Withanolides:
Elucidating steroidal lactone biosynthesis in Nightshades

Von der Naturwissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades
Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation
von
Annika Stein, M. Sc.

2022

Referent: Prof. Dr. rer. nat. Jakob Franke
Korreferent: Prof. Dr. Russell J. Cox
Tag der Promotion: 04.10.2022

Abstract

Key Words: withanolides • biosynthesis • *Withania somnifera* • cytochrome P450 oxygenases

Withania somnifera (Solanaceae) is well known in ayurvedic medicine as a strengthening tonic for various medical purposes. Its effects are mainly due to withanolides, a class of steroidal lactones with diverse oxidation patterns present in various nightshade plants. Pharmacological studies attributed anti-proliferative and anti-inflammatory properties to withanolides. However, obtaining medicinally active withanolides can be complicated, as complex mixtures are present in producing plants and total synthesis of withanolides is costly and time consuming. Therefore, investigation of their biosynthesis is important to enable biotechnological enhancement and to provide novel insights into plant steroid biochemistry.

This work aimed to investigate withanolide biosynthesis in *Physalis peruviana* and *Withania somnifera.* Both plants were investigated for their main withanolides, as producers can accumulate a diverse array of withanolides, depending on the cultivation conditions. Here, besides several known withanolides, two yet unknown, truncated withanolides (irinan A (**1**) and B (**2**)) were isolated from *P. peruviana* and their structures were elucidated.

As intermediates of withanolide biosynthesis were needed for enzyme assays but are neither known, nor commercially available, metabolic engineering in yeast was attempted to divert yeast ergosterol biosynthesis towards production of 24-methyldesmosterol (**3**), the last known precursor in withanolide biosynthesis. However, while production of the precursor 24-methylenecholesterol (**4**) was temporarily observed, **3** did not accumulate.

Furthermore, based on the biosynthetic hypothesis, 21 cytochrome P450 (P450) and 14 dehydratase (DH) gene candidates were selected after analysis of three withanolide-producing species. Of those, 17 P450 and 6 DH candidates could be cloned and evaluated by gene silencing in *W. somnifera,* identifying 5 P450 and 2 DH gene candidates where silencing evoked significant decrease of the main withanolide (withaferin A, **5**). Those candidates were further examined by heterologous expression experiments in the model plant *Nicotiana benthamiana*. Here activity on the substrate 24-methyldesmosterol (synthesized by Dave Biedermann) was detected for one candidate (P450-7), while another exhibited activity on native cycloartenol (**6**) from the host plant (P450-17). Further investigation of P450-17 revealed that orthologs were present in tomato and potato, both non-producers of withanolides. In both plants P450-17 homologous genes are arranged in gene clusters, with neither the genes nor the cluster being reported before. In conclusion, this work provides insights into oxidations involved in withanolide biosynthesis and yet unknown phytosterol pathways in Solanaceae plants.

Zusammenfassung

Schlagwörter: Withanolide • Biosynthese • *Withania somnifera* • Cytochrom P450 Oxygenasen

Withania somnifera (Solanaceae) ist eine in der ayurvedischen Medizin weithin als vielseitiges Heilmittel bekannte Pflanze. Ihre Wirksamkeit beruht auf Withanoliden, eine Gruppe hoch oxidierter Steroidlactone mit anti-proliferativen und anti-inflammatorischen Eigenschaften, welche in diversen Nachtschattengewächsen vorkommen. Aufgrund komplexer Withanolidgemische ist eine Aufreinigung medizinisch wirksamer Verbindungen für Anwendungszwecke schwierig, während Totalsynthese durch die komplexen Strukturen zeitaufwendig und teuer ist. Somit ist die Aufklärung der Withanolidbiosynthese von Interesse, um biotechnologische Produktion und medizinische Anwendung zu ermöglichen.

Diese Arbeit befasst sich mit der Withanolidbiosynthese in *Physalis peruviana* und *Withania somnifera.* Beide Pflanzen wurden auf ihre Hauptwithanolide untersucht. Hierbei wurden neben bekannten Withanoliden zwei bisher unbekannte, verkürzte Androwithanolide (Irinan A (**1**) und B (**2**)) aus *P. peruviana* isoliert und charakterisiert.

Da Intermediate der Withanolidbiosynthese für Assays benötigt wurden, jedoch nicht bekannt bzw. kommerziell erhältlich sind, wurde versucht einen Hefestamm zu generieren, der das letzte bekannte Intermediat, 24-Methyldesmosterol (**3**), durch Manipulation der Ergosterol-Biosynthese produziert. Hierbei konnte das Intermediat 24-Methylenecholesterol (**4**) detektiert werden, die Akkumulation von **3** wurde jedoch nicht beobachtet.

Basierend auf einer biosynthetischen Hypothese und der Analyse von Transkriptomdaten dreier Withanolidproduzenten wurden 21 Cytochrom P450 (P450) Oxygenase-Kandidaten und 14 Dehydratase (DH)-Kandidaten zur weiteren Testung ausgewählt. Davon konnten 17 P450- und 6 DH-Kandidaten kloniert und in Gensilencing-Experimenten in *W. somnifera* untersucht werden. Silencing bewirkte bei 5 P450- und 2 DH-Kandidaten eine Verringerung des Hauptwithanolides (Withaferin A, **5**). Des Weiteren wurden diese Kandidaten in der Modellpflanze *Nicotiana benthamiana* exprimiert. Hierbei wurde Aktivität von P450-7 an dem Substrat (24-Methyldesmosterol) festgestellt, während P450-17 das native Cycloartenol (**6**) der Wirtspflanze oxidierte. Weitere Untersuchungen von P450-17 zeigten orthologe Gene in Kartoffeln und Tomate, beides keine Withanolidproduzenten. In beiden Pflanzen sind die homologen Gene in Genclustern arrangiert; weder die Gene noch das Cluster wurden zuvor beschrieben. Insgesamt bietet diese Arbeit neue Einblicke in Oxidationen der Withanolidbiosynthese sowie in weiteren unbekannten Phytosterolbiosynthesewegen in Solanaceen.

