 mL/min
aqueous solvent:	0.1 % (v/v) formic acid in water
organic solvent:	acetonitrile
column temperature:	30°C
injection volume:	10 µL of plant extracts and reference metabolites (0.1 mg/mL)
spectrum range:	200-400 nm

MS parameters

electrospray ionization mode:	positive (ESI+) and negative (ESI-)
capillary voltage:	2.75 kV (ESI+), 3.00 kV (ESI-)
cone voltage:	25 V
cone gas flow:	10 L/h
source temperature:	120°C
desolvation gas:	nitrogen
desolvation temperature:	400°C
desolvation gas flow:	800 L/h

mass range:	m/z 50-1000

MS/MS parameters

Daughter ions of (m/z):	273, 291, 325
electrospray ionization mode:	positive
capillary voltage:	2.75 kV
cone voltage:	25 V
cone gas flow:	50 L/h
source temperature:	120°C
collision energy:	20 eV
collision gas:	argon
collision gas flow rate	0.3 mL/min
dwell time:	0.5 s
mass range:	m/z 30-500

2.5.5.4. Characterization of flavin-dependent tryptophan halogenases by UPLC-MS

After verification of synthesized metabolites by replication of HPLC runs on the UPLC-MS system (see 2.5.5.3), large-scale screening of compounds was performed using the C18 Acquity UPLC BEH column (Waters, USA) and a nonlinear mobile-phase gradient (tab 2.11) with the same UPLC-MS system. The following parameters and instrument settings were used for LC-MS analyses:

UPLC parameters

column:	Acquity UPLC BEH, C18, 2.1x100 mm, 1.7 µm
flow rate:	0.25 mL/min
aqueous solvent:	0.1% (v/v) formic acid in water
organic solvent:	acetonitrile
column temperature:	30°C
injection volume:	5 µL of plant extracts and reference metabolites (0.1 mg/mL)
spectrum range:	200-400 nm

Tab. 2.11: Composition of aqueous and organic phase for sample separation on the C18 BEH UPLC column for the analyses of tryptophan, tryptamine and their derivatives.

time [min]	0.1 % formic acid	acetonitrile	
0.0 – 2.0	90 %	10 %	isocratic
2.0 – 4.5	70 %	30 %	gradient
4.5 – 6.0	70 %	30 %	isocratic
6.0 – 6.5	0 %	100 %	gradient
6.5 – 8.0	0 %	100 %	isocratic
8.0 – 8.01	90 %	10 %	gradient
8.01 – 10.0	99 %	1 %	isocratic

MS parameters

MS parameters were set as described in chapter 2.5.5.3.

2.5.5.5. Analyses of tryptophan bromination and biosynthesis of chloroindican

To analyze the potential of tryptophan halogenases to brominate tryptophan *in planta*, the leaf extracts were analyzed by means of the reversed-phase C18 Acquity UPLC BEH column and a nonlinear mobile-phase gradient (tab 2.12) upon the same LC-MS instrument (see 2.5.5.3). Additionally, biosynthesis of chloroindican through modification of the indoxyl biosynthetic pathway was verified using this method and the following instrument settings.

UPLC parameters

column:	Acquity UPLC BEH, C18, 2.1x50 mm, 1.7 µm
flow rate:	0.25 mL/min
aqueous solvent:	0.1% (v/v) formic acid in water
organic solvent:	methanol
column temperature:	40°C
injection volume:	5 µL of plant extracts and reference metabolites (0.1 mg/mL)

Tab. 2.12: Composition of aqueous and organic phase for sample separation on the C18 BEH UPLC column for the analyses of tryptophan, tryptamine and their derivatives.

time [min]	0.1 % formic acid	methanol	
0.0	90 %	10 %	initial
0.0 - 2.5	70 %	30 %	gradient
2.5 - 5.0	70 %	30 %	isocratic
5.01 - 7.5	0 %	100 %	isocratic
7.51 - 10.0	90 %	10 %	isocratic

MS parameters

MS parameters were set as described in chapter 2.5.5.3.

2.5.5.6. Analyses of metabolites of strictosidine biosynthetic pathway by UPLC-MS

The analyses of metabolites of strictosidine pathway from transgenic *N. tabacum* PaGPPS-IS plants as well as transiently transformed *N. benthamiana* plants were performed upon the UPLC-MS system (see 2.5.5.3), using the C18 BEH column. Metabolites were separated in a nonlinear mobile phase gradient according to [100] (tab. 2.13).

UPLC parameters

column:	Aquity UPLC BEH, C18, 2.1x100 mm, 1.7 µm
flow rate:	0.3 mL/min
aqueous solvent:	0.1 % (v/v) formic acid in water
organic solvent:	acetonitrile
column temperature:	30°C
injection volume:	5 µL of plant extracts and reference metabolites (0.1 mg/mL)
spectrum range:	200-400 nm

Tab. 2.13: Composition of aqueous and organic phase during sample separation on the C18 BEH UPLC column for analyses of metabolites of strictosidine biosynthetic pathway.

time [min]	0.1 % formic acid	acetonitrile	
0.0 -0.5	99 %	1 %	isocratic
0.5-5.0	92 %	8 %	gradient
5.0-6.5	70 %	30 %	gradient
6.5-7.2	50 %	50 %	gradient
7.2-7.5	70 %	30 %	gradient
7.5-8.0	92 %	8 %	gradient
8.0-8.5	99 %	1 %	gradient
8.5-11	99 %	1 %	isocratic

MS parameters

MS parameters were set as described in 2.5.5.3, with the exception of the capillary voltage of 3.0 kV in positive ionization mode.

2.5.5.7. Analyses of metabolites of strictosidine biosynthetic pathway with a high definition mass spectrometer

Metabolites of strictosidine pathway extracted from transiently transformed *N. benthamiana* plants were analyzed using the Acquity UPLC system connected to the SYNAPT G2-Si high definition mass spectrometer (Waters, USA). UPLC parameters were set as described before (see 2.5.5.6). The measurements were performed in the research group of Dr. Heiko Rischer under supervision of Dr. Tuulikki Seppänen-Laakso at the Technical Research Center of Finland (VTT) in Espoo. The following MS setting were used for the analyses.

MS parameters

mass range:	m/z 100-1000
electrospray ionization mode:	negative
capillary voltage:	1.0 kV
cone voltage:	35.0 V
cone gas flow:	50 L/h
source temperature:	150°C

desolvation gas:	nitrogen
desolvation temperature:	500°C
desolvation gas flow :	1000 L/h

3. Results and discussion

3.1. Functional characterization of three flavin-dependent tryptophan halogenases

3.1.1. Molecular cloning of DNA encoding flavin-dependent tryptophan halogenases

For functional characterization of flavin-dependent tryptophan halogenases, RebH wt, Stth and RebH Y455W, all coding sequences were incorporated into the GoldenBraid (GB) cloning system. Therefore, gene sequences of *rebH wt* and *rebH Y455W* were used as templates for adding appropriate fusion sites and *Bsm*BI recognition sites by PCR (see 2.2.1). Additionally, internal *Bsm*BI and *Bsa*I recognition sites were removed by introduction of silent mutations to avoid erroneous or ineffective assembly during the cloning procedure. All primers used for gene amplification are listed in tab 2.3. Following purification, the PCR products were ligated into the universal domesticator, pUPD *via* a GB reaction (see 2.2.3, 2.2.6). For different approaches, 3'-ends of the halogenase genes were either fused to a *yfp* CDS during TU assembly or to an 8x His-tag. In this regard, fusion of *rebH wt* and *rebH Y455W* to the His-tag sequence was achieved by adding annealed and phosphorylated His-tag oligonucleotides, pGB2_His-tag MN 30 and 40 (tab. 2.7), during the domestication reaction (see 2.2.5). The resulting GB plasmids harboring the genes of interest were amplified in *E. coli* and verified by sequencing (see 2.2.11). Furthermore, the CDS of the tryptophan 6-halogenase, Stth and the flavin reductase, RebF were synthesized for direct fusion in transcriptional units. Therefore, *rebH wt*, *rebH Y455W* and *stth* were ligated to a Cauliflower mosaic virus 35S promoter (P35S) and a nopaline synthase terminator (TNos). For enzyme targeting to the desired cell compartments, the 5'-end of each CDS was fused to a chloroplast (*cp*) or an apoplast (*er*) targeting sequence. To facilitate cytosolic localization of enzymes, 5'-ends of encoding CDSs were directly fused to the P35S_ATG promoter. Moreover, the 3'end of *stth* was fused to previously annealed and phosphorylated His-tag oligonucleotides, pGB2_His-tag MN 50 and 60 (tab. 2.7) during TU assembly. For verification of cellular localization and accumulation of recombinant enzymes, the 3'-end of each halogenase CDS was alternatively fused to a yellow fluorescent protein gene (*yfp*). Finally, *rebF* was ligated to the ubiquitin3 promoter (PAtUBq3), the actin terminator (Tactin) and either *cp* or *er* targeting sequence. For cytosolic localization of the reductase, *rebF* was fused to the CaMV P35S_ATG promoter and a nopaline synthase terminator (TNos). Furthermore, all reductase TUs contained a blue fluorescent protein gene (*bfp*), which was fused to the 3'-end of *rebF*.

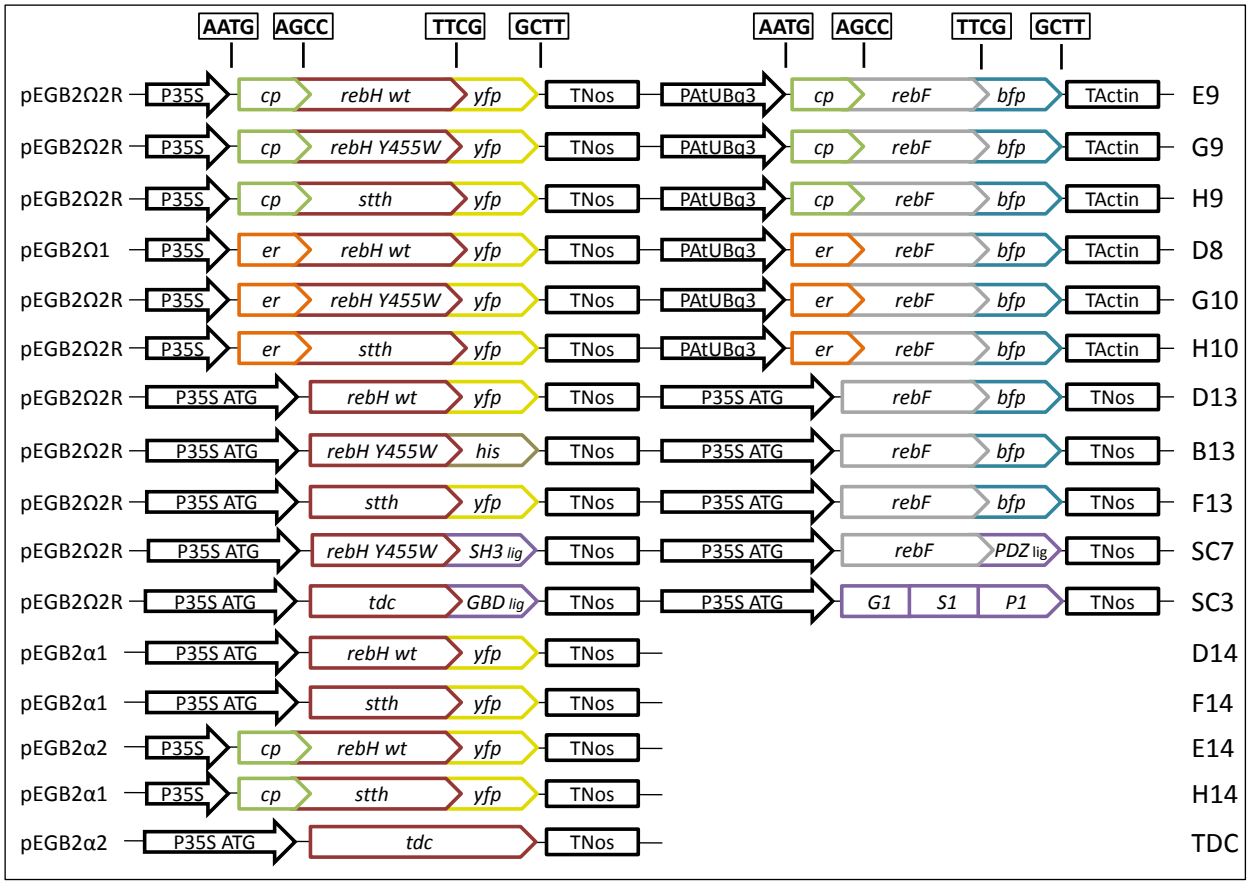

Figure 3.1: Schematic diagram of generated DNA constructs for the characterization of tryptophan halogenases RebH wt, RebH Y455W and Stth. Resistance genes of each GoldenBraid backbone can be found in chapter 2.2.6. Abbreviation of each construct is listed on the right.

For the assembly of multigene constructs, each halogenase TU was fused to the *rebF* TU harboring an identical targeting sequence in Ω level plasmids. The multigene constructs of RebH wt, RebH Y455W and Stth harboring the chloroplast targeting sequence were termed E9, G9 and H9, respectively (fig. 3.1). Furthermore, halogenase/reductase plasmids of RebH wt, RebH Y455W and Stth with incorporated apoplast targeting sequence were termed D8, G10 and H10, respectively (fig. 3.1). Finally, DNA constructs designed for cytosolic localization of RebF and either RebH wt, RebH Y455W or Stth were designated D13, B13 and F13 (fig. 3.1). Additionally, α level plasmids of RebH wt and Stth constructs fused to the *cp* signal sequence were termed E14 and H14, while constructs without a targeting sequence were designated D14 and F14 (fig. 3.1).

To facilitate biosynthesis of tryptamine, the tryptophan decarboxylase gene, *tdc*, was introduced into the GB system. Therefore, appropriate fusion sites and *Bsm*BI recognition sites were incorporated into the gene sequence *via* PCR using primers listed in tab 2.4. Following domestication, *tdc* was fused to the CaMV 35S_ATG promoter and Nos terminator into an α level plasmid (fig. 3.1, TDC).

For optimization of metabolite flux, RebH Y455W, TDC and RebF were anchored within a protein scaffold. Therefore, synthesized GBD, SH3 and PDZ ligand sequences were integrated into the GB cloning system by a domestication reaction (see 2.1.6). For assembly of TUs, 3'-ends of *rebH Y455W*, *rebF* and *tdc* were ligated to *SH3*, *PDZ* and *GBD* ligand sequences, respectively. Additionally, all CDS were fused to the CaMV P35S_ATG promoter and Nos terminator. The sequence of the synthetic protein scaffold $G_1S_1P_1$ was amplified by PCR for addition of *Bsm*BI recognition sites and GB fusion sites using primers listed in tab. 2.5. Following purification, the CDS was ligated into the universal domesticator plasmid *via* a GB reaction. Subsequently, $G_1S_1P_1$ was fused to the CaMV P35S_ATG promoter sequence and the TNos sequence during TU assembly. Finally, α level plasmids of *rebH Y455W-SH3* and *rebF-PDZ* were combined to the multigene construct SC7 (fig. 3.1). In addition, the *tdc-GBD* and the scaffold TUs were assembled to the Ω level plasmid SC3 (fig. 3.1). The construction of the scaffold transcriptional units and multigene constructs was performed by Bastian Wagner in course of a scientific internship.

Each cloning step involved a GB reaction for either domestication of different parts or assembly of TUs and multigene constructs (see 2.2.6). Subsequently, *E. coli* TOP10 or DH5α competent cells were transformed with the total amount of GB reaction *via* heat shock and cultivated on LB agar supplemented with X-Gal for blue-white screening as well as antibiotics according to the selection marker genes of the particular destination vector (see 2.2.8.1). Following overnight incubation at 37°C, white colonies were inoculated into LB medium supplemented with proper antibiotics and cultivated overnight at 37°C and 160 rpm. After plasmid DNA purification, correct assembly of the GB constructs was verified by DNA digestion and agarose gel electrophoresis (see 2.2.2, 2.2.7, 2.2.10). Overnight cultures of positive clones were stored as glycerol stocks at -80°C (see 2.2.13). For transient transformation of *N. benthamiana*, *A. tumefaciens* EHA105 or GV3101 competent cells were transformed with the desired GB plasmids *via* heat shock (see 2.2.8.2). Following the transformation procedure, bacteria were transferred to LB agar supplemented with appropriate antibiotics and subsequently incubated for 48 h at 28°C. Furthermore, *A. tumefaciens* colonies were screened by colony

PCR using the cr3 and cr4 primers binding specifically to the 35S promoter sequence (tab. 2.6, see 2.2.8.2). Finally, overnight cultures of positive clones were stored as glycerol stocks at -80°C.

3.1.2. Functional characterization of wild type flavin-dependent halogenases

Chlorination and bromination of tryptophan by Stth or RebH wt was investigated in transiently transformed *N. benthamiana* plants (see 2.4.3). If not stated otherwise, all infiltrations were performed in three biological replicates. Additionally, transformation with the pDGB2α2 plasmid served as a negative control (EV). After verification of sufficient enzyme accumulation, metabolites were extracted from leaf tissue and analyzed by HPLC regarding the biosynthesis of new products (see 2.5.1, 2.5.4.1). Further verification of m/z values of particular metabolites was done by LC-MS using the HPLC column and settings for reproduction of the initial results. Additionally, LC-MS parameters were optimized for the benefit of time-saving and efficient measurements.

3.1.2.1. Enzymatic activity of Stth and RebH wt in different cell compartments

Enzyme activity in the cytosol, chloroplasts and apoplast was analyzed by HPLC and LC-MS in terms of production of chlorinated tryptophan (see 2.5.5.1, 2.5.5.3). Localization of RebH wt in the cytosol (fig. 3.1, D13) and chloroplasts (fig. 3.1, E9) resulted in the formation of a new metabolite, absent in control samples, which exhibited an R_t value of 7.9 min (fig. 3.2., A). UV-absorption spectrum of this molecule was similar to tryptophan (220 nm, 280 nm) featuring two maxima at 224 nm and 282 nm (fig. 3.2, B). This UV-absorption shift could be attributed to substitution of chlorine into the heteroaromatic ring system. Further LC-MS analyses revealed an m/z value of 239 in the positive ionization mode for the putative 7-clorotryptophan, which eluted after 7.49 min (fig. 3.2., C, D). However, no RebH wt activity was detected after translocation to the apoplast (fig. 3.1, D8).

Besides biosynthesis of 7-chlorotryptophan, formation of a second, unknown molecule was observed when the halogenase was localized either in the cytosol or chloroplasts. Characterization of this putative product of RebH wt revealed an R_t value of 11.39 min and UV maxima at 226 nm and 284 nm (fig. 3.2, A)

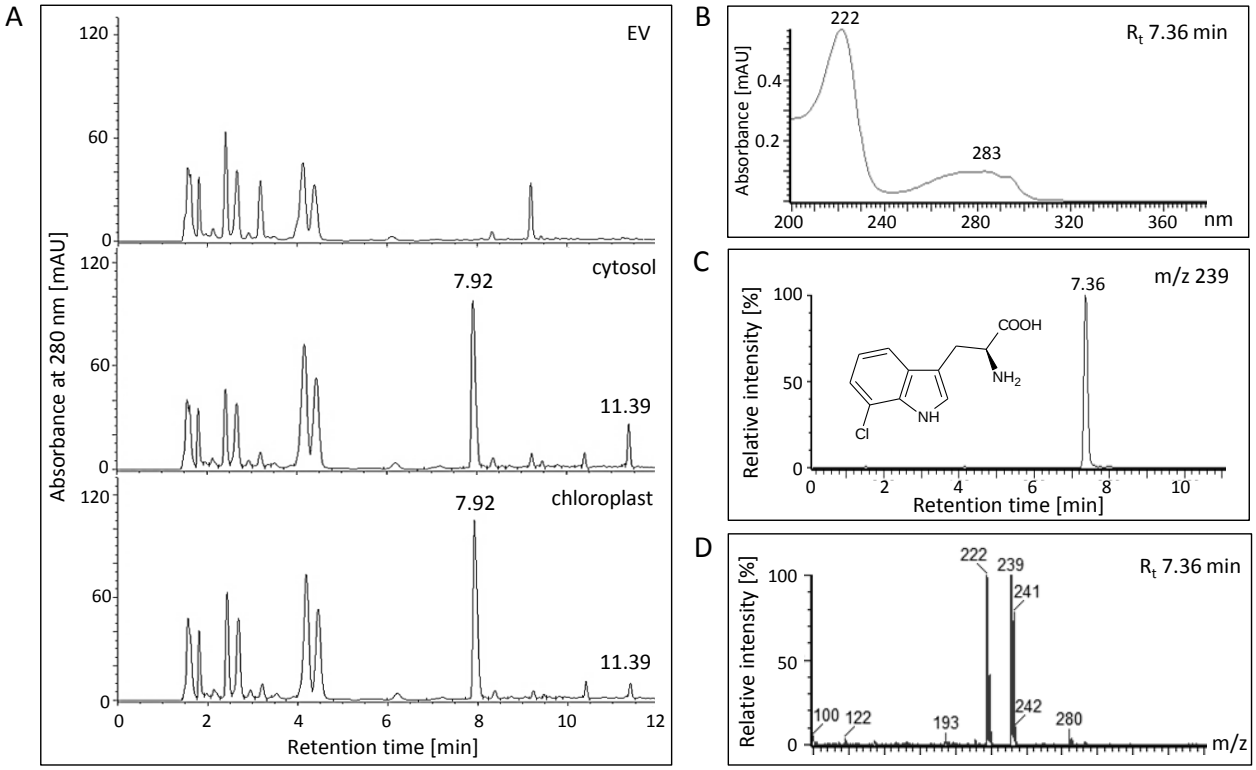

Figure 3.2: Functional characterization of RebH wt in different cell compartments of transiently transformed *N. benthamiana* leaves. A: RebH wt localization in the cytosol and chloroplasts correlated with the biosynthesis of two new metabolites (R_t 7.92 min, R_t 11.39 min) absent in control samples (EV) in HPLC analyses **B:** UV-spectrum of the putative 7-chlorotryptophan (R_t 7.36 min, LC-MS analysis) **C:** The putative 7-chlorotryptophan exhibited an m/z value of 239 in LC-MS analyses. **D:** MS spectrum of the putative 7-chlorotryptophan showed two m/z values of 239 and 241 consistent with mono-chlorinated tryptophan.

Mirroring the outcomes of characterization of RebH wt, localization of Stth in the cytosol (fig. 3.1, F13) and chloroplasts (fig. 3.1, H9) resulted in the formation of a new, highly concentrated metabolite undetectable in the control samples (EV). The putative 6-chlorotryptophan eluted after 8.3 min, exhibiting two UV maxima at 227 nm and 284 nm (fig. 3.3, A, B). Compared to the characteristic tryptophan UV-spectrum, the slightly shifted UV maxima might result from chlorine substitution of arenes. Subsequent UPLC-MS analyses revealed an m/z value of 239 in positive ionization mode for the investigated metabolite eluting after 7.7 min (fig. 3.3, C). Moreover, the mass spectrum exhibited two m/z values of 239 and 241 consistent with mono-chlorinated tryptophan. However, no Stth activity was detected after translocation of the halogenase to the apoplast (fig. 3.1, H10).

Further, localization of Stth in the cytosol and chloroplasts correlated with the formation of a second, unknown metabolite which eluted after 11.55 min in HPLC analyses (fig. 3.3, A). The R_t value diverged slightly from the RebH wt product, probably due to halogenation at a different position of the aromatic ring system. Moreover, UV maxima at 231 nm and 285 nm displayed an absorption shift comparable to that characteristic of the tryptophan spectrum.

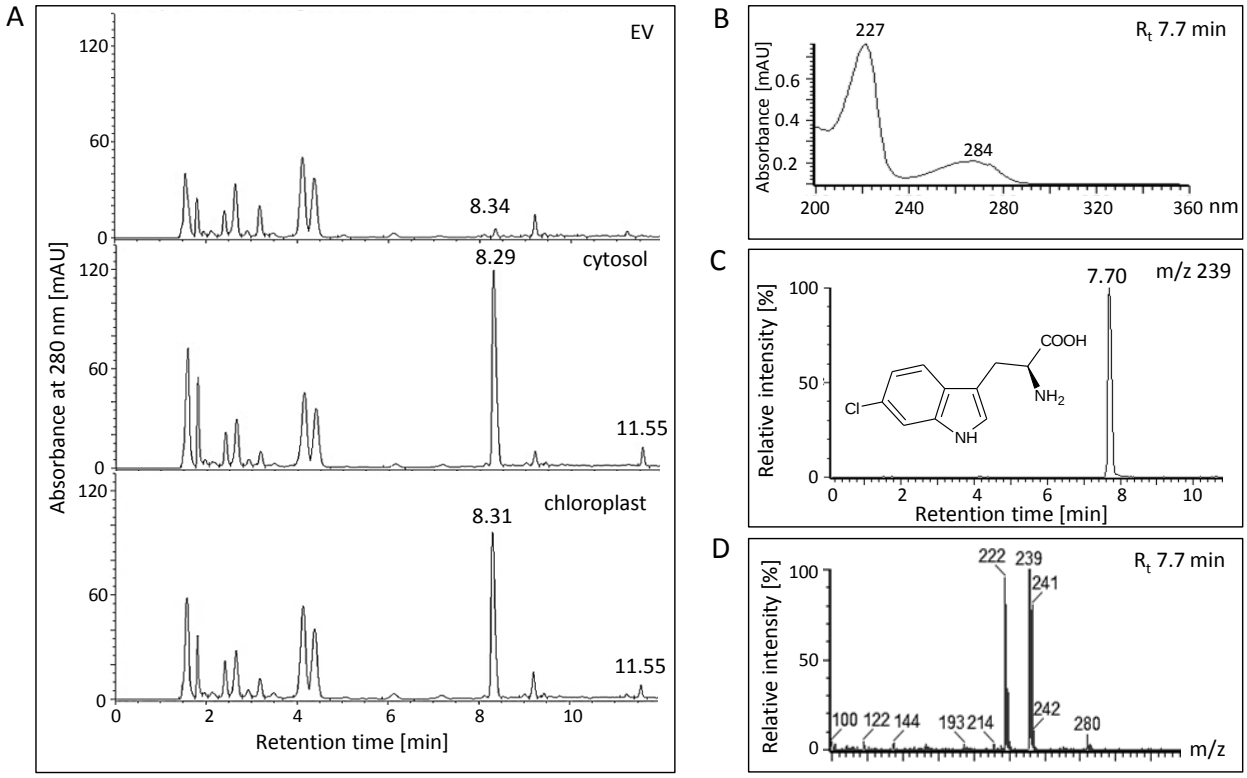

Figure 3.3: Functional characterization of Stth in different cell compartments of transiently transformed *N. benthamiana* plants. A: Localization of Stth in the cytosol and chloroplasts correlated with the accumulation of two halogenase products (R_t 8.3 min, R_t 11.6 min) absent in negative controls (EV) in HPLC analyses **B:** UV-spectrum of the putative 6-chlorotryptophan (R_t 7.7 min, LC-MS analyses) **C:** The Stth product exhibited an m/z value of 239 in LC-MS analyses **D:** The MS spectrum of the putative 6-chlorotryptophan showed two m/z values of 239 and 241 consistent with a mono-chlorinated tryptophan molecule.

Co-localization of bacterial halogenases, RebH wt or Stth with flavin reductase, RebF resulted in the biosynthesis of 7-or 6-chlorotryptophan in the cytosol and chloroplasts, indicating that the enzymes are suitable for the production of secondary metabolites *in planta*. With respect to the huge variation of metabolite concentrations within RebH wt and Stth samples and absent internal controls, calculation of the most efficient halogenase localization is challenging. However, peak integration of chlorotryptophan

derivatives indicated that the activity of both halogenases is similar in the cytosol and chloroplasts. Moreover, $FADH_2$, generated by RebF from *L. aerocolonigenes* was utilized by the tryptophan 6-halogenase, Stth from *S. toxytricini*. The exchange of flavin reductases from different organisms was previously demonstrated by Zeng *et al*. who characterized the catalytic properties of Stth using the flavin reductase Fre from *E. coli* BL21 for $FADH_2$ biosynthesis in *in vitro* assays [29]. Given that both halogenated tryptophan molecules represent the most prominent peaks within the plant extracts, RebH wt and Stth are eminently suitable for an efficient and inexpensive production of these rare metabolites. Moreover, characterization of both halogenases revealed their potential for flexible application in biosynthetic pathways of halogenated secondary metabolites by localization of enzymes in different cell compartments. However, no activity was observed after translocation of the investigated halogenases to the apoplast. One reason for this finding could be the absence of tryptophan from the apoplast, demonstrated previously in four weeks old tomato plants [140]. Additionally, degradation of halogenases by proteases located in the apoplast might account for the activity loss of both enzymes. Proteases play a key role in physiological development and plant defense mechanisms against pathogens, such as fungi or bacteria [141]. To stabilize heterologous enzymes in the apoplastic space, biosynthesis of specific protease inhibitors can support accumulation of recombinant proteins as exemplified by Goulet *et al.* in transiently transformed *N. benthamiana* [142]. Finally, both halogenases seem to convert an endogenous molecule. The UV spectra of the resulting products were similar to halogenated tryptophan derivatives, indicating halogenation of an indole ring system. Based on the distinct retention times of both products, halogenation by RebH wt and Stth is probably stereospecific. Accordingly chlorination of tryptophan at C6 and C7 seems to affect the interaction strengths of the molecules with the stationary phase and results in slightly different R_t values. Further characterization of these unidentified products of RebH wt and Stth is given in chapter 3.1.3.

3.1.2.2. Functional analyses of RebH wt and Stth without the partner reductase, RebF

Next, the recruitment of endogenous flavin-reductases by RebH wt and Stth was investigated in terms of chlorotryptophan biosynthesis. Therefore, RebH wt and Stth were localized, without their partner reductase, RebF in the cytosol (fig. 3.1, D14, F14) and chloroplasts (fig. 3.1, E14, H14) in transiently transformed *N. benthamiana* leaves. Biosynthesis of 6- and 7-chlorotryptophan was analyzed by HPLC to verify halogenase activity (see 2.5.5.1).

Targeting of RebH wt and Stth to chloroplasts resulted in the biosynthesis of significant amounts of 7- and 6-chlorotryptophan, respectively. In contrast, only trace amounts of the metabolites were detectable when both halogenases were present in the cytosol (fig. 3.4). Comparison of integrated peak areas of halogenated tryptophan revealed only 1.7 ± 0.6 % activity of RebH wt in the cytosol compared to chloroplast localization. Moreover, Stth activity level in the cytosol reached only 4.3 ± 1.4 % of the chloroplast localized enzyme.

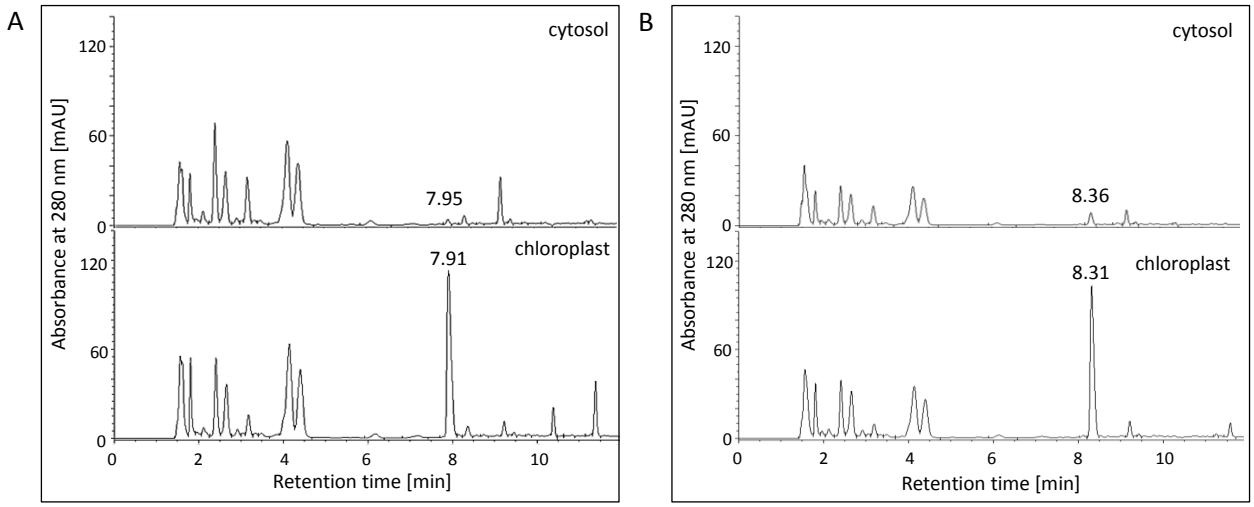

Figure 3.4: Functional analyses of RebH wt and Stth without partner reductase, RebF in the cytosol and chloroplasts. Formation of 6- and 7-chlorotryptophan in transiently transformed *N. benthamiana* leaves was analyzed by HPLC. **A:** 7-chloro-tryptophan (R_t 7.9 min) biosynthesis in the cytosol was dramatically reduced compared to chloroplast localization of RebH wt. **B:** Enzymatic activity of Stth in chloroplasts resulted in the accumulation of 6-chlorotryptophan (R_t 8.3 min) in high amounts, whereas cytosolic localization of the enzyme correlated with significantly decreased tryptophan halogenation.

The efficient biosynthesis of chlorotryptophan by RebH wt and Stth in chloroplasts in absence of a partner reductase implies presence of FAD as well as its reduction to $FADH_2$ by other mechanisms. FAD is an essential cofactor of various enzymes, such as oxidases, reductases and dehydrogenases [143, 144]. This molecule is synthesized by flavin synthases in chloroplasts of plants and might be reduced to $FADH_2$ by reductases or other proteins involved in the electron transfer of the photosystem [145]. However, supply of free $FADH_2$ was not sufficient in the cytosol, indicating a low or absent flavin reductase activity.

3.1.2.3. Formation of dichlorotryptophan *via* double halogenation by RebH wt and Stth

The ability of tryptophan halogenases to introduce a second chlorine atom to a chlorotryptophan molecule was investigated by transient co-transformation of *N. benthamiana* with *rebH wt* and *stth*. To evaluate the biosynthesis of dichlorotryptophan in the cytosol, both halogenases were co-localized with the partner reductase, RebF (fig. 3.1, D13 + F13). Additionally, tandem-halogenation of tryptophan in chloroplasts by RebH wt and Stth was analyzed without RebF (fig. 3.1, D14 + H14), as previous analyses had revealed that co-localization of the reductase in the organelles does not have any advantage over endogenous enzymes (see 3.1.2.2). Besides formation of 6- and 7-chlorotryptophan, co-localization of Stth and RebH wt correlated with accumulation of a third metabolite eluting after 10.36 min in HPLC analyses (fig. 3.5, A). In addition, this molecule featured two UV maxima, at 227 nm and 285 nm, indicating chlorine substitution to an indole ring system. Interestingly, the molecule was also detectable in control samples of RebH wt, displaying an identical UV-absorption spectrum. Moreover, biosynthesis of the observed metabolites occurred independently of co-localization of both halogenases in chloroplast or the cytosol. The new halogenase product revealed an m/z value of 273 in positive ionization mode eluting after 5.59 min (fig. 3.5, B) in LC-MS analyses (see 2.5.5.4). These results gave final proof of dichlorotryptophan biosynthesis by RebH wt and Stth. However, regioselectivity of halogenation could not be ultimately clarified using this approach.

These results indicate that 7-chlorotryptophan is processed by RebH wt leading to unspecific halogenation at a second carbon atom of the heteroaromatic ring. The regioselectivity of the second halogenation as well as the exact molecule structure of the resulting dichlorotryptophan needs to be analyzed using other methods, such as nuclear magnetic resonance (NMR). However, arene substitution at C6 and C7 seems to be most likely, as R_t values and UV-spectra of both dichlorotryptophan molecules are identical. Double halogenation of substrate molecules by RebH wt was previously observed by Payne *et al.* in *in vitro* assays [6]. Characterization of regioselective halogenation of a diverse assortment of aromatic metabolites had resulted in double halogenation of metabolites by RebH wt in case of long incubation periods [6].

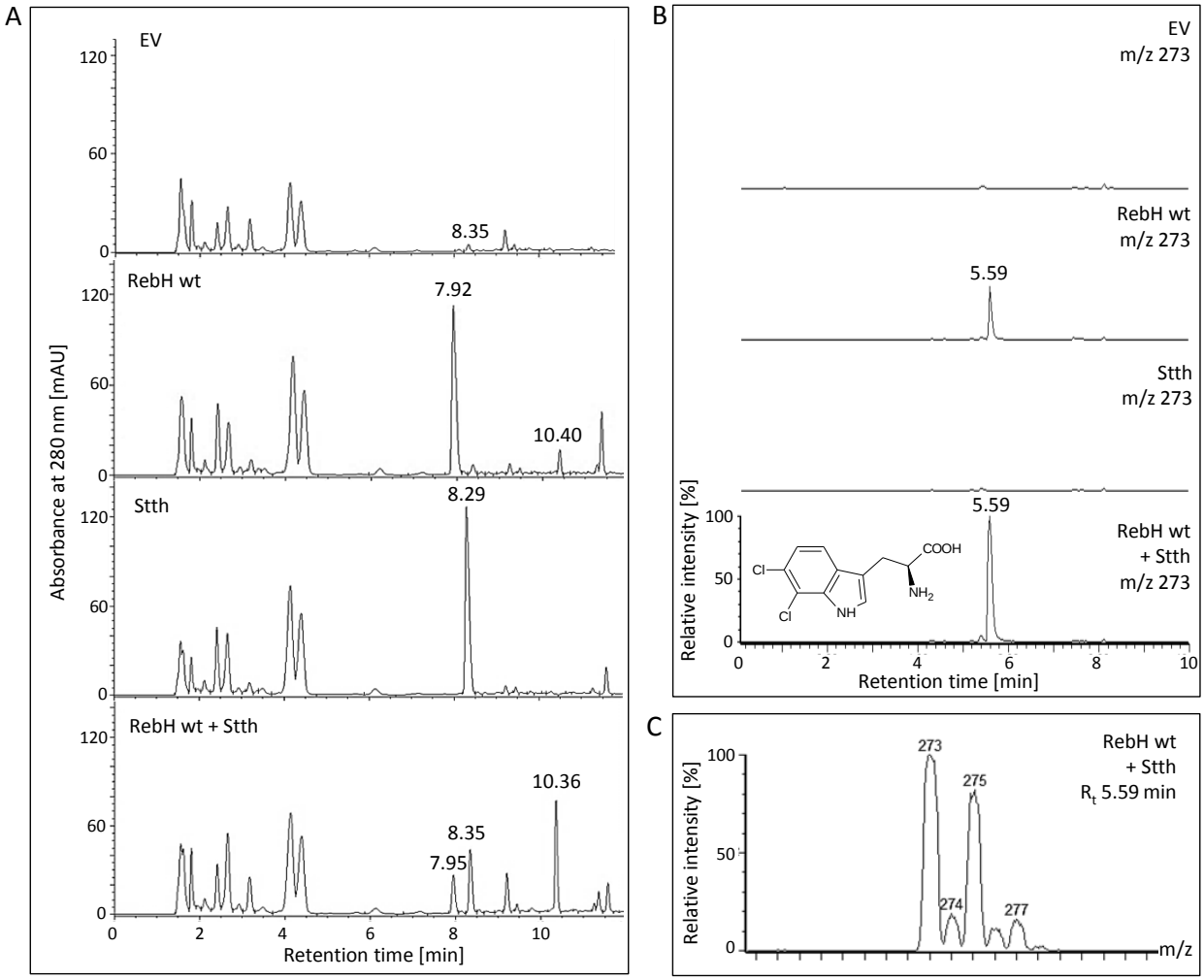

Figure 3.5: Biosynthesis of dichlorotryptophan by RebH wt and Stth. A: HPLC analyses revealed the biosynthesis of a new metabolite (R_t 10.36 min) by Stth and RebH wt compared to the negative control (EV), exemplified for chloroplastic localization of enzymes. **B:** The putative dichlorotryptophan, synthesized in chloroplasts, exhibited an m/z value of 273 in LC-MS analyses. **C:** Mass spectrum of the putative 6-,7-dichlorotryptophan.

