Metabolic engineering of biosynthesis and sequestration of artemisinin

Bo Wang
Thesis committee

Promotor
Prof. Dr Harro J. Bouwmeester
Professor of Laboratory Plant Physiology
Wageningen University

Co-promotor
Prof. Dr.ir. AR (Sander) van der Krol
Associate Professor of Laboratory Plant Physiology
Wageningen University

Other members
Prof. Dr AHJ Bisseling, Wageningen University
Prof. Dr A Tissier, Leibniz Institute of Plant Biochemistry, German
Prof. Dr SC de Vries, Wageningen University
Prof. Dr HJ Bosch, Utrecht University / Plant Research International Wageningen

This research was conducted under the auspices of
the Graduate School of Experimental Plant Sciences
Metabolic engineering of biosynthesis and sequestration of artemisinin

Bo Wang

Thesis
submitted in fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 18 March 2015
at 11.00 a.m. in the Aula.
Bo Wang
Metabolic engineering of biosynthesis and sequestration of artemisinin

212 pages.

PhD thesis, Wageningen University, Wageningen, NL (2016)
With references, with summaries in Dutch and English

ISBN 978-94-6257-672-8
CONTENTS

Chapter 1  
General Introduction  

Chapter 2  
The metabolite chemotype of *Nicotiana benthamiana* transiently expressing artemisinin biosynthetic pathway genes is a function of CY-P71AV1 type and relative gene dosage  

Chapter 3  
The *A. annua* Lipid Transfer Protein AaLTP3 prevents reflux of dihydroartemisinic acid that was transported to the apoplast and increases production of artemisinin in *Nicotiana benthamiana*  

Chapter 4  
Interaction between ectopically induced flavonoid and sesquiterpenoid biosynthetic pathways in *Nicotiana benthamiana*  

Chapter 5  
Higher terpene production from ectopic expressed terpene synthase genomic clones than from terpene synthase cDNAs in transgenic *Arabidopsis*  

Chapter 6  
General Discussion  

Summary  
Samenvatting  
Acknowledgements  
Curriculum vitae  
Education certificate
Chapter 1

General Introduction
1 Plant secondary metabolites and terpenes

1.1 plant secondary metabolites

In 1803, the first bio-active plant natural product, morphine, was isolated in its crystalline pure form from the medicinal plant opium poppy by Friedrich Sertturner (1783-1841). Although medicinal plants had been used for many centuries by many different cultures, this marked the beginning of modern scientific phytochemistry (Goerig and Schulte am Esch, 1991). In the next two centuries, more and more bio-active plant products were isolated and their structure elucidated. At present over 200,000 different plant derived metabolites have been reported (Hartmann, 2007). This provided the chemical basis for the research on plant secondary metabolism and lead to the development of a natural products industry, which over the last 50 years has begun to exploit the rich diversity of plant secondary metabolites. On the human side, research addresses the understanding of the bioactivity of secondary metabolites, trying to identify their molecular action, improve their target specificity and target delivery. On the plant side, the research mainly addresses the question why and how plant cells synthesize such a rich diversity of secondary metabolites. Moreover, scientists study the possibility to use biotechnology, including genetic engineering, to introduce or enhance the production of secondary metabolites in crops, to improve, for example, nutritional value, environmental stress tolerance or disease resistance. An excellent example of such genetic engineering is the purple tomato, which has strongly enhanced levels of the health-promoting anthocyanin (Butelli et al., 2008). This was achieved through genetic modification by introducing the Antirrhinum majus transcription factors Delila and Rosea1 into tomato. As many plant secondary metabolites are toxic to the plant itself, research into the use of biotechnology to increase or introduce secondary metabolite formation in plants should also involve the question how plants can escape or tolerate these potentially toxic compounds.

Secondary metabolites may be classified into different families based on their biosynthetic pathways. The phenolics, which are products of the phenylpropanoid biosynthesis pathway; the terpenoids and steroids, which are products of the terpenoid or isoprenoid biosynthesis pathways; and the alkaloids, which are products of the alkaloid biosynthesis pathway (Bourgaud et al., 2001). Some classes of secondary metabolites are specific to certain plant species. For instance, glucosinolates occur as secondary metabolites in almost all plant species of the order Brassicales (e.g. families Crucifer-
ae, Capparidaceae, and Caricaceae) and in the genus Drypetes (family Euphorbiaceae) (Abdel-Farid et al., 2009; Rodman et al., 1996). While each of these biosynthesis pathways alone can already produce a myriad of complex molecules, the biochemical diversity of plant-derived compounds is further extended by the production of combined products from different pathways. An important example of such hybrid product formation is indole-alkaloid biosynthesis in Catharanthus roseus. Indole-alkaloids are composed of a monoterpenoid and an alkaloid moiety. Examples are vincristine and vinblastine, which are important anti-cancer drugs, making them interesting targets for metabolic engineering. Indole-alkaloid biosynthesis involves many enzymatic steps, catalyzed by enzymes that belong to a wide range of protein families (El-Sayed and Verpoorte, 2007). Metabolic engineering approaches for this type of compounds form a good illustration of the importance of sequestration and transport. When the first enzyme of the C. roseus monoterpenoid indole alkaloid biosynthetic pathway, geraniol synthase, was heterologously expressed in tobacco this resulted in the formation of multiple geraniol derivatives, including further oxidized and double bond reduced geraniol derivatives and their glycosides (Dong et al., 2013). When subsequently a large part of the terpenoid branch of the pathway was reconstructed through transient expression in N. benthamiana only a very low flux to the end product was observed (Miettinen et al., 2014). The authors showed that competing side reactions occurred in N. benthamiana draining away carbon from the main pathway. Suitable compartmentation and improved transport and sequestration could therefore potentially solve these issues.

1.2 Terpenoids

In my thesis, I will focus on the terpenoids, one of the most important families of plant derived natural secondary metabolites. The terpenoids have a large diversity of functions in plants, for example as plant hormones (gibberellins, ABA, strigolactones and cytokinins), to attract pollinators or in defense against insects (Abel et al., 2009; Chen et al., 2003; Paré and Tumlinson, 1999). They are extensively used as flavors, fragrances, and drugs for humans (Ikram et al., 2015; Rohloff, 2004; Russo, 2011; Singh and Sharma, 2015; Tholl and Lee, 2011). Terpenes are formed by the condensation of 5-carbon isoprene units, the ubiquitous precursors isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). Depending on the number of isoprene building blocks, terpenes may be classified into mono-, sesqui-, di-, sester-, tri-, tetra- and polyterpenes (Table 1). The precursors of terpenes are synthesized from two separate pathways: the
mevalonic acid (MVA) pathway, localized in the cytosol, endoplasmic reticulum and peroxisomes and the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway, which is localized in the plastids (Figure 1) (Vranová et al., 2013). Terpene biosynthesis in the mitochondria may be supported by import of IPP from the cytosolic MVA pathway (Figure 1) (Rodríguez-Concepción and Boronat, 2002; Tholl and Lee, 2011; Vranová et al., 2013).
Figure 1. The biosynthesis pathway of MVA and MEP pathway. Acetoacetyl-CoA thiolase (AACT), HMG-CoA synthase (HMGS), HMG-CoA reductase (HMGR), mevalonate kinase (MVK), phosphomevalonate kinase (PMK), and mevalonate diphosphate decarboxylase (MVD), isopentenyl diphosphate isomerase (IDI), farnesyl diphosphate synthase (FPPS), geranylgeranyl diphosphate synthase (GGPPS), squalene synthase (SQS), deoxy-d-xylo-lulose-5-phosphate synthase (DXS), 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR), 2-C-methyl-d-erythritol 4-phosphate cytidylyltransferase (MCT), 4-(cytidine 5’-diphospho)-2-C-methyl-d-erythritol kinase (CMK), 2-C-methyl-d-erythritol 2,4-cyclodiphosphate synthase (MECPS), (E)-4-hydroxy-3-methylbut-2-enyl diphosphatesynthase (HDS) and isopentenyl diphosphate synthase (IDS), Geranyl diphosphate synthase (GPPS), terpene synthases (TPS).

Table 1. Classification of the terpenes

<table>
<thead>
<tr>
<th>Terpenes</th>
<th>Isoprene units</th>
<th>Carbon atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monoterpene</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Sesquiterpene</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>Diterpene</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>Sesterterpenes</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>Triterpenes</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>Tetraterpenes</td>
<td>8</td>
<td>40</td>
</tr>
<tr>
<td>Polyterpenes</td>
<td>&gt; 40</td>
<td>&gt; 40</td>
</tr>
</tbody>
</table>

The MEP pathway starts with the condensation of pyruvate and d-glycer-aldehyde 3-phosphate (G3P) to 1-deoxy-xylulose-5-phosphate (DXP) by DXP synthase (DXS) (Rodríguez-Concepción, 2006). Then, the formation of the C5 unit MEP is catalyzed by DXP reducto-isomerase (DXR) (Rodríguez-Concepción and Boronat, 2002). The downstream steps from MEP to IPP are catalyzed by 2-C-methyl-d-erythritol 4-phosphate cytidylyltransferase (MCT), 4-(cytidine 5’-diphospho)-2-C-methyl-d-erythritol kinase (CMK), 2-C-methyl-d-erythritol 2,4-cyclodiphosphate synthase (MECPS), (E)-4-hydroxy-3-methylbut-2-enyl diphosphatesynthase (HDS) and isopentenyl diphosphate synthase (IDS). Cytosolic and plastidal IPP can be exchanged, and in both compartments can be isomerized into DMAPP by isopentenyl diphosphate isomerase (IDI), providing an intrinsic link between the plastidal MEP pathway and the cytosolic MVA pathway (Vickers et al., 2014; Vranová et al., 2013). Geranyl diphosphate synthase (GPPS) catalyzes the condensation of one molecule each of IPP and DMAPP to produce GPP (C10), the universal precursor of the monoterpenes (Hemmerlin et al., 2012; Tholl and Lee, 2011; Vickers et al., 2014; Vranová et al., 2013). In addition, DMAPP and three molecules of IPP is catalyzed by Geranylgeranyl diphosphate synthase (GGPPS) into GGPP (C20) in plastids, cytosol, endoplasmic reticulum (ER) and mitochondria (Okada et al., 2000).
Chapter 1

The first step in the MVA pathway is the condensation of two molecules of acetyl-CoA in the cytosol to acetoacetyl-CoA by the acetoacetyl-CoA thiolase (AACT) (Lombard and Moreira, 2011). HMG-CoA synthase (HMGS) and HMG-CoA reductase (HMGR) subsequently catalyze the formation of MVA (C6) (Gardner and Hampton, 1999) (Figure 1). Mevalonate is further converted to IPP by mevalonate kinase (MVK), phosphomevalonate kinase (PMK), and mevalonate diphosphate decarboxylase (MVD), respectively. IPP is transformed into DMAPP by IDI, a divalent metal ion-requiring enzyme in the cytosol (Dorsey and Porter, 1968; Vranová et al., 2013). Condensation of one DMAPP and two IPP molecules catalyzed by farnesyl diphosphate synthase (FPPS) leads to the formation of FPP (C15), the natural precursor of sesquiterpenes (Cunillera et al., 1996).

Following the formation of the prenyl diphosphate precursors GPP and FPP, the monoterpenes and sesquiterpenes skeleton is generated by the action of terpene synthases (TPS). The primary products of these terpene synthases are often volatile. Subsequent enzymatic steps catalyzed by cytochrome P450s, reductases, dehydrogenases, acyl transferases, glycosyl transferases, etc then create the enormous structural diversity in the mono- and sesquiterpenoids which may be less volatile and with a wide range of hydrophilic or hydrophobic. Diterpenes are formed from GGPP. For the biosynthesis of triterpenes, FPP is converted into 2,3-oxidosqualene by the squalene synthase (SQS) and squalene monooxygenase or epoxidase (SQE). Then 2,3-oxidosqualene is cyclized by oxidosqualene cyclase (OSC) to form triterpenes (Abe et al., 1993). Also in the di- and triterpenes an enormous structural diversity is subsequently created through the action of additional enzymes as described above for the mono- and sesquiterpenes.

**Function of induced terpene biosynthesis**

Terpenes are often produced and stored in specialized glands or trichomes below the cuticular wax layer of the cells, so that also potential terpenes do not evaporate easily. However, when a plant is damaged by herbivorous insects, the leaves release stored volatiles and newly produced volatiles from the induction of terpene synthases (Paré and Tumlinson, 1999). This blend of volatile terpenes functions in communication between and among species, and as messengers between members of the same species (Paré and Tumlinson, 1999; Witzany, 2006). They may be toxic or repellent to the attacking insect, or they attract parasitic and predatory enemies of the insects (Law and Regnier, 1971; Paré and Tumlinson, 1999). Also flowers release
volatiles. These are important for plant reproduction as they attract pollinators (Chen et al., 2003; Dudareva et al., 2013; Dudareva et al., 2006; Huang et al., 2012). (E)-β-caryophyllene, for example, of which the biosynthesis is catalyzed by caryophyllene synthase (CST), AtTPS21, that is specifically expressed in the flowers, is the major volatile organic compound emitted from Arabidopsis thaliana flowers (Chen et al., 2003; Huang et al., 2012). In addition to its role in attracting pollinators it was shown to be an effective defense compound against a bacterial pathogen (Abel et al., 2009).

1.3 Metabolic engineering of secondary metabolism

With the identification of the biosynthetic pathway genes of secondary metabolites, metabolic engineering to increase their production comes within reach. However, initial attempts showed that production of engineered products was low (Farhi et al., 2011; Lange et al., 2011; Wallaart et al., 2001). Therefore strategies should be investigated to improve this production, for example by increasing the precursor pool, eliminating competing pathways, targeting the enzymes to a different subcellular compartment with higher precursor pool and intron-enhanced gene transcription.

Increasing the precursor pool

Increasing the precursor pool could provide more substrate for the biosynthesis genes. So far, most efforts have focused on ectopic expression of genes encoding the early, rate-limiting, steps of the pathway. HMGR is a key rate-limiting step in the MVA pathway (Goldstein and Brown, 1990). Overexpression of a truncated HMGR resulted in increased production of endogenous sterols (Chappell et al., 1995) as well as the heterologous sesquiterpenoid, patchoulol, in tobacco (Wu et al., 2006; Zhan et al., 2014). Overexpression of DXS, the key rate-determining enzyme in the MEP pathway (Rodríguez-Concepción, 2006), in Arabidopsis, increased accumulation of many plastidic isoprenoids including chlorophyll, α-tocopherol, carotenoids, and abscisic acid (Estévez et al., 2001; Wright et al., 2014).

Eliminating competing pathways

Endogenous enzymes in plants cells may modify the pathway intermediates of end products that are produced by ectopic expression of a pathway in a heterologous host, such conversions may decrease the production of the
target compound (Staniek et al., 2013). Competing modification enzymes may be reductases, dehydrogenases and oxidases, as well as conjugating enzymes such as glycosyl transferases and glutathione S-transferases (Dudareva et al., 2004; Renault et al., 2014; Shi and Xie, 2014; Staniek et al., 2013; Strasser et al., 2014). These modifications and conjugations formation has been reported for many metabolically engineered compounds, such as geraniol (formation of glycosides), costunolide (formation of glutathione conjugates) and artemisinic acid (formation of glycosides) (Dong et al., 2013; Liu et al., 2011b; Ting et al., 2013). A successful example of blocking the competing enzymes was carried out in *E. coli*. A dehydrogenases, *YjgB*, was knocked out to increase heterologous geraniol production by preventing geraniol dehydrogenation. This resulted in increased purity and production of geraniol by more than 20% (Zhou et al., 2014).

**Targeting expression**

Plant cells contain multiple membrane-delimited compartments (Small et al., 1998). Enzymes and products of secondary metabolism may be in different compartments of the cell. For example, flavonoid metabolism occurs on the cytoplasmic side of the endoplasmic reticulum but the products of the flavonoid biosynthesis pathway are transported to the vacuole by transporters (Zhao and Dixon, 2010). In addition, enzymes in one metabolic biosynthesis pathway may be localized in different compartments. Moreover, alkaloids are produced within multiple cell types and precursors in an alkaloid biosynthesis pathway may be translocated intracellularly - between cytosol, vacuole, ER, vesicles, and thylakoid membrane - as well as intercellularly (Deus-Neumann and Zenk, 1986; McCaskill et al., 1988; Qu et al., 2015; Roytrakul and Verpoorte, 2007; St-Pierre et al., 1999).

Enzymes may show targeting to different compartments in different plant species. In terpene biosynthesis, for example, geraniol synthases have been reported with both plastidic and cytosolic localization (Dong et al., 2013). Using metabolic engineering, a default subcellular compartmentation may be artificially altered or circumvented. For example, artificially plastid targeted geraniol synthase produced 30% more than upon cytosolic targeting (Dong et al., 2013). Furthermore, in an exciting metabolic engineering approach, (Kumar et al., 2012) remodeled the cytosolic MVA pathway into the tobacco chloroplast, which increased the synthesis of isoprenoid metabolites including mevalonate, carotenoids, sterols, and squalene.
**Introns-enhanced gene transcription**

Introns have a dramatic effect on gene expression in plants because they may contain enhancer elements or alternative promoters, or may increase mRNA accumulation by intron-mediated enhancement (IME) (Rose, 2008). For example, a heterologous intron increased expression in *Arabidopsis* relative to the intronless control, even for tissue specific genes, which indicates that different genes respond similarly to the same stimulatory intron (Emami et al., 2013). This approach has not fully been exploited yet to increase the production of secondary metabolites in metabolic engineering approaches but the results suggest that it could be worthwhile to study the possible effect of introns in enhancing of secondary metabolites biosynthesis upon heterologous expression of pathway genes in plant cells.

**1.4 Transport in secondary metabolism**

The formation of side products in metabolic engineering, as described above, suggests that transport and sequestration may be important issues to solve if we want to be able to exploit metabolic engineering to the full. Many secondary metabolites are produced in certain cell types only, such as trichomes in tobacco or laticifers in chicory and sequestered potential toxic secondary metabolites outside the cells in subcuticular space or in laticifers. Biotechnological approaches should therefore at least consider this, as sequestration outside of the cell could potentially improve the yield of the engineered natural product. Several membrane transport events may be involved in secondary metabolite transport and/or sequestration, such as glycosylation formation and then transport to the vacuole (1), transport into the vacuole directly (2), transport the secondary molecules into ER for the further modification (3), secretion of metabolites outside the cell by vesicles or non-vesicular trafficking (4), export from the cytosol to the apoplast through the plasma membrane - a process often carried out by ABC transporters (5), and export of plastid-produced secondary metabolites to ER or cytosol (6) (Figure 2). In addition, we considered a potential role of LTPs which reside in the cell wall (7).

**1.4.1 ATP-binding cassette (ABC) transporters**

ABC proteins are large, ubiquitous, and diverse membrane transporters that carry their substrates over biological membranes by an ATP-dependent reaction. Plant ABC transporters can be localized in most membranes, such as the vacuolar membrane, the plasma membrane, and the membranes of plastids,
peroxisomes, mitochondria and endoplasmic reticulum. In plants, ABC transporters contribute to multiple functions, ranging from detoxification and a role in resistance against pathogens and abiotic stress, to a role in plant development.

The core structure of ABC-transporters consists of two different types of domains, a nucleotide-binding domain (NBD, or ATP-binding cassettes) and a transmembrane domain (TMD). The NBD has a highly conserved sequence. Two NBD monomers form a closed dimer which can bind two ATP molecules at the dimer interface in the cytoplasm. This will result in the closed dimer to open up and ATP hydrolysis and release of ADP and Pi (Higgins and Linton, 2004). This sustained cyclic opening and closing of the NBD dimer during ATP binding and release of ADP and Pi provides energy for transporting a substrate across the membrane (Doshi et al., 2010). The TMD is embedded in the membrane bilayer to connect the cytosolic and apoplastic side. The substrate binding site of the TMD recognizes the substrate. Then two TMD modules form a solvent-accessible space to transport the substrate across the membrane.
The ABCG subfamily contains both half-size proteins, called White–Brown Complex (WBC) transporters, and full-size proteins coined Pleiotropic Drug Resistance (PDR) transporters. WBC transporters are involved in transport of eye pigment precursors in *Drosophila melanogaster* (Mackenzie et al., 2000). NtWBC1 was the first reported transporter of this subfamily in plants. It was shown to play an important role in plant reproduction probably because it is involved in exudate production in the stigma and anther dehiscence (Otsu et al., 2004). *Arabidopsis* ABCG11/WBC11 interacts with ABCG12/CER5 and together these transporters form a key component of the export pathway for cutin and wax secretion through the plasma membrane of Arabidopsis epidermis cells to the cell wall and cuticle (Bird et al., 2007; Panikashvili et al., 2007).

PDRs play a role in the export of metabolites before their concentration reaches cellular toxic levels. Goossens et al reported that PDR genes can be used to stimulate the secretion of secondary metabolites in plant cell suspension cultures (Goossens et al., 2003) and when *Saccharomyces cerevisiae* PDR5 was expressed in tobacco BY2 (*Nicotiana tabacum* L. cv Bright-Yellow 2) cells, this resulted in improved secretion of alkaloids. NpPDR1 transports the antifungal diterpene, sclareol, onto the leaf surface of *Nicotiana plumbaginifolia*, and NpPDR1 knock out lines displayed increased susceptibility to a group of fungal and oomycete pathogens (Bultreys et al., 2009; Stukkens et al., 2005). Treatment of suspension-cultured tobacco BY-2 cells with the bacterial elicitor, flagellin, induced *NtPDR1* expression (Sasabe et al., 2002). The plasma membrane localized NtPDR1 transports diterpenes involved in the plant defense response in *Nicotiana tabacum* (Crouzet et al., 2013). Another PDR transporter, from *C. roseus*, CrTPT2, is expressed predominantly in the epidermis of young leaves and the level of catharanthine–vindoline drug dimers was increased when it was silenced (Yu and De Luca, 2013). The suppression of *CrTPT2* expression blocks the efflux of catharanthine, resulting in a decline in the leaf-surface catharanthine levels while increasing catharanthine levels within the leaf. Despite this scattered evidence for the role of ABC transporters in secondary metabolite transport, still little is known about the (possible) involvement of PDR transporters in secondary metabolite formation, and especially in the transport of mono- and sesquiterpenes.

### 1.4.2 Lipid transfer proteins (LTPs)

LTPs are abundant, small, basic proteins that are able to bind and exchange hydrophobic compounds between membranes (Yeats and Rose, 2008). This transfer process is distinct from the secretory vesicles of the ER (Holthuis...
and Menon, 2014; Lev, 2010). The size of LTPs ranges from 6.5 ~ 10.5 kD (Liu et al., 2015) and an eight-cysteine residue conserved motif forms the backbone of the LTPs (Liu et al., 2015; Yeats and Rose, 2008). The LTP protein forms a large internal hydrophobic cavity, containing lipid-binding sites (Liu et al., 2015; Wang et al., 2012). Four disulfide bonds between the cysteine residues stabilize this hydrophobic cavity structure (Liu et al., 2015; Samuel et al., 2002). In mammals and yeast, LTPs play an important role in non-vesicular trafficking of cholesterol, phospholipids, and sphingolipids between biological membranes (Lev, 2010; Prinz, 2007). LTPs regulate diverse lipid-mediated cellular processes in ER-to-Golgi and Golgi-to-plasma membrane trafficking (De Matteis et al., 2007; Lev, 2010; Prinz, 2007). The first report about LTPs playing an essential role in non-vesicular trafficking was in yeast, in which Sec14 loss of function produced defects in membrane trafficking from the Golgi complex to the plasma membrane (Bankaitis et al., 1990; Bankaitis et al., 1989). Sec14 was shown to be involved in controlling the production of diacylglycerol and phosphatidylinositol-4-P that play a key role in the formation of plasma-membrane-destined transport carriers at the Golgi complex (Bankaitis et al., 2012; Mousley et al., 2007). In mammalian cells, the lipid transfer protein CERT (CERamide Transfer protein) specifically binds ceramide from the ER membrane and delivers it to the Golgi membrane (Hanada et al., 2003; Kawano et al., 2006; Simanshu et al., 2013). In addition to the activity at the surface of the Golgi complex, the trafficking from the outer to the inner Golgi membrane by LTPs was reported in mammals and yeast (Gatta et al., 2015; Georgiev et al., 2011; Mesmin et al., 2011). Deleting all OSBP (sterol specific LTPs, OxySterol Binding Protein) homologs in yeast reduced the sterol trafficking from the plasma membrane to the ER (Georgiev et al., 2011). Apart from OSBPs, STeroidogenic Acute Regulatory Transfer (START) proteins are also sterol specific LTPs, and they can transfer sterols from the plasma membrane to the ER in both mammals and yeast (Gatta et al., 2015; Mesmin et al., 2011).

In plants, not much is known about the roles of LTPs, with a few exceptions. Almost all plant LTPs carry an N-terminal signal peptide directing the protein to the apoplast (Edstam et al., 2011). Arabidopsis DIR1 is an example of such a putative apoplastic LTP. It seems to interact with lipophilic molecules to promote long distance pathogen defense signaling during systemic acquired resistance and a plasma membrane receptor seems to be required to activate the resistance response (Lascombe et al., 2006; Maldonado et al., 2002). Blein et al. (Blein et al., 2002) showed that wheat LTP1 binds with elicitin receptors in the plasma membrane to induce plant cell responses (agonist behavior). However, whether this LTP1 forms a com-
plex with lipids to interact with the elicitin receptor is still unknown. MtN5 belongs to a group of LTPs having a role in plant-microbe interaction (Pii et al., 2010). Overexpression of MtN5 in Medicago truncatula led to an increased number of nodules when induced by Sinorhizobium meliloti (Pii et al., 2010). Silencing of OsC6, one of the LTPs in rice, resulted in defective development of lipidic orbicules (i.e. Ubisch bodies) and pollen exine during anther development and reduced pollen fertility (Zhang et al., 2010). Although OsC6 displays lipid binding activity in vitro, there is no clue of which specific lipid molecules it could bind to for regulating anther development.

In addition to the function to regulate cell signalling, plant LTPs are also active in lipid secretion. Tobacco NtLTP1, a glandular-trichome specific LTP, is required for diterpene secretion from glandular trichomes and transgenic tobacco plants overexpressing NtLTP1 displayed increased resistance against aphids (Choi et al., 2012). Interestingly, LTP genes show very high transcriptional activity in glandular trichomes, which produce terpenes, in many plant species (peppermint (Lange et al., 2000), alfalfa (Aziz et al., 2005), Artemisia annua (Berte et al., 2005), hops (Wang et al., 2008), Salvia fruticosa (Chatzopoulou et al., 2010), tomato (Schilmiller et al., 2010) and tobacco (Harada et al., 2010). This suggests that LTPs may play a role in terpene transport and sequestration.

1.4.3 Glutathione S-transferases (GSTs)

GSTs are abundant soluble proteins encoded by a highly divergent, ancient gene family, which can be divided on the basis of sequence identity into the phi, tau, theta, zeta and lambda classes (Edwards and Dixon, 2005; Sheehan et al., 2001). Plant GSTs conjugate electrophilic xenobiotics with glutathione (GSH) for vacuolar sequestration and catalyze the reduction of toxic organic hydroperoxides (Edwards et al., 2000). The detoxification process of electrophilic compounds is mediated by three groups of enzymes (Ishikawa, 1992; Marrs, 1996; Sandermann Jr, 1992). Substrate specific enzymes such as cytochrome P450 monooxygenases introduce functional groups to substrates in the transformation phase; enzymes such as UDP-glucosyltransferases and GSTs use this functional group as a site of conjugation resulting in a less toxic and more water-soluble conjugate in the conjugation stage; transporter enzymes, often ATP-dependent membrane pumps such as the ABC transporters described above, recognize and transfer conjugates across membranes for excretion or sequestration in the compartmentation stage.

GSTs might also have non-catalytic roles and serve as carrier proteins for transport of phytochemicals between cellular compartments (Edwards et al.,
Since cytosolic accumulation of anthocyanin is toxic to cells and results in feed-back inhibition (Licciardello et al., 2014), GSH conjugation is required for anthocyanin transport and sequestration into the vacuole (Licciardello et al., 2014; Mueller et al., 2000). *Bronze-2 (Bz2)* is a GST involved in vacuolar transfer of flavonoids in maize (Marrs et al., 1995). Bz2 conjugates GSH to cyanidin 3-glucoside and transports the resulting product into the vacuole via a tonoplast Mg-ATP–requiring glutathione pump (GS-X pump) (Alfenito et al., 1998). In *Arabidopsis*, the GST, Transparent testa19 (TT19), functions as a carrier to transport cyanidin and/or anthocyanins across the tonoplast (Sun et al., 2012). Multidrug Resistance-associated Protein (MRP)-type ABC transporters can transport glutathione conjugates into the vacuole in maize (Goodman et al., 2004; Klein et al., 2006). ABCC1 from grape berry, which shares a high amino acid sequence identity with Zm-MRP3 (Francisco et al., 2013; Goodman et al., 2004), acts as an anthocyanin transporter that depends on GSH without the formation of an anthocyanin-GSH conjugate (Francisco et al., 2013). Except MRP/ABCC transporters, multidrug and toxic extrusion (MATE) transporter, Transparent testa12 (TT12), was shown to be involved in flavonoid sequestration in the vacuole of cells in the seed coat endothelium (Debeaujon et al., 2001). In plants, GSTs bind indole-3-acetic acid (IAA) as carrier proteins while only a small amount of IAA-GSH conjugates was detected (LeClere et al., 2002; Pěnčík et al., 2013), which allow temporary storage or IAA uptake from membranes and trafficking to receptors (Marrs, 1996).

## 2 Artemisinin and its biosynthesis pathway

Malaria is the most devastating infectious disease in the tropics (Talisuna et al., 2004). Current estimates indicate that up to three million people die of malaria each year in Africa (Breman et al., 2004; Ntoumi F, 2004). The medicinal plant *Artemisia annua* (commonly called Sweet wormwood, Qinghao) has been used for treatment of malaria symptoms for over 2000 years in China. The medical use of *A. annua* as herbal tea to treat malaria was recorded in “Compendium of Materia Medica” by Li Shenzhen in 1596. The main bio-active compound in *A. annua* is artemisinin, a sesquiterpene lactone containing an endoperoxide bridge. Artemisinin was extracted and for the first time purified from leaves of *A. annua* by Project 523 between 1969 and 1972 under the supervision of Youyou Tu, and was named *Qinghaosu* (Chinese: 青蒿素) (Faurant, 2011). Artemisinin is widely used for the treatment of malaria, and has saved millions of lives. Therefore, this year the Nobel Prize in Physiology or Medicine was awarded to Youyou Tu. The World
health Organization (WHO) has recommended artemisinin-based combined therapy (ACT) for the treatment of malaria as the first choice from 2001. However, the production of artemisinin in *A. annua* varies from 0.01% to 0.8% DW (Liu et al., 2011a) and even the high artemisinin yielding hybrid selected by Mediplant in Switzerland only produces up to 1.95% DW of artemisinin (Simonnet et al., 2008). Because of these low contents artemisinin supply by isolation from *A. annua* does not satisfy the global demand (Durante et al., 2011) and several initiatives to solve this using biotechnology in both plants as well as micro-organisms were developed as I will explain below.

### 2.1 Biosynthesis of artemisinin

Artemisinin is produced from the precursor FPP, which is generated in the MVA pathway (Figure 3). The first committed step of artemisinin biosynthesis is the cyclization of FPP to amorpha-4,11-diene (AD) by amorphaadiene synthase (ADS) (Bouwmeester et al., 1999). AD is subsequently oxidized to artemisinic alcohol (AAOH), artemisinic aldehyde (AAA) and artemisinic acid (AA) by a cytochrome P450 enzyme, CYP71AV1, which is anchored to the endoplasmic reticulum (ER) (Covello et al., 2007; Teoh et al., 2006). In *A. annua*, AA is non-enzymatically converted to arteannuin B (AB) (Figure 3) (Brown and Sy, 2007). However, the exocyclic double bond of AAA can also be reduced by the soluble cytosolic protein DBR2, resulting in dihydroartemisinic aldehyde (DHAAA) (Zhang et al., 2008). DHAAA is oxidized either by CYP71AV1 or an alcohol dehydrogenase (ALDH1) to dihydroartemisinic acid (DHAA), which is the direct precursor of artemisinin (Teoh et al., 2009). It is assumed that the oxidation from DHAA to artemisinin occurs non-enzymatically (Figure 3) (Brown, 2010).

#### Chemotypes of *A. annua*

*A. annua* has two different chemotypes according to the relative level of the end products in the branched biosynthesis pathway: arteannuin B and artemisinin (Figure 3). The high artemisinin production (HAP) chemotypes also have a high DHAA level and a relatively low AA level, while the low artemisinin production (LAP) chemotypes have a relatively low level of DHAA compared with AA (Wallaart et al., 2000). On the other hand, these two chemotypes also show a differential response to elicitation by methyl jasmonate (MeJA) (Wu et al., 2011). As a result of exogenous application of MeJA, HAP plants display increased accumulation of DHAA and artemisinin while only a trace amount of AA can be detected. For the HAP chemotype there is a cor-
responding activation of most artemisinin pathway genes by MeJA except for the cytochrome P450 reductase (CPR) (Wu et al., 2011). In contrast, AA and arteisinin decreased dramatically under MeJA elicitation for the LAP plants in favor of AA and arteannuvin B production (Wu et al., 2011), which matches with the fact that the expression of the biosynthesis genes ADS, CYP71AV1 and DBR2 (but not ALDH1) were downregulated at various time points after JA treatment. The accumulation of AA in LAP plants and DHAA in HAP plants suggests that the conversion of these precursors represents a rate-limiting step in the biosynthesis of both arteannuvin B and artemisinin, respectively.
2.2 Metabolic engineering of artemisinin

2.2.1 Metabolic engineering in A. annua

Because of its highly valued product artemisinin, many attempts have been made to increase artemisinin production in A. annua. Efforts have been aimed at selecting HAP chemotypes (Wallaart et al., 2000; Wu et al., 2011), higher numbers of glandular trichomes (Lommen et al., 2006; Maes et al., 2011) and stimulation of the pathway genes by (Me)JA (Maes et al., 2011; Wu et al., 2011). Also, a number of genes involved in artemisinin biosynthesis have been overexpressed in A. annua to enhance artemisinin production (Tang et al., 2014). Indeed, artemisinin production was increased by overexpression of the genes encoding ADS and CYP71AV1 (Alam and Abdin, 2011; Ma et al., 2009; Shen et al., 2012). Alternatively, it was studied whether production may be increased by blocking competing reactions. By blocking the FPP consuming enzyme squalene synthase (SQS), for example, artemisinin production was increased 3.14-fold compared with the wildtype (Zhang et al., 2009). Since DHAAA reductase 1 (RED1) competes with ALDH1 for DHAAAs (Figure 3), it was suggested that knocking-down RED1 may lead to increased production of artemisinin in A. annua (Rydén et al., 2010).

The possibility to increase the expression of all artemisinin biosynthetic genes at once was explored by identifying transcription factors involved in the coordinated expression of the pathway genes. The transcription factor AaWRKY1 was reported to regulate artemisinin production (De Muynck et al., 2009) and overexpression of AaWRKY1 in A. annua resulted in an artemisinin content in transgenic lines of 24.5 mg/g (DW), which is 4.4-fold that of the control lines (De Muynck et al., 2009; Tang et al., 2014). The jasmonate-responsive AP2/ERF-type transcription factors AaERF1 and AaERF2 also positively elevate the transcript levels of both ADS and CYP71AV1, and both artemisinin and AA content were reduced in transgenic lines of A. annua in which AaERF1 and AaERF2 expression were suppressed through RNAi (Yu et al., 2012). Up-regulation or down-regulation of the expression of AaORA, a trichome-specific AP2/ERF transcription factor, in A. annua resulted in a significant increase or decrease, respectively, in artemisinin and DHAA (Lu et al., 2013). AaORA did not only positively regulate the expression of ADS, CYP71AV1 and DBR2 but also stimulated the expression of AaERF1 in A. annua.

Because it had been shown that JA can increase the expression of the artemisinin biosynthesis pathway genes attempts were made to improve artemisinin production in A. annua by constitutive expression of allene
oxide cyclase (AOC), the key enzyme in the biosynthetic pathway of jasmonates. Expression of the *A. annua* AOC resulted in an almost 2-fold increase (87.9 %) in artemisinin in transgenic lines (Tang et al., 2014).

2.2.2 Metabolic engineering in heterologous plants

Although all artemisinin biosynthetic pathway genes have been identified and cloned, there is no or only very low levels of artemisinin detected when these biosynthetic genes are overexpressed in heterologous plant hosts, either using transient expression (Ting et al., 2013; Zhang et al., 2011) or in stably transformed tobacco plants (Farhi et al., 2011). Heterologous expression of *ADS*, *CYP71AV1*, *DBR2* and *ALDH1* in tobacco resulted in the accumulation of amorpha diene and AAOH and DHAAOH, but not AA and DHAA (Zhang et al., 2011). Transient expression of a multi-gene construct containing *ADS*, *FPPS*, *t-HMGR* and *CYP71AV1* - with ribosomal skipping sequences and mitochondrial targeting of *ADS* - in *N. benthamiana* resulted in the formation of AA which was efficiently conjugated by endogenous glycosyl transferases (Van Herpen et al., 2010). The transient expression of *CYP71AV1* (that was isolated from a LAP chemotype) together with the other artemisinin biosynthetic genes in *N. benthamiana* resulted in high production of glycoside-conjugated artemisinin precursors (Ting et al., 2013). Transgenic tobacco, expressing five plant- and yeast-derived genes involved in the MVA and artemisinin pathways in a single vector produced 0.48 - 6.8 mg/g dry weight artemisinin, which is comparable to the production in *A. annua* (Farhi et al., 2011).