Acknowledgement

First, I would like to thank Prof. Jakob Franke for the excellent guidance through this challenging project. Building a new lab takes time and effort, thank you for including me in this process and giving me valuable insights. You were always there when I needed help and gave me the opportunity to develop many new skills and grow my confidence. I am grateful for having you not only as a scientific mentor, but also as a role model on how to be a leader.

I would also like to thank Prof. Russell Cox, who sparked my interest in natural products during my master's studies and for being the co-examiner.

I am grateful for my students turned co-workers Dave Biedermann and Samuel E. Hakim for contributing to this project with their chemical skills. Further I would like to thank my dear co-worker Ling Chuang for suffering through various lab problems with me; monkeys are stronger together indeed!
In research you often need help from other groups. Therefore, I would like to thank Bianka Karge and Prof. Mark Brönstrup for collaboration on the Irinan project. Further I would like to thank Dr. Marco Herde and Prof. Claus-Peter Witte for giving me access on their Orbitrap and providing us the Agrobacteria strains needed to infect my plants.
Moreover, I would like to thank all my students that contributed to this project: Lara Tenge, Katrin Duman, Marcel Arndt, Arne Bültemeier and Mahsum Kadah; and further all past and current co-workers in the Franke and Cox group, especially Sara, Karan, Shenyu, Lukas and Carsten for creating a great work environment together.

Special appreciation has to be given to all technical staff involved in this project. This includes BMWZ technicians under the lead of Katja Körner, I am sure this building would not function without you. Further I would like to thank the NMR department for always delivering fast measurements of my samples. Moreover, special appreciation goes out to Oliver Stahlschmidt and the ZVA Team and Yvonne Leye and Miriam Fent for not only taking care of my plants, but also sparking my interest in gardening.

Lastly, I would like to thank my family for their constant support throughout years of studying.

Abbreviations and Units

24ISO	24-isomerase	m-	milli-
µ	micro-	*m/z*	mass to charge ratio
A. thaliana	*Arabidopsis thaliana*	MEP	2-C-methyl-D-erythritol 4-phosphate
aa	amino	MES	2-(N-morpholino)ethanesulfonic acid
ATP	adenosine triphosphate	min	minute
BLAST	basic local alignment search tool	MS	mass spectrometry
bp	basepair	MVA	mevalonic acid
carb	carbenicillin	n	nano-
CoA	coenzyme A	*N. benthamiana*	*Nicotiana benthamiana*
COSY	correlation spectroscopy	NADPH	nicotinamide adenine dinucleotide phosphate
d	days	nat	nourseothricin
DAD	diode array detector	NMR	nuclear magnetic resonance
DH	dehydratase	ORF	open reading frame
DMAPP	dimethylallyl pyrophosphate	*P. alkekengi*	*Physalis alkekengi*
DNA	deoxyribonucleic acid	*P. floridana*	*Physalis floridana*
dpi	days post infiltration	*P. peruviana*	*Physalis peruviana*
E. coli	*Escherichia coli*	P450	cytochrome P450
EDTA	ethylenediaminetetraacetic acid	PCR	polymerase chain reaction
EI	electron ionization	PDS	phytoene desaturase
EIC	extracted ion chromatogram	PEG	polyethylene glycol
ELSD	evaporative light scattering detector	ppm	parts per million
EV	empty vector	qPCR	quantitative PCR
FPP	farnesyl pyrophosphate	rif	rifampicin
GCMS	gas chromatography mass spectrometry	RNA	ribonucleic acid
gDNA	genomic DNA	rpm	revolutions per minute
gent	gentamycin	rRNA	ribosomal RNA
GOI	gene of interest	RT	room temperature
GRAS	generally recognised as safe	RT-PCR	reverse transcriptase-PCR
h	hours	*S. cerevisiae*	*Saccharomyces cerevisiae*
HPLC	high performance liquid chromatography	*S. lycopersicum*	*Solanum lycopersicum*
HRMS	high resolution mass spectrometry	*S. tuberosum*	*Solanum tuberosum*
Hz	Hertz	TAE	tris-acetate-EDTA
IPP	isopentenyl pyrophosphate	TIC	total ion current
kan	kanamycin	tR	retention time
kDa	kilo Dalton	Tris	2-Amino-2-(hydroxymethyl)propane-1,3-diol
KO	knockout	UV	ultraviolet
L	Litre	VIGS	virus induced gene silencing
LCMS	liquid chromatography mass spectrometry	*W. somnifera*	*Withania somnifera*
M	molar	WT	wild-type