Accumulation of 6- and 7-chlorotryptophan by Stth and RebH wt was evaluated by peak integration of HPLC chromatograms. Since no effects on metabolite biosynthesis could be detected in terms of enzyme localization, values of chloroplast and cytosol localized halogenases were summarized in one group. Due to absence of internal controls, interpretation of metabolite levels in different samples is challenging. Nevertheless, chlorotryptophan biosynthesis by either Stth or RebH wt was compared to the accumulation of these metabolites after co-localization of both halogenases. In this regard, 6-chloro-tryptophan levels reached only 25.1 ± 11.5 % after co-expression of *rebH wt* and *stth* compared to

expression of *stth* alone. Moreover, co-localization of both halogenases resulted in 7-chlorotryptphan amounts up to only 11.7 ± 7 % of that synthesized by RebH wt alone. These findings indicate that both 6- and 7- chlorotryptophan were converted to dichlorotryptophan in a highly efficient manner by RebH wt and Stth, respectively. Moreover, 7-chlorotryptophan concentration reached only 42.8 ± 26 % of that characteristic of 6-chlorotryptophan after co-transformation, suggesting that this metabolite is more efficiently converted to dichlorotryptophan than 6-chlorotryptophan. Given that RebH wt was shown to halogenate 7-chlorotryptophan, whereas Stth does not catalyze the halogenation of 6-chloro-tryptophan, reduced 7-chlorotryptophan levels probably result from conversion by both halogenases.

3.1.2.4. Halogenation of tryptamine by RebH wt and Stth

Tryptamine, as a building block of the universal precursor of MIAs, strictosidine, is a potential target for modifications like halogenation, leading to the biosynthesis of novel metabolites [11, 28]. Especially the addition of halogens to the indole ring system enables easy modifications of the skeleton *via* substitutions [5, 10]. Biosynthesis of 6- and 7-chlorotryptamine was analyzed by co-expression of *rebH wt* and *stth* with the tryptophan decarboxylase gene, *tdc* in transiently transformed *N. benthamiana* leaves. In this regard, chlorotryptamine formation in the cytosol was investigated by co-localization of TDC with either RebH wt (fig. 3.1, D13 + TDC) or Stth (fig. 3.1, F13 + TDC). Additionally, RebH wt and Stth were targeted to the chloroplasts, whereas TDC was localized in the cytosol to avoid any toxic effects on the tryptophan metabolism (fig. 3.1, E14 + TDC, H14 + TDC). Plant extracts were analyzed by HPLC (see 2.5.5.1) and m/z values of the metabolites of interest were determined by subsequent LC-MS measurements (see 2.5.5.4). The combined presence of RebH wt and TDC resulted in the formation of tryptamine (R_t 4.0 min) and 7-chlorotryptophan (R_t 7.95 min) as well as a third metabolite eluting after 9.27 min in HPLC analyses (fig. 3.6, A). The latter exhibited two UV maxima at 222 nm and 283 nm indicating halogenation of an indole ring system. Moreover, biosynthesis of this putative 7-chlorotryptamine was not affected by cellular localization of the halogenase. Additional LC-MS analyses of this new product revealed an R_t value of 5.07 min and an m/z value of 195 anticipated for 7-chlorotryptamine in positive ionization mode (fig. 3.6, B). Similar results were obtained analyzing products of TDC and Stth in leaf tissue. Besides tryptamine and 6-chlorotryptophan (R_t 8.36 min), an additional metabolite eluted after 9.6 min in HPLC analyses (fig. 3.6, A). The putative 6-chlorotryptamine exhibited two UV maxima at 227 nm and 285 nm. Subsequent LC-MS experiments revealed an R_t value of 5.21 min and an m/z value of 195 anticipated for chlorotryptamine (fig. 3.6, B). Furthermore, mass

spectra of the putative 6- and 7-chlorotryptamine exhibited two m/z values of 195 and 197 consistent with a mono-chlorinated tryptamine molecule. To investigate the biosynthesis of dichlorotryptamine, TDC was co-localized with both halogenases after triple infiltration in *N. benthamiana*. Biosynthesis of TDC, RebH wt and Stth resulted in the formation of dichlorotryptophan (fig. 3.6, A, R_t 10.37 min) as demonstrated previously (see 3.1.2.3), whereas dichlorotryptamine could not be detected in any of the samples (fig. 3.6, A). Therefore, leaf extracts were screened for the anticipated m/z value of 229 and compared to control infiltrations of the empty vector in LC-MS analyses. However, accumulation of dichlorotryptamine was not observed in any of the samples.

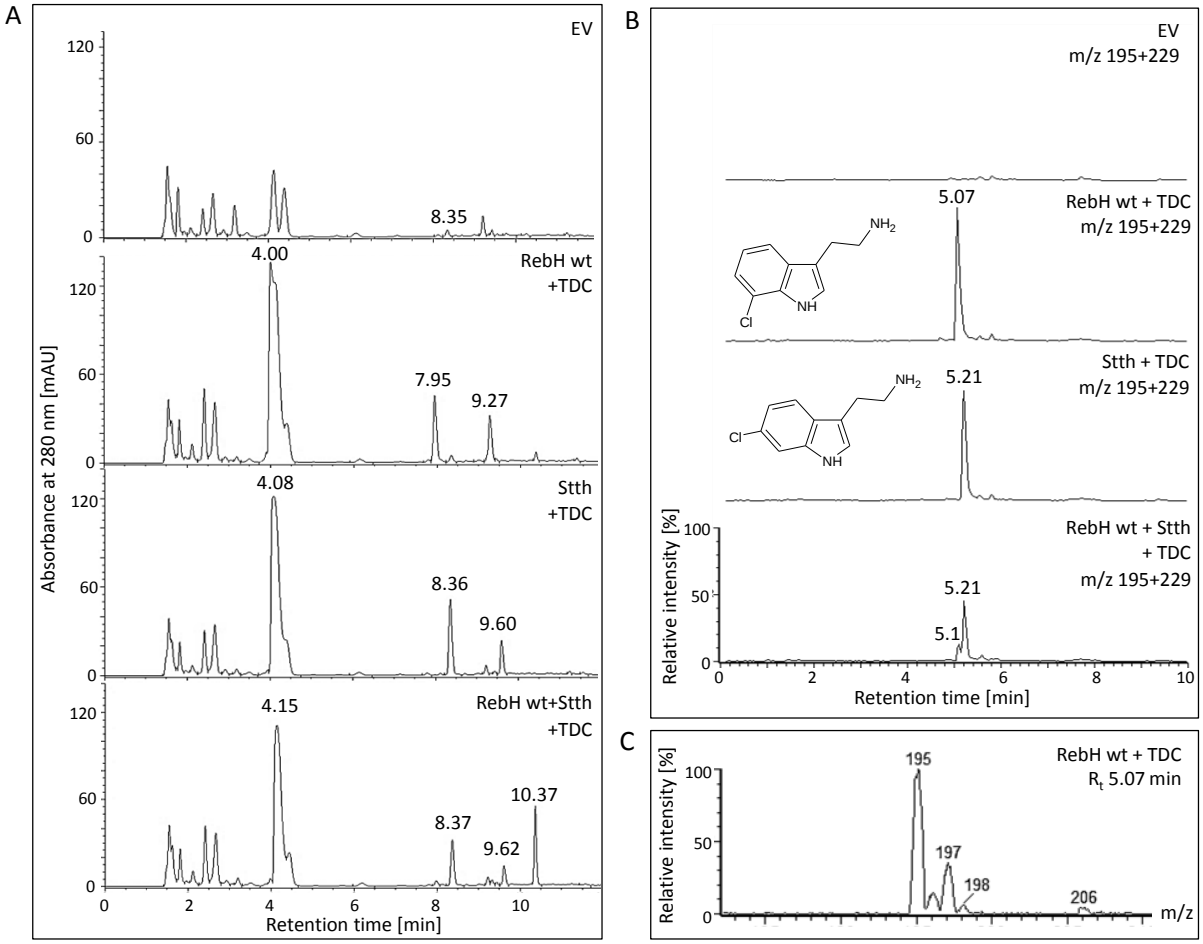

Figure 3.6.: Biosynthesis of halogenated tryptamine derivatives by TDC, Stth and RebH wt exemplified for cytosolic localization of halogenases. A: The biosynthesis of two new molecules eluting after 9.27 min (RebH wt + TDC) and 9.6 min (Stth + TDC) was observed in HPLC analyses. tryptamine (R_t: 4.0, 4.1 min); 7-chlorotryptophan (R_t: 7.95 min), 7-chlorotryptamine (R_t: 9.27 min), 6-chlorotryptophan (R_t: 8.36 min), 6-chlorotryptamine (R_t: 9.6 min) **B:** LC-MS analyses revealed the biosynthesis of chlorotryptamines (m/z 195) after co-expression of halogenase genes and *tdc*, whereas dichlorotryptamine (m/z 229) was not observed within leaf extracts. C: Mass spectrum of putative 7-chlorotryptamine synthesized by RebH wt and TDC.

These results demonstrate, that neither TDC converts dichlorotryptophan nor do Stth or RebH wt halogenate chlorotryptamine. For the evaluation of chlorotryptamine content in leaf tissue, peak areas were compared to a 5-chlorotryptamine standard of known concentration. The calculated 6-chlorotryptamine levels were 6.17 ± 2 ng per mg fresh weight of infiltrated leaf tissue. However, 7-chlorotryptamine concentrations were not determined due to co-eluting endogenous metabolites. Therefore, peak integration should be repeated and calculated according to the specific mass of the metabolite by LC-MS.

Two potential biosynthetic pathways can lead to the formation of halogenated tryptamine derivatives. First, tryptophan halogenases RebH wt and Stth convert tryptamine to chlorotryptamine or second, halogenated tryptophan derivatives are converted by TDC. Glenn *et al.* demonstrated conversion of 40 % of initially applied tryptamine by RebH wt after 12 h in *in vitro* assays [28]. Moreover, TDC converts 7-chlorotryptophan at only 3 % efficiency, as compared to the natural substrate, tryptophan [11]. Accordingly, 7-chlorotryptamine biosynthesis most likely results from tryptamine halogenation by RebH wt. However, the conversion rate of tryptamine by Stth has not been reported, yet. Therefore, biosynthesis of 6-chlorotryptamine was investigated in more detail to identify the reaction mechanisms. For this purpose, transiently transformed *N. benthamiana* leaves accumulating Stth in the cytosol (fig. 3.1, F13) were infiltrated with 400 µM tryptamine (see 2.4.6). Co-expression of *stth* and *tdc* served as a control (fig. 3.1, F13 + TDC) and both infiltration approaches were performed in two biological replicates. Two days after tryptamine infiltration, metabolites were extracted and analyzed by LC-MS (see 2.5.5.3). In case of substrate supplementation, only trace amounts of 6-chlorotryptamine could be detected, demonstrating the low conversion rate of tryptamine by Stth. In contrast, co-transformation of *stth* and *tdc* resulted in accumulation of relatively high levels of 6-chlrotryptamine (fig. 3.7). However, tryptamine concentration was considerably lower in samples supplemented with this metabolite compared to those synthesizing TDC. This finding indicates that the tryptamine content differed significantly within these two groups during the entire incubation period. Therefore, low 6-chloro-tryptamine concentration probably resulted from limitation of the substrate rather than inefficient tryptamine halogenation by Stth.

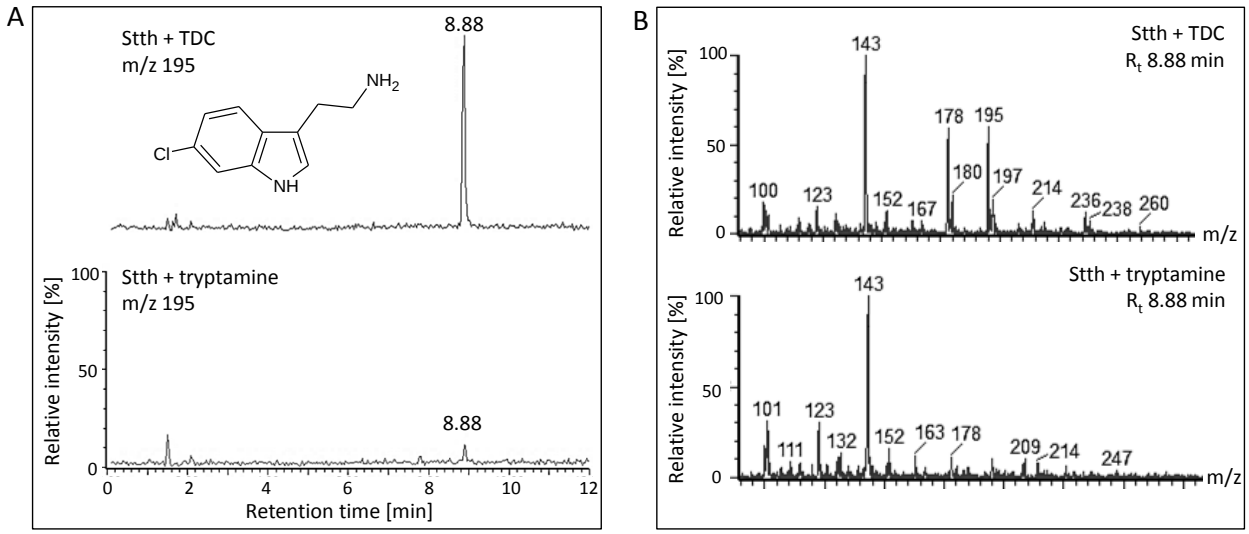

Figure 3.7: LC-MS analyses of 6-chlorotryptamine biosynthesis by Stth and TDC. A: LC-MS analyses revealed low conversion rate of supplemented tryptamine to 6-chlorotryptamine (m/z 195) by Stth compared to samples with TDC activity. **B:** Mass spectrum of 6-chlorotryptamine (m/z 195) displayed the low concentration of this metabolite after infiltration of tryptamine.

Given that the low conversion rate of tryptamine by Stth might be influenced by low availability of substrate after infiltration, the order of enzymatic steps leading to the biosynthesis of 6-chlorotryptamine remains unclear. To evaluate if either tryptamine is halogenated by Stth or 6-chlorotryptophan is converted by TDC, purified enzymes need to be analyze in terms of catalytic activities on the corresponding substrate. Determination of catalytic constants, such as of K_m and k_{cat} values, could give further insights into potential metabolite fluxes. Purification of Stth fused to a C-terminal His-tag can be easily achieved from leaf tissue after transient transformation. However, recombinant TDC does not include any tags for enzyme purification and requires reconstruction of the TU. Moreover, 6-chloro-tryptophan is a rare chemical in contrast to tryptamine and probably needs to be synthesized chemically or purified from leaf tissue after transient expression of *stth.* Therefore, catalytic activity of Stth in terms of tryptamine halogenation can be determined with feasible effort, whereas characterization of the catalytic properties of TDC requires considerable expenditure. However, given that TDC converts 5- and 7-chlorotryptophan at 5 % and 3 % efficiency [11], conversion of 6-chlorotryptophan by this enzyme might prove inefficient as well. Therefore, biosynthesis of 6-chlorotryptamine probably results from tryptamine halogenation by Stth.

3.1.2.5. Bromination of tryptophan by RebH wt and Stth

The tryptophan halogenases RebH wt and Stth both catalyze the substitution of bromide on the heteroaromatic ring of tryptophan in a regiospecific manner [25, 29]. The potential of both halogenases to synthesize these expensive metabolites *in planta* was examined after transient transformation of *N. benthamiana* leaves. In this regard, bromination of tryptophan was investigated by localization of either RebH wt (fig. 3.1, D13) or Stth (fig. 3.1, F13) in the cytosol. Furthermore, 40 mM KBr solution was supplemented into transiently transformed leaves one day after the initial infiltration. Transformation with the pDGB2α2 plasmid and subsequent KBr infiltration served as a control. Three days after KBr supplementation, metabolites were extracted from leaf tissue and analyzed by LC-MS (see 2.5.4.1, 2.5.5.5).

Compared to negative controls, activity of RebH wt and Stth correlated with the biosynthesis of a new metabolite exhibiting an R_t value of 2.40 min and an m/z value of 283 (fig. 3.8, A). In contrast to previously analyzed halogenated metabolites, no shift of the retention times was observed in terms of C6 or C7 bromination. Given that regiospecific bromination by RebH wt and Stth was demonstrated in published results [25, 29], identical retention times of both metabolites probably result from the short column length rather than an identical halogenation pattern. Moreover, mass spectra of both molecules exhibited m/z values of 283 and 285 with identical signal intensities consistent with a mono-brominated tryptophan molecule (fig. 3.8, B).

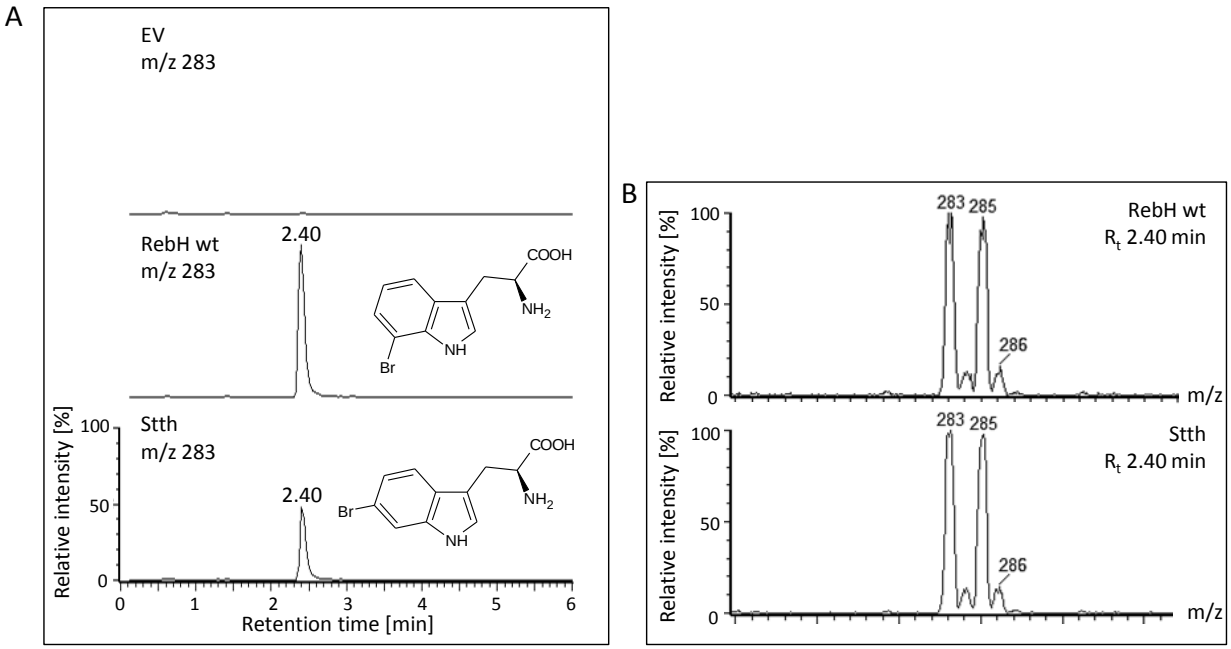

Figure 3.8: Biosynthesis of bromotryptophan by RebH wt and Stth in transiently transformed *N. benthamiana*. A: LC-MS analyses revealed the biosynthesis of 7-bromotryptophan by RebH wt as well as 6-bromotryptophan by Stth after KBr supplementation **B:** Mass spectra of both bromotryptophan isomers exhibited m/z values of 283 and 285 with identical signal intensities representing ion distribution characteristic of bromide. EV=empty vector.

Additionally, dibromination by cytosolically localized RebH wt and Stth (fig. 3.1, D13 + F13) was examined after co-transformation of *N. benthamiana* leaves and subsequent infiltration of 40 mM KBr. Metabolites were extracted from leaf tissue and analyzed by LC-MS (see 2.5.4.1, 2.5.5.5). Compared to negative controls, co-expression of *rebH wt* and *stth* correlated with the formation of a new metabolite that eluted after 2.82 min and revealed m/z values of 361, 363 and 362 with signal intensity ratios of 1:2:1, which is consistent with a double-brominated tryptophan molecule (fig. 3.9). Moreover, a second metabolite exhibiting an R_t value of 2.82 min and m/z values of 317, 319 and 321 was observed (fig. 3.9). According to the detected mass and the characteristic mass spectrum, this molecule was most likely, a mono-chlorinated and mono-brominated tryptophan molecule. However, stereospecific substitution of both halogens could not be determined in this approach. In this regard, the determined mass peak presumably represents a mixture of 6-bromo-7-chlorotryptophan and 6-chloro-7-bromotryptophan. NMR analyses could give further insights into the chemical structure of the observed molecule and its halogenation pattern.

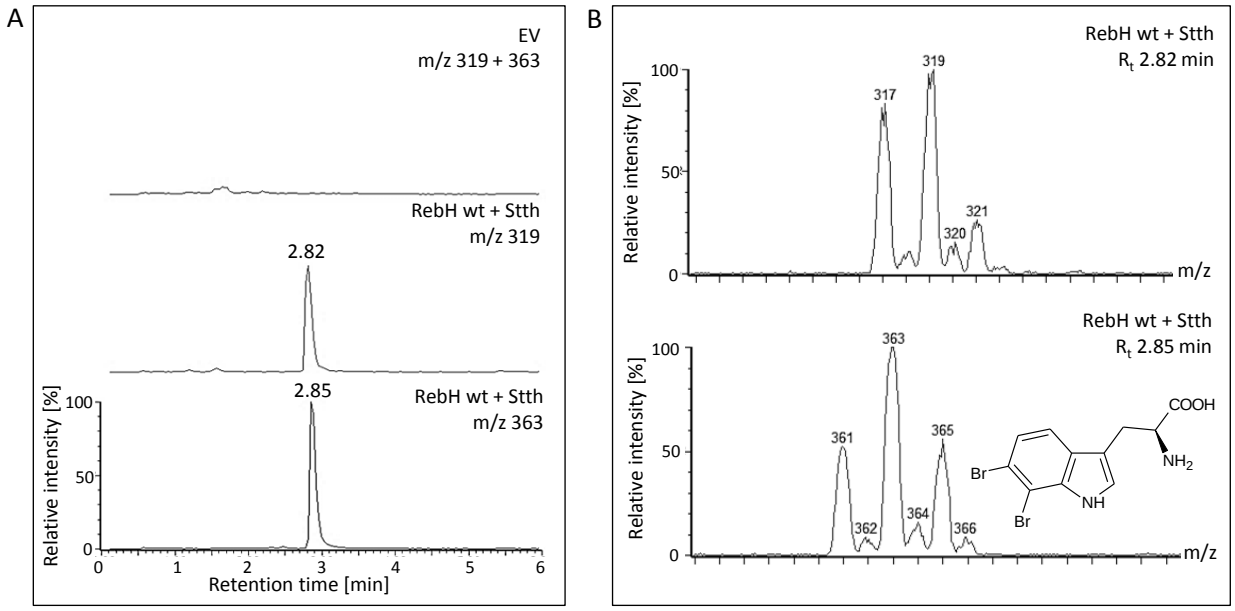

Figure 3.9: Biosynthesis of mono-brominated and mono-chlorinated tryptophan as well as dibromotryptophan. A: Co-localization of RebH wt and Stth in the cytosol correlated with the formation of a mono-chlorinated and mono-brominated tryptophan molecule (m/z 319) exhibiting an R_t value of 2.82 min. Moreover, biosynthesis of dibromotryptophan (m/z 363) eluting after 2.85 min was observed in LC-MS analyses. **B:** Mass spectra revealed the mass units' distribution characteristic of mono-chlorinated and mono-brominated tryptophan (R_t 2.82 min) as well as dibromotryptophan (R_t 2.85 min).

The presented transient approach demonstrates the broad variety of possible applications of RebH wt and Stth for modification of metabolites *in planta*. The high catalytic activity of both enzymes enabled efficient bromination of tryptophan leading to high amounts of this molecule in leaves of *N. benthamiana* after no more than four days. Final concentrations of the newly synthesized molecules could not be determined without standard metabolites. However, integration of peak areas revealed that 7-bromotryptophan levels were approximately 1.5 fold of that of 7-chlorotryptophan after KBr supplementation. Moreover, Stth synthesized equal amounts of bromo- and chlorotryptophan. According to published data, RebH wt brominates tryptophan with a k_{cat} value of 0.4 min^{-1}, about three times lower than for chlorination, indicating that bromide concentrations exceeded those of chloride within the cytosol after infiltration [25]. The average chloride concentration in the cytoplasm of higher plants ranges between 3 and 10 mM [146]. Moreover, anions are actively transported into the cell by anion/proton symporters while being released by passive fluxes through anion channels located in the plasma membrane [146]. The increased bromination of tryptophan suggests active transport of bromide into the cytoplasm after leaf infiltration, resulting in bromide levels that probably exceed chlorine

concentration, thereby promoting biosynthesis of bromotryptophan. Given that both halogenases efficiently chlorinated tryptamine it is likely, that RebH wt and Stth might also brominate this metabolite, which can be analyzed by co-localization of halogenases and TDC in the cytosol and subsequent supplementation of KBr.

In summary, characterization of the tryptophan 7-halogenase, RebH wt and the 6-halogenase, Stth unveiled various possible applications of these enzymes to facilitate the biosynthesis of expensive and rare molecules, such as chlorotryptophan, dichloroptryptophan, chlorotryptamine, bromotryptophan and dibromotryptophan. Besides the herein described transient approaches, stable transgenic tobacco plants would facilitate constitutive production of these molecules. Moreover, target-oriented biosynthesis of particular metabolites could be achieved through establishment of cell suspension cultures. By modification of the culture medium composition through chlorine elimination or bromide supplementation, biosynthesis of one specific molecule could be fostered. Additionally, including tryptophan to the culture medium might further increase yields of final products. Finally, the variety of halogenated products can be magnified by introduction of an additional tryptophan 5-halogenase, such as PyrH from *Streptomyces rugosporus*, which was shown to be functional in *C. roseus* [11, 24]. This assortment could even be extended by integration of a tryptophan 4-halogenase. However, no flavin-dependent halogenase has been described until today that catalyzes halogen substitution at C4 position of tryptophan. Finally, the combination of different halogenases, the addition of TDC and supplementation of bromide could facilitate the biosynthesis of a huge variety of mono-, di- or tri halogenated tryptophan and tryptamine derivatives *in planta*.

3.1.3. Characterization of the unidentified metabolite halogenated by RebH wt and Stth

Initial functional analyses of RebH wt and Stth *in planta* had revealed a halogenated product synthesized by both enzymes in addition to the anticipated tryptophan derivatives (see 3.1.2.1). Due to the stereospecific halogenation by RebH wt and Stth R_t values of both molecules differed slightly. Both halogenation metabolites and the endogenous precursor were characterized in more detail by UPLC-MS (see 2.5.5.3). In this regard, RebH wt and Stth products revealed an m/z value of 325 in positive ionization mode and an R_t value of 10.44 min and 10.65 min, respectively. Furthermore, mass spectra exhibited two m/z values of 279 and 281, an indicator for mono-chlorinated molecules (fig. 3.10). Accordingly, plant extracts were screened for a potential precursor with a molecular weight of

290 g/mol. In all samples, including negative controls, one metabolite featuring the expected m/z value of 291 and an R_t value of 8.43 min was observed in high concentrations (fig. 3.10). To investigate if the halogenated metabolite synthesized by RebH wt and Stth originates from the detected endogenous molecule, fragmentation patterns of the putative substrate and the product were analyzed *via* MS-MS (see 2.5.5.3). The obtained m/z values of the fragments of both molecules displayed high similarity, revealing nine fragments of common origin (tab. 3.1). Interestingly, all fragments of the halogenated metabolites exhibited an increase of 34 mass units compared to precursor fragments, indicating addition of a chlorine atom (fig. 3.10).

Tab. 3.1: MS/MS analyses of the endogenous tobacco metabolite and its halogenated product. Comparison of m/z values of the potential precursor molecule and corresponding m/z values of the halogenated product.

m/z values after fragmentation of the endogenous metabolite	130	146	159	170	185	200	209	227	245
associated m/z values after fragmentation of the halogenated RebH wt product	164	180	193	204	219	234	243	261	279

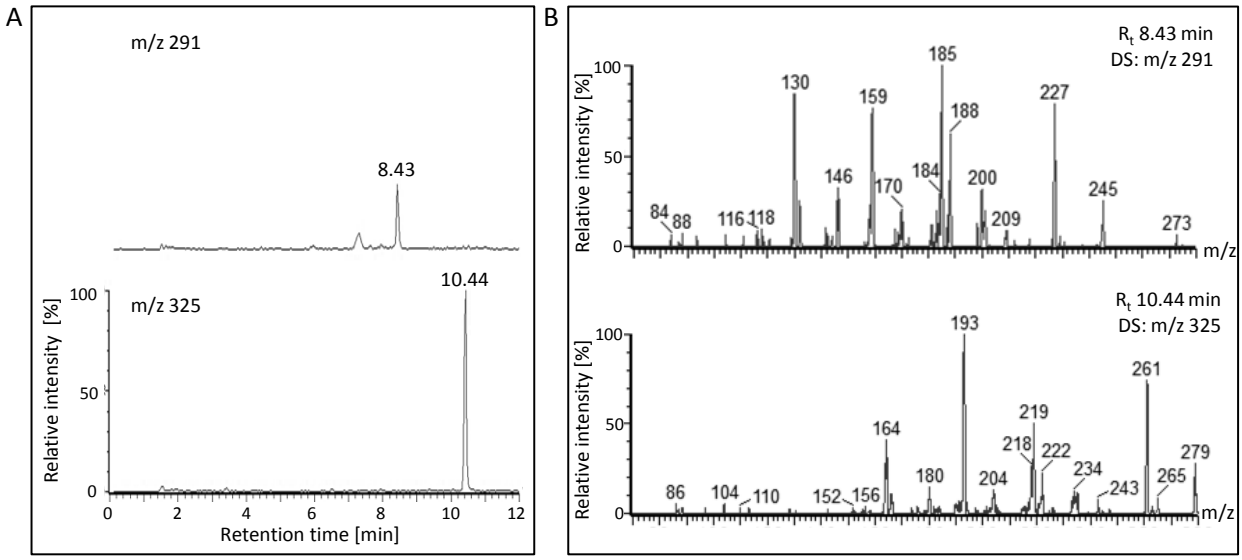

Figure 3.10: UPLC-MS analyses of the unidentified halogenation product and its proposed precursor exemplified for tryptophan halogenase RebH wt localized in chloroplasts. A: LC-MS analyses of the halogenated RebH wt product (m/z 325) and its putative endogenous precursor (m/z 291) **B:** Daughter scan of the unidentified RebH wt product (m/z 325) and its putative endogenous precursor (m/z 291).

To provide final evidence, that the precursor molecule, halogenated by RebH wt and Stth, originated from tobacco rather than infiltrated *Agrobacterium* cells, extracts from four and six weeks old *N. benthamiana* wt plants were analyzed in two replicates by HPLC (see 2.5.5.1). The extracts were derived from several leaves to exclude any influence of their developmental stages. The unidentified metabolite was detected in four weeks old *N. benthamiana* plants, while only marginal amounts were observed in leaves of six weeks old tobacco (fig. 3.11).

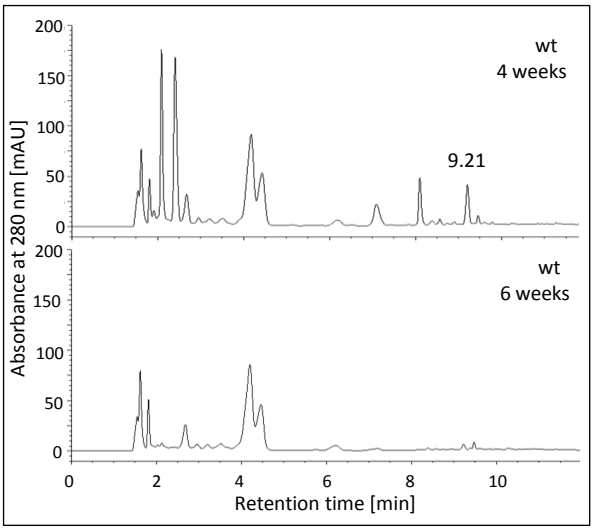

Figure 3.11: HPLC analyses of *N. benthamiana* wt plants of two different developmental stages. The investigated metabolite (R_t 9.21 min) was observed in leaves of four weeks old tobacco, indicating that the putative precursor of the unidentified halogenase product originates from tobacco.

The accumulation the putative precursor molecule in untreated *N. benthamiana* leaves confirms the initial assumption that the unidentified product of both halogenases originates from an endogenous metabolite. Moreover, these results demonstrate variation in the concentration of the endogenous molecule according to the developmental stage of the plant which indicates its physiological function. For the observed biosynthesis of the halogenated molecule originating from tobacco, two possibilities were considered. First, tryptophan halogenases RebH wt and Stth might directly chlorinate the endogenous tobacco molecule. Second, the halogenated product might originate from chlorotryptophan, as a precursor in the biosynthetic pathway of the endogenous metabolite. The latter assumption implies conversion of halogenated tryptophan by endogenous tobacco enzymes of this unknown biosynthetic pathway. To verify this hypothesis, wild type plants could be supplemented with chlorotryptophan and analyzed for biosynthesis of the halogenated metabolite. In case of conversion of 6- and 7-chlorotryptophan by endogenous enzymes, formation of the halogenated product should be observed.

One promising candidate molecule for Stth- and RebH wt-catalyzed halogenation in tobacco is indole-3-acetic acid–aspartate (IAA-Asp) first identified in *Nicotiana sp.* by Sitbon *et al.* [147] (fig. 3.12, A). Many IAA conjugates known from plants represent an inactive or storable form of the phytohormone, auxin [148]. While metabolic function of IAA-Asp in plants is not completely understood, several studies provided evidence for its involvement in inactivation of free auxin, since IAA-Asp is not well hydrolyzed by amidohydrolases [149]. This metabolite has the same molecular mass as the precursor molecule of 290 g/mol and accumulates up to 7.8 ± 0.4 ng/g fresh weight in *A. thaliana* [150]. Therefore, IAA-Asp seemed to be a highly promising candidate for further characterization. For this reason plant extracts were supplemented with 690 mM indole-3-acetic acid–aspartate, to exclude any effects of the tobacco matrix and its mass fragmentation was compared to the endogenous tobacco metabolite by LC-MS/MS (see 2.5.5.3). However, analytical results revealed a vast discrepancy between both fragment patterns, giving evidence that IAA-Asp was not the endogenous precursor halogenated by RebH wt and Stth.

An alternative molecule taken into account as a potential substrate for tryptophan halogenases was the flavonoid, catechin (fig. 3.12, B). This secondary metabolite is found in many plant species like in *Camellia sinensis* or in fruits of *Vitis sp.* [151, 152]. It has two chiral centers leading to four diastereoisomers with a molecular weight of 290 g/mol. Moreover, UV spectra of catechin stereoisomers feature two maxima at 210 nm and 270-280 nm [153]. Therefore, catechin seemed to be a good candidate for further LC-MS/MS analyses to identify the endogenous halogenase substrate. However, this hypothesis could not be proved within the time scope of this research project. Besides analytical characterization of catechin enantiomers, further studies of the endogenous metabolite, like identification of functional groups or NMR analyses can help to unveil the molecule structure. For this investigation, the metabolite needs to be purified from plant extracts in mg amounts, e.g. by preparative HPLC. Identification of the natural plant-derived metabolite could give further insights into the substrate scope of RebH wt. As demonstrated previously by Payne *et al.*, RebH wt converts a huge variety of bicyclic and tricyclic molecules substituted with diverse side-chains in contrast to other tryptophan halogenases, such as PrnA [6, 7].

As a structural analog of tryptophan, indole-3-acetic acid (IAA) seemed to be an interesting target for tryptophan halogenases (fig. 3.12). In this regard, conversion of IAA by either RebH wt (fig. 3.1, D13) or

Stth (fig. 3.1, F13) localized in the cytosol was investigated in transiently transformed *N. benthamiana* leaves (see 2.4.3). Transformation with the pDGB2α2 plasmid served as a negative control. Furthermore, transformed leaves were infiltrated with 400 µM IAA or with a control H_2O:MeOH solution (see 2.4.6). Following incubation, metabolites were extracted from leaf tissue and analyzed by HPLC (see 2.5.4.1 2.5.5.1).

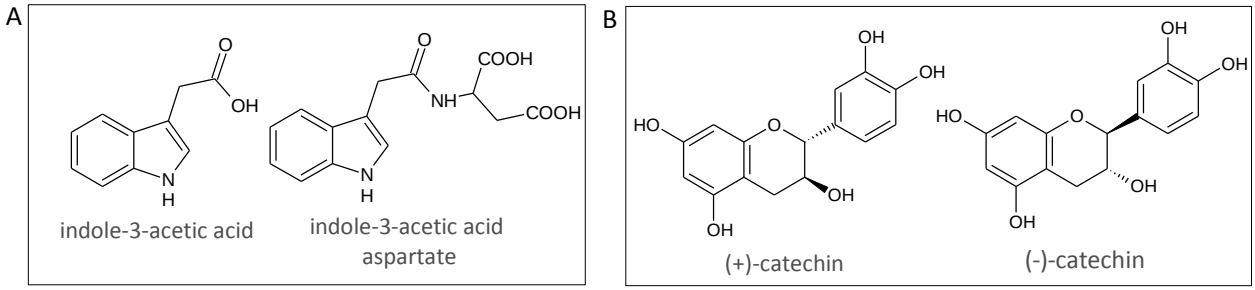

Figure 3.12: Potential substrates for tryptophan halogenases RebH wt and Stth in *Nicotiana sp*. A: IAA-conjugate, indole-3-acetic acid-aspartate. **B:** Enantiomers of the flavonoid, catechin.