2.2.3 Metabolic engineering in microbes

The heterologous production of artemisinin in micro-organisms is another potential solution for the shortage of artemisinin. Plasmid-based expression of *S. cerevisiae* genes encoding the MVA pathway in combination with a codon-optimized *ADS* in *Escherichia coli* resulted in the estimated production of amorpha-4,11-diene up to 22.6 to 112.2 mg/L (Martin et al., 2003). The same group reported the engineering of *Saccharomyces cerevisiae* to produce up to 100 mg per liter of AA using an engineered MVA pathway, in combination with *ADS* and *CYP71AV1* (Ro et al., 2006). The conversion of AA into artemisinin was achieved by photochemical transformation involving a singlet-oxygen-induced ene reaction and the addition of triplet oxygen, which triggers the reaction cascade that incorporates the endoperoxide group (Lévesque and Seeberger, 2012). The biosynthetic pathway was further optimized by including an NAD-dependent alcohol dehydrogenase
(ADH1), and a ubiquitous electron transport protein of cytochrome P450s in the ER membrane, cytochrome b5 (CYb5). This further improved AA yield and fermentation titres reached 25 grams per liter of AA in *Saccharomyces cerevisiae* (Paddon et al., 2013). In addition, a chemical source of singlet oxygen was developed to convert DHAA to artemisinin, thus avoiding the need for specialized photochemical equipment (Paddon et al., 2013). The successfully produced semi-synthetic artemisinin is now in commercial production, demonstrating the potential of synthetic biology for the development and production of pharmaceutical compounds (Paddon and Keasling, 2014).

### 2.3 Transport in artemisinin biosynthesis pathway

Although the metabolic engineering work in plants improved the production of artemisinin, it is still far away from being economically attractive. Upon over-expression of the artemisinin biosynthesis genes in heterologous plants only a very low production of artemisinin was detected. This may result from the subcellular compartmentalization of the artemisinin biosynthesis genes, and the therefore suboptimal operation of the pathway in a heterologous system. For example, DBR2 is a soluble cytosolic protein while CYP71AV1 is embedded in the ER, so the rate of intracellular transport of intermediates between these enzymes may influence the metabolic flow (Olofsson et al., 2011). Thus, it is important to know how these molecules are transported intracellularly, but also intercellularly and/or how they are sequestered in the apoplast, where artemisinin is stored naturally, in the glandular trichomes, in *A. annua*. However, so far there is little knowledge about artemisinin transport. As far as I know there is just one reference about transporters in *A. annua*. After identifying the ABC transporter unigenes in the NCBI database, the transcription levels over the leaf development and under methyl jasmonate or abscisic acid treatments were tested, however, no any transport activity and products information were included in this literature (Zhang et al., 2012).

### 3 Thesis outline

The objective of my PhD project is to gain knowledge artemisinin biosynthesis pathway and to improve the production of artemisinin using different metabolic engineering strategies. After the general introduction in **Chapter 1**, **Chapter 2** compares the artemisinin biosynthesis genes in two different *A. annua* chemotypes. The ectopic expression of the artemisinin biosynthesis
pathway in *N. benthamiana* is established. Then two different biosynthesis genes, CYP71AV1, are identified and their contribution to the production difference between the chemotypes investigated. In Chapter 3, I explore the sequestration of artemisinin and its biosynthetic pathway intermediates in the apoplast. Two PDRs and three LTPs are identified and isolated from *A. annua*. PDR2 and LTP3 stimulate the sequestering of intermediates DHAA and AA in the apoplast. Chapter 4 describes how the activity of the flavonoid transcription factor *Rosea1* (ROS) stimulates sesquiterpene metabolic flux. Combining ROS with the *A. annua* transporter genes, *PDR2* and *LTP3*, improves sequestration of DHAA and AA in the apoplast and the final products artemisinin and arteannuin B are enhanced in *N. benthamiana*. Chapter 5 explores the role of introns in enhancing sesquiterpene production. Two sesquiterpene synthases, *CST* from *Arabidopsis*, and *ADS* from *A. annua* are included in both stable transgenic and transient expression assays with and without introns. Genomic sesquiterpene synthase overexpression yields higher production of the anticipated products. The results of this thesis are integrated and discussed in a broader perspective in Chapter 6. The future perspectives for metabolic engineering of natural secondary metabolites and the importance of transport are discussed as well.

Reference


General Introduction


Butelli, E., Titta, L., Giorgio, M., Mock, H.-P., Matros, A., Peterek, S., Schijlen,


but difference in subcellular localization. Metabolic Engineering 20, 198-211.


Holthuis, J.C.M., and Menon, A.K. (2014). Lipid landscapes and pipelines in mem-


189.


imental Botany 55, 1643-1654.


Rodríguez-Concepción, M., and Boronat, A. (2002). Elucidation of the methyerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A met-
abiotic milestone achieved through genomics. Plant Physiology 130, 1079-1089.


Tholl, D., and Lee, S. (2011). Terpene Specialized Metabolism in Arabidopsis thali-


Wu, S., Schalk, M., Clark, A., Miles, R.B., Coates, R., and Chappell, J. (2006). Redirec-
tion of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. Nature Biotechnology 24, 1441-1447.


Chapter 2

The metabolite chemotype of *Nicotiana benthamiana* transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage

Hieng-Ming Ting¹, Bo Wang¹, Anna-Margareta Rydén¹², Lotte Woittiez¹, Teun van Herpen¹, Francel W.A. Verstappen¹, Carolien Ruyter-Spira¹, Jules Beekwilder¹², Harro J. Bouwmeester¹ and Alexander van der Krol¹

¹Laboratory of Plant Physiology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
²Plant Research International, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.

Abstract

The plant *Artemisia annua* that produces the anti-malaria compound artemisinin, occurs as high artemisinin production (HAP) and low artemisinin production (LAP) chemotypes. Understanding the molecular basis of *A. annua* chemotype may help optimising artemisinin biosynthesis in heterologous production platforms. We present the first systematic comparison of artemisinin biosynthesis genes to determine factors that contribute to artemisinic acid (AA) or dihydroartemisinic acid (DHAA) chemotype of agro-infiltrated leaves with ADS, *CYP71AV1/AMO, DBR2* and *ALDH1*. Results show that the enzyme activity of DBR2 and ALDH1 from the two chemotypes does not differ. The Amorphadiene Oxidase from HAP (AMOHAP) showed reduced activity compared to that from LAP chemotype (AMOLAP), which relates to a seven amino acid N-terminal extension in AMOLAP compared to AMOHAP. The GFP fusion of both proteins show equal localization to the ER, but AMOLAP may be more stable. Product profile characterisation by LC-QTOF-MS/MS, UPLC-MRM-MS and GC-MS of transient expression in *Nicotiana benthamiana* show that AMOLAP not only displayed a higher enzyme activity but also affected the ratio of end products (e.g. leaf chemotype), which could be mimicked by reduced gene dosage of AMOLAP in the pathway. However, expression in combination with the *DBR2* and *ALDH1* also resulted in a qualitatively different product profile (‘chemotype’) when DBR2 infiltration dosage was diluted, shifting saturated (dihydro) branch toward unsaturated branch and of the pathway.

Keywords

Introduction

Based on the content of artemisinin and its precursors, two chemotypes of \textit{A. annua}, can be distinguished; the low-artemisinin production (LAP) chemotype and a high-artemisinin production (HAP) chemotype (Wallaart \textit{et al.}, 2000). Both chemotypes contain artemisinin and arteannuin B, but the HAP chemotype has a relatively high content of artemisinin and its presumed precursor dihydroartemisinic acid (DHAA), while the LAP chemotype has a high content of arteannuin B and its presumed precursor artemisinic acid (AA).

The artemisinin biosynthesis pathway has been largely elucidated and the genes required for production of dihydroartemisinic acid, the most likely precursor of artemisinin (\textit{ADS}, \textit{CYP71AV1}, \textit{DBR2} and \textit{ALDH1}) have all been described (Bouwmeester \textit{et al.}, 1999, Rydén \textit{et al.}, 2010, Teoh \textit{et al.}, 2009, Teoh \textit{et al.}, 2006, Zhang \textit{et al.}, 2008). Artemisinin is a sesquiterpene lactone endoperoxide, which is synthesized in the cytosol from the general isoprenoid precursors IPP and DMAPP. These are converted to FPP and the first committed step in the artemisinin biosynthetic pathway is the cyclization of FPP to amorpha-4,11-diene (AD) by amorphadiene synthase (Fig. 1) (Bouwmeester \textit{et al.}, 1999, Mercke \textit{et al.}, 2000). In the subsequent step, AD is oxidized by the cytochrome P450 enzyme, CYP71AV1/AMO, to artemisinic alcohol (AAOH), artemisinic aldehyde (AAA) and artemisinic acid (AA) (Fig. 1) (Ro \textit{et al.}, 2006, Teoh \textit{et al.}, 2006). However, the latter mainly occurs in the LAP chemotype. In the HAP chemotype only very little of the AAA is converted to AA, as most of the AAA is converted to dihydroartemisinic aldehyde (DHAAA) by DBR2, the enzyme that reduces the exocyclic double bond of AAA (Fig. 1) (Bertea \textit{et al.}, 2005, Zhang \textit{et al.}, 2008). Supposedly, DHAAA is subsequently oxidized by alcohol dehydrogenase ALDH1 to the final intermediate dihydroartemisinic acid (DHAA) (Bertea \textit{et al.}, 2005, Zhang \textit{et al.}, 2008). The conversion of DHAA to artemisinin, is believed to be a non-enzymatic and spontaneous photo-oxidation reaction (Sy&Brown, 2002, Wallaart \textit{et al.}, 1999). Similarly, in the LAP chemotype, AA is likely spontaneously converted to arteannuin B.

Recently we reported on the (transient) reconstruction of the artemisinin biosynthetic pathway in \textit{Nicotiana benthamiana} leaves, resulting in up to 39.5 mg.kg\(^{-1}\) FW of AA (van Herpen \textit{et al.}, 2010). In the present work we analyse the role of DBR2, ALDH1 and CYP71AV1 in determining the ‘chemotype’ (as defined by the AA and DHAA ratio) of \textit{N. benthamiana} leaves agro-infiltrated with artemisinin biosynthesis genes. Results show that the chemotype is a function of the CYP71AV1 type and relative dosage of DBR2 and ALDH1.
Chapter 2

Materials and Methods

Cloning of the ADS+FPS+HMGR expression construct AmFH.

Cloning of the AmFH expression construct which contains ADS with CoxIV mitochondrial targeting signal, mitochondrial targeted FPS and a cytosolic (truncated) HMGR under control of the CaMV35S promoter has been described previously (van Herpen et al., 2010).

Identification and cloning of CYP71AV1 from HAP and LAP chemotypes

The AMOHAP EST sequence was identified in the sequence database of an Artemisia annua L. (HAP chemotype) glandular trichome cDNA library (Bertea et al., 2006). This sequence has been deposited in the GenBank database (JQ254992). The full length sequence of AMOHAP was obtained by RACE PCR (Clontech, Mountain View, CA, USA). We note that the race PCR did not yield any sequences that indicated the presence of a AMOLAP
version in our cDNA library. Subsequently, the full length coding region was amplified from *Artemisia annua* trichome cDNA by PCR using Phusion polymerase (Finnzymes, Espoo, Finland) using primers 1+2 (Table S1). The full length AMOLAP coding region was cloned by primers 3+4. After confirmation of the correct sequences both the AMOHAP and AMOLAP coding region were isolated from pGEMT/AMOHAP and pGEMT/AMOLAP using BamHI and KpnI for cloning into the yeast expression vector pYEDP60 (Pompon *et al.*, 1996), resulting in pYEDP60-AMOHAP and pYEDP60-AMOLAP.

For the cloning into a plant binary expression vector, the genes were first introduced into ImpactVectorC3.1 (http://www.wageningenur.nl/en/show/Productie-van-farmaceutische-en-industriele-eiwitten-door-plante.htm). For this a Kpn1 site was introduced into C3.1, resulting in C3.1/Kpn1. Hereto, two oligo’s (GATCCATTTCGGTACCAATTAGC and GGC-CGCTAATTGGTACCCGAATG) were hybridized, kinase-treated and ligated into C3.1 digested with BamHI and NotI. The BamHI/KpnI fragment of pGEMT/AMOHAP and pGEMT/AMOLAP was isolated and ligated into the vector C3.1/Kpn1 between the CaMV35S promoter and Rbcs1 terminator. The resulting plasmids C3.1/AMOHAP and C3.1/AMOLAP were digested with Ascl and PacI and the full gene sequence was cloned into the Ascl and PacI site of the pBinPlus binary vector (van Engelen *et al.*, 1995).

**Identification and cloning of DBR2 from HAP and LAP chemotypes**

We cloned a DBR2HAP from *Artemisia annua* HAP chemotype (Bertea *et al.*, 2006) and DBR2LAP was isolated from *Artemisia annua* LAP chemotype from Iran (Table S2). Both DBR2 cDNA sequences were isolated by RT-PCR from cDNA constructed from RNA extracted from *Artemisia annua* flowers isolated from either chemotype. For amplification of the DBR2 sequence, primers 5+6 were used (Table S1), of which sequences were based on the published DBR2 sequence from a HAP chemotype (Zhang *et al.*, 2008). These primer sets were also able to amplify a DBR2 sequence from the LAP chemotype. The primers used introduce BamHI and NotI restriction sites which were used for cloning into ImpactVectorplV1A_2.1 (http://www.wageningenur.nl/en/show/Productie-van-farmaceutische-en-industriele-eiwitten-door-plante.htm). The resulting plV1A_2.1/DBR2 was digested with Ascl and PacI and the full gene was cloned into the Ascl and PacI sites of the pBinPlus binary vector (van Engelen *et al.*, 1995). DBR2 sequences identified here have been deposited in the GenBank database (KC505370, JX898526 and JX898527).
Identification and cloning of ALDH1 from HAP and LAP chemotypes

ALDH1 sequences were isolated by RT-PCR on cDNA from floral RNA isolated from the Artemisia annua HAP and LAP chemotype using primers 7+8 (Table S1). The sequences of these primers were based on the published ALDH1 sequence from a HAP chemotype (Teoh et al., 2009) and could amplify an ALDH1 sequence from both HAP and LAP cDNA. Inspection of the coding sequence revealed that there were no specific amino acid residue differences in ALDH1 from HAP and LAP (both identical to ALDH1 GenBank: FJ809784). The ALDH1 cDNA was cloned into pBinPlus binary vector (van Engelen et al., 1995).

Cloning of CYP71AV1 GFP reporter constructs

For the GFP fusion protein expression constructs (NtermAMOHAP:GFP, NtermAMOLAP:GFP, AMOHAP:GFP and AMOLAP:GFP): the N-terminal domains of AMOHAP (first 43 codons) and AMOLAP (first 50 codons) were amplified using primers 9+10 and 11+12, respectively. After digestion with BamHI and KpnI the fragments were cloned into C3.1/Kpn1 to form C3.1/NtermAMOHAP and C3.1/NtermAMOLAP. The GFP coding sequence was amplified by PCR using primers 13+14 (Table S1) using pBin-Egfp as template (http://www.wageningenur.nl/en/show/Productie-van-farmaceutische-en-industriele-eiwitten-door-planten.htm). After digesting with KpnI and NotI, GFP was subcloned into plasmid C3.1/NtermAMOHAP, C3.1/NtermAMOLAP, C3.1/AMOHAP and C3.1/AMOLAP to form 35S-NtermAMOHAP:GFP, 35S-NtermAMOLAP:GFP, 35S-AMOHAP:GFP and 35S-AMOLAP:GFP.

To remove the stop codon of AMOHAP and AMOLAP and fuse them in-frame to the ATG start-codon of the GFP coding sequence in the 35S-AMOHAP:GFP and 35S-AMOLAP:GFP constructs, the 3’ends of the AMOLAP and AMOHAP were PCR amplified using Phusion DNA polymerase (Finnzymes) by primers 15+16, and 17+18, respectively.

The PCR products and plasmids 35S-AMOHAP:GFP and 35S-AMOLAP:GFP were digested with KpnI and EcoRI, followed by gel purification. The PCR products were ligated into the plasmids and transformed into Escherichia coli. Positive colonies were analyzed and sequenced to confirm that inserts were correct.
Transient expression in leaves of *Nicotiana benthamiana*

Agro-infiltration for transient expression in leaves of *Nicotiana benthamiana* Domin was done as described (van Herpen et al., 2010). Briefly, individual *Agrobacterium tumefaciens* strains with different expression constructs (or empty vector as control) were co-infiltrated into *Nicotiana benthamiana* leaves using a syringe without needle. After seven days of transient expression, leaves were harvested for chemical analysis. In each set of experiments the total dosage of *Agrobacterium tumefaciens* between treatments was the same by diluting with *Agrobacterium tumefaciens* with empty vector where necessary. We note that leaves infiltrated with AmF-H+AMOLAP developed necrotic lesions, indicating that at this time certain compounds started to accumulate to toxic levels in the infiltrated leaves.

**CYP71AV1 subcellular localization studies**

For subcellular localization of the AMO:GFP fusion proteins, *Arabidopsis thaliana* (L.) Heynh protoplasts were isolated and transfected with expression constructs 35S-NtermAMOHAP:GFP, 35S-NtermAMOLAP:GFP, 35S-AMOHAP:GFP and 35S-AMOLAP:GFP, based on a published protocol (Yoo et al., 2007). As reference for ER subcellular localization the ER-YFP construct was used (Aker et al., 2006). After transfection, protoplasts were analyzed by a Carl-Zeiss Confocal Scanning Laser Microscopy, with excitation of GFP at 488 nm and YFP at 514 nm. The fluorescence was detected via a band pass filter (GFP: 505-530 nm, YFP: 535-590 nm). Chlorophyll was detected using a 650 nm long pass filter.

**Analysis of non-volatile metabolites by LC-QTOF-MS/MS**

Seven days after agro-infiltration of *Nicotiana benthamiana*, the infiltrated leaves were harvested and immediately frozen in liquid nitrogen and ground to a fine powder. From each infiltrated leaf, 100 mg of powder was extracted in 300 µl methanol: formic acid (1000:1, v/v). Non-volatile compounds from the infiltrated leaves were analyzed by LC-QTOF-MS as described (van Herpen et al., 2010).

Data were processed using the protocol for untargeted metabolomics of plant tissues as described (De Vos et al., 2007, van Herpen et al., 2010). Briefly, LC-QTOF-MS data were analyzed using Masslynx 4.0 (Waters) and processed using MetAlign version 1.0 (www.metAlign.nl) for base-
line correction, noise elimination and subsequent spectral data alignment (De Vos et al., 2007). The processing parameters of MetAlign for LC-QTOF-MS data were set to analyze from scan numbers 60-2590 (corresponding to retention time 1.15-49.16 min) with a maximum amplitude of 25,000. After MetAlign processing, masses were clustered using the Multivariate Mass Spectra Reconstruction (MMSR) approach (Tikunov et al., 2005) to elucidate which mass signals originate from the same metabolite. The mass signal intensity differences between treatments were compared using the student’s t-test. Mass-directed LC-QTOF-MS/MS analysis for further elucidation of metabolite identities was done on differential compounds with signal intensities higher than 500 ion counts per scan.

**Quantification of artemisinin precursors by UPLC-MRM-MS**

Targeted analysis of artemisinin precursors in agro-infiltrated *Nicotiana benthamiana* leaves was performed with a Waters Xevo tandem quadrupole mass spectrometer equipped with an electrospary ionization source and coupled to an Acquity UPLC system (Waters) as described (Kohlen et al., 2011) with some modifications. For details and instrument settings see supplemental data (Methods S1).

**Analysis of volatile metabolites by GC-MS**

Extracts were analysed by GC-MS using a gas chromatograph (7890A; Agilent, Amstelveen, the Netherlands) equipped with a 30-m x 0.25-mm i.d., 0.25-mm film thickness column with 5-m guard column (Zebron ZB5-MS; Phenomenex, Utrecht, The Netherlands) and a mass selective detector (model 5975c, Agilent). The GC was programmed at an initial temperature of 80°C for 1min, with a ramp of 5°C min⁻¹ to 235°C and then a ramp of 25 °C min⁻¹ to 280°C with a final time of 5 min. The injection port temperature was 250°C, and the He inlet pressure was controlled with electronic pressure control to achieve a constant column flow of 1.0 mL min⁻¹. 1 µl of the extracts was injected in split mode with a split flow set at 9 ml min⁻¹. Scanning was performed from 45 to 450 atomic mass units.

**Glycosidase treatment**

Viscozyme L (Sigma) was used as glycosidase treatment to hydrolyze hexose-conjugated compounds for subsequent quantification us-
ing GC-MS. Hereto, 200 mg infiltrated leaf material from each treatment was incubated in 1 ml citrate phosphate buffer, pH 5.4 containing 200 µl of Viscozyme L as previously described (van Herpen et al., 2010).

**Glutathione conjugation assay**

*In vitro* conjugation of metabolites to glutathione by glutathione transferase activity (GST) was performed as described (Liu et al., 2011). In brief, glutathione (GSH) (150 mM) in 7 µl potassium phosphate buffer (100 mM; pH 6.5), and 30 mM of artemisinin precursor (AAA, AAOH, AA, DHAAOH, DHAAA and DHAA) in 7 µl ethanol were added to 200 µl potassium phosphate buffer (100 mM; pH 6.5). The reaction was initiated by adding 7 µl of glutathione transferase (GST) (1g L⁻¹, in 100 mM KH₂PO₄ potassium phosphate buffer; pH 6.5) into the mixture. The controls were complete assay mixtures without GST enzyme or either of the substrates. After 15 min incubation at room temperature, samples were cooled to -20°C until LC-QTOF-MS analysis.

**Results**

**Comparison of artemisinin biosynthesis protein sequences from HAP and LAP chemotypes reveals only relevant differences for CYP71AV1**

Because the difference between the HAP and LAP chemotypes must arise after the ADS step in the biosynthesis pathway, and because no expression differences were found for *ADS* genes in HAP and LAP chemotypes (Maes et al., 2011), here we focussed on analysis of putative differences in biosynthesis genes downstream of *ADS* (e.g. CYP71AV1, DBR2, ALDH1) in search of an explanation for the two different chemotypes of *A. annua*. For this purpose CYP71AV1, DBR2 and ALDH1 were isolated from an *A. annua* HAP and LAP chemotypes and the encoded protein sequences compared.

**CYP71AV1:** Analysis of the different CYP71AV1 sequences that have been deposited in GenBank shows the occurrence of two major types of CYP71AV1, encoding two proteins which differ by a seven amino acids extension at the N-terminus of the protein (Fig. 2). The long version of CYP71AV1 (which we refer to as AMOLAP) has been isolated from *A. annua* Tanzania (Sandeman seed), which is a LAP chemotype (Ro et al., 2006). The other long version of CYP71AV1 (which we refer to as AMOLAP.1) was isolated from a different *A. annua* LAP chemotype (Kim et al., 1992) (Soon-Un Kim,
personal communication). Two versions of the CYP71AV1 (here referred to as AMOHAP and AMOHAP.1) were cloned from two different HAP chemotypes (Bertea et al., 2006, Teoh et al., 2006). The alignment of the AMOHAP and AMOLAP variants shows that none of the other single amino acid substitutions between the different AMOLAP and AMOHAP sequences are specific to the long or the short version of CYP71AV1 and therefore likely do not play a role in determining the LAP or HAP chemotypes (Fig. 2).

RACE-PCR was used to analyse multiple CYP71AV1 5’sequences amplified from RNA isolated from a HAP chemotype and only AMOHAP 5’sequences were found (Fig. S1). Variation in the 5’untranslated region (nt 56-60) could be an indication that two different alleles of AMOHAP are present in this chemotype, both translating into the short AMOHAP. The RACE-sequence data were consistent with the recently published sequence of the CYP71AV1 promoter cloned from an A. annua HAP chemotype (Wang et al., 2011).

DBR2: We cloned DBR2 from the A. annua HAP chemotype (here referred to as DBR2HAP.1) and the sequence we obtained was similar to the recently described DBR2 (here referred to as DBR2HAP) (Zhang et al., 2008).

Figure 2. Alignment of the AMOLAP and AMOHAP deduced amino acid sequences. All amino acid sequences used in the alignment were retrieved from GenBank (GenBank number given in front of each sequence). Two of the sequences were from confirmed LAP chemotypes: ABB82944, referred to as AMOLAP (Ro et al., 2006) and ACF74516, referred to as AMOLAP.1 (Kim et al., 1992). Two of the sequences were from confirmed HAP chemotypes: AFP19100, referred to as AMOHAP (Bertea et al., 2006) and ABC41927, referred to as AMOHAP.1 (Teoh et al., 2006). For the other sequences deposited in GenBank the chemotype of the plant from which they were isolated was not given.
with the exception of a one amino acid difference (Fig. 3). Using primers based on DBR2HAP we isolated two variants of DBR2 from an A. annua LAP chemotype (see methods). Five of the eleven clones showed few amino acid differences with the published DBR2HAP (here referred to as DBR2LAP). However, in the remaining six clones (here referred to as DBR2LAP), the encoded protein sequence showed a number of amino acid residue differences with DBR2HAP, including two additional amino acids in position 295 (Fig. 3).

**ALDH1:** Cloning of ALDH1 from A. annua has been described (Teoh et al., 2009). We isolated ALDH1 from both the HAP and LAP chemotypes. Alignment of the AA-sequence showed no differences between the two proteins (data not shown).

Both AMOHAP and AMOLAP are localized to the Endoplasmic Reticulum

Since AMOLAP and AMOHAP only consistently differ in their N-terminal amino acid residues which supposedly encode the ER anchoring domain (Fig. 2), we investigated whether this difference causes altered subcellular targeting or difference in protein stability. Expression constructs encoding either full-length protein-GFP fusions or truncated N-terminal domain-GFP fusions of AMOLAP and AMOHAP were transiently expressed in A. thaliana protoplasts. Confocal microscopy of Arabidopsis protoplasts co-transfected with the full length AMOLAP protein fused to GFP (AMOLAP::GFP) showed co-localization of the GFP fluorescence signal with the fluorescence signal of the ER marker (ER::YFP) (Fig. 4a). Similarly, full length AMOHAP protein fused to
GFP (AMOHAP:GFP) also showed co-localization with the ER marker (ER:YFP) (Fig. 4c). In addition, the truncated N-terminal portion of AMOLAP and AMOHAP were fused to GFP. When transfected into Arabidopsis protoplasts the NtermAMOLAP:GFP and NtermAMOHAP:GFP both showed co-localization with the ER:YFP ER marker (Figs. 4c,d). Localization experiments with 35S expression constructs may lead to artefacts such as cytosolic localization when ER import is saturated, however this was not observed in these experiments. Combined, the results demonstrate that AMOLAP and AMOHAP do not differ in subcellular targeting, as both proteins localize to the ER. Although the fluorescence signal varies between transfection assays, there were indications that AMOLAP may be more stable. For instance we do find a higher fluorescence signal for NtermAMOLAP:GFP than for the NtermAMOHAP:GFP.

**Figure 4.** Subcellular localization of AMOLAP and AMOHAP. Confocal microscopy analysis of Arabidopsis protoplasts co-transfected with (a) AMOLAP:GFP + ER:YFP; (b) NtermAMOLAP:GFP + ER:YFP; (c) AMOHAP:GFP + ER:YFP; (d) NtermAMOHAP:GFP + ER:YFP. Artificial colors were given to GFP fluorescence (green), YFP fluorescence (red), and auto fluorescence of chloroplasts (purple). Merging of the pictures results in a yellow color for GFP-YFP overlap. The scale bar = 10 µm.
Different product profiles in planta from AMOLAP and AMOHAP

Both the AMOLAP and the AMOHAP enzymes have been characterised in a yeast expression system and both were shown to be able to produce AAOH, AAA and AA from AD (Ro et al., 2006, Teoh et al., 2006). However, a direct comparison of variants AMOLAP and AMOHAP in the same expression system has not been performed until now. The different in planta expression studies using heterologous plant hosts (tobacco and N. benthamiana) have only been reported for the longer AMOLAP (van Herpen et al., 2010, Zhang et al., 2011). To test the effect of the seven AA extension of the AMOLAP protein we made two expression constructs, both based on the AMOHAP sequence but in one construct we introduced the seven AA extension to the protein sequence as found in AMOLAP, thus limiting the difference between the two forms to the N-terminal extension. The activity of these AMOLAP and AMOHAP genes was subsequently compared in planta by co-expression with ADS, using transient expression in N. benthamiana leaves. To achieve high levels of artemisinin precursor production, ADS was expressed with a mitochondrial targeting signal, and overexpression was combined with a mitochondrial targeted FPS and a truncated, cytosolic form of HMGR. ADS, FPS and HMGR were combined into a single 2A expression construct (AmFH) as described before (van Herpen et al., 2010). Each expression construct (AmFH and AMOLAP or AMOHAP) was introduced into A. tumefaciens and N. benthamiana leaves were infiltrated with AmFH+AMOLAP or AmFH+AMOHAP.

Particularly with AmFH+AMOLAP, infiltrated leaves developed symptoms of necrosis around seven days post infiltration (Fig. S2), suggesting the production of a toxic compound. Necrosis symptoms were stronger in leaves infiltrated with AmFH+AMOLAP than with AmFH+AMOHAP, suggesting that the products of both treatments may not be the same. Because necrosis started to appear after seven days, in all our experiments the leaves were harvested at day seven instead of day ten after infiltration, as previously done (van Herpen et al., 2010).

Analysis of free products: To quantify the products from the ADS and AMOLAP/HAP transient enzyme activity in the infiltrated N. benthamiana, leaves were extracted with aqueous methanol for UPLC-MRM-MS analysis. Intriguingly, the distribution over the entire product range was different in leaves infiltrated with AmFH+AMOLAP and AmFH+AMOHAP (Table 1). Leaves infiltrated with AmFH+AMOLAP produced predominantly AA, while in leaves infiltrated with AmFH+AMOHAP AA levels were 50-fold lower. However, leaves infiltrated with AmFH+AMOHAP contained 3-fold higher levels of AAOH and AAA than leaves expressing AmFH+AMOLAP (Table 1).
Also DHAAOH, DHAAA and DHAA were detected in the AmFH+AMOLAP leaf samples (Table 1), suggesting the presence of an endogenous N. benthamiana enzyme with carbon double bond reducing activity (catalysing the conversion of AAA to DHAAA just as DBR2 in A. annua), and enzymes similar to A. annua RED1 (catalysing the formation of DHAAOH from DHAAA) and A. annua ALDH1 (catalysing the conversion of DHAAA to DHAA) (Fig. 1). Free DHAAOH levels were higher in AmFH+AMOLAP compared to AmFH+AMOHAP infiltrated leaves, suggesting that formation of DHAAOH is not directly related to the free DHAAA, AAOH or AAA levels in leaves, which were lower in AmFH+AMOLAP. DHAA was only detected in AmFH+AMOLAP infiltrated leaves.

Table 1. Unconjugated artemisinin precursors produced in Nicotiana benthamiana as identified and quantified by UPLC-MRM-MS

<table>
<thead>
<tr>
<th>(ng g⁻¹ FW)</th>
<th>AmFH+AMOLAP</th>
<th>AmFH+AMOHAP</th>
<th>AmFH+AMOLAP+DBR2</th>
<th>AmFH+AMOHAP+DBR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAOH</td>
<td>16709 ± 3977</td>
<td>46946 ± 5692</td>
<td>3596 ± 1247</td>
<td>16554 ± 3233</td>
</tr>
<tr>
<td>AAA</td>
<td>5501 ± 1486</td>
<td>19564 ± 6314</td>
<td>781 ± 246</td>
<td>3418 ± 890</td>
</tr>
<tr>
<td>AA</td>
<td>3969 ± 1391</td>
<td>74 ± 15</td>
<td>53 ± 16</td>
<td>ND</td>
</tr>
<tr>
<td>DHAAOH</td>
<td>1821 ± 576</td>
<td>597 ± 19</td>
<td>87972 ± 15014</td>
<td>57289 ± 9455</td>
</tr>
<tr>
<td>DHAAA a</td>
<td>(5966 ± 1646)</td>
<td>(8067 ± 1341)</td>
<td>(220347 ± 65373)</td>
<td>(124787±26145)</td>
</tr>
<tr>
<td>DHAA</td>
<td>17 ± 8</td>
<td>ND</td>
<td>838 ± 517</td>
<td>25 ± 9</td>
</tr>
</tbody>
</table>

AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAAOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid. ND: not detectable. a The values for DHAAA are shown in brackets as they represent peak intensities and not concentrations.

Analysis of glycosylated products: Previous results showed that most of the products of the ADS and AMOLAP activity in N. benthamiana agro-infiltration are present as glycosylated conjugates, mainly of AA (van Herpen et al., 2010). Therefore, leaf material was also analysed by LC-QTOF-MS. N. benthamiana leaves co-expressing AmFH+AMOLAP indeed contain AA-12-β-di-glucoside, as previously reported (van Herpen et al., 2010). However, in addition several other AA-glycoside conjugates were detected, including conjugates with additional hexose units as well as malonylated hexoses. Also for the glycosylated products the distribution over the entire product range was different between AmFH+AMOLAP and AmFH+AMOHAP (Fig. 5a, Table 2 (Page 80)). Leaves infiltrated with AmFH+AMOLAP produced more AA conjugates, while leaves infiltrated with AmFH+AMOHAP produced more AAOH conjugates (Table 2, Fig. S3a). For both treatments also several DHAAOH
and DHAA conjugates with hexose and malonyl groups were detected, but no DHAAA conjugates (Table 2). Table 2 shows the mass fragmentation profiles of the detected products and their putative identification. MS/MS analysis was used to further confirm product identity and an example of the identification of one of the DHAA-hexose conjugates is shown in Fig. S5.

To quantify the levels of glycosylated products, samples were treated with a mix of glycosidases (Viscozyme L) and deglycosylated products were quantified using GC-MS. Note that the Viscozyme treatment only cleaves hexose conjugates but not malonylated hexose conjugates (Fig. S4). Results show that leaves infiltrated with AmFH+AMOLAP contained c. 40 mg.kg⁻¹ FW of AA [consistent with the previously reported 39.5 mg.kg⁻¹ FW of AA (van Herpen et al., 2010)], while the sensitivity of the GC-MS was not sufficient to detect any AA in leaves infiltrated with AmFH+AMOHAP (Table 3). GC-MS analysis after Viscozyme treatment confirmed that AAOH was the major glycosylated product in leaves infiltrated with AmFH+AMOHAP (as was suggested by Table 2) at 24 mg.kg⁻¹ FW.

Table 3. Artemisinin precursors in Nicotiana benthamiana agro-infiltrated with artemisinin biosynthetic pathway genes.

<table>
<thead>
<tr>
<th>(mg kg⁻¹ FW)</th>
<th>AmFH+AMOLAP</th>
<th>AmFH+AMOHAP</th>
<th>AmFH+AMOLAP+DBR2</th>
<th>AmFH+AMOHAP+DBR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAOH</td>
<td>8.1 ± 1.6</td>
<td>24.0 ± 3.5</td>
<td>5.1 ± 0.5</td>
<td>9.4 ± 2.1</td>
</tr>
<tr>
<td>AAA</td>
<td>1.6 ± 0.1</td>
<td>1.6 ± 0.1</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>AA</td>
<td>39.9 ± 9.8</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>DHAAOOH</td>
<td>1.6 ± 0.2</td>
<td>2.0 ± 0.2</td>
<td>42.9 ± 14.9</td>
<td>22.8 ± 8.6</td>
</tr>
<tr>
<td>DHAAA</td>
<td>ND</td>
<td>ND</td>
<td>4.0 ± 2.3</td>
<td>1.3 ± 0.6</td>
</tr>
<tr>
<td>DHAA</td>
<td>ND</td>
<td>ND</td>
<td>7.3 ± 2.2</td>
<td>ND</td>
</tr>
</tbody>
</table>

Agro-infiltrated leaves were treated with glycosidase (Viscozyme L.) and hydrolysed metabolites extracted and analysed by GC-MS. AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAAOOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid; ND: not detectable. Results are means ± S.D. of three co-infiltrated leaves.