Table of Contents

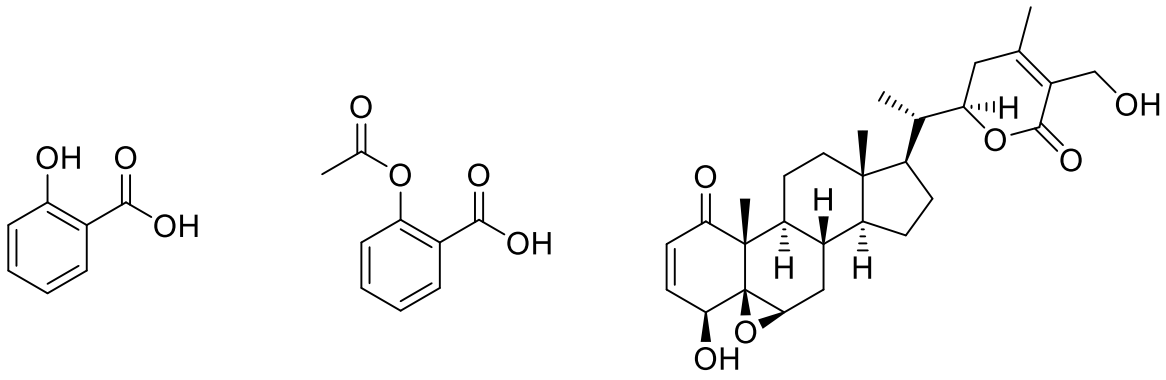
1 Introduction

1.1 Plant natural products provide an important source for drug development

As sessile organisms, plants require a high line of defense against microbial attacks, virus infections and animal or insect feeding. Over 100,000 natural products are known from plants, which present a wide array of bioactivity. [1] Since many of these compounds are naturally active against intruders, it is not surprising that humankind has used their properties for centuries. To date, drugs that are inspired by natural products still comprise 41% of medicinally used compounds. [2]

In an attempt to transfer the knowledge of traditional medicine to modern drugs, many of the well-known household remedies have been screened for their active agents. A popular example for such a success story is acetylsalicylic acid (**7**, Figure 1.1), commonly known as aspirin. The leaves and barks of willow trees were used to treat fever and pain in many cultures across many centuries. In the 19th century, salicylic acid (**8**) was discovered as the molecule of action and at the end of the century a derivative with reduced side effects, acetylsalicylic acid, conquered the market. [3] Like in this example, natural products often function as a template for further modification to reduce side effects or improve the bioavailability of the respective drug.

Another example for plant secondary metabolites discovered from traditional medicine are withanolides (e.g. withaferin A, **5**), steroidal lactones present in various nightshades. Due to their various medicinal properties, complex structure and unsolved biosynthesis, this class of compounds has sparked attention for further research in our group and will be discussed in the following section.



salicylic acid (**8**) acetylsalicylic acid (**7**) withaferin A (**5**)

Figure 1.1 Natural products such as salicylic acid (**8**) and withaferin A (**5**) can provide templates for derivatives with enhanced properties (e.g., acetylsalicyclic acid, **7**).

1.2 Withanolides are medicinally important plant natural products

In traditional ayurvedic medicine, the nightshade plant *Withania somnifera* is known as Ashwaghanda. The powdered root is used as a strengthening tonic for various medical purposes. [4] Similar properties have been described in traditional Chinese medicine for *Physalis angulata* or in traditional South American medicine for *Physalis peruviana.* [5,6]

Those properties can be ascribed to a class of compounds called withanolides. [5,6] The first withanolide was isolated from *W. somnifera* in 1965 and was called withaferin A (**5**). [7] During the last two decades, pharmacological studies attributed various medicinal properties to withanolides, with anti-proliferative and anti-inflammatory properties being the most prominent ones. [6,8,9]

The metabolite profiles of various cultivars of *W. somnifera* have been extensively analyzed, providing numerous different compounds. [7,10–13] Furthermore, withanolides are found in various genera of the Solanaceae family, as for example *Datura*, *Dunalis*, *Iochroma*, *Jaborosa*, *Lycium*, *Physalis* and *Withania*. [14] Withanolides can exhibit various modifications on the skeleton and further diverse oxidations, providing a wide array of different subclasses. The number of known withanolides is growing rapidly, more than 360 known to date. [14,15]

Withanolides are C_{28} steroidal lactone triterpenoids that, in most cases, are based on an ergostane skeleton and furthermore exhibit a complex oxidation pattern. The modifications of the carbocyclic skeleton as well as in the side chain divide them into different subclasses. A major distinction is made with the substitution pattern of the side chain with δ-lactones or δ-lactols comprising the type A withanolides while γ-lactones or γ-lactols form the subclass of type B withanolides (Figure 1.2A). [14] The more common class of type A withanolides (termed 22-hydroxyergostane-26-oic acid 26,22-lactones) can be further divided into compounds with an unmodified or modified skeleton. Examples for type A withanolides with a modified skeleton include withaphysalins, physalins or withajardins (Figure 1.2B). [14] Each subclass contains many reported compounds, e.g., with unmodified type A withanolides harboring more than 40 different reported structures to date. [4,16]