Although IAA was considered as an appropriate candidate for activity assays, HPLC analyses of plant extracts revealed no additional metabolites compared to leaves infiltrated with the empty vector or H_2O:MeOH solution. Furthermore, no remaining IAA could be detected in any of the samples, indicating rapid conversion of IAA or halogenated IAA derivatives by regulatory processes within the plant. As a phytohormone, auxin concentrations are strictly regulated within the plant cell and additionally supplemented IAA might be directly converted into a non-active form [148]. Consequently, reduced availability of substrate might contribute to decreased levels of halogenated products. For the final evaluation of IAA processing by RebH wt and Stth to chlorinated derivatives, enzyme activity should be determined in *in vitro* assays. In contrast to a plant-based approach, this method benefits form the absence of competing enzymes and facilitates detection of small product amounts, which might be masked by the plant matrix in leaf extracts.

3.1.4. Functional characterization of mutant tryptophan 7-halogenase, RebH Y455W

The investigated 7-halogenase mutant, RebH Y455W was generated by Glenn *et al.* to predominantly halogenate tryptamine instead of tryptophan [28]. This variation in substrate specificity can be used for biosynthesis of halogenated indole alkaloids originating from tryptamine [28]. In contrast to wild type halogenases, RebH Y455W does not deplete the tryptophan pool and prevents chlorotryptophan

accumulation within the plant [28]. Therefore, RebH Y455W was characterized in more detail in transiently transformed *N. benthamiana* leaves. Since this halogenase converts tryptophan at relatively low rates [28], *rebH Y455W* was co-expressed with the tryptophan decarboxylase gene, *tdc* (fig. 3.1, B13 + TDC). Infiltrations were performed in two biological replicates and transformation of the pDGB2α2 plasmid served as negative control. To evaluate chlorotryptamine accumulation in leaf tissue, metabolites were extracted and analyzed by HPLC (see 2.5.4.1, 2.5.5.1). However, halogenated tryptamine was not detected in plant samples while tryptamine accumulated in high amounts. Therefore, RebH Y455W activity was further investigated in LC/MS analyses (see 2.5.5.3). Thereby, low amounts of 7-chlorotryptamine were detected, which eluted after 8.60 min with an m/z value of 195 in positive ionization mode (fig. 3.13).

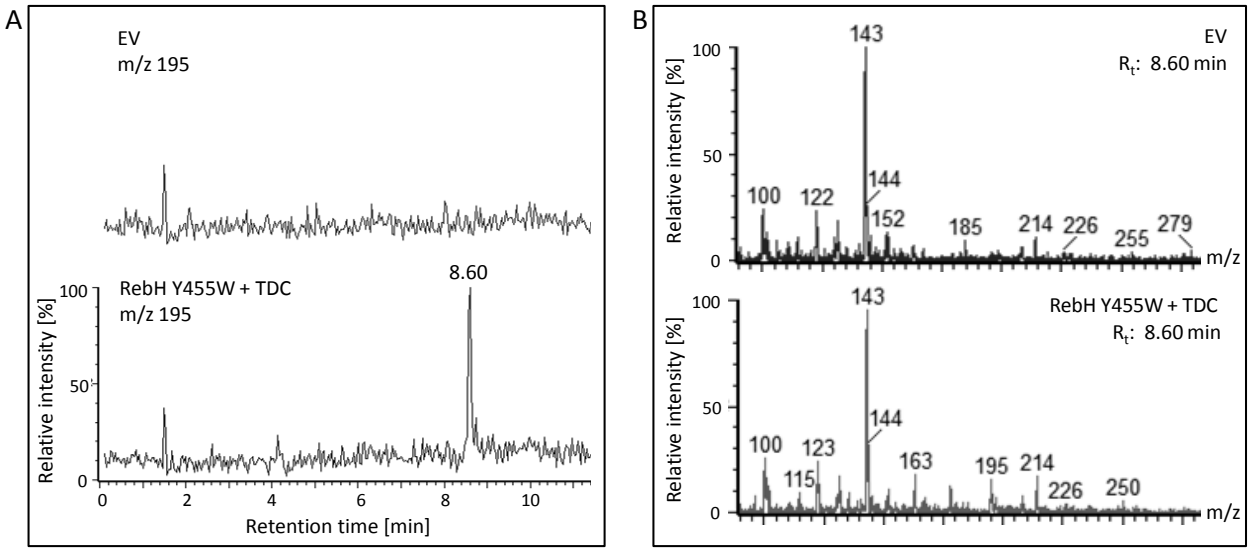

Figure 3.13: Biosynthesis of 7-chlorotryptamine by RebH Y455W in transiently transformed *N. benthamiana* leaves. A: 7-chlorotryptamine eluted after 8.60 min at extremely low concentration in LC-MS analyses. **B:** MS spectrum of the RebH Y445W product, 7-chlorotryptamine (m/z 195).

According to these results, RebH Y455W converts only marginal amounts of tryptamine to the halogenated product. This finding is in agreement with activity assays of Glenn *et al.* who analyzed tryptophan and tryptamine conversion by the wt and mutant enzyme [28]. This approach revealed a conversion rate of less than 5 % of initially supplemented tryptophan by RebH Y455W within 12 h in an *in vitro* assay. In contrast, 10 % of tryptamine was halogenated within the same time-frame by the mutant. However, RebH wt converted 60 % of tryptophan within 12 h and, even more remarkably, 40 %

of tryptamine [28]. This result corroborates the low conversion rate of RebH Y455W compared to the wt enzyme and questions its beneficial use in plant biotechnological approaches.

3.1.4.1. Optimization of 7-chlorotryptmanine biosynthesis by RebH Y455W through anchoring within a protein scaffold

Accumulation of metabolites synthesized by recombinant enzymes can be optimized by channeling of intermediates [19]. This principle was implemented to the halogenation of tryptamine by RebH Y455W, RebF and TDC. Therefore, *N. benthamiana* leaves were transiently transformed with the corresponding genes fused to ligand sequences as well as the TU encoding for protein scaffold (fig. 3.1, SC3 + SC7). Transformation with the pDGB2α2 plasmid was used as a negative control. After verification of sufficient enzyme accumulation, metabolites were extracted from leaf tissue and analyzed by HPLC (see 2.5.4.1, 2.5.5.1). Transient transformation of *N. benthamiana* and analytical studies were performed by Bastian Wagner in course of a scientific internship.

In this regard, HPLC analyses revealed high tryptamine content as well as biosynthesis of a new metabolite in two of four samples exhibiting an R_t value of 9.2 min. According to initial analyses of tryptamine halogenation by RebH wt, this retention time is characteristic for 7-chlorotryptamine (see 3.1.2.4). However, another endogenous tobacco metabolite co-eluted with the chlorinated molecule of interest (tab. 3.2, see 3.1.5). Given that only half of the samples showed an increased metabolite accumulation, final evidence for an enhanced tryptamine halogenation by RebH Y455W through metabolite channeling needs to be provided. Consequently, this transient approach should be repeated in more replicates and analyzed by LC-MS to determine the m/z value of the newly synthesized metabolite. One reason for insufficient metabolite flux could be an unequal ratio of enzymes binding to the protein scaffold. This refers especially to vacant binding domains diminishing metabolite channeling. To exclude these negative effects, protein scaffold biosynthesis could be down-regulated on transcriptional level. Gene expression under regulation of weak promoter-terminator combinations would lead to an excess of binding enzymes and thereby increasing probability of fully packed protein scaffold generation. Moreover, dimerization of enzymes might block binding domains of other enzymes. This applies especially to TDC and RebH Y455W which form homodimers [27, 108]. Optimization of chlorotryptamine biosynthesis could be achieved by reassembly of binding domains with regard to catalytic efficiencies of both enzymes. The relatively high enzymatic activity of one TDC molecule might

be sufficient to supply substrate for several RebH Y455W enzymes featuring low conversion rate. Evaluation of the most efficient stoichiometric composition of each scaffold binding domain can increase product yields dramatically as previously exemplified for the mevalonate biosynthetic pathway by Dueber *et al*. [19].

3.1.5. Summary of metabolite analyses

Characterization of wt and mutant halogenases was performed in diverse approaches and with the use of different instruments. For this reason, the determined R_t values differed in each experiment. A final overview of evaluated R_t and m/z values of all analyzed metabolites is given in tab. 3.2.

Tab. 3.2: Summary of R_t values of metabolites of interest as evaluated by different HPLC and LC-MS approaches.

	HPLC 300SB (2.5.5.1) R_t [min]	HPLC Bonus RP (2.5.5.2) R_t [min]	LC-MS 300SB (2.5.5.3) R_t [min]	LC-MS BEH C18 (2.5.5.4) R_t [min]	LC-MS BEH C18 (2.5.5.5) R_t [min]	LC-MS ES+ m/z
tryptophan	3.1 (s)	2.3 (s)	3.3	2.5	n.a.	205
7-chlorotryptophan	7.9	5.2	7.5	4.7	2.3	239
6-chlorotryptophan	8.3	5.8	7.7	4.8	2.3	239
6-,7-dichlorotryptophan	10.36	n.a.	9.4	5.6	n.a.	273
tryptamine	4.2 (s)	1.3	4.4	3.1	n.a.	161
5-chlorotryptamine	9.4 (s)	n.a.	n.a.	n.a.	n.a.	n.a.
6-chlorotryptamine	9.6	1.7	8.9	5.2	n.a.	195
7-chlorotryptamine	9.3	n.a.	8.6	5.1	n.a.	195
6-bromortryptophan	n.a.	6.4	n.a.	n.a.	2.4	283, 285
7-bromortryptophan	n.a.	6.0	n.a.	n.a.	2.4	283, 285
bromo-chlorotryptophan	n.a.	n.a.	n.a.	n.a.	2.82	319
dibromotryptophan	n.a.	7.3	n.a.	n.a.	2.85	363
halogenated unidentified metabolite (C7)	11.4	n.a.	10.4	6.5	n.a.	325
(C6)	11.6	n.a.	10.7	6.7	n.a.	325
endogenous unidentified metabolite	9.2	n.a.	8.4	5.4	n.a.	291

(s), reference metabolite used as standard; n.a., not available

3.2. Molecular engineering of the indoxyl biosynthetic pathway

3.2.1. Molecular cloning of indoxyl biosynthetic pathway genes

The biosynthesis of halogenated derivatives originating from the indoxyl biosynthetic pathway should be examined in stable and transient transformed *Nicotiana sp.* Therefore, all pathway genes needed to be integrated into the GoldenBraid (GB) cloning system. Additionally, modification of this pathway was achieved through integration of tryptophan halogenase transcriptional units (TUs).

Plasmid DNA of the indoxyl biosynthetic pathway genes, *bx1* and *2A6 L240C/N297Q*, the latter is referred to as *2A6mut* from here on, was used as template for the introduction of GB fusion sites *via* PCR. A synthetic *2A6mut* cDNA was generated within the scope of a doctoral dissertation of Dr. Susan Urbank [154]. For the integration of *bx1* and *2A6mut* into the GB system, endogenous targeting sequences were excluded during gene amplification and final domestication of both CDSs was accomplished by Dr. Agata Staniek. For assembly of transcriptional units in α level plasmids, the 5'-end of *bx1* was fused to the chloroplast targeting sequence, *cp*. Additionally, *2A6mut* was ligated to its endogenous targeting sequence, *2A6targ* (*2A6t*), which was generated by domestication of annealed oligonucleotides (see 2.2.5). Both genes were regulated by a *Cauliflower mosaic virus* 35S promoter (P35S) and a nopaline synthase terminator (TNos).

The newly generated TUs were assembled in an Ω level plasmid (fig. 3.14, I1) and fused to different halogenase/reductase multigene constructs. To accomplish the biosynthesis of chlorinated indole derivatives in chloroplasts, halogenase constructs E9 (*rebH wt-rebF*), G9 (*rebH Y455W-rebF*) and H9 (*stth-rebF*) were fused to the multigene construct I1 resulting in the α level plasmids IE9, IG9 and IH9, respectively (fig. 3.14). Additionally, to investigate indoxyl halogenation in the cytosol, TUs of *bx1* and *2A6mut* were assembled with D13 (*rebH wt-rebF*) and F13 (*stth-rebF*) multigene constructs, resulting in the α level plasmids ID13 and IF13, respectively (fig. 3.14). However, fusion of the rebH Y455W-RebF construct, B13, and I1 could not be achieved. Therefore, *A. tumefaciens* cells, carrying one of both plasmids, were co-infiltrated for transient transformation of *N. benthamiana*.

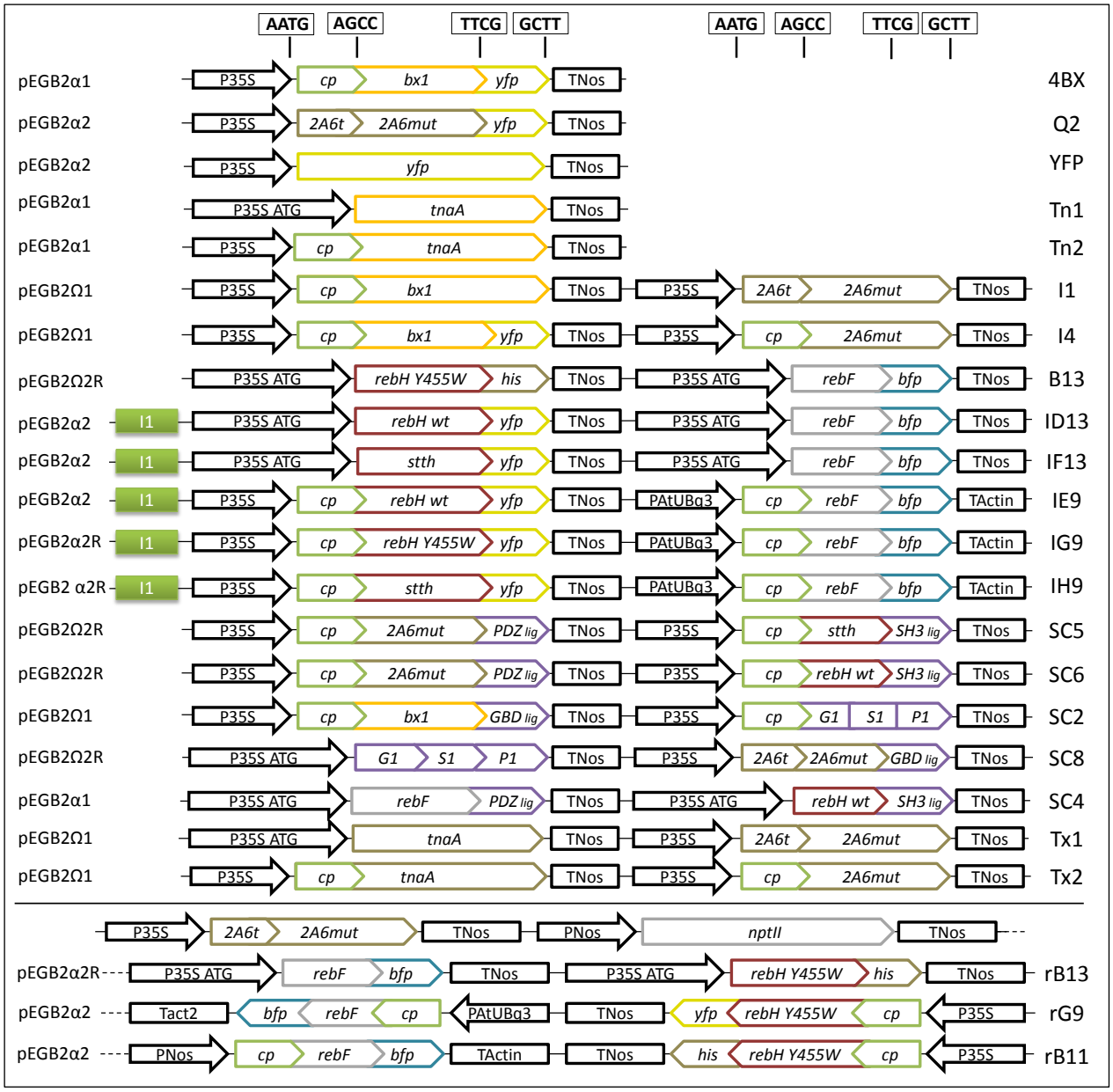

Figure 3.14: GoldenBraid2.0 DNA constructs used for molecular engineering of the indoxyl biosynthetic pathway through introduction of flavin-dependent tryptophan halogenases. Resistance genes of all GB destination vectors can be found in chapter 2.2.6. Abbreviation of each construct is listed on the right.

For verification of correct localization and sufficient accumulation of recombinant enzymes, 3'-ends of *bx1* and *2A6mut* were fused to *yfp* during TU assembly to generate α level constructs 4BX and Q2, respectively.

For stable transformation of transgenic *N. tabacum* BX1#12 plants, the *2A6mut* TU was fused to a *nptII* selection marker gene, which was under control of a nopaline synthase promoter (PNos) and terminator (TNos). Subsequently, this multigene construct was assembled with three different *rebH Y455W-rebF* multigene constructs into α level plasmids. To enable indole halogenation in chloroplasts, G9 (see 3.1.1) was fused to the *2A6-nptII* multigene construct to generate the α level plasmid rG9 (fig. 3.14). Furthermore, to accomplish the halogenation of indoxyl in the cytosol, the *rebh Y455W-rebF* multigene construct, B13 (see 3.1.1) was combined with the TUs of *2A6mut* and *nptII* and the resulting plasmid was designated rB13 (fig. 3.14). To determine the effect of a reduced *rebF* expression on RebH Y455W activity, an additional construct composing of *rebF* under control of a weak promoter-terminator combination was made by fusing *rebF* to nopaline synthase promoter (PNos) and actin2 terminator (Tact2) sequences. This regulation variant leads to 14 times reduced gene expression compared to the standard CaMV 35S promoter and Nos terminator combination [21]. Additionally, this *rebF* TU comprised a chloroplast targeting sequence and was fused to TUs of *cp-rebY455W*, *2A6* and *nptII* leading to the α level plasmid rB11 (fig. 3.14).

For the enhancement of metabolite biosynthesis through binding of recombinant enzymes to a protein scaffold, *bx1*, *2A6mut*, *rebH wt*, *stth* and the $G_1S_1P_1$ protein scaffold CDS were fused to a chloroplast targeting sequence (*cp*). Additionally, 3'-ends of *bx1*, *2A6mut* and halogenase genes were fused to the *GBD*, *PDZ* and *SH3* ligand sequences, respectively. Expression of all CDSs was regulated by CaMV 35S promoter (P35S) and nopaline synthase terminator (TNos). Subsequently, the *2A6mut* TU was assembled with either *rebH wt* or *stth* resulting in the multigene constructs SC6 and SC5, respectively (fig. 3.14). Finally, the TUs of *bx1* and the protein scaffold were combined to Ω level plasmid, SC2 (fig. 3.14). Assembly of transcriptional units and multigene constructs, for anchoring of recombinant enzymes within a scaffold, was primarily accomplished within the scope of a scientific internship by Bastian Wagner.

An alternative to the plant BX1 enzyme for indole biosynthesis is the tryptophanase, TnaA from bacteria, which hydrolyses tryptophan to generate indole, pyruvate and ammonia [155]. To analyze this reaction in a plant-based expression system, the *tnaA* sequence (Gen ID 948221) was amplified from *E. coli* cell lysate and integrated into the GB cloning system after adding appropriate fusion sites and *Bsm*BI recognition sites. For the TU assembly, *tnaA* was fused to the CaMV 35S promoter and nopaline synthase

terminator sequence in the α level plasmid Tn1 (fig. 3.14). Additionally, the Tn2 construct included a chloroplast targeting sequence to validate the effect of chloroplast-localized TnaA on indole biosynthesis (fig. 3.14). Furthermore, different multigene constructs composing of TUs of *tnaA* and *2A6*mut with varying targeting sequences were assembled to facilitate indican biosynthesis *in planta*. In this regard, the multigene construct, Tx1 contained the *tnaA* and *2A6targ-2A6mut* TUs to locate both enzymes in the cytosol (fig. 3.14). A second Ω level plasmid, designated Tx2, composed of both genes fused to a chloroplast targeting sequence (fig. 3.14). Domestication of *tnaA* as well as all following cloning steps were performed by Bastian Wagner within the scope of a bachelor thesis research project [156].

A detailed description of all chronological steps performed during assembly of GB constructs and transformation of *A. tumefaciens* used for transient transformation of *N. benthamiana* is given in chapter 3.1.1

3.2.2. Docking

The probability of RebH wt catalyzed halogenation of molecules originating from the indoxyl biosynthetic pathway was evaluated through docking of molecule structures into the active site of the enzyme using AutoDock 4.2 (see 2.3.2). Reliability of the applied Lamarckian genetic algorithm was verified by an initial docking of the natural substrate tryptophan, which was compared to the localization of the bound tryptophan of the crystal structure of the RebH complex. Thereby, 99 of 100 docking runs resulted in a similar arrangement of the ligand within the active site (fig. 3.15, A). Especially the heteroaromatic ring was perfectly arranged, whereas the localization of amine and carboxylic acid groups varied slightly. Only one calculation led to a wrong prediction of the tryptophan localization. Since this result confirmed the reliability of the applied algorithm, dockings of indole and indican molecules were performed using identical parameters. In a next step, predictions of the different arrangements of each ligand within the active site were again compared to the tryptophan position. In this regard, docking of the indole molecule led to 100 identical predictions that displayed a perfect overlay of the indole molecule and the tryptophan heteroaromatic ring (fig. 3.15, B). However, docking of the indican molecule structure revealed three different clusters of calculated conformations that differed significantly from the tryptophan ligand position (fig. 3.15, C). This result indicates that stereospecific halogenation of indican by RebH wt might be inhibited, but halogenation might occur at other carbon atoms resulting in formation of the desired chloroindican derivatives.

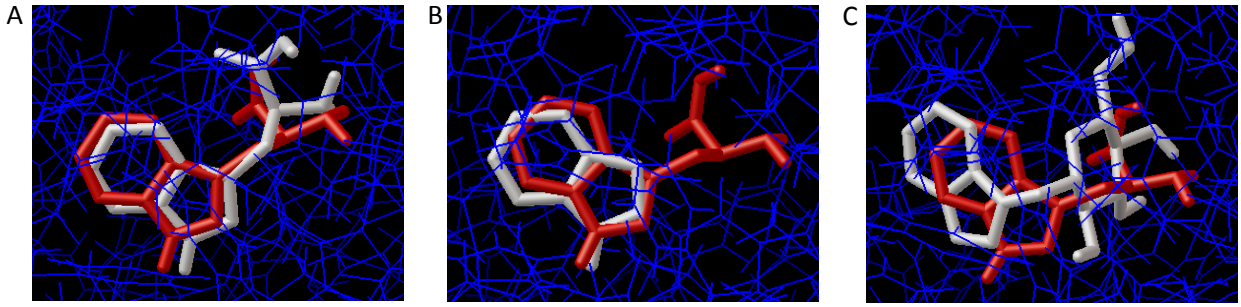

Figure 3.15: Ligand docking of tryptophan (A), indole (B) and indican (C) into the active site of RebH wt. The position of the ligands (red) was compared to the validated localization of tryptophan (grey).

In summary, docking of indole and indican into the active site of RebH wt suggests that indole is the most promising candidate for halogenation. However, docking calculations mimic only simplified conditions of the actual reaction. Even though, predictions led to 100% correct arrangements of indole in the active site, any conclusion has to be analyzed critically. This refers especially to the docking of ligands using ridged amino acid side chains. Therefore, docking of indole needs to be recalculated using flexible residues to get a more reliable result. Moreover, docking results indicate that indican might not be converted by RebH wt. Nevertheless, halogenation of indole might lead to the biosynthesis of the corresponding indican derivatives. To proof this assumption, indole halogenation was analyzed initially in transient approaches by co-expression of indoxyl biosynthetic pathway genes, *rebH wt*, *rebH Y455W* and *stth*.

3.2.3. Biosynthesis of halogenated indole and indican derivatives after transient transformation of *N. benthamiana*

Molecular engineering of the indoxyl biosynthetic pathway was accomplished by transient transformation of *N. benthamiana* with *bx1*, *2A6*, *rebF* and halogenase genes (see 2.4.3). To facilitate chlorination of indole, RebH wt, RebH Y455W or Stth were localized in chloroplasts (fig. 3.14, IE9, IG9, IH9). Moreover, halogenation of indoxyl by RebH wt or Stth in the cytosol was investigated (fig.3.14, ID13, IF13). Additionally, conversion of indoxyl by RebH Y455W was examined after co-transformation of tobacco leaves with *bx1-2A6mut* and *rebH Y455W-rebF* constructs (fig. 3.14, I1 + B13). Even though tryptamine chlorination was catalyzed insufficiently by the halogenase mutant (see 3.1.4), processing of other small molecules could benefit from the fact that these ligands do not compete with the natural substrate tryptophan [28]. Metabolites were extracted from leaf tissue three days after transformation

and analyzed by HPLC (see 2.5.1, 2.5.5.2). This transient approach was accomplished by Kim Röder within the scope of a bachelor thesis research project [157].

Localization of RebH wt and Stth in the cytosol and chloroplasts resulted in the formation of 7- and 6-chlorotryptophan, respectively (fig. 3.16). Determined R_t values of 7-chlorotryptophan (R_t 5.1 min) and 6-chlorotryptophan (R_t 5.8 min) correlated with the results of previous analyses (tab. 3.2). However, concentrations of both metabolites were moderate compared to initial characterization of halogenases, which probably resulted from a shortened incubation period (see 3.1.2.1). Interestingly, small amounts of 7-chlorotryptophan could be detected in case of chloroplast localized RebH Y455W, even though, this enzyme is supposed to convert tryptophan at low rate [28]. However, no halogenated indole was detectable in any of the samples (tab 3.3). Moreover, indole was not clearly distinguishable from tobacco metabolites, probably due to its low concentration within the plant tissue. Likewise, indican was detectable in no more than two samples of cytosolic localized RebH wt and Stth exhibiting an R_t value of 3.5 min (fig. 3.16, A). Nevertheless, two metabolites, absent in untreated samples, could be determined after transformation of halogenase genes. The first newly synthesized molecule eluted after 4.2 min and displayed UV maxima at 230 nm and 286 nm (fig. 3.16). This molecule certainly features some polar groups due to its early elution between indican and chlorotryptophan. The second metabolite associated with halogenase activity exhibited an R_t value of 13.3 min suggesting a non-polar molecule (fig. 3.16). The determined UV maxima were 236 nm and 281 nm. Even though these metabolites displayed UV spectra comparable to halogenated indole derivatives (tab 3.3), it is to be questioned if the biosynthesis of these molecules resulted from tryptophan halogenase activity. The R_t values of the newly synthesized metabolites did not differ in terms of a putative 7- or 6-halogenation. Slightly distinct retention times were observed for all investigated metabolites halogenated by either RebH wt or Stth, such as chloroindole, chlorotryptophan and chlorotryptamine (tab. 3.2, 3.3). Therefore, these compounds could be endogenous metabolites that might be synthesized as a reaction to *Agrobacterium* infection. For further characterization, molecular masses of the investigated compounds could be determined by mass spectrometry.

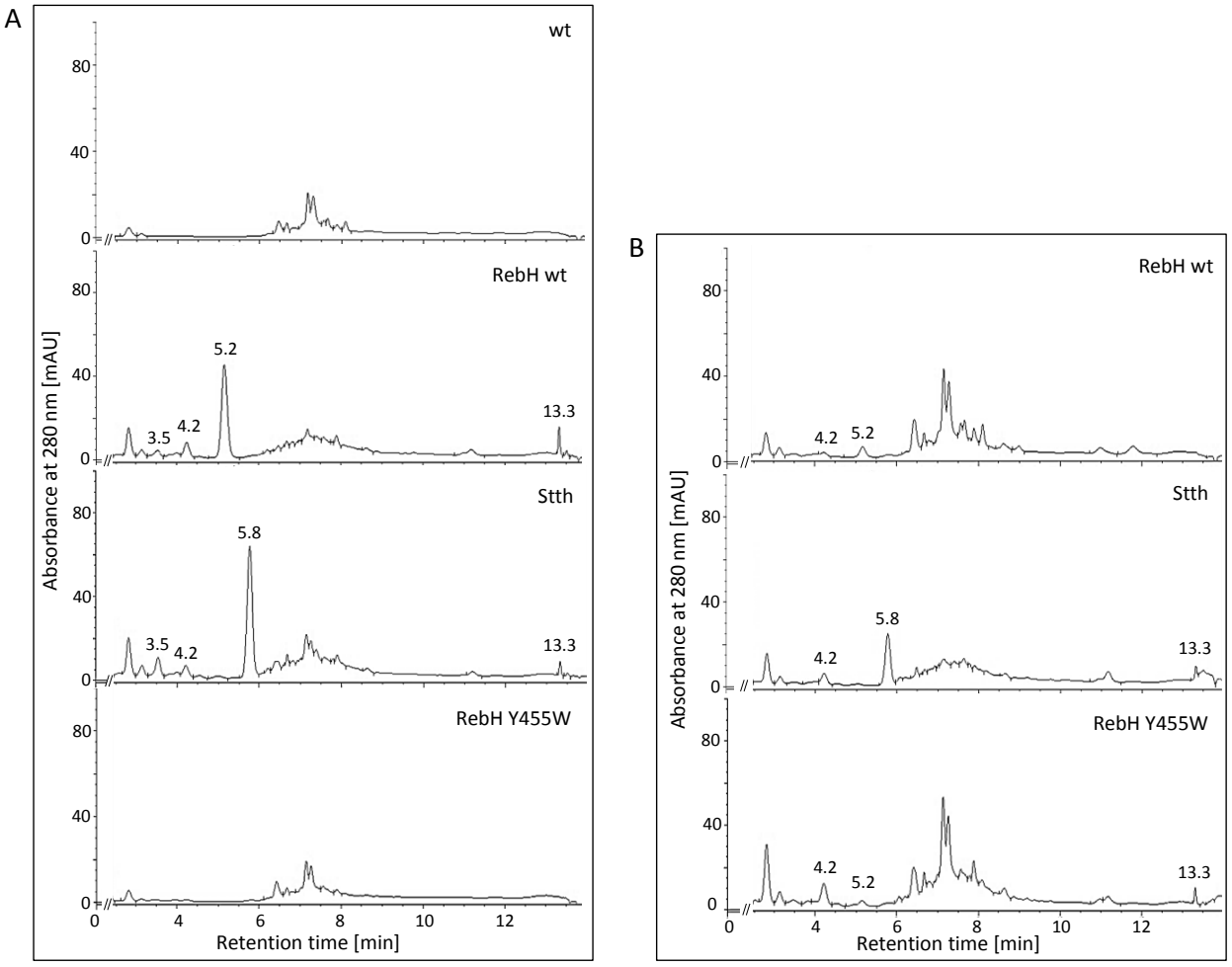

Figure 3.16: HPLC analyses of metabolites from leaf extracts after construction of the indoxyl biosynthetic pathway and co-localization of tryptophan halogenases. A: Cytosolic localization of halogenases correlated with biosynthesis of two new metabolites (R_t 4.2 min, 13.3 min), 7- and 6-chlorotryptophan (R_t 5.2 min, 5.8 min) as well as indican (R_t 3.5 min) compared to the wild type (wt) sample. **B:** Localization of tryptophan halogenases in chloroplasts resulted in the formation of chlorotryptophan and two new metabolites, whereas indican did not accumulate in detectable amounts.

The biosynthesis of halogenated indole and indican derivatives was not observed in this transient approach. Moreover, the concentration of substrate molecules such as indole and indican was relatively low three days after infiltration. This observation suggests a less efficient substrate conversion by either BX1 or 2A6mut in contrast to wild type halogenases. Consequently, the incubation period needs to be increased for sufficient accumulation of indican within the cells. Moreover, BX1 and 2A6mut could be affected by secondary metabolites in tobacco -like nicotine- resulting in a reduced catalytic efficiency [73]. Given that both wild type halogenases were active in transiently transformed plant tissue in terms

of chlorotryptophan biosynthesis, a general adaption of the incubation time for infiltrated *N. benthamiana* should be taken into consideration. The onset of necrosis caused by either accumulation of toxic end products or high protein biosynthesis after overexpression of transgenes might be reduced by these measures. In fact, a short incubation time of only three days might be insufficient in case of a low catalytic capacity of recombinant enzymes or limited availability of substrate.

Tab. 3.3: Determined R_t values and UV maxima of reference metabolites in HPLC analyses (see 2.5.5.2).

reference metabolites	R_t value [min]	UV max [nm]
indole	7.9	220, 280
5-chloroindole	9.7	222, 278
6-chloroindole	9.6	222, 278
7-chloroindole	9.2	218, 271
indican	3.5	220, 280

Biosynthesis of halogenated metabolites could not be detected in HPLC analyses. Also the prediction of R_t values in HPLC was challenging and chlorinated indican standards were not available for localization of putative product peaks. Therefore, some samples were selected for thin-layer chromatography (TLC) (see 2.5.3). In contrast to HPLC, TLC enables additional treatments of the separated compounds, like deglycosylation and formation of pigments, leading to visualization of components. The TLC analyses were partially performed within the scope of a bachelor thesis research project by Kim Röder [157].

The separation of an indican standard revealed an R_f value of 0.58, whereas indole migrated in the solvent front (fig. 3.17). For this reason the R_f value of indole was not determined in the approach. Halogenated indigo derivatives, indicating the presence of halogenated indican, were not detected by TLC in any of the samples after deglycosylation. Similar to the HPLC analyses, accumulation of indican was not observed within the analyzed leaf extracts. A repetition of this transient approach revealed two samples with indigo levels sufficient for TLC detection (fig. 3.17). This refers to co-transformation of the I1 and ID13 constructs comprising *bx1*, *2a6, rebH wt* and *rebF* (fig. 3.14). Two different *A. tumefaciens* strains (AGL1 and EHA105) were compared in this approach regarding the efficiency of *N. benthamiana*

transformation. The results of these analyses were described in more detail by Kim Röder [157] and are not discussed in this context.

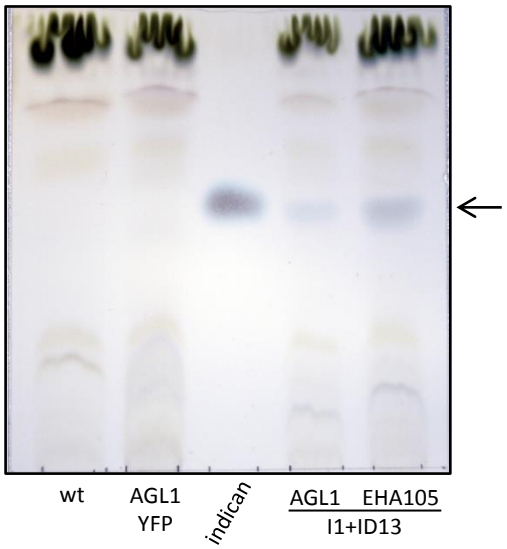

Figure 3.17: TLC analysis of indican biosynthesis after transient transformation *of N. benthamiana*. Indigo could be detected after deglucosylation of indican in samples co-transformed with the I1 and ID13 constructs compared to a wild type plant (wt) and a control infiltration (YFP); (S) 5 µg indican standard. TLC analyses were performed by Kim Röder during a bachelor thesis research project [157].

The low biosynthesis rate of indole and indican after transient expression of *bx1* and *2A6mut* indicates an inefficient substrate conversion by recombinant enzymes. Especially formation indole was not observed in HPLC analyses. Given that chlorination of this metabolite by RebH wt was observed in *in vitro* assays, the low availability of substrate in a transient approach might negatively affect the conversion rate of RebH wt and Stth [6]. However, the RebH wt catalyzed halogenation of indole occurred at the electronically most activated position, C2 [6]. Consequently, further oxidation by 2A6mut is inhibited. However, Stth and RebH Y455W might halogenate indole at other positions enabling subsequent conversion by 2A6mut. Moreover, synthesized chloroindoles could be detected by HPLC and their absence might result from an inefficient indole biosynthesis by BX1 or diffusion of indole through cell membranes [158]. To increase the efficiency of indican biosynthesis and circumvent loss of indole molecules by diffusion, 2A6mut and BX1 can be co-localized in chloroplast.

3.2.4. Co-localization of 2A6mut and BX1 in chloroplasts of transiently transformed *N. benthamiana* leaves

One reason for the inefficient biosynthesis of indican in transient approaches could be a low availability of indole within the cells. Due to its small size and lipophilic nature, indole easily diffuses through cell membranes [158]. Moreover, a low enzymatic efficiency or gene silencing of the overexpressed *bx1* gene

could decrease the indole biosynthesis rate [131]. To circumvent this loss of precursor molecules and compensate inefficient biosynthesis, 2A6mut was targeted to chloroplasts to directly oxidize indole at the site of production and thereby increase indoxyl concentrations. In addition, absence of glycosyltransferases in chloroplasts might result in dimerization of indoxyl to indigo [159]. However, one basic requirement for substrate oxidation by cytochrome P450 is the availability of electrons, which are transferred by NADPH-cytochrome P450 oxidoreductases anchored within the ER membrane. These enzymes interact with several cytochrome P450s and probably contribute to the transfer of electrons to the recombinant 2A6mut enzyme [41]. In chloroplasts, electrons need to be supplied by other mechanisms as recently demonstrated by Nielsen *et al.* in transiently transformed *N. benthamiana* leaves [160]. After translocation of CYP79A1, CYP71E1 and an UDP-glucosyltransferase from *Sorghum bicolor* into tobacco chloroplasts, the final product, dhurrin, was synthesized in a light-dependent manner in isolated thylakoids [160]. This indicates the direct transfer of electrons from the photosystem I (PSI) to ferredoxin and further to both cytochrome P450s [160]. However, this system was established for plant cytochrome P450s containing an endogenous signal peptide, which is responsible for anchoring within the ER membrane [160].

In this regard, electron availability and interaction of 2A6mut with PSI in chloroplasts was evaluated after transient gene expression (see 2.4.3). Therefore, *N. benthamiana* plants were transformed with the *2A6mut* and *bx1* genes each fused to a chloroplast targeting sequence (fig. 3.14, I4). Moreover, indican biosynthesis was compared to 2A6mut anchored within the ER membrane and chloroplastic localized BX1 (fig. 3.14, I1). Given that indole is also oxidized to a certain degree by endogenous cytochromes in tobacco, *N. benthamiana* leaves were additionally transformed with *bx1* as reference (fig. 3.14, 4BX). All transformations were performed in two biological replicates and metabolites were analyzed by HPLC (see 2.5.4.1, 2.5.5.2). This analytical approach was performed by Bastian Wagner in course a scientific internship.