**No difference in DBR2 activity from HAP and LAP A. annua chemotypes**

We compared the activity of the two variants of DBR2 by comparing product profiles of N. benthamiana leaves infiltrated with AmFH+AMOLAP in combination with either DBR2HAP or DBR2LAP. In addition we tested the
two DBR2 variants in combination with AmFH+AMOHAP. Analysis of the conjugated products show that there is no difference in product profile between DBR2HAP and DBR2LAP (Fig. 5b, S3b, Table S3), indicating that

Figure 5. Artemisinic acid (AA) and dihydroartemisinic acid (DHAA) related compounds in leaves of *N. benthamiana* agro-infiltrated with artemisinin biosynthesis genes as identified by LC-QTOF-MS. (a) Agro-infiltrated leaves with AmFH+AMOLAP and AmF-H+AMOHAP. For all products identified by LC-QTOF-MS see Table 2. (b) Agro-infiltrated leaves with AmFH+AMOLAP/HAP+DBR2LAPI/HAP. For all products identified by LC-QTOF-MS see Table S3. (c) Agro-infiltrated leaves with AmFH+AMOLAP/HAP+DBR2LAPI/HAP+ALDH1. For all products identified by LC-QTOF-MS see Table S3. Data represent peak intensities for each of the compounds in LC-QTOF-MS analysis. Peak intensities are the mean of three agro-infiltrated leaves.
the two forms of DBR2 do not differ in enzymatic activity. The co-infiltration with DBR2 relieved the necrosis symptoms caused by expression of AmF-H+AMOLAP or AmFH+AMOHAP alone (Fig. S2), suggesting that additional DBR2 enzyme activity lowered the level of the product(s) that cause necrosis. Product analysis in leaves agro-infiltrated with AmFH+AMOLAP+DBR2 or AmFH+AMOHAP+DBR2 showed that DBR2 activity resulted in a significant increase in DHAAOH, DHAAA and DHAA levels (Table 1) and this is also clear from LC-QTOF-MS analysis that shows a strong increase in DHAAOH and DHAA conjugates to hexose and malonyl groups (Table 2).

The analysis of deglycosylated extracts by GC-MS confirmed that co-expression of DBR2 increased the levels of DHAAOH, DHAAA and DHAA at the expense of AAOH, AAA and AA levels (Table 3). The total yield of DHAA in leaves agro-infiltrated with AmFH+AMOLAP+DBR2 as released by glycosidase treatment was c. 7.3 mg.kg⁻¹ FW while DHAA in leaves agro-infiltrated with AmFH+AMOHAP+DBR2 was below the level of detection by GC-MS. Combined, these results show that DBR2 further enhances the double bond reduction of the CYP71AV1 products that is also already catalysed by endogenous tobacco reductase activity. In addition, endogenous tobacco glycosyl and malonyl transferases modify these double-bond-reduced products which leads to DHAAOH and DHAA conjugates.

**Increased DHAA and AA by combining ADS, AMO and DBR2 with ALDH1**

As described above, no differences were found between the ALDH1 protein sequence from LAP and HAP A. *annua* chemotypes. To test how the addition of ALDH1 activity affects the product profile of the artemisinin HAP and LAP biosynthesis pathway, leaves were agro-infiltrated with AmFH+AMOLAP+DBR2+ALDH1 or AmFH+AMOHAP+DBR2+ALDH1. After seven days leaves were extracted and products were profiled by LC-QTOF-MS. The levels of conjugated DHAAOH products significantly decreased when ALDH1 was added to AmFH+AMOLAP+DBR2 (Table S3, Fig. S3b), coinciding with a substantial increase in glycosylated AA and DHAA product levels (Fig. 5c). This suggests that ALDH1 may be more efficient in the conversion of AAA to AA and DHAAA to DHAA than AMOLAP and AMOHAP as already suggested by the work of Teoh (Teoh *et al*., 2009). Although the level of the presumed direct precursor of artemisinin (DHAA) was substantially increased by ALDH1 (c. 13-fold by adding ALDH1 to AmFH+AMOLAP+DBR2 and c. 110-fold by adding ALDH1 to AmFH+AMOHAP+DBR2), no artemisin-
inin could be detected in *N. benthamiana* by UPLC-MRM-MS (Table 1, 4).

**Qualitative effects on product profile by AMOLAP dosage**

The comparison of the total product levels produced by the reconstituted pathway with AMOLAP or AMOHAP suggests that AMOLAP has a higher enzyme activity than AMOHAP, in combination with a different product profile (Fig. 5a, Table 2). This difference could not be related to different subcellular localization (Fig. 4). To test if differences in relative enzyme activity within the pathway can affect the product profile in a qualitative way, we tested the effect of different dilutions of AMOLAP in combination with the rest of the biosynthesis pathway (*ADS+AMOLAP+DBR2HAP*). This was achieved by diluting the *Agrobacterium* strain carrying the *AMOLAP* expression construct with a suspension of an *Agrobacterium* strain carrying an empty expression vector to keep the total *Agrobacterium* dosage for infiltration the same. Results show that reduction of the AMOLAP agro-infiltration dosage to 1/2 had only little effect, but that dilution up to 1/10 strongly decreased DHAA-glycoside production (Figs. 6a, S3c; Table S4). The AA-glycoside levels were also reduced but to a much lower extent.

**More AAA-related glutathione-conjugate from AMOHAP than from AMOLAP**

Recently we described the reconstruction of the biosynthetic pathway of costunolide in *N. benthamiana* (Liu *et al.*, 2011) in which it was shown that the exocyclic carbon double bond of costunolide conjugates to glutathione (GSH). Since some of the products of ADS (AD) and AMOLAP/HAP enzyme activities also contain such an exocyclic double bond (AAA), but lack the hydroxyl or acid group used for glycosylation in AAOH and AA, we specifically looked for GSH conjugates of artemisinin intermediates in *N. benthamiana* leaves agro-infiltrated with the biosynthetic pathway genes. A putative GSH-conjugated compounds was detected by LC-QTOF-MS, both in negative mode (*m/z*=542.25) and positive mode (*m/z*=544.24) (Fig. S6). The level of the GSH-conjugate (*m/z*=542.25) was higher in *AmFH+AMOHAP* than in *AmFH+AMOHAP+DBR2* agro-infiltrated leaves (Table 2). We tested the artemisinin biosynthetic pathway intermediates (AAA, AAOH, DHAAOH, AA, DHAA and DHAAA) in *in vitro* reactions for spontaneous or glutathione-S-transferase (GST) driven GSH conjugation. Only AAA formed an AAA-GSH conjugate similar to that extracted from agro-infiltrated leaves express-
Figure 6. Artemisinic acid (AA) and dihydroartemisinic acid (DHAA) conjugated compounds in leaves of *N. benthamiana* agro-infiltrated with AMOLAP/HAP and different dosage of DBR2 and ALDH1. (a) Comparison AA and DHAA conjugated compounds in agro-infiltrated leaves with dilution of AMOLAP/HAP. For all products identified by LC-QTOF-MS see Table S4. (b) Comparison AA and DHAA conjugated compounds in agro-infiltrated leaves with dilution of DBR2. For all products identified by LC-QTOF-MS see Table S5. (c) Comparison AA and DHAA conjugated compounds in agro-infiltrated leaves with dilution of ALDH1. For all products identified by LC-QTOF-MS see Table S5. (d) Comparison AA and DHAA conjugated compounds in agro-infiltrated leaves with dilution of DBR2 and ALDH1. For all products identified by LC-QTOF-MS see Table S5. Data represent peak intensities for each of the compounds in LC-QTOF-MS analysis.
ing the pathway genes \((m/z=526.22)\) (Fig. S6 and S7). The mass of the major GSH conjugate formed \textit{in planta} is 18 D higher, suggesting an additional two protons and one oxygen atom, which could be explained by hydroxylation of the endocyclic double bond in AAA (Fig. S6). The level of the putative AAA glutathione conjugate was higher in leaves infiltrated with \(AmFH+AMOHAP\) than in leaves infiltrated with \(AmFH+AMOLAP\) (Table 2). Dilution of AMOLAP resulted in an increase in the level of the AAA glutathione conjugate, although not reaching the level of the full dosage of AMOHAP (Fig. 7).

**Effects on product profile by DBR2 or ALDH1 dosage**

![Graph showing the effects of DBR2 and ALDH1 dosage on product profile.](image)

Figure 7. Artemisinic aldehyde (AAA) conjugated compounds in leaves of \(N. benthamiana\) agro-infiltrated with dilution of AMOLAP/HAP. Data represent peak intensities for each of the compounds in LC-QTOF-MS analysis. Peak intensities are the mean ± S.D. of three agro-infiltrated leaves.

We also tested the effect of \(DBR2\) and \(ALDH1\) gene dosage on DHAA:AA related product ratio by testing dilutions of DBR2 and ALDH1 in combination with AMOLAP or AMOHAP. Product analysis by LC-QTOF-MS of glycosylated products showed that with full DBR2 and ALDH1 agro-infiltration dosage the ratio of DHAA:AA was more skewed towards the DHAA branch of the pathway. However, when the DBR2 infiltration dosage was diluted 10-fold, the ratio of DHAA:AA decreased, shifting the pathway activity more towards the AA branch, with relatively little effect on the total product level (Fig. 6b, Fig. 1).

In contrast, 10-fold dilution of ALDH1 agro-infiltration dosage resulted in a strong decrease in glycosylated AA conjugates, while the glycosylated
DHAA conjugates were hardly affected (Fig. 6c), suggesting that ALDH1 has a preference for the DHAAA substrate over the AAA substrate (Fig. 1).

When both DBR2 and ALDH1 agro-infiltration dosage were diluted 10-fold the DHAA:AA related products ratio came close to one, with slight preference for the DHAA branch of the pathway with AMOHAP and slight preference for the AA branch of the pathway with AMOLAP (Fig. 6d). Combined these results suggest that the agro-infiltrated leaf ‘chemotype’ is determined by a combination of the AMOLAP/AMOHAP catalytic effectivity in combination with especially DBR2 gene dosage (Figs. 6, S3d; Table S5).

Discussion

Here we have compared the different proteins from the artemisinin biosynthesis pathway encoded by genes isolated from high and from low artemisinin producing *A. annua* chemotypes. The expression levels of artemisinin biosynthesis genes (*ADS, AMO, DBR2* and *ALDH1*) were recently analysed in *A. annua* HAP and LAP chemotypes showing that the chemotype identities of *A. annua* did not correlate with differences in expression level of these genes in the absence of stress (Maes *et al.*, 2011). Nevertheless, total product yield in 5-week-old *A. annua* leaves is higher in LAP than HAP chemotypes (Maes *et al.*, 2011). Therefore, differences in protein activity rather than differences in gene expression level may account for differences in chemotype. In the present work, two DBR2 variants were identified (Fig. 3) but characterization of the *in planta* activity showed they have equal activity (Fig. 5b, S3b). No difference was found for the ALDH1 amino acid sequence from *A. annua* HAP and LAP chemotypes. Therefore, the only consistent difference in artemisinin biosynthesis proteins in the branched pathway after ADS is in the AMOHAP and AMOLAP from the *A. annua* HAP and LAP chemotype, respectively (Fig. 2).

In the LAP chemotype CYP71AV1 consistently is seven amino acids longer than CYP71AV1 from the HAP chemotype (Fig. 2). When expressed together with ADS (+FPS+HMGR) the total product yield (based on the cumulative levels of AAOH, AAA, AA, DHAAOH and DHAA released from conjugated products) in leaves co-infiltrated with AMOLAP was approximately twice as high as in leaves co-infiltrated with AMOHAP (Table 3). This suggests a lower enzyme activity for AMOHAP than for AMOLAP (Fig. 5a). AMOHAP and AMOLAP seem to anchor equally well to the ER membrane (Fig. 4). The observed difference in efficiency may be caused by a different stability of the two proteins. Nevertheless, dilution of AMOLAP gene dosage in agro-infiltration experiments did not fully mimic the AMOHAP phenotype (Figs. 6a,
S3c). In theory, the lower efficiency of AMOHAP could also be due to a less efficient interaction with the cytochrome P450 reductase (CPR) (Lengler et al., 2006). Alignment of the AMOLAP and AMOHAP protein sequence with other related sesquiterpene oxidases like germacrene A oxidase (GAO) from several different Asteraceae (Nguyen et al., 2010) shows that the N-terminal extension in AMOLAP is the exception. However, the CPR interaction domain does not map to the N-terminus (Sevrioukova et al., 1999). Finally, it may be that the substrate entry or release of AMOHAP is compromised.

Variable ‘chemotype’ of leaves expressing artemisinin biosynthesis genes

We defined the ‘chemotype’ of the *N. benthamiana* agro-infiltrated leaves based on the DHAA and AA glycoside conjugates (Fig. 5). We note that the peak intensity is a relative quantification as detection efficiency (e.g. ionisation) may differ between compounds. However, comparison of the relative quantifications based on peak intensities from LC-QTOF-MS and the absolute quantification of DHAA and AA products by UPLC-MRM-MS or GC-MS (Table 1-3) indicates a good correlation between the two analytical techniques. Expression of *ADS+AMOHAP* results in a HAP chemotype (more DHAA than AA) and expression of *ADS+AMOLAP* in a LAP chemotype (more AA than DHAA) (Fig. 5a). However, when *DBR2* is included, the chemotype for both the combination *ADS+AMOHAP+DBR2* and *ADS+AMOLAP+DBR2* is changed to a HAP chemotype with relatively higher DHAA level (Fig. 5b). In all these combinations, the overall yield of DHAA was always lower for AMOHAP (c. 15-fold). When *ALDH1* was included, the chemotype remained that of HAP, but the relative yield of the gene combinations that include AMOHAP was substantially increased and was now comparable to that of the gene combinations with AMOLAP (Fig. 5c). Because in the agro-infiltration assay in *N. benthamiana* addition of new genes to the pathway leads to a reduction in the relative dosage of the other genes infiltrated into the leaf, we tested whether the relative gene dosage affects the product profile. Lowering the dosage of AMOLAP with similar dosage of ADS and DBR2 resulted in a profile more closely related to that of AMOHAP, but also resulted in lower product yield of AA and DHAA conjugates (Fig. 6a). Lowering the relative dosage of DBR2 resulted in a reversion of the infiltrated leaf chemotype from HAP to LAP (Fig. 6b), while lowering the relative dosage of ALDH1 did not change the chemotype, but did decrease AA conjugates more than the DHAA conjugates (Fig. 6c). When both DBR2 and ALDH1 dosage were reduced, AMOHAP resulted in a more HAP related chemotype, while the com-
bination with AMOLAP resulted in a more LAP related chemotype (Fig. 6d).

*N. benthamiana* enzyme activities both enhance and limit the DHAA ‘chemotype’

AMOHAP seems to be less efficient in the conversion to AA, presumably resulting in an early release of AAA. Indeed free AAA and an AAA glutathione conjugate were present at higher levels when the pathway was expressed in combination with AMOHAP than with AMOLAP (Table 1, Fig. 7). The AAA released by, particularly, AMOHAP may subsequently be substrate for double-bond reductases (endogenous from *N. benthamiana* or the co-expressed DBR2) to produce DHAAA and DHAAA derived products (Table 1, 2 and 3). The endogenous DBR2-like activity is far from saturating, as introduction of *A. annua* DBR2 greatly enhanced the conversion to DHAAOOH, DHAAA and DHAA (Table 1). The early release of AAA by AMOHAP also reveals the activity of an endogenous *N. benthamiana* reductase, similar to the *A. annua* aldehyde reductase (RED1) that catalyses the conversion of DHAAA to DHAAOOH (Bertea et al., 2005, Rydén et al., 2010). Indeed, the presence of a RED1-like activity in tobacco was previously demonstrated through feeding experiments: tobacco leaves supplied with DHAAA and AAA form DHAAOOH and AAOH, respectively (Zhang et al., 2011). This suggests that in *N. benthamiana* the elevated pools of AAOH and DHAAOOH (and glycosylated derivatives) may be the result of a reverse product flux from AAA back to AAOH and DHAAA to DHAAOOH (Fig. 1). If this is the case the affinity of AMOHAP for the AAOH substrate may be underestimated.

Table 4. Artemisinic acid (AA) and dihydroartemisinic acid (DHAA) produced in agro-infiltrated *Nicotiana benthamiana* as identified and quantified by UPLC-MRM-MS.

<table>
<thead>
<tr>
<th>(ng g⁻¹ FW)</th>
<th>AmFH+AMOLAP+DBR2+ALDH1</th>
<th>AmFH+AMOHAP+DBR2+ALDH1</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>8836 ± 1730</td>
<td>2589 ± 563</td>
</tr>
<tr>
<td>DHAA</td>
<td>10792 ± 341</td>
<td>2756 ± 547</td>
</tr>
</tbody>
</table>

Results are means ± S.D. of three co-infiltrated leaves.

DHAA (conjugates) were detected in the leaves infiltrated with *AmFH+AMOLAP*. This suggests the presence of an endogenous aldehyde dehydrogenase, similar to ALDH1 from *A. annua*, which can produce DHAA from DHAAA (Teoh et al., 2009). Alternatively, the low levels of DHAA could be the result of AMOLAP catalysing oxidation of DHAAA. Work in yeast showed that AMOLAP is far less effective in the conversion of DHAAA to DHAA than in the conversion
of AAA to AA (Teoh et al., 2009). If we assume that also AMOHAP catalyses this step less effectively, expression of ALDH1 in a HAP background should further enhance the production of DHAA in the HAP chemotype. Indeed, our data show that product flow towards DHAA increased c. 110-fold when expression of AmFH+AMOHAP+DBR2 was combined with ALDH1 (Table 4).

Figure 8. Schematic representation of the artemisinin precursor biosynthetic pathway in Nicotiana benthamiana. NbDBR2-like: Nicotiana benthamiana artemisinic aldehyde double-bond reductase-like enzyme; NbRED1-like: Nicotiana benthamiana dihydroartemisinic aldehyde reductase 1-like enzyme; NbALDH1-like: Nicotiana benthamiana aldehyde dehydrogenase 1-like enzyme; NbGTase: Nicotiana benthamiana glycosyl transferase; NbMTase: Nicotiana benthamiana malonyl transferase; NbGSTase: Nicotiana benthamiana glutathione-S-transferase; Enzyme names in round brackets indicates low activity of the enzyme for the corresponding step.
Conjugating activities limit precursor pool for artemisinin production

Most of the products formed upon agro-infiltration of the artemisinin pathway genes are present in the form of hexose and/or hexose/malonyl conjugates. This explains why a previous heterologous expression study in tobacco detected only limited levels of free AD, AAOH, DHAAOH and no AA or DHAA (Zhang et al., 2011). The deglycosylation experiments show that the conjugated forms such as AA and DHAA accumulate to levels up to 8 to 10-fold higher than the corresponding free forms (Tables 1-3). The glycosylation/conjugation is most likely a response to the production of potentially toxic compounds, as leaves expressing AmFH+AMOLAP/HAP showed signs of necrosis (Fig. S2). Necrosis was stronger in leaves containing higher levels of free (and conjugated) AA and indeed presence of DBR2 reduced both the necrotic phenotype and the free and conjugated AA levels (Fig. S2, Table 1). In addition to the AA/DHAA glycosides, AAA glutathione conjugates were detected. All these conjugating activities limit accumulation of DHAA, the direct precursor of artemisinin (Fig. 8). Interestingly, in extracts from A. annua HAP flowers, no glycosides of AAOH, AA, DHAAOH and DHAA or glutathione conjugates of AAA were detected (Fig. S8). This indicates that the cells of A. annua that produce artemisinin either do not have competing glycosyl/malonyl/GSH transferase activity or, perhaps more likely, that product flux in these specialised A. annua cells is protected from such competing activities.

Concluding remarks

Our results show that the chemical profile of agro-infiltrated N. benthamiana leaves is a function of both type and relative dosage of the expression constructs. Results still do not fully explain the difference in HAP and LAP chemotypes found in A. annua. The expression level of the different biosynthesis genes do not differ between A. annua chemotypes (Maes et al., 2011) and therefore, at present, the difference in CYP71AV1 catalytic efficiency between HAP and LAP is the only identified factor that contributes to this difference in chemotype. However, other, as yet unidentified, factors in A. annua may further contribute to the chemical difference in the HAP and LAP varieties.
Chapter 2

Acknowledgements

H-M.T. was funded by the graduate school of Experimental Plant Sciences (EPS). T.v.H. and J.B. were supported through the NWO-CW/ACTS IBOS programme (053.63.305) which was co-sponsored by Dafra Pharma, Turnhout, Belgium. We thank Ric de Vos for helpful discussions on LC-QTOF-MS/MS data, Bert Schipper for assistance in LC-QTOF-MS analysis, Desalegn Woldes Etalo, Ting Yang and Lemeng Dong for help in MS data analysis and Qing Liu for support in the glutathione conjugation assay. We also would like to acknowledge Sajad Rashidi Manfard and Peter E. Brodelius (Linneaeus University, Sweden) for providing seeds of an A. annua LAP chemotype.

Reference


Academy of Sciences 96: 1863-1868.


Supporting Information

Methods S1. Details of fractionation and UPLC-MRM-MS analysis.

Chromatographic separation was obtained on an Acquity UPLC BEH C18 column (100 × 2.1 mm, 1.7 μm; Waters) by applying a water/acetonitrile gradient to the column, starting from 5% (v/v) acetonitrile in water for 1.25 min and rising to 50% (v/v) acetonitrile in water in 2.35 min, followed by an increase to 90% (v/v) acetonitrile in water in 3.65 min, which was maintained for 0.75 min before returning to 5% acetonitrile in water using a 0.15 min gradient. Finally, the column was equilibrated for 1.85 min using this solvent composition. Column temperature and flow rate were 50°C and 0.5mL min⁻¹, respectively. Injection volume was 5 µl. The mass spectrometer was operated in positive electrospray ionization mode. Cone and desolvation gas flows were set to 50 and 1000 L h⁻¹, respectively. The capillary voltage was set at 3.0 kV, the source temperature at 150°C, and the desolvation temperature at 650°C. The cone voltage was optimized for individual artemisinin precursors using the Waters IntelliStart MS Console. Argon was used for fragmentation by collision-induced dissociation in the ScanWave collision cell. Multiple Reaction Monitoring (MRM) was used for identification and quantification by comparing retention times and MRM mass transitions with that of authentic AAOH, AAA, AA, DHAAOH, DHAAA, DHAA and artemisinin standards. MRM transitions for artemisinin precursors were optimized using the Waters IntelliStart MS Console as follows:

<table>
<thead>
<tr>
<th></th>
<th>Transitions (m/z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAOH</td>
<td>221.16&gt;203.27</td>
</tr>
<tr>
<td>AAA</td>
<td>219.16&gt;145.08</td>
</tr>
<tr>
<td>AA</td>
<td>235.16&gt;189.22</td>
</tr>
<tr>
<td>DHAAOH</td>
<td>223.22&gt;205.27</td>
</tr>
<tr>
<td>DHAA</td>
<td>221.16&gt;105.14</td>
</tr>
<tr>
<td>DHAAA</td>
<td>237.16&gt;163.17</td>
</tr>
<tr>
<td>Artemisin</td>
<td>283.19&gt;219.21</td>
</tr>
</tbody>
</table>

Table S1. Primers used in this study.

<table>
<thead>
<tr>
<th>No. primer</th>
<th>Sequence 5’ to 3**</th>
<th>Restriction site</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5’CGGCGGATCCATGGCACTCTCACTGACCACCTTCCA 3’</td>
<td>Bam HI</td>
</tr>
<tr>
<td>2</td>
<td>5’CGGCGGTACCCTAGAAAACCTGGAACGAGTAACACT-CAGGCTTTTC 3’</td>
<td>Kpn I</td>
</tr>
</tbody>
</table>
Table S2. Unconjugated artemisinin precursors as identified and quantified by UPLC-MRM-MS in *Artemisia anuua*.

<table>
<thead>
<tr>
<th>(µg mg-1 FW)</th>
<th>A. annua-HAP</th>
<th>A. annua -LAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>0.16</td>
<td>0.35</td>
</tr>
<tr>
<td>DHAA</td>
<td>8.14</td>
<td>0.55</td>
</tr>
<tr>
<td>Artemisinin</td>
<td>0.95</td>
<td>0.04</td>
</tr>
</tbody>
</table>


* Restriction sites are shown in italics.
The Metabolites Chemotype in *N. benthamiana*

Figure S1. Alignment of the 5’RACE sequences of *AMOHAP* in Clustal W. Seven 5’RACE clones from a HAP chemotype are aligned with *AMOLAP* (DQ268763), *AMOHAP.1* (DQ315671) and the recently published *CYP71AV1* promoter isolated from an *Artemisia anua* HAP chemotype, here referred to as *AMOHAP.5* (Wang et al., 2011). The start codons of *AMOLAP* and *AMOHAP* are indicated by red lines. Note that *AMOLAP* has an ATG at the same position as the start codon of *AMOHAP* but is preceded by another in-frame ATG.
Chapter 2

Figure S2. Necrotic symptoms in *Nicotiana benthamiana*. Leaves agro-infiltrated with (a) pBin empty vector, (b) AmFH+AMOLAP, (c) AmFH+AMOLAP+DBR2, (d) AmFH+AMOHAP, and (e) AmFH+AMOHAP+DBR2 seven days after infiltration.

Figure S3. All conjugated compounds grouped according to free precursor detected in agro-infiltrated *Nicotiana benthamiana*. (a) Agro-infiltrated leaves with AmFH+AMOLAP/HAP and AmFH+AMOLAP/HAP+DBR2. For all products identified by LC-QTOF-MS see Table 2. (b) Agro-infiltrated leaves with AmFH+AMOLAP/HAP+DBR2LAP/HAP+ALDH1. For all products identified by LC-QTOF-MS see Table S3. (c) Agro-infiltrated leaves with dilution of AMOLAP/HAP. For all products identified by LC-QTOF-MS see Table S4. (d) Agro-infiltrated leaves with dilution of DBR2 and ALDH1. For all products identified by LC-QTOF-MS see Table S5. Data shown with peak intensity which analysed by LC-QTOF-MS. Peak intensities are the mean of three agro-infiltrated leaves. (*: different conjugated compounds grouped according to free precursor)
Figure S4. Malonylated glycosylated compounds are not cleaved by Viscozyme L. (a) LC-QTOF-MS chromatogram (total ion count) of methanol extract of *Nicotiana benthamiana* leaves infiltrated with *AmFH+AMOLAP+DBR2* without Viscozyme-treatment. (b) LC-QTOF-MS chromatogram (total ion count) of methanol extract of *Nicotiana benthamiana* leaves infiltrated with *AmFH+AMOLAP+DBR2* after Viscozyme-treatment. Peak 1: DHAA-Hex2, Peak 2: DHAA-Hex2-Mal, Peak 3: DHAAOH-Hex2, Peak 4: AA-Hex2-Mal, Peak 5: DHAAOH-Hex2-Mal. (DHAA: dihydroartemisinic acid; DHAAOH: dihydroartemisinic alcohol; AA: artemisinic acid; Hex: compound conjugated with hexose; Mal: compound conjugated with malonate)
Figure S5. MS/MS spectrum and MS/MS fragmentation of DHAA-Hex3. (a) MS/MS spectrum of the compound eluting at 23.87 min, DHAA-Hex3 (m/z 721.32; [M-H]). (b) MS/MS fragmentation of mass 721.32 eluting at 23.87 min. (DHAA: dihydroartemisinic acid, Hex: compound conjugated with hexose)

Figure S6. Comparison of in vitro and in planta formed AAA-GSH conjugates. (a) Mass spectrum (positive mode) of AAA-GSH*, as formed in planta (Table 2). The mass suggests a AAA-GSH derivative with 18 D mass increase, suggesting an additional OH and H. This is most easily explained by a reduction of the endocyclic double bond in AAA and hydroxylation. (b) Mass spectrum (positive mode) of AAA-GSH, as formed in vitro (see Fig. S6c). (AAA: artemisinic aldehyde, GSH: glutathione)
Figure S7. *In vitro* conjugation of artemisinic aldehyde (AAA) with glutathione. LC-QTOF-MS analysis of (a) Glutathione (GSH) incubated with Glutathione S transferase (GST), (b) AAA+GST, (c) AAA+GSH and (d) AAA+GSH+GST. Peak 1: AAA, Peak 2: AAA-GSH conjugate.

Figure S8. Absence of artemisinin precursor conjugates in *Artemisia annua* HAP chemotype. (a) LC-QTOF-MS chromatogram (total ion count) of methanol extract of flowers of *Artemisia annua* HAP chemotype. (b) LC-QTOF-MS chromatogram (total ion count) of methanol extract of *Nicotiana benthamiana* leaves infiltrated with AmFH+AMO-LAP+DBR2. Peak 1: DHAA-Hex3, Peak 2: DHAAOH-Hex3, Peak 3: DHAA-Hex2, Peak 4: DHAAOH-Hex2, Peak 5: AA-Hex2-Mal, Peak 6: DHAAOH-Hex2-Mal. (DHAA: dihydroartemisinic acid; DHAAOH: dihydroartemisinic alcohol; AA: artemisinic acid; Hex: compound conjugated with hexose; Mal: compound conjugated with malonate)
Table S3 Peak intensity of conjugated artemisinin precursors produced in agro-infiltrated *Nicotiana benthamiana* leaves with AmF-H+AMOLAP/HAP+DBR2LAP/HAP and AmFH+AMOLAP/HAP+DBR2LAP/HAP+ALDH1 were analysed by LC-QTOF-MS.

<table>
<thead>
<tr>
<th>Putative ID ++</th>
<th>Intensity *</th>
<th>AmFH+AMOLAP+DBR2HAP</th>
<th>AmFH+AMOLAP+DBR2LAP</th>
<th>AmFH+AMOHAP+DBR2HAP</th>
<th>AmFH+AMOHAP+DBR2LAP</th>
<th>AmFH+AMOLAP+DBR2HAP+ALDH1</th>
<th>AmFH+AMOLAP+DBR2LAP+ALDH1</th>
<th>AmFH+AMOHAP+DBR2HAP+ALDH1</th>
<th>AmFH+AMOHAP+DBR2LAP+ALDH1</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAOH-Hex2</td>
<td>229</td>
<td>220</td>
<td>1383</td>
<td>1312</td>
<td>158</td>
<td>156</td>
<td>1448</td>
<td>1545</td>
<td></td>
</tr>
<tr>
<td>AAOH-Hex2-Mal (I) a</td>
<td>1042</td>
<td>933</td>
<td>1350</td>
<td>1339</td>
<td>1029</td>
<td>997</td>
<td>1737</td>
<td>1716</td>
<td></td>
</tr>
<tr>
<td>AAOH-Hex2-Mal (II)</td>
<td>579</td>
<td>525</td>
<td>13994</td>
<td>14298</td>
<td>157</td>
<td>158</td>
<td>4589</td>
<td>4801</td>
<td></td>
</tr>
<tr>
<td>AAOH-Hex2-Mal (III)</td>
<td>ND</td>
<td>ND</td>
<td>339</td>
<td>313</td>
<td>2</td>
<td>ND</td>
<td>682</td>
<td>718</td>
<td></td>
</tr>
<tr>
<td>AAA-GSH-H$_2$O b</td>
<td>68</td>
<td>67</td>
<td>1168</td>
<td>968</td>
<td>98</td>
<td>98</td>
<td>1055</td>
<td>1135</td>
<td></td>
</tr>
<tr>
<td>(AA-Hex3) FA c</td>
<td>10</td>
<td>10</td>
<td>ND</td>
<td>ND</td>
<td>4923</td>
<td>5101</td>
<td>3318</td>
<td>3231</td>
<td></td>
</tr>
<tr>
<td>AA-Hex2 (I) d</td>
<td>77</td>
<td>62</td>
<td>48</td>
<td>50</td>
<td>4067</td>
<td>3969</td>
<td>2285</td>
<td>2441</td>
<td></td>
</tr>
<tr>
<td>AA-Hex2 (II)</td>
<td>ND</td>
<td>ND</td>
<td>3</td>
<td>3</td>
<td>6212</td>
<td>6210</td>
<td>3157</td>
<td>3320</td>
<td></td>
</tr>
<tr>
<td>AA-Hex2-Mal (I)</td>
<td>710</td>
<td>729</td>
<td>355</td>
<td>353</td>
<td>5969</td>
<td>5912</td>
<td>5008</td>
<td>5525</td>
<td></td>
</tr>
<tr>
<td>AA-Hex2-Mal (II)</td>
<td>32</td>
<td>29</td>
<td>21</td>
<td>22</td>
<td>3249</td>
<td>3703</td>
<td>1214</td>
<td>1150</td>
<td></td>
</tr>
<tr>
<td>(DHAOOH-Hex3) FA</td>
<td>8994</td>
<td>9033</td>
<td>6724</td>
<td>6254</td>
<td>205</td>
<td>188</td>
<td>302</td>
<td>340</td>
<td></td>
</tr>
<tr>
<td>DHAOOH-Hex2 (I)</td>
<td>12703</td>
<td>12612</td>
<td>10560</td>
<td>10243</td>
<td>261</td>
<td>249</td>
<td>330</td>
<td>346</td>
<td></td>
</tr>
<tr>
<td>DHAOOH-Hex2-Mal (I)</td>
<td>24565</td>
<td>24544</td>
<td>24521</td>
<td>24534</td>
<td>1374</td>
<td>1039</td>
<td>1637</td>
<td>1598</td>
<td></td>
</tr>
<tr>
<td>DHAOOH-Hex2-Mal (II)</td>
<td>1761</td>
<td>1749</td>
<td>1485</td>
<td>1529</td>
<td>215</td>
<td>238</td>
<td>100</td>
<td>117</td>
<td></td>
</tr>
<tr>
<td>(DHA-A-Hex3) FA</td>
<td>13317</td>
<td>12376</td>
<td>366</td>
<td>364</td>
<td>17940</td>
<td>17983</td>
<td>14736</td>
<td>14812</td>
<td></td>
</tr>
<tr>
<td>DHA-A-Hex2</td>
<td>3160</td>
<td>2876</td>
<td>23</td>
<td>18</td>
<td>7441</td>
<td>7599</td>
<td>4506</td>
<td>4660</td>
<td></td>
</tr>
<tr>
<td>DHA-A-Hex2-Mal (I)</td>
<td>100</td>
<td>102</td>
<td>890</td>
<td>867</td>
<td>284</td>
<td>268</td>
<td>2133</td>
<td>2453</td>
<td></td>
</tr>
<tr>
<td>DHA-A-Hex2-Mal (II)</td>
<td>5123</td>
<td>4800</td>
<td>113</td>
<td>95</td>
<td>8563</td>
<td>7343</td>
<td>7025</td>
<td>6975</td>
<td></td>
</tr>
</tbody>
</table>

* Peak intensities are the mean of three agro-infiltrated leaves. **AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAOOH: dihydroartemisinic alcohol; DHAAO: dihydroartemisinic aldehyde; DHAO: dihydroartemisinic acid. ND: not detectable. *Hex: compound conjugated with hexose; Mal: compound conjugated with malonate; (I~III): different isobaric forms (i.e. identical accurate mass, but different retention times). gGSH: glutathione. (FA): formic acid adduct. dAA-Hex2: The structure of artemisinic acid-12-β-diglucoside was confirmed by NMR (van Herpen et al., 2010).
Table S4 Peak intensity of conjugated artemisinin precursors produced in agro-infiltrated *Nicotiana benthamiana* leaves with AmF-H+AMOLAP/HAP+DBR2 compared with different dilution of AMOLAP/HAP were analysed by LC-QTOF-MS.

<table>
<thead>
<tr>
<th>Putative ID**</th>
<th>Intensity*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AmFH</td>
</tr>
<tr>
<td></td>
<td>+AMOLAP+DBR2</td>
</tr>
<tr>
<td>AAOH-Hex2</td>
<td>220</td>
</tr>
<tr>
<td>AAOH-Hex2-Mal (I)(^a)</td>
<td>1057</td>
</tr>
<tr>
<td>AAOH-Hex2-Mal (II)</td>
<td>ND</td>
</tr>
<tr>
<td>AAOH-Hex2-Mal (III)</td>
<td>717</td>
</tr>
<tr>
<td>AAA-GSH-H(_2)O(^b)</td>
<td>64</td>
</tr>
<tr>
<td>(AA-Hex3) FA(^c)</td>
<td>39</td>
</tr>
<tr>
<td>AA-Hex2 (I)(^d)</td>
<td>60</td>
</tr>
<tr>
<td>AA-Hex2 (II)</td>
<td>ND</td>
</tr>
<tr>
<td>AA-Hex2-Mal (I)</td>
<td>681</td>
</tr>
<tr>
<td>AA-Hex2-Mal (II)(_{[2M-H]})</td>
<td>390</td>
</tr>
<tr>
<td>(DHAOOH-Hex3) FA</td>
<td>11759</td>
</tr>
<tr>
<td>DHAOOH-Hex2 (I)</td>
<td>12104</td>
</tr>
<tr>
<td>DHAOOH-Hex2-Mal (I)</td>
<td>24571</td>
</tr>
<tr>
<td>DHAOOH-Hex2-Mal (II)</td>
<td>1640</td>
</tr>
<tr>
<td>(DHA-Hex3) FA</td>
<td>12449</td>
</tr>
<tr>
<td>DHA-Hex2</td>
<td>2632</td>
</tr>
<tr>
<td>DHA-Hex2-Mal (I)</td>
<td>122</td>
</tr>
<tr>
<td>DHA-Hex2-Mal (II)</td>
<td>3664</td>
</tr>
</tbody>
</table>

* Peak intensities are the mean of three agro-infiltrated leaves. **AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAOOH: dihydroartemisinic alcohol; DHAOOA: dihydroartemisinic aldehyde; DHA: dihydroartemisinic acid. ND: not detectable.

\(^a\)Hex: compound conjugated with hexose; Mal: compound conjugated with malonate; (I–III): different isobaric forms (i.e. identical accurate mass, but different retention times).