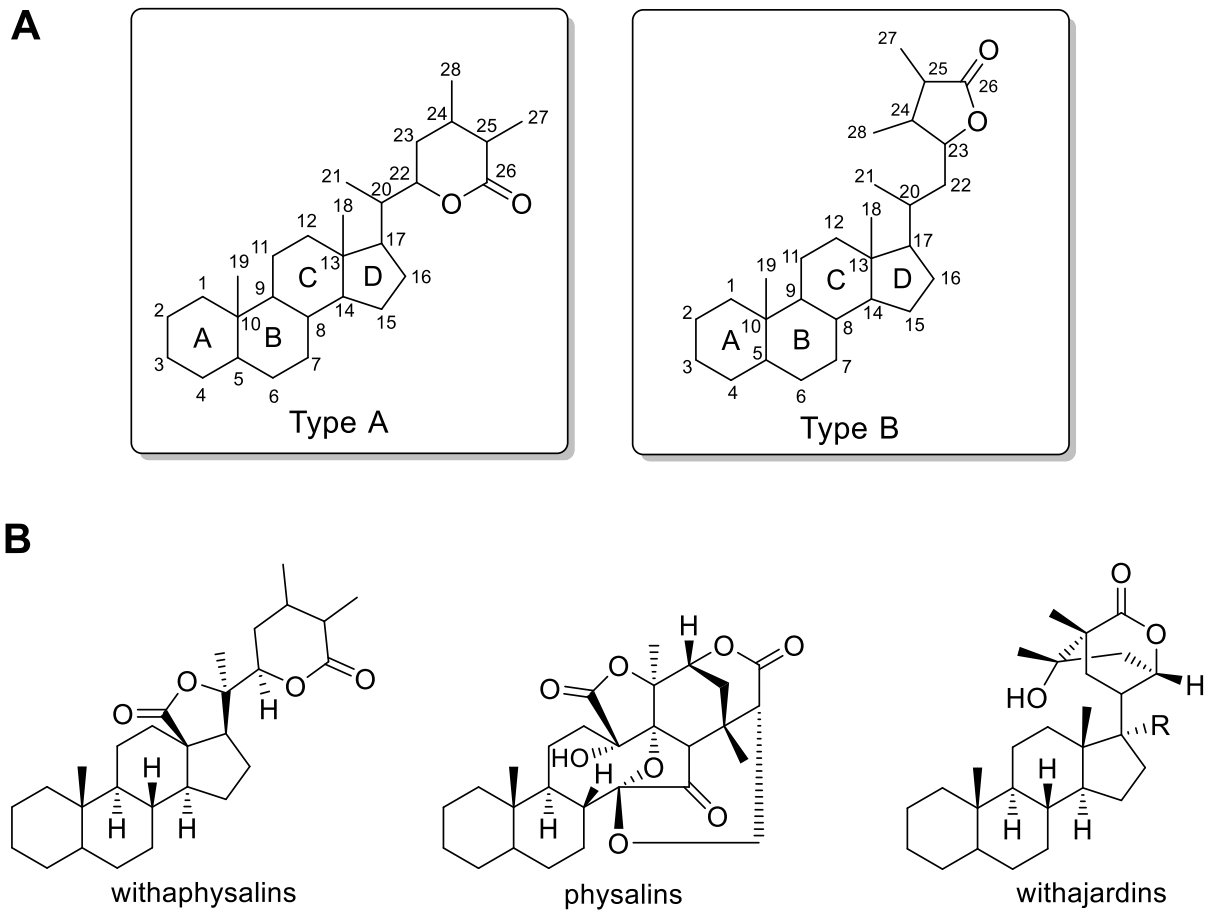


Figure 1.2 Withanolides can be divided into two major subclasses: Type A and Type B (A); Examples for subclasses of Type A withanolides with a modified skeleton include withaphysalins, physalins and withajardins (B).

As withanolides are known for their antitumor activity, the relationship between chemical structure and bioactivity has been investigated. In general, withanolides featuring an α,β-unsaturated ketone in ring A, a 5β,6β-epoxy group in ring B and a nine-carbon side chain with an α,β-unsaturated δ-lactone display the highest antitumor activity (for ring nomenclature see Figure 1.2A). [17,18] Their medicinal properties make them desired compounds for further drug development. However, yields in plants are not sufficient for clinical studies or large-scale medicinal application. Furthermore, chemical synthesis can be very complex due to multiple stereocenters, resulting in high costs if applied for a larger scale. However, total synthesis of withanolide A (**9**) [19] or of the complex right part of physalins has been reported. [20] Another approach that has come into focus is production by fermentation. However, in contrast to the medicinal properties of withanolides, few research has been conducted on their biosynthesis (see section 1.3). However, their complex oxidation pattern sparked interest in this group. As this work will aim to gather further insights into withanolide biosynthesis, this topic will be discussed in the following section.

1.3 Biosynthesis of Terpenes, Phytosterols and Withanolides

In general, natural products can be divided into primary metabolites, essential for the survival of an organism, and specialized metabolites, which are not essential, but provide an advantage for the organism. Specialized metabolites can be further classified by many criteria, with the most common system being the classification by structure. The main classes are polyketides like lovastatin (**10**), [21] nonribosomal peptides like penicillin G (**11**), [22] alkaloids like morphine (**12**) [23] and terpenes like artemisinin (**13**, Figure 1.3). [24] Of those classes, only nonribosomal peptides have not been identified in plants.