With the exception of one sample, all replicates revealed a low indican content in leaf tissue. Moreover, 2A6mut activity did not differ in terms of a chloroplastic localization or anchoring within the ER membrane. Contrary to expectations, indican levels of 2A6mut samples were similar to those producing BX1 only. Therefore, indoxyl biosynthesis might rather be catalyzed by an endogenous cytochrome P450 than by recombinant 2A6mut. Co-localization of both enzymes was thought to enhance the oxidation of

indole in chloroplasts by circumventing substrate loss through transport across the chloroplast envelope or diffusion into the intercellular space. However, this pathway engineering did not yield in an improved indican biosynthesis. One reason for this result might be the low availability of precursor molecules, such as indole-3-glycerol phosphate, which is converted by both BX1 and the tryptophan synthase [68, 69]. The biosynthesis of this amino acid is feedback regulated by inhibition of the anthranilate synthase which catalyzes the first step in this biosynthetic pathway [161]. Therefore, tryptophan accumulation can directly influence the biosynthesis of precursor molecules of the indole and indican biosynthetic pathway. Another reason for the inefficient biosynthesis rate of indican might be a low catalytic activity of 2A6mut in chloroplasts. Cytochrome P450s require electrons, which are provided by co-localized NADPH-cytochrome P450 oxidoreductases [41]. Thereby, electrons are transferred from NADPH through FAD and FMN-cofactors to the heme-group of the P450s [41]. Therefore, 2A6mut requires additional electron sources in chloroplasts due to absent NADPH-cytochrome P450 oxidoreductases [41]. This electrons can be provided by the electron transport chain of the photosystem I (PSI) as demonstrated by Nielsen *et al.* [160]. However, the translocation of electrons depends on the co-localization of heterologous P450s in the chloroplast membrane. Therefore, loss of 2A6mut activity might be associated with inefficient electron transport from PSI or other reductases.

3.2.5. Optimization of indican biosynthesis through metabolite channeling

Anchoring of recombinant enzymes within a scaffold potentially increases the biosynthesis of final products through effective metabolite channeling between pathway enzymes as exemplified by Dueber *et al.* in *E. coli* [19]. In this regard, BX1, 2A6mut and tryptophan halogenases were attached to a protein scaffold in chloroplasts or the cytosol to analyze the effect of this pathway optimization on indican biosynthesis. Therefore, *N. benthamiana* was co-transformed with the *bx1-*$G_1S_1P_1$ construct (fig. 3.14, SC2) as well as *2A6mut* and either *rebH wt* (fig. 3.14, SC6) or *stth* (fig. 3.14, SC5). Indican biosynthesis was compared to the non-scaffolded pathway localized in chloroplasts. In this context *N. benthamiana* was co-transformed with *bx1*, *2A6mut*, *rebF* and either *rebH wt* (fig. 3.14, I1 + E9) or *stth* (fig. 3.14, I1 + H9). Additionally, RebH wt and RebF were linked to the protein scaffold in the cytosol in combination with 2A6mut, which was additionally anchored within the ER membrane (fig. 3.14, 4BX + SC4 + SC8). Further, indican biosynthesis was compared to the non-scaffolded pathway (fig. 3.14, I1 + D13). Following extraction, metabolites were analyzed by HPLC (see 2.5.4.1, 2.5.5.2). These analytical experiments were performed by Bastian Wagner within the scope of a scientific internship.

The anchoring of BX1 and 2A6mut together with RebH wt in the cytosol resulted in biosynthesis of 7-chlorotryptophan (R_t 5.17 min). Moreover, 6-chlorotryptophan biosynthesis was observed in case of chloroplast localized Stth bound to the protein scaffold. In contrast, 7-chlorotryptophan was not detected after scaffold attachment of RebH wt in chloroplasts, indicating inhibition of the enzyme (fig. 3.18). Besides chlorotryptophan derivatives, small amounts of indican could be detected in samples transformed with *bx1* and *2A6mut*. However, concentrations did not differ from controls of the non-scaffolded pathway. Furthermore, biosynthesis of chloroindole and chloroindican by RebH wt or Stth was not observed after anchoring of recombinant enzymes within a protein scaffold either in the cytosol or chloroplasts.

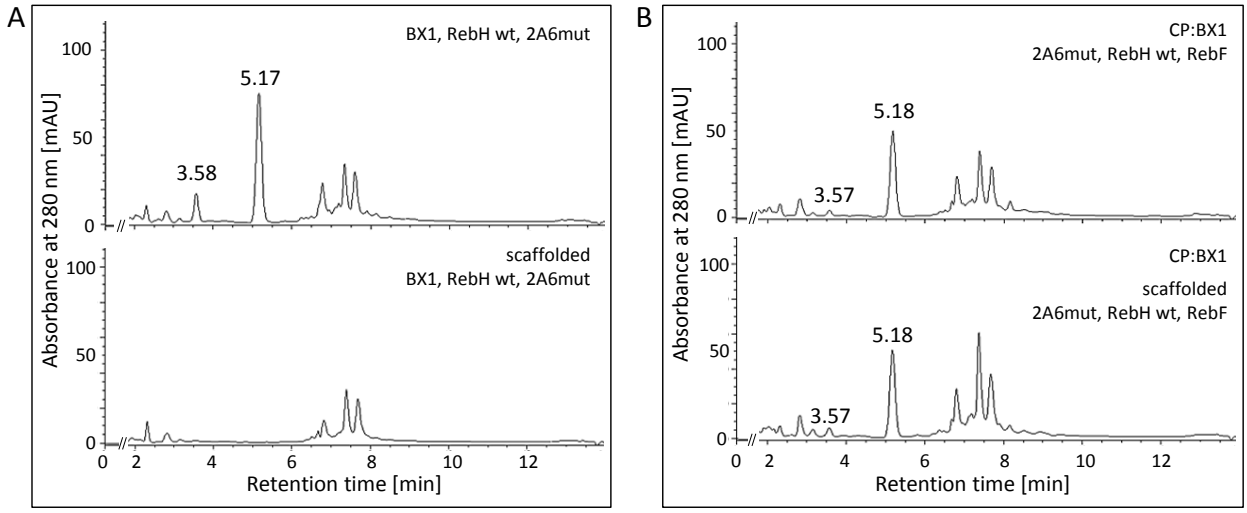

Figure 3.18: Anchoring of RebH wt within the protein scaffold resulted in an abolished enzymatic activity in chloroplasts. A: HPLC analyses revealed the inhibition of RebH wt in chloroplasts after attachment to the protein scaffold in terms of 7-chlorotryptophan biosynthesis (R_t 5.18 min) **B:** HPLC analyses of cytosolic localized RebH wt revealed no changes of enzyme activity by anchoring within the protein scaffold compared to control samples. Besides indican (R_t 3.57 min) and 7-chlorotryptophan, accumulation of additional metabolites was not observed. CP: chloroplast signal peptide

Although linkage of recombinant enzymes to a protein scaffold seems to be a highly promising tool to increase biosynthesis of desired metabolites, this principle could not be adapted to the indoxyl biosynthetic pathway with the chosen scaffold and enzyme combinations. Moreover, anchoring of RebH wt within the scaffold resulted in the entire loss of enzymatic activity in chloroplasts. Interestingly, RebH wt was still active in the cytosol after attachment to the corresponding binding domain. To rule out any mutation, which might have occurred during the cloning procedure, the halogenase CDS

incorporated into the α level plasmid was confirmed by sequencing. Loss of RebH wt activity was reported previously by Payne *et al.* after fusing the enzyme to a maltose-binding protein [6]. Therefore, C-terminal fusion of RebH wt seems to be a possible reason for its abolished activity. However, combination of the halogenase and YFP had no effect on tryptophan conversion efficiency. Moreover, biosynthesis of 7-chlorotryptophan by anchored RebH wt was not affected in the cytosol. One reason for these findings could be an inefficient binding of RebH wt to the protein scaffold due to the formation of homodimers [27]. Moreover, 2A6mut, which is anchored within the ER membrane, might place the protein scaffold at an unfavorable position for ligand interactions. Consequently, enzymatic activity of non-scaffolded RebH wt would be maintained. Another reason for 7-chlorotryptophan biosynthesis by the attached RebH wt in the cytosol could be an efficient transfer of reduced flavin by the co-localized reductase RebF. This enzyme was not linked to the scaffold in chloroplasts, because initial characterization of RebH wt had demonstrated the sufficient $FADH_2$ supply by endogenous enzymes in this compartment (see 3.1.2.2). If the binding of $FADH_2$ is negatively affected by linkage of RebH wt to the scaffold, co-localization of a flavin reductase might be required to maintain the enzymatic activity. Finally, activity loss of RebH wt in chloroplasts could be associated with erroneous protein folding due to the C-terminal fusion to the SH3 ligand. Proteins carrying a chloroplast transit peptide are synthesized by free ribosomes in the cytosol [162]. To facilitate the translocation of these proteins through both chloroplast membranes, they are kept in an unfolded structure by chaperones [162]. After transport into the stroma, proteins fold to form their tertiary structure [162]. Even though this process in assisted by chaperones, incorrect folding might occur due to the foreign ligand peptide [162].

In summary, construction of the indoxyl biosynthetic pathway did not result in sufficient biosynthesis of indole and indican after transient transformation of *N. benthamiana* leaves. Moreover, optimization of metabolite flux by translocation of 2A6mut into chloroplasts and anchoring of all pathway enzymes within a scaffold did not result in a sufficient biosynthesis of indole, indican or halogenated derivatives. Finally, the low availability of substrate resulted in an inefficient chlorination of indole and indoxyl by tryptophan halogenases.

Since indican accumulation was demonstrated previously in transgenic *N. tabacum* plants [14], molecular engineering of the indoxyl biosynthetic pathway by introduction of tryptophan halogenases was further investigated by constitutive gene expression *in planta*.

3.2.6. Biosynthesis of halogenated indole and indoxyl derivatives in transgenic *N. tabacum* plants

For the constitutive biosynthesis of halogenated indole, indoxyl and indican derivatives, transgenic *N. tabacum* lines were generated by *Agrobacterium*-mediated gene transfer (see 2.4.1). Therefore, *N. tabacum* BX1#12 plants were transformed with *2A6*, *rebH Y455W* and *rebF*. Additionally, the neomycin phosphotransferase encoded by *nptII*, was used as a selection marker. Although a low catalytic activity of RebH Y445W was proven previously in transient approaches (see 3.1.2.1), this enzyme was selected to avoid toxic effects on the tryptophan metabolism through constitutive gene expression. Furthermore, the impact of RebH Y455W compartmentalization was determined by transformation of three different constructs leading to cytosolic (fig. 3.14, rB13) or chloroplastic (fig. 3.14, rG9, rB11) enzyme localization. In this context, the effect on a decreased reductase gene expression in chloroplasts was evaluated by fusion of *rebF* to either a weak promoter-terminator combination (rB11) or a moderate one (rG9) [21]. In a first selection process, calli and plants were cultivated on growth medium supplemented with kanamycin and phosphinothricin. After regeneration of whole plants, promising lines were selected based on the indican content in leaves, indicating a high 2A6mut activity, by HPLC (see 2.5.5.2).

3.2.6.1. Characterization of transgenic *N. tabacum* rB13, rG9 and rB11 plants

Following transformation, a total of 27 plants from all three events were recovered from tissue culture. Metabolites were analyzed by HPLC and R_t values were compared to indole, indican and chloroindole reference metabolites (tab. 3.3.). In this context, seven rB13 plants emerging from six different calli were analyzed in terms of indican biosynthesis. While all of those revealed high indican concentrations, halogenated indole or indican derivatives were not detected in any individual plant (fig. 3.19). The same applied for chloroplast localization of RebH Y455W. Here, 13 plants deriving from six different calli were successfully regenerated after transformation of the rG9 construct. Nine of those exhibited high indican content, but no halogenated derivatives (fig. 3.19). Moreover, transformation of the rB11 construct resulted in successful recovery of nine plants. However, only three of those produced indican at high levels (fig. 3.19). Based on previous results regarding RebH wt activity in chloroplast without a partner reductase, RebF, transgenic lines rB11 and rG9 are supposed to feature the same halogenase activity (see 3.1.2.2). Therefore, reduced *rebF* expression level in rB11 plants might not influence catalytic

activity of the halogenase. Since halogenated metabolites were not observed in HPLC analyses, some rG9, rB11 and rB13 plants were analyzed by more sensitive LC-MS (see 2.5.5.6). In this regard, leaf extracts were screened for halogenated indole and indican derivatives and compared to *N. tabacum* wt control samples. However, the expected halogenated metabolites were not present in detectable amounts in the investigated samples.

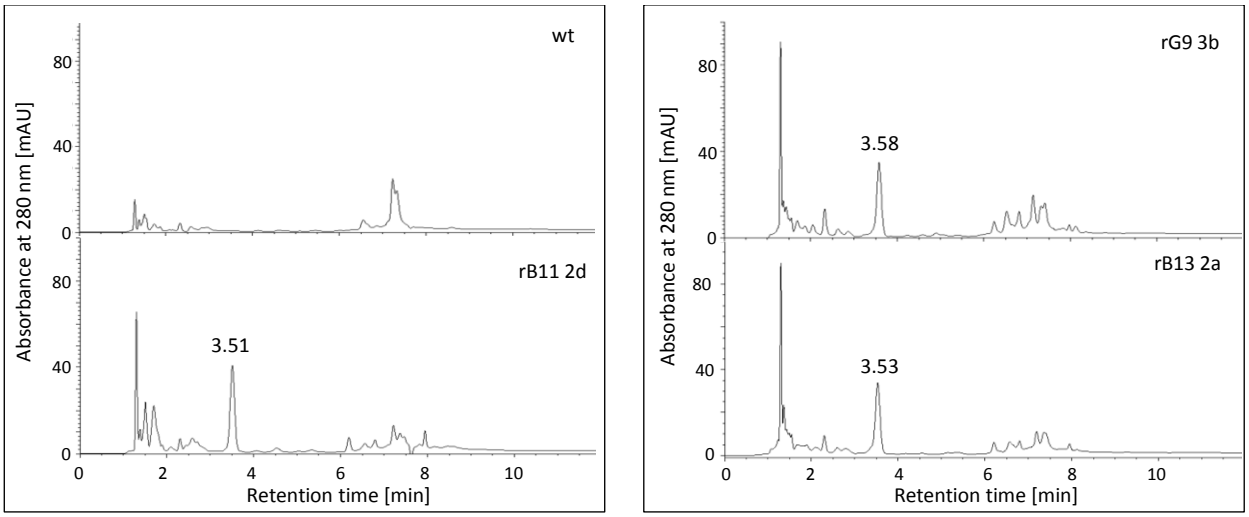

Figure 3.19: Biosynthesis of indican in transgenic *N. tabacum* plants. All synthesized metabolites were surveyed for an indole-like spectrum and compared to standard metabolites. Besides indican biosynthesis (R_t 3.5 min), formation of additional, probably halogenated metabolites was not observed.

The relatively small number of plants of each line, which could be regenerated, indicates a negative effect of gene overexpression on the plant viability. Especially halogenated indole derivatives were shown to have toxic effects on organisms which was exemplified by Wu *et al.* in *E. coli* cell cultures [75]. Therefore, successful cell growth and plant regeneration might be inhibited by accumulation of toxic metabolites. As already discussed above, halogenation of indole at C3 might also inhibit subsequent oxidation by 2A6mut and consequently results in an inefficient indican biosynthesis. However, formation of 2-chloroindole could not be detected in LC-MS analyses, demonstrating the low catalytic efficiency of RebH Y455W. This is consistent with initial characterization of RebH Y455W *in planta* (see 3.1.4) leading to the conclusion that this halogenase mutant might be not suitable for application in plant synthetic biology.

Although additional metabolites, which could be indicating formation of halogenated indole, indoxyl or indican derivatives, were not detected by HPLC and LC-MS, plant extracts of several transgenic plants were analyzed by thin-layer chromatography (see 2.5.3). Therefore, metabolites of three biological samples of *N. tabacum* wt as well as from F1 generation of BX1#12, BX1/2A6#2-13 and rB11 2a were extracted and analyzed by TLC using two different solvent mixtures for separation of either indole or indican derivatives (see 2.5.3). Moreover, 5 µg of indican and 2 µg of 7-chloroindole served as positive controls and R_f values were calculated (formula (1)). Separation of plant extracts using a mobile phase composing of chloroform and ethyl acetate revealed an R_f value of 0.91 for 7-chloroindole (fig 3.20, metabolite a) whereas halogenated indole derivatives were not observed in rB11#2a plant extracts.

The separation of indican and its derivatives using chloroform, methanol and water, revealed an R_f value of 0.53 for the indican standard (fig. 3.20, metabolite b). Moreover, this metabolite could be visualized in relatively high amounts in extracts of 2A6/BX12-13 compared to BX1#12 and rB11 2a samples. The accumulation of indican in 2A6/BX1 plants in concentrations easily detectable by TLC was already demonstrated by Warzecha *et al.* [14]. According to initial HPLC analyses, rB11 2a plants exhibited lower indican levels, visualized as faint blue line on the TLC plate (fig. 3.20). This refers especially to rB11 2a#3 revealing an indican content comparable to BX1#12 plants. This individual might represent the genotype of BX1#12, which was used for the generation of the transgenic line rB11. However, chloroindican derivatives, which should exhibit similar R_f values to indican were not observed in rB11 2a leaf extracts.

Finally, three additional metabolites were present in some individual plants. This refers to metabolite (b) extracted from rB11 2a #1, which exhibited an R_f value of 0.82. Furthermore, biosynthesis of a reddish colored molecule (c) was observed in rB11 2a F1 #2 and #3 extracts (fig. 3.20). This metabolite revealed an R_f value of 0.88 and 0.87, respectively, which is similar to the one of 7-chloroindole (R_f 0.91). Additionally, biosynthesis of a third, more hydrophilic metabolite (d) with an R_f value of 0.12 was observed in extracts of rB11 2a #2 and #3 (fig. 3.20). In contrast to indican and indole, these newly synthesized metabolites featured no fluorescence after excitation at 366 nm.

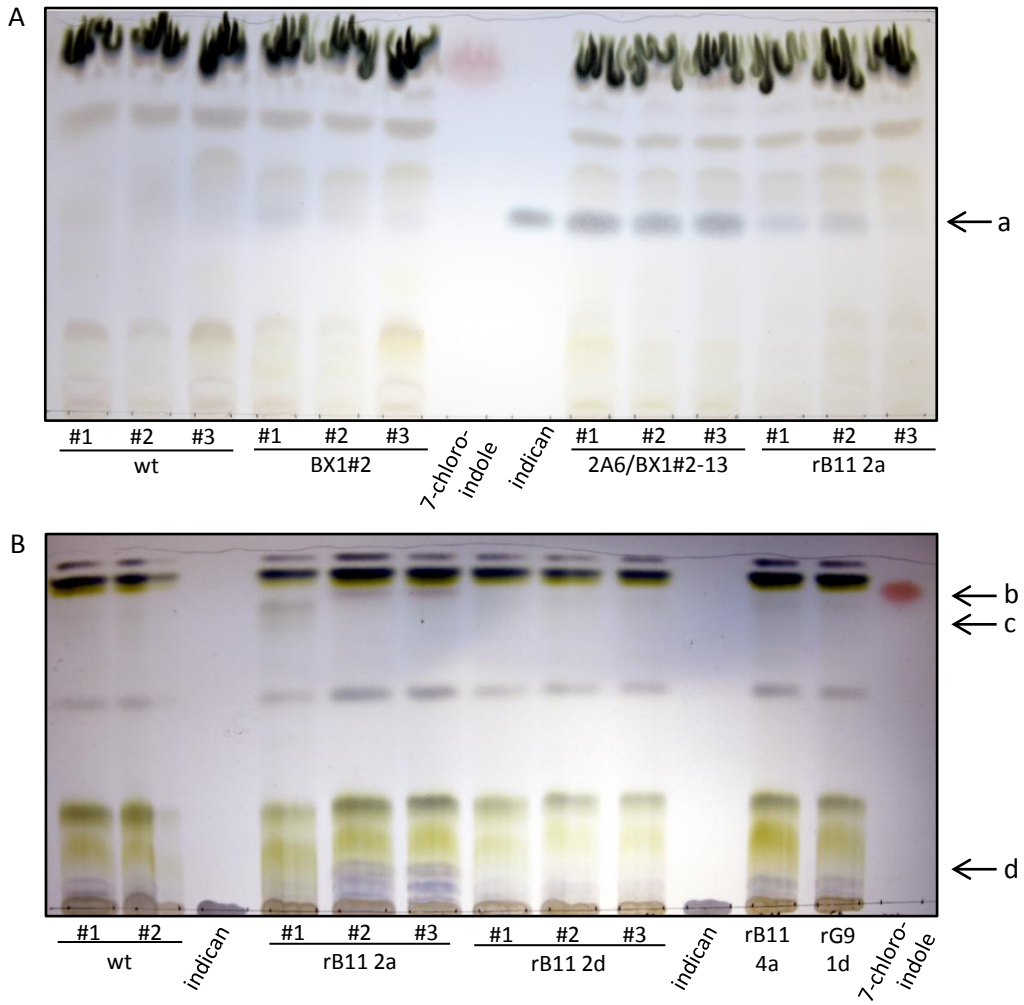

Figure 3.20: Thin-layer chromatography of transgenic *N. tabacum* plants for identification of halogenated indican (A) and indole (B) derivatives. Indican biosynthesis was verified in rB11 2a plants (a). Moreover, three additional pigments (b, c, d) were synthesized in this transgenic line.

In summary, three different transgenic tobacco lines expressing genes of the indoxyl biosynthetic pathway as well as *rebH Y455W* and *rebF* were successfully generated. According to previous analyses (see 3.1.4), characterization of these lines by HPLC revealed a low catalytic activity of the halogenase mutant. Moreover, this enzyme was not beneficial in terms of halogenation of alternative substrates. Nevertheless, formation of two new metabolites was observed by TLC, which were not detected by HPLC. Since all generated transgenic plants were first analyzed by HPLC, further characterization by TLC might reveal more pigments accumulating within the plants. For further characterization of the detected blue and reddish compounds in line rB11 2a, molecules can be extracted from TLC plates to analyzed UV spectra or to identify functional groups. Moreover, substitution of halogen atoms could be verified by

mass spectrometry and determination of the molecular mass could help to identify the individual compounds. However, for the final clarification of molecule structures, metabolites need to be analyzed by NMR.

3.2.6.2. Chloroindican biosynthesis by 2A6mut after supplementation of chlorinated indole derivatives to transgenic *N. tabacum* leaves

Since halogenated indole and indican derivatives were not detected in transgenic RebH Y455W plants, it needs to be evaluated if these metabolites are stable *in planta* or might be degraded or further metabolized. Therefore, chloroindole was supplemented to different transgenic tobacco plants to analyze the biosynthesis of chloroindican by 2A6mut. Also, R_t values should be determined to simplify metabolite identification in HPLC analyses. In this regard, leaves of *N. tabacum* wt, BX1#12, 2A6/BX1#2-13 and rB1 2a were infiltrated with 400 µM 5-, 6- and 7-chloroindole (see 2.4.6). Injection of a mixture of water and methanol served as negative controls. Following incubation, metabolites were extracted from leaf tissue and investigated by HPLC (see 2.5.5.2). The biosynthesis of halogenated indican by oxidation of chloroindole was anticipated exclusively for the 2A6 double mutant present in rB11 plants. Accordingly, recombinant 2A6 (BX1/2A6#2-13 plants) should not convert chloroindole to chloroindoxyl. As already shown for *N. tabacum* BX1#12 plants, oxidation of indole by an endogenous cytochrome P450 leads to formation of indican leaf tissue. In this regard, conversion of halogenated derivatives by endogenous enzymes is hardly predictable.

Interestingly, chloroindole derivatives were detected exclusively in BX1/2A6#2-13 and rB11 2a plants after infiltration. In addition, some samples revealed the biosynthesis of new metabolites in comparison to H_2O:MeOH control infiltrations. In this regard, 6-chloroindole infiltration resulted in biosynthesis of 6-chlorotryptophan (R_t 5.8 min) in wt and transgenic plants suggesting that this metabolite was synthesized by an endogenous enzyme (fig. 3.21). However, biosynthesis of 7-chlorotryptophan with an expected R_t value of 5.2 min was not observed after infiltration of 7-chloroindole. It is unclear if 5-chlorotryptophan was synthesized from 5-chloroindole as the R_t value for this metabolite was not determined before. Nevertheless, accumulation of two additional metabolites correlated with 5-chloroindole infiltration. These molecules could be detected in *N. tabacum* wt as well as BX1#12 plants and exhibited R_t values of 5.3 min and 5.4 min, respectively. These findings suggest that an endogenous enzyme, most likely the tryptophan synthase (TS), catalyzes the biosynthesis of chlorotryptophan from

chloroindole. TS forms heterotetramers composing of two α- and two β-subunits [69]. In a first reaction step indole-3-glycerolphosphate is converted to indole by the α-subunit. Subsequently, tryptophan is synthesized from the indole intermediate, which is never released due to metabolite channeling between both subunits [69]. The biosynthesis of chlorotryptophan from infiltrated chloroindole implies transport of this molecule through the α-unit to be directly converted by the β-subunit. Until today, this fact was not described for plant-derived TS. Moreover, it can be speculated if indole, synthesized by BX1 in transgenic tobacco, is converted by TS accordingly. This assumption could explain the relatively low amount of indole within leaf tissue of transgenic plants (see 3.2.6.1). However, formation of additional metabolites, indicating the biosynthesis of chloroindican, was not observed in rB11 leaf extracts.

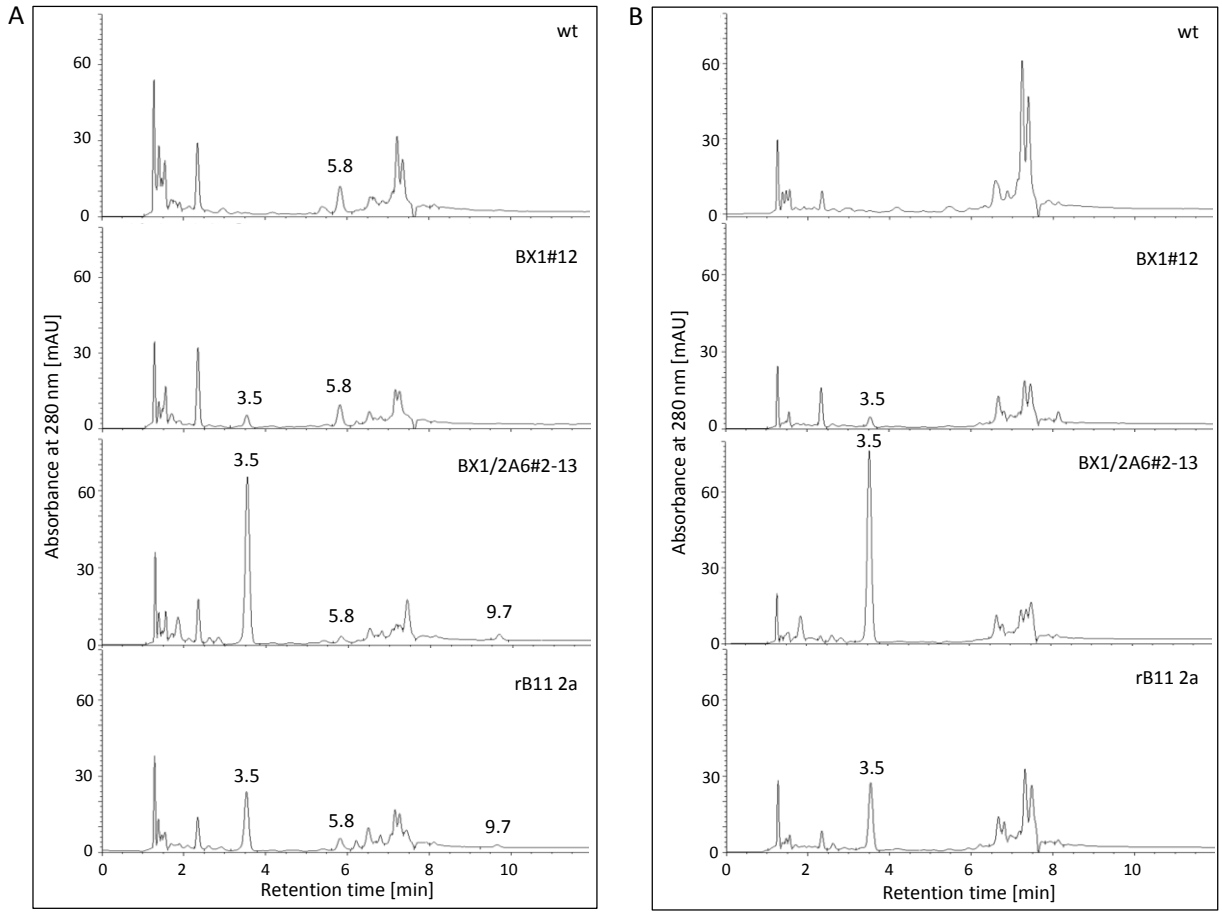

Figure 3.21: Biosynthesis of 6-chlorotryptophan from supplemented 6-chloroindole in *N. tabacum* wt, BX1#12, BX1/2A6#2-13 and rB11 plants. A: HPLC analyses revealed the formation of 6-chlorotryptophan (R_t 5.8 min) in all plant samples after infiltration of 6-chloroindole. However, biosynthesis of chloroindican by 2A6mut in rB11 plants was not confirmed. Moreover, 6-chloroindole (R_t 9.7 min) was detected in marginal amounts in BX1/2A6 and rB11 plant extracts. indican: R_t 3.5 min. **B:** Formation of chloroindole or chlorotryptophan was not observed in negative controls infiltrated with the H_2O:MeOH solution.

To determine R_t values of chloroindican derivatives synthesized by 2A6mut and simplify their identification in HPLC analyses, leaves of rB11 2a plants were infiltrated with chloroindoles. Even though 2A6mut was proven to convert halogenated indoles [74, 75], corresponding indican derivatives were not detectable in this approach. Instead, infiltration of 6-chloroindole resulted in the biosynthesis of 6-chlorotryptophan, whereas 7-chloroindole was not converted to 7-chlorotryptophan. These findings indicate that 6-chloroindole is converted by endogenous TS, whereas 7-chloroindole was not processed by this enzyme. In this regard, Goss *et al.* previously demonstrated the inhibition of bacterial TS by 7-chloroindole in cell lysates [163]. Moreover, they could show the conversion of 5- and 6-chloroindole to the chlorotryptophan equivalents [163]. These findings correlate with the biosynthesis of putative 5- and 6-chlorotryptophan by the plant-derived TS as exemplified in *N. tabacum*. The conversion of particular halogenated indole derivatives by TS from tobacco was never described before and suggests that catalytic properties of the enzyme are similar to bacterial TS. Given that chloroindole is converted by the endogenous tobacco TS, this enzyme limits the availability of substrate of 2A6mut. Consequently, concentration of synthesized chloroindican molecules might be masked by other tobacco metabolites. Therefore, R_t values of chloroindican should be determined by more sensitive LC-MS. Moreover, validation of m/z values could give final proof of chlorotryptophan biosynthesis from chloroindole by the endogenous tryptophan synthase.

3.2.6.3. Increase of enzyme activity by nicotine depletion through grafting

Tobacco is one of the main model plants for plant biotechnology due to a relative simple and highly efficient transformation and regeneration process [164]. However, the nicotine content within tobacco leaves could affect the activity of heterologous enzymes. This refers especially to 2A6 which was shown to be inhibited by nicotine at K_i values of 4.4 ± 0.6 µM in an *in vitro* assay [73]. This secondary metabolite is synthesized in the roots of the plant and subsequently transported to the leaves *via* the xylem [165, 166]. It is then probably stored in vacuoles and acts as protection against herbivore attacks [167, 168]. The accumulation of nicotine within tobacco leaves can easily be eliminated by rootstock removal and grafting of tobacco scions onto the rootstock of a variety of other *Solanaceae sp.* (see 2.4.5). This principle was applied for the elimination of nicotine from transgenic tobacco plants expressing the indoxyl biosynthetic pathway genes as well as the halogenase gene *rebH Y455W*. In this regard, scions of transgenic *N. tabacum* lines BX1#12, BX1/2A6#2-13, rG9 6a, rB11 2a and rB13 2a were grafted onto *Solanum lycopersicum* var. TA234 tomato rootstocks (fig. 3.22). After four weeks metabolites were

extracted from leaves tissue. Additionally, leaf extracts from transgenic plants of these lines as well as *N. tabacum* wt plants were used as controls in LC-MS analyses (see 2.5.5.6). Due to the grafting process, developmental stages of tobacco scions and control plants were not comparable. Therefore, samples were prepared from plants of similar sizes but different ages.

Figure 3.22: Tobacco scion grafted onto tomato rootstock for removal of nicotine in tobacco leaves.

LC-MS analyses of extracts from tobacco plant scions revealed a significant nicotine reduction up to complete elimination from leaf tissue. The alkaloid eluted after 1.5 min with an m/z value of 163 in positive ionization mode in control samples (fig. 3.23). However, chloroindole and chloroindican derivatives were not observed in leaf extracts of transgenic tobacco scions.

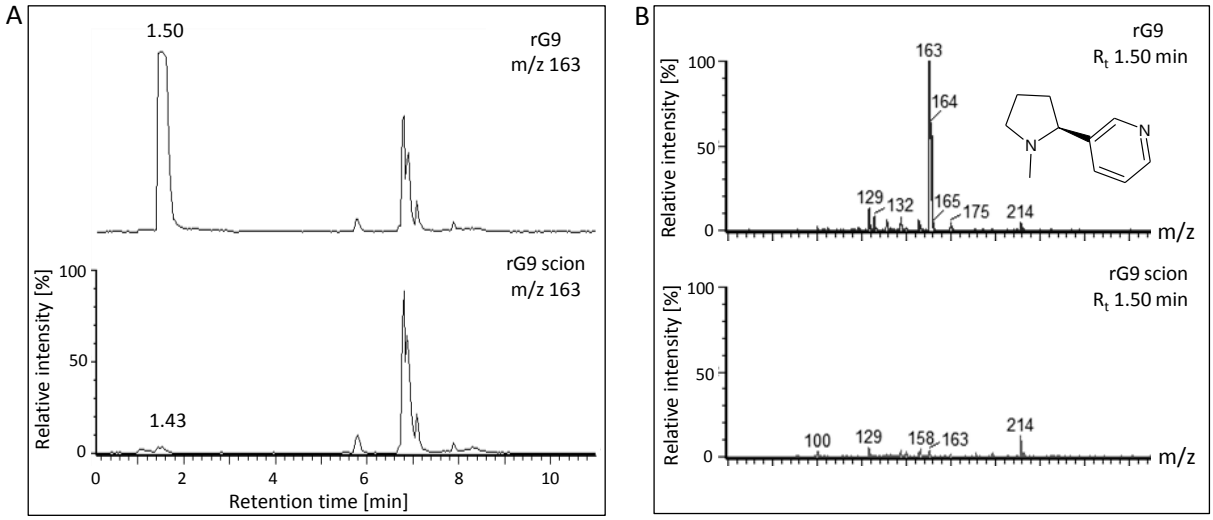

Figure 3.23: Removal of nicotine by grafting of transgenic *N. tabacum* plants exemplified for line rG9 6a. A: LC-MS analyses could prove almost complete elimination of nicotine in leaves of tobacco scions. **B:** Mass spectra of rG9 scion revealed the successful depletion of nicotine compare to control rG9 plants.

These results demonstrate the efficient elimination of nicotine from transgenic *N. tabacum* plants by grafting. Since nicotine has a proven inhibitory effect on P450 2A6 [73], catalytic activity of this cytochrome, as well as other susceptible heterologous enzymes should be increased in tobacco scions. However, elimination of nicotine had no beneficial effect on the biosynthesis of halogenated indole and indican derivatives suggesting that 2A6mut activity is not the limiting factor. A low conversion rate of tryptamine by RebH Y455W determined in previous studies [28] and within the scope of the presented study (see 3.1.3), could account for the inefficient halogenation of indole and indoxyl. Given that RebH wt was shown to sufficiently chlorinate indole at C3, it can be assumed that the halogenase mutant catalyzes the same reaction. However, synthesized chloroindole derivatives were not detected in LC-MS analyses. Besides an inefficient catalytic activity of RebH Y455W, diffusion of volatile chloroindole molecules could be another reason for their absence in leaf extracts. Accordingly, indole was usually not detected in transgenic *N. tabacum* BX1#12 plants. Only the glycosylated indoxyl molecule, stored in the vacuole as indican, indicated the biosynthesis of indole within the plants. Given that 2-chloroindole is not processed by 2A6mut or endogenous enzymes, diffusion of this volatile molecule might complicate its detection in leaf tissue. On the contrary, if 2-chloroindole is further processed by endogenous tobacco enzymes, the resulting product is hardly detectable in LC-MS analyses if the kind of modification and consequently the molecular mass is unknown. Moreover, low concentrated products might by hidden by the metabolite matrix, which complicates any detection by UV-VIS measurements in HPLC analyses.

In summary, molecular engineering of the indoxyl biosynthetic pathway by introduction of the 7-halogenase, RebH Y455W and the flavin-reductase, RebF in transgenic *N. tabacum* lines did not result in the biosynthesis of halogenated indole or indican derivatives. Moreover, nicotine elimination in leaves through grafting of transgenic tobacco scions was not sufficient to increase the biosynthesis of the desired metabolites. Therefore, an alternative biosynthetic pathway was designed to efficiently synthesize these halogenated molecules *in planta*.

3.2.7. Improvement of chloroindole biosynthesis through application of tryptophanase TnaA from *E. coli*

Biosynthesis of indole by the transgenic *N. tabacum* plants used in course of this present study is catalyzed by indole synthase, BX1 from maize that converts indole-3-glycerolphosphate to indole in chloroplasts. However, indole biosynthesis can be accomplished using another natural metabolite as a

substrate. A huge diversity of bacterial species synthesize indole through hydrolysis of tryptophan [169]. Indole functions as a signaling molecule, regulating many physiological functions, such as formation of biofilms, motility or antibiotic resistance [170-172]. In *E. coli*, this metabolite is synthesized by tryptophanase TnaA, which hydrolyses tryptophan to generate indole, pyruvate and ammonia [169]. Therefore, it is not only an alternative enzyme for indole biosynthesis, but also a promising candidate for biosynthesis of halogenated indole derivatives from chlorinated and brominated tryptophan. In addition, the proposed biosynthetic track helps circumvent halogen incorporation at position C3 by virtue of RebH wt, which would inhibit indole conversion by 2A6mut. Therefore, after tryptophan halogenation by Stth or RebH wt and subsequent hydrolysis by TnaA, chloroindole should be available for oxidation by 2A6mut, resulting in the formation of chloroindoxyl. This unstable molecule could then be glycosylated by endogenous glycosyltransferases, as shown for indican biosynthesis (fig. 3.24). TnaA requires pyridoxyl 5-phosphate, synthesized in plants by the pyridoxine synthase 1 and PDX2 [173], as a cofactor [169]. Moreover, conversion of halogenated tryptophan derivatives by TnaA was previously described by Lee and Phillips [174], which makes this enzyme an even more promising candidate for the postulated biosynthetic approach. Till now, production of TnaA in a plant-based system has not been described. Therefore, enzymatic activity of TnaA was determined by transient transformation of *N. benthamiana* leaves to evaluate its potential for plant synthetic biology applications. Characterization of TnaA as well as analysis of the modified indoxyl biosynthetic pathway in *N. benthamiana* were performed by Bastian Wagner within the scope of a bachelor thesis research project [156].