\(^b\)GSH: glutathione. \(^{\text{c(FA)}}\): formic acid adduct. \(^{\text{d}AA\text{-Hex2}}\): The structure of artemisinic acid-12-β-diglucoside was confirmed by NMR (van Herpen et al., 2010).
Table S5 Peak intensity of conjugated artemisinin precursors produced in agro-infiltrated *Nicotiana benthamiana* leaves with AmF-H+AMOLAP/HAP+DBR2+ALDH1 with different dosage of DBR2 and ALDH1 were analysed by LC-QTOF-MS.

<table>
<thead>
<tr>
<th>Putative ID**</th>
<th>Intensity*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AmFH+AMOLAP +DBR2</td>
</tr>
<tr>
<td>AAOH-Hex2</td>
<td>222</td>
</tr>
<tr>
<td>AAOH-Hex2-Mal (I) a</td>
<td>935</td>
</tr>
<tr>
<td>AAOH-Hex2-Mal (II)</td>
<td>568</td>
</tr>
<tr>
<td>AAOH-Hex2-Mal (III)</td>
<td>ND</td>
</tr>
<tr>
<td>AAA-GSH-H₂O c</td>
<td>180</td>
</tr>
<tr>
<td>(AA-Hex3) FA</td>
<td>15298</td>
</tr>
<tr>
<td>AA-Hex2 (I)</td>
<td>8126</td>
</tr>
<tr>
<td>AA-Hex2 (II)</td>
<td>1631</td>
</tr>
<tr>
<td>AA-Hex2-Mal (I)</td>
<td>23925</td>
</tr>
<tr>
<td>AA-Hex2-Mal (II)</td>
<td>2061</td>
</tr>
<tr>
<td>(DHAOOH-Hex3) FA</td>
<td>1247</td>
</tr>
<tr>
<td>DHAOOH-Hex2 (I)</td>
<td>49/</td>
</tr>
<tr>
<td>DHAOOH-Hex2-Mal (I)</td>
<td>2288</td>
</tr>
<tr>
<td>DHAOOH-Hex2-Mal (II)</td>
<td>285</td>
</tr>
<tr>
<td>(DHA-Aex3) FA</td>
<td>24592</td>
</tr>
<tr>
<td>DHAA-Hex2</td>
<td>15312</td>
</tr>
<tr>
<td>DHAA-Hex2-Mal (I)</td>
<td>512</td>
</tr>
<tr>
<td>DHAA-Hex2-Mal (II)</td>
<td>22905</td>
</tr>
</tbody>
</table>

* Peak intensities are the mean of three agro-infiltrated leaves. **AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; DHAAOH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid. ND: not detectable. 

aHex: compound conjugated with hexose; Mal: compound conjugated with malonate; (I~III): different isobaric forms (i.e. identical accurate mass, but different retention times). bGSH: glutathione. (FA): formic acid adduct. cAA-Hex2: The structure of artemisinic acid-12-β-diglucoside was confirmed by NMR (van Herpen et al., 2010).
Table S6. Metabolite profiling of *Artemisia annua* plants. (Data retrieved from Maes *et al.*, 2011).

<table>
<thead>
<tr>
<th>Metabolite</th>
<th><em>A. annua</em>-HAP (5-wk)</th>
<th><em>A. annua</em>-LAP (5-wk)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAOH*</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>AAA</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>AA</td>
<td>9</td>
<td>1050</td>
</tr>
<tr>
<td>AB</td>
<td>11</td>
<td>320</td>
</tr>
<tr>
<td>DHAOAH</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>DHAA</td>
<td>188</td>
<td>35</td>
</tr>
<tr>
<td>DHAA</td>
<td>450</td>
<td>55</td>
</tr>
<tr>
<td>AN</td>
<td>600</td>
<td>80</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1293</strong></td>
<td><strong>1562</strong></td>
</tr>
</tbody>
</table>

AAOH: artemisinic alcohol; AAA: artemisinic aldehyde; AA: artemisinic acid; AB: arteannuin B; DHAOAH: dihydroartemisinic alcohol; DHAAA: dihydroartemisinic aldehyde; DHAA: dihydroartemisinic acid; AN: artemisinin.
Table 2: Conjugated artemisinin precursors produced in agro-infiltrated *Nicotiana benthamiana* leaf extracts. Non-volatile metabolites with mass intensity higher than 500 in LC-QTOF-MS, which were significantly increased in leaves agro-infiltrated with AmFH+AMOLAP, AmFH+AMOHAP, AmFH+AMOLAP+DBR2, or AmFH+AMOHAP+DBR2 were targeted for analysis by LC-QTOF-MS/MS fragmentation.

<table>
<thead>
<tr>
<th>Ret (min)</th>
<th>Detected Mass(D)</th>
<th>MS-MS fragments</th>
<th>Mol Form</th>
<th>ΔMass (ppm)</th>
<th>Putative ID</th>
<th>Intensity</th>
<th>AmFH +AMOLAP</th>
<th>AmFH +AMOHAP</th>
<th>AmFH +AMOLAP+DBR2</th>
<th>AmFH +AMOHAP+DBR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>27.85</td>
<td>543.279</td>
<td>381[M+hex-H]</td>
<td>C₆H₆O₁₁</td>
<td>1.3</td>
<td>AAOH-Hex2</td>
<td>319±57</td>
<td>3715±610</td>
<td>264±51</td>
<td>175±3823</td>
<td></td>
</tr>
<tr>
<td>20.36</td>
<td>620.2810</td>
<td>381[M-Mal-H]</td>
<td>C₆H₆O₁₁</td>
<td>0</td>
<td>AAOH-Hex2-Mal (II)</td>
<td>606±37</td>
<td>2425±70</td>
<td>627±100</td>
<td>1021±5097</td>
<td></td>
</tr>
<tr>
<td>23.80</td>
<td>542.2536</td>
<td>272.25410.179.143.128</td>
<td>C₄H₄N₂O₈</td>
<td>1.0</td>
<td>AAA-GSH-H₂O⁺</td>
<td>246±323</td>
<td>789±65</td>
<td>130±22</td>
<td>1970±395</td>
<td></td>
</tr>
<tr>
<td>24.02</td>
<td>765.3181</td>
<td>719[M+H]-399[M-2Hex-H]</td>
<td>C₄H₆O₁₀</td>
<td>0.1</td>
<td>(AA-Hex3) FA</td>
<td>21190</td>
<td>612±35</td>
<td>100±80</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td>24.86</td>
<td>753.3545</td>
<td>707[M-H]-545[M-Hex-H], 393[M-2Hex-H]-221[M-3Hex-H]</td>
<td>C₄H₆O₁₈</td>
<td>0.9</td>
<td>(DAHAO-Hex3) FA</td>
<td>145±46</td>
<td>85±8</td>
<td>124±57</td>
<td>1157±206</td>
<td></td>
</tr>
<tr>
<td>27.28</td>
<td>554.2755</td>
<td>397[M-1.599[M-CO₂-H]-235[M-2Hex-H]-235[M-Hex-H]</td>
<td>C₄H₆O₁₂</td>
<td>0.3</td>
<td>DAAHO-Hex2</td>
<td>ND</td>
<td>5188±2039</td>
<td>120±54</td>
<td>ND</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 3

The A. annua Lipid Transfer Protein AaLTP3 prevents reflux of dihydroartemisinic acid that was transported to the apoplast and increases production of artemisinin in Nicotiana benthamiana

Bo Wang¹,⁴, Arman Beyraghdar Kashkooli¹,⁴, Adrienne Sallets², Hieng-Ming Ting¹, Norbert C.A. de Ruijter³, Marc Boutry², Harro Bouwmeester¹, Alexander R. van der Krol¹

¹Laboratory of Plant Physiology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
²Institut des Sciences de la Vie, Université catholique de Louvain, Croix du Sud, 4-5, Box L7.07.14, 1348 Louvain-la-Neuve, Belgium.
³Laboratory of Cell Biology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
⁴These authors contributed equally to this paper
Abstract

Metabolic engineering in plants for the production of high value terpenes is currently limited by our understanding of transport and sequestration of the heterologous products. For instance, although all biosynthesis genes for the production of artemisinin have been identified, ectopic expression of these genes in *Nicotiana benthamiana* yielded mostly glycosylated pathway intermediates and only very little free (dihydro)artemisinic acid [(DH)AA] or artemisinin. Here we tested three *Lipid Transfer Protein (LTP)* genes from *Artemisia annua* for their role in the sequestration of free artemisinin pathway (AN-PW) products – produced by co-expression of the AN-PW genes in the apoplast of *N. benthamiana* leaves. Ectopically produced non-glycosylated AN-PW products like (DH)AA accumulate in the apoplast, indicative of a low intrinsic transport activity in *N. benthamiana* leaf cells. The amount of apoplastic (DH)AA is slightly enhanced by the transient co-expression of an ABC membrane transporter, *AaPDR2*, while co-expression of three *AaLTPs in combination with AN-PW+AaPDR2*, results in even higher levels of free (DH)AA in the apoplast. Subsequent assays with individual LTPs show that AaLTP3 is most effective for accumulation of (DH)AA in the apoplast. When (DH)AA is infiltrated into the apoplast of *N. benthamiana* leaves it is very rapidly taken up by and converted inside the cells to (DH)AA-glycosides, indicating that active retention in the apoplast is needed. The rapid intracellular glycosylation of infiltrated (DH)AA was reduced by 27% by transient expression of *AaPDR2*, by 74% by expression of *AaLTP3* and by 82% by the combined expression of *AaPDR2* and *AaLTP3*. Free DHAA and AA were converted to artemisinin and arteannuin B, respectively, in necrotic *N. benthamiana* leaves and co-expression of AN-PW+AaLTP3+AaPDR2 resulted in the highest yield of artemisinin and arteannuin B.

Keywords

Lipid transfer protein, ABC transporter, Pleiotropic Drug Resistance protein, artemisinin, *Artemisia annua*, *Nicotiana benthamiana*
Introduction

Artemisinin is a highly oxygenated sesquiterpene containing a unique 1,2,4-tri-oxane ring structure produced by the plant *Artemisa annua*. Artemisinin is used for the treatment of malaria and research of the past decade has led to the identification of all biosynthesis steps that are required for the production of the artemisinin precursor dihydroartemisinic acid (DHAA) (Bouwmeester et al., 1999; Cheng et al., 2004; Covello et al., 2007; Teoh et al., 2009; Teoh et al., 2006; Zhang et al., 2008). The final conversion of DHAA to artemisinin is thought to be the result of non-enzymatic oxidation (Brown, 2010).

DHAA accumulates in the subcuticular space of capitate glandular secretory trichomes on the surface of leaves and flowers of *A. annua* (Duke et al., 1994). These glandular trichomes consist of 10-celled structures with two basal cells, two stalk cells, and three pairs of secretory cells and a subcuticular sac (Duke and Paul, 1993). The subapical cells contain chloroplasts and seem to have a function in the overall metabolism of glandular trichomes (Olsson et al., 2009). The apical cells contain extensive ER membrane structures, typical of secretory cells, and the genes encoding the artemisinin biosynthesis enzymes (*ADS, ADO, DBR2, ALDH1*) are expressed in these apical cells (Olsson et al., 2009; Soetaert et al., 2013). The biosynthesis of the direct precursor of artemisinin (DHAA) takes place inside the apical cells, while artemisinin is formed outside the cells in the subcuticular space, possibly upon rupture of the cuticle of the trichome (Olsson et al., 2009). This implies that DHAA is transported out of the apical secretory cells into the subcuticular space. Studies with *A. annua* cell suspensions which produce artemisinin have shown that most of the produced artemisinin is in the medium (Durante et al., 2001), which could support that DHAA after being transported out of the cell is rapidly converted in the cell wall or medium.

When the artemisinin pathway (AN-PW) genes are transiently expressed in *N. benthamiana* pathway intermediates (artesiminic alcohol (AAOH), artemisinic aldehyde (AAA), artemisinic acid (AA), dihydroartemisinic alcohol (DHAAOH); dihydroartemisinic aldehyde (DHAAA); dihydroartemisinic acid (DHAA)) are produced but they are present in glycosylated form presumably present in the cytosol and/or vacuole (Ting et al., 2013). Only low levels of non-glycosylated, free AN-PW intermediates are detected, suggesting that the capacity of *N. benthamiana* leaf cells to sequester these free compounds away from the glycosyl transferases is very limited. This raises the question whether sequestration of AN-PW intermediates can be stimulated by co-expression of the AN-PW genes with oth-
er genes from *A. annua* which may encode relevant transport proteins.

Here we characterized the activity of *A. annua* LTPs which may be involved in the sequestration of DHAA into the apoplast. LTPs are small, basic proteins which are characterized in *in-vitro* assays by their ability to transfer lipids between membranes. LTPs contain a hydrophobic cavity which may bind lipophilic molecules, but the exact function of LTP proteins *in vivo* has not been fully elucidated. Most LTP proteins contain a signal peptide sequence for secretion to the apoplast and extracellular localization of LTP proteins has been confirmed by LTP-GFP localization studies (Huang et al., 2013; Kader, 1997; Yeats and Rose, 2008). Suggestions for a putative role of LTPs in terpene transport come from the fact that LTP genes show very high transcriptional activity in glandular trichomes that are also known to produce terpenes (peppermint: (Lange et al., 2000), alfalfa: (Aziz et al., 2005), Artemisia annua: (Bertea et al., 2006), hop: (Wang et al., 2008), Salvia fruticosa: (Chatzopoulou et al., 2010), tomato: (Schilmiller et al., 2010) and tobacco: (Harada et al., 2010)). Moreover, recently it was shown that over-expression of NtLTP1 in transgenic Nicotiana tabacum resulted in increased secretion of diterpenes from the glandular trichomes (Choi et al., 2012).

The *AaLTP* genes were used in a transient expression assay together with AN-PW genes and the previously characterized ABC transporters for AN-PW intermediates, *AaPDR1* and *AaPDR2* (Boutry, personal communication), to study the effect on apoplastic accumulation of (DH)AA in leaves of agroinfiltrated *N. benthamiana*. When AN-PW genes are expressed in *N. benthamiana* leaves, only small amounts of free (non-glycosylated) (DH)AA are detected in leaf extracts. Here we show that free AA and DHAA and other non-glycosylated AN intermediates are in the apoplast, suggesting the presence of a low intrinsic transport activity for these compounds in *N. benthamiana*. Co-expression of the AN-PW genes with either the two *AaPDR* genes or the three *AaLTP* genes had only a small effect on the total level of free AN-PW intermediates in the apolplast. However, co-expression of the AN-PW genes with both the *PDR* and *LTP* genes resulted in a significant increase in free AN-PW intermediates in the apoplast. Using the transient expression system we demonstrate different specificity in transport activity for the two *PDR* genes and we show specificity in LTP function, with co-expression of *PDR2+LTP3* being the most effective combination to accumulate free AN-PW intermediates into the apoplast. Combined, the results reveal new opportunities to enhance production of valuable medicinal terpenoids in heterologous production systems.
Materials and methods

Transient expression in *N. benthamiana* leaves

*N. benthamiana* was grown in soil in 7 cm pots in greenhouse kept at 22 ºC. The agro-infiltration was in 4th leaves of 5 week old *N. benthamiana* plants. Agroinfiltration of *N. benthamiana* leaves was done as described (van Herpen et al., 2010). The final OD_{600} value was adjusted to 0.5 with the agroinfiltration buffer, including 10 mM MES (2-(N-morpholino) ethanesulfonic acid, Duchefa Biochemie), 10 mM MgCl₂ and 100 µM acetosyringone (4’-hydroxy-3’,5’-dimethoxyacetophenone, Sigma). then individual *A. tumefaciens* strains transformed with different expression constructs were co-infiltrated into *N. benthamiana* leaves using a 1 ml syringe without needle. PW means AN-PW genes including V5, ADO, DBR2, ALDH1 (Ting et al., 2013; van Herpen et al., 2010). Since HMGR is a key rat-limiting step in the sesquiterpene biosynthesis pathway, we add the HMGR to boost the AN production in *N. benthamiana*. PW* means AN-PW genes with an extra HMGR in the agroinfiltration. The relative dosage of each *A. tumefaciens* strain between treatments within the same experiment was kept the same by diluting with an *A. tumefaciens* strain carrying an empty vector construct (EV). *A. tumefaciens* harboring TBSV p19 (P19) was included to maximize the protein production by suppression of gene silencing (Voinnet et al., 2003). Infiltrated leaves were harvested at the days indicated (varying from 3-18 days post agroinfiltration (dpi)).

*A. annua* RT-PCR gene expression analysis

The high-artemisinin production (HAP) chemotype *A. annua* were grown in the greenhouse chamber with 12 h light at 20ºC and 12 h darkness at 18 ºC. The shoot apex and leaves from top (Leaf1) to bottom (Leaf10) were collected from 2 month old *A. annua* plants and pooled per leaf number for RNA extraction. From each pooled and frozen grinded sample 100 mg was extracted for gene expression analysis. RNA was isolated from the frozen grinded samples and DNase I (Invitrogen) was used to remove remaining genomic DNA. Reverse transcription was carried out with the SuperScript First-Strand Synthesis System for the cDNA synthesis (Invitrogen). Real-time quantitative PCR was performed using the iQ5 RT-PCR (BioRad). Primers used are listed in Table S1. The *A. annua* β-actin gene was used for the normalization (Lu et al., 2013; Olofsson et al., 2011). ΔΔCT was calculated as follows: ΔCT=CT(Target)-CT(Actin), ΔΔCT was normalized using ΔCT. The fold change value was calculated by 2^{-ΔΔCT}. 
Cloning of the different expression vectors

Cloning of the LTP overexpression genes

The full-length *AaLTP1*, *AaLTP2* and *AaLTP3* were amplified from cDNA made from RNA isolated from *A. annua* flowers by RACE PCR (Clontech, USA) (Bertea et al., 2006). The primer pairs *AaLTP1*-F/*AaLTP1*-R, *AaLTP2*-F/*AaLTP2*-R and *AaLTP3*-F/*AaLTP3*-R (Table S1) introduce a Ncol site at the start codon and a NotI restriction site after the stop codon of each LTP coding sequence. The PCR fragments were cloned into pIV1A2.1 vector (www.impactvector.com) between the CaMV35S promoter and RBCS1 terminator fragment. pIV1A2.1/AaLTP1, pIV1A2.1/AaLTP2, pIV1A2.1/AaLTP3 were cloned into binary vector pBinPlus by LR reaction (Invitrogen) and the resulting pBIN/AaLTP1, pBIN/AaLTP2, and pBIN/AaLTP3 were transformed into the *A. tumefaciens* AGL-O which was used for the agroinfiltration experiments.

Cloning of the LTP-GFP and LTP-RFP fusion protein genes

For the LTP-GFP fusion protein studies we first made an entry vector with EGFP. The EGFP coding sequence was amplified by PCR using primers EGFP_c-term-F/EGFP_c-term-R (Table S1) by using pBinEGFP as template (http://www.wageningenur.nl/en/show/Productie-van-farmaceutische-en-industriële-eiwitten-door-planten.htm). After digesting with NotI and SacI, GFP was cloned into ImpactVector pIV1A_2.1 (www.impactvector.com) under the control of CaMV 35S promoter to generate an entry vector pIV1A_2.1/EGFP. To remove the stop codon of AaLTP1, AaLTP2 and AaLTP3, DNA fragments were amplified using the primer pairs AaLTP1_gfp-F/AaLTP1_gfp-R, AaLTP2_gfp-F/AaLTP2_gfp-R and AaLTP3_gfpF/AaLTP3_gfp-R (Table S1) by using the pIV1A_2.1/AaLTP1, pIV1A_2.1/AaLTP2 and pIV1A_2.1/AaLTP3 vector as template, respectively. The PCR products were digested with BamHI and NotI and subcloned into vector pIV1A_2.1/EGFP. The resulting pIV1A_2.1/AaLTP1-GFP, pIV1A_2.1/AaLTP2-GFP and pIV1A_2.1/AaLTP3-GFP containing the CaMV 35S promoter were cloned into pBinPlus binary vector (Invitrogen) using LR recombination (van Engelen et al., 1995). The pBinPlus constructs containing the 35S:AaLTP1-EGFP, 35S:AaLTP2-EGFP and 35S:AaLTP3-EGFP were transferred to *A. tumefaciens* AGL-O using electroporation.

For the LTP-RFP fusion protein genes the LTP coding sequence without stop-codon were cloned in frame in TOPO vector (Invitrogen). The resulting TOPO/AaLTP1, TOPO/AaLTP2 and TOPO/AaLTP3 were cloned under control
of CaMV 35S promoter into the pK7RWG2 binary vector with the RFP coding sequence (http://gateway.psb.ugent.be/) using LR recombination (Invitrogen).

**Metabolite analysis using LC-MS**

For the AN-PW free compound analysis, the agro-infiltrated leaves were harvested, grinded in liquid nitrogen. Frozen grinded samples (100 mg) were extracted in 300 ul methanol using 10 min sonication, after which the samples were centrifuged at 18,625 x g at 4 °C. The supernatant was collected and filtered through 0.45 μm PTEE-membrane filter and diluted 10-fold with 10% methanol for analysis by a Waters Xevo tandem quadrupole mass spectrometer equipped with an electrospray ionization source and coupled to an Acquity UPLC system (Waters) as described (Ting et al., 2013). There is no standard for (DH)AAA, therefore, the data of (DH)AAA was given as relative peak intensity. For the AN-PW glycosides and conjugates, 100 mg of infiltrated leaf material was extracted in 300 ul methanol / 0.137% formic acid using 10 min sonication, after which the samples were centrifuged at 18,625 x g at 4 °C and filtered through a 0.45μm PTEE-membrane filter. The supernatant was analyzed on a LC-QTOF-MS (Waters). Since leaves showed signs of necrosis after 7 dpi, for samples collected from 7 dpi to 18 dpi the samples were freeze-dried and yield was determined by dry-weight rather than by fresh weight. For analysis 100 mg of dried samples was used following the same extraction procedure as for fresh-weight samples.

**Apoplast fluid analysis using LC-MS**

Apoplastic fluids were collected using the infiltration centrifugation technique with modifications (Joosten, 2012; Witzel et al., 2011). For the vacuum infiltration, the fresh leaves were cut into 5 cm strips and stack in a beaker filled with distilled water, and then applied vacuum with a vacuum pump. The vacuum was released slowly by opening the vent on the desiccator jar. For the apoplast fluid collection, leaf strips were placed on a grid resting on the conical bottom part of 50 ml plastic centrifuge tube (Greiner bio-one). Tubes were centrifuged at 400 g, 15 min at 4°C to collect the fluid in the bottom of the tube. After collection of the released apoplast fluid the leaf strips were weighted again, and subsequently shock frozen in liquid nitrogen for the metabolites analysis. The apoplast wash fluid was filtered through a 0.45 μm PTEE-membrane filter and the filtered fluid was dilute 5 times with 12.5% MeOH. For analysis of extracted compounds, the diluted apoplast fluid was injected into the UPLC-MRM-MS (Waters). The
concentration of the different products in the apoplast fluid was used to compare the different gene combinations in the transient expression assays.

**Semi-quantification of LTP-EGFP fusion protein levels in agro-infiltrated leaves**

We used LTP-GFP fluorescence as a semi-quantitative measure of LTP-EGFP protein accumulation. Leaves of *N. benthamiana* were infiltrated with *A. tumefaciens* carrying the *CaMV 35S::LTP-EGFP* expression constructs or with empty vector construct (control). Infiltrated leaves were harvested at 6 dpi and four regions of equal size at the same position were imaged per leaf under UV illumination in equal exposure time with a Leica MZIII fluorescence stereomicroscope and a Nikon Optiphot-2 coupled to a mercury lamp. Color images were split into individual R, G or B channels and the signal in the G channel was quantified in ImageJ 1.49n. Four biological replicate leaves (16 individual areas) were used to determine the average fluorescence signal in the G channel for each treatment.

**Subcellular Localization of LTP-GFP and GFP-PDR**

The LTP-GFP and GFP-PDR proteins were transiently expressed in the *N. benthamiana* leaves. Subcellular targeting analysis was carried out at different dpi by confocal laser scanning. In order to remove air from leaf sections, the leaf pieces were cut into 0.5 cm discs and mounted in perfluorodecalin (Sigma) (Littlejohn and Love, 2012). A Zeiss LSM 5 PASCAL laser scanning microscope (Carl Zeiss) and a 63x (N.A. 1.4 oil) Plan Apochromat objective (Zeiss) was used for collect the images as described (Dong et al., 2013). The same samples were analyzed with laser excitation settings of 488-nm for GFP detection and 561-nm for RFP detection. Images were processed with Zeiss LSM image brower version 5.

**Exclusion assay**

4 weeks old *N. benthamiana* leaves were infiltrated with *A. tumefaciens* harboring AaLTP3, AaPDR2, AaLTP3+AaPDR2 or EV as the control, each together with P19. In order to achieve the highest expressed protein levels, infiltrated leaves were kept on plant for 5 days and then harvested. For each treatment, four leaf disks of 1 cm diameter were excised from four individual leaves and were vacuum infiltrated with 3500 µl of 500 µM Na2PO4 buffer to make a con-
constant flow between the buffer and apoplastic region. Then 2500 µl of solution was taken out and a solution of AA was added into each well in a manner that each well got 0.625 mg of free compound. The 16-well plate was shaken gently for 2h on an orbital shaker and then the leaf disks were extracted with 300 µl of MeOH and 0.1% formic acid. Extracts were analyzed by LC-QTOF-MS.

Results

Isolation of *A. annua* glandular trichome LTP genes

Six different LTP sequences were identified in our *A. annua* trichome cDNA library (Bertea et al., 2006). Three of them (AaLTP1/2/3) were consistently detected, also in other published *A. annua* glandular trichome enriched cDNA libraries (Covello et al., 2007; Teoh et al., 2009), while three other LTP sequences (AaLTP4/5/6) were detected in an *A. annua* filamentous trichome enriched cDNA library (Soetaert et al., 2013) (Table 1). The expression of the six LTP genes was analyzed in different stages of *A. annua* leaf development and the apex of *A. annua* plants and compared to that of ADS, which encodes the first step in artemisinin biosynthesis. The expression profiles of AaLTP4/5/6 are different from the expression profile of ADS (Figure S1), which suggests they are not involved in AN-PW product transport, but in another process possibly in filamentous trichomes. AaLTP1 and AaLTP2 have a similar profile to that of ADS (Figure S1), while the expression profile of AaLTP3 is different, showing a relative low expression in the shoot apex and much higher activity in older leaves. Since the pathway for artemisinin is active in glandular trichomes, LTP1/2/3 were selected for further functional analysis in combination with AN-PW genes.

For the functional analysis of the AaLTP genes we also used two characterized *A. annua* ABC-G type transporter genes (AaPDR1 and AaPDR2; Boutry, personal communication). Both AaPDR1 and AaPDR2 are present in multiple glandular trichome enriched cDNA libraries (Table 1). Expression profiling shows that the profile of AaPDR1 is similar to that of ADS and AaLTP1/2, while the expression profile of AaPDR2 is more similar to that of AaLTP3 (Figure S1). The expression profile of AaLTP3 and AaPDR2 matches with the accumulation profile of DHAA over these developmental stages (Zhang et al., 2012). The three LTP genes with expression in glandular trichomes were cloned into binary expression vectors for functional analysis in combination with the AN-PW genes and the two PDRs.
Table 1. LTPs and PDRs expression in trichome databases. F: Filamentous trichomes; G: Glandular trichomes; ND: not detected; +: detected.

<table>
<thead>
<tr>
<th>Data type</th>
<th>Tissues</th>
<th>LTP1</th>
<th>LTP2</th>
<th>LTP3</th>
<th>LTP4</th>
<th>LTP5</th>
<th>LTP6</th>
<th>PDR1</th>
<th>PDR2</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNAseq data</td>
<td>G vs. F</td>
<td>G</td>
<td>G</td>
<td>ND</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>G</td>
<td>ND</td>
<td>(Soetaert et al., 2013)</td>
</tr>
<tr>
<td>ESTs</td>
<td>G</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>ND</td>
<td>(Covello et al., 2007)</td>
</tr>
<tr>
<td>454 pyrosequencing</td>
<td>G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ND</td>
<td>(Wang et al., 2009)</td>
</tr>
<tr>
<td>transcriptome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cDNA library</td>
<td>G + F</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>ND</td>
<td>+</td>
<td>(Bertea et al., 2006)</td>
</tr>
<tr>
<td>Trichome database</td>
<td>G + F</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>+</td>
<td>(Dai et al., 2010)</td>
</tr>
<tr>
<td>Tang’s RNAseq</td>
<td>G vs. F</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Personal communication</td>
</tr>
<tr>
<td>Brodelius’ RNAseq</td>
<td>G vs. F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td>G</td>
<td></td>
<td></td>
<td>Personal communication</td>
</tr>
</tbody>
</table>

**Subcellular localization of AaLTP1/2/3 upon transient expression in N. benthamiana**

All three glandular trichome expressed AaLTPs contain a N-terminal signal peptide sequence, which is indicative for targeting to the ER and secretion to the apoplast. We tested the subcellular targeting of these AaLTPs by construction *LTP-GFP* fusion genes under control of the CaMV 35S promoter. Transient expression of these genes showed that at 4 dpi all LTP1-GFP, LTP2-GFP and LTP3-GFP were detected in internal membrane structures (presumably ER) (Figure 1A and S2). However, at 6-7 dpi the strongest GFP signal was observed outside the cells. Especially in the spongy mesophyll bridges of
fluorescent material between cells were observed, indicating local accumulation of the LTP-GFP fusion proteins in the apoplast (Figure 1B). Remarkably, for all three LTP-GFP fusion proteins, the extracellular accumulation seemed to be coordinated between cells in spongy mesophyll, resulting in organization of LTP-GFP proteins surrounding an intracellular cavity (Figure 1C).

AaLTPs+AaPDRs enhance free form AN-PW products in *N. benthamiana* leaves

We attempted to proof a putative role of AaLTP1 to 3 in the accumulation of AN-PW products by inhibiting their expression in *A. annua* using Virus Induced Gene Silencing (VIGS). However, our VIGS constructs did not dis-
play silencing activity in *A. annua* (data not shown). Therefore, we performed a functional analysis of the AaLTPs using transient co-expression with AN-PW and *AaPDR* genes in *N. benthamiana*. First, the AN-PW genes were co-infiltrated with either the three AaLTP genes or the two AaPDR genes. In all these experiments the relative dosage of agro-infiltrated AN-PW genes was kept constant, such that product accumulation was constant and the contribution of LTPs and PDRs could be accurately determined. Leaves were harvested six days post-agroinfiltration (dpi) and un-conjugated AN-PW intermediates were quantified by LC-MRM-MS as described (Ting et al., 2013). Figure 2 shows that co-expression of the AN-PW biosynthesis with the three *LTP* genes or the two *PDR* genes did not have a significant effect on the levels of (DH)AA, (DH)AAOH or (DH)AAA (Figure 2). However, the combination of AN-PW+AaLTPs+AaPDRs genes resulted in a significant increase in the accumulation of free AN-PW products compared with the pathway alone (Figure 2). LC-QTOF-MS analysis of the leaf extracts showed that the level of AN-PW glycosides and glutathione conjugates, which are more abundant than the free AN-PW products and intermediates, in the agroinfiltrated leaves, were not significantly affected by the AaLTPs and AaDPRs (Figure S3). The results suggest that AaLTPs, together with AaDPRs, can enhance sequestration of free form AN-PW intermediates away from conjugating enzyme activities. This raised the question whether these free form compounds reside inside or outside the cellsAN-PW.
AN-PW products accumulate in the apoplast of *N. benthamiana* leaves

To investigate whether the increased level of free form AN-PW intermediates and products mediated by co-expression of *AaLTPs+AaPDRs* with AN-PW is due to increased transport to the apoplast, we extracted and analyzed the apoplast. Here, leaves were harvested at 6 dpi and infiltrated with distilled water under vacuum. The infiltrated water was subsequently collected by mild centrifugation according to (Witzel et al., 2011). Analysis of this apoplast extract by UPLC-MRM-MS showed that AAOH, AA, DHAAOH and DHAA can be detected in the apoplast of *N. benthamiana* leaves expressing only the AN-PW genes (Figure 3). This suggests there is intrinsic transport activity for AN-PW intermediates in *N. benthamiana*. Rutin is part of the metabo-

![Figure 3 Artemisinin pathway intermediates and products in the apoplast of *N. benthamiana* leaves infiltrated with AN-PW genes with the *LTPs and PDRs* as identified by LC-MRM-MS at 6dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. 10 constructs were co-expressed in this experiment. Error bar is SE; p<0.05, n=4.](image)

lome of *N. benthamiana* whole leaf extracts and it was reported to be only present in the cytoplasm and/or vacuole of plants (Markham et al., 2001; Marrs et al., 1995; Ökmen et al., 2013). As expected, rutin was not detected in the apoplast wash fluid (Figure S4). Analysis of apoplast samples by LC-QTOF-MS occasionally showed the presence of very low levels of DHAA-Hex3 (at most 5% of the total amount found in whole leaf extracts) (Figure S5). Combined, this indicates that only low level of cell damage occurred during the apoplast wash procedure and that one specific glycosylated AN-PW intermediate is transported to the apoplast at a low rate (Figure S5).

Analysis of the apoplast of leaves expressing either PW+*LTPs* or PW+*D-PRs* did not show a significant increase in AN-PW intermediates, con-
firming results from total leaf extracts (Figure 2). In contrast, expression of AN-PW+AaPDRs+AaLTP genes resulted in a significant increase in apoplastic levels of AAOH, AA, DHAAOH and DHAA (Figure 3), and DHAAA and AAA cannot detected in the apoplast. Having confirmed that the combined LTPs and PDRs can enhance apoplastic sequestration of AN-PW intermediates, we subsequently tested the specificity of the individual LTPs and PDRs in AN-PW product sequestration.

AaPDR2 is more effective in AN-PW product sequestration than AaPDR1

To determine the transport/sequestration activity of AaPDR1 and AaPDR2 in the transient expression assay in *N. benthamiana*, the AN-PW genes were co-expressed with all three LTP genes in combination with either AaPDR1 or AaPDR2. The infiltrated leaves were harvested at 7 dpi and the apoplast was isolated and analyzed by UPLC-MRM-MS. Results show that for AA and DHAA AaPDR2 was more effective than AaPDR1 in transport to the apoplast (Figure 4). This difference in AaPDR activity could be related to difference in individual AaPDR structure or protein stability. Because AaPDR2 turned out to be most active, we subsequently tested the specificity of the individual AaLTPs in combination with the AN-PW genes and AaPDR2. In addition, co-expression of all LTPs and PDRs is less efficient than expressing specifically LTPs+PDR2 with the AN-PW genes as these...
genes are co-expressed in the same cell in *N. benthamiana*, indicating the transport membrane dosage affects the sequestering of the metabolites as well. Transporter genes compete membrane sites, thus prevent the transport efficiency. The sequestering of (DH)AA for PDRs + LTPs + AN-PW is less efficient than PDR2+LTPs+PW at the same relative gene dosage.

**AaLTP3 is most active in enhancing (DH)AA levels in *N. benthamiana***

To determine whether there is a difference in transport/sequestration activity of the individual AaLTPs, the AN-PW genes were co-expressed with AaD-PR2 and the individual AaLTP genes in *N. benthamiana* leaves. Agro-infiltrated leaves were harvested at 7 dpi and free form AN-PW products were quantified in total leaf extracts. The results show that all three AaLTPs enhance the level of free (DH)AA compared with expression of AN-PW+AaP-DR2 alone. However, AaLTP3 had the strongest effect on enhancing AN-PW product sequestration (Figure 5). A control experiment showed that the transcript levels of the individual AaLTPs and AaPDR2 were very

![Figure 5](image-url)
similar in these experiments (Figure S6), indicating that the difference in AaLTP sequestering activity must be either related to difference in individual AaLTP structure/function or difference in individual LTP protein stability.

**Relative transport efficiency for artemisinin pathway intermediates**

The full combination of $AN-PW+PDR2+LTP3$ gave the highest AN-PW product levels in the apoplast, while $PW+PDR2$, or $PW+LTP3$ only gave a small increase in AN-PW product levels in the apoplast compared with $PW+2EV+P19$ (Figure 6). The ratio between the product level in the apoplast
and in a total leaf extract was calculated for each of the AN-PW intermediates. This showed a higher specificity for transport of (DH)AA (up to 33%) to the apoplast than for (DH)AAOH (up to 10%). However, this specificity is the same as for the intrinsic transport activity in *N. benthamiana* as well as for leaves that have additional overexpression of AaPDR2 (Figure 6). The fact that not all non-glycosylated AN-PW products are in the apoplast means some free compounds are still inside the cells, or extraction of the apoplast with water is not complete. Indeed, control experiments show that the relative transport efficiency of (DH)AA is very similar to the partitioning of these products between polar water phase and an inorganic chloroform phase (Figure S7), while the relative transport efficiency of (DH)AAOH is actually lower than the partitioning of (DH)AAOH between water and chloroform.

To determine whether the increased sequestration of free AN-PW products in the apoplast has a positive effect on total AN-PW activity we also measured the (DH)AA glycosides and glutathione conjugates using LC-QTOF-MS in leaves expressing the different combinations of AN-PW, AaLTP3 and AaPDR2. These measurements show that the levels of glycosylated (DH)AAOH and glutathione AAA are lower than of the glycosylated (DH)AA products (Figure S8), consistent with a specific transport activity for the (DH)AA products. The total flux through the AN-PW is increased by AaPDR2 and even stronger by AaLTP3 or AaLTP3+AaPDR2 (Figure S8).