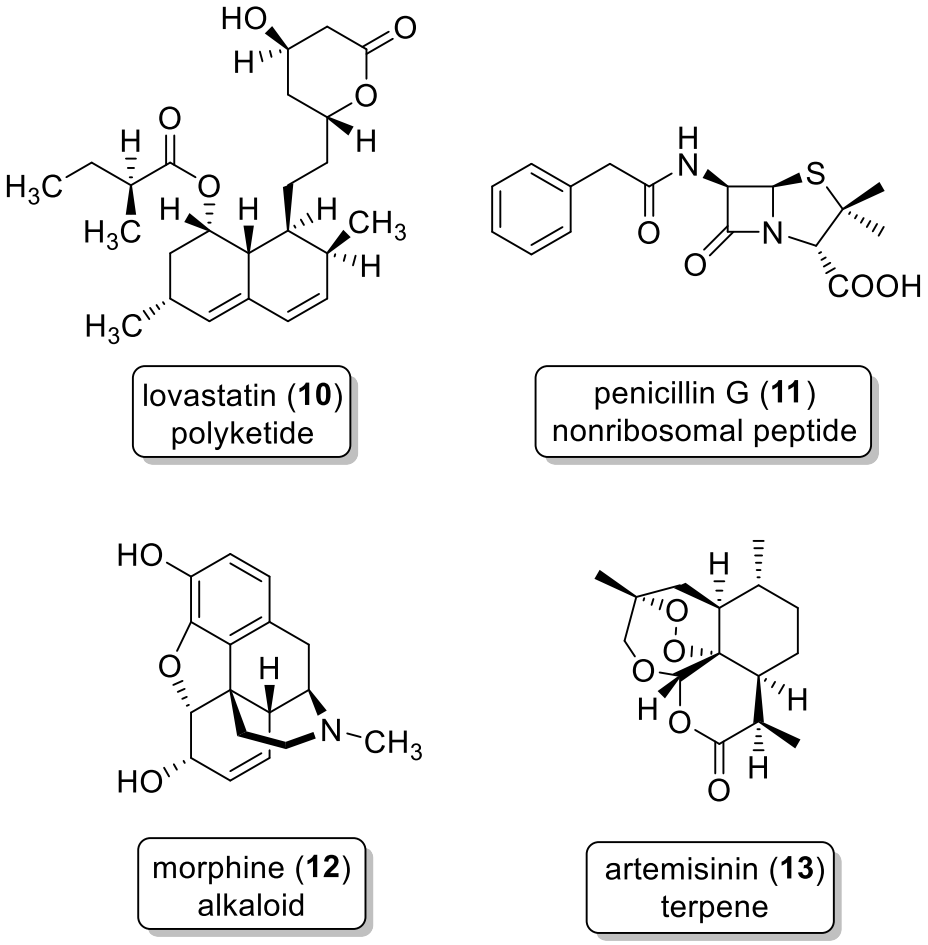


Figure 1.3 Secondary metabolites are typically classified into polyketides (e.g, lovastatin, **10**), [21] nonribosomal peptides (e.g., penicillin G, **11**), [22] alkaloids (e.g., morphine, **12**) [23] and terpenes (e.g., artemisinin, **13**). [24]

In plants, terpenes are the structurally most diverse group of specialized metabolites. [25] And while the biosynthesis of withanolides has not been entirely elucidated, research showed that their steroidal backbone is based on the phytosterol biosynthetic pathway and consequently involve terpene building block formation. In the following chapter, this pathway should be

discussed in detail to provide necessary information for the following research and integrate research about withanolide biosynthesis into a biosynthetic context.

1.3.1 Terpene biosynthesis builds the base for phytosterol and withanolide biosynthesis

The sterol skeleton is a structure with biological significance in all eukaryotes (e.g. for membrane fluidity). [26] However, different compounds of this group can be observed in plants (e.g. β-sitosterol **14**) compared to animals (cholesterol **15**) and fungi (ergosterol **16**, Figure 1.4A). [26] In any case, sterols can be classified as terpenoids (terpenes with additional functional groups) based on their structure.