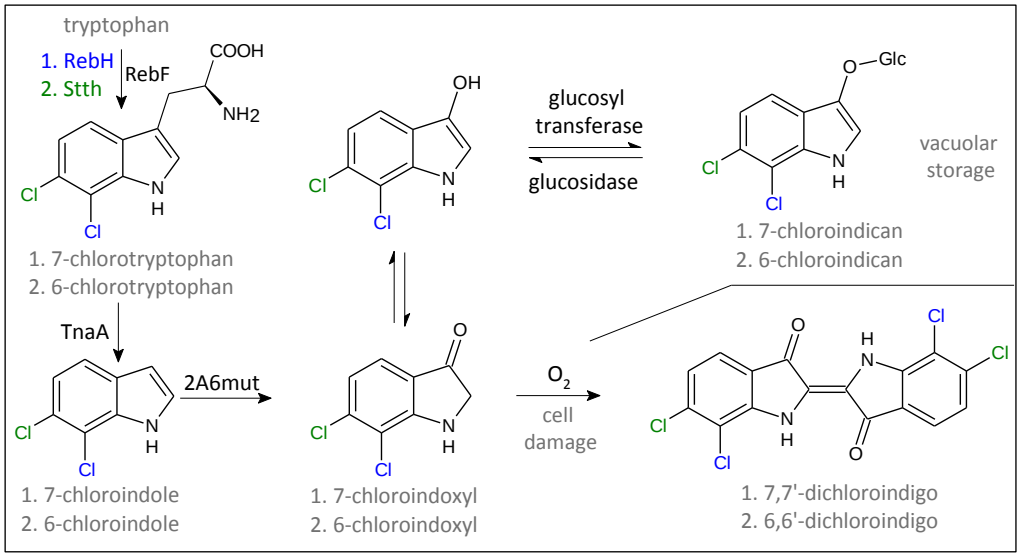

Figure 3.24: Proposed biosynthesis of chloroindican through optimization of the indoxyl biosynthetic pathway. Following halogenation of tryptophan by RebH wt or Stth, chlorotryptophan is hydrolyzed by the tryptophanase, TnaA. Thus synthesized chloroindole is subsequently oxidized by the 2A6 double mutant. After glycosylation of chloroindoxyl by endogenous glycosyltransferases, halogenated indican derivatives accumulate in the vacuole. In case of cell disruption, chloroindican is deglycosylated and indoxyl derivatives dimerize in the presence of molecular oxygen to form dichloroindigo.

To analyze the biosynthesis of indole by TnaA, *N. benthamiana* plants were transiently transformed with different TnaA constructs resulting in localization of the enzyme in the cytosol (fig. 3.14, Tn1) or chloroplasts (fig. 3.14, Tn2). Additionally, cytosolic TnaA and ER-anchored 2A6mut (fig. 3.14, Tx1) as well as chloroplast localization of both enzymes (fig. 3.14, Tx2) was validated in terms of synthesized indican levels. Transformation of pDGBα2 plasmid served as a negative control and all infiltrations were performed in three biological replicates. Finally, metabolites were extracted from leaf tissue and analyzed by HPLC (see, 2.5.4.1, 2.5.5.2). In this regard, biosynthesis of indole (R_t 8.0 min) was observed in samples with cytosol- or chloroplast-localized TnaA (fig. 3.25, A). Moreover, co-localization with 2A6mut resulted in the accumulation of indican (R_t 3.5 min) in high levels, as compared to previous transient approaches using BX1 and 2A6mut (see 3.2.3). Remarkably, biosynthesis of indican by chloroplast-localized 2A6mut was comparable to that of the ER anchored P450, suggesting that electrons required for substrate oxidation were supplied by proteins of the photosystem I (PSI). As discussed previously, PSI was shown to transfer electrons over ferredoxin to recombinant cytochrome P450s, leading to the biosynthesis of pathway products in a light-dependent manner [160]. Accordingly, efficient electron transfer from PSI to 2A6mut might contribute to its high enzymatic activity in chloroplasts. To prove this

hypothesis, light-dependent biosynthesis of indican can be determined through variation of illumination periods of transiently transformed tobacco. In contrast to 2A6mut anchored within the ER membrane, indican levels should correlate with the duration of light exposure applied to plants expressing chloroplast-localized *2A6mut*.

Subsequently, biosynthesis of chloroindican derivatives through hydrolysis of 6- or 7-chlorotryptophan by TnaA and subsequent oxidation by 2A6mut was investigated. Therefore, *N. benthamiana* leaves were co-transformed with different *tnaA-2A6mut* constructs as well as *stth* and *rebH wt*. First of all, the effect on chloroindican biosynthesis after translocation of TnaA, 2A6mut, RebF and either RebH wt (fig. 3.14, Tx2 + E9) or Stth (fig. 3.14, Tx2 + H9) into chloroplasts was analyzed by co-transformation. Additionally, TnaA, 2A6mut, RebF were co-localized with either RebH wt (fig. 3.14, Tx1 + D13) or Stth (fig. 3.14, Tx1 + F13) in the cytosol. Finally, cellular arrangement was modified to optimize the metabolic flux and reduce the accumulation of byproducts, such as indole and indican. Therefore, RebH wt or Stth were transported into chloroplasts to increase the biosynthesis of chlorotryptophan (fig. 3.1, E9, H9). Moreover, TnaA was localized in the cytosol, whereas 2A6mut was anchored within the ER membrane (fig. 3.14, Tx1). Due to this enzyme separation in different cell compartments, halogenases and TnaA did not compete for the same substrate.

Compared to TnaA and 2A6mut control samples, biosynthesis of four additional metabolites correlated with the activity of RebH wt (fig. 3.25). The first molecule exhibited an R_t value of 5.2 min characteristic of 7-chlorotryptophan (tab. 3.3). The second metabolite, accumulating in high concentrations, eluted after 7.0 min and absorbed UV light at 224 nm and 287 nm. The third eluted after 9.3 min and revealed UV maxima of 218 nm and 271 nm, identical to those of the 7-chloroindole standard metabolite (tab. 3.3). Finally, the fourth compound synthesized by rebH wt exhibited an R_t value of 10.2 min and UV maxima at 224 nm and 285 nm. Accordingly, Stth activity correlated with the formation of three additional metabolites (fig. 3.25). First, 6-chlorotryptophan, eluting after 5.6 min, followed by a second molecule exhibiting an R_t value of 7.0 min and UV maxima at 230 nm and 285 nm. The concentration of this compound was low as compared to the aforementioned products resulting from RebH wt infiltration experiments. The last Stth product eluted after 9.7 min and revealed UV maxima at 222 nm and 278 nm, characteristic of 6-chloroindole (tab. 3.3). Moreover, the highest levels of the putative 6-chloroindican eluting after 7.0 min were obtained when Stth was localized in chloroplasts, whereas TnaA and 2A6mut

were co-localized in the cytosol. A reason for this finding could be the previously observed conversion of 6-chloroindole to the corresponding tryptophan derivative by tryptophan synthase (TS), (see 3.2.6.2). Thus, biosynthesis of 6-chloroindole by TnaA in the cytosol presumably prevents its subsequent conversion by TS. In contrast, 7-chloroindole was not converted by this enzyme and, therefore, 2A6mut substrate levels were not reduced by TS. In summary, the proposed biosynthetic pathway including TnaA, 2A6mut and RebH wt is more flexible in terms of cellular localization and leads to higher product yields compared to the equivalent pathway with incorporated Stth.

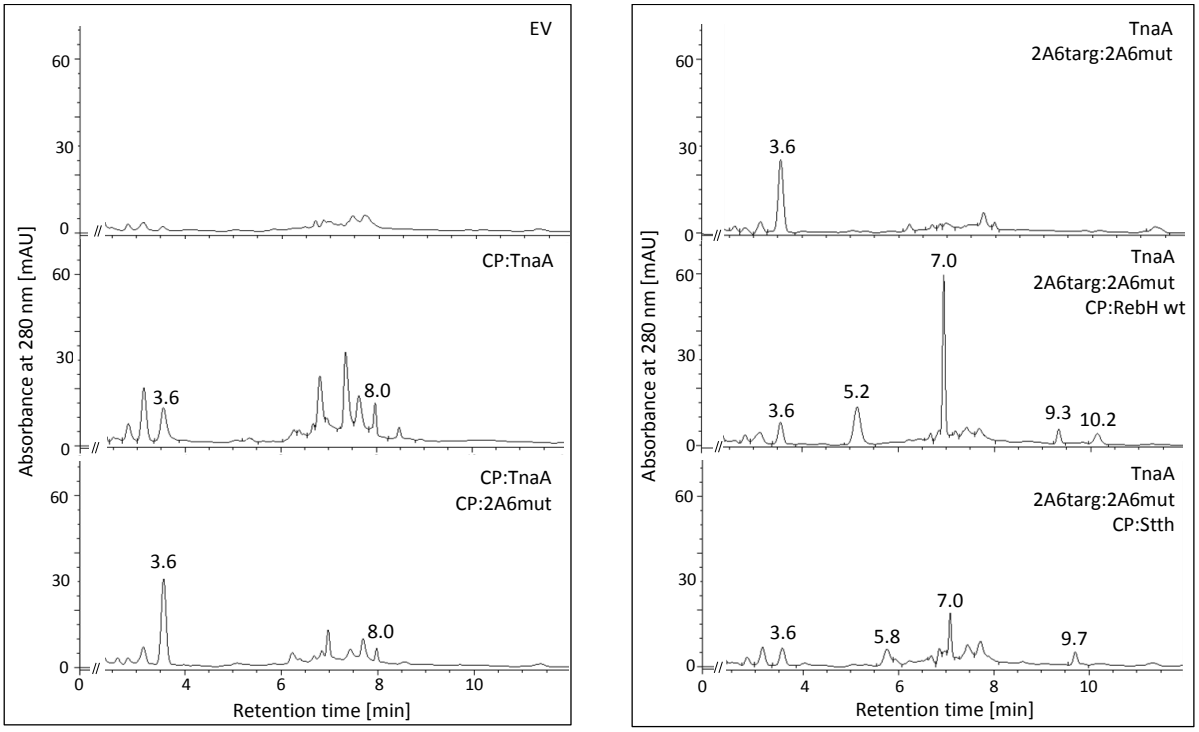

Figure 3.25: Biosynthesis of halogenated indole and indican derivatives in transiently transformed *N. benthamiana* leaves. Biosynthesis of indole (R_t 8.0 min) and indican (R_t 3.6 min) by TnaA and 2A6mut in chloroplasts was verified in HPLC analysis. Introduction of tryptophan halogenases, RebH wt and Stth, resulted in formation of 7-chloroindole (R_t 9.3 min), 6-chloroindole (R_t 9.7 min) and putative chloroindican (R_t 7.0 min), as exemplified for co-localization of TnaA and 2A6mut in the cytosol and translocation of halogenases to the chloroplasts. An additional metabolite (R_t 10.2 min) was detected in RebH wt samples. EV = empty vector, CP = chloroplast targeting, 2A6targ = anchoring of 2A6mut within the ER membrane.

Additionally, leaf extracts were analyzed by LC-MS to determine m/z values of the putative chloroindican derivatives (see 2.5.5.5). In accord with the initial results, co-localization of TnaA and 2A6mut in chloroplasts correlated with the formation of indican, that eluted after 1.69 min and exhibited an m/z value of 294 in negative ionization mode (fig. 3.26, A). As discussed above, biosynthesis of indican was also observed in absence of 2A6mut, indicating that, to a certain degree, indole was converted by an endogenous monooxygenase. Moreover, co-localization of RebH wt, TnaA and 2A6mut in chloroplasts resulted in the formation of a new metabolite eluting after 2.42 min (fig. 3.26, A). The mass spectrum of this compound revealed two m/z values of 328 and 330 in negative ionization mode with a signal intensity ratio of 3:1, consistent with a mono-chlorinated indican molecule (fig. 3.26, B). Accordingly, biosynthesis of 6-chloroindican was observed in samples with chloroplast-localized Stth, cytosolic TnaA and ER-anchored 2A6mut. The metabolite exhibited an R_t value of 2.39 min and two m/z values of 328 and 330 with a signal intensity ratio of 3:1 (fig. 3.26). Similar to initial HPLC analyses, R_t values of chlorinated indican derivatives differed only slightly in terms of C6 and C7 halogenation. One of the reasons for the highly similar retention times of both metabolites could be insufficient separation on the 50 mm UPLC column. However, the two molecules were not separated in the HPLC analyses, performed with a 100 mm column, either. Therefore, it is more likely that stereospecific halogenation of the relatively hydrophilic indican does not significantly affect the retention time parameters of its substituted derivatives. Consequently, the characteristic shift in terms of C6- and C7-halogenated metabolites, observed for chlorinated tryptophan, indole and tryptamine was not detected (tab. 3.2, 3.3). Surprisingly, small amounts of chloroindican were also detected in the sample with no recombinant halogenase activity. Accordingly, HPLC analyses of a substantial set of samples revealed the accumulation of chloroindole in relatively low levels within control leaf extracts. Due to the lipophilic nature of indole, it easily diffuses through cell membranes [158]. Therefore, it can be assumed that halogenated indole was transported within the plant and accumulated in leaves that served as controls. When 2A6mut was present in those leaves, chloroindican was synthesized. Consequently, in future experiments, infiltrated constructs need to be carefully divided between disparate plants in terms of indole and chloroindole biosynthesis. Moreover, an additional metabolite, exhibiting an m/z value of 328, eluted after 3.4 min. According to the mass spectrum, this molecule was not halogenated and might have been a side-product synthesized by TnaA or 2A6mut, or an endogenous metabolite.

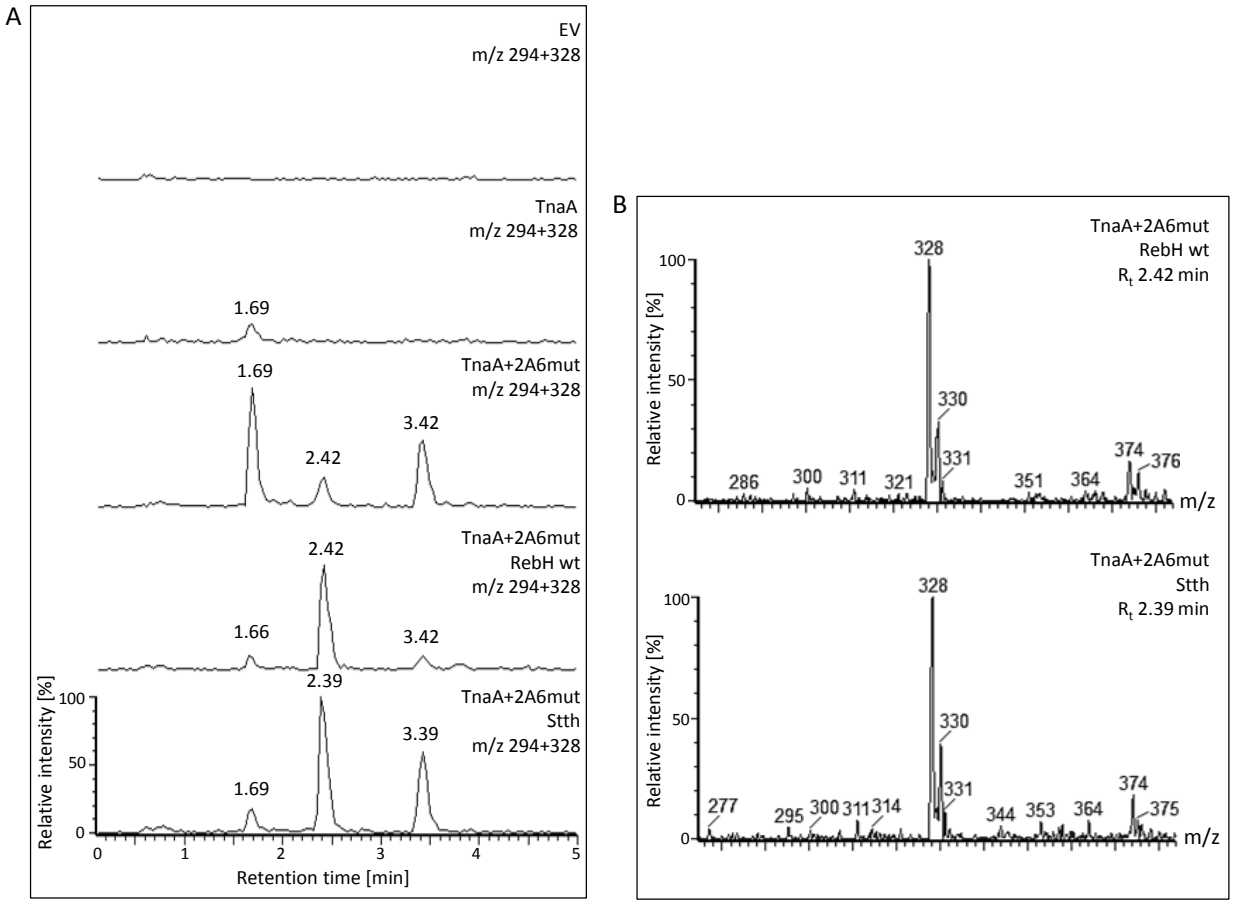

Figure 3.26: Biosynthesis of 6- and 7-chloroindican in transiently transformed *N. benthamiana*. A: Co-localization of TnaA, 2A6mut and RebH wt in chloroplasts correlated with accumulation of indican (R_t: 1.66 min, m/z: 294) and 7-chloroindican (R_t: 2.42, m/z: 328). Accordingly, co-localization of TnaA and 2A6mut in the cytosol and translocation of Stth to chloroplasts resulted in biosynthesis of indican and 6-chloroindican (R_t: 2.39 min, m/z: 328). An additional metabolite, exhibiting an m/z value of 328, eluted after 3.39 min, which was, according to its mass spectrum, not halogenated. **B:** Mass spectra of 7-chloroindican (RebH wt) and 6-chloroindican (Stth) exhibited two m/z values of 328 and 330 with a signal intensity ratio of 3:1 consistent with a mono-chlorinated indican.

All described metabolites were detected independently of the cellular arrangement of the involved enzymes, suggesting that this biosynthetic pathway is robust and highly efficient, even though halogenases and TnaA compete for the same substrate. However, comparing RebH wt and Stth constructs and, therefore, regiospecific halogenation at C6 and C7, chloroindican biosynthesis efficiency differed significantly. In this regard, more 7-chloroindican than the 6-chlorinated derivative accumulated in tobacco leaves. This might be the result of 6-chloroindole reuse by tryptophan synthase leading to the formation of 6-chlorotryptophan [163]. This assumption is supported by the fact that infiltration of

6-chloroindole into the leaves of wild type and transgenic *N. tabacum* plants resulted in the biosynthesis of chlorotryptophan. In contrast, 7-chloroindole was not converted to the corresponding tryptophan derivative (see 3.2.6.2). Consequently, 6-chlorotryptophan and 6-chloroindole might be retained in a circular pathway, limiting 2A6mut substrate availability.

Interestingly, accumulation of indican was observed when all pathway enzymes were localized in chloroplasts. It was anticipated that the low or absent glycosyltransferase activity in chloroplasts would result in biosynthesis and storage of indoxyl and chloroindoxyl in plastids [159, 175]. However, it needs to be assumed that the instable indoxyl is glycosylated in an efficient manner by chloroplast localized enzymes. Li *et al*. investigated glycosyltransferases encoded in the genome of in *A. thaliana* in phylogenetic analyses [159]. In this regard, none of the analyzed enzymes comprised an N-terminal signal peptide suggesting that these enzymes are exclusively located in the cytosol [159]. In chloroplasts, transfer of an activated glycosyl moiety of ADP-glucose is catalyzed by starch synthases contributing to biosynthesis of amylopectin and amylose [176]. However, processing of small molecules was never described for this class of enzymes and, therefore, glucosylation of indoxyl might be catalyzed by other enzymes which have not yet been characterized. To get further insights into the subcellular localization and storage of indican, metabolites can be extracted from isolated chloroplasts. If indican is detected in within the organelles, the processing of indoxyl is most likely catalyzed by a chloroplast-localized enzyme. Even though indican was highly concentrated in samples of chloroplast localized 2A6mut, a small percentage indoxyl molecules could dimerize to form indigo within the organelles. However, purification of indigo from leaf tissue and further separation of this pigment from chloroplast chlorophylls might be challenging due to the highly lipophilic nature of these pigments [177]. Moreover, marginal amounts of the blue indigo might be masked by co-migrating chlorophylls in TLC analyses.

In summary, the exchange of the indole synthase, BX1 by the bacterial tryptophanase, TnaA resulted in significant increase in indole biosynthesis, as compared to *N. benthamiana* transiently transformed with *bx1* (see 3.2.3) or transgenic *N. tabacum* BX1#12 lines (see 3.2.6.1). These findings demonstrate the significant potential of TnaA for application in plant synthetic biology. Moreover, TnaA sufficiently hydrolyzed 6- and 7-chlorotryptophan to synthesize the corresponding chloroindole molecules. Therefore, regiospecific halogenation of tryptophan by Stth or RebH wt and subsequent hydrolysis by TnaA allows the biosynthesis of specifically halogenated indole derivatives. In contrast, direct

halogenation of indole by RebH wt results in chlorine substitution at C3, which inhibits oxidation by 2A6mut.

For further improvement of chloroindole and chloroindoxyl biosynthesis, the most efficient cellular arrangement of all pathway enzymes should be validated in terms of metabolic flux optimization. Due to a high variation of metabolite concentrations within biological replicates, a reliable comparison of biosynthetic tracks, resulting in the highest chloroindican levels, could not be determined in the aforementioned study. In general, protein biosynthesis is more efficient in young leaves leading to high yields of synthesized metabolites [178]. Moreover, infiltration efficiency varies in each experiment and depends on the age of the plant and the leaves. In this regard, comparing metabolite biosynthesis in disparate plants and leaves is challenging. These difficulties can be circumvented by vacuum infiltration of tobacco, which allows efficient and consistent infiltration of leaves. Thus, variation of infiltration efficiency, typically occurring in case of *Agrobacterium* application with syringes, is eliminated. Moreover, extraction of compounds from leaves at all developmental stages of one plant is a simple method to equalize metabolite concentrations. Moreover, biosynthesis of the valuable chloroindican derivatives by recombinant enzymes can be increased by elongation of the incubation period from four up to seven or ten days. However, a critical aspect of such an extension is the continuous reduction of soluble tryptophan levels within the cells. Therefore, the optimal incubation time for metabolite biosynthesis needs to be determined in terms of plant vitality. Besides modification of the localization of enzymes as well as the incubation time, reduction of toxic intermediate content can help increase pathway productivity. This refers especially to halogenated indole derivatives, which were proven to have toxic effects on bacteria at a concentration of 1 mM [75]. By linking TnaA and 2A6mut to a protein scaffold, halogenated indole derivatives would be (presumably) immediately oxidized to the corresponding indoxyl molecules. Through reduced accumulation of toxic intermediates within the cells, plant vitality might be positively affected over several days. However, enzyme activity needs to be maintained after anchoring within the scaffold, which could not be demonstrated for 2A6mut (see 3.2.5). To increase the catalytic activity of this enzyme, an NADPH-cytochrome P450 oxidoreductase could be co-located on the protein scaffold to directly transfer required electrons to the cytochrome.

Furthermore, chloroindican concentrations can be increased through enhancement of precursor biosynthesis. In this regard, a potential target is anthranilate synthase (AnS) which catalyzes the first step

of the tryptophan biosynthetic pathway to form anthranilate from chorismate [161]. This enzyme is a heterotetramer consisting of two α- and β-subunits and is feedback regulated by binding of tryptophan to the α-subunit [161]. Moreover, Salcher *et al*. demonstrated that 7-chlorotryptophan inhibited AnS activity in *Pseudomonas aureofaciens* [179]. Given that AnS in plants is similarly regulated, accumulation of 7-chlorotryptophan could inhibit the biosynthesis of the parent amino acid and thereby limit the concentration of the final product. Especially if chlorotryptophan is not further processed and accumulates in high amounts within the plants, feedback regulation of AnS might limit tryptophan availability (see 3.1.2.1). In this regard, Li and Last described an AnS α-subunit mutant (Trp5) in *A. thaliana*, which was not feedback regulated and led to a 3-fold increase in tryptophan concentrations *in planta* [180]. Moreover, co-localization of the AnS α-subunit mutant, Trp5 and the wt β-subunit, Asb1 resulted in an enhanced tryptophan accumulation in *C. roseus* hairy root cultures compared to a Trp5 control [181]. Consequently, co-expression of *trp5, asb1, 2A6mut* and halogenase genes might enhance chloroindican biosynthesis and improve the profitability of the postulated plant-based production system.

Moreover, indican biosynthesis could be improved by introduction of a heterologous glycosyltransferase (GT) in order to facilitate indoxyl glycosylation by a recombinant enzyme rather than an endogenous one. One promising candidate is arbutin synthase (AS), first described by Hefner *et al.,* which catalyzes glycosylation of a broad range of aromatic compounds [182]. Additionally, establishing transgenic cell cultures could enable constitutive production of halogenated indican derivatives *in planta*. Furthermore, integration of the 5-halogenase, PyrH would probably result in accumulation of 5-chloroindican and thereby widen the product spectrum. Biosynthesis of brominated indican derivatives represents another possibility to enlarge the pathway variety. Given that 2A6mut is capable of bromoindole oxidation, the substrate scope of TnaA needs to be determined. In case of an inefficient conversion of bromotryptophan by TnaA, enzyme mutants generated by amino acid exchange within the active site can be tested [175, 183].

Even though indican biosynthesis by BX1 and 2A6mut was highly efficient in transgenic *N. tabacum* plants [14], the outcome of transient reconstruction of this biosynthetic pathway was rather poor (see 3.2.6.1). Initially, it was assumed that 2A6mut might be rate-limiting, probably due to nicotine inhibition. The supposition was supported by similar indican contents in plants transformed with *bx1* and

2A6mut and those expressing only *bx1*. This finding indicated that indican biosynthesis was catalyzed by endogenous enzymes rather than 2A6mut. However, the P450 together with TnaA synthesized sufficient amounts of indican in transient approaches, indicating that BX1, and not 2A6mut, might be the limiting factor. Initial characterization of the indole synthase showed that BX1 converts indole-3-glycerol phosphate (indole-3-GP) to indole in a highly efficient manner [68]. Determination of catalytic constants of this reaction revealed a K_m value of 0.027 mM and a k_{cat} value of 2.8 s^{-1} [68]. However, in a plant-based production system, BX1 and tryptophan synthase (TS) compete for the same substrate. Consequently, indole-3-GP might not accumulate within the cells, thereby limiting the substrate availability of BX1. Given that TS probably synthesizes 6-chlorotryptophan from 6-chloroindole, it can be assumed that the enzyme also converts indole and, thus, decreases 2A6mut substrate. Therefore, indoxyl biosynthesis might be decisively influenced by TS, that not only reduces indole3-GP levels but also converts indole synthesized by BX1 (fig. 3.27, A). In contrast, TnaA does not compete with TS and, therefore, has higher substrate levels available as compared to BX1. Still, TS might reduce product yields through indole conversion while, at the same time, synthesizing fresh substrate supplies for TnaA, which might lead to cellular equilibrium of indole and tryptophan levels. Moreover, through TnaA-catalyzed conversion, amounts of cellular tryptophan are continuously reduced. Therefore, the probability of AnS inhibition by this molecule is decreased, leading to enhancement indole-3-GP biosynthesis (fig. 3.27, B). The restriction of BX1 activity by low substrate levels could be circumvented through increased precursor levels. Introduction of a tryptophan-insensitive AnS would probably lead to accumulation of indole-3-GP and an enhanced biosynthesis of indoxyl by BX1 and 2A6mut in transient approaches [161]. The limitation of indole-3-GP in transgenic tobacco might have no negative effect on final product yields, as indican accumulates slowly, but continuously, over days and weeks. To facilitate a more flexible and efficient application of BX1 in transient transformation approaches, co-localization of tryptophan-insensitive AnS, Trp5 might prove beneficial.

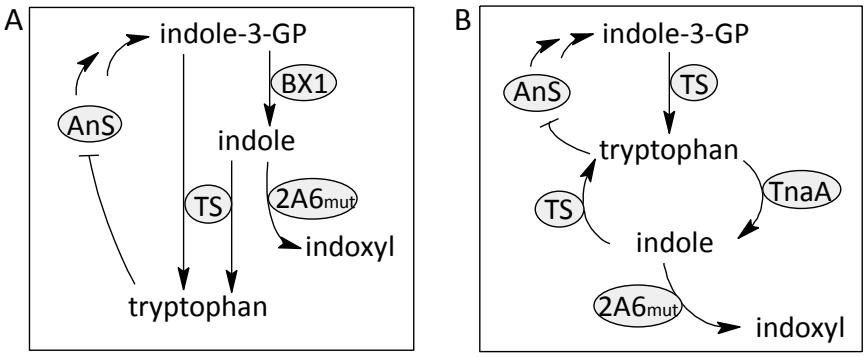

Figure 3.27: Proposed metabolic flux of the initially proposed [14] (A) and modified (B) indoxyl biosynthetic pathway. A: Tryptophan synthase (TS) and BX1 compete for the same substrate, indole-3-glycerol phosphate (indole-3-GP) [69]. Moreover, TS presumably converts indole to tryptophan and thereby reduces 2A6mut substrate levels. **B:** The continuous conversion of tryptophan and indole by TnaA and TS, might lead to equilibrium of cellular content of both metabolites. Further, tryptophan levels are reduced by TnaA, which decreases the probability of anthranilate synthase (AnS) inhibition and enables enhanced biosynthesis of precursor molecules, leading to increased indole-3-GP levels [161].

3.3. Molecular engineering of the strictosidine biosynthetic pathway

3.3.1. Molecular cloning of DNA constructs for reconstitution of the strictosidine biosynthetic pathway

For co-expression of all strictosidine biosynthetic pathway genes in an alternative host, DNA sequences were introduced into the GoldenBraid (GB) cloning system. Gene sequences of the geraniol synthase (*ges*), geraniol 8-oxidase (*g8o*), 8-hydroxygeraniol oxidoreductase (*8-hgo*), iridoid synthase (*is*), iridoid oxidase (*io*), 7-deoxyloganetic acid glucosyltransferase (*7-dlgt*), 7-deoxyloganic acid hydroxylase (*7-dlh*), loganic acid O-methyltransferase (*lamt*), secologanin synthase (*sls*), tryptophan decarboxylase (*tdc*) and strictosidine synthase (*str*) were used as templates for the introduction of all relevant GB fusion sites as well as *BsmBI* recognition sites by PCR using primers listed in tab. 2.4 (see 2.2.1). Additionally, all internal *Bsa*I and *BsmB*I recognition sites were removed by introduction of silent mutations. The endogenous chloroplast targeting sequence of *ges* as well as the vacuolar targeting sequence of *str* were excluded from the amplification reaction. According to previous studies, the strictosidine synthase mutant STRV214M catalyzes the conversion halogenated tryptamine in contrast to the wild type enzyme [117]. Therefore, *str* gene sequence was modified by PCR using primers listed in tab 2.4 to introduce the desired amino acid exchange. Following amplification and purification, DNA fragments were ligated into the universal domesticator *via* GB reaction (see 2.2.3, 2.2.6). During assembling of transcriptional units in α level plasmids all CDSs were fused to a CaMV 35S promoter (P35S) and a nopaline synthase terminator comprising an 8x His-tag (his-TNos). Additionally, the 5'-end of *ges* was ligated to the chloroplast targeting sequence, *cp.* Moreover, the 5'-end of the *strV214M* CDS was either fused to the vacuolar targeting sequence, *erV* or to the P35S_ATG promoter leading to localization of enzymes in the cytosol. For verification of successful biosynthesis of recombinant enzymes by fluorescence microscopy, additional transcriptional units were assembled. In this regard, a *bfp* CDS was ligated to the 3'-end of *strV214M* CDS which was either fused to a vacuolar or none targeting sequence.

For the biosynthesis of secologanin by transient transformation, TUs of the first pathway genes, -*ges*, *g8o*, *8-hgo* and *is*- were assembled with *yfp* TU in an Ω level plasmid (fig. 3.28, Sec7). The second part of the pathway included the *io*, *7-dlgt*, *7-dlh*, *lamt* and *sls* TUs (fig. 3.28, Sec8). Additionally, assembly of the entire pathway was achieved by fusion of the Sec7 and Sec8 TUs into the pHUGE plasmid (fig. 3.28, Sec10). To facilitate strictosidine biosynthesis from secologanin and tryptamine, the *tdc* TU was either

fused to the *erV:strV214M:bfp* (fig. 3.28, T13) construct or the *strV214M:bfp* (fig. 3.28, T12)construct. For subsequent cloning, the same multigene constructs were assembled using both *strV214M* TUs but without *bfp* fusion. For the modification of the strictosidine biosynthetic pathway through introduction of the tryptophan 7-halogenase RebH Y455W, two additional constructs were generated. Therefore, the *rebY455W-rebF:bfp* multigene construct (fig. 3.1, B13) was fused to the *erV:strV214M-tdc* multigene construct (fig. 3.28, ST2) to analyze biosynthesis of halogenated strictosidine in the vacuole. Additionally, the same halogenase/reductase construct was fused to a *strV214M-tdc* construct (fig. 3.28, ST3) to investigate the formation of chlorostrictosidine in the cytosol.

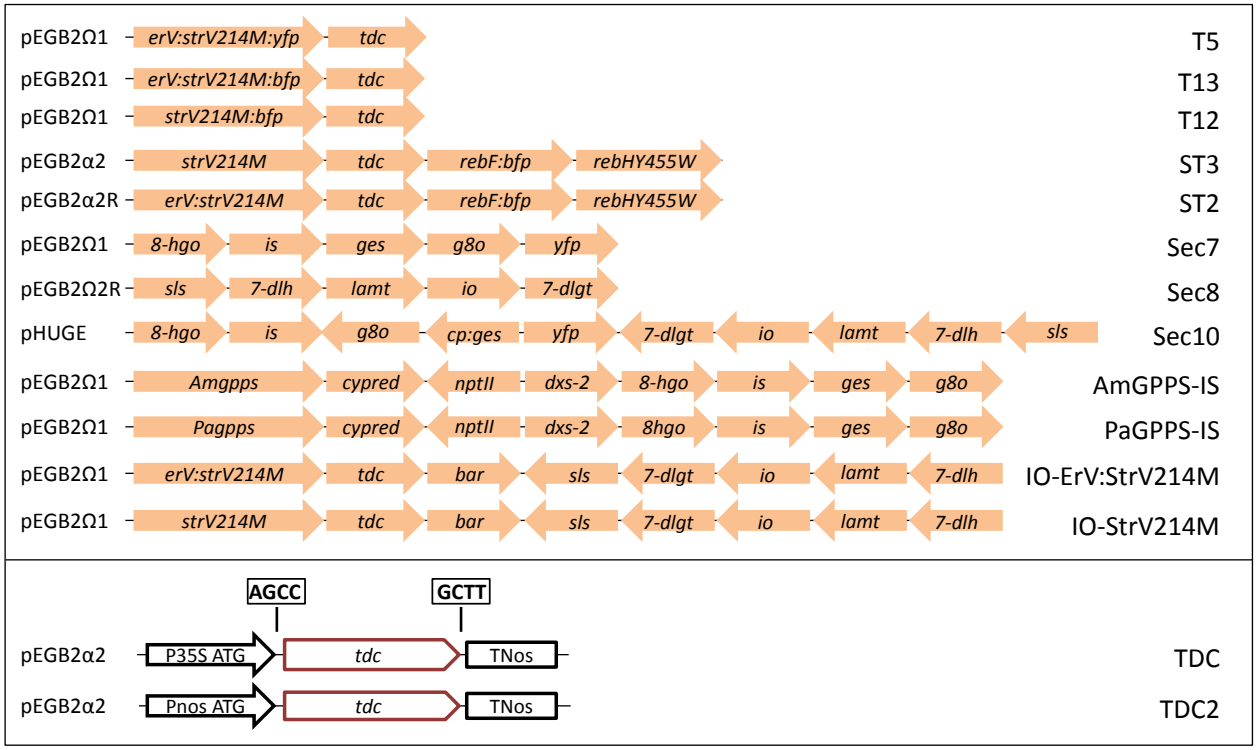

Figure 3.28: GoldenBraid 2.0 DNA constructs used for molecular engineering of the strictosidine biosynthetic pathway in *Nicotiana sp*. Illustrated transcriptional units (arrows) include a CaMV 35S promoter and Nos terminator. Targeting sequences and DNA parts of fusion proteins are separated from CDSs by double dots. Resistance genes of each GB backbone can be found in chapter 2.2.6 and abbreviations are specified on the right.

To increase the biosynthesis rate of final products, the strictosidine biosynthetic pathway was extended by three additional synthetic genes. In this regard, the gene encoding for the 1-deoxy-D-xylulose synthase from *Solanum lycopersicum* (*dxs-2*) was ligated into the universal domesticator *via* GB reaction. Moreover, geranyl diphosphate synthases (GPPS) from *Picea abies* (*Pagpps*) or the GPPS small subunit of

Antirrhinum majus (*Amgpps*) were both synthesized without the endogenous chloroplast targeting sequence and introduced to the GB cloning system by ligation into the universal domesticator. Finally, the electron transport to recombinant cytochrome P450s anchored within the ER membrane was enhanced by introduction of a NADPH-cytochrome P450 oxidoreductase from *Rauvolfia serpentina*. For this purpose, cDNA encoding the reductase was integrated into the GB cloning system by Kai Göpfer within the scope of a bachelor thesis research project [130]. For assembly of transcriptional units, these four genes were fused to a CaMV 35S promoter and a Nos terminator. Additionally, 5'-ends of *gpps* genes were fused to a chloroplast targeting sequence. For the selection process of transgenic *N. tabacum* plants, the resistance marker genes *nptII* encoding for a neomycin phosphotransferase II and *bar* encoding for a phosphinothricin acetyl transferase were selected. The first part of the optimized pathway was assembled by fusion of *ges*, *g8o*, *8-hgo* and *is* TUs to *dxs2*, *cypred*, *nptII* and either *Amgpps* (AmGPPS-IS) or *Pagpps* (PaGPPS-IS). The second part of the pathway was reconstructed by assembly of the *io*, *7-dlgt*, *7-dlh*, *lamt*, *sls*, *bar*, *tdc* and either *erV:strV214M* (IO-ErV:STRV214M) or *strV214M* (IO-STRV214M) TUs (fig. 3.28).

Moreover, to decrease tryptophan decarboxylase levels within the plant cells, *tdc* expression was reduced by fusion to the nopaline synthase promoter and terminator (fig. 3.28, TDC2).