**High sequestration activity of LTP3 relates to high AaLTP3 protein stability**

AaLTP3 was more effective in combination with AaPDR2 in sequestering (DH)AA than AaLTP1 or AaLTP2 at equal transcriptional activity (Figure 5). This could be due to higher specific activity of AaLTP3 towards (DH)AA or could be the result of differences in intrinsic protein stability. In order to investigate the intrinsic protein stability of the three *A. annua* LTPs we expressed AaLTP-GFP fusion proteins in *N. benthamiana* leaves and imaged...
leaf fluorescence under an epi-fluorescence stereo-microscope at 5 dpi. At this stage the fluorescence of the AaLTP-GFP fusion protein was clearly above the background fluorescence of chlorophyll in the leaf samples and the AaLTP-GFP fluorescence was semi-quantified from the intensity in the green channel of RGB images taken at different positions of agro-infiltrated leaves. Results show that fluorescence of leaves expressing AaLTP3-GFP is significantly higher than that of leaves expressing AaLTP1-GFP or AaLTP2-GFP (Figure S9), while control experiments show that transcriptional activity of these genes is very similar. The higher sequestration activity of AaLTP3 therefore seems to directly relate to intrinsic stability of AaLTP3.

We also tested whether the interaction between AaLTP3 and AaPDR2 in sequestration of (DH)AA (Figure 7) is having an effect on AaLTP3 protein stability. We previously noticed that the synergistic interaction between the set of three AaLTPs +AaPDR2 for sequestration of AN-PW intermediates, coincided with an overall increase in GFP fluorescence signal when AaLTP-GFP fusion genes were used (Figure 7a). We therefore also tested whether AaLTP3 protein stability is affected by active AN-PW product transport. For this, AaLTP3-GFP was expressed with either AaPDR2 alone or with the AN-PW genes or with the full set of AN-PW+AaPDR2 (keeping the AaLTP3-GFP gene dosage the same between different treatments). AaLTP3-GFP fluorescence significantly increased in the presence of both AaPDR2 and AN-PW gene expression, suggesting that AaLTP3-GFP protein levels are

![Figure 7](image.png)
stabilized in the presence of active AN-PW product transport (Figure 7b).

**Conversion of (DH)AA in *N. benthamiana* leaves is most effectively blocked by AaLTP3**

So far we have shown that AaLTP3+AaPDR2 can enhance transport of (DH)AA to the apoplast of *N. benthamina* leaves expressing the AN-PW genes. Another way by which transport and sequestration activity of proteins may be tested is by a product exclusion assay where the product is presented outside of the cells expressing the candidate proteins involved in transport. We therefore wanted to test the activity of AaLTP3 and AaPDR2 in excluding influx of (DH)AA infiltrated into the apoplast of *N. benthamiana* leaves. First we determined the fate of (DH)AA which was infiltrated into the apoplast of *N. benthamiana* leaves without transgene expression. Unexpectedly, this showed a very rapid conversion of AA and DHAA into (DH)AAA and (DH)AAOH, indicating (1) that the substrates are rapidly internalized into the leaf cells, (2) that endogenous enzymes very actively reduce the infiltrated substrates, leading to (DH)AAA and (DH)AAOH within minutes of (DH)AA infiltration) and full consumption of (DH)AA at 4 hr post apoplast infiltration (Figure 8). Next, we used *N. benthamiana* leaves at 6 dpi expressing either EV, AaLTP3+EV, AaPDR2+EV or AaLTP3+AaPDR2 for infiltration of AA. We used the intracellular conversion to glycosides at two hours post apoplast infiltration as a measure of exclusion from the cells. Results show that leaves expressing EV were not able to exclude the substrate, resulting in high levels of AA glycosides and DHAA glycosides. Expression of only AaPDR2 was able to prevent ~27% of product conversion to glycosides, AaLDP3 was able to prevent ~74% of product conversion to glycosides and the combined action of AaLTP3+AaPDR2 prevented ~82% of product conversion. Combined, the results indicate that transport of compounds to the apoplast is not effective where there it is not combined with a retention system in the apoplast and that AaLTP3 can function in the retention of artemisinin products in the apoplast.

**AaLTP3 + AaPDR2 enhance accumulation of artemisinin and arteanuin B**

Previous attempts to detect the desired end products artemisinin (AN) or arteanniun B (AB) in *N. benthamiana* leaves expressing the AN-PW genes were unsuccessful when leaves were harvested at 7 dpi (Ting et al., 2013). However, we found that in leaves harvested at 9 dpi both AN and AB were
detected and that levels of AN and AB increased until 13 dpi, after which AN and AB declined again (Figure 9). Both for AN and AB the accumulation was highest when the AN-PW genes were co-expressed with AaLT-P3+AaPDR2. The accumulation of the different pathway intermediates (AAOH, AAA, AA, DHAAOH, DHAAA, DHAA) is shown in Figure S10. Results show a general trend of decrease in the alcohol (AAOH, DHAAOH) and acid products (AA, DHAA) but an increase in the aldehyde products (AAA, DHAAA) over time. The change in product profile coincides with an onset of necrosis in the agroinfiltrated leaves and raises the question wheth-

Figure 8. Bioconversion dynamics of DHAA or AA in *N. benthamiana* leaves. (a) Apoplast infiltration of DHAA (500 ng/mg) and harvest whole leaf discs at 1-240 min; (b) Apoplast infiltration of AA (500 ng/mg) and harvest at 0-4 h; Unit of AA, DHAA, AAOH, DHAAOH is ng/mg except for AAA and DHAAA (intensity) in a and b. (c) Apoplast infiltration of AA (2.67 mM) at 3 dpi of *N. benthamiana* leaves expressing AaPDR2+EV, AaLTP3+EV or AaPDR2+AaLTP3 at 3 dpi and harvest whole leaf discs 120 min after infiltration of AA into the apoplast.
er conversion of (DH)AA to AN or AB is actually caused by leaf necrosis. The onset of necrosis after 7 dpi in leaves expressing the AN-PW genes (with or without LTP3+PDR2) prevents a reliable analysis of the apoplast content by the apoplast–wash method. Therefore, at present it is not known whether the accumulation of AN and AB is a consequence of continued (DH) AA transport to the apoplast between day 7-13 in agro-infiltrated leaves.

Discussion

AaLTP3 affects the AN biosynthesis pathway activity in N. benthamiana

Here we have investigated the role for AaLTPs in AN-PW product accumulation using transient expression in N. benthamiana. We have collected mul-

Figure 9 Production of AB and AN in leaf of N. benthamiana infiltrated AN-PW genes with the LTP3 and PDR2 as identified by LC-MRM-MS at different time points. 8 constructs were co-expressed in this experiment. AB, arteannuin B; AN, artemisinin. Error bar is SE; p<0.01, n=4.
tiple lines of evidence that AaLTP3 functions with AaPDR2 in the sequestration of AN-PW products: (1) AaLTP3 is coexpressed with AN-PW genes and PDR transporter genes in glandular trichomes (Table 1); (2) the AaLTP3 expression profile is similar to that of AaPDR2 (Figure S1) and AaLTP3 activity on free AN-PW compound accumulation is highest in combination with AaPDR2 (Figure 5); (3) AaLTP3 protein is extracellular (Figure 1), while AaLTP3 enhances AN-PW product accumulation in the apoplast; (4) AaLTP3 functions in retention of AA in the apoplast (Figure 8); (5) AaLTP3 protein stability is increased by AN-PW activity (Figure S9) suggesting some kind of functional interaction between AaLTP3 and AN-PW activity.

The need of AaLTP3 function in planta

The experiments with apoplast infiltration of AA and DHAA have shown that both of these compounds are rapidly taken up by the cells of N. benthamiana and that endogenous enzyme activity in N. benthamiana causes a strong reverse flux of (DH)AA to (DH)AAA and (DH)AAOH. This is supposedly caused by endogenous alcohol dehydrogenases (ADHs) or oxidoreductases that reduce (DH)AA. Such activity may compromise accumulation of desired AN-PW end-products in N. benthamiana. A similar activity was detected in A. annua by Ryden et al., who characterized an oxidoreductase Red1, expressed in A. annua flowers that catalyses the reduction of DHAAA to DHAAOH (Rydén et al., 2010). This suggests that for heterologous expression systems and possibly also for A. annua the removal of (DH)AA from the cytosol through transport to and retention in the apoplast by AaLTP3 may be an important driver for the AN-PW flux towards (DH)AA. Indeed, the specific removal of (DH)AA from the endogenous AN-PW in N. benthamiana increased the overall flux through the AN-PW, as both free (DH)AA and glycosylated (DH)AA product accumulation were increased by co-expression of PDR2 and LTP3 (Figure 5, 6 and S8), while there was much less effect on free and glycosylated (DH)AAOH (Figure S8).

Specificity of AaLTPs for AN-PW products

There are several lines of evidence that this effect of AaLTP3 is specific for (DH)AA and is not caused by generation of a general (lipid) sink for sesquiterpenes in the apoplast: (1) in contrast to AaLTP3, AaLTP1 and AaLTP2 have little or no effect on free (DH)AA accumulation in the apoplast. (2) AaLTP3 protein stability is increased by AN-PW activity. Finally, transient co-expression of LTP3 with the costunolide biosynthetic pathway from feverfew (a sesquiter-
pene PW very similar to the AN-PW; (Liu et al., 2011)) in *N. benthamiana* did not affect costunolide accumulation, suggesting specificity for (DH)AA and not any other lipid molecule (A. Beyraghdar Kashkooli, personal communication).

**Putative model of AaLTP3 function**

The retention activity of AaLTP3 for AA cannot be explained by a simple one to one binding of LTP3 to AA. If we assume an extreme high AaLTP3 production at 3 dpi of 1 mg/gr FW in *N. benthamiana*, the estimated Molar ratio of (AaLTP3):(infiltrated AA) is approximately 1: 635. Although at present we can only speculate on how AaLTP3 fulfills the retention activity in the apoplast, we favor the model in which AaLTP3 transports AN-PW products from the space between the plasma-membrane and cell wall to the outer side of the cell wall in the intercellular space. Once deposited at this site a reflux of (DH)AA back into the cell is then largely prevented by LTP3. The empty LTP may subsequently diffuse back to the plasma membrane to be re-loaded by another (DH)AA molecule.

**Concluding remark**

In conclusion, we here suggest that LTP3 plays a role in the transport of artemisinin pathway products. In combination with PDR2, LTP3 facilitates sequestration of DH(AA) into the apoplast and prevents reflux back into the cells. This newly discovered function of LTPs opens up new possibilities for the engineering of biosynthesis pathways of high value terpenes in heterologous expression systems.

**Acknowledgements**

We thank Miriam Goedbloed, Marielle Schreuder and Jacqueline Busscher-Lange for helpful assistance in the laboratory, Bert Schipper for assistance in LC-QTOF-MS analysis and Francel Verstappen and Kristyna Flokova for help in LC-MRM-MS analysis.

**References**


Bertea, C.M., Voster, A., Verstappen, F.W.A., Maffei, M., Beekwilder, J., and Bouw-


Joosten, M.A.J. (2012). Isolation of apoplastic fluid from leaf tissue by the vacuum infiltration-centrifugation technique. In Plant Fungal Pathogens, M.D. Bolton, and
Artemisia Lipid Transfer Protein

B.P.H.J. Thomma, eds. (Humana Press), pp. 603-610.


**Supplementary information**

Figure S1 Relative gene expression of *AaPDRs* and *AaLTPs* genes in shoot apex and different stages leaves of *A. annua*. Values indicate the mean fold relative to sample shoot apex, which was normalized to 1.
Figure S2 Bright field of AaLTP1/2/3 subcellular localization. Confocal microscopy analysis of *N. benthamiana* leaf infiltrated with *pBin:GFP, AaLTP1:GFP, AaLTP2:GFP* and *AaLTP3: GFP* separately in (A) intercellular at 4 dpi; (B) bridges between spongy mesophyll at 4 dpi; (C) cavity between the mesophyll cells at 6-7 dpi. Bar: 20 µm.
Figure S3 Artemisinin glycosides conjugations in leaves of *N. benthamiana* infiltrated AN-PW genes with the LTPs and PDRs as identified by LC-QTOF-MS at 6dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. 10 constructs were co-expressed in this experiment. Error bar is SE; p<0.05, n=4.
Figure S4 Detection of rutin in the apoplast samples under the LC-Orbitrap-MS. A: apoplast sample, B: remaining leaf sample after apoplast wash, C: mass spectrum of rutin.

Figure S5 Detection of (DHAA-Hex3)FA in the apoplast samples under the LC-Orbitrap-MS. A: apoplast sample, B: remaining leaf sample after apoplast wash, C: mass spectrum of (DHAA-Hex3)FA. DHAA, dihydroartemisinic acid.
Figure S6 The relative gene expression of *AaPDR2* and *AaLTP1/2/3* in the *N. benthamiana* leaves. Error bar is SE; \( p < 0.01 \), \( n = 4 \).

Figure S7 Partitioning of artemisinin-intermediates standards in water vs. chloroform. AAOH, artemisinic alcohol; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAA, dihydroartemisinic acid.
Figure S8 Artemisinin-related glycosides and glutathione conjugates production in leaf extraction of *N. benthamiana* infiltrated AN-PW genes with LTP3 and PDR2 as identified by LC-QTOF-MS at 6 dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. 7 constructs were co-expressed in this experiment. Error bar is SE; p<0.05, n=5.
Figure S9 Intrinsic fluorescence of individual LTP-GFP, transiently expressed in *N. benthamiana*. LTP1: AaLTP1-GFP. LTP2: AaLTP2-GFP. LTP3: AaLTP3-GFP. EV: Empty Vector. **: P < 0.01
Figure S10 Production of AN-PW compounds in leaf of *N. benthamiana* infiltrated AN-PW genes with the *LTP3* and *PDR2* as identified by LC-MRM-MS at different time points. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. 8 constructs were co-expressed in this experiment. Error bar is SE; p<0.01, n=4.
# Table S1 Primers used in this paper

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence (5’ &gt; 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>β-actin_Fw</td>
<td>CCAGGCTGTTTCAGTCTCTGTAT</td>
</tr>
<tr>
<td>β-actin_Rev</td>
<td>CGCTCGGTAGGATCCTCATCA</td>
</tr>
<tr>
<td>AaADS_Fwd</td>
<td>GGAATGACCCCAACCTTGA</td>
</tr>
<tr>
<td>AaADS_Rev</td>
<td>TCGTCTCCCCATACGTGTGAA</td>
</tr>
<tr>
<td>AaLTP1_Fwd</td>
<td>GCAATAGAGGAGGAGGTGAC</td>
</tr>
<tr>
<td>AaLTP1_Rev</td>
<td>AAGCAGCTGCTGAGGCTTA</td>
</tr>
<tr>
<td>AaLTP2_Fwd</td>
<td>AACTCCGCTGCTAAACGA</td>
</tr>
<tr>
<td>AaLTP2_Rev</td>
<td>TCGGTCTAAGAGGATCCTT</td>
</tr>
<tr>
<td>AaLTP3_Fwd</td>
<td>AGCACCCTATGCTGAGGCTA</td>
</tr>
<tr>
<td>AaLTP3_Rev</td>
<td>AGTTTAGGACAGGAGGTG</td>
</tr>
<tr>
<td>AaPDR1_Fwd</td>
<td>AGAAGCAGCTGCTGAGGAT</td>
</tr>
<tr>
<td>AaPDR1_Rev</td>
<td>GAGCAACGAGCAAACATGA</td>
</tr>
<tr>
<td>AaPDR2_Fwd</td>
<td>GTTGGCTTCTCAGGTTT</td>
</tr>
<tr>
<td>AaPDR2_Rev</td>
<td>CCGAGCTTGGTGGT</td>
</tr>
<tr>
<td>AaLTP1-F</td>
<td>TTCCATGTTGAGTACATGAG</td>
</tr>
<tr>
<td>AaLTP1-R</td>
<td>GCGGCCGCTCATTGTATCGG</td>
</tr>
<tr>
<td>AaLTP2-F</td>
<td>TTCCATGGAAGATGAGGAG</td>
</tr>
<tr>
<td>AaLTP2-R</td>
<td>GCGGCCGCTCATTGTATCGG</td>
</tr>
<tr>
<td>AaLTP3-F</td>
<td>TTCCATGGAAGATGAGGAG</td>
</tr>
<tr>
<td>AaLTP3-R</td>
<td>GCGGCCGCTCATTGTATCGG</td>
</tr>
<tr>
<td>EGFP_c-term-F</td>
<td>GAAGGAGCGGCGGCTGCCAG</td>
</tr>
<tr>
<td>EGFP_c-term-R</td>
<td>CACAAAGAGCTCTTTATAAG</td>
</tr>
<tr>
<td>AaLTP1_gfp-F</td>
<td>AACACCGGATCCAATGTTGAAAG</td>
</tr>
<tr>
<td>AaLTP1_gfp-R</td>
<td>GACCAAGCTACACAGGAGGAG</td>
</tr>
<tr>
<td>AaLTP2_gfp-F</td>
<td>AACACCGGATCCAATGTTGAAAG</td>
</tr>
<tr>
<td>AaLTP2_gfp-R</td>
<td>GACCAAGCTACACAGGAGGAG</td>
</tr>
<tr>
<td>AaLTP3_gfp-F</td>
<td>AACACCGGATCCAATGTTGAAAG</td>
</tr>
<tr>
<td>AaLTP3_gfp-R</td>
<td>GACCAAGCTACACAGGAGGAG</td>
</tr>
<tr>
<td>AaLTP1_rfp-F</td>
<td>CACCATGGTTGAGTACATGAG</td>
</tr>
<tr>
<td>AaLTP1_rfp-R</td>
<td>GACCAAGCTACACAGGAGGAG</td>
</tr>
<tr>
<td>AaLTP2_rfp-F</td>
<td>CACCATGGTTGAGTACATGAG</td>
</tr>
<tr>
<td>AaLTP2_rfp-R</td>
<td>GACCAAGCTACACAGGAGGAG</td>
</tr>
<tr>
<td>AaLTP3_rfp-F</td>
<td>CACCATGGTTGAGTACATGAG</td>
</tr>
<tr>
<td>AaLTP3_rfp-R</td>
<td>GACCAAGCTACACAGGAGGAG</td>
</tr>
</tbody>
</table>
Chapter 4

Interaction between ectopically induced flavonoid and sesquiterpenoid biosynthetic pathways in *Nicotiana benthamiana*

Bo Wang¹, Nikolay Ouchkouroí ¹, Jules Beekwilder¹,², Harro J. Bouwmeester¹, Alexander R. van der Krol¹

¹Laboratory of Plant Physiology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
²Plant Research International, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands
Summary

The sesquiterpenoid biosynthetic pathway in the glandular trichomes of *Artemisia annua* is responsible for the production of the important anti-malarial compound artemisinin. Also the flavonoid biosynthesis pathway is expressed in *A. annua* glandular trichomes, however, the functional significance of this has not been studied so far. Here we investigated the interaction between these two pathways using transient expression assays in *Nicotiana benthamiana*. The endogenous *N. benthamiana* flavonoid biosynthetic pathway was stimulated by transient expression of the *Antirhinum majus* transcription factor *Rosea1 (ROS)* and the artemisinin (AN) biosynthesis pathway (PW) by expression of AN-PW genes. Upon coexpression of AN-PW genes and ROS product formation through the AN-PW increased. Gene expression analysis showed that this can be attributed to stimulation by ROS of the expression of endogenous mevalonate kinase (*NbMVK*), which catalyzes a rate-limiting step in the mevalonate pathway which supplies the substrate for sesquiterpenoid biosynthesis. Finally, we showed that flavonoid biosynthesis competes with product formation by the AN-PW. Together these observations match with the result that highest accumulation of artemisinin was obtained by co-expression of AN-PW, ROS, and transporter genes with or without *CHS*_{RNAi}.

Key words

Flavonoids, sesquiterpenoids, *Artemisia annua*, *Nicotiana benthamiana*
**Introduction**

Flavonoids and terpenoids are two super families of secondary metabolites in plants and here we studied the interaction between these two major biosynthesis pathways. Flavonoids are synthesized in the phenylpropanoid biosynthesis pathway. The first committed step for flavonoid biosynthesis is catalyzed by the enzyme chalcone synthase (CHS) which combines 3 molecules of malonyl-CoA with one p-coumaroyl-CoA (formed in the shikimate biosynthesis pathway) to produce naringenin chalcone. This basic structure of naringenin chalcone can be further modified by reductases, isomerases, hydroxylases, and glycosyltransferases to make various subclasses of flavonoids, such as flavones, flavonols, anthocyanins, proanthocyanidins (Ferrer et al., 2008; Glas et al., 2012). Most of the structural genes in the flavonoid biosynthesis pathway are coordinately regulated by a ternary complex comprising three types of transcription factors (TFs); a R2R3-MYB TF, a basic helix–loop–helix (bHLH) TF and a WD-repeat (WDR) TF (Patra et al., 2013). Especially the MYB and bHLH TFs have been successfully used to ectopically induce the flavonoid biosynthesis pathway and enhance anthocyanin content in several different plant species (Outchkourov et al., 2014; Schwinn et al., 2014; Stommel et al., 2009; Tai et al., 2013). Here we used transient expression of the *Antirhinum majus* MYB TF Rosea1 (ROS) to ectopically enhance production of flavonoids in *Nicotiana benthamiana*.

Terpenoids are composed of isoprene units, which are synthesized in two separate pathways: the mevalonic acid (MVA) pathway in the cytosol, and the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway in plastids (Vranová et al., 2013). These isoprene units are converted by terpene synthases into different types of terpenoids (monoterpenoids, sesquiterpenoids, diterpenoids, etc), which may again be further modified by hydroxylation, oxidation, reduction, glycosylation, acetylation or other modifying enzyme activities, which mostly occur in the cytosol and on the ER membrane. Here we used ectopic expression of the sesquiterpenoid biosynthesis pathway genes from *Artemisia annua* in *N. benthamiana* as representative of a sesquiterpene biosynthesis pathway, to study the interaction with the flavonoid biosynthesis pathway as activated by ROS.

Previous research in tomato has indicated a requirement of flavonoid biosynthesis for type VI trichome development, and sesquiterpene biosynthesis is located in type VI trichomes as well (Kang et al., 2014; Kang et al., 2010). Moreover, a common transcriptional control of the flavonoid and terpene biosynthesis pathway by a WRKY TF has been demonstrated in
Catharanthus roseus (Suttipanta et al., 2011) and by the Arabidopsis MYB transcription factor PAP1 in Rosa hybrida (Zvi et al., 2012). In A. annua, the localization of flavonoid biosynthesis seems to overlap with that of sesquiterpene biosynthesis, i.e. in the glandular trichomes, as both the genes from the artemisinin biosynthesis pathway (AN-PW) and the flavonoid biosynthesis pathway were identified in a glandular trichome cDNA library (Wang et al., 2009). A common transcriptional regulation in A. annua has not been demonstrated, but both pathways seem to show similar activity profiles in A. annua as flavonoids and artemisinin (AN) reach their highest contents at the full blooming stage in the aerial parts of A. annua (Baraldi et al., 2008) and both accumulation of anthocyanin and AN is promoted by overexpression of the blue light receptor CRY1 (Hong et al., 2009).

It has been suggested that the co-localization of two major secondary metabolite pathways in the same glandular trichome cell may serve to optimize the function of trichome glands in a dynamic environment (Kang et al., 2014). However, at present it is not clear whether molecules of the sesquiterpene biosynthesis pathway affect flavonoid biosynthesis or vice-versa. An interaction between these two types of molecules is perhaps suggested by skin penetration studies, where terpenoids are routinely used to enhance the penetration of flavonoids (Saija et al., 1998), indicating the these molecules together may affect penetration of biological membranes.

To determine if there is an interaction between the flavonoid biosynthesis and sesquiterpene biosynthesis pathways, we expressed the AN-PW from A. annua (with transporters) and the flavonoid biosynthesis pathway (as ectopically activated by the ROS TF) alone or together by agroinfiltration of N. benthamiana leaves. Co-expression of the endogenous flavonoid biosynthesis pathway, as activated through ectopic expression of ROS, with the AN-PW genes resulted in 3-fold increase in the level of free and glycosylated artemisinic acid, while other AN-PW intermediates showed an up to 2-fold increase upon co-expression with ROS. This could not be attributed to differences in AN-PW gene transcription. Subsequent experiments showed that the increase in the AN-PW activity may be caused by the upregulation by ROS of mevalonate kinase (NbMVK), which catalyzes the rate-limiting step in the mevalonate pathway which supplies substrate for AN biosynthesis. Rather than enhancing transport of sesquiterpenoids over membranes, flavonoid biosynthesis actually reduces AN-PW product levels. Indeed, experiments with CHSRNAi indicate a competition between the AN-PW and the ROS-activated flavonoid-PW. Finally, we demonstrate that the highest level of AN and arteannuin B in N. benthamiana are reached by ectopic expres-
sion of the AN-PW, ROS, and transporter genes with or without CHS\textsuperscript{RNAi}.

**Materials and methods**

**Gene expression analysis**

100mg of infiltrated fresh leaf materials were extracted and four biological replicates were prepared for RNA isolation and gene expression analysis. Real-time quantitative PCR was performed using the iQ5 RT-PCR (BioRad). The RT-PCR cycling was: 95 °C (3min); 40cycles at 95 °C (10s), 60 °C (30s) for PCR real-time acquisition; 95 °C (1min); 55 °C (1min); and 81cycles at 55 °C (10sec) for melt curve acquisition. Primers were list in table S1. Three technical replicates were run for each cDNA sample. ΔΔCT was calculated as follows: ΔCT=CT(Target)-CT(Actin), ΔΔCT was normalized using ΔCT. The fold change value was calculated by 2-ΔΔCT.

**Transient expression in *N. benthamiana* leaves**

Transient expression in leaves of *N. benthamiana* was done as described (Van Herpen et al., 2010). The agro-bacteria cells were collected at 3000rpm, 15min, 20°C. The dosage of each construct was resuspend-ed and adjusted to a final OD600 of 0.5 in the agro-infiltration buffer. Individual expression constructs transformed *A. tumefaciens* strains were co-infiltrated into *N. benthamiana* leaves with 1ml syringes. The total dosage of *A. tumefaciens* between treatments was the same by dilution with empty vector in each set of experiments. Infiltrated leaves were harvested for chemical analysis at the 5 day post infiltration (dpi).

**Metabolites analysis using LC-MS**

Infiltrated leaves were harvested and ground to a fine powder in liquid nitrogen for metabolite analysis of infiltrated leaves in *N. benthamiana*. For the free compounds, 100mg of leaf powder were weight and extracted in 300ul methanol. After centrifuge and filtering, 10 times dilution of the supernatant with 10% methanol is performed on a Waters Xevo tandem quadrupole mass spectrometer equipped with an electROSpray ionization source and coupled to an Acquity UPLC system (Waters). For the glycosides and glutathione conjugations, 100mg of grinding infiltrated leaves were weight and extracted
in 300ul methanol with 0.137% formic acid. The samples were centrifuge at 12000rpm, 10min at 4 °C. The supernatant was performed on the LC-QTOF-MS. When necrosis was detected, the fresh samples were dried by the freezer dry after the grinding in the liquid nitrogen for the AN measurement.

**Analysis apoplastic fluid using LC-MS**

Apoplastic fluids were collected using the infiltration centrifugation technique with modifications (Joosten, 2012; Witzel et al., 2011). In order to get rid of the leaf necrosis, the leaves after the agro-infiltration were harvested at 6 or 7 dpi for the apoplast wash. For the vacuum infiltration, the fresh leaves were cut into 5 cm strips and stack in a beaker filled with demi water. A weight was placed on the leaves to prevent the leaves from floating. The vacuum was carried out 10 min in a desiccator using a vacuum pump. To prevent the cell broken, the vacuum was released slowly until the leaf become water-soaked and dark in colour. For the apoplast collection, the centrifuge tubes consist two part, the bottom part is a 50 ml plastic centrifuge tube (Greiner bio-one); the upper part has a perforated bottom and loosely fits on the bottom part. The vacuum infiltrated leaves was placed inside the upper centrifuge tube and centrifuged 400g, 15min at 4°C to collect the fluid in the bottom centrifuge tube. After the extraction of the apoplastic fraction, the residual leaf tissue was shock frozen in liquid nitrogen for the metabolites analysis. The apoplast fluids were centrifuged and filtered with 0.45μm PTEE-membrane filter. The filtered fluid was dilute into 4 times with 10% MeOH. The diluted apoplast fluid was injected into the LC-MS.

**Electrolyte leakage assay**

The electrolyte leakage assay was carried out as described (Bouwmeester et al., 2014). 8mm Leaf discs were punched from agro-infiltrated *N. benthamiana* leaves. Six leaf discs were floated on 4mL MQ water at room temperature under 250 rpm shaker for 2 hours. Four biology replicates were prepared. Conductivity was measured using a Mettler Toledo InLab 741 ISM conductivity probe (Mettler Toledo, Tiel, Netherlands). Then, samples with the leaf discs were autoclaved and measured for conductivity as well. Relative electrolyte leakage ratios were calculated by dividing conductivity values of samples before autoclaving with the conductivity values of samples after autoclaving.
Results

Co-expression of ROS+DEL+AN-PW genes leads to leaf necrosis

The artemisinin biosynthesis pathway (AN-PW) gene set was used for ectopic expression in *N. benthamiana* (N-terminal truncated 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase (*thmgr*), farnesyl diphosphate synthase (*fps*), amorphadiene synthase (*ads*), amorphadiene oxidase (*amo*), artemisinic aldehyde double-bond reductase (*dbr2*), aldehyde dehydrogenase 1 (*aldh1*) (Bouwmeester et al., 1999; Rydén et al., 2010; Teoh et al., 2009; Teoh et al., 2006; Ting et al., 2013; Zhang et al., 2008), while p19 was co-infiltrated to suppress possible silencing of transgenes (Voinnet et al., 2003). For activation of the flavonoid biosynthesis pathway in *N. benthamiana* we tested transient expression of the *A. majus* transcription factors ROS (MYB) and DEL (bHLH) (Outchkourov et al., 2014). However, combined ectopic expression of the AN-PW genes with the ROS and DEL TFs resulted in necrosis within 3 days post agro-infiltration (dpi) in *N. benthamiana* (Figure 1A). Previous results indicated that ROS is sufficient to induce

![Figure 1](image.png)

Figure 1. Necrosis in *N. benthamiana* expressing artemisinin pathway (PW), Rosea1 (ROS) and Della (DEL). a, time course onset of necrosis symptom, the black arrow indicates the necrosis; b, leaves infiltrated with different gene combination at 5 dpi.
low level accumulation of flavonoids, while DEL is only required to enhance functions of ROS upon transient expression in *N. benthamiana* (Outchkourov et al., 2014). Subsequently ectopic expression of only ROS with the AN-PW gene set was tested, in which case the onset of necrosis was delayed until about 6 dpi (Figure 1B). Leaves expressing AN-PW+ROS+P19 in *N. benthamiana* showed accumulation of red anthocyanins at 3 dpi and only mild necrotic symptoms (Figure 1B). Therefore, in the following experiment the harvest time was set at 5 dpi for AN-PW product analysis and at 13 dpi for measurement of the final products artemisinin (AN) and arteannuin B (AB).

**Co-expression of ROS + AN-PW genes enhances AN sesquiterpene biosynthesis**

AN-PW genes were co-expressed with or without ROS in *N. benthamiana*, such that the relative dosage of the AN-PW genes was constant in the two treatments. Leaves were harvested at 5 dpi and the freeform AN intermediates were extracted in methanol and extracts were analyzed by the LC-trip-quad-MS as described (Ting et al., 2013). Figure 2a shows that both AA and DHAA production were significantly stimulated by the co-expression with ROS. Since we did not have proper standards for (DH)AAA, we used the (DH)AAA peak intensities for semi-quantification. AAA production was also strongly stimulated by co-expression with ROS, while ROS had only limited

![Graph](image-url)

Figure 2. AN and its precursors in leaves of *N. benthamiana* infiltrated with artemisinin pathway (PW) and Rosea1 (ROS) and analyzed by LC-triple-quad-MS at 5 dpi (a and b). AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.
Interaction between Flavonoid and Sesquiterpenoid in *N. benthamiana*

Effect on accumulation of DHAAA (Figure 2b). In order to test the effect of ROS on the glycosides and glutathione conjugates of AN-PW intermediates, leaf tissue was extracted with methanol with 0.137% formic acid and extracts were analyzed by LC-QTOF-MS (Ting et al., 2013). Similar to the free AN-PW product levels, also the AN-PW related glycosides and glutathione conjugates are affected.
gates showed a 2- to 3-fold increase upon co-expression with ROS (Figure 3). Combined, these results show that ROS enhances all AN-PW product levels.

**AN-PW has little effect on major ROS-induced flavonoids**

The effect of AN-PW products on flavonoid PW accumulation was quantified by comparing ROS+5EV+P19 with ROS+AN-PW+P19 and 2EV+P19, in which the ROS gene dosage was kept constant in two of the treatments. For testing the effect of AN-PW activity on flavonoids we analyzed the flavonoid rutin and a single anthocyanin, delphinidin-3-rutinoside (D3R). Previously it was shown that ectopic ROS expression in *N. benthamiana* results in the production of D3R and a number of unknown flavonoid products (Outchkourov et al., 2014). The concentration of rutin was not affected by ROS expression alone or ROS+AN-PW (Figure 4). However, the levels of D3R were greatly enhanced by the ectopic expression of ROS, while additional expression of AN-PW did not have an effect. In contrast, accumulations of the unknown compounds that are induced by ROS in *N. benthamiana* leaves were slightly enhanced by co-expression of ROS+AN-PW. Combined, the results show that ROS stimulates accumulation of AN-PW products, but that the AN-PW has no or little effect on accumulation of major flavonoids in *N. benthamiana* leaves, only minor compounds induced by ROS are slightly en-

![Figure 4. Flavonoids in leaves of *N. benthamiana* infiltrated with artemisinin pathway (PW) and *Rosea1* (ROS) and analyzed by LC-triple-quad-MS at 5 dpi (a and b). a, major flavonoids rutin and delphinidin-3-rutinoside (D3R); b, minor unknown putative flavonoid compounds induced by ROS. Error bar is SE, n=4.](image)
hanced by AN-PW activity. Subsequently, we addressed the question whether the effect of ROS on AN-PW product accumulation is due to improved transport of the AN-PW products to the apoplast of *N. benthamiana* leaves.

**ROS activity mainly enhances sequestration of AN-PW free products inside the cells**

To investigate whether the increased concentration of freeform AN-intermediates upon co-expression of ROS and AN-PW was due to storage in the apoplast, we analyzed the apoplast fluid of agro-infiltrated leaves for free (non-glycosylated) AN-PW product accumulation and compared this to the level of free AN-PW compounds in total extracts of the leaves (see Chapter 3). Because at 5 days there were already signs of necrosis, which could lead to contamination of apoplast fluid with cell content, the apoplast assay was performed at 4 dpi. Apoplastic AA and DHAA only increased 1.5-fold

![Figure 5. artemisinin pathway (PW) intermediate distribution in apoplast and leaf cells of *N. benthamiana* infiltrated with PW and *Rosea1 (ROS)* and analyzed by LC-triple-quad-MS at 5 dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; AB, arteannuin B; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.](image-url)
and 4.8-fold respectively, while inside the leaf AA and DHAA increased 9.8-fold and 6.2-fold respectively, when AN-PW genes were co-expressed with ROS (Figure 5). The effect of ROS on the free form of AAOH and DHAAOH was less pronounced with only a small increase in the apoplast and total leaf extract concentration (Figure 5). Expression of only AN-PW genes in N. benthamiana leaves did not result in detectable levels of AAA or DHAAA in the apoplast, but upon coexpression of AN-PW with ROS, both AAA and DHAAA could be extracted from the apoplast. Rutin has been reported to reside in the cytosol and vacuole, and can be used to test for cell leakage (Markham et al., 2001; Marrs et al., 1995; Ökmen et al., 2013). Control measurements of apoplast fluid indicated that no rutin is present, indicating that (DH)AAA in the apoplast is not from leaking cells (Figure S1). Overall, the results indicate that ectopic expression of ROS may affect sequestration of free AN-PW products inside the cells, but that ROS has little effect on transport capacity of (DH)AA. However, ROS does seem to enhance transport of (DH)AAA to the apoplast.