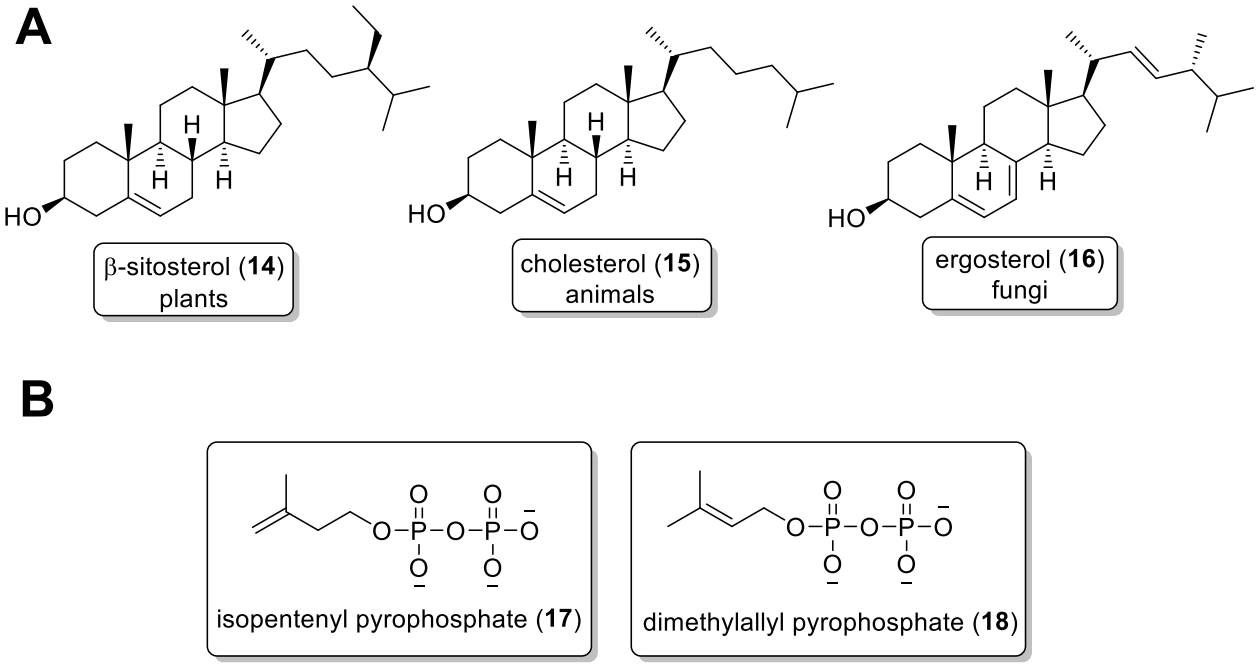
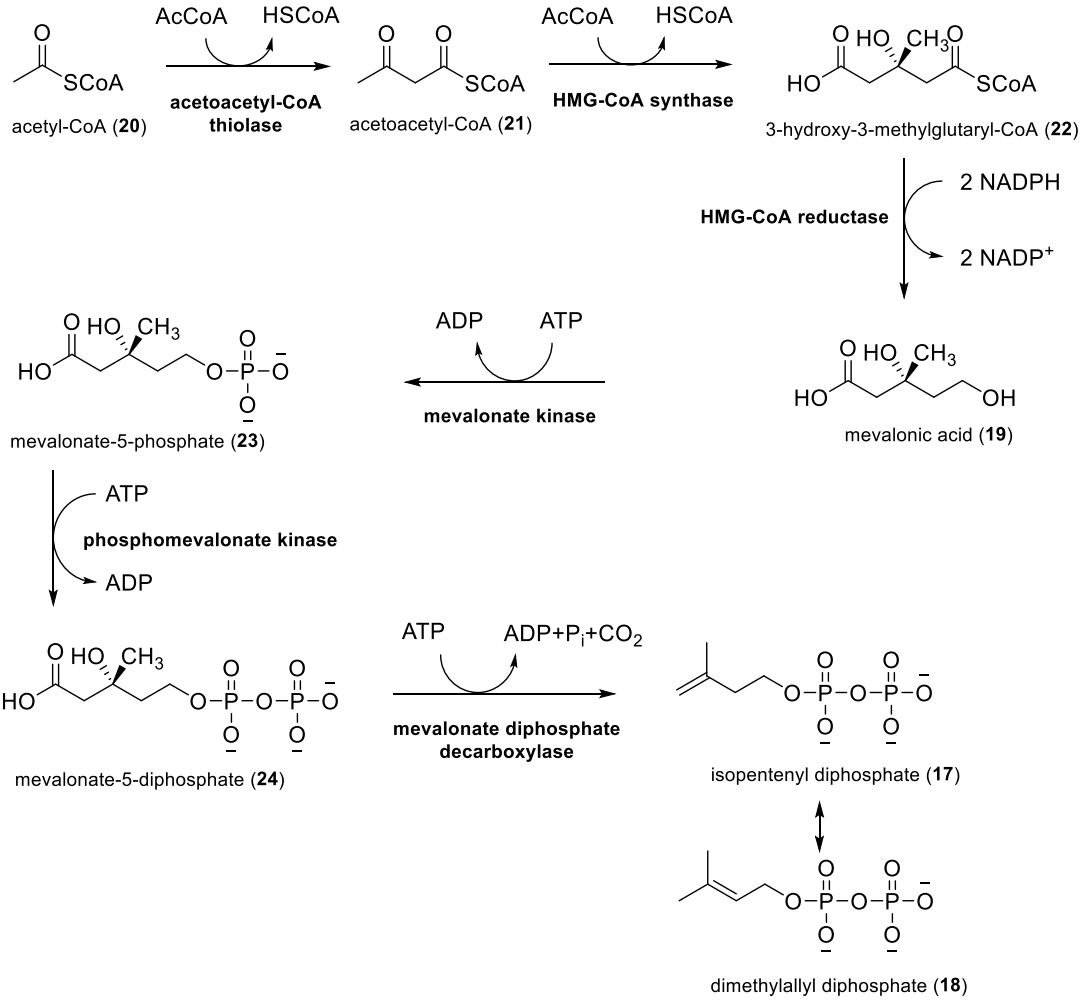


Figure 1.4 Different sterols are observed in plants (e.g. β-sitosterol, **14**), animals (cholesterol, **15**) and fungi (ergosterol, **16**, A); isopentenyl pyrophosphate (**17**, IPP) and dimethylallyl pyrophosphate (**18**, DMAPP) are the building blocks for C_5 isoprene units in terpene biosynthesis (B).

Terpenes and terpenoids are derived from C_5 isoprene units, with isopentenyl pyrophosphate (**17**, IPP) and dimethylallyl pyrophosphate (**18**, DMAPP) being the building blocks for those isoprene units (Figure 1.4B). In plants those isoprene units can be built via two independent pathways in separate cellular compartments.[25] In the cytosol, the building block isopentenyl diphosphate (**17**) is produced via the mevalonic acid (**19**, MVA) pathway (Scheme 1.1). [27] Here, two units of acetyl-CoA (**20**) are condensed to acetoacetyl-CoA (**21**) by the acetoacetyl-CoA thiolase, followed by a second condensation of an additional acetyl-CoA unit by the hydroxymethylglutaryl-CoA synthase (HMG-CoA synthase). [28,29] Afterwards the resulting 3-