Finally, for verification of vacuolar localized STRV214M, the 3'-end of erV:*strV214M* was fused to *yfp* CDS. Furthermore, this TU was assembled with the *tdc* TU (fig. 3.28, T5).

A detailed description of all chronological steps performed during assembly of GB constructs and transformation of *A. tumefaciens* cells used for transient transformation of *N. benthamiana* is given in chapter 3.1.1

3.3.2. Reconstitution of the strictosidine biosynthetic pathway in *Nicotiana benthamiana* by transient transformation

For the biosynthesis of metabolites of the strictosidine biosynthetic pathway, *N. benthamiana* leaves were co-transformed with all pathway genes through *Agrobacterium*-mediated gene transfer in transient approaches (see 2.4.3). These transformation approaches as well as all following analyses of metabolite composition were performed during an international exchange in the group of Dr. Heiko Rischer at the

Technical Research Center of Finland VTT. To facilitate the biosynthesis of secologanin, leaves were transformed with all pathway genes starting from *ges* up to *sls* by either co-infiltration (fig. 3.28, Sec7 + Sec8) or transformation with the pHUGE construct (fig. 3.28, Sec10). Furthermore, strictosidine accumulation was investigated in terms of localization of STRV214M in the vacuole (fig. 3.28, Sec7 + Sec8 + T13) or the cytosol (fig. 3.28, Sec7 + Sec8 + T12). Finally, the biosynthetic pathway was modified by introduction of the tryptophan 7-halogenase, RebH Y455W and the flavin reductase, RebF, both localized in the cytosol. As described above, biosynthesis of chlorostrictosidine was analyzed with regards to localization of STRV214M in the vacuole (fig. 3.28, Sec7 + Sec8 + ST2) or cytosol (fig. 3.28, Sec7 + Sec8 + ST3). Infiltration of leaves with ICON buffer or transformation with the pDGB2α2 plasmid, ST2 or ST3 served as negative controls. Moreover, to increase metabolite yields, leaves were supplemented with 400 µM geraniol, whereas infiltration of a H_2O:MeOH solution served as a control (see 2.4.6). To analyze subsequent processing of secologanin or strictosidine by endogenous enzymes, leaves were separately infiltrated with 400 µM of each metabolite (see 2.4.6). Newly synthesized molecules were analyzed in extracts of leaf tissue by UPLC-MS (see 2.5.4.2, 2.5.5.7). Iridodial glycoside, iridotrial glycoside, 11-OH-iridodial glycoside as well as loganic acid, loganin, secologanin and strictosidine standard metabolites were additionally analyzed to determine specific R_t and m/z values (tab 3.5). Due to the fact that many secondary metabolites as well as xenobiotics are glycosylated by endogenous glycosyltransferases to increase water solubility and stability or for detoxification [184, 185], glycosylated iridoids were used as standard metabolites for the analyses. Moreover, glycosylation of iridoids simplifies LC-MS measurements, whereas corresponding aglycons are relatively volatile and lipophilic and therefore need to be analyzed by gas chromatography (GC).

Tab. 3.5: Determined R_t and m/z values of reference metabolites in LC-MS analyses in negative ionization mode.

standard metabolite	R_t value	m/z value	adduct
iridodial glycoside	7.98	375	formic acid
11-OH-iridodial glycoside	7.31	391	formic acid
iridotrial glycoside	7.35	389	formic acid
7-deoxyloganic acid	7.27	359	---
loganin	6.89	435	formic acid
secologanin	7.24	433	formic acid
strictosidine	7.60	575	formic acid

Moreover, leaf extracts were screened for biosynthesis of new metabolites associated with transgene expression using the TransOmics™ informatics software (waters). In this regard, samples were subdivided in three groups (tab. 3.4). The first treatment group composed of leaves transformed with transgenes that were additionally supplemented with geraniol to boost secologanin biosynthesis. The second treatment group included leaves that were transformed with pathway genes but infiltrated with a control solution. Metabolites of both treatment groups were compared to one control group, which composed of control infiltrations supplemented with geraniol. Consequently, all metabolites synthesized through conversion of geraniol by endogenous enzymes were automatically excluded from the set of newly synthesized metabolites.

Tab. 3.4: Overview of designated treatment and one control groups for categorization of newly synthesized metabolites evaluated by TransOmics™ informatics software analyses.

treatment group geraniol supplementation		**treatment group H_2O:MeOH infiltrated**		**control group geraniol supplementation**	
GB construct	**transformed genes**	**GB construct**	**transformed genes**	**GB construct**	**transformed genes**
Sec7+Sec8	*ges-sls*	Sec7+Sec8	*ges-sls*	buffer	---
Sec7+Sec8 +T12	*ges-sls* *strV214M:bfp* *tdc*	Sec7+Sec8 +T12	*ges-sls* *strV214M:bfp* *tdc*	EV	---
Sec7+Sec8 +T13	*ges-sls* *erV:strV214M:bfp* *tdc*	Sec7+Sec8 +T13	*ges-sls* *erV:strV214M:bfp* *tdc*	ST2	*erV:strV214M* *tdc* *rebH Y455W* *rebF:bfp*
Sec7+Sec8 +ST2	*ges-sls* *erV:strV214M:bfp* *tdc* *rebH Y455W* *rebF:bfp*	Sec7+Sec8 +ST2	*ges-sls* *erV:strV214M:bfp* *tdc* *rebH Y455W* *rebF:bfp*	ST3	*strV214M* *tdc* *rebH Y455W* *rebF:bfp*
Sec7+Sec8 +ST3	*ges-sls* *strV214M:bfp* *tdc* *rebH Y455W* *rebF:bfp*	Sec7+Sec8 +ST3	*ges-sls* *strV214M:bfp* *tdc* *rebH Y455W* *rebF:bfp*		
Sec10	*ges-sls*	Sec10	*ges-sls*		

The software-based analyses of the LC-MS dataset revealed 14 new metabolites that biosynthesis correlated with transgene expression (tab 3.5). These metabolites were not present in negative controls infiltrated with geraniol and therefore do not result from geraniol conversion by endogenous enzymes. Surprisingly, none of the 14 synthesized molecules was present in extracts of leaves transformed with the pHUGE construct Sec10. Therefore, plasmid preparations of the Sec7, Sec8 and pHUGE constructs from GV3101 cells used for this transient approach were analyzed in terms of integration of pathway genes *via* PCR using gene specific primers (tab. 2.4). Thereby, all anticipated genes were verified in the Sec7 and Sec8 constructs. However, only the *yfp* CDS was detected in the pHUGE plasmid Sec10, indicating loss of all other TUs putatively by homologous recombination in GV3101 cells due to repetitive promoter and terminator sequences. DH5α and TOP10 cells are deficient for recombination events due to a single point mutation of the *recA* gene encoding for the recombinase A [186]. In contrast, GV3101 cells are not *recA*- and therefore recombination of large recombinant plasmids with repeating sequence stretches might be a likely event. Consequently, for transformation of large gene constructs in *A. tumefaciens*, usage of a *recA* deficient strain, such as AGL1 should be taken into account [187]. Another negative aspect of gene constructs of this enormous size is the significantly decreased transformation efficiency into *Agrobacterium*, which was observed using plasmids that size exceeded 20,000 bp. Therefore, co-transformation of tobacco using plasmids with up to eight TUs seems to be a much more efficient and timesaving method. Due to the deletion of TUs in the pHUGE plasmids in GV3101 cells, *N. benthamiana* leaves were most likely transformed with the *yfp* TU alone. Therefore, transformation of the Sec10 constructs served as an additional negative control.

Four out of 14 newly synthesized metabolites were associated with geraniol supplementation and were absent or relatively low concentrated in samples infiltrated with the control solution (tab. 3.6). Moreover, geraniol supplementation resulted in increased amounts of five metabolites, indicating that these molecules were synthesized by recombinant enzymes and originated from geraniol (tab. 3.6). Accordingly, concentrations of five metabolites were relatively stable upon geraniol treatment, indicating that these molecules represent side products of endogenous molecules converted by recombinant enzymes. Only one of the newly synthesized molecules displayed a close similarity to one of the standard metabolites exhibiting an R_t value of 7.34 min and m/z value of 391 in negative ionization mode, which is similar to the determined values of 11-OH-iridodial glycoside (tab 3.5, fig. 3.29). However, R_t values of both molecules slightly differed from each other about 0.03 s. Moreover,

11-OH-iridodial was not described as an intermediate after expression of the first pathway genes in *N. benthamiana* [12].

Tab. 3.6: R_t values and m/z values of metabolites associated with reconstitution of the strictosidine biosynthetic pathway evaluated by LC-MS in negative ionization mode.

metabolites synthesized independently of geraniol supplementation		metabolites synthesized upon geraniol supplementation only	
R_t [min]	m/z	R_t [min]	m/z
7.38	717	7.34	391
7.39 (+)	155	7.45	347
7.39 (+)	199	7.60	185
7.42	185	7.65	389
7.43	393		
7.51	347		
7.51 (+)	363		
7.60	187		
7.63 (+)	349		
7.78 (+)	391		

(+) biosynthesis was significantly increased upon geraniol supplementation

Furthermore, biosynthesis of a second metabolite correlated with transgene expression. This molecule eluted after 7.78 min, exhibiting an m/z value of 391 (fig. 3.29). Moreover, biosynthesis of this metabolite was increased upon geraniol supplementation. These findings indicate the specific conversion of geraniol or another metabolite originating from geraniol by recombinant enzymes to an unknown product. Additionally, biosynthesis of a third metabolite exhibiting the same mass as 11-OH-iridodial glycoside correlated with geraniol supplementation. This molecule revealed an R_t value of 7.28 min and was found in control and treatment groups indicating additional conversion by endogenous tobacco enzymes. Moreover, concentration of this compound was significantly increased in samples synthesizing recombinant enzymes upon geraniol treatment (fig. 3.29). These findings suggest the specific conversion of geraniol by recombinant and endogenous enzymes, which presumably catalyze the same reaction. This molecule was not detected by the software-based analyses, because it was present in negative controls. However, due to increased biosynthesis by recombinant enzymes upon geraniol treatment it

could represent an intermediate originating from geraniol. Given that R_t values of these two molecules differed significantly from 11-OH-iridodial glycoside, it is possible that they represents glycosylated derivatives of the strictosidine pathway that might be further processed by endogenous enzymes. For the unequivocal identification of the new compounds, synthesis of additional standard metabolites would be beneficial to evaluate retention times and mass spectra. Since concentrations of molecules synthesized by recombinant enzymes are rather low, their preparation in mg scale for NMR analyses is tedious and might not be suitable for this approach.

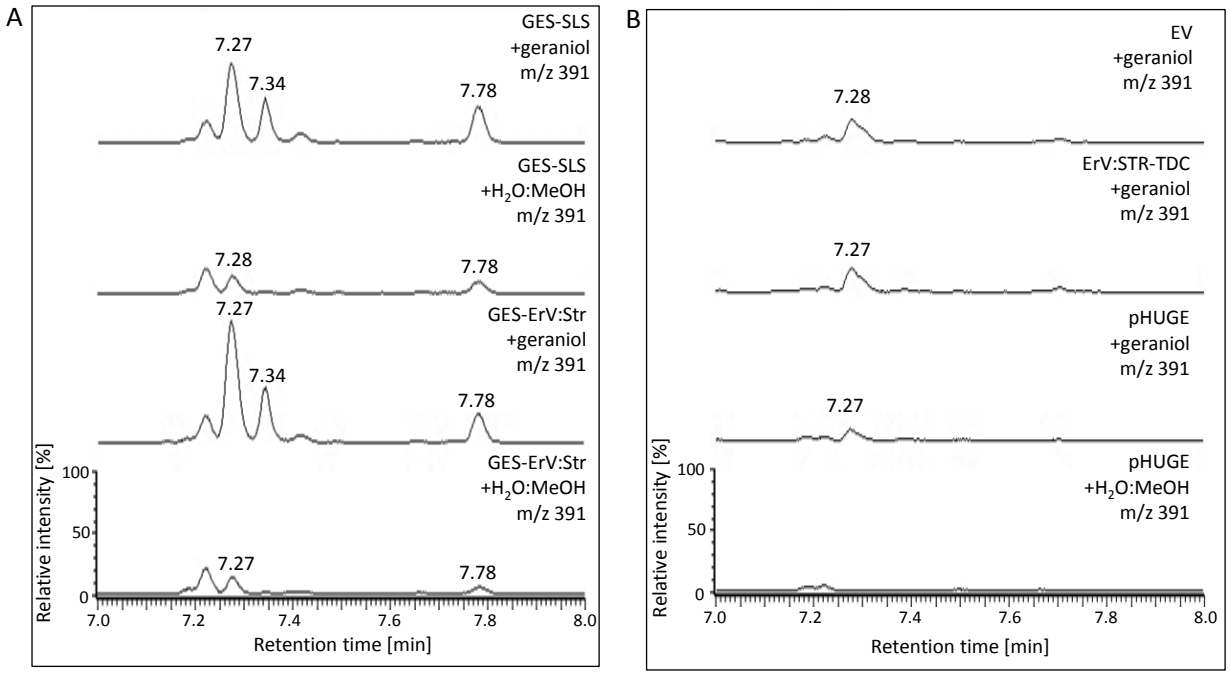

Figure 3.29: Biosynthesis of the putative 11-OH-iridodial glycoside and two additional metabolites in transiently transformed *N. benthamiana*. Newly synthesized metabolites of both treatment groups (A) were compared to those present in the control group (B). Transgene expression and geraniol supplementation correlated with the biosynthesis of putative 11-OH-iridodial glycoside (R_t 7.34 min) as exemplified for two different transformations. Biosynthesis of a second putative pathway intermediate exhibiting an m/z value of 391 and R_t value of 7.78 min correlated with transgene expression and was increased upon geraniol supplementation. Moreover, biosynthesis of a third metabolite exhibiting an R_t value of 7.27 min and m/z value of 391 was associated with geraniol infiltration and was found in treatment and control groups.

Since secologanin and strictosidine could be further metabolized by endogenous tobacco enzymes, structure and molecular mass of these derivatives are hardly predictable and might not be identified within the group of molecules specified by Transomics software analyses. Therefore, 400 µM secologanin

and strictosidine were infiltrated separately into *N. benthamiana* leaves and analyzed by LC-MS (see 2.4.6, 2.5.5.7). The resulting LC-MS data were screened for the occurrence of all metabolites that correlated with transgene expression (tab. 3.6). However, none of the 14 detected molecules were present in these positive controls suggesting that the newly synthesized metabolites do not represents metabolized secologanin or strictosidine derivatives.

Given the fact that R_t and m/z values of merely one metabolite detected in leaf extracts correlated with one of the standard metabolites, further conversion of pathway intermediates by endogenous tobacco enzymes might contribute to the biosynthesis of the remaining 13 molecules. However, evaluation of possible metabolic flux within the host plant is a challenging task. Moreover, the detected metabolites might represent products of the heterologous enzymes converting metabolites originating from the host. This refers especially to 8-HGO and IO, which were shown to convert a variety of substrates like the primary alcohols *trans*-2-hexenol, farnesol and nerol (8-HGO) or lavandulol, nerol and citronellol (IO) [12]. Moreover, side-reactions and conversion of pathway intermediates by both enzymes represents another possible reason for the biosynthesis of side products. In this regard, 8-HGO was shown to convert also geraniol, and IO additionally catalyzes the conversion of 8-oxogeraniol [12]. To gain further insights into pathway efficiency and potential bottlenecks, first, the accumulating intermediates need to be identified. Especially G8O, which revealed a low conversion rate of geraniol to 8-hydroxygeraniol in yeast, might catalyze a rate-limiting step in the plant-based expression system [115]. Moreover, overexpression of *dxs* and *g10h* in *C. roseus* hairy roots resulted in a significantly increased indole alkaloid accumulation, indicating that both G10H and DXS are potential bottlenecks [188]. In this regard, volatile and highly lipophilic metabolites synthesized during the first steps of the pathway, could be analyzed by GC-MS to unveil potential bottlenecks. Subsequently, gene expression can be enhanced by introduction of several copies of one gene or by fusion to a strong promoter-terminator combination like the CaMV 35S promoter and terminator, which was shown to increase gene expression almost 3 fold compared to the combination used in this approach [21]. Additionally, introduction of alternative pathway enzymes should be taken into account. Krithika *et al.* recently characterized a 10-hydroxy-geraniol dehydrogenase (10HGO) from *C. roseus*, which catalyzes the biosynthesis of 10-oxogeranial from 10-hydroxygeraniol [95]. This reaction was reversible by forming the intermediates 10-oxogeraniol or 10-hydroxygeranial. In contrast to the 8-hydroxygeraniol oxidoreductase (8-HGO) used in course of the undertaken studies, Cr10HGO revealed low enzymatic activity on geraniol, nerol and farnesol. Due to

this small substrate spectrum, formation of side products could be reduced for the benefit of final product yields. Given that Cr10HGO together with GPPS, GES and IS forms *cis*-*trans*-nepetalactol from isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), this enzymes could be advantageous for production of iridoids in tobacco [95].

Since transient expression of all strictosidine biosynthetic pathway genes had not resulted in strictosidine formation, biosynthesis of chlorostrictosidine from chlorotryptamine was analyzed in a second approach. Therefore, *N. benthamiana* leaves were infiltrated with 400 µM secologanin, whereas infiltration of a H_2O:MeOH solution served as negative control (see 2.4.6). Subsequently, leaves were additionally transformed with *tdc*, *rebH Y455W*, *rebF* and either *erV:strV214M* for vacuolar or *strV214M* for a cytosolic STRV214M localization (fig. 3.28, ST2 , ST3). Biosynthesis of new metabolites was analyzed in leaf extracts by LC-MS (see 2.5.4.2, 2.5.5.7). Moreover, chromatograms were compared to leaves previously infiltrated with 400 µM secologanin or strictosidine to analyze further conversion of these molecules by endogenous tobacco enzymes. In this regard, expression of *strV214M* and *tdc* correlated with the biosynthesis of marginal amounts of potential strictosidine, which eluted in line with the infiltrated standard metabolite exhibiting an R_t value of 7.6 min and m/z value of 575. However, putative strictosidine was only observed when STRV214M was localized in the vacuole, indicating that this enzyme is not functional in the cytosol. Moreover, conversion of tryptamine and secologanin, which are processed by STRV214M in the vacuole, implies the transport of both metabolites from the cytosol across the tonoplast. These findings suggest that the biosynthesis of strictosidine in *Nicotiana sp.* is theoretically possible but needs to be optimized to increase final product yields. Secologanin was relatively high concentrated in transiently transformed leaf tissue compared to a wt leaf infiltrated with this metabolite (fig. 3.30). This observation highlighted the low catalytic activity of STRV214M and correlates with determined kinetic constants of this variant. In this regard the mutant enzyme revealed a 30 fold increased K_m value for tryptamine and 200 times decreased V_{max}/K_m value compared to the wild type enzyme [117]. Besides strictosidine biosynthesis, extracts were screened for halogenated tryptamine and strictosidine. However, compared to metabolite composition in wt leaves infiltrated with strictosidine and secologanin no additional compounds were detected which exhibited the anticipated m/z values of halogenated tryptamine of strictosidine. These findings correlated with the initial functional characterization of RebH Y455W, which was shown to possess low catalytic activity (see 3.1.4). Given that, supplemented secologanin is converted to strictosidine in a transient approach,

co-expression of wild type halogenases could result in the biosynthesis of chlorostrictosidine derivatives on the condition that halogenated tryptamine is also transported into the vacuole.

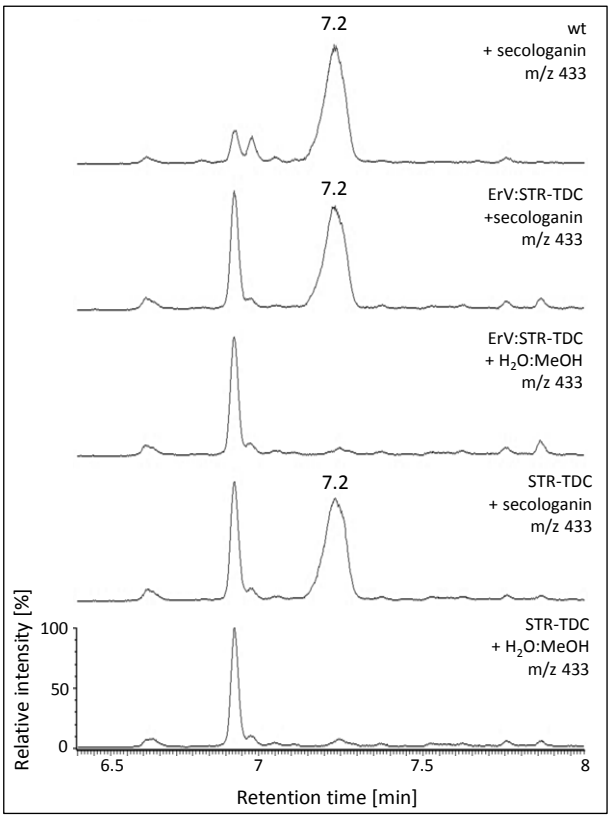

Figure 3.30: Secologanin conversion by STRV214M in transiently transformed *N. benthamiana*. Secologanin concentrations (R_t 7.2 min, m/z 433) in extracts of leaves expressing *strV214M* and *tdc* were comparable to the wt control, indicating a low conversion rate of this metabolite by the strictosidine synthase mutant.

Due to the inefficient biosynthesis of strictosidine and intermediates, accumulation of precursor molecules was enhanced by introduction of two additional enzymes originating from the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway. The MEP pathway is localized in chloroplasts of plant cells and leads to the biosynthesis of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), which are precursors of mono-, di- and tetraterpenes [189]. In this regard, the pathway was extended by introduction of the 1-deoxy-D-xylulose-5-phosphate synthase (DXS), which converts pyruvate and glyceraldehyde 3-phosphate to 1-deoxy-D-xylulose-5-phosphate [190, 191]. This initial biosynthetic step of isoprenoids was shown to be rate-limiting in *A. thaliana* [192]. Moreover, overexpression of *dxs* together with *g10h* resulted in enhanced biosynthesis of several indole alkaloids in *C. roseus* hairy roots, which emphasizes the rate-limiting nature of this enzyme [188]. Two isoforms, DXS1 and DXS2, were identified in many plant species that share about 70% identity of their amino acid

sequences and catalyze the same reaction [193, 194]. To increase biosynthesis of strictosidine and its precursors, the *dxs2* gene from *Solanum lycopersicum* was integrated into the biosynthetic pathway. Additionally, geranyl diphosphate (GPP) biosynthesis was enhances by introduction of geranyl diphosphate synthases (GPPS). These enzymes belong to the prenyltransferases and catalyze the condensation of DMAPP and IPP to form geraniol diphosphate in chloroplasts [189, 195]. In this regard, GPPS from *Picea abies* was incorporated to the biosynthetic pathway, a homodimer, which synthesizes both GPP and geranylgeranyl diphosphate (GGPP) in a ratio of 9:1 [196]. Additionally, GPP biosynthesis was enhanced using the small subunit of the heterotetrameric GPPS from *Antirrhinum maju*s binding to geranylgeranyl diphosphate synthases (GGPPS) to form a functional GPPS [197]. It was demonstrated that this modification of chain length specificity of GGPPS causes decreased levels of di- and tetraterpenes in transgenic tobacco [198]. Consequently, these plants suffered from chlorosis and were shown to be more light-sensitive and growth-restricted due to limited chlorophyll, carotenoids and gibberellins [198]. Therefore, GPPS small subunit from *A. majus* might be beneficial for GPP biosynthesis in transient approaches but inappropriate for establishing stable transgenic tobacco lines. Finally, efficiency of heterologous cytochrome P450s was increased by introduction of a NADPH-cytochrome P450 oxidoreductases from *R. serpentina.* Given that the strictosidine biosynthetic pathway comprises four P450s, electron transfer from endogenous P450 reductases might be insufficient. Therefore, introduction of a heterologous NADPH-cytochrome P450 oxidoreductase potentially increases P450 efficiency and final product yields.

To evaluate the most promising pathway variation, different construct combinations were analyzed in terms of most efficient GPPS and cellular localization of STRV214M. Therefore, first pathway genes, starting from *gpps* up to *is*, were co-expressed in *N. benthamiana* leaves by transformation with the PaGPPS-IS or AmGPPS-IS construct (fig. 3.28). Additionally, leaves were co-transformed with the last pathway genes, from *io* to *strV214M*, to facilitate strictosidine biosynthesis in the vacuole (fig. 3.28, IO-ErV:STRV214M) or the cytosol (fig. 3.28, IO-STRV214M). As a negative control, tobacco cells were infiltrated with cells carrying the unmodified pDGB2α2 plasmid. Each infiltration was performed in three biological replicates and biosynthesis of strictosidine and pathway intermediates was investigated by HPLC and LC-MS (see 2.5.5.1, 2.5.5.6).

Except for tryptamine (R_t 4.1 min), biosynthesis of new metabolites in samples expressing all pathway genes was not observed in HPLC analyses. Therefore, some samples were further analyzed in terms of anticipated m/z values of standard metabolites by LC-MS (tab. 3.5). Even though biosynthesis of precursor molecules was optimized by introduction of two enzymes of the MEP pathway, formation of intermediates or strictosidine was not observed. These findings indicate that a potential bottleneck or imbalances of metabolite flux account for the inefficient biosynthesis of strictosidine rather than the limitation of precursor molecules. In each biosynthesis step, metabolite flux can break down due to competing pathways of the heterologous host [115]. Additionally, identification of accumulating metabolites of the biosynthetic pathway might unveil potential bottlenecks. Therefore, determination of initial pathway metabolites, such as geraniol, 8-oxogeranial or 8-hydroxygeranial by GC-MS could help to identify potential bottlenecks. This refers especially to G8O, which catalyzes most likely one rate-limiting step within the pathway as demonstrated by Brown *et al.* in yeast and Peebles *et al.* in *C. roseus* hairy roots [115, 188]. Given that strictosidine biosynthesis was not observed after construction of an optimized biosynthetic pathway, transgenic tobacco lines expressing initial or all pathway genes were established to facilitate constitutive gene expression and thereby increase product yields.

3.3.3. Reconstitution of the strictosidine biosynthetic pathway by stable transformation of *Nicotiana tabacum*

To increase the yields of strictosidine synthesized by heterologous enzymes, stable transgenic tobacco plants were generated by transformation with corresponding genes. In contrast to transient approaches, persistence of pathway enzymes over the whole growth period probably results in accumulation of the desired metabolites or intermediates. In this regard, *N. tabacum* plants were transformed with the first seven or all fourteen genes of the optimized strictosidine biosynthetic pathway. To increase precursor production, as previously described for transient approaches, DXS-2 and GPPS were incorporated into the pathway. Given that the small subunit of *Antirrhinum majus* GPPS was shown to decrease di- and tetraterpene levels, the GPPS from *Picea abies* was selected to increase GPP concentration in transgenic plants [198]. Due to the enormous size of all 15 TU including one selection marker gene, stable transformation was performed with two plasmids using two different resistance genes, *nptII* and *bar*. In this regard, *N. tabacum* plants were transformed by biolistic bombardment with the PaGPPS-IS plasmid to facilitate iridodial biosynthesis (fig. 3.28, see 2.4.2.1). After regeneration of transgenic tobacco plants and characterization of high-producing lines, a second transformation with the IO-ErV_STRVM plasmid

was expected to lead to a transgenic *N. tabacum* line finally synthesizing strictosidine. However, this two-step transformation is a longsome process, due to time consuming regeneration of transgenic plants. To facilitate construction of the entire pathway in a more time-saving manner, tobacco leaves were co-transformed with the PaGPPS-IS and IO-ErV_STRVM plasmids to integrate all 14 pathway genes into the tobacco genome at once (fig. 3.28, see 2.4.2.1).

Following transformation with the PaGPPS-IS gene construct, 24 transgenic lines were successfully regenerated from calli. To verify integration of the investigated genes, genomic DNA was extracted from leaf tissue and screened for integration of the corresponding DNA sequences by PCR using gene specific primers (tab 2.6). In this regard, one or more transgenes could be detected in 20 transgenic lines. Due to low purity of extracted genomic DNA, polymerase-based amplification can be inhibited leading to a false negative result. Therefore, lines exhibiting genomic integration of one or more transgenes were expected to inherit all pathway genes. For 15 plants the altered phenotype was observed and 13 of those plants were positive for genomic integration of transgenes (fig. 3.31). Leaves of these plants had a coriaceous morphology and were thicker compared to the wild type. This particular phenotype appeared independently of leaf age. Furthermore, vascular tissue was prominent on both sides of leaves compared to wild type plants. The altered phenotype was also associated with a dark green coloration of leaves, indicating increased chlorophyll levels. Transgenic line #31 was completely different from all the other phenotypes. It featured a succulent morphology with thick, coriaceous leaves and a chaotic arrangement of vascular tissue. However, it is uncertain if this atypical phenotype is caused by high transgene expression levels or by recombination of genes into a particular site of the genome, causing the deleterious effects.

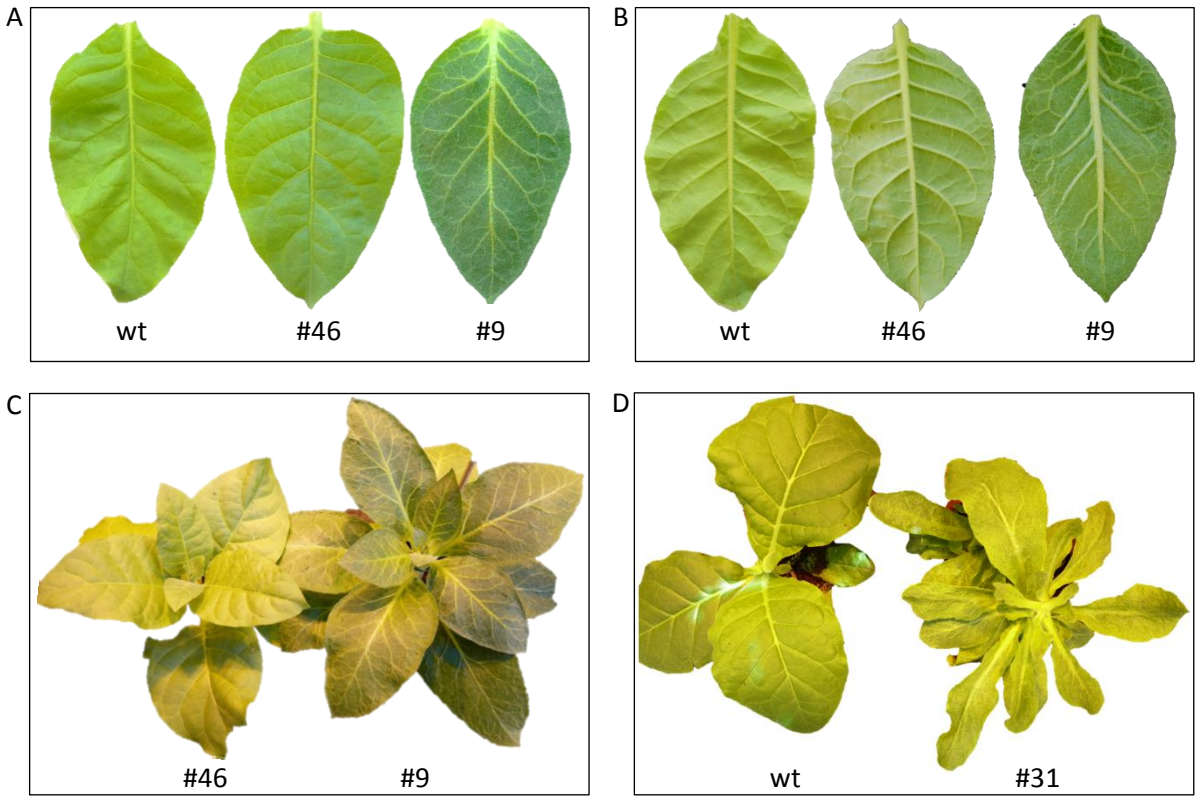

Figure 3.31: Construction of the first part of the optimized strictosidine biosynthetic pathway correlated with an altered phenotype in 15 *N. tabacum* plants. A: Adaxial side of leaves from transgenic plants exhibiting the altered phenotype (#9) had coriaceous and thick leaves with prominent and asymmetric vascular tissue compared to wt or trangenic plants featuring a normal phenotype (#46) **B:** The altert phenotype (#9) correlated with thick and prominent vascular tissue on the abaxial side of leaves. **C:** Comparission of trangenic plants with normal (#46) and the altered (#9) phenotype demonstrated the asymetric organisation of vascular tissue as well as the thick and dark green leaves. **D:** Transgenic line #31 displayed an succulent morphology with narrow leaves and chaotic arrangement of vascular tissue.

To verify the biosynthesis of pathway intermediates, metabolites were extracted from leaf tissue and analyzed by LC-MS (see 2.5.4.1, 2.5.5.6). Chromatograms were compared to *N. tabacum* wt samples and screened for anticipated m/z values of 8-hydroxygeraniol, 8-hydroxygeranial, 8-oxogeraniol, oxogeranial, iridodial and iridodial glycoside in negative and positive ionization mode. Additionally, plant extracts were screened for initially evaluated m/z values of glycosylated iridotrial and 11-OH iridodial (tab 3.5). However, none of these molecules was detected in plant extracts, except for a putative iridodial glycoside, which was present in line #1, 2, 3, 5, 23, 26, 46. This molecule eluted after 7.94 min and exhibited an m/z value of 331 in positive ionization mode. Though, this result was not reproducible,

indicating that the biosynthesis of first pathway intermediates in transgenic tobacco was not successful. Given that the altered phenotype observed in some transgenic plants might result from high transgene expression levels, extracts were again compared to transgenic lines with normal appearance. To exclude any influence of potentially increased water content in altered leaves, analyses were repeated using extracts of freeze-dried leaf tissue (see 2.4.7). Even though it is very likely that the altered phenotype of 15 plants is closely related with accumulation of molecules synthesized by heterologous enzymes, formation of new metabolites was not observed in this approach. A software-based analysis of all detected compounds could give further insights into metabolomic changes in transgenic lines compared to wild type plants. Moreover, increased concentrations of precursor molecules synthesized by DXS-2 and GPPS could also imbalance metabolite composition within the plants. This refers especially to phytohormones that derive from the MEP pathway, such as gibberellins, cytokinins and abscisic acid. Gibberelins are C_{19} or C_{20} tetracyclic diterpenoids that regulate plant growth and development by induction of stem and hypocotyl elongation [199]. Cytokinins, on the other hand, are substituted adenine derivatives, which mainly consists of isoprenoid chains originating from DMAP synthesized by the MEP pathway [200]. They regulate many physiological mechanisms including growth, cell division, vascular development and influence auxin transport and signaling as well as ethylene biosynthesis [201]. Finally the sesquiterpene abscisic acid (ABA) is synthesized from tetraterpenes in chloroplasts. It regulates not only stomata closure during draught stress but also increases leaf size and, together with other phytohormones, controls developmental processes [202]. Consequently, overexpression of *dxs-2* and *gpps* could affect phytohormone composition leading to an altered morphology. However, levels of these phytohormones are not only regulated by their biosynthesis rate but also by endogenous mechanisms resulting in modification of phytohormones to inactive forms, such as 2β-hydroxylation of gibberellins or glycosylation of cytokinins [199, 200]. Since complex regulatory mechanisms control phytohormone activities and altered levels of particular phytohormones have impacts on other phytohormones, a specific regulatory mechanism leading to the altered phenotype of PaGPPS-IS plants is not predictable.

One possible reason for the absence of iridodial from leaf extracts of transgenic plants could be the immediate conversion of this reactive and unstable molecule by endogenous enzymes to prevent any damage through accumulation. Further conversion, glycosylation and vacuolar storage represent possible detoxification processes. Therefore, identification of initial pathway metabolites would help to

identify transgenic IS lines, applicable for further transformation with the final pathway genes. Thereby, storage of unstable molecules like iridodial could be circumvented by direct conversion to 7-deoxy-loganetic acid.

Besides transformation of the first genes, reconstitution of the entire strictosidine biosynthetic pathway was achieved by co-transformation of *N. tabacum* with the PaGPPS-IS and IO-ErV:StrVM constructs (fig. 3.28, see 2.4.2.1). However, only three transgenic tobacco lines were successfully regenerated from emerging calli. Given that both plasmids need to be delivered into one nucleus at the same time, double transformation of tobacco cells might be a rare event. Moreover, two of three calli (#1, #3) displayed a low growth rate, indicating a negative effect of transgene expression on the metabolism. Due to the constitutive expression of the tryptophan decarboxylase gene *tdc*, low tryptophan levels within the cells presumably account for this growth inhibition. Therefore, calli and emerging shoots were partially transferred onto growth media supplemented with 200 µM tryptophan to enhance growth. As expected, transferred calli and plants exhibited a normal growth speed after tryptophan supplementation, indicating that TDC has an immense impact on tryptophan metabolism and protein biosynthesis. Moreover, growth of calli cultivated on standard medium was still inhibited. Screening of genomic DNA from callus tissue by PCR using gene specific primers (tab. 2.6) revealed the integration of genes located on both plasmids in transgenic lines #1 and #3. To investigate the biosynthesis of metabolites from the strictosidine biosynthetic pathway, leaf extracts were analyzed by HPLC and LC-MS (see 2.5.5.1, 2.5.5.6). However, expected metabolites were not observed in leaf extracts, indicating that a rate-limiting step is located upstream of IO and leads to inefficient metabolite flux rather than storage of glycosylated iridoids. Even though integration of transgenes was verified by PCR, transgene expression could be further investigated on a transcriptional level. However, even if all transcripts could be determined, presence of recombinant enzymes cannot be verified by this approach.

In summary, despite initial reconstitution of the strictosidine biosynthetic pathway and further optimization by an increased precursor biosynthesis, strictosidine was not synthesized in transiently transformed *N. benthamiana*. A constitutive gene expression in transgenic *N. tabacum* lines was supposed to increase metabolite biosynthesis and to result in accumulation of strictosidine or other intermediates. However, characterization of these transgenic lines did not reveal any newly synthesized molecules that might originate from the strictosidine biosynthetic pathway. These findings indicate that

the metabolite flux might be retarded by one or more bottlenecks. As discussed above, not only DXS might catalyze a rate-limiting step but also G10H as exemplified by Peebles *et al*. in *C. roseus* hairy roots and Brown *et al*. in yeast [115, 188]. Therefore, low catalytic activity of recombinant G10H potentially leads to geraniol accumulation in transgenic tobacco lines, which could be further investigated by GC-MS. To circumvent this bottleneck, gene expression can be increased either through regulation by strong promoter and terminator or by introduction of multiple copies of this particular gene [21]. Moreover, replacement of G8O by the Cr10HGO recently characterized by Krithika *et al.* could be beneficial for strictosidine biosynthesis [95]. Cr10HGO catalyzes the same reaction as G8O but revealed a smaller substrate scope and therefore, accumulation of side-products could be reduced [12, 95]. Moreover, strictosidine biosynthesis can be increased through replacement of the STRV214M by the wt enzyme. The introduced amino acid exchange in STRV214M is beneficial for chlorotryptamine conversion, a reaction that is not catalyzed by the wt enzyme. At the same time, the mutation leads to a significantly decreased turnover number of tryptamine [117]. Low catalytic activity of STRV214M was also demonstrated previously by transient transformation of *N. benthamiana* (see 3.3.2). Therefore, STR might be beneficial for strictosidine accumulation, if biosynthesis of halogenated derivatives is not intended. Moreover, overexpression of transgenes correlated with an altered phenotype, which presumably results from imbalances of phytohormone composition. However, except for one transgenic line, all plants revealed a vital growth and efficient reproduction. Therefore, constitutive expression of genes of the optimized strictosidine biosynthetic pathway is in principle possible, although further optimization effort is necessary to facilitate strictosidine biosynthesis *in planta*. Moreover, product yields can be up scaled by reconstitution of the entire pathway into chloroplasts. Given that enzymatic activity of recombinant P450s is presumably maintained by electron transfer from PSI, as previously demonstrated for chloroplastic localized 2A6 (see 3.2.7), chloroplast transformation represents a promising tool for secologanin or strictosidine biosynthesis *in planta*.