**ROS does not alter transcriptional activity of AN-PW genes in transient expression assays**

Because ROS enhanced AN-PW product formation, this raised the question whether transcription of the different AN-PW transgenes (which are all under control of the CaMV 35S promoter) is somehow stimulated by ROS. Therefore the transcript levels of AN-PW genes were measured by qRT-PCR, but this showed that ROS does not affect their transcription (Figure 6). Alternatively, the overall increased AN-PW product level in the presence of ROS could be due to effects of ROS on endogenous N. benthamiana genes that encode enzymes involved in the biosynthesis of precursors for the sesquiterpene pathway. Previously, the transcriptional response to ectopic expression of ROS in tomato callus was analysed by RNAseq (Ouchkourova et al., unpublished). We screened these tomato RNAseq data for all known genes related to either the MVA, MEP, terpenoid, flavonoid and shikimate pathways (Figure 12) (Dudareva et al., 2013; Falcone Ferreyra et al., 2012; Maeda and Dudareva, 2012; Petruzza et al., 2013; Saito et al., 2013; Vranová et al., 2013), to determine if any of these genes are affected by ROS. This analysis showed that in tomato callus the expression of a geranylgeranyl diphosphate synthase (GGPPS, Solyc07g064660) was downregulated 68-fold upon expression of ROS, while a mevalonate kinase (MVK, Solyc01g09884) was upregulated 2.2-fold. Downregulation of the GGPPS could potentially reduce competition for substrate flux towards sesquiterpene biosynthesis, while upregulation of MVK could potentially enhance the flux towards sesquiterpene biosynthesis.
Interaction between Flavonoid and Sesquiterpenoid in *N. benthamiana* (Hinson et al., 1997). Therefore, if similar genes in *N. benthamiana* are affected by ROS, this could explain the overall higher AN-PW product level as result of AN-PW+ROS. The expression of the homologs of GGPPS and MVK in *N. benthamiana* was quantified by rt-PCR in the leaf samples expressing either AN-PW or AN-PW+ROS. Results show that *NbMVK* is indeed enhanced 3.0-fold by ROS (Figure 6), while *NbGGPPS* transcripts were not detected in leaves (not shown). Although HMGR was reported to be the rate-limiting step in the MVA pathway (Chappell et al., 1995; Stermer and Bostock, 1987; Tholl and Lee, 2011), *NbHMGR2* was not up-regulated by ROS activity in *N. benthamiana* (Figure 6). Thus, boosting of AN-PW products by ROS expression could be the direct result of boosting of *NbMVK* gene activity by ROS.

![Figure 6](image_url)

**Figure 6.** Relative gene expression of amorphadiene synthase (*ADS*), amorphadiene oxidase (*AMO*), artemisinic aldehyde double-bond reductase (*DBR2*) and mevalonate kinase (*NbMVK*) in leaves of *N. benthamiana* infiltrated with artemisinin pathway (PW) and ROS described as (ROS) and analyzed at 5 dpi. Gene expression analysis was carried out with 4 biological replicates and each of the biological replicates has three technical replicates. Error bar is SE, n=3.
Chapter 4

Blocking flavonoid biosynthesis induced by ROS through CHS\textsuperscript{RNAi} enhances flux through the AN biosynthesis pathway

Besides an effect of ROS on endogenous gene activity, which could affect flux through the AN-PW, it is also of interest to determine if flavonoids themselves affect AN-PW product accumulation. To test for this we expressed the AN-PW+ROS with and without a construct which blocks chalcone synthase (CHS), the first enzymatic step of the flavonoid pathway. CHS transcription was silenced by a CHS\textsuperscript{RNAi} construct (Schijlen et al., 2007). Analysis of the ROS-induced flavonoid D3R and the endogenous flavonoid rutin, showed that the CHS\textsuperscript{RNAi} construct was very effective in suppression of D3R and had only limited effect on rutin (Figure 7 and Figure S2). The silencing of flavonoid biosynthesis (in the presence of ROS) resulted in an increase in both freeform and glycosylated AN-PW related compounds (Figure 7 and Figure 8), indicating that activity of the flavonoid pathway somehow competes with the sesquiterpene pathway in cells. We note that the ratio of different AN-PW related glycosides is different for \textit{N. benthamiana} leaves expressing AN-PW+ROS (Figure 3) and AN-PW+ROS+CHS\textsuperscript{RNAi} (Figure 8), indicating that suppression of the flavonoid biosynthesis alters the AN-PW related leaf chemotype.

![Figure 7](image_url)

**Figure 7.** Targeted analysis of extracts of \textit{N. benthamiana} leaves infiltrated with artemisinin pathway (PW) and \textit{Rosea1} (ROS) with and without CHS\textsuperscript{RNAi} and analyzed by LC-triple-quad-MS at 5 dpi. a, AN intermediates; b, flavonoids. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; AB, arteannuin B; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihy droartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.
Figure 8. AN glycosides and glutathione conjugates in leaves of *N. benthamiana* infiltrated with artemisinin pathway (PW), *Rosea 1 (ROS)* and CHS\textsuperscript{RNAi} and analyzed by LC-triple-quad-MS at 5 dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; AB, arteannuin B; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.
**AaLTP3 and AaPDR2** have no influence on the effect of ROS on AN-PW product accumulation, but alter the glycoside chemotype of *N. benthamiana* leaves.

Because previously we have shown that AaPDR2 and AaLTP3 enhance sequestration of free AN-PW products in the apoplast (Chapter 3), we subsequently tested the effect of ROS on AN-PW product accumulation in the presences of AaPDR2 and AaLTP3. To this end, the product profile in leaves agro-infiltrated with AN-PW+ROS+AaPDR2+AaLTP3 was compared with that of AN-PW+EV+AaPDR2+AaLTP3. Note that, due to the additional two gene constructs (35S::AaPDR2 and 35S::AaLTP3) in these experiments the relative dosage of AN-PW genes was slightly lower than in those used in Figures 5 and 7. Results show again that ROS has more effect on sequestration of (DH)AA inside the leaf cells than on accumulation in the apoplast (Figure 9 and Figure S3), and, similar as in Figure 5,

![Figure 9](image-url)

**Figure 9.** AN intermediate distribution in the apoplast and in leaf extracts of *N. benthamiana* infiltrated with artemisinin pathway (PW), *Rosea1* (ROS) and the *Artemisia annua* transporter genes (P-DR2, LTP3) and analyzed by the LC-triple-quad-MS at 5 dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; AB, arteannuin B; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.
Figure 10. artemisinin pathway related glycosides and glutathione conjugates in leaves of N. benthamiana infiltrated with artemisinin pathway (PW), Rosea1 (ROS) and transporter genes (PDR2, LTP3) analyzed by LC-triple-quad-MS at 5 dpi. AAOH, artemisinic alcohol; AAA, artemisinic aldehyde; AA, artemisinic acid; AB, arteannuin B; DHAAOH, dihydroartemisinic alcohol; DHAAA, dihydroartemisinic aldehyde; DHAA, dihydroartemisinic acid. Error bar is SE, n=4.
ROS stimulated accumulation of (DH)AAA in the apoplast. ROS also had a stimulating effect on glycosylated AN-PW product accumulation, as indicated by the increased accumulation of (DH)AA and (DH)AAOH glycosides (Figure 10). The effect of ROS on overall AN-PW product accumulation was not further influenced by coexpression of AaPDR2 and AaLTP3. However, we did note that the ratio of different AN-PW related glycosides is different for N. benthamiana leaves expressing AN-PW+ROS (Figure 3) and AN-PW+ROS+AaPDR2+AaLTP3 (Figure 10), indicating that the co-expression with AaPDR2+AaLTP3 changes the AN-PW related leaf chemotype.

Figure 11. Final product accumulation and quantification of necrosis at different time points after infiltration with artemisinin pathway (PW*), Rosea1 (ROS) and Artemisia transporter genes (PDR2, LTP3). a, Relative electrolyte leakage assay from 5 dpi to 18 dpi, b, arteannuin B(AB) detection from 11dpi to 18 dpi; c, artemisinin (AN) detection from 11 dpi to 18 dpi. Error bar is SE, n=4.
**ROS enhances AN and AB accumulation from AN-PW+AaPDR2 + AaLTP3 expression in N. benthamiana leaves**

Previously we have shown that accumulation of the final products AN-PW products AN and AB only occurs after onset of necrosis in *N. benthamiana* leaves expressing the AN-PW genes (Chapter 3). Since leaves expressing AN-PW+ROS show visible signs of necrosis at day 6 we semi-quantified necrosis levels by ion-leakage assay (Bouwmeester et al., 2014). Results show that agro-infiltration of EV or AN-PW+AaPDR2+AaLTP3 both resulted in ~20% ion leakage at 5 dpi, with a gradual increase to 40% at 18 dpi for EV, but to 80% for leaves expressing the AN-PW+AaPDR2+AaLTP3 genes. When AN-PW+AaPDR2+AaLTP3 were coexpressed with ROS the leakage at 5 dpi was already 40%, but was more or less steady at 60% from 7-13 dpi (Figure 11a). Subsequently we quantified AN and AB accumulation in *N. benthamiana* leaves agro-infiltrated with the different gene sets. The highest levels of AN and AB were reached at 13 dpi and product accumulation was highest for the expression of AN-PW+ROS+AaPDR2+AaLTP3+CHSRNAi and AN-PW+ROS+AaPDR2+AaLTP3. AN or AB accumulation did not correlate with the level of necrosis, which was highest for leaves expressing PW+AaPDR2+AaLTP3+EV at 13 dpi (Figure 11 and Figure S4).

**Discussion**

**ROS activity enhanced the AN metabolic flux**

Here we have tested the interaction between the induced flavonoid biosynthesis pathway of *N. benthamiana* and an ectopically expressed AN sesquiterpene biosynthesis pathway. ROS activity resulted in a substantial and similar increase in both free form and glycosylated AN-PW products (Figure 2 and Figure 3). The effect of ROS on AN-PW+AaPDR2+AaLTP3 (Figure 9) was similar to that of ROS on AN-PW alone (Figure 5 and Figure 11). Both results suggest that overall AN-PW activity is increased by ROS, while ROS does not affect transgene transcript levels (Figure 6). However, analysis of the expression of (precursor) pathway genes demonstrated a 2-fold upregulation of *NbMVK* (Figure 6) by ROS, which could result in an increased precursor flow into the AN-PW and which thus could account for the higher content of AN-PW products in the presence of ROS (Figure 2, 3 and 11). Previously, *HMGR* was reported as the rate-limiting step in the biosynthesis of isoprenoid compounds derived from the cytosolic MVA pathway (Chappell et al., 1995; Stermer and Bostock, 1987; Tholl and Lee, 2011). Indeed, co-infiltration of a
tHMGR with ADS resulted in strongly enhanced amorphadiene production in *N. benthamiana* (Van Herpen et al., 2010). However, ROS did not enhance transcriptional activity of endogenous *NbHMGR* (Figure 9). Since the AN-PW genes include tHMGR we propose that under transient expression conditions *HMGR* is no longer limiting for MVA pathway activity and that other enzymatic steps such as MKV become limiting. This could then explain why upregulation of endogenous MKV can further enhance the flux through the AN-PW (Hinson et al., 1997). Next to the effect of ROS on MKV expression, it is still possible that as result of ROS activity and higher flavonoid levels in the *N. benthamiana* leaves, proteins are less prone to oxidative damage. Indeed ROS expression does seem to lower and stabilize cell leakage between 9 dpi and 13 dpi (Figure 11a) and it could be that the increased flavonoid levels in these leaves also protect enzymes of the AN-PW, resulting in more prolonged activity, which than could contribute to increase AN-PW product levels.

**Competition between the flavonoid and terpene pathway**

In the test to determine if flavonoid compounds themselves affect AN-PW product accumulation, we found that inhibition of flavonoid biosynthesis, resulted in a 20% increase in AN-PW product levels (Figure 7 and Figure 8). However, final product accumulation of AB and AN at 13 dpi was not significantly affected by *CHS*R**RNAi** in the present of ROS (Figure 11). The effect of ROS+CHS*R**RNAi** on AN-PW product intermediates at 5 dpi may be explained in different ways. For instance, phosphoenol pyruvate feeds both into the AN-PW and the flavonoid pathway (see Figure 12) and both pathways compete with each other for this compound. Thus the lower competition for phosphoenol pyruvate upon coexpression of *CHS*R**RNAi** could result in the increase in AN-PW product formation. Alternatively, it could be that upon *CHS*R**RNAi** expression, modifying P450 enzymes of the flavonoid biosynthesis pathway no longer consume NADPH nor utilize cytochrome P450 oxidoreductase (CPR) activity for P450 regeneration (Renault et al., 2014). This would leave more NADPH and CPR activity for the P450s of the AN-PW. We note that the apparent competition between the AN-PW and flavonoid biosynthesis PW could also be the result of exceptional high expression of the AN-PW genes in the context of the transient expression in *N. benthamiana* and therefore may be of less relevance to situations with more natural levels of PW gene expression.
Interaction between Flavonoid and Sesquiterpenoid in *N. benthamiana*

We previously defined the chemotype of the *N. benthamiana* leaves expressing the AN-PW as the ratio between DHAA-hexose and AA-hexose, as quantified by peak intensities in MS chromatograms of leaf extract analysis (Ting et al., 2013). Table 1 lists this *N. benthamiana* leaf chemotype for the different treatments. Seasonal changes in growth conditions may affect flux through the PW and therefore in table 1 experiments executed at different times of the year.}

**AN-PW chemotype in *N. benthamiana* leaves affected by ROS and CHS**

We previously defined the chemotype of the *N. benthamiana* leaves expressing the AN-PW as the ratio between DHAA-hexose and AA-hexose, as quantified by peak intensities in MS chromatograms of leaf extract analysis (Ting et al., 2013). Table 1 lists this *N. benthamiana* leaf chemotype for the different treatments. Seasonal changes in growth conditions may affect flux through the PW and therefore in table 1 experiments executed at different times of the year.
the year are indicated. The results show that ROS lowers the DHAA:AA leaf chemotype in two separate set of experiments (Figure 3 and Figure 10).

Table 1 ratio of DHAA-hex : AA-hex for the different gene combinations. Treatments performed at different times of the year are indicated by same number. AN-PW= V5+AMO+DBR2+ALDH1.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>DHAA-hex : AA-hex</th>
<th>Relative PW gene dosage</th>
<th>Harvest day</th>
</tr>
</thead>
<tbody>
<tr>
<td>1AN-PW +EV+ p19 (this chapter)</td>
<td>3.21 (Figure 3)</td>
<td>4/6</td>
<td>(dpi)</td>
</tr>
<tr>
<td>1AN-PW+ROS +p19</td>
<td>1.74 (Figure 3)</td>
<td>4/6</td>
<td>5</td>
</tr>
<tr>
<td>2AN-PW+ROS+CHSRNAi</td>
<td>1.07 (Figure 8)</td>
<td>4/6</td>
<td>5</td>
</tr>
<tr>
<td>2AN-PW+ROS+EV</td>
<td>1.74 (Figure 8)</td>
<td>4/6</td>
<td>5</td>
</tr>
<tr>
<td>3AN-PW+EV+ AaPDR2+AaLT-P3+p19</td>
<td>0.95 (Figure 10)</td>
<td>4/8</td>
<td>5</td>
</tr>
<tr>
<td>3AN-PW+ROS +AaPDR2+AaLT-P3+p19</td>
<td>0.61 (Figure 10)</td>
<td>4/8</td>
<td>5</td>
</tr>
</tbody>
</table>

The ratio of DHAA:AA decreased from 3.21 to 1.74 by the ROS activity, and even further by additional coexpressing with CHSRNAi (to 1.07). Coexpression of the AN-PW with the Artemisia transporter genes combination AaP-DR2+AaLTP3 resulted in a DHAA:AA ration of 0.95, while coexpression of the AN-PW with transporters and ROS lowered the ratio to 0.61 (Table 1). In most of these treatments the relative gene dosage of the AN-PW was similar and also the inhibitor of transgene silencing (P19) does not seem to affect the leaf chemotype (e.g. compare AN-PW+ROS+P19 with AN-PW+ROS+EV). The effect of the transporter genes indicates that AaP-DR2+AaLTP3 have more effect on AA and therefore they are more efficient in stimulating the flux through the AA branch than through the DHAA branch of the AN-PW. Besides an effect of the treatments on the DHAA:AA chemotype, we also note difference in the ratio of the different types of AA-hex glycosides in *N. benthamiana* leaves expressing AN-PW+EV or AN-PW+ROS (Figure 3). At present it is not clear how ROS is affecting DHAA:AA or glycosylation chemotype. It could be that ROS affects expression of glycosyltransferases, which preferably act on AA. Alternatively, it could be due to a direct interaction between flavonoid and terpene molecules that somehow affects AA more than DHAA. However, this is not supported by the experiment with inhibition of flavonoid biosynthesis in the presence of ROS, which leads to even stronger effect on AA compared to DHAA. Although P19 did not affect the overall DHAA:AA chemotype, the ratio of AA-Hex2-Mal to AA-
Hex$_2$ is higher in leaves expressing AN-PW+ROS+CHS$^{RNAi}$ compared to that in leaves expressing AN-PW+ROS+P19. This indicates that either silencing or inhibition of flavonoid biosynthesis activity affects glycosylation. More careful investigation in the future is needed to determine whether and how the ratio of AA glycosides is affected by AN-PW agroinfiltration gene dosage, harvest dpi, ROS and or inhibition of transgene silencing by P19.

**Concluding remarks**

In conclusion, we here report that the TF ROS not only stimulates flavonoid biosynthesis genes but also stimulates gene expression for precursor input into the AN sesquiterpene pathway, presumably by upregulating MVK. In addition we have shown that the stimulated flavonoid biosynthesis activity competes with total sesquiterpene production at 5 dpi, but does not affect overall accumulation of AN/AB at 13 dpi in planta. The effect of activated flavonoid biosynthesis pathway activity on the type of AN-PW glycosylated products may be the consequence of interaction between flavonoid and sesquiterpenoid and needs further investigation.

**Acknowledgement**

We thank Bert Schipper for assistance in LC-QTOF-MS analysis, Kristyna Floko-va for help in LC-triple-quad-MS analysis, Arnaud Bovy for providing the CHS$^{RNAi}$ construct and Klaas Bouwmeester for support on the electrolyte leakage assay.

**References**


catalyzed by 3-hydroxy-3-methylglutaryl coenzyme A reductase a rate-limiting step for isoprenoid biosynthesis in plants? Plant Physiology 109, 1337-1343.


Interaction between Flavonoid and Sesquiterpenoid in *N. benthamiana*

375, 397-400.


Chapter 4


**Supporting information**

**Table S1 Primers**

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence from 5’ to 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>35S-prom-F</td>
<td>GCTCCTACAAATGCCATCA</td>
</tr>
<tr>
<td>35S-prom-R</td>
<td>GATAGTGGGATTGTGCGTCA</td>
</tr>
<tr>
<td>NbHMGR2-F</td>
<td>CCAGGGATAACATGATGATTC</td>
</tr>
<tr>
<td>NbHMGR2-R</td>
<td>CAGCCAAGCGTATGTTGCG</td>
</tr>
<tr>
<td>Ubi3-F</td>
<td>GCCGACTACACATCCAGAAGG</td>
</tr>
<tr>
<td>Ubi3-R</td>
<td>TGCAACACAGCGAGCTTAACC</td>
</tr>
<tr>
<td>NbMVK_3F</td>
<td>ATGGCGTGGAGATTTCAATCC</td>
</tr>
<tr>
<td>NbMVK_3R</td>
<td>GGCACGTTTTGATGTTTCT</td>
</tr>
<tr>
<td>NbGGPPS_1F</td>
<td>CTGGGCTCTATCGAGTCTG</td>
</tr>
<tr>
<td>NbGGPPS_1R</td>
<td>CATGTGCTGGGTGTTAAG</td>
</tr>
</tbody>
</table>
Figure S1 LC-orbitrap-MS chromatogram of \([M+H]^+ = 611.160\) in apoplastic fluid of *N. benthamiana* infiltrated with *AN-PW+ROS+P19*. a, apoplast sample and remaining leaf sample; b, Rutin MS spectrum of peak at 21.37 for *N. benthamiana* leaves infiltrated with *AN-PW+ROS*.

Figure S2 Characterization of flavonoid production in leaves of *N. benthamiana*. a: LC-orbitrap-MS chromatogram of \([M+H]^+ = 611.160\) at retention time=8.505min, b: MS spectrum of D3R of peak at 8.505 for *N. benthamiana* leaves infiltrated with *AN-PW+ROS+EV*. 
Interaction between Flavonoid and Sesquiterpenoid in *N. benthamiana*

Figure S3 LC-orbitrap-MS chromatogram of AN-PW+ROS+PDR2+LTP3+P19 at [M+H]^+ = 611.160 in apoplastic fluids of *N. benthamiana*. A, apoplast sample and remaining leaf sample; b, Rutin MS spectrum of peak at 21.37 for *N. benthamiana* leaves infiltrated with AN-PW+ROS.

Figure S4 AN-intermediates accumulation from 11 dpi to 18 dpi in *N. benthamiana* infiltrated with artemisinin pathway (PW*), *Rosea1* (ROS) and Artemisia transporter genes (*PDR2*, *LTP3*); Error bar is SE, n=4.
Chapter 5

Higher terpene production from ectopic expressed terpene synthase genomic clones than from terpene synthase cDNAs in transgenic *Arabidopsis*

Bo Wang, Melida Rojas Chango, Thierry Delatte, Harro J. Bouwmeester, and Alexander R. van der Krol

Laboratory of Plant Physiology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
Abstract

Intron splicing and splicing complexes remaining on processed mRNA may stimulate gene expression in eukaryotes in multiple ways. However, in metabolic engineering of plants, mostly ectopic expression of cDNAs under control of a strong promoter is used. Here we tested whether ectopic expression of a genomic clone consistently results in higher expression than the corresponding cDNA when under control of the same promoter. For this purpose the genomic and cDNA clone of the sesquiterpene synthases Caryophyllene Synthase (gCST and cCST) and Amorphadiene Synthase (gADS and cADS) were isolated from Arabidopsis thaliana and Artemisia annua respectively and expressed from the CaMV 35S promoter in Arabidopsis. For each expression construct 6-10 independent stable transformed plants were generated and overexpression (OE) in leaves was quantified. Average CST expression in gCST-OE lines is 27-fold higher than in cCST-OE lines, while average ADS expression in gADS-OE lines is 11-fold higher than in cADS-OE lines. The headspace volatiles of rosette plants were quantified by GC-MS, showing that average β-caryophyllene emission of the gCST overexpression (OE) lines is 58-fold higher than that of cCST-OE lines, while average amorphadiene emission of the gADS-OE lines is 5.4-fold higher than that of cADS-OE lines. Therefore, for both CST and ADS results indicate that genomic clones with introns boost transgene expression compared to the cDNA. The four expression constructs were also tested in transient expression assays using agro-infiltration of N. benthamiana leaves. However, in this context the stimulating effect of intron containing transgenes is largely lost.

Keywords

Amorphadiene, β-caryophyllene, cDNA, Genomic DNA, Introns, Sesquiterpene
Introduction

Introns in eukaryotic pre-mRNAs are removed by splicing complex to generate mature mRNA. During intron splicing, exon junction complexes (EJC) bind to the pre-mRNA and next to a role in joining the 5’and 3’ intron splice junctions sites, these complexes may have a role in maturated mRNA export and in decay of faulty nonsense-RNAs (Le Hir et al. 2001). Both in plants, insects, mice and humans some introns in genes are known to enhance expression and this effect has been named Intron-Mediated Enhancement (IME) (Mascarenhas et al. 1990). Introns and intron-splicing may stimulate the transcription in multiple ways (Le Hir et al. 2003; Niu and Yang 2011). For instance, (1) introns may contain transcriptional enhancer elements, which help stimulate expression from the upstream promoter. (2) the intron-splicing complex may interact with the RNA polymerase II and enhance reloading of the transcriptional machinery onto the promoter (Moabbi et al. 2012). (3) the intron splicing complex may interact with the 7-methyl guanosine (m7G) 5’-end capping and 3’-end polyadenylation, enhancing the splicing efficiency and polyadenylation on either side of the terminal exon (Gunderson et al. 1998; Komarnitsky et al. 2000; Lewis and Izaurriéde 1997; Proudfoot et al. 2002; Vagner et al. 2000). In some instances, (4) intron splicing affects exon RNA editing before splicing and may thus affect the actual protein product (Le Hir et al. 2003; Reenan 2001; Sowden and Smith 2001). (5) Residual protein complexes on the intron splice site on mature mRNA may promote export of mRNA from the nucleus (Legrain and Rosbash 1989; Luo and Reed 1999). (6) In addition, residual protein complexes on the intron splice site on mature mRNA may increase translational efficiency of mRNAs by affecting mRNA stability (Le Hir et al. 2003). (7) It has been demonstrated that exon splice-junctions near the 5’-end of some transcripts stimulate translation relative to an intronless mRNA, whereas splice-junctions in the 3’-UTR significantly reduce translation in Xenopus oocytes (Braddock et al. 1994; Matsumoto et al. 1998). Models have been developed to predict the enhancement of introns on gene expression, and the effect of an intron sequence and position on gene expression can thus be calculated and expressed as an IME score (Parra et al. 2011; Rose et al. 2008).

In our previous research on engineering of sesquiterpene biosynthesis in plants we produced Arabidopsis plants with overexpression of the Arabidopsis caryophyllene synthase gene (CST), using a genomic clone (gCST). While a previously reported Arabidopsis line expressing CST cDNA (cCST) emits ~ 4.0 ng·g⁻¹ FW·h⁻¹ (Huang et al. 2012), our line with overexpression of the gCST emits ~ 80 ng·g⁻¹ FW·h⁻¹ (Delatte, unpublished). In both cases
the transgene was driven by the same CaMV 35S promoter and both the cCST and gCST encode for identical proteins. Because in metabolic engineering of terpene production in heterologous plant hosts we aim for high production, we set out to determine whether there is a consistent higher expression from intron containing genes, as additional strategy to boost ectopic terpene production in plants. Here we show a proof of principle for the Arabidopsis CST and another sesquiterpene biosynthesis gene, encoding the first committed step in artemisinin biosynthesis in Artemisia annua, Amorphadiene synthase (ADS) (Bertea et al. 2006; Bouwmeester et al. 1999; Ting et al. 2013; Van Herpen et al. 2010). Indeed, in stable transformants the genomic fragment consistently results in higher expression and terpene production, but this boosting effect of an intron-containing gene is more or less lost in the context of transient expression in N. benthamiana.

Materials and methods

Quantitative reverse transcriptions PCR (qRT-PCR)

RNA was extracted from rosette leaves using RNeasy Plant Mini kit (Qiagen, Germany) with RNAase-Free DNase digestion (Qiagen, Germany). The quality and quantity of the RNA were assessed using the agarose gel electrophoresis and Nanodrop (Nanodrop Technologies, USA). cDNA was transcribed from 1000 ng of RNA using iScript cDNA Synthesis kit (BioRad, USA). Gene specific primer pairs for the qPCR analysis were actin (5’- CTA AGC TCT CAA GAT CAA AGG CTT A -3’/ 5’- ACT AAA ACG CAA AAC GAA AGC GGT T -3’), ADS-qF/ ADS-qR (5’- GGA GTA TGC CCA AAC CTT GA -3’/ 5’- TCG TCT CCC ATA CGT GTG AA -3’) and CST-qF/ CST-qR (5’- TCA AGG ACG AAA ATG GGA AG -3’/ 5’- ATT GCG AGA TGA GGG CTA GA -3’). The primer pairs of house-keeping gene are NbUbi3_q/NbUbi3_qR (5’- GCC GAC TAC AAC ATC CAG AAG G-3’/ 5’-TGCAACACACGCGTGTTAAACC) and Atactin2_qF / Atactin2_qR (5’-CTAAGCTCTCAAGATCAAAGGCTTA-3’/ 5’-ACTAAAAACGCAAAAAGCGTTT -3’) were used for Nicotiana benthamiana and Arabidopsis samples, respectively. The qRT-PCR was performed on the iCycler iQ5 system (BioRad) as following program: 95 °C for 10 min, 40 cycles of 95 °C for 10 s, 55° C for 30 s, 95 °C for 1 min and followed by melting curve analysis from 65 to 95 °C. Relative gene expression was calculated using ∆∆CT.
**Vector construction**

**35S:gCST** and **35S: cCST.** Genomic DNA was extracted from *Arabidopsis* (Col-0) leaves by the cetyltrimethylammonium bromide (CTAB) method. *gCST* and *cCST* were amplified separately using primer pair **CST-F** (5’- TTT TCC ATG GGG AGT GAA GTC AAC CGTC -3’) and **CST-R** (5’- TTG GTT GCG GCC GCT CAA ATG GGT ATA G -3’). The NcoI and NotI restriction sites were introduced and cloned into the Impact Vector pIV1A2.1, which contains the CaMV35S promoter and Rbcs terminator ([www.impactvector.com](http://www.impactvector.com)). The resulting pIV1A2.1/gCST and pIV1A2.1/cCST were cloned into pBinPlus vector using the LR reaction. The pBinPlus containing 35S: gCST and 35S: cCST were transformed to *A. tumefaciens* AGL-0 using the electroporation, separately.

**35S:gADS** and **35S: cADS.** Genomic DNA was extracted from A. annua flowers by the CTAB method. *gADS* and *cADS* were amplified separately using primer pair **ADS-F** (5’- CTA CTC TAG AAT GTC ACT TAC AGA AGA -3’) and **ADS-R** (5’- GAC AAA CCA TGG TCA TAT ACT CAT AGG -3’). The XbaI and NcoI restriction sites were introduced and cloned into the Impact Vector pIV1A2.1, which contains the CaMV35S promoter and Rbcs terminator ([www.impactvector.com](http://www.impactvector.com)). The resulting pIV1A2.1/gADS and pIV1A2.1/cADS were cloned into pBinPlus vector using the LR reaction. The pBinPlus containing 35S:gADS and 35S:cADS were transformed into *A. tumefaciens* AGL-0 using the electroporation.

**Transformation of Arabidopsis**

The expression constructs 35S:gCST, 35S:cCST, 35S:gADS and 35S:cADS were transformed into *Arabidopsis* Col-0 using the floral dipping method as described (Harrison et al. 2006). Transformed seeds were selected on MS medium containing kanamycin and 6-10 independent transformed lines for each expression construct were selected after 5 days on selective medium and transfer to soil for further analysis. Transformation was confirmed by qPCR on cDNA isolated from rosette leaves harvested from 3 weeks old primary transformants using primer pairs specific to each of the four expression constructs.

**Plant growth conditions**

Transformed seeds were sown on the MS medium with the kanamycin and stored in the dark at cold room (4°C). After 2 days, seeds were transferred in the room temperature until germination, and then sown in soil
in growth chamber (12 h light at 22 °C in the day and 12 h dark at 18 °C in the night). 3 week old plants were prepared for the further analysis.

**Transient expression in N. benthamiana**

The agro-infiltration was carried out in the 4th leaf of 5 week old *N. benthamiana* plants as described (Van Herpen et al., 2010). Briefly, *A. tumefaciens* strains transformed with individual expression constructs were infiltrated into *N. benthamiana* leaves using a 1 ml syringe without needle. Dosage compensation of infiltrated *A. tumefaciens* was done by diluting with an *A. tumefaciens* strain carrying an empty vector (EV) construct. Agrobacterium with a P19 expression construct was added to suppress post-transcriptional gene silencing (Voinnet et al. 2003). Infiltrated leaves were harvested at the 5 days post infiltration (dpi).

**Headspace trapping and analysis of volatiles by GC-MS**

Headspace analysis of stable transformed *Arabidopsis* was done using 3 weeks old primary transformants grown on soil. The volatiles were collected use dynamic headspace trapping as described (Houshyani et al. 2013). After *Arabidopsis* were enclosed in 1 L glass jar, the vacuum pump was used to draw air through the glass jar and the incoming air was purified through Tenax TA at 200 ml/min inlet. The outlet was trapped at 300 ml/min to collect the volatiles. The gCST-OE lines were trapped 4 hours from 10.00-13.00 at 22 °C. Wildtype plants and rest of the OE lines were trapped 24 hours (12 h light at 22 °C in the day and 12 h dark at 18 °C in the night). The headspace samples were analyzed by Thermodesorption GC-MS with a thermal desorber (Unity, Markes International Limited) and a Trace GC Ultra (Thermo Electron Corporation) coupled with DSQ mass spectrometer (Thermo Electron Corporation) as described (Houshyani et al. 2013). The GC-MS results were analyzed by the Xcalibur software (Thermo, Waltham). Quantification of terpene levels was accomplished by determining the peak area of the characteristic m/z (133) for β-caryophyllene, the characteristic m/z (93) for α-humulene, the characteristic m/z (119) for α-copene and the characteristic m/z (189) for amorphadiene. As the comparison, 5 open and 5 more closed *Arabidopsis* flowers were trapped for 24 h by twister in 10 mL vial in a glass vial and measured on GC-MS.

**Quantification of untargeted metabolites by UPLC-MS analysis**

After headspace analysis rosette leaves 3 and 4 were harvested from
Introns in Terpene Synthase Genes Boost Sesquiterpene Yield

3 week old primary transformants and immediately frozen in liquid Nitrogen. From this, 100 mg of frozen ground samples were extracted using 300 ul 99.9% MeOH with 0.133% formic acid. Analysis of non-volatile compounds in transgenic *Arabidopsis* rosette extracts was carried out by LC-QTOF-MS as described (Ting et al. 2013). The raw mass data was processed using metalign (www.metalign.nl). Multiple mass signals from the same compound were grouped using MSClust software (biotools.wur.net.nl). Significance of average intensity differences of aligned mass-signals between samples were assessed using student t-test (p< 0.05).

**Results**

**Isolation of gCST and gADS genomic sequences**

In order to investigate whether the introns affect the sesquiterpene production, the genomic DNA and cDNA of *Arabidopsis* CST (gCST and cGST) and *Artemisia annua* ADS (gADS and cADS) were isolated by PCR on (Arabidopsis or A. annua) genomic DNA or (Arabidopsis or A. annua) leaf cDNA, respectively. The fragments (gCST, cCST, gADS, cADS) were cloned downstream of the CaMV 35S promoter in a binary expression vector and *Arabidopsis* plants were transformed with the expression constructs 35S:gCST, 35S:cCST, 35S:gADS and 35S:cADS by the conventional floral dipping method (Harrison et al. 2006). Sequence analysis of the genomic clones showed that both the *Arabidopsis* gCST and A. annua gADS gene have six introns (Fig. 1A and B). The *Arabidopsis* TAIR database indicates two splice variants for CST mRNA, with a small amino acid sequence difference in exon 5 (Fig. 1A). However, the alternative CST splice product was not detected in our leaf cDNA, also not using primer pairs specific for the alternative splice product (Fig. 1A). The intron sequences of gCST and gADS were evaluated using IMEter v2.0 (http://korflab.ucdavis.edu/cgi-bin/web-imeter2.pl) (Parra et al. 2011). The analysis indicates that intron 4 of gCST and intron 2 of gADS have the highest IME score of the six introns in each gene (Table 1). An IME score of 20 predicts a 5-fold increase in expression (Parra et al. 2011; Rose et al. 2008). Since the highest IME score is lower, but similar for both gCST and gADS, the prediction is that both may show a similar, but lower than 5-fold stimulation of gene expression, when compared to expression of their respective cDNAs.
Table 1. IMEter score of individual introns for CST and ADS genes

<table>
<thead>
<tr>
<th>Intron</th>
<th>AT5G23960.1</th>
<th>AT5G23960.2</th>
<th>ADS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.70</td>
<td>1.70</td>
<td>0.97</td>
</tr>
<tr>
<td>2</td>
<td>3.02</td>
<td>3.02</td>
<td>13.12</td>
</tr>
<tr>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>1.79</td>
</tr>
<tr>
<td>4</td>
<td>10.27</td>
<td>11.93</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>6</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Fig.1. Sequence comparison of gDNA and mRNA of caryophyllene synthase (CST) and amorphadiene synthase (ADS) genes. a, CST; b, ADS.

Introns increased the transcription expression in Arabidopsis

For each of the four expression (35S:gCST, 35S:cCST, 35S:gADS and
Introns in Terpene Synthase Genes Boost Sesquiterpene Yield

35S:cADS) constructs 6-10 independent Arabidopsis transformants were selected, based on seedling resistance to kanamycin. After 5 days on selective medium, transformed seedlings were transferred to soil and plants were grown for 3 weeks at which RNA was isolated from two rosette leaves for quantification of transgene expression level by qPCR. On average the transcription level of CST was increased 74-fold for lines expressing cCST, compared to expression of the endogenous CST in untransformed Arabidopsis. In contrast, average transcription level of CST was 1978-fold higher for lines expressing gCST, compared to endogenous CST activity (Fig. 2A). Therefore, on average, expression in gCST transformants is 26.7-fold higher than in cCST transformants. Similarly, ADS transcript levels in individual gADS-OE lines were consistently higher than the ADS transcript levels in cADS-OE lines (Fig. 2B), with average expression level of ADS in gADS-OE lines 11-fold higher than the average expression in cADS-OE lines. Therefore, for both CST and ADS, expression was higher from the intron-containing genomic DNA than from the corresponding cDNA. For
both the CST and ADS the increase in expression of the genomic clone compared to the cDNA clone was higher than predicted by the IME score.