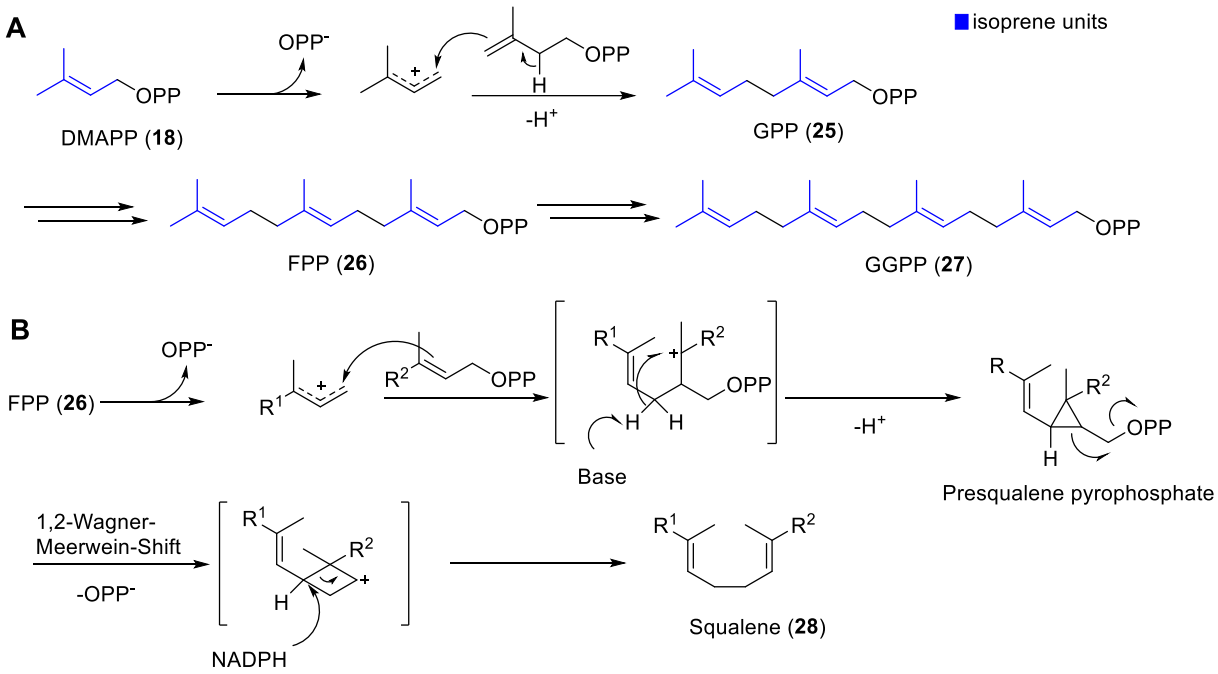
hydroxy-3-methyl-glutaryl-CoA (**22**) is reduced to mevalonate (**19**), with this reaction being the rate limiting step in sterol pathways. [30] Subsequently, two phosphate moieties are transferred onto the mevalonic acid by kinases, forming mevalonate-5-phosphate (**23**), then mevalonate-5-diphosphate (**24**). [28,29] Lastly, conversion towards IPP (**17**) is catalyzed by a mevalonate diphosphate decarboxylase under consumption of another molecule of ATP. [28,29] Alternatively, IPP (**17**) can be formed via pyruvate and glyceraldehyde-3-phosphate in the plastids, following the methylerythritol phosphate (MEP) pathway. [31] There is evidence that isoprene units are exchanged between the two pathways. [32] Tracing experiments of ^{13}C-D-glucose lead to the assumption, that both pathways are involved in withaferin A (**5**) synthesis, with a ratio of 25:75 (MEP:MVA), as different patterns of ^{13}C labeling are created in the two pathways. [32] Furthermore a HMG-CoA reductase of the MVA pathway and the 1-deoxy-D-xylulose-5-phosphate synthase and 1-deoxy-D-xylulose-5-phosphate reductase of the MEP pathway have been characterized in *W. somnifera*. [33,34]



Scheme 1.1 Mevalonic acid pathway for terpene biosynthesis.

In the following steps of terpene biosynthesis, IPP (**17**) and its allylic isomer dimethylallyl diphosphate (**18**) can be condensed in a head-to-tail reaction. The reactions are catalyzed by the respective prenyl transferases (e.g. farnesyl diphosphate synthase, FPPS) and result in geranyl diphosphate (**25**, GPP), farnesyl diphosphate (**26**, FPP) and geranylgeranyl diphosphate (**27**, GGPP), the precursors for monoterpenes, sesquiterpenes and diterpenes (Scheme 1.2A).[35] The *W. somnifera* FPPS cDNA has been isolated, and expression levels in different tissues seem to mirror withanolide content. [36,37]

Furthermore, head-to-head (also termed tail-to-tail) reactions can be catalyzed by prenyltransferases to form C_{30} or C_{40} chains, as for example sterols and carotenoids (Scheme 1.2B). [38] The condensation of two farnesyl pyrophosphate units to squalene (**28**) marks a branching point towards phytosterol biosynthesis and is catalyzed by the squalene synthase (SQS). [39] The *W. somnifera* SQS has been characterized and expression experiments and tissue specific expression analysis underlined the involvement of SQS in withanolide biosynthesis. [40]