Finally, TDC was shown to negatively affect vitality and growth of transgenic plants. Therefore, *tdc* expression levels need to be down regulated to facilitate a normal and vital growth after stable transformation of tobacco.

3.3.4. Optimization of *tdc* gene expression

The constitutive gene expression of *tdc* in transgenic *N. tabacum* PaGPPS-ErV:STRV214M plants had resulted in growth inhibition, which was compensated by tryptophan supplementation to culture medium. To reduce the toxic effect of TDC on the tryptophan metabolism, transgene expression needs to be reduced to reach the optimal level for constitutive gene expression in transgenic plants. This can be achieved by variation of the promoter and terminator strengths. In this regard, Sarrion-Perdigones and Vazquez-Vilar *et al.* analyzed the regulatory effects of a huge variety of promoter and terminator combinations [21]. Referring to these results, *tdc* gene expression was adapted using the nopaline synthase promoter and terminator combination, which is supposed to reduce gene expression about 14 fold compared to the standard CaMV 35S promoter and nopaline synthase terminator combination [21]. To analyze the effect on tryptamine biosynthesis and tryptophan availability, *N. benthamiana* leaves were transformed with the GB constructs TDC and TDC2 (fig. 3.28), leading to a moderate or reduced *tdc* gene expression, respectively (see 2.4.3). Transformation with the pDGB2α2 plasmid served as a control (EV). Finally, biosynthesis of tryptamine and tryptophan was analyzed by HPLC (see 2.5.5.1). Compared to control infiltrations, *tdc* expression correlated with biosynthesis of tryptamine, which exhibited an R_t value of 3.9 min and 4.1 min (fig. 3.32). Furthermore, tryptamine concentration was significantly reduced when *tdc* expression was regulated by the weak promoter-terminator combination. However, tryptophan concentration (R_t 3.1 min) was still reduced compared to negative controls suggesting that the reduced *tdc* gene expression might not be sufficient for vital growth of transgenic plants.

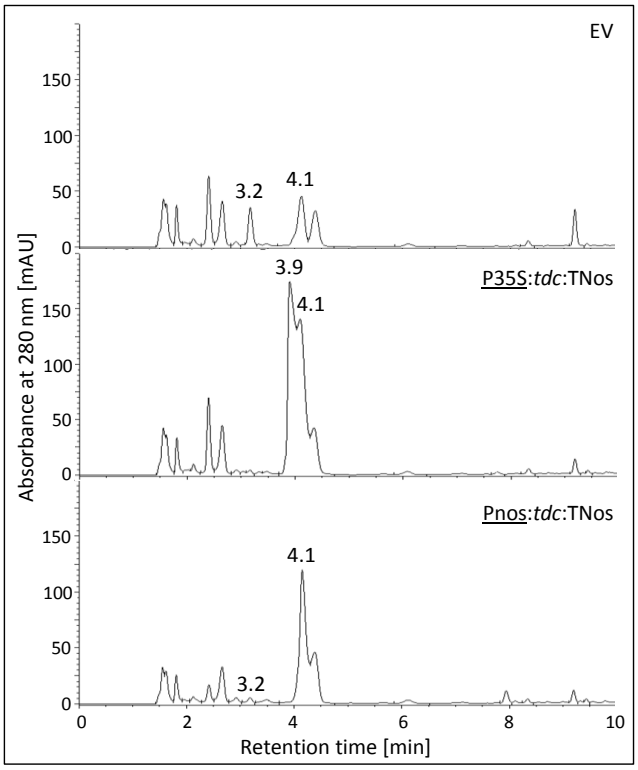

Figure 3.32: Biosynthesis of tryptamine as a function of *tdc* gene expression. Regulation of *tdc* gene expression by a weak nopaline synthase promoter and terminator combination resulted in reduced tryptamine accumulation (R_t 3.9 min, 4.1 min) compared to the standard CaMV 35S promoter and Nos terminator combination. However, tryptophan (R_t 3.2 min) levels were still reduced. EV=empty vector

Although tryptamine biosynthesis rate was reduced, tryptophan levels remained relatively low. To further improve gene expression, coding sequences can be fused to inducible promoter sequences. In general, an efficient inducible expression system must not have pleiotropic effects on the plant and need to induce transgenes expression exclusively. One regulatory system established for application in plant biotechnology is the bacterial tetracycline inducible expression system as exemplified by Gatz *et al.* in transgenic tobacco plants [203]. This system is based on the transgene repression by a Tet repressor, which is constitutively synthesized. In absence of tetracycline this repressor binds to a modified CaMV 35S promoter containing two *tet* operators. Thereby, transgene expression was reduced 80 fold compared to tetracycline induction [203]. However, this reduction was shown only on mRNA level, indicating a leaky repression system. Moreover, induction of gene expression was achieved by tetracycline infiltration into leaves of transgenic plants, which is inappropriate for an industrial production of monoterpene indole alkaloids *in planta*. Another regulatory system successfully established in transgenic tobacco plants and cell culture is the glucocorticoid-induced gene expression [204, 205]. This system is based on a chimeric transcription factor, GVG, which comprises a hormone-binding domain of the vertebrate glucocorticoid receptor (GR), a transactivating domain of the herpes

viral protein VP16 and the DNA-binding domain of the transcription factor GAL4 from yeast. Upon treatment with the artificial glucocorticoid hormone, GAL4 transcription factor acts in *trans* on six copies of the GAL4 upstream activating sequence and thereby induces gene expression. One benefit of this system is the specific gene activation at low inducer concentrations [204, 206]. Moreover, gene expression is regulated in a dose dependent manner, which offers not only the activation of specific genes but also tight regulation of expression levels. Additionally, this system is relatively efficient as demonstrated by expression of a luciferase reporter gene in transgenic tobacco, which was increased up to 100 times of the basal expression level upon glucocorticoid treatment [204]. Gene expression was also induced after spraying of glucocorticoid on leaves of transgenic tobacco plants. However, induction of defense-related genes and growth inhibition observed in transgenic *A. thaliana* lines was supposed to correlate with glucocorticoid treatment [207]. Besides application in whole plants, this inducible system works also in cell culture. In this regard, gene expression in transgenic tobacco cell suspension cultures was induced using relatively low glucocorticoid levels of only 0.1 µM [205]. In summary the GVG expression system represents an extremely flexible and efficient system for induction of transgene expression and might be promising for regulation of *tdc* in transgenic tobacco plants and cell cultures.

Besides the aforementioned regulation of *tdc* gene expression, tryptophan levels can be increased by enhanced precursor biosynthesis through introduction of a recombinant anthranilate synthase (AnS). This enzyme catalyzes the conversion of chorismate to anthranilate and plays a key role in tryptophan biosynthesis [161]. It is feedback regulated by tryptophan, which binds to the α-subunit of the enzyme [161]. This regulation is inhibited in some AnS mutants, such as *trp5* from *A. thaliana* leading to increased tryptophan levels *in planta* [180]. The biosynthesis of this particular AnS mutant in transgenic tobacco potentially restocks the tryptophan pool continuously and thereby antagonizes the toxic effects of TDC.

3.4. Verification of localization of recombinant enzyme by confocal laser scanning microscopy

Molecular engineering of biosynthetic pathways as well as characterization of flavin-dependent halogenases was performed by localization of recombinant enzymes in different cell compartments. In this regard, chloroplast (CP) and apoplast (ER) signal peptides, obtained from other sources, were introduced to the GB system (see 2.1.6). Moreover, endogenous targeting sequences of 2A6 and STR were adapted to the GB cloning technique. For verification of correct localization of enzymes, different GB constructs, each encoding one of the aforementioned signal peptides, were used for transient transformation of *N. benthamiana* leaves in three biological replicates (see 2.4.3). Abaxial sides of leaf sections were analyzed by confocal laser scanning microscopy and representing areas of each sample were documented on photographs (see 2.5.2). Due to intense background fluorescence all samples were compared to control infiltrations of the pDGB2α2 plasmid (fig. 3.32, F).

In this regard, cytosolic, chloroplast and apoplast compartmentalization of the Stth-YFP fusion protein was verified after transient transformation with the F13, H9 and H10 constructs, respectively (fig. 3.1). Non-targeted Stth revealed an intense fluorescence signal in the cytosol and nucleus of epidermal cells indicating a nucleocytosolic localization (fig 3.33, A). However, with a molecular weight of 86647 g/mol the fusion protein might not be able to enter the nucleus by passive diffusion through nucleoporins. It is assumed that proteins with a size up to 60 kDa (1 Da = 1 g/mol) passively diffuse into the organelle, whereas larger proteins need to contain a nuclear localization signal to be actively imported [208]. Therefore, Stth might rather accumulate around the organelle. However, Wang and Brattain suggested that the maximal protein size for passive diffusion might be larger than expected at least for mammalian cells [209]. Final evidence of Stth localization could be provided by co-transformation with a specific marker, such as CFP nucleocytoplasmic marker [107]. Moreover, nuclei, purified from disrupted cells by density centrifugation, can be investigated separately in terms of YFP fluorescence to exclude any passive diffusion of the fusion protein into the organelle. However, final localization of smaller non-fused Stth cannot be examined in this approach and requires immunodetection in purified nucleus fractions. Furthermore, targeting of the Stth-YFP fusion protein by the chloroplast signal peptide resulted in a bright fluorescence signal of chloroplasts of spongy mesophyll cells (fig. 3.33, B). Therefore, it was assumed that recombinant enzymes fused to the CP signal peptide were successfully transported into

chloroplasts. In addition, only a faint fluorescence signal was observed in the apoplastic space of epidermal cells after translocation of Stth-YFP fused to the ER signal peptide (fig. 3.33, C). This result correlates with observations of previous analyses concerning enzymatic activity of tryptophan halogenases *in planta* (see 3.1.2.1). One reason for this weak fluorescence signal could be the sensitivity of YFP to acidic conditions, which can prevail in the extracellular space. In this regard, Young *et al.* could show extinction of the YFP fluorescence signal at pH of 5.5 [210]. Given that the apoplastic pH can decrease to 5 [209], localization studies should be undertaken using a fluorescent dye insensitive for such acidic conditions like the red fluorescent protein (RFP) [211, 212]. Additionally, protein degradation by proteases localized in the apoplast, could account for the poor signal intensity. Proteases play a key role in the physiological development and plant defense mechanisms and therefore might affect the stability of recombinant proteins [141]. To stabilize heterologous enzymes in the extracellular compartment, biosynthesis of specific protease inhibitors can support their accumulation [142]. Additionally, localization of the strictosidine synthase STRV214M fused to its endogenous vacuolar targeting signal ErV was determined by C-terminal fusion to the YFP protein (fig. 3.28, T5). The N-terminal signal peptide facilitates translocation into the vacuole as exemplified by Guirimand *et al.* [107]. However, no fluorescence signal was observed for the vacuolar targeted STRV214M-YFP fusion protein (fig 3.32, D). As already discussed, YFP fluorescence might diminish in acidic conditions of the vacuole. Moreover, slight BFP fluorescence signal of STRV214M fusion protein (fig. 3.28, T14) was detectable after transformation of tobacco leaves, when examining appropriate accumulation of recombinant enzymes *in planta* (see 3.3.2). Therefore, ErV fusion to RFP or BFP might be beneficial to verify the functionality of the GB adapted targeting signal ErV. Finally, localization of 2A6mut anchored within the ER membrane was analyzed by N-terminal fusion to its endogenous signal peptide 2A6targ as well as C-terminal fusion to the YFP protein (fig. 3.14, Q2). Microscopy revealed an intense fluorescence signal of nuclei in epidermal cells, indicating that 2A6mut is associated to the nucleus or ER membrane (fig. 3.31, E). However, this observation differs from other fluorescent microscopy results addressing the subcellular localization of cytochrome P450s, such as SLS [108]. Given, that the enzymatic activity of 2A6mut was confirmed by indican accumulation after P450 catalyzed oxidation of indole, co-localization of 2A6mut and NADPH-cytochrome P450 oxidoreductases in the ER-membrane is very likely (see 3.2.6.1, 3.2.7).

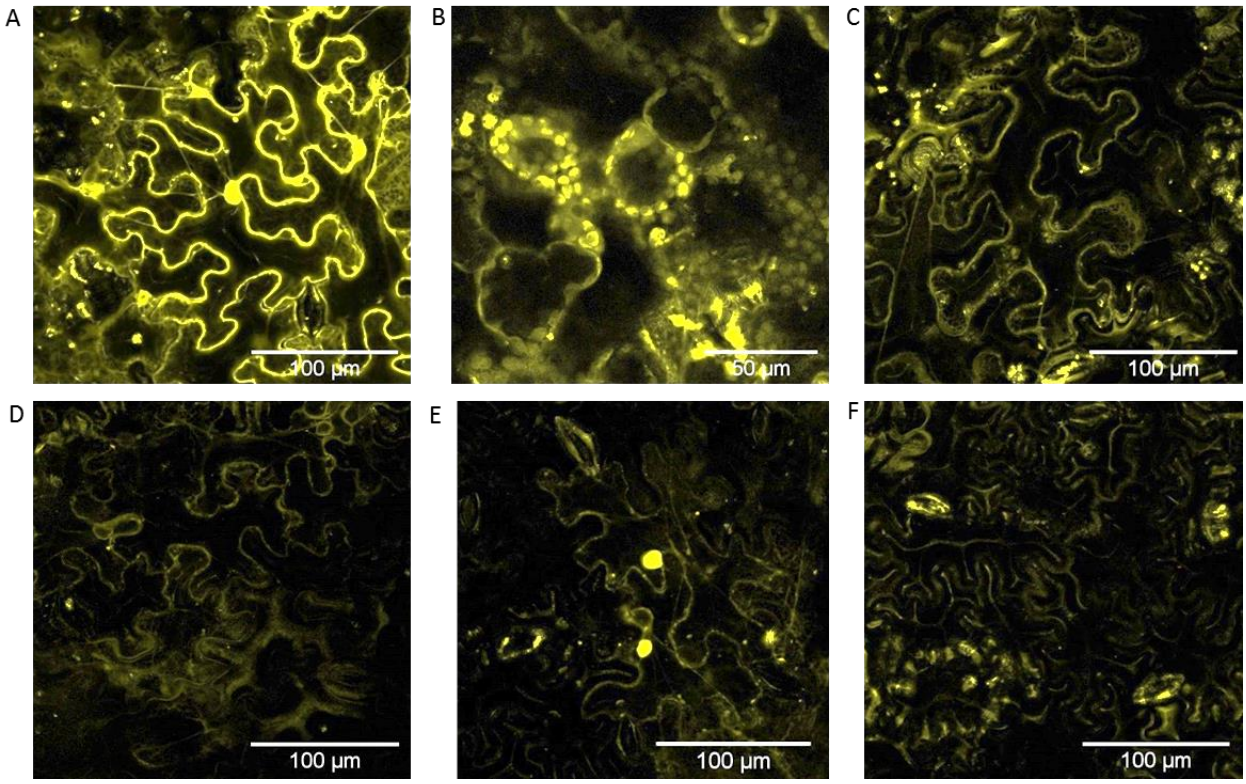

Figure 3.33: Verification of enzymes fused to different targeting signals. A: The non-targeted Stth-YFP fusion protein exhibited cytosolic localization. **B:** Chloroplast localized Stth-YFP fusion protein **C:** A faint fluorescence signal was detected after translocation of Stth-YFP to the apoplast. **D:** Vacuolar targeted STRV214M-YFP fusion protein revealed no fluorescence signal compared to negative controls. **E:** 2A6mut-YFP fusion protein anchored within the ER membrane. **F:** Negative control of the empty vector transformation.

The investigation of transport of recombinant enzymes to different cellular compartments by CLSM provided evidence for correct localization of Stth in the cytosol and chloroplasts as well as anchoring of 2A6mut by its endogenous signal peptide within the ER membrane. However, successful translocation of recombinant enzymes to the apoplast and vacuole remains unclear. To exclude any influence on fluorescence signal by acidic conditions that might prevail in both compartments YFP can be replaced by other less sensitive proteins, such as RFP [211, 212]. Due to the fact that all signal peptides used in course of the presented studies were characterized by previous studies, specific cellular markers were not co-localized in this approach. However, for final verification of a nucleocytosolic localization of non-targeted halogenases as well as ER-anchored 2A6mut, co-localization of appropriate markers, such as “ER"-mcherry (CD3-960) or the CFP nucleocytoplasmic marker might be useful [107, 213].

4. Conclusion

Characterization of the flavin-dependent tryptophan halogenases, RebH wt and Stth, demonstrated the extremely efficient halogenation of tryptophan and tryptamine in cytosol and chloroplasts. Moreover, in chloroplasts both enzymes were functional without the partner reductase RebF, indicating reduction of flavin by other reductases or through the photosystem (PS). The halogenases also catalyzed substitution of bromide to tryptophan, which leads to the assumption that RebH wt and Stth might also be capable of brominating tryptamine. Due to these characteristics, RebH wt and Stth represent promising tools for molecular engineering of biosynthetic pathways and the production of valuable metabolites *in planta*. Moreover, the variety of halogenated products can be augmented by introduction of an additional tryptophan 5-halogenase, such as PyrH from *Streptomyces rugosporus* [24]. The combination of different halogenases, the addition of TDC, and the supplementation of bromide could facilitate the biosynthesis of a huge variety of di- or tri-halogenated tryptophans and tryptamines. In contrast to the efficient biosynthesis of halogenated products by wild type enzymes, the 7-halogenase mutant, RebH Y455W, revealed only low catalytic activity *in planta*. This finding led to the conclusion that this modified halogenase might be not suitable for application in plant synthetic biology.

The modification of valuable products through introduction of tryptophan halogenases was demonstrated for the indoxyl biosynthetic pathway. Following transient transformation of *N. benthamiana*, 6-and 7-chloroindican accumulated in high levels within the cells. Final product yields can be further increased by introduction of a tryptophan insensitive anthranilate synthase [161, 182]. This modification could also be beneficial to enhance the BX1 and 2A6mut catalyzed pathway, which had only poor product outcome in transient approaches. Moreover, indican biosynthesis could be improved by integration of a recombinant glycosyltransferases, such as arbutin synthase in order to facilitate indoxyl glycosylation by a recombinant enzyme rather than endogenous ones [182]. Additionally, establishing transgenic cell cultures could enable a constitutive production of halogenated indican derivatives *in planta*. Furthermore, integration of the 5-halogenase PyrH probably results in accumulation of 5-chloroindican and thereby widens the product spectrum. The biosynthesis of brominated indican derivatives represents another possibility to increase the pathway variety. Given that 2A6mut is capable of bromoindole oxidation, the substrate scope of TnaA needs to be determined. In case of an inefficient conversion of bromotryptophan by TnaA different enzyme variants, established by amino acid exchange within the active site of the enzyme, can be tested [75, 175, 183]. Finally,

co-localization of TnaA and 2A6mut in chloroplasts resulted in high indican levels. This finding indicates that the required electrons, usually supplied by co-localized NADPH-cytochrome P450 oxidoreductases, are channeled to the cytochrome P450 in a highly efficient manner presumably from the PSI [160]. This fact offers completely new possibilities for molecular engineering of P450 catalyzed reactions like reconstitution of biosynthetic pathways through chloroplast transformation. Besides the huge number of these organelles in mesophyll cells, accumulation of recombinant enzymes is not affected by post transcriptional gene silencing. In general, establishing transgenic tobacco plants by either nuclear or chloroplast transformation simplifies and probably enhances chloroindican biosynthesis through constitutive gene expression. Taken together, the modified indoxyl biosynthetic pathway offers several possibilities for additional optimization in terms of an increased metabolic flux or enlargement of the product range. Moreover, it represents a promising alternative to chemical synthesis of halogenated indican derivatives and enables an efficient and economic biosynthesis of these valuable metabolites *in planta*.

Finally, modification of the strictosidine biosynthetic pathway was investigated by introduction of the halogenase mutant RebH Y455W. However, reconstitution of this pathway did not result in biosynthesis of the anticipated intermediates or strictosidine. Moreover, in contrast to published results, RebH Y455W and STRV21M revealed only low catalytic activity. To facilitate the production of halogenated strictosidine *in planta*, potential bottlenecks need to be circumvented to increase the metabolic flux. Additionally, constitutive strictosidine biosynthesis is potentially more efficient using STR. However, halogenated tryptamine derivatives will not be converted by the wild type enzyme and presumably requires the transformation of several copies of *strV214M* to facilitate chlorostrictosidine accumulation [117]. Furthermore, halogenation of the tryptamine precursor could be enhanced by introduction of wild type halogenases, which were proven to efficiently convert this metabolite in contrast to RebH Y455W. Finally, product yields can be increased by reconstitution of the entire pathway into chloroplasts. Given that enzymatic activity of recombinant P450s could be maintained by electron transfer from PSI, as previously demonstrated for 2A6mut, chloroplast transformation represents a promising tool for secologanin or strictosidine biosynthesis *in planta*. Due to the high value of strictosidine as precursor of all MIAs like the pharmacologically important secondary metabolites Vincristine and Vinblastine, production of this molecule is of main interest. Even though first reconstitution was not successful, further optimization might lead to efficient biosynthesis of either strictosidine or halogenated derivatives *in planta*.

5. References

1. **Gribble, G.W.,** The diversity of naturally produced organohalogens. Chemosphere, 2003. 52: p. 289-97.

2. **van Wageningen, A.M., Kirkpatrick, P. N., Williams, D. H., Harris, B. R., Kershaw, J. K., Lennard, N. J., Jones, M., Jones, S. J. & Solenberg, P. J.,** Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. Chemistry & Biology, 1998. 5: p. 155–162.

3. **Piraee, M.R.L.W., Leo C. Vining,** Biosynthesis of the dichloroacetyl component of chloramphenicol in Streptomyces venezuelae ISP5230: genes required for halogenation. Microbiology, 2004. 150: p. 85-94.

4. **César Sánchez, I.A.B., Alfredo F. Braña, Jürgen Rohr, Carmen Méndez, José A. Salas,** The biosynthetic gene cluster for the antitumor rebeccamycin: characterization and generation of indolocarbazole derivatives. Chemistry & Biology, 2002. 9: p. 519–531.

5. **Neumann, C.S., D.G. Fujimori, and C.T. Walsh,** Halogenation strategies in natural product biosynthesis. Chem Biol, 2008. 15: p. 99-109.

6. **Payne, J.T., M.C. Andorfer, and J.C. Lewis,** Regioselective arene halogenation using the FAD-dependent halogenase RebH. Angew Chem Int Ed Engl, 2013. 52: p. 5271-4.

7. **Holzer, M., W. Burd, H.U. Reissig, and K.H. van Pee,** Substrate specificity and regioselectivity of tryptophan 7-halogenase from Pseudomonas fluorescens BL915. Advanced Synthesis & Catalysis, 2001. 343: p. 591-595.

8. **Blasiak, L.C. and C.L. Drennan,** Structural perspective on enzymatic halogenation. Acc Chem Res, 2009. 42: p. 147-55.

9. **Smith, K. and G.A. El-Hiti,** Regioselective control of electrophilic aromatic substitution reactions. Current Organic Synthesis, 2004. 1: p. 253-274.

10. **Paterson, I. and E.A. Anderson,** The renaissance of natural products as drug candidates. Science, 2005. 310: p. 451-453.

11. **Runguphan, W., X. Qu, and S.E. O'Connor,** Integrating carbon-halogen bond formation into medicinal plant metabolism. Nature, 2010. 468: p. 461-4.

12. **Miettinen, K., L. Dong, N. Navrot, T. Schneider, V. Burlat, J. Pollier, L. Woittiez, S. van der Krol, R. Lugan, T. Ilc, R. Verpoorte, K.M. Oksman-Caldentey, E. Martinoia, H. Bouwmeester, A. Goossens, J. Memelink, and D. Werck-Reichhart,** The seco-iridoid pathway from *Catharanthus roseus*. Nat Commun, 2014. 5: p. 3606.

13. **Farhi, M., E. Marhevka, J. Ben-Ari, A. Algamas-Dimantov, Z. Liang, V. Zeevi, O. Edelbaum, B. Spitzer-Rimon, H. Abeliovich, B. Schwartz, T. Tzfira, and A. Vainstein,** Generation of the potent anti-malarial drug artemisinin in tobacco. Nat Biotechnol, 2011. 29: p. 1072-4.

14. **Warzecha, H., A. Frank, M. Peer, E.M. Gillam, F.P. Guengerich, and M. Unger,** Formation of the indigo precursor indican in genetically engineered tobacco plants and cell cultures. Plant Biotechnol J, 2007. 5: p. 185-91.

15. **Liu, C.J., J.W. Blount, C.L. Steele, and R.A. Dixon,** Bottlenecks for metabolic engineering of isoflavone glycoconjugates in *Arabidopsis*. Proc Natl Acad Sci U S A, 2002. 99: p. 14578-83.

16. **Cankar, K., E. Jongedijk, M. Klompmaker, T. Majdic, R. Mumm, H. Bouwmeester, D. Bosch, and J. Beekwilder,** (+)-Valencene production in *Nicotiana benthamiana* is increased by down-regulation of competing pathways. Biotechnology Journal, 2015. 10: p. 180-189.

17. **Capell, T. and P. Christou,** Progress in plant metabolic engineering. Current Opinion in Biotechnology, 2004. 15: p. 148-154.

18. **Lucker, J., H.J. Bouwmeester, W. Schwab, J. Blaas, L.H. van der Plas, and H.A. Verhoeven,** Expression of Clarkia S-linalool synthase in transgenic petunia plants results in the accumulation of S-linalyl-beta-D-glucopyranoside. Plant J, 2001. 27: p. 315-24.

19. **Dueber, J.E., G.C. Wu, G.R. Malmirchegini, T.S. Moon, C.J. Petzold, A.V. Ullal, K.L. Prather, and J.D. Keasling,** Synthetic protein scaffolds provide modular control over metabolic flux. Nat Biotechnol, 2009. 27: p. 753-9.

20. **Sarrion-Perdigones, A., E.E. Falconi, S.I. Zandalinas, P. Juarez, A. Fernandez-del-Carmen, A. Granell, and D. Orzaez,** GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. PLoS One, 2011. 6: p. e21622.

21. **Sarrion-Perdigones, A., M. Vazquez-Vilar, J. Palaci, B. Castelijns, J. Forment, P. Ziarsolo, J. Blanca, A. Granell, and D. Orzaez,** GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. Plant Physiol, 2013. 162: p. 1618-31.

22. **Dorrestein, P.C., E. Yeh, S. Garneau-Tsodikova, N.L. Kelleher, and C.T. Walsh,** Dichlorination of a pyrrolyl-S-carrier protein by FADH2-dependent halogenase PltA during pyoluteorin biosynthesis. Proc Natl Acad Sci U S A, 2005. 102: p. 13843-8.

23. **Vaillancourt, F.H., E. Yeh, D.A. Vosburg, S. Garneau-Tsodikova, and C.T. Walsh,** Nature's inventory of halogenation catalysts: oxidative strategies predominate. Chem Rev, 2006. 106: p. 3364-78.

24. **Zehner, S., A. Kotzsch, B. Bister, R.D. Sussmuth, C. Mendez, J.A. Salas, and K.H. van Pee,** A regioselective tryptophan 5-halogenase is involved in pyrroindomycin biosynthesis in Streptomyces rugosporus LL-42D005. Chem Biol, 2005. 12: p. 445-52.

25. **Yeh, E., S. Garneau, and C.T. Walsh,** Robust in vitro activity of RebF and RebH, a two-component reductase/halogenase, generating 7-chlorotryptophan during rebeccamycin biosynthesis. Proc Natl Acad Sci U S A, 2005. 102: p. 3960-5.

26. **Heemstra, J.R., Jr. and C.T. Walsh,** Tandem action of the O2- and FADH2-dependent halogenases KtzQ and KtzR produce 6,7-dichlorotryptophan for kutzneride assembly. J Am Chem Soc, 2008. 130: p. 14024-5.

27. **Bitto, E., Y. Huang, C.A. Bingman, S. Singh, J.S. Thorson, and G.N. Phillips, Jr.,** The structure of flavin-dependent tryptophan 7-halogenase RebH. Proteins, 2008. 70: p. 289-93.

28. **Glenn, W.S., E. Nims, and S.E. O'Connor,** Reengineering a tryptophan halogenase to preferentially chlorinate a direct alkaloid precursor. J Am Chem Soc, 2011. 133: p. 19346-9.

29. **Zeng, J. and J. Zhan,** Characterization of a tryptophan 6-halogenase from *Streptomyces toxytricini*. Biotechnol Lett, 2011. 33: p. 1607-13.

30. **Dong, C., S. Flecks, S. Unversucht, C. Haupt, K.H. van Pee, and J.H. Naismith,** Tryptophan 7-halogenase (PrnA) structure suggests a mechanism for regioselective chlorination. Science, 2005. 309: p. 2216-9.

31. **Zhu, X., W. De Laurentis, K. Leang, J. Herrmann, K. Ihlefeld, K.H. van Pee, and J.H. Naismith,** Structural insights into regioselectivity in the enzymatic chlorination of tryptophan. J Mol Biol, 2009. 391: p. 74-85.

32. **Yeh, E., L.J. Cole, E.W. Barr, J.M. Bollinger, Jr., D.P. Ballou, and C.T. Walsh,** Flavin redox chemistry precedes substrate chlorination during the reaction of the flavin-dependent halogenase RebH. Biochemistry, 2006. 45: p. 7904-12.

33. **Yeh, E., L.C. Blasiak, A. Koglin, C.L. Drennan, and C.T. Walsh,** Chlorination by a long-lived intermediate in the mechanism of flavin-dependent halogenases. Biochemistry, 2007. 46: p. 1284-92.

34. **Seibold, C., H. Schnerr, J. Rumpf, A. Kunzendorf, C. Hatscher, T. Wage, A.J. Ernyei, C. Dong, J.H. Naismith, and K.-H. Van Pée,** A flavin-dependent tryptophan 6-halogenase and its use in modification of pyrrolnitrin biosynthesis. Biocatalysis and Biotransformation, 2006. 24: p. 401-408.

35. **Wani, M.C., H.L. Taylor, M.E. Wall, P. Coggon, and A.T. Mcphail,** Plant Antitumor Agents .6. Isolation and Structure of Taxol, a Novel Antileukemic and Antitumor Agent from Taxus-Brevifolia. Journal of the American Chemical Society, 1971. 93: p. 2325-&.

36. **Onoyovwe, A., J.M. Hagel, X. Chen, M.F. Khan, D.C. Schriemer, and P.J. Facchini,** Morphine biosynthesis in opium poppy involves two cell types: sieve elements and laticifers. Plant Cell, 2013. 25: p. 4110-22.

37. **van Der Heijden, R., D.I. Jacobs, W. Snoeijer, D. Hallard, and R. Verpoorte,** The *Catharanthus* alkaloids: pharmacognosy and biotechnology. Curr Med Chem, 2004. 11: p. 607-28.

38. **Wink, M.**, Introduction, in Annual Plant Reviews Volume 39: Functions and Biotechnology of Plant Secondary Metabolites. 2010, Wiley-Blackwell. p. 1-20.

39. **Ferrer-Miralles, N., P. Saccardo, J.L. Corchero, Z. Xu, and E. Garcia-Fruitos,** General introduction: recombinant protein production and purification of insoluble proteins. Methods Mol Biol, 2015. 1258: p. 1-24.

40. **Morant, M., S. Bak, B.L. Moller, and D. Werck-Reichhart,** Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation. Current Opinion in Biotechnology, 2003. 14: p. 151-162.

41. **Jensen, K. and B.L. Moller,** Plant NADPH-cytochrome P450 oxidoreductases. Phytochemistry, 2010. 71: p. 132-41.

42. **Gillam, E.M., Z. Guo, and F.P. Guengerich,** Expression of modified human cytochrome P450 2E1 in *Escherichia coli*, purification, and spectral and catalytic properties. Arch Biochem Biophys, 1994. 312: p. 59-66.

43. **Blake, J.A., M. Pritchard, S. Ding, G.C. Smith, B. Burchell, C.R. Wolf, and T. Friedberg,** Coexpression of a human P450 (CYP3A4) and P450 reductase generates a highly functional monooxygenase system in Escherichia coli. FEBS Lett, 1996. 397: p. 210-4.

44. **Gillam, E.M., T. Baba, B.R. Kim, S. Ohmori, and F.P. Guengerich,** Expression of modified human cytochrome P450 3A4 in *Escherichia coli* and purification and reconstitution of the enzyme. Arch Biochem Biophys, 1993. 305: p. 123-31.

45. **Siddiqui, M.S., K. Thodey, I. Trenchard, and C.D. Smolke,** Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools. FEMS Yeast Res, 2012. 12: p. 144-70.

46. **Farhi, M., E. Marhevka, T. Masci, E. Marcos, Y. Eyal, M. Ovadis, H. Abeliovich, and A. Vainstein,** Harnessing yeast subcellular compartments for the production of plant terpenoids. Metab Eng, 2011. 13: p. 474-81.

47. **Oey, M., M. Lohse, B. Kreikemeyer, and R. Bock,** Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. Plant J, 2009. 57: p. 436-45.

48. **Fischer, R. and N. Emans,** Molecular farming of pharmaceutical proteins. Transgenic Res, 2000. 9: p. 279-99; discussion 277.

49. **Walmsley, A.M. and C.J. Arntzen,** Plants for delivery of edible vaccines. Curr Opin Biotechnol, 2000. 11: p. 126-9.

50. **Schillberg, S., R. Fischer, and N. Emans,** Molecular farming of recombinant antibodies in plants. Cell Mol Life Sci, 2003. 60: p. 433-45.

51. **Krenek, P., O. Samajova, I. Luptovciak, A. Doskocilova, G. Komis, and J. Samaj,** Transient plant transformation mediated by *Agrobacterium tumefaciens*: Principles, methods and applications. Biotechnol Adv, 2015. 33: p. 1024-42.

52. **Zupan, J., T.R. Muth, O. Draper, and P. Zambryski,** The transfer of DNA from *Agrobacterium tumefaciens* into plants: a feast of fundamental insights. Plant J, 2000. 23: p. 11-28.

53. **Sanford, J.C., T.M. Klein, E.D. Wolf, and N. Allen,** Delivery of Substances into Cells and Tissues Using a Particle Bombardment Process. Particulate Science and Technology, 1987. 5: p. 27-37.

54. **Maliga, P.,** Plastid transformation in higher plants. Annu Rev Plant Biol, 2004. 55: p. 289-313.

55. **Buchanan, B.B., W. Gruissem, and R.L. Jones,** Biochemistry & molecular biology of plants. 2000, Rockville, Md.: American Society of Plant Physiologists. xxxix, 1367 p.

56. **Daniell, H., M.S. Khan, and L. Allison,** Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology. Trends Plant Sci, 2002. 7: p. 84-91.

57. **Glowacki, E.D., G. Voss, L. Leonat, M. Irimia-Vladu, S. Bauer, and N.S. Sariciftci,** Indigo and Tyrian Purple - From Ancient Natural Dyes to Modern Organic Semiconductors. Israel Journal of Chemistry, 2012. 52: p. 540-551.

58. **Clark, R.J.H., C.J. Cooksey, M.A.M. Daniels, and R. Withnall,** Indigo, Woad, and Tyrian Purple - Important Vat Dyes from Antiquity to the Present. Endeavour, 1993. 17: p. 191-199.

59. **Xia, Z.Q. and M.H. Zenk,** Biosynthesis of Indigo Precursors in Higher-Plants. Phytochemistry, 1992. 31: p. 2695-2697.

60. **Epstein, E., M.W. Nabors, and B.B. Stowe,** Origin of Indigo of Woad. Nature, 1967. 216: p. 547-549.

61. **Robertson, A.,** CCLI.-Syntheses of glucosides. Part I. The synthesis of indican. Journal of the Chemical Society (Resumed), 1927: p. 1937-1943.

62. **Cooksey, C.J.,** Tyrian purple: 6,6 '-dibromoindigo and related compounds. Molecules, 2001. 6: p. 736-769.

63. **Baker, J.T. and Sutherla.Md,** Pigments of Marine Animals .8. Precursors of 6,6'-Dibromoindigotin (Tyrian Purple) from Mollusc *Dicathais orbita* Gmelin. Tetrahedron Letters, 1968: p. 43-&.

64. **Westley, C. and K. Benkendorff,** The distribution of precursors and biosynthetic enzymes required for Tyrian purple genesis in the hypobranchial gland, gonoduct, an egg masses of *Dicathais orbita* (Gmelin, 1791) (Neogastropoda: Muricidae). Nautilus, 2009. 123: p. 148-153.

65. **Christophersen, C., F. Watjen, O. Buchardt, and U. Anthoni,** Revised Structure of Tyriverdin - Precursor of Tyrian Purple. Tetrahedron, 1978. 34: p. 2779-2781.

66. **Mcgovern, P.E. and R.H. Michel,** Royal Purple-Dye - the Chemical Reconstruction of the Ancient Mediterranean Industry. Accounts of Chemical Research, 1990. 23: p. 152-158.

67. **Jannun, R. and E.L. Coe,** Bromoperoxidase from the Marine Snail, Murex-Trunculus. Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology, 1987. 88: p. 917-922.

68. **Frey, M., P. Chomet, E. Glawischnig, C. Stettner, S. Grun, A. Winklmair, W. Eisenreich, A. Bacher, R.B. Meeley, S.P. Briggs, K. Simcox, and A. Gierl,** Analysis of a chemical plant defense mechanism in grasses. Science, 1997. 277: p. 696-9.

69. **Hyde, C.C., S.A. Ahmed, E.A. Padlan, E.W. Miles, and D.R. Davies,** Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from *Salmonella typhimurium*. J Biol Chem, 1988. 263: p. 17857-71.

70. **Kharasch, E.D., D.C. Hankins, and K.E. Thummel,** Human kidney methoxyflurane and sevoflurane metabolism. Intrarenal fluoride production as a possible mechanism of methoxyflurane nephrotoxicity. Anesthesiology, 1995. 82: p. 689-99.