**Introns of CST increased the sesquiterpene emission in Arabidopsis**

To test how the increased expression in transgenic gCST-OE and gADS-OE lines (compared to cCST-OE and gADS-OE lines, respectively) is affecting sesquiterpene emission, volatiles were trapped in the headspace of 3-week old primary transformants and analyzed by GC-MS. β-caryophyllene emission in eight selected gCST-OE lines was consistently higher than in the six selected cCST-OE lines (Fig. 3A). The average β-caryophyllene emission in the gCST-OE lines was 429.7-fold higher than in WT plants, while average β-caryophyllene emission in the cCST-OE lines was 7.4-fold higher than in WT plants. Therefore, average β-caryophyllene emission in gCST-OE lines was 58-fold higher than in cCST-OE lines, which is substantially more than the 26.7-fold difference in CST transcript levels between these two groups of lines (Fig. 3A and Fig. 4A-C). In individual lines, the amount of β-caryophyllene emitted matched well with the CST mRNA level. For instance, gCST-3 OE showed highest β-caryophyllene emission (Fig. 3A) and also the highest CST mRNA level (Fig. 2A). Besides β-caryophyllene, CST enzyme activity also produces some minor products, i.e. α-humulene and α-copaene (Chen et al., 2003). Figure 4A shows representative mass-traces of the headspace analysis of WT, gCST-OE and cCST-OE lines and the mass-spectra of the three products from CST enzyme activity. Although a peak is visible in the headspace of WT and cCST-OE at the position of the α-copaene peak in gCST-OE headspace (peak 4, 5, 3 respectively, Fig. 4A), the mass spectrum of peak 4 in WT is not from α-copaene. The predicted molecular composition of compound 4 in WT is C_{15}H_{24}, thus compound 4 is most likely another sesquiterpene from endogenous enzyme activity (Fig. 4B). In cCST-OE lines the mass spectrum of peak 4 position shows a combination of the mass spectra of the unknown compound and the mass spectrum of α-copaene (Fig. 4B). Curiously, in the gCST-OE lines the mass spectrum of the unknown compound is no longer detected and only the mass spectrum of α-copaene is detected. This may indicate that high ectopic expression in gCST-OE plants results in such high CST enzyme activity that it outcompetes the other endogenous sesquiterpene synthase which produces unknown compound 4. Therefore, the semi-quantification of the α-copaene peak intensity may be overestimated in cCST-OE lines (Fig. 3C). For the gCST-OE lines the increase in β-caryophyllene production was higher than the increase in CST transcript level compared with that in cCST-OE lines, suggesting that addition post-tran-
Fig. 3. Volatiles emission of caryophyllene synthase overexpression (CST-OE) and amorphadiene synthase overexpression (ADS-OE) lines in *Arabidopsis*. a, β-caryophyllene emission; b, α-humulene emission; c, α-copaene emission; d, amorphadiene. The headspace of WT plants (Col-1, Col-2), cCST-OE lines, cADS-OE lines and gADS-OE lines was trapped for 24 hours, while the headspace of gADS-OE lines was trapped for 4 hours, using Tenax TA.
Fig. 4. GC-MS mass-traces for cDNA of caryophyllene synthase overexpression (cCST-OE) lines, genomic caryophyllene synthase overexpression (gCST-OE) lines at m/z=119, for cDNA of amorphadiene synthase overexpression (cADS-OE) lines and genomic amorphadiene synthase overexpression (gADS-OE) lines at m/z=189 in Arabidopsis. a, chromatogram of WT, for cCST-OE lines, gCST-OE line at m/z=189; Mass spectrum for compounds: 1, β-caryophyllene, 2, α-humulene and 3, α-copaene, 4, unknown, 5, unknown; b, chromatogram of cADS-OE and gADS-OE lines at m/z=189; Mass spectrum for compound: 6, amorphadiene. Note: mass-trace at m/z=119 was shown to clear show the α-copaene peak.
scriptional mechanisms may contribute to the higher emission rates from plants expressing genomic clones compared with plants expressing cDNAs.

Analysis of the headspace volatiles of the gADS-OE and cADS-OE lines shows that amorphadiene emission was also consistently higher in gADS-OE lines (Fig. 3D and Fig. 4F-G). The average amorphadiene emission in the gADS-OE lines was 5.4-fold higher than in cADS-OE lines (Fig. 3D), while the average ADS transcript level in gADS-OE lines was 11-fold higher than in cADS-OE lines (Fig. 2B). The A. annua ADS in Arabidopsis therefore seems to show less benefit from a post-transcriptional intron effect on overall gene expression, than the post-transcriptional intron effect on the Arabidopsis CST gene.

We note that (based on peak area in mass-traces), amorphadiene emission is only 1.2% of the average β-caryophyllene emission in transgenic Arabidopsis (Fig. 3A-D), while standards of β-caryophyllene and amorphadiene indicate that both compounds are detected with similar efficiency. We therefore considered that amorphadiene may be converted by endogenous Arabidopsis Cytochrome P450 or other enzymes to non-volatile amorphadiene derived products, just as in A. annua. In A. annua, amorphadiene is modified by an Amorphadiene Oxidase P450 enzyme (ADO) to non-volatile artemisinic acid (Teoh et al. 2006). Moreover, in a heterologous expression system oxidized amorphadiene could be further subject to glycosylation, as happens during transient expression of ADS and ADO in N. benthamiana (Ting et al. 2013; Van Herpen et al. 2010). Therefore, leaf extracts from both gADS-OE and cADS-OE Arabidopsis lines were analyzed by LC-QTOF-MS. However, in leaf extracts, no ADS-derived compounds were detected (Fig. S1).

Introns show less effect in the transient assay in N. benthamiana

We have shown that in stable transformants the expression of sesquiterpene synthases may be boosted by making use of genomic clones, which contain introns, rather than making use of a cDNA. The transient expression in N. benthamiana by agro-infiltration of terpene pathway genes is an efficient and fast way to reconstruct whole metabolic pathways in planta (Ting et al. 2013; Van Herpen et al. 2010). However, we have also experienced that when multiple individual constructs are co-expressed, the yield decreases due to gene dosage dilution effects. Because transient expression for pathway elucidation could benefit from intrinsic higher expression in agro-infiltrated leaves, we tested whether intron-mediated-enhancement also works in the context of transient assays in N. benthamiana. For this, N. benthamiana leaves were agro-infiltrated with either the gCST, cCST,
Chapter 5

**gADS or cADS expression constructs and leaves harvested at 5 dpi were used to measure volatile emission. Results show that transient expression of gCST in N. benthamiana results in similar β-caryophyllene emission as for cCST expression (Fig. 6A). Also, there is no significant difference in CST transcript levels between the two treatments (Fig. 5A). When the dosage of agro-infiltrated cCST or gCST is reduced, by diluting agro-infiltrated bacteria with Agrobacterium carrying an empty vector (EV), overall emission levels are reduced, but caryophyllene emission in leaves with gCST expression are slightly higher than in leaves with cCST expression (Fig. 5B).**

**Fig. 5.** Transcription levels of caryophyllene synthase gene (CST) and amorphadiene synthase gene (ADS) in leaves of N. benthamiana infiltrated with cCST, gCST, cADS and gADS. a, gene expression of CST with the non-diluted gene dosage, n=4; b, gene expression of CST with 5 x diluted gene dosage, n=9; c, gene expression of ADS, n=3.

**Fig. 6.** Volatile analysis on leaves of N. benthamiana infiltrated with cDNA of caryophyllene synthase (cCST), genomic caryophyllene synthase (gCST), cDNA of amorphadiene synthase (cADS) and genomic amorphadiene synthase (gADS) genes as identified by GC-MS at 5dpi. a, β-caryophyllene, α-humulene and α-copaene emission with the non-diluted gene dosage, n=4; b, β-caryophyllene, α-humulene and α-copaene emission with 5 x diluted gene dosage, n=9; c, amorphadiene emission, n=3. The headspace of infiltrated leaves was trapped for 4 hours using Tenax TA.
Introns in Terpene Synthase Genes Boost Sesquiterpene Yield

However, there is still no significant difference between CST mRNA levels in gCST and cCST agro-infiltrated leaves (Fig. 5B). In contrast, amorphadiene emission was increased 2-fold in N. benthamiana leaves expressing gADS compared with leaves expressing cADS (Fig. 6C), although ADS transcript levels were not significantly different between the two treatments (Fig. 5C).

Discussion

Introns enhance the expression of sesquiterpene biosynthesis genes

Here we have shown that engineering of ectopic terpene production in plants may benefit from the use of genomic clones containing introns, compared to ectopic expression of cDNAs. The intron-enhancement effect was demonstrated for two sesquiterpene synthase genes, CST and ADS (Fig. 3). In the gCST-OE lines the enhancement in product formation was higher than the increase in CST expression, suggesting additional post-transcriptional effects besides an effect on promoter activity as has been described before for IME (Parra et al. 2011; Rose et al. 2008). However, although ADS expression was also enhanced by the presence of introns, the increase in amorphadiene production was lower than the increase in ADS transcription. This indicates that for this gene the post-transcriptional effects of IME do not seem to work.

When we assume that for both the Arabidopsis gCST and A. annua gADS the Arabidopsis pre-mRNA maturation protein complexes and exon-junction protein complexes on mRNA are identical, this suggest an additional role for mRNA exon sequences or a role for spatial distribution of exon-junction protein complexes on mRNA in post-transcriptional IME effects. Apparently this post-transcriptional IME effect is less effective for the heterologous ADS mRNA sequence than for the endogenous CST mRNA sequence.

Lower ectopic amorphadiene production than ectopic β-caryophyllene production in Arabidopsis

When quantified by peak area, the average amorphadiene emission amounts to only 1.2% of the average β-caryophyllene emission in transformed Arabidopsis (Fig. 3A-D). This difference in emission is not explained by conversion of amorphadiene to non-volatile products, as no ADS-related compounds were detected by LC-QTOF-MS in a leaf extract (Fig. S1). The $K_m$ of CST recombinant enzyme for farnesyl diphosphate (FPP) (2.1 μM (Chen et al. 2003; Tholl et al. 2005) is higher than that of ADS for FPP
(0.6-0.9 μM: (Bouwmeester et al. 1999; Mercke et al. 2000). This difference in \( K_m \) CST and \( K_m \) ADS does not explain the quantitative difference in sesquiterpene headspace production between the \( g_{ADS-OE} \) and \( g_{CST-OE} \) lines. An alternative explanation for the low ADS activity in \( Arabidopsis \) may be that endogenous enzymes (e.g. CST) form a metabolon complex with FPPS, thereby limiting availability of FPP for ectopically produced ADS.

**Introns have no clear effect in transient expression**

Although introns enhance gene expression and sesquiterpene emission in stable transformed transgenic \( Arabidopsis \), in transient assays in \( N. benthamiana \) leaves the effect of introns is much less obvious. There is no clear difference in transcript levels for ectopic expression of genomic and cDNA fragments. However, for \( ADS \) in the transient assay in \( N. benthamiana \) there is a 2-fold increase in amorphaadiene production, indicating that there still is a small post-transcriptional effect from IME, which is not observed in stable transformed \( Arabidopsis \) (Fig. 3D). Also, when gene dosage was reduced, the effect of IME for in \( N. benthamiana \) became more apparent for \( g_{CST} \) and \( c_{CST} \) (Fig. 6B). Therefore, the artificial high expression levels under transient expression conditions seem to mask effects of IME manifested at normal gene dosage in stable transformed plants.

**Variable product ratios from CST activity in stable transformants**

The ratio of the three products (\( \beta \)-caryophyllene:\( \alpha \)-humulene:\( \alpha \)-copaene) was different in WT and \( g_{CST-OE} \) and \( c_{CST-OE} \) lines (Fig. 7A-B). The \( \alpha \)-humulene proportion in the headspace of rosette plants was significantly increased in \( g_{CST-OE} \) lines compared with \( c_{CST} \)-lines (Fig. 7A). In untransformed WT plants CST activity was highest in flowers (Huang et al. 2012) and headspace analysis from inflorescences showed that the \( \beta \)-caryophyllene:\( \alpha \)-humulene:\( \alpha \)-copaene ratio in the headspace of WT flowers is similar to the \( \beta \)-caryophyllene:\( \alpha \)-humulene:\( \alpha \)-copaene ratio in the headspace of \( g_{CST-OE} \) rosette leaves (Fig. 7A-B). The relative high proportion of \( \alpha \)-humulene in the rosette leaf headspace of untransformed WT \( Arabidopsis \) plants may be due to overestimation of the \( \alpha \)-humulene peak areas which was close to background levels (Fig. 4A).

In addition, product ratios may be influenced by saturation of the trapping agent. Indeed different trapping methods (Tenax vs. twister) resulted in different product ratios (Fig. 7B). Saturation of the trapping agent may be pre-
Introns in Terpene Synthase Genes Boost Sesquiterpene Yield

vented by varying the headspace trapping interval time or by varying the
gene dosage (e.g. in the transient assays). The peak intensities for caryophyllene were decreased 5- to 8-fold when the cCST and gCST gene dosages were diluted 5-fold upon transient expression in N. benthamiana (Fig. 6 A-B), however, expression of cCST or gCST at different gene dosage does not affect product ratio (Fig. 7C), suggesting that variation in product ratio is not related to volatile trapping saturation. Besides putative technical reasons

Fig. 7. Ratio of β-caryophyllene, α-humulene and α-copaene. a, WT, cDNA of caryophyllene synthase (cCST-OE), genomic caryophyllene synthase (gCST-OE) overexpression lines; b, wildtype flower tissues using different trapping agents (Tenax TA and twister) with 24 hour trapping; c, transient expression in leaves of N. benthamiana infiltrated with cCST, gCST. The headspace of WT plants (Col-1, Col-2), cCST-OE lines, cADS-OE lines and gADS-OE lines was trapped for 24 hours, while the headspace of gADS-OE lines was trapped for 4 hours using Tenax TA.
for variation in headspace product ratio from CST activity, it could also be that variation in products is related to different forms of CST derived from alternative splicing of the genomic CST gene. Two gene isoforms for CST, AT5G23960.1 and AT5G23960.2, have been reported in the TAIR database. These isoforms proteins show a small difference near the conserved catalytic side and therefore the two forms could differ in product specificity. However, no CST splice variant AT5G23960.2 was detected by us either in leaf RNA from untransformed WT plants or in leaf RNA from gCST-OE plants.

Conclusions

Here we have demonstrated that the use of intron containing genes is an efficient way to up-regulate heterologous terpenoid production in stably transformed transgenic plants, while the IME effect in transient expression is virtually absent.

Supplementary data

Figure S1. LC-QTOF-MS chromatograms for WT and cDNA of amorphadiene synthase (cADS-OE), genomic amorphadiene synthase (gADS-OE) lines in Arabidopsis. No ADS-related peaks were detected, Peak 1 and 2 are not significantly different between the samples (n=3).
Acknowledgments

We thank Francel Verstappen and Iris Kappers for their assistance with headspace analysis, Jacques Davies for the help with GC-MS analysis, Hieng-Ming Ting and Esmer Jongedijk for helpful discussion of GC-MS data analysis and Bert Schipper for assistance in LC-QTOF-MS analysis.

Reference


Gunderson SI, Polycarpou-Schwarz M, Mattaj IW (1998) U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase Mol Cell 1:255-264 doi:http://dx.doi.org/10.1016/S1097-2765(00)80026-X


Dev 14:2452-2460
Introns in Terpene Synthase Genes Boost Sesquiterpene Yield

Teoh KH, Polichuk DR, Reed DW, Nowak G, Covello PS (2006) Artemisia annua L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin FEBS Lett 580:1411-1416 doi:http://dx.doi.org/10.1016/j.febslet.2006.01.065
Vagner S, Vagner C, Mattaj IW (2000) The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3′-end processing and splicing Gene Dev 14:403-413
Chapter 6

General Discussion
Artemisinin (AN) is an important sesquiterpenoid molecule for the treatment of malaria produced in the plant *Artemisia annua*. The biosynthesis pathway of AN is branched, yielding both AN and arteannuin B (AB) and different *A. annua* chemotypes are recognized based on either high AN and low AB (high artemisinin plants, HAP) or low AN and high AB content (low artemisinin plants, LAP). However, due to the low natural concentration of AN, even in *A. annua* of the HAP chemotype, yields of AN do not meet the pharmaceutical demand. In order to improve the production of AN, it is important to understand AN biosynthesis and transport and storage of AN and its precursors in order to optimally apply available technologies for metabolic engineering. In this thesis, options for AN production in a heterologous plant host (*Nicotiana benthamiana*), using transient expression of AN pathway (AN-PW) genes, were explored and used to validate various strategies which may enhance AN product accumulation. In Chapter 2 I transiently expressed AN-PW genes in *N. benthamiana* to identify the mechanistic basis for the presence of an AB and AN branch in the pathway and showed that there are gene variants that are responsible for these different ‘chemotypes’. In Chapter 3 I used transient expression of AN-PW genes to characterize putative AN-PW product related transporter and carrier genes, leading to the identification of LTPs with specific sequestering activity. In Chapter 4 I used transient expression of AN-PW genes to study the interaction between the flavonoid and sesquiterpenoid biosynthesis pathways. Finally, In Chapter 5 I used transient expression of AN-PW genes to study the function of introns for enhancement of gene expression efficiency.

Here, in Chapter 6, I discuss the results in my thesis in a broader perspective. I will discuss major limitations we have encountered and potential solutions to metabolic engineering that still may be explored. Moreover, some of our findings go beyond engineering of the AN-PW and have implications for metabolic engineering in general (e.g. control of glycosylation, compartmentalization and sequestering of secondary metabolites, choice of host plant for heterologous production of secondary metabolites). Also, the current yields of AN in wildtype *A. annua* and in different other, heterologous, hosts through metabolic engineering will be compared. Finally, I will discuss how some results may also be of relevance to transport of plant hormones and therefore link my work to a more general understanding of plant physiology and development.
Evaluation of transient expression for metabolic pathway analysis

Throughout this thesis we have made use of the transient expression of genes in *N. benthamiana*. Most of the secondary metabolites are synthesized by multiple enzymes, and the reconstruction of an entire secondary metabolite pathway therefore requires the coordinated expression of multiple transgenes in a single cell. The transient expression system of *N. benthamiana* allows for quick assessment of PW genes as not all genes need to be in the same expression construct, but can be cloned in individual expression vectors and subsequently *N. benthamiana* leaves can be agro-infiltrated with all constructs at once by mixing the different *A. bacterium* strains containing these expression vectors. The fact that PW end products are produced after co-agro-infiltration in *N. benthamiana* leaves indicates that cells are transformed by multiple *A. bacterium* strains, which results in expression of multiple transgenes and full reconstitution of the PW within a single cell. Examples the use of multiple *A. bacterium* with different expression vectors are the reconstitution of the AN pathway in *N. benthamiana* through coexpression of 4 PW genes (this thesis), the parthenolide biosynthesis pathway in *N. benthamiana* though co-expression of 6 PW genes (Liu et al., 2014) and reconstitution of the secologain biosynthesis PW in *N. benthamiana* though co-expression of even 12 PW genes (Miettinen et al., 2014). However, in the latter case the total flux through the PW was very much reduced and needed to be boosted by infiltration of PW intermediates to obtain detectible levels of the end product (Miettinen et al., 2014). It could be argued that the more different *A. bacteria* are co-infiltrated into the *N. benthamiana* leaf, the higher the chances are that some of the cells will not be expressing all transgenes, resulting in a mixture of full and partial PW activity.

In this thesis, the combination of (1) *AaLTP3+AaPDR2+AN-PW* showed the highest free product formation, while the combination of (2) *AaLTP1+AaLT-P2+AaLTP3+2xEV+AN-PW* and (3) *AaPDR1+AaPDR2+3xEV+AN-PW* did not increase the (DH)AA (Chapter 3). The difference is that in combination (1) the PW genes are co-infiltrated with just 2 additional agrobacterium strains, while in combinations (2) and (3) the PW genes are co-infiltrated with 5 additional agrobacterium strains. The final relative gene dosage of the AN-PW genes may therefore affect total end-product accumulation and that is why in a single experiment, the relative dosage of the AN-PW genes was always kept constant.
Since in all experiments we find both early (AAOH, AAA, DHAAOH, DHAAA) and late intermediates of the AN pathway (AA, DHAA), this raises the question whether early intermediate accumulation is caused by incomplete PW reconstitution in cells. Also, because in Chapter 2 we did experiments in *N. benthamiana* expressing only AN-PW genes, while in Chapter 4 we did experiments expressing AN-PW+*AaPDR2+AaLTP3+ROS*, the relative dosage of AN-PW genes between these experiment is different. Therefore, if comparing the product levels obtained in different experiments we need to take into account the different relative AN-PW gene dosage that has been used.

To obtain insight into the effect of relative gene dosage we tested the AN-PW gene set, consisting of the *V5* construct (*ADS+FPS+HMGR*) in combination

![Figure 1. Effect of relative gene dosage on the product accumulation in *N. benthamiana*. A, Effect of relative dosage of agrobacterium carrying AN-PW genes (*V5+ADO*) by combining with different dosage of agrobacterium carrying EV on artemisinic acid (AA) product accumulation in *N. benthamiana*. B, Same data as in A but each value was corrected for the decrease in AN-PW gene dosage. Dilution with up to 7xEV can be explained by simple gene dosage effects. Beyond that correction for gene dosage is not sufficient and additional factors contribute to lower product levels (e.g. lack of overlap in expression of cells expressing both *ADS* and *ADO*).](image-url)
with amorphadiene oxidase (ADO) + P19 with different dosages of agrobacterium carrying EV through agro-infiltration in N. benthamiana. Results indicate that relative AN-PW gene dosage indeed has an effect on AA accumulation (Figure 1a). When product yield is corrected for the difference in AN-PW-gene dosage the results show that up to a dilution with 7xEV the change in yield is explained by the reduced AN-PW gene dosage. Beyond that point the reduction in PW product level is greater than is explained by the EV dilution and other factors interfere with efficient PW activity. For instance, beyond di-lution with 7xEV the diminished PW product accumulation may be explained by decreased overlap of cells expressing both ADS and ADO. This confirms that there is a dilution effect from co-expressing increasingly more trans-genes and this should be taken into account when comparing yields between experiments where different number of transgenes are used. Perhaps at a low dilution of the AN-PW there may also be already some lack of coexpression of ADS and ADO in the same cell, where exchange of amorphadiene between cells expressing ADS and cells expressing ADO compensates for that. At stronger dilution of AN-PW genes the distance between ADS and ADO expressing cells may become too big for this compensating effect.

**Competing reactions in the AN pathway in N. benthamiana leaves**

One of the most revealing experiments in this thesis is the experiment in which I infiltrated (DH)AA into the N. benthamiana leaves because, first of all, subsequent analysis of leaves shows that (DH)AA is very rapidly taken up by cells. This made us realize that, although transporters are involved in the secretion of (DH)AA produced inside the cells, these products very rapidly diffuse back into the cell, making the action of the PDR transporters futile if there is no additional mechanism for sequestering the product away from the plasma membrane. Secondly, analysis of what happens to infiltrated (DH)AA showed a rapid conversion towards (DH)AAA and (DH)AAOH inside cells, revealing endogenous enzymes that act in a reverse flux through the AN-PW (Chapter 3). Presumably these enzymes are non-specific alcohol dehydrogenase enzymes. A similar reverse AN-PW activity has also been reported for A. annua, as the enzyme oxidoreductase1(RED1) converts DHAAA into DHAAOH and possibly also AAA to AAOH (Rydén et al., 2010). Such reverse pathway flux is limiting AN-PW product accumulation and could contribute to competing reactions in N. benthamiana by providing (DH)AAOH.
for glycosyltransferases. Glycosylation of (DH)AAOOH and (DH)AA may be a way of detoxification of these heterologous compounds in *N. benthamiana*. Glycosylation makes the compound more hydrophilic and therefore more water soluble (Rivas et al., 2013). It could be that glycosylation of the end product (DH)AA enhances the metabolic flux through the AN-PW in *N. benthamiana* as the bulk of AN-PW products is in the form of glycosides (Chapter 2). For instance, treatment of *N. benthamiana* leaf extracts with the glycosidase (Viscozyme L.) showed that up to 10-fold more (DH)AA is present in the form of glycosides compared to free (DH)AA (Chapter 2). *In vitro* biochemical studies indicate that multiple glycosyltransferases from the same plant can glycosylate the same compounds, and that individual glycosyltransferases have the capacity to glycosylate multiple substrates (Osborn et al., 2009). Lipophilic terpenes may diffuse passively over the membranes (Widhalm et al., 2015) as also suggested by our infiltration studies of (DH)AA (Chapter 3). Glycosylation of AN-PW products may prevent free diffusion of these modified compounds over membranes separating the different cellular compartments and therefore glycosylation affects where products eventually accumulate and whether they are still available for final conversion to AN or AB (if de-glycosylation also occurs). However, glycosylation of pathway intermediates such as DHAAOOH also prevents further flux through the pathway. Therefore, suppression of glycosylation in *N. benthamiana* could be one of the targets for future research on AN-PW engineering, but the lack of specificity and presence of multiple glycosyltransferases in cells will make it challenging to control AN precursor glycosylation. The presence of a very active reverse PW flux and the presence of competing side reactions like the glycosylation which occurs in *N. benthamiana* both stress the need for efficient AN-PW product sequestering activity to escape from these competing reactions and to give a directional flow to the AN-PW. In addition, efficient sequestering of terpenes into the apoplast may be needed to prevent toxic effects of these compounds inside the cell. Therefore, future researched aimed at increasing AN-PW yield in the *N. benthamiana* production platform should target identifying and inhibiting genes encoding enzymes involved in reverse AN-PW activity, identification and inhibition of genes encoding glycosyltransferases acting on AN-PW products, and by enhancing sequestering activity into the apoplast, e.g. by enhancing transport out of the cell by PDRs and enhancing apoplast sequestering activity with LTPs (Chapter 3).
Effect of ROS on AN-PW glycosylation profile

In Chapter 4 we studied the interaction between the AN-PW and the flavonoid biosynthesis pathway, as activated by ectopic expression of ROS. An indirect effect of flavonoids on terpene biosynthesis was suggested by studies in tomato where it was shown that a mutation in the gene encoding the flavonoid biosynthetic enzyme chalcone isomerase (CHI) also results in reduced terpene production in leaves (Kang et al., 2014). However, this effect was based on an effect of flavonoids on type VI trichome development in which terpene biosynthesis takes place and in plants with mutated CHI the type VI trichomes never reach a stage at which the terpene biosynthesis pathway is activated. Therefore, the reduced terpene production in CHI mutants cannot be ascribed to a direct interaction between these two pathways in the same cell (Kang et al., 2014). In our transient expression studies in N. benthamiana leaves ROS can stimulate AN-PW product accumulation. Because expression of all these genes is in already fully established leaves we do not expect that at this stage ROS is having any effect on trichomes where AN-PW genes could be more effective. Therefore, the stimulating effects of ROS on AN-PW must have other reasons and indeed I identified induction of Mevalonate Kinase (MVK) by ROS, which is involved in the sesquiterpene precursor pathway (Chapter 4). Curiously, the concentration of glycosides of (DH)AA and (DH)AAOH was enhanced 2- to 3-fold in leaves expressing AN-PW+ ROS compared to expression of only the AN-PW (Chapter 4). This effect of ROS was reduced when at the same time flavonoid biosynthesis was blocked by co-expression with the CHS<sub>RNAi</sub> construct. This suggests that it may be the flavonoid compounds themselves which affect glycosylated AN-PW product accumulation (Chapter 4). It is known from cosmetics studies that terpenes and flavonoid molecules interact to increase skin penetration (Saija et al., 1998) but more research is needed to determine whether a direct interaction between sesquiterpene and flavonoid molecules is responsible for the effect on glycosylated product accumulation.

Host organism selection for engineering

During my PhD study, I used different plant host species for overexpression of AN-PW genes: transient expression of AN-PW in N. benthamiana (Chapter 2, 3, 4) and transient expression of ADS in Lactuca sativa (lettuce; data not shown) and stable overexpression of ADS in Arabidopsis (Chapter 5) and sta-
Chicory, lettuce and *A. annua* belong to the *Asteraceae* family and produce sesquiterpenes derived from amorphadiene (*A. annua*) or germacrene A (lettuce, chicory). The germacrene A derived sesquiterpene lactones in lettuce and chicory are responsible for the bitter taste (Bennett et al., 2002; Bouwmeester et al., 2002; Graziani et al., 2015; Ikezawa et al., 2011; Mahmoud et al., 1986). Initial biosynthesis of germacrene A derived sesquiterpenes is very similar to AN biosynthesis (Figure 2). Germacrene A synthase (GAS) is the first committed step for sesquiterpene biosynthesis in chicory and lettuce followed by oxidation by Germacrene A oxidase (GAO), a cytochrome P450 enzyme. In fact, it has been shown that chicory or lettuce GAO can act on amorphadiene as well (Nguyen et al., 2010), suggesting that expression of only ADS in either chicory or lettuce may be sufficient to produce ADS-derived sesquiterpenes. However, upon transient expression of the *AN-PW* in lettuce, no ADS derived products could be detected, while control experiments with a 35S:*luciferase* or 35S:*CST* indicate that expression from an agro-infiltrated gene is possible in lettuce (data not shown). Also chicory, stably transformed with the *ADS* gene, shows that *ADS* is transcribed, but no ADS-derived sesquiterpenes were detected in chicory leaves. One explanation for the lack
of ADS-products in lettuce and chicory may be that endogenous enzymes form a metabolon complex with FPPS, thereby limiting availability of FPP for ectopically producing ADS-product (Figure 2). The lack of apparent ADS activity in lettuce and chicory is not related to affinity for the substrate FPP as the Km for FPP of ADS is about 10-fold lower (0.6 ~ 0.9 μM: (Bouwmeester et al., 1999; Mercke et al., 2000), compared to the Km for FPP of GAS (6.6 ~ 6.9 μM; (Bouwmeester et al., 2002; de Kraker et al., 1998). Alternatively, the amorphadiene from ADS activity in chicory and lettuce is perhaps converted to reactive compounds which crosslink to compounds and/or proteins such that they are no longer visible in the MS chromatograms. Therefore, additional experiments are needed to determine the bottleneck in AN-PW production in lettuce and chicory. For instance, infiltration experiments with AA and/or DHAA in leaves of chicory or lettuce could be used to determine intrinsic bioconversion activity. Clearly, our initial choice to exploit plants with a similar pathway for ectopic expression of the AN-PW (chicory, lettuce) did not give the expected results. In contrast, expression of the AN-PW genes was far more successful in *N. benthamiana*. It needs further investigation to determine whether established sesquiterpene metabolons in plants with a similar sesquiterpene pathway may actually hinder introduction of competing enzyme activities, compared to plants not expressing such pathways.
Efficiency and capacity for (DH)AA export from plant cells

With the substrate exclusion assay I found a new way to test functionality of LTPs in plants, provided that influx of the substrate into the cells can be quantified. I showed that characterized transporters in the AN-PW, such as AaPDR2, are barely sufficient to prevent reflux of (DH)AA back into the cell and that extracellular LTPs are far more efficient in preventing reflux (Chapter 3). The relative efficiency for transport to the apoplast for each of the AN-PW intermediates was calculated as the ratio of product level in the apoplast and product level in a total leaf extract (Figure 3). Such calculations show that the endogenous transport activity in *N. benthamiana* (AN-PW gene expression without *AaLTP* or *AaPDR*) is very similar to the relative transport activ-

Figure 3 Transport efficiency (ng apoplast/ ng total leaf) in *N. benthamiana* infiltrated AN-PW genes with LTP3 and PDR2 at 6dpi. 7 constructs were co-expressed in this experiment. Error bar is SE; p<0.05, n=5.

Figure 4 Detection of artemisinin and its intermediates during 0, 10 and 60 min sonication extraction of *N. benthamiana* leaf extracts. AA, artemisinic acid; DHAA, dihydroartemisinic acid; AAOH, artemisinic alcohol; DHAAOH, dihydroartemisinic alcohol; AAA, artemisinic aldehyde; DHAAA, dihydroartemisinic aldehyde; AB, arteannuin B; AN, artemisinin.
ity when additional AaLTP and/or AaPDR proteins are expressed. However, both for the endogenous transport activity as for the ectopically produced transport activity there is a higher preference for (DH)AA products than for (DH)AAOH products. Although the specificity for transport is not changed upon co-expression of AaPDR2, the capacity for transport did increase as demonstrated by higher absolute levels of (DH)AA in the apoplast (Chapter 3). Based on partitioning of (DH)AA over water and chloroform, we can speculate that extraction of (DH)AA from the apoplast using water will not be complete. However, for N. benthamiana leaves expressing AN-PW+DPR2+LTP3 harvested at 5 or 6 dpi, the increase in free (DH)AA is only partly explained by the increase of these compounds in the apoplast, as at this stage there also still seems to be more free (DH)AA in the remaining leaf than can be explained by partitioning of (DH)AA over two phases in the apoplast. Therefore, I think that at 5-6 dpi there still is a substantial amount of (DH)AA inside the cells. This could be explained by intracellular sequestering of (DH)AA to a different subcellular compartment where there is no competing glycosylation activity. It will be worthwhile to investigate in the future whether AaLTP3 and AaDPR2 are already relevant inside the cell, for instance in loading of vesicles with (DH)AA, which subsequently fuse with the plasma membrane to release both (DH)AA and LTP into the extracellular space and AaPDR on the plasma membrane.

We also considered that higher levels of free form (DH)AA in whole leaf extracts could be due to sample processing. That is, whole leaf extracts are sonicated for several minutes and it could be that AN-PW glycosides are not stable under these conditions and are hydrolysed into free (DH)AA. However, control experiments show that sonification has no effect on free AN-PW product accumulation from whole leaf extracts (Figure 4).

Remaining questions on the specificity of AaPDRs and AaLTPs

In our analysis of AaLTPs we have focused on the most active combination of AaLTP3 and AaPDR2. AaDPR1 shows a lower activity towards AN-PW products even though the expression profile of PDR1 over the leaf development in A. annua fits better with that of AN-PW biosynthesis genes (Chapter 3). AaPDR1 seem to have more effect on early AN-PW intermediates like (DH)AAOH but it could also be that AaPDR1 is involved in the transport of one or some of the many other (sesqui)terpenes that are produced and stored in A. annua trichomes (Markus Lange and Turner, 2013). In A. annua, AN levels are increased by Jasmonic Acid (JA) treatment, due to stim-
ulation of AN biosynthesis genes. However, *PDR1* transcription in glandular trichomes is not increased by JA (Maes et al., 2011) (Soetaert et al., 2013). Therefore, more work needs to be done on characterization of the role of AaPDR1, also in combination with the other AaLTPs. At present it is also not clear why co-expression of the set of AaLTPs and AaPDRs with AN-PW genes is less efficient than expressing specific combinations like AaLTP3+AaPDR2 with the AN-PW genes (Chapter 3). Actually, the function of AaDPR activity and AaLTP activity also needs validation in the context of *A. annua* itself. For this we tried to use VIGS on *A. annua*, but did not succeed in knock-down of gene function in *A. annua*. It would be good to proof the function of AaPDRs and AaLTPs in *Artemisia annua* in the future, for instance by generating transformants with AaPDR1/2 RNAi or AaLTP1/2/3 RNAi constructs or with overexpression constructs, to test whether and how manipulation of these genes affects AN-PW product accumulation.

**Comparing AN yields from *A. annua* and from metabolic engineering in yeast and *N. benthamiana***

The yields of AN and AB in *N. benthamiana* reached 18 ng/mg DW when the AN-PW was expressed in combination with AaPDR2, AaLTP3 and ROS (Chapter 4). Agro-infiltration can be done using 4-week *N. benthamiana* plants, and leaves can be harvested at 13 dpi. Therefore, such a full cycle could be repeated 8-9 times per year. Allowing for space to grow *N. benthamiana* plants, an industrial scale production of AN through agro-infiltration of *N. benthamiana* plants could yield around 60 kg AN /ha/year. In comparison, the best *Artemisia* cultivars yield ca. 1.5% DW AN, and agricultural yields may be as high as 74 kg DW/ha every year (Elfawal et al., 2012; Kumar et al., 2004). Thus, the potential year yield for industrial scale production of AN through agro-infiltration is currently at the range of the best *A. annua* cultivars yields. Alternatively, AN may be produced in *A. annua* hairy roots which can produce up to 1.12 mg AN/g root DW, equivalent to 25.78 mg AN/liter hairy root culture in 36 days (Patra and Srivastava, 2015). Therefore, in order to reach yields similar to that from *A. annua*, a total of 2868 liter hairy root culture would be needed per year. AA can be produced in yeast with yields up to 100 mg AA per liter but then needs further photochemical conversion to AN (Paddon and Keasling, 2014; Paddon et al., 2013; Ro et al., 2006). To reach the same yield of AN per year as from one hectare of *A. annua* plants 100 liters of yeast culture would be necessary (with subsequent
100% efficient conversion of AA to AN) to reach a similar yield. The potential yields of the different production platforms are summarized in Table 1.

A putative model of AaLTP3 function in the AN-PW in *N. benthamiana*

In this thesis work I showed that AaLTP3 enhances sequestering of (DH)AA in the apoplast of leaves expressing the AN-PW genes and I confirmed the extracellular localization of AaLTPs with LTP-GFP fusion studies. AaLTP3 helps to retain apoplastic (DH)AA outside the cells and is more effective than AaPDR2, suggesting that AaLTP3 manages to keep (DH)AA away from the plasma membrane from where it is easily imported back into the cell, as judged from the

<table>
<thead>
<tr>
<th>Species</th>
<th>AN or AA production</th>
<th>time (weeks)</th>
<th>cycles</th>
<th>Special consuming</th>
<th>year yield per ha or 100 liter bioreactor</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Nicotiana benthamiana</em></td>
<td>18 ng/mg DW AN</td>
<td>5-6</td>
<td>8-9</td>
<td>Medium, <em>A. tumefaciens</em></td>
<td>57.6-64.8 kg/ha</td>
<td>This thesis</td>
</tr>
<tr>
<td>Wildtype <em>Artemisia annua</em></td>
<td>0.01%–1% DW AN</td>
<td>≥30</td>
<td>1-2</td>
<td>-</td>
<td>5-25 kg/ha</td>
<td>Kumar et al., 2004</td>
</tr>
<tr>
<td>Best <em>Artemisia annua</em> cultivars</td>
<td>ca. 1.5% DW AN</td>
<td>≥30</td>
<td>1-2</td>
<td>-</td>
<td>28.5-74.2 kg/ha</td>
<td>Elfawal et al., 2012;</td>
</tr>
<tr>
<td>Hairy root bioreactor</td>
<td>25.78 mg/L AN</td>
<td>5-6</td>
<td>8-9</td>
<td>MS media</td>
<td>20.6-23.2 g/100 L</td>
<td>Patra et al, 2015</td>
</tr>
<tr>
<td>Yeast</td>
<td>25g/L AA</td>
<td>2-3</td>
<td>21-23</td>
<td>Medium, yeast, chemical syntheses</td>
<td>52.5-57.5 kg/100 L</td>
<td>Paddon and Keasling, 2014; Paddon et al., 2013; Ro et al., 2006</td>
</tr>
</tbody>
</table>
(DH)AA apoplast infiltration assays. It could be that AaLTP3 transports lipids to the outside of the cell and forms an apoplastic lipid sink where protonated (DH)AA molecules can accumulate. However, that would not explain the specific effect of AaLTP3 on (DH)AA but not (DH)AAOH (Chapter 3). Moreover, specificity of AaLTP3 was also demonstrated by testing the effect of AaLTP3 in combination with another but very similar sesquiterpene biosynthesis pathway from feverfew: the costunolide pathway. Coexpression of the feverfew costunolide biosynthesis genes in *N. benthamiana* with or without AaLTP3 did not change the accumulation of costunolide in the apoplast (personal communication A.B. Kashkooli). Therefore, it is not likely that AaLTP3 functions by creating a general apoplastic lipid sink, as other hydrophobic molecules in the apoplast should show a similar benefit from such general lipid sink.

Figure 5 Putative model of the role of ALTP3 in apoplastic sequestering of (DH)AA. (DH)AA is produced in the cytosol by P450 enzymes on the ER membrane. From there (DH)AA is substrate for glycosylation (and presumably subsequent sequestration into the vacuole) or (DH)AA somehow reaches the AaPDR2 membrane transporter and is transported to the other side of the plasma membrane. Without sufficient AaLTP3 in the cell wall, (DH)AA may diffuse back into the cell. With AaLTP3 present in the cell wall, (DH)AA may be bound by AaLTP3 and the complex may diffuse over the cell wall to reach the apoplast where (DH)AA is deposited and empty AaLTP3 diffuses back to the inner side of the cell wall to repeat the carrier cycle.
LTP proteins have a hydrophobic cavity with space sufficient for the binding of lipid molecules (Cheng et al., 2004; Gomar et al., 1996; Guerbette et al., 1999). Therefore, one possibility is that AaLTP3 binds the (DH)AA to form a LTP-(DH)AA complex which prevents reflux of (DH)AA back into the cell. It has been shown that one LTP protein can bind one (Buhot et al., 2001; Zachowski et al., 1998) or even 4 lipophilic molecules (Cheng et al., 2004).

If the effect of the LTP on preventing pathway reflux is solely based on binding to the substrate (DH)AA, this would require an excess of LTP protein molecules over the substrate molecules. If we assume an extremely high AaLTP3 protein level of 1 mg/g FW in the transient expression in N. benthamiana leaves, and we compare the estimated number of LTP proteins with the number of DHAA molecules produced by ectopic expression of the AN-PW genes in N. benthamiana leaves [3 ng/mg FW (DH)AA] the number of AaLTP3 proteins are indeed in excess compared to the number of (DH)AA molecules. However, in the (DH)AA apoplast infiltration assay in leaves expressing AaLTP3, we estimate that the number of infiltrated (DH)AA molecules is > 600-fold higher than the number of AaLTP3 protein molecules, and the efficient apoplast sequestering activity of AaLTP3 can therefore not solely be explained by AaLTP3 binding one or even 10 molecules of (DH)AA.

Therefore, we favor a model in which (DH)AA produced inside the cells is transported across the plasma membrane by AaPDR2 and where subsequently (DH)AA is carried over the cell wall by AaLTP3 to be deposited at the other side of the cell wall, away from the plasma-membrane and thus preventing reflux of (DH)AA back into the cell (Figure 5). Without AaLTP3 expression, free (DH)AA exported from the cell may rapidly diffuse back into the cell and may still contribute to feedback inhibition of ADO or become substrate for glycosyltransferases inside the cell. This model suggests a shuttling of AaLTP3 between the inner surface of the cell wall near the plasma membrane and outer surface of the cell wall at the intracellular space and could explain the preferred accumulation of AaLTPs at the intercellular space as observed for the AaLTP-GFP fusion proteins. The model does raise some specific questions. First of all, it suggests a transient binding of (DH)AA to AaLTP3, something that should be tested in the near future. Moreover, it raises the question what triggers the release of (DH)AA by loaded AaLTP3 at the outer side of the cell wall. At present this suggested movement of LTPs over the cell wall is below the resolution limit of confocal microscopy. However, to test a role for LTP mobility in the cell wall, LTP could be
immobilized at the plasma membrane surface, for instance by addition of a glycosylphosphatidylinositol (GPI) anchor site to the AaLTP3 protein, to keep it bound to the plasma membrane (Alberts B, 2002; Sharom FJ, 2002).

Removal of the free (DH)AA from the cytosol seems to have consequences for the overall flux through the pathway, as overall AN-PW product levels are increased upon co-expression with AaPDR2+AaLTP3 (Chapter 3). Effects on PW flux are strongest for (DH)AA products, while (DH)AAOH glycosides and AAA glutathione conjugates are less effected (Chapter 3). Despite the improved transport and sequestering of free (DH)AA, still the majority of AN-PW products are glycosylated in *N. benthamiana*. Therefore, it will be important to determine whether any transport components are still missing for efficient export of (DH)AA from the cell, and whether escape from side reactions can be enhanced. Alternatively, we may need to spend more research on how to control the competing side reactions.

**Are LTPs also relevant for plant hormone homeostasis?**

Some major plant hormones like gibberellins (GA), abscisic acid (ABA) and strigolactones (SL) are produced by the plastidic terpenoid biosynthesis pathway. Moreover, ABCG-type membrane transporters have been implicated in the transport of these terpenoid plant hormones over the plasma membrane. For example in *Arabidopsis* ABA is transported by an ABCG-type transporter (Kuromori and Shinozaki, 2010), while in petunia strigolactones (SLs) are transported by the ABCG-type transporter PhPDR1. Moreover, this petunia PDR1 shows apical localization in root cells, suggesting that they may be involved in apical transport of SLs which are produced in the root tip. In petunia roots there are also passage cells in which PDR1 protein is localized laterally, indicating a role of this transporter in secretion of SLs to the soil where this molecule may activate arbuscular mycorhizal fungi (Cheng et al., 2013; Kretzschmar et al., 2012). Therefore, questions on transport of terpenoid plant hormones are very similar to the questions I have addressed on transport of sesquiterpenes of the AN-PW. Consequently, since a specific AaLTP3 has been implemented in sequestering of (DH) AA into the apoplast, this raises the intriguing question whether other LTPs in plants act in apoplast sequestering of terpenoid plant hormones and whether LTPs contribute to directional flux of terpenoid plant hormones. If LTPs are involved in transport of terpenoid plant hormones we predict hormone relat-
ed phenotypes for LTP mutants or overexpression plants. The Arabidopsis genome contains over 70 annotated LTP genes (Beisson et al., 2003) which need to be studied in a systematic way. However, recently indeed an effect of LTP on hormone homeostasis was reported for Arabidopsis AtLTP3 OE plants, which have higher ABA levels, while salicylic acid (SA) levels are decreased (Gao et al., 2015). Although it is still open whether there is a role for LTPs in plant hormone homeostasis, this paper and our own preliminary results with newly constructed Arabidopsis LTP overexpression lines do seem to confirm a role for LTPs in plant hormone related phenotypes (Table 2).

Overexpression of the AaLTPs showed that these extracellular proteins tend to show polar accumulation on cells surrounding a common central cavity (Chapter 3). It will be interesting to test in the future whether expressions of LTP-GFP fusion proteins under control of their own promoter actually also show polar localization in the whole plant tissue context and how this relates to cellular expression and localization of PDR transporters implicated in terpenoid hormone transport over plasma membranes.

Table 2. Arabidopsis LTPs show the hormone-related phenotypes accumulation. “–”, means; “+”, increase; n.d, not detected; n.c, not changed significantly.

<table>
<thead>
<tr>
<th>Transgenic lines</th>
<th>ABA</th>
<th>SA</th>
<th>leaf size</th>
<th>Leaf numbers</th>
<th>flowering time</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtLTP3 OE vs. WT</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>preliminary result, Gao et al., 2015</td>
</tr>
<tr>
<td>AtLTP3 KO vs. WT</td>
<td>n.c</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>Ting, 2014; Gao et al., 2015</td>
</tr>
</tbody>
</table>

Concluding remarks and future perspectives

The results described in this thesis present the overview of the progress that has been made towards understanding the transport and metabolic engineering of terpenoids in planta and most importantly, progress in the understanding of transport of terpenoid pathway products and intermediates to the apoplast was achieved, revealing a novel role for LTPs in trafficking of terpenoids. These results should be exploited in novel engineering approaches to industrial production of AN or other valuable (sesqui)terpenes.
Nevertheless, many questions still remain, and we still do not have full control over the flux of an engineered metabolic pathway in a heterologous host.

Reference


Elfawal, M.A., Towler, M.J., Reich, N.G., Golenbock, D., Weathers, P.J., and Rich, S.M.


synthase, a key enzyme of artemisinin biosynthesis in *Artemisia annua* L. Archives of Biochemistry and Biophysics 381, 173-180.


Summary
The sesquiterpenoid artemisinin (AN) is the most important medicine for the treatment of malaria in humans. The industrial production of AN still mainly depends on extraction from the plant *Artemisia annua*. However, the concentration of AN in *A. annua* is low. Although different engineering strategies have been used in both *A. annua* and heterologous plant and yeast production platforms, the worldwide capacity and production costs for AN are not in balance with its demand (*Chapter 1*). Although the genes encoding for the entire AN biosynthesis pathway (AN-PW) of the AN precursor dihydroartemisinic acid (DHAA) have been identified, the application of these genes in pathway engineering seem to be limited by lack of control over product transport and sequestration. At the onset of this thesis project there was no information on transport in the AN-PW. However, it was known that DHAA is converted into AN outside the glandular trichome cells of *A. annua*. Therefore, in this thesis I tried to gain more knowledge on transport within the AN-PW and the use of different metabolic engineering strategies to improve the production of AN.

At the onset of my PhD project, the AN-PW genes from two different *A. annua* chemotypes were compared to understand the basis of different relative activities in the two branches of the AN-PW (*Chapter 2*). For these assays we used transient expression in *N. benthamiana*. In the AN-PW, artemisinic aldehyde (AAA) is at a branch point as it can be converted to artemisinic acid (AA) by amorphadiene oxidase (AMO), or to dehydroartemisinic aldehyde (DHAAA) by artemisinic aldehyde Δ11 (13) reductase (DBR2). AA is the precursor for arteannuin B (AB) while DHAAA may be converted by a CYP71AV1 or an ALDH1 to dehydroartemisinic acid (DHAA), the precursor for AN. In this chapter we demonstrate that the CYP71AV1 from a high AN production (HAP) chemotype has reduced activity in the AB branch of the pathway compared to the CYP71AV1 from a low AN production (LAP) chemotype. In addition, we show that the relative expression levels of DBR2 and ALDH1 also affect the AN/AB chemotype. The low catalytic efficiency of AMO from the HAP chemotype may be caused by a deletion of seven amino acids at the N-terminus of the protein compared to CYP71AV1 from LAP. Ectopic expression of the AN-PW genes in *N. benthamiana* showed that the bulk of the PW products are modified by glycosylation and glutathione conjugations. These side reactions therefore compete with the biosynthesis flux towards the AN precursor DHAA. At this point in my thesis the ectopic expression of AN-PW genes in *N. benthamiana* had not yield-
ed any AN. At a later stage it became clear that this was due to harvest of leaves at 5-7 days post agro-infiltration (dpi), while AN in \textit{N. benthamiana} leaves expressing AN-PW genes only becomes detectable after 7 dpi.

Glycosylation of the bulk of the AN-PW products in \textit{N. benthamiana} stresses the need for an efficient transport of (DH)AA to the outside of cells in order to escape from the glycosylation reactions. In Chapter 3, transport and sequestration of AN precursors was investigated by studying the effect of membrane transporters (PDRs) and lipid transfer proteins (LTPs). Hereto, two membrane transporters with activity towards AN-PW products were made available by the group of Prof. Marc Boutry and we isolated three \textit{LTP} genes from \textit{Artemisia annua} which showed expression in the glandular trichomes. In this chapter we show that AaLTP3 displays specific activity, together with AaPDR2 towards transport of (DH)AA to the apoplast in \textit{N. benthamiana}. Moreover, infiltration experiments with (DH)AA in \textit{N. benthamiana} leaves revealed that these compounds are rapidly taken up by the cells and that inside the cells there is a strong reverse flux in the AN-PW by conversion of (DH)AA towards (DH)AAA and (DH)AAOH. Subsequently we demonstrated that AaLTP3 has a stronger activity in keeping products in the apoplast than the AaPDR2 membrane transporter. Therefore, I suggest that by removal of (DH)AA from the cytosol through transport over the plasma membrane by AaPDR2 and subsequent sequestration in the apoplast by AaLTP3, AaLTP3 creates sink activity which prevents reflux of (DH)AA from the apoplast back into the cells. AaLTP3 therefore contributes to a directional flux through the AN-PW towards the end product (DH)AA. Finally, in this work we could also for the first time detect AN and AB in \textit{N. benthamiana} leaves by extraction of necrotic leaves at 13 dpi.

Because in \textit{A. annua} glandular trichome cells both the AN sesquiterpene biosynthesis pathway and the flavonoid biosynthesis pathway are active, we explored whether there is a functional interaction between these two major secondary metabolite biosynthesis pathways. In Chapter 4 we describe how we manipulate the flavonoid biosynthesis pathway in \textit{N. benthamiana} leaves using the \textit{Antirrhinum majus} transcription factor \textit{Rosea1 (ROS)} and test coexpression of \textit{ROS} with AN-PW genes. The co-expression of \textit{ROS} stimulates AN-PW product accumulation. Subsequent analysis indicates that this is most likely from transcriptional activation of the enzyme Mevalonate Kinase.
Summary

(MVK) in the mevalonate pathway, which provides precursors for the sesqui-terpene biosynthesis pathway. In addition, we demonstrate that production of flavonoids competes with AN-PW product accumulation, as co-expression of AN-PW genes with ROS, but simultaneous inhibition of chalcone synthase (CHS) by a CHS<sub>RNAi</sub> construct, results in higher AN-PW product levels. However, accumulation of the end products AN and AB was not affected significantly. Finally, the combined expression of AN-PW+ROS+AaPDR2+AaLTP3+CHS<sub>RNAi</sub> results in highest sequestration of (DH)AA in the apoplast and highest accumulation of the end products AN and AB in <i>N. benthamiana</i>.

During my thesis work, in a related project it was found that expression of another sesquiterpene biosynthesis gene (caryophyllene synthase; CST) in transgenic <i>Arabidopsis</i> resulted in higher caryophyllene emission for a transformant expressing a genomic DNA of CST, compared with a similar transformant expressing a CST cDNA described in literature. This suggested that ectopic expression of intron containing genes is more efficient than ectopic expression of cDNAs. To test whether in the context of metabolic engineering the use of genomic (intron-containing) genes is more efficient than the use of the corresponding cDNA we generated a set of stable transformed <i>Arabidopsis</i> lines with either genomic CST (gCST), cDNA CST (cCST), genomic amorphadiene synthesis (gADS) and cDNA ADS (cADS). In chapter 5 we show that indeed the lines with overexpression of the genomic clones yield higher levels of the anticipated products (caryophyllene or amorphadiene) than the lines with overexpression of the corresponding cDNAs. Transcript analysis showed that for gCST the increase in caryophyllene production was higher than can be explained solely by the increase in CST transcription. In the context of transient expression in <i>N. benthamiana</i> leaves the intron-mediated-enhancement effect was less pronounced.

In the final discussion chapter 6 I review limitations and potential solutions to metabolic engineering of the AN-PW in plants, and I discuss the impact of our findings on AN production capacity using transient expression versus natural production in <i>A. annua</i>. Moreover, I discuss how the finding of this thesis go beyond just insights into the AN-PW as especially the identification of the role of LTPs in sequestration of (sesqui)terpenes into the apoplast may have an impact on the metabolic engineering efforts.
of many other (sesqui)terpene pathways. Because some plant hormones are also terpenoid products the newly identified role of LTPs may also have impact on a deeper understanding of hormone signalling in plants. I have already started exploring this path by generating a set of *Arabidopsis* plants with overexpression of different *Arabidopsis* LTP genes to test whether any hormone related traits are altered (Chapter 6). Preliminary results do indeed confirm a role of LTPs in endogenous plant hormone balance, something worthwhile to be further explored in future research.
Samenvatting
Samenvatting

De sesquiterpenoïde artemisinine (AN) is het belangrijkste geneesmiddel voor de behandeling van malaria bij de mens. De industriële productie van AN nog steeds hangt vooral af van extractie uit de plant Artemisia annua. De concentratie van AN in A. annua laag. Hoewel verschillende technische strategieën zijn gebruikt in zowel A. annua en heterologe planten en gist productieplatforms, de wereldwijde capaciteit en productiekosten voor AN niet in evenwicht met zijn verzoek (hoofdstuk 1). Hoewel de genen die coderen voor de gehele AN biosyntheseroute (AN-PW) van de AN voorloper dihydroartemisininezuur (DHAA) zijn geïdentificeerd, de toepassing van deze genen in pathway engineering lijken te worden beperkt door gebrek aan controle over transportproduct en opslag. Aan het begin van dit proefschrift project was er geen informatie over het vervoer in de AN-PW. Er werd echter bekend dat DHAA wordt omgezet in een buiten de klier trichoom cellen van A. annua. Daarom is in dit proefschrift heb ik geprobeerd om meer kennis op vervoer binnen de AN-PW en het gebruik van verschillende metabolic engineering strategieën te krijgen tot de productie van een verbetering.

Bij het begin van mijn doctoraatsproject werden de AN-PW genen van twee verschillende A. annua chemotypes opzichte basis van verschillende relatieve activiteiten van de twee takken van de AN-PW (hoofdstuk 2) te begrijpen. Voor deze testen gebruikten we tijdelijke expressie in N. benthamiana. In de AN-PW, artemisinine aldehyde (AAA) op een vertakkingspunt als het kan worden geconverteerd naar artemisinine zuur (AA) van amorphadiene oxidase (AMO), of aldehyde (DHAAA) van artemisinine aldehyde Δ11 (13) reductase (DBR2). AA is de voorloper van arteannuin B (AB), terwijl DHAAA worden omgezet door CYP71AV1 of ALDH1 tot dehydroartemisinic acid (DHAA), de voorloper van AN. In dit hoofdstuk tonen we aan dat de CYP71AV1 vanaf een hoog AN productie (HAP) chemotype activiteit in de AB tak van de route is verminderd in vergelijking met de CYP71AV1 van een lage AN productie (LAP) chemotype. Bovendien tonen wij dat de relatieve expressieniveaus van DBR2, ALDH1 en beïnvloeden ook de AN / AB chemotype. De lage katalytische efficiëntie van AMO van de HAP chemotype kan worden veroorzaakt door een deletie van zeven aminozuren aan de N-terminus van het eiwit vergeleken met CYP71AV1 van LAP. Ectopische expressie van de AN-PW genen in N. benthamiana bleek dat het grootste deel van de producten PW worden gemodificeerd door glycosylatie en glutathion vervegingen. Deze nevenreacties dus concurreren met de biosynthese flux in de richting van de AN voorloper DHAA. Op dit punt in mijn proefschrift de ectopische expressie van AN-PW genen in N. benthamiana niet had opgeleverd geen AN. In een later stadium werd het duidelijk dat dit te wijten was aan oogst van bladeren op 5-7 dagen na agro-infiltratie (dpi), terwijl een in N. benthami-
Samenvatting

ana bladeren uitdrukken AN-PW genen wordt pas waarnembaar na 7 dpi. Glycosylering van het grootste deel van de AN-PW producten in *N. benthamiana* benadrukt de noodzaak van een efficiënt transport van (DH) AA aan de buitenkant van de cellen om te ontsnappen uit de glycosylering reacties. In hoofdstuk 3, het vervoer en de opslag van AN voorlopers werd onderzocht door het bestuderen van het effect van membraan transporters (PDRs) en lipide overdracht eiwitten (LTPs). Hiervoor werden twee membraan transporters met activiteit ten opzichte van AN-PW producten ter beschikking gesteld door de groep van prof Marc Boutry en we geïsoleerd drie LTP genen van Artemisia annua, die uitdrukking in het glandulaire trichomen toonde. In dit hoofdstuk laten we zien dat AaLTP3 weergeeft specifieke activiteit, samen met AaPDR2 de richting van het transport van (DH) AA naar de apoplast in *N. benthamiana*. Bovendien infiltratie experimenten (DH) AA in *N. benthamiana* bladeren aangetoond dat deze verbindingen snel door de cellen worden opgenomen en dat in de cellen is er een sterke omgekeerde omzet in de AN-PVW door omzetting van (DH) AA richting (DH) en AAA (DH) AAOH. Vervolgens hebben we aangetoond dat AaLTP3 heeft een sterkere activiteit in het houden van producten in de apoplast dan de AaPDR2 membraan transporter. Daarom stel ik voor dat door het verwijderen van (DH) AA van het cytoplasma door middel van transport over de plasmamembracoan door AaPDR2 en de daaropvolgende vastlegging in de apoplast door AaLTP3, AaLTP3 creëert wastafel activiteit die reflux van (DH) AA uit de apoplast voorkomt terug in de cellen. AaLTP3 draagt derhalve bij tot een gerichte transport van AN-PW in de richting van het eindproduct (DH) AA. Tenslotte, in dit werk kunnen we ook voor de eerste keer detecteert AN en AB in *N. benthamiana* bladeren door extractie van necrotische bladeren bij 13 dpi.

Omdat in *A. annua* glandulaire trichome cellen zowel de AN sesquiterpene biosynthese route en de flavonoïde biosynthese route actief, onderzochten we of er een functionele interactie tussen deze twee belangrijke secundaire metaboliet biosynthese routes. In hoofdstuk 4 beschrijven we hoe we manipuleren van de flavonoïde biosynthese route in *N. benthamiana* bladeren met behulp van de Antirhinum majus transcriptiefactor Rosea1 (ROS) en test co-expressie van ROS met AN-PW genen. De co-expressie van ROS stimuleert AN-PW product accumulatie. Daaropvolgende analyse geeft aan dat dit waarschijnlijk van transcriptionele activatie van het enzym Mevalonaat Kinase (MVK) in mevalonaat pathway De precursors voorziet sesquiterpene biosyntheseroute. Verder tonen we aan dat de productie van flavonoïden concurreert met AN-PW productaccumulatie, als co-expressie van AN-PW genen met ROS, maar gelijktijdige remming van chalcon synthase (CHS)
Samenvatting

van een CHS\textsubp{RNAi} bouwen, tot hogere AN-PW product levels. Accumulatie van de eindproducten AN en AB werd niet significant beïnvloed. Tenslotte, de gecombineerde expressie van AN-PW + ROS + AaPDR2 + AaLTP3 + CHS\textsubp{RNAi} resulteert in hoogst sekwestratie van (DH) AA in de apoplast en hoogste accumulatie van de eindproducten AN en AB in \textit{N. benthamiana}.

Tijdens mijn proefschrift, in een gerelateerd project bleek dat de expressie van een sesquiterpeen biosynthese gen (caryofyleen synthase, CST) in transgene \textit{Arabidopsis} leidde tot hogere caryofyleen emissie voor een transformant expressie brengen van een genomisch DNA van CST, vergeleken met een soortgelijke transformant expressie brengen van een CST cDNA beschreven in de literatuur. Dit suggereerde dat ectopische expressie van intron bevattende genen efficiënter dan ectopische expressie van cDNA. Om te testen of in de context van metabolic engineering het gebruik van genoom (intron bevattende genen) efficiënter dan het gebruik van de overeenkomstige cDNA genereerden wij een aantal stabiele getransformeerde \textit{Arabidopsis}-lijnen met ofwel genomische CST (gCST), cDNA CST (cCST), genomische amorphadiene synthese (gADS) en cDNA ADS (cADS). In hoofdstuk 5 laten we zien dat inderdaad lijnen met overexpressie van het genomische klonen op hogere niveaus van de verwachte producten (caryophyleen of amorphadiene) dan de lijnen met overexpressie van de overeenkomstige cDNAs. Transcript analyse toonde dat gCST de toename caryofyleen productie hoger dan alleen kan worden verklaard door de toename in CST transcriptie. In de context van tijdelijke expressie in \textit{N. benthamiana} bladeren de intron-gemedieerde versterking effect was minder uitgesproken.

In het laatste hoofdstuk discussie te beoordelen 6 I beperkingen en mogelijke oplossingen voor metabolic engineering van de AN-PW in planten, en bespreek ik de impact van onze bevindingen over AN productiecapaciteit met behulp van transiënte expressie versus natuurlijke productie in \textit{A. annua}. Bovendien heb ik bespreken hoe de vondst van dit proefschrift gaan dan alleen inzichten in de AN-PW als met name de identificatie van de rol van LTPs in sesquestration van (sesqui) terpenen in de apoplast kan een impact hebben op de metabolic engineering inspanningen van vele andere hebben (sesqui) terpenen paden. Omdat sommige plantaaardige hormonen zijn ook terpenoid producten van het onlangs geïdentificeerd rol van LTPs kan ook invloed hebben op een dieper begrip van hormoon signalering in planten. Ik heb al begonnen met het verkennen van deze weg door het genereren van een reeks van \textit{Arabidopsis} planten met overexpressie van verschillende Arabidopsis LTP genen om te testen of een hormoon-gerelateerde kenmerken worden gewijzigd (hoofdstuk 6). Voorlopige resultaten wel degelijk
een rol van LTPs in endogeen plantenhormoon balans, iets wat de moeite waard verder worden onderzocht in de toekomst onderzoek te bevestigen.
Acknowledgements
Acknowledgements

This thesis could never been accomplished without the help from so many people. Here, I would like to express my deepest gratefulness and appreciation to my supervisors, colleagues, friends, family and those who have contributed to this thesis book in any possible way.

First of all, thank you Harro, for giving me the opportunity to be a member of Plant Physiology. Your optimistic leadership supported me to carry this PhD journey to the final line. Thanks to my daily supervisor, Sander. Thank you for your input and innovating ideas, for teaching me how to think, how to present, how to write and for all these interesting discussions we had. Thanks to my external supervisor, Jules, for encouraging me to continue when I feel depressed, for your nice advises.

I would like to express my sincere thanks to the people who helped or collaborated in this thesis. Thanks Francel for helping me the LC-trip-quad-MS analysis at the beginning of my PhD, Bert for the assistance of the LC-MS for the untargeted analysis, Norbert for the assistance for the subcellular localization work, Klass Bouwmeester for the help in the ML-II lab, Jacqueline and Marielle, Miriam for all of support in the lab during my PhD study. Thanks to Kristyna for sharing your knowledge of LC-MS and for the technical support of the LC-MS. Thanks to Thierry for the discussion and suggestions for the project. Thanks to Nikolay for the sharing important constructs. Thanks to Teade, Gerrit and Casper for all your support in preparing the soil and taking care of the plants in the greenhouse.

Thanks to my previous colleagues and friends: Jimmy, Liu Qing, Dong Lemeng, Yang Ting. Thanks to my current colleagues and friends: Arman, Esmer, Umidjon, Mark, Jun, Yuanyuan, Karen. Thanks to all members in the group of Plant Physiology.

To my lovely friends from the Chinese community in Wageningen (Du Yu, Yanxia, Hanzi, Yuanyuan, Yanting, Yunmeng, Junwei, Song wei, Bailing, Xi, Jianhua, Wei zhen, Yanli, Guiling, Tingting, Huchen, Zeng Tian, Cheng Xu, Zhu Feng, Song yin, Jinling, Wu jinbing, Sun Kaile and all other Chinese Friends. Thanks for all of my friends with who shared party or dinner together and accompanied me during my life in Netherlands.

致我亲爱的家人:
谢谢爸爸妈妈对我的无尽的爱和支持，千言万语凝聚在心头，却不能表达我
对你们的思念，很遗憾千里迢迢不能时常陪伴左右。自2011年开始，已是第五个春节漂泊在荷兰，谢谢公公婆婆对我的理解和爱护，婚姻的一种美好是有你们的爱和关心。谢谢哥哥，姐姐，妹妹，有了你们让我安心的远行，哥哥的思虑，姐姐的叮咛，妹妹的细腻一直鼓励着我一路前行。最后，感谢我的爱人，李耀旺，感谢你真诚的爱，感谢你在我失落时对我的鼓励和一路上风雨同舟的相守。我亲爱的嘉致，有你在我身边，生活才会绚丽。

Bo Wang
Leiden
07-02-2016
Curriculum vitae
Bo Wang was born on 10th of October 1984 in Anhui, China. She finished her high school education in 2004, and started her 4 years Bachelor education in Anhui Agricultural University. During her 4 years study in the Food quality and safety, she explored the interesting towards the science. In 2008, she got the Bachelor degree and moved to Beijing and started her Master education in China Agricultural University. There she studied Food Biotechnology and worked on the biosynthesis of hypericin in the medicinal plant, St. John’s wort. In 2011, she obtained her Master degree and got the fellowship from China Scholarship Council (CSC) for supporting her PhD study in Wageningen University, Netherlands. There she worked in the laboratory of Plant Physiology under the supervision of Harro Bouwmeester and Sander van der krol. During her PhD study she mainly focused on the project of biosynthesis and transport of artemisinin.
# Education Statement of the Graduate School

## Experimental Plant Sciences

### Issued to:
Bo Wang

### Date:
18 March 2016

### Group:
Laboratory of Plant Physiology

### University:
Wageningen University & Research Centre

### 1) Start-up phase

<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>First presentation of your project</td>
<td></td>
</tr>
<tr>
<td>Metabolic Engineering of Artesiminin (Ginghaosu) in Heterologous Host Plants</td>
<td>Nov 20, 2011</td>
</tr>
<tr>
<td>Writing or rewriting a project proposal</td>
<td></td>
</tr>
<tr>
<td>Metabolic Engineering of Artesiminin (Ginghaosu) in Heterologous Host Plants</td>
<td>Nov 20, 2011</td>
</tr>
<tr>
<td>Writing a review or book chapter</td>
<td></td>
</tr>
<tr>
<td>Laboratory use of isotopes</td>
<td></td>
</tr>
</tbody>
</table>

Subtotal Start-up Phase: 7.5 credits*

### 2) Scientific Exposure

<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPS PhD student days</td>
<td></td>
</tr>
<tr>
<td>EPS PhD student day, Leiden</td>
<td>Nov 21, 2013</td>
</tr>
<tr>
<td>EPS PhD Student Days “Größteher”, Soest</td>
<td>Jan 29, 2015</td>
</tr>
<tr>
<td>EPS theme symposia</td>
<td></td>
</tr>
<tr>
<td>EPS Theme 3 Symposium “Metabolism and Adaptation”, Utrecht University</td>
<td>Apr 26, 2012</td>
</tr>
<tr>
<td>EPS Theme 4 Symposium “Genome Biology”, Wageningen University</td>
<td>Dec 03, 2014</td>
</tr>
<tr>
<td>EPS Theme 1 Symposium “Developmental Biology of Plants”, Leiden University</td>
<td>Jan 08, 2015</td>
</tr>
<tr>
<td>EPS Theme 3 Symposium “Metabolism and Adaptation” Wageningen University</td>
<td>Feb 10, 2015</td>
</tr>
<tr>
<td>NWO Lunteren days and other National Platforms</td>
<td></td>
</tr>
<tr>
<td>ALW meeting Experimental Plant Scienes, Lunteren</td>
<td>Apr 02-03, 2012</td>
</tr>
<tr>
<td>ALW meeting Experimental Plant Scienes, Lunteren</td>
<td>Apr 14-15, 2014</td>
</tr>
<tr>
<td>ALW meeting Experimental Plant Scienes, Lunteren</td>
<td>Apr 13-14, 2015</td>
</tr>
<tr>
<td>Seminars (lunchees), workshops and symposia</td>
<td></td>
</tr>
<tr>
<td>Plant Engine W1 workshop: Modelling and databases for the enhancement of terpenes in plants</td>
<td>Feb 15-16, 2012</td>
</tr>
<tr>
<td>Seminar prof. Marc Bouvy</td>
<td>Mar 13, 2013</td>
</tr>
<tr>
<td>Seminar prof. Kazuki Saito</td>
<td>Apr 08, 2013</td>
</tr>
<tr>
<td>Seminar prof. Alain Tissier</td>
<td>May 03, 2013</td>
</tr>
<tr>
<td>Seminar prof. Eric Schrezen</td>
<td>Nov 21, 2013</td>
</tr>
<tr>
<td>Seminar Prof. Claus Schwechheimer</td>
<td>Oct 21, 2014</td>
</tr>
<tr>
<td>Phytopathology MINI-symposium</td>
<td>Nov 20, 2014</td>
</tr>
<tr>
<td>“Omics advances for academic and industry-towards true molecular plant breeding” Symposium</td>
<td>Dec 11, 2014</td>
</tr>
<tr>
<td>EPISAWAS seminar on disease resistance</td>
<td>Mar 10, 2015</td>
</tr>
<tr>
<td>EPS seminar Dr. Francois Percy</td>
<td>Oct 15, 2015</td>
</tr>
<tr>
<td>EPS Flying Seminar Dr. Alain Goossens</td>
<td>Dec 08, 2015</td>
</tr>
<tr>
<td>Seminar plus</td>
<td></td>
</tr>
<tr>
<td>International symposia and congresses</td>
<td></td>
</tr>
<tr>
<td>TerpNet meeting 2013, Crete, Greece</td>
<td>Jan 01-05, 2013</td>
</tr>
<tr>
<td>3rd international conference on “Plant Metabolic conference”, Kamer, China</td>
<td>Apr 02-05, 2014</td>
</tr>
<tr>
<td>Presentations</td>
<td></td>
</tr>
<tr>
<td>poster: Engineering artemisinin biosynthesis pathway: the next steps...</td>
<td>2013</td>
</tr>
<tr>
<td>poster: Transport in Artemisinin Biosynthesis Pathway</td>
<td>2014</td>
</tr>
<tr>
<td>IAB interview</td>
<td></td>
</tr>
<tr>
<td>Meeting with a member of the International Advisory Board of EPS</td>
<td>Jan 05, 2015</td>
</tr>
</tbody>
</table>

Subtotal Scientific Exposure: 11.7 credits*

### 3) In-Depth Studies

<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPS courses or other PhD courses</td>
<td></td>
</tr>
<tr>
<td>Jornal club</td>
<td></td>
</tr>
<tr>
<td>Plant physiology literature discussion group</td>
<td>2011-2015</td>
</tr>
<tr>
<td>Individual research training</td>
<td></td>
</tr>
<tr>
<td>Standardized Modular Cloning for Plant Synthetic Biology: Applications in metabolic engineering and fuel biotechnology, laboratory of Diego Orriols, Valencia, Spain</td>
<td>Sep 2013</td>
</tr>
</tbody>
</table>

Subtotal In-Depth Studies: 7.5 credits*

### 4) Personal development

<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skill training courses</td>
<td></td>
</tr>
<tr>
<td>ExPerceptionS Career Day</td>
<td>Nov 18, 2011</td>
</tr>
<tr>
<td>PhD Competence Assessment</td>
<td>Mar-Apr, 2012</td>
</tr>
<tr>
<td>Project and Time Management</td>
<td>Sep-Oct, 2012</td>
</tr>
<tr>
<td>Techniques for writing and presenting a scientific paper</td>
<td>Jul 2013</td>
</tr>
<tr>
<td>Organisation of PhD students day, course or conference</td>
<td></td>
</tr>
<tr>
<td>Membership of Board, Committee or PhD council</td>
<td></td>
</tr>
</tbody>
</table>

Subtotal Personal Development: 3.3 credits*

TOTAL NUMBER OF CREDIT POINTS*: 30.0

*Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS

*A credit represents a normative study load of 28 hours of study.
This work was performed at the Laboratory of Plant Physiology, Wageningen University, and financially supported by China Scholarship Council (CSC) and Wageningen University. Finalcial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover image: Artemisia annua; foot print of Jiazhi Li
Cover design and layout:
Yaowang Li
Bo Wang

Printed by: GVO drukkers & vormgevers B.V.