71. **Spracklin, D.K., K.E. Thummel, and E.D. Kharasch,** Human reductive halothane metabolism in vitro is catalyzed by cytochrome P450 2A6 and 3A4. Drug Metab Dispos, 1996. 24: p. 976-83.

72. **Messina, E.S., R.F. Tyndale, and E.M. Sellers,** A major role for CYP2A6 in nicotine C-oxidation by human liver microsomes. J Pharmacol Exp Ther, 1997. 282: p. 1608-14.

73. **Denton, T.T., X. Zhang, and J.R. Cashman,** Nicotine-related alkaloids and metabolites as inhibitors of human cytochrome P-450 2A6. Biochem Pharmacol, 2004. 67: p. 751-6.

74. **Nakamura, K., M.V. Martin, and F.P. Guengerich,** Random mutagenesis of human cytochrome p450 2A6 and screening with indole oxidation products. Arch Biochem Biophys, 2001. 395: p. 25-31.

75. **Wu, Z.L., P. Aryal, O. Lozach, L. Meijer, and F.P. Guengerich,** Biosynthesis of new indigoid inhibitors of protein kinases using recombinant cytochrome P450 2A6. Chem Biodivers, 2005. 2: p. 51-65.

76. **Gillam, E.M., A.M. Aguinaldo, L.M. Notley, D. Kim, R.G. Mundkowski, A.A. Volkov, F.H. Arnold, P. Soucek, J.J. DeVoss, and F.P. Guengerich,** Formation of indigo by recombinant mammalian cytochrome P450. Biochem Biophys Res Commun, 1999. 265: p. 469-72.

77. **Trifonov, S., Y. Yamashita, M. Kase, M. Maruyama, and T. Sugimoto,** Overview and assessment of the histochemical methods and reagents for the detection of beta-galactosidase activity in transgenic animals. Anat Sci Int, 2015.

78. **de Siloniz, M.I., M.J. Valderrama, and J.M. Peinado,** A chromogenic medium for the detection of yeasts with beta-galactosidase and beta-glucosidase activities from intermediate moisture foods. J Food Prot, 2000. 63: p. 651-4.

79. **Kiernan, J.A.,** Indigogenic substrates for detection and localization of enzymes. Biotech Histochem, 2007. 82: p. 73-103.

80. **Horwitz, J.P., J. Chua, R.J. Curby, A.J. Tomson, M.A. Darooge, B.E. Fisher, J. Mauricio, and I. Klundt,** Substrates for Cytochemical Demonstration of Enzyme Activity. I. Some Substituted 3-Indolyl-Beta-D-Glycopyranosides. J Med Chem, 1964. 7: p. 574-5.

81. **Juers, D.H., B.W. Matthews, and R.E. Huber,** LacZ beta-galactosidase: structure and function of an enzyme of historical and molecular biological importance. Protein Sci, 2012. 21: p. 1792-807.

82. **Cotson, S. and S.J. Holt,** Studies in Enzyme Cytochemistry .4. Kinetics of Aerial Oxidation of Indoxyl and Some of Its Halogen Derivatives. Proceedings of the Royal Society Series B-Biological Sciences, 1958. 148: p. 506-519.

83. Copyright © 2015 Sigma-Aldrich Co. LLC. cited [23.10.2015] Available from: http://www.sigmaaldrich.com/catalog/product/sial/b4252?lang=de&region=DE.

84. Copyright © 2015 Sigma-Aldrich Co. LLC. cited [23.10.2015] Available from: http://www.sigmaaldrich.com/catalog/product/sigma/93546?lang=de&region=DE.

85. **Guirimand, G., V. Courdavault, B. St-Pierre, and V. Burlat**, Biosynthesis and Regulation of Alkaloids, in Plant Developmental Biology - Biotechnological Perspectives, E.C. Pua and M.R. Davey, Editors. 2010, Springer Berlin Heidelberg. p. 139-160.

86. **Stavrinides, A., E.C. Tatsis, E. Foureau, L. Caputi, F. Kellner, V. Courdavault, and S.E. O'Connor,** Unlocking the diversity of alkaloids in *Catharanthus roseus*: nuclear localization suggests metabolic channeling in secondary metabolism. Chem Biol, 2015. 22: p. 336-41.

87. **O'Connor, S.E. and J.J. Maresh,** Chemistry and biology of monoterpene indole alkaloid biosynthesis. Nat Prod Rep, 2006. 23: p. 532-47.

88. **Simkin, A.J., K. Miettinen, P. Claudel, V. Burlat, G. Guirimand, V. Courdavault, N. Papon, S. Meyer, S. Godet, B. St-Pierre, N. Giglioli-Guivarc'h, M.J. Fischer, J. Memelink, and M. Clastre,** Characterization of the plastidial geraniol synthase from Madagascar periwinkle which initiates the monoterpenoid branch of the alkaloid pathway in internal phloem associated parenchyma. Phytochemistry, 2012.

89. **Morgan, J.A. and J.V. Shanks,** Determination of metabolic rate-limitations by precursor feeding in *Catharanthus roseus* hairy root cultures. J Biotechnol, 2000. 79: p. 137-45.

90. **Lee-Parsons, C.W. and A.J. Royce,** Precursor limitations in methyl jasmonate-induced *Catharanthus roseus* cell cultures. Plant Cell Rep, 2006. 25: p. 607-12.

91. **Banerjee, A. and T.D. Sharkey,** Methylerythritol 4-phosphate (MEP) pathway metabolic regulation. Nat Prod Rep, 2014. 31: p. 1043-55.

92. **Höfer, R., L. Dong, F. Andre, J.F. Ginglinger, R. Lugan, C. Gavira, S. Grec, G. Lang, J. Memelink, S. Van der Krol, H. Bouwmeester, and D. Werck-Reichhart,** Geraniol hydroxylase and hydroxygeraniol oxidase activities of the CYP76 family of cytochrome P450 enzymes and potential for engineering the early steps of the (seco)iridoid pathway. Metab Eng, 2013. 20: p. 221-32.

93. **Collu, G., N. Unver, A.M. Peltenburg-Looman, R. van der Heijden, R. Verpoorte, and J. Memelink,** Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis. FEBS Lett, 2001. 508: p. 215-20.

94. **Ikeda, H., N. Esaki, S. Nakai, K. Hashimoto, S. Uesato, K. Soda, and T. Fujita,** Acyclic monoterpene primary alcohol:NADP+ oxidoreductase of *Rauwolfia serpentina* cells: the key enzyme in biosynthesis of monoterpene alcohols. J Biochem, 1991. 109: p. 341-7.

95. **Krithika, R., P.L. Srivastava, B. Rani, S.P. Kolet, M. Chopade, M. Soniya, and H.V. Thulasiram,** Characterization of 10-hydroxygeraniol dehydrogenase from *Catharanthus roseus* reveals cascaded enzymatic activity in iridoid biosynthesis. Sci Rep, 2015. 5: p. 8258.

96. **Geu-Flores, F., N.H. Sherden, V. Courdavault, V. Burlat, W.S. Glenn, C. Wu, E. Nims, Y. Cui, and S.E. O'Connor,** An alternative route to cyclic terpenes by reductive cyclization in iridoid biosynthesis. Nature, 2012. 492: p. 138-42.

97. **Chen, F., D. Tholl, J. Bohlmann, and E. Pichersky,** The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. Plant J, 2011. 66: p. 212-29.

98. **Degenhardt, J., T.G. Kollner, and J. Gershenzon,** Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry, 2009. 70: p. 1621-37.

99. **Salim, V., B. Wiens, S. Masada-Atsumi, F. Yu, and V. De Luca,** 7-deoxyloganetic acid synthase catalyzes a key 3 step oxidation to form 7-deoxyloganetic acid in *Catharanthus roseus* iridoid biosynthesis. Phytochemistry, 2014. 101: p. 23-31.

100. **Asada, K., V. Salim, S. Masada-Atsumi, E. Edmunds, M. Nagatoshi, K. Terasaka, H. Mizukami, and V. De Luca,** A 7-deoxyloganetic acid glucosyltransferase contributes a key step in secologanin biosynthesis in *Madagascar periwinkle*. Plant Cell, 2013. 25: p. 4123-34.

101. **Salim, V., F. Yu, J. Altarejos, and V. De Luca,** Virus-induced gene silencing identifies *Catharanthus roseus* 7-deoxyloganic acid-7-hydroxylase, a step in iridoid and monoterpene indole alkaloid biosynthesis. Plant J, 2013. 76: p. 754-65.

102. **Murata, J., J. Roepke, H. Gordon, and V. De Luca,** The leaf epidermome of *Catharanthus roseus* reveals its biochemical specialization. Plant Cell, 2008. 20: p. 524-42.

103. **Irmler, S., G. Schroder, B. St-Pierre, N.P. Crouch, M. Hotze, J. Schmidt, D. Strack, U. Matern, and J. Schroder,** Indole alkaloid biosynthesis in *Catharanthus roseus*: new enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase. Plant J, 2000. 24: p. 797-804.

104. **Treimer, J.F. and M.H. Zenk,** Purification and properties of strictosidine synthase, the key enzyme in indole alkaloid formation. Eur J Biochem, 1979. 101: p. 225-33.

105. **Mizukami, H., H. Nordlov, S.L. Lee, and A.I. Scott,** Purification and properties of strictosidine synthetase (an enzyme condensing tryptamine and secologanin) from *Catharanthus roseus* cultured cells. Biochemistry, 1979. 18: p. 3760-3.

106. **De Luca, V., C. Marineau, and N. Brisson,** Molecular cloning and analysis of cDNA encoding a plant tryptophan decarboxylase: comparison with animal dopa decarboxylases. Proc Natl Acad Sci U S A, 1989. 86: p. 2582-6.

107. **Guirimand, G., V. Courdavault, A. Lanoue, S. Mahroug, A. Guihur, N. Blanc, N. Giglioli-Guivarc'h, B. St-Pierre, and V. Burlat,** Strictosidine activation in Apocynaceae: towards a "nuclear time bomb"? BMC Plant Biol, 2010. 10: p. 182.

108. **Guirimand, G., A. Guihur, O. Ginis, P. Poutrain, F. Hericourt, A. Oudin, A. Lanoue, B. St-Pierre, V. Burlat, and V. Courdavault,** The subcellular organization of strictosidine biosynthesis in *Catharanthus roseus* epidermis highlights several trans-tonoplast translocations of intermediate metabolites. FEBS J, 2011. 278: p. 749-63.

109. **de Waal, A., A.H. Meijer, and R. Verpoorte,** Strictosidine synthase from *Catharanthus roseus*: purification and characterization of multiple forms. Biochem J, 1995. 306 (Pt 2): p. 571-80.

110. **Ma, X., S. Panjikar, J. Koepke, E. Loris, and J. Stockigt,** The structure of *Rauvolfia serpentina* strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins. Plant Cell, 2006. 18: p. 907-20.

111. **van der Fits, L.,** ORCA3, a Jasmonate-Responsive Transcriptional Regulator of Plant Primary and Secondary Metabolism. Science, 2000. 289: p. 295-297.

112. **Aerts, R.J., D. Gisi, E. Decarolis, V. Deluca, and T.W. Baumann,** Methyl Jasmonate Vapor Increases the Developmentally Controlled Synthesis of Alkaloids in *Catharanthus* and *Cinchona* Seedlings. Plant Journal, 1994. 5: p. 635-643.

113. **Gantet, P., N. Imbault, M. Thiersault, and P. Doireau,** Inhibition of alkaloid accumulation by 2,4-D in *Catharanthus roseus* cell suspension is overcome by methyl jasmonate. Acta Botanica Gallica, 1997. 144: p. 501-508.

114. **El-Sayed, M. and R. Verpoorte,** Methyljasmonate accelerates catabolism of monoterpenoid indole alkaloids in *Catharanthus roseus* during leaf processing. Fitoterapia, 2005. 76: p. 83-90.

115. **Brown, S., M. Clastre, V. Courdavault, and S.E. O'Connor,** De novo production of the plant-derived alkaloid strictosidine in yeast. Proc Natl Acad Sci U S A, 2015. 112: p. 3205-10.

116. **Ziegler, J. and P.J. Facchini,** Alkaloid biosynthesis: metabolism and trafficking. Annu Rev Plant Biol, 2008. 59: p. 735-69.

117. **Bernhardt, P., E. McCoy, and S.E. O'Connor,** Rapid identification of enzyme variants for reengineered alkaloid biosynthesis in periwinkle. Chem Biol, 2007. 14: p. 888-97.

118. **Luo, Y., B.Z. Li, D. Liu, L. Zhang, Y. Chen, B. Jia, B.X. Zeng, H. Zhao, and Y.J. Yuan,** Engineered biosynthesis of natural products in heterologous hosts. Chem Soc Rev, 2015. 44: p. 5265-90.

119. **Engler, C., R. Kandzia, and S. Marillonnet,** A one pot, one step, precision cloning method with high throughput capability. PLoS One, 2008. 3: p. e3647.

120. **Weber, E., C. Engler, R. Gruetzner, S. Werner, and S. Marillonnet,** A modular cloning system for standardized assembly of multigene constructs. PLoS One, 2011. 6: p. e16765.

121. **Hellens, R.P., E.A. Edwards, N.R. Leyland, S. Bean, and P.M. Mullineaux,** pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. Plant Molecular Biology, 2000. 42: p. 819-832.

122. **Chi-Ham, C.L., S. Boettiger, R. Figueroa-Balderas, S. Bird, J.N. Geoola, P. Zamora, M. Alandete-Saez, and A.B. Bennett,** An intellectual property sharing initiative in agricultural biotechnology: development of broadly accessible technologies for plant transformation. Plant Biotechnol J, 2012. 10: p. 501-10.

123. **Roberts C, R., S., Smith,L.M., Nguyen,T.A., Yang,W.,, Nugrohu S, Ravi,K.S., Vijayachandra,K., Harcourt,R.L.,, Dransfield L, Desamero,N., Slamet,I., Hadjukiewicz,P., Svab,Z.,, Maliga P, Mayer,J.E., Keese,P.K., Kilian, A., Jefferson,R.A.**, A comprehensive set of modular vectors for advanced manipulations and efficient transformation of plants, in *pCAMBIA Vector Release Manual*. 1997.

124. **Kizer, L., D.J. Pitera, B.F. Pfleger, and J.D. Keasling,** Application of functional genomics to pathway optimization for increased isoprenoid production. Appl Environ Microbiol, 2008. 74: p. 3229-41.

125. **Pitera, D.J., C.J. Paddon, J.D. Newman, and J.D. Keasling,** Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. Metab Eng, 2007. 9: p. 193-207.

126. **Bloom, J.D., M.M. Meyer, P. Meinhold, C.R. Otey, D. MacMillan, and F.H. Arnold,** Evolving strategies for enzyme engineering. Curr Opin Struct Biol, 2005. 15: p. 447-52.

127. **Bertani, G.,** Studies on Lysogenesis .1. The Mode of Phage Liberation by Lysogenic *Escherichia coli*. Journal of Bacteriology, 1951. 62: p. 293-300.

128. **Murashige, T. and F. Skoog,** A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. Physiologia Plantarum, 1962. 15: p. 473-497.

129. **Elizabeth E. Hood, S.B.G., Leo S. Melchers, Andre Hoekema** New *Agrobacterium* helper plasmids for gene transfer to plants. 1993. 2: p. 208-218.

130. **Göpfer, K.**, Application of the GoldenBraid 2.0 modular cloning system for metabolic engineering of ajmaline biosynthesis. 2013, Technische Universität Darmstadt.

131. **Voinnet, O., S. Rivas, P. Mestre, and D. Baulcombe,** An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J, 2003. 33: p. 949-56.

132. **Mullis, K.B. and F.A. Faloona,** Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol, 1987. 155: p. 335-50.

133. **Emanuelsson, O., H. Nielsen, S. Brunak, and G. von Heijne,** Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol, 2000. 300: p. 1005-16.

134. **Morris, G.M., R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, and A.J. Olson,** AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem, 2009. 30: p. 2785-91.

135. **O'Boyle, N.M., M. Banck, C.A. James, C. Morley, T. Vandermeersch, and G.R. Hutchison,** Open Babel: An open chemical toolbox. J Cheminform, 2011. 3: p. 33.

136. **Wall, P.E.,** Thin-layer Chromatography. 2005.

137. **Lloyd R. Snyder, J.J.K., John W. Dolan**, The Chromatographic Process, in Introduction to Modern Liquid Chromatography, 3rd Edition. 2010, A John Wiley & Sons, Inc. p. 20-24.

138. **Sparkman, J.T.W.a.O.D.**, Introduction, in Introduction to Mass Spectrometry: Instrumentation, Applications, and Strategies for Data Interpretation. 2007, John Wiley & Sons, Ltd. p. 3-9.

139. **J. Throck Watson, O.D.S.**, Tandem-in-Space Mass Spectrometry (MS/MS), in Introduction to Mass Spectrometry: Instrumentation, Applications, and Strategies for Data Interpretation, 4th Edition. 2007, John Wiley & Sons Ltd. p. 186-192.

140. **Solomon, P.S. and R.P. Oliver,** The nitrogen content of the tomato leaf apoplast increases during infection by Cladosporium fulvum. Planta, 2001. 213: p. 241-9.

141. **Doran, P.M.,** Foreign protein degradation and instability in plants and plant tissue cultures. Trends Biotechnol, 2006. 24: p. 426-32.

142. **Goulet, C., M. Khalf, F. Sainsbury, M.A. D'Aoust, and D. Michaud,** A protease activity-depleted environment for heterologous proteins migrating towards the leaf cell apoplast. Plant Biotechnol J, 2012. 10: p. 83-94.

143. **Joosten, V. and W.J. van Berkel,** Flavoenzymes. Curr Opin Chem Biol, 2007. 11: p. 195-202.

144. **Walsh, C.T. and T.A. Wencewicz,** Flavoenzymes: versatile catalysts in biosynthetic pathways. Nat Prod Rep, 2013. 30: p. 175-200.

145. **Sandoval, F.J., Y. Zhang, and S. Roje,** Flavin nucleotide metabolism in plants: monofunctional enzymes synthesize fad in plastids. J Biol Chem, 2008. 283: p. 30890-900.

146. **Barbier-Brygoo, H., M. Vinauger, J. Colcombet, G. Ephritikhine, J. Frachisse, and C. Maurel,** Anion channels in higher plants: functional characterization, molecular structure and physiological role. Biochim Biophys Acta, 2000. 1465: p. 199-218.

147. **Sitbon, F., A. Ostin, B. Sundberg, O. Olsson, and G. Sandberg,** Conjugation of Indole-3-Acetic Acid (IAA) in Wild-Type and IAA-Overprodcing Transgenic Tobacco Plants, and Identification of the Main Conjugates by Frit-Fast Atom Bombardment Liquid Chromatography-Mass Spectrometry. Plant Physiol, 1993. 101: p. 313-320.

148. **Korasick, D.A., T.A. Enders, and L.C. Strader,** Auxin biosynthesis and storage forms. J Exp Bot, 2013. 64: p. 2541-55.

149. **Ljun, K., A.K. Hul, M. Kowalczyk, A. Marchant, J. Celenza, J.D. Cohen, and G. Sandberg,** Biosynthesis, conjugation, catabolism and homeostasis of indole-3-acetic acid in *Arabidopsis thaliana*. Plant Mol Biol, 2002. 50: p. 309-32.

150. **Tam, Y.Y., E. Epstein, and J. Normanly,** Characterization of auxin conjugates in *Arabidopsis*. Low steady-state levels of indole-3-acetyl-aspartate, indole-3-acetyl-glutamate, and indole-3-acetyl-glucose. Plant Physiol, 2000. 123: p. 589-96.

151. **Graham, H.N.,** Green tea composition, consumption, and polyphenol chemistry. Prev Med, 1992. 21: p. 334-50.

152. **Liszt, K.I., R. Eder, S. Wendelin, and V. Somoza,** Identification of Catechin, Syringic acid and Procyanidin B2 in Wine as Stimulants of Gastric Acid Secretion. J Agric Food Chem, 2015.

153. **Bronner, W.E. and G.R. Beecher,** Method for determining the content of catechins in tea infusions by high-performance liquid chromatography. J Chromatogr A, 1998. 805: p. 137-42.

154. **Urbank, S.**, Charakterisierung humaner Cytochrom-P450-Enzyme in transplastomen Tabakpflanzen am Beispiel des CYP2A6. 2012, Technische Universität Darmstadt.

155. **Deeley, M.C. and C. Yanofsky,** Nucleotide sequence of the structural gene for tryptophanase of *Escherichia coli* K-12. J Bacteriol, 1981. 147: p. 787-96.

156. **Wagner, B.**, Development of novel pathways for the biosynthesis of halogenated indican derivatives in *Nicotiana benthamiana*. 2015, Technische Universität Darmstadt.

157. **Röder, K.**, Molecular Engineering in Nicht-Modellpflanzen. 2014, Technische Universität Darmstadt.

158. **Pinero-Fernandez, S., C. Chimerel, U.F. Keyser, and D.K. Summers,** Indole transport across *Escherichia coli* membranes. J Bacteriol, 2011. 193: p. 1793-8.

159. **Li, Y., S. Baldauf, E.K. Lim, and D.J. Bowles,** Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. J Biol Chem, 2001. 276: p. 4338-43.

160. **Nielsen, A.Z., B. Ziersen, K. Jensen, L.M. Lassen, C.E. Olsen, B.L. Moller, and P.E. Jensen,** Redirecting photosynthetic reducing power toward bioactive natural product synthesis. ACS Synth Biol, 2013. 2: p. 308-15.

161. **Romero, R.M., M.F. Roberts, and J.D. Phillipson,** Anthranilate synthase in microorganisms and plants. Phytochemistry, 1995. 39: p. 263-76.

162. **Buchanan, B.B., Gruissem, W. and Russell L. J.**, Protein sorting and vesicle traffic., in Biochemistry and molecular biology of plants. 2000, American Society of Plant Physiologists, : Rockville, MD. p. 160-201.

163. **Goss, R.J. and P.L. Newill,** A convenient enzymatic synthesis of L-halotryptophans. Chem Commun (Camb), 2006: p. 4924-5.

164. **Dhingra, A., V.A. James, H.U. Koop, M.C. Mok, R. De Paepe, M. Gallo, and K.M. Folta**, Tobacco, in Compendium of Transgenic Crop Plants. 2009, John Wiley & Sons, Ltd.

165. **Dawson, R.F.,** An experimental analysis of alkaloid production in *Nicotiana*: the origin of nornicotine. American Journal of Botany, 1945. 32: p. 416-423.

166. **Guthrie, F.E., Campbell, W.V., Baron, R.L.,** Feeding sites of the green peach aphid with respect to its adaptation to tobacco. Entomological Society of America, 1962. 55: p. 42-46.

167. **Shitan, N., M. Morita, and K. Yazaki,** Identification of a nicotine transporter in leaf vacuoles of *Nicotiana tabacum*. Plant Signal Behav, 2009. 4: p. 530-2.

168. **Steppuhn, A., K. Gase, B. Krock, R. Halitschke, and I.T. Baldwin,** Nicotine's defensive function in nature. PLoS Biol, 2004. 2: p. E217.

169. **Newton, W.A., Y. Morino, and E.E. Snell,** Properties of Crystalline Tryptophanase. J Biol Chem, 1965. 240: p. 1211-8.

170. **Lee, J.H., Y.G. Kim, M.H. Cho, T.K. Wood, and J. Lee,** Transcriptomic analysis for genetic mechanisms of the factors related to biofilm formation in *Escherichia coli* O157:H7. Curr Microbiol, 2011. 62: p. 1321-30.

171. **Hirakawa, H., Y. Inazumi, T. Masaki, T. Hirata, and A. Yamaguchi,** Indole induces the expression of multidrug exporter genes in *Escherichia coli*. Mol Microbiol, 2005. 55: p. 1113-26.

172. **Bansal, T., D. Englert, J. Lee, M. Hegde, T.K. Wood, and A. Jayaraman,** Differential effects of epinephrine, norepinephrine, and indole on *Escherichia coli* O157:H7 chemotaxis, colonization, and gene expression. Infect Immun, 2007. 75: p. 4597-607.

173. **Ehrenshaft, M. and M.E. Daub,** Isolation of PDX2, a second novel gene in the pyridoxine biosynthesis pathway of eukaryotes, archaebacteria, and a subset of eubacteria. Journal of Bacteriology, 2001. 183: p. 3383-3390.

174. **Lee, M. and R.S. Phillips,** The mechanism of *Escherichia coli* tryptophan indole-lyase: substituent effects on steady-state and pre-steady-state kinetic parameters for aryl-substituted tryptophan derivatives. Bioorg Med Chem, 1995. 3: p. 195-205.

175. **Blume, D.E., J.G. Jaworski, and J.W. McClure,** Uridinediphosphate-glucose: Isovitexin 7-O-glucosyltransferase from barley protoplasts: Subcellular localization. Planta, 1979. 146: p. 199-202.

176. **Orzechowski, S.,** Starch metabolism in leaves. Acta Biochimica Polonica, 2008. 55: p. 435-445.

177. **Wu, E., K. Komolpis, and H.Y. Wang,** Chemical extraction of indigo from *Indigofera tinctoria* while attaining biological integrity. Biotechnology Techniques, 1999. 13: p. 567-569.

178. **Wroblewski, T., A. Tomczak, and R. Michelmore,** Optimization of *Agrobacterium*-mediated transient assays of gene expression in lettuce, tomato and *Arabidopsis*. Plant Biotechnol J, 2005. 3: p. 259-73.

179. **Salcher, O. and F. Lingens,** Regulation of phospho-2-keto-3-deoxy-heptonate aldolase (DAHP synthase) and anthranilate synthase of *Pseudomonas aureofaciens*. J Gen Microbiol, 1980. 121: p. 473-6.

180. **Li, J. and R.L. Last,** The *Arabidopsis thaliana* trp5 mutant has a feedback-resistant anthranilate synthase and elevated soluble tryptophan. Plant Physiol, 1996. 110: p. 51-9.

181. **Hong, S.B., C.A. Peebles, J.V. Shanks, K.Y. San, and S.I. Gibson,** Expression of the *Arabidopsis* feedback-insensitive anthranilate synthase holoenzyme and tryptophan decarboxylase genes in *Catharanthus roseus* hairy roots. J Biotechnol, 2006. 122: p. 28-38.

182. **Hefner, T., J. Arend, H. Warzecha, K. Siems, and J. Stockigt,** Arbutin synthase, a novel member of the NRD1 beta glycosyltransferase family, is a unique multifunctional enzyme converting various natural products and xenobiotics. Bioorganic & Medicinal Chemistry, 2002. 10: p. 1731-1741.

183. **Ku, S.Y., P. Yip, and P.L. Howell,** Structure of *Escherichia coli* tryptophanase. Acta Crystallogr D Biol Crystallogr, 2006. 62: p. 814-23.

184. **Jones, P. and T. Vogt,** Glycosyltransferases in secondary plant metabolism: tranquilizers and stimulant controllers. Planta, 2001. 213: p. 164-74.

185. **Gachon, C.M., M. Langlois-Meurinne, and P. Saindrenan,** Plant secondary metabolism glycosyltransferases: the emerging functional analysis. Trends Plant Sci, 2005. 10: p. 542-9.

186. **Kawashima, H., T. Horii, T. Ogawa, and H. Ogawa,** Functional Domains of *Escherichia coli* Reca Protein Deduced from the Mutational Sites in the Gene. Molecular & General Genetics, 1984. 193: p. 288-292.

187. **Lazo, G.R., P.A. Stein, and R.A. Ludwig,** A DNA Transformation-Competent *Arabidopsis* Genomic Library in Agrobacterium. Bio-Technology, 1991. 9: p. 963-967.

188. **Peebles, C.A., G.W. Sander, E.H. Hughes, R. Peacock, J.V. Shanks, and K.Y. San,** The expression of 1-deoxy-D-xylulose synthase and geraniol-10-hydroxylase or anthranilate synthase increases terpenoid indole alkaloid accumulation in *Catharanthus roseus* hairy roots. Metab Eng, 2011. 13: p. 234-40.

189. **Wang, K. and S. Ohnuma,** Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution. Trends in Biochemical Sciences, 1999. 24: p. 445-451.

190. **Lange, B.M., M.R. Wildung, D. McCaskill, and R. Croteau,** A family of transketolases that directs isoprenoid biosynthesis via a mevalonate-independent pathway. Proc Natl Acad Sci U S A, 1998. 95: p. 2100-4.

191. **Lois, L.M., N. Campos, S.R. Putra, K. Danielsen, M. Rohmer, and A. Boronat,** Cloning and characterization of a gene from *Escherichia coli* encoding a transketolase-like enzyme that catalyzes the synthesis of D-1-deoxyxylulose 5-phosphate, a common precursor for isoprenoid, thiamin, and pyridoxol biosynthesis. Proc Natl Acad Sci U S A, 1998. 95: p. 2105-10.

192. **Estevez, J.M., A. Cantero, A. Reindl, S. Reichler, and P. Leon,** 1-deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. Journal of Biological Chemistry, 2001. 276: p. 22901-22909.

193. **Walter, M.H., J. Hans, and D. Strack,** Two distantly related genes encoding 1-deoxy-d-xylulose 5-phosphate synthases: differential regulation in shoots and apocarotenoid-accumulating mycorrhizal roots. Plant J, 2002. 31: p. 243-54.

194. **Kim, S.M., T. Kuzuyama, Y.J. Chang, K.S. Song, and S.U. Kim,** Identification of class 2 1-deoxy-D-xylulose 5-phosphate synthase and 1-deoxy-D-xylulose 5-phosphate reductoisomerase genes from *Ginkgo biloba* and their transcription in embryo culture with respect to ginkgolide biosynthesis. Planta Med, 2006. 72: p. 234-40.

195. **Burke, C.C., M.R. Wildung, and R. Croteau,** Geranyl diphosphate synthase: Cloning, expression, and characterization of this prenyltransferase as a heterodimer. Proceedings of the National Academy of Sciences of the United States of America, 1999. 96: p. 13062-13067.

196. **Schmidt, A., B. Wachtler, U. Temp, T. Krekling, A. Seguin, and J. Gershenzon,** A Bifunctional Geranyl and Geranylgeranyl Diphosphate Synthase Is Involved in Terpene Oleoresin Formation in *Picea abies*. Plant Physiology, 2010. 152: p. 639-655.

197. **Tholl, D., C.M. Kish, I. Orlova, D. Sherman, J. Gershenzon, E. Pichersky, and N. Dudareva,** Formation of monoterpenes in *Antirrhinum majus* and *Clarkia breweri* flowers involves heterodimeric geranyl diphosphate synthases. Plant Cell, 2004. 16: p. 977-992.

198. **Orlova, I., D.A. Nagegowda, C.M. Kish, M. Gutensohn, H. Maeda, M. Varbanova, E. Fridman, S. Yamaguchi, A. Hanada, Y. Kamiya, A. Krichevsky, V. Citovsky, E. Pichersky, and N. Dudareva,** The small subunit of snapdragon geranyl diphosphate synthase modifies the chain length specificity of tobacco geranylgeranyl diphosphate synthase in planta. Plant Cell, 2009. 21: p. 4002-17.

199. **Yamaguchi, S.,** Gibberellin metabolism and its regulation. Annu Rev Plant Biol, 2008. 59: p. 225-51.

200. **Kieber, J.J. and G.E. Schaller,** Cytokinins. Arabidopsis Book, 2014. 12: p. e0168.

201. **El-Showk, S., R. Ruonala, and Y. Helariutta,** Crossing paths: cytokinin signalling and crosstalk. Development, 2013. 140: p. 1373-83.

202. **Raghavendra, A.S., V.K. Gonugunta, A. Christmann, and E. Grill,** ABA perception and signalling. Trends Plant Sci, 2010. 15: p. 395-401.

203. **Gatz, C., A. Kaiser, and R. Wendenburg,** Regulation of a modified CaMV 35S promoter by the Tn10-encoded Tet repressor in transgenic tobacco. Mol Gen Genet, 1991. 227: p. 229-37.

204. **Aoyama, T. and N.H. Chua,** A glucocorticoid-mediated transcriptional induction system in transgenic plants. Plant J, 1997. 11: p. 605-12.

205. **Nara, Y., H. Kurata, M. Seki, and K. Taira,** Glucocorticoid-induced expression of a foreign gene by the GVG system in transformed tobacco BY-2 cells. Biochem Eng J, 2000. 6: p. 185-191.

206. **McNellis, T.W., M.B. Mudgett, K. Li, T. Aoyama, D. Horvath, N.H. Chua, and B.J. Staskawicz,** Glucocorticoid-inducible expression of a bacterial avirulence gene in transgenic *Arabidopsis* induces hypersensitive cell death. Plant J, 1998. 14: p. 247-57.

207. **Kang, H.G., Y. Fang, and K.B. Singh,** A glucocorticoid-inducible transcription system causes severe growth defects in *Arabidopsis* and induces defense-related genes. Plant J, 1999. 20: p. 127-33.

208. **Mattaj, I.W. and L. Englmeier,** Nucleocytoplasmic transport: The soluble phase. Annual Review of Biochemistry, 1998. 67: p. 265-306.

209. **Wang, R.W. and M.G. Brattain,** The maximal size of protein to diffuse through the nuclear pore is larger than 60 kDa. Febs Letters, 2007. 581: p. 3164-3170.

210. **Young, B., R. Wightman, R. Blanvillain, S.B. Purcel, and P. Gallois,** pH-sensitivity of YFP provides an intracellular indicator of programmed cell death. Plant Methods, 2010. 6: p. 27.

211. **Kimura, S., T. Noda, and T. Yoshimori,** Dissection of the autophagosome maturation process by a novel reporter protein, tandem fluorescent-tagged LC3. Autophagy, 2007. 3: p. 452-60.

212. **Yu, Q., C. Tang, and J. Kuo,** A critical review on methods to measure apoplastic pH in plants. Plant and Soil, 2000. 219: p. 29-40.

213. **Nelson, B.K., X. Cai, and A. Nebenfuhr,** A multicolored set of in vivo organelle markers for co-localization studies in *Arabidopsis* and other plants. Plant J, 2007. 51: p. 1126-36.

6. Appendix

The attached CD-ROM comprises all relevant raw data and DNA sequences.

7. Danksagung

An erster Stelle möchte ich mich bei Herrn Prof. Dr. Heribert Warzecha für die Möglichkeit bedanken dieses abwechslungsreiche und interessante Thema im Rahmen einer Promotion bearbeiten zu dürfen. Vielen Dank für deine Unterstützung, deine Geduld und Hilfe sowie die Möglichkeit an zwei *Short-Term Scientific Missions* (STSMs) teilzunehmen.

Des Weiteren möchte ich mich sehr herzlich bei Herrn Prof. Dr. Gerhard Thiel für die Erstellung des Zweitgutachtens bedanken.

Ein großes Dankeschön geht auch an Herrn Dr. Markus Krischke für seine Zeit, Unterstützung und hilfreichen Ratschläge bei der Durchführung und Auswertung zahlreicher LC-MS Messungen an der Julius-Maximilians-Universität Würzburg.

Auch möchte ich mich bei Herrn Dr. Heiko Rischer für die Möglichkeit bedanken Versuche im Rahmen einer STSM in seiner Arbeitsgruppe am *Technical Research Centre of Finland* durchzuführen. Dabei gilt mein ganz besonderer Dank Frau Dr. Tuulikki Seppänen-Laakso und Frau Dr. Heli Nygren, die mich bei allen LC-MS Messungen vor Ort unterstützt haben. Auch bedanke ich mich bei der gesamten Arbeitsgruppe, besonders bei Frau Tuuli Teikari, für ihre Hilfe. Des Weiteren danke ich sehr herzlich Herrn Dr. Diego Orzáez für die Möglichkeit die GoldenBraid Technik im Rahmen einer STSM in seinem Labor zu erlernen. Ganz besonders möchte ich mich in diesem Zusammenhang auch bei Dr. Alejandro Sarrión Perdigones für die Betreuung während meiner Klonierungsarbeit bedanken, sowie bei der gesamten Arbeitsgruppe für ihre Hilfe und die leckere Paella. Auch bedanke Ich mich bei der *European Cooperation in Science and Technology* (COST) für die finanzielle Unterstützung und die Möglichkeit diese beiden STSMs durchführen zu können.

Des Weiteren möchte ich mich bei Prof. Dr. Johan Memelink für die zur Verfügung gestellten Gene des Strictosidin Synthesewegs bedanken. Auch danke ich Prof. Sarah E. O'Connor für die Halogenasegene.

Mein ganz besonderer Dank geht an Agata Staniek, nicht nur für die Korrektur dieser Arbeit und zahlreicher weiterer Texte, sondern auch für die vielen hilfreichen Ratschläge und Ideen. Ebenso danke ich Sabine Knorr für die Unterstützung und Hilfe bei der Durchführung der Docking Simulationen. Ganz herzlich möchte ich mich auch bei meinen beiden Bachelor Studenten Bastian Wagner und Kim Röder, sowie bei allen Masterpraktikanten für ihre gute Arbeit bedanken.

Außerdem danke ich Birgit und Renate für ihre zuverlässige und gute Pflege und Betreuung der Tabakpflanzen.

Ganz herzlich möchte ich mich auch bei Prof. Dr. Joachim Stöckigt für die bereitgestellten Standardsubstanzen zur Analyse bedanken.

Ein großes Dankeschön geht auch an Lucia Carrillo, für ihre Geduld und Zeit am CLSM. Ganz besonders freue ich mich aber, dass ich die Promotionszeit mit ihr und María Mancheño Solano zusammen erleben durfte. Ohne euch wär es einfach nicht dasselbe gewesen. Auch bedanke ich mich bei Henning Pennekamp und Jascha Volk sowie der gesamten Arbeitsgruppe für die schöne Zeit, die entspannte Atmosphäre und die vielen gemeinsamen Abende und Mittagspausen mit leckeren Salaten.

Zuletzt möchte ich mich bei meinen Eltern bedanken, die mir stets helfend zur Seite standen und mich immer uneingeschränkt unterstützt haben. Mein ganz besonderer Dank gilt dir, Moritz, für deine liebevolle Unterstützung, deine unendliche Geduld und dein Verständnis in dieser spannenden und manchmal auch schwierigen Zeit.

8. Curriculum Vitae

Sabine Fräbel,
Kiesstraße 91, 64287 Darmstadt

Born 12.04.1987 in Darmstadt
German

Academic career

10/2011 - present **PhD thesis: "Characterization of flavin-dependent tryptophan halogenases and their application in plant metabolic engineering"**
Technische Universität Darmstadt, Germany
Plant Biotechnology and Metabolic Engineering

10/2006 – 09/2011 **Biology Diploma**
Technische Universität Darmstadt, Germany
Thesis: "Plant induced gene silencing in *Nosema ceranae*"

- acquired specialized knowledge of plant physiology, microbiology and ecology

01-06/2010 **Semester abroad,**
University of Skövde, Sweden

- master courses in biomedicine

1999 - 2006 **University-Entrance Diploma**
Albert-Einstein-Schule, Groß-Bieberau, Germany

9. 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 02.11.2015