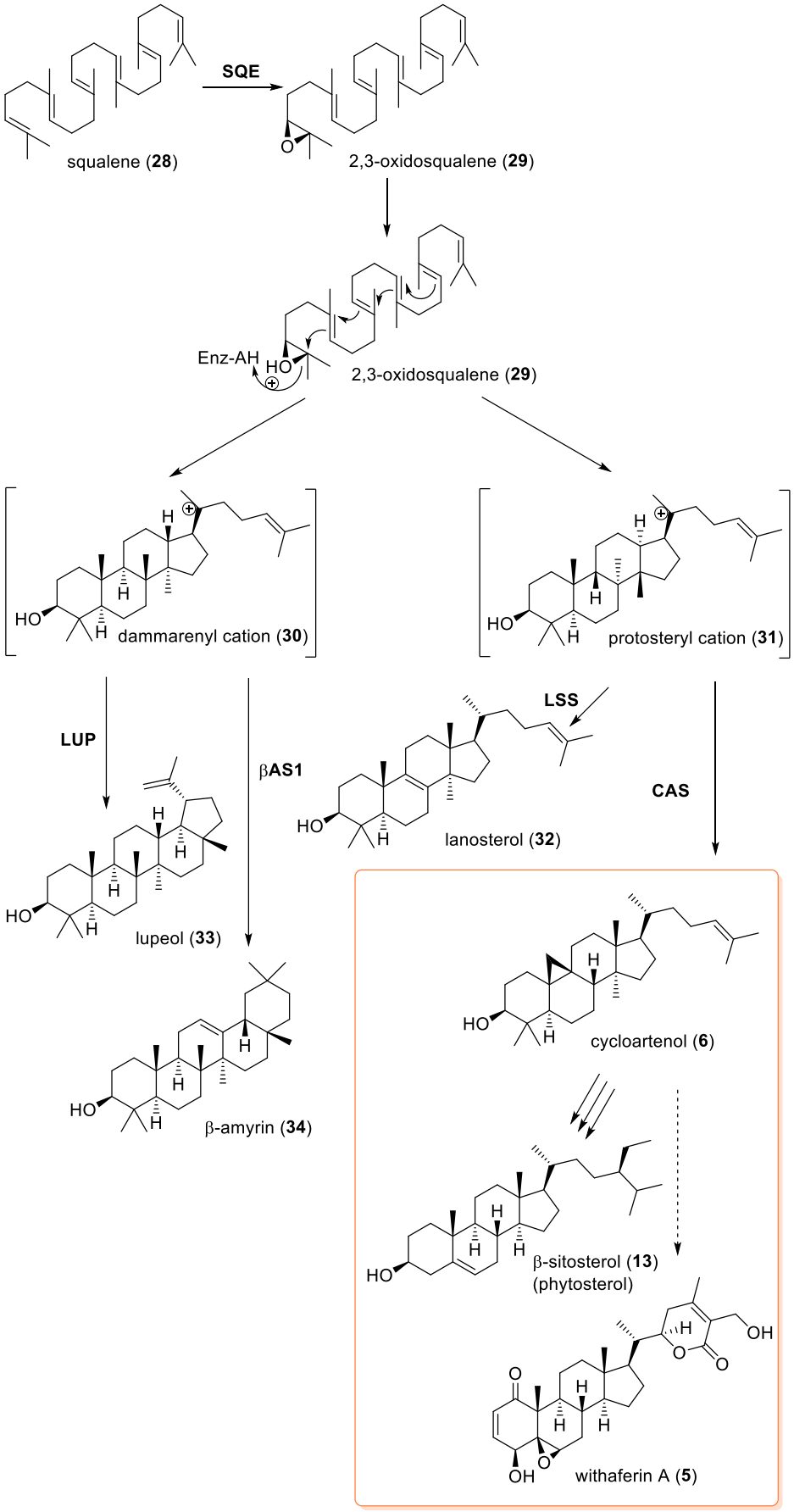
Scheme 1.2 Head-to-tail condensation of IPP (**17**) and DMAPP (**18**, A) [41] and head-to-head condensation for the formation of squalene (**28**, B).[42,43]

1.3.2 Withanolide biosynthesis eventually branches from the phytosterol pathway

After formation of squalene (**28**), the sterol biosynthesis is continued with the conversion of squalene to 2,3-oxidosqualene (**29**), a rate limiting step catalyzed by the squalene epoxidase (SQE). [44] The *Withania somnifera* SQE was cloned and characterized and tissue specific expression analysis suggested involvement of SQE in withanolide biosynthesis. [45]

2,3-Oxidosqualene (**29**) is an important intermediate in sterol biosynthesis, that is cyclized by protonation of the epoxide, triggering a cascade of carbon-carbon bond formations (Scheme 1.3). [46] This mechanism is executed by a class of enzymes termed oxidosqualene cyclases (OSC). Depending on the folding of 2,3-oxidosqualene (**29**) substrate in the enzyme, different structures are created. The chair-chair-chair conformation yields the dammarenyl cation (**30**), while the chair-boat-chair conformation results in formation of the protosteryl cation (**31**, Scheme 1.3). [47]

In animals and fungi 2,3-oxidosqualene (**29**) is mainly converted to lanosterol (**32**) for cholesterol (**15**) or ergosterol (**16**) biosynthesis via the protosteryl cation. [48] In plants however, 2,3-oxidosqualene (**29**) is processed to a variety of triterpenoid skeletons. The dammarenyl cation (**30**) yields structures such as lupeol (**33**) or β-amyrin (**34**), while the protosteryl cation results in the intermediate cycloartenol (**6**, Scheme 1.3). [49] This compound is the first step towards phytosterol and phytohormone (brassinolide) biosynthesis. [49] Against common belief, cholesterol (**15**) biosynthesis exists in plants as well, and cross talk between the cholesterol (**15**) biosynthetic pathway and phytosterol metabolism has been reported. [50]



Scheme 1.3 Cyclization of 2,3-oxidosqualene (**29**) results in cycloartenol formation among others, building the base for phytosterol and withanolide biosynthesis. (adapted from Phillips *et al.*) [49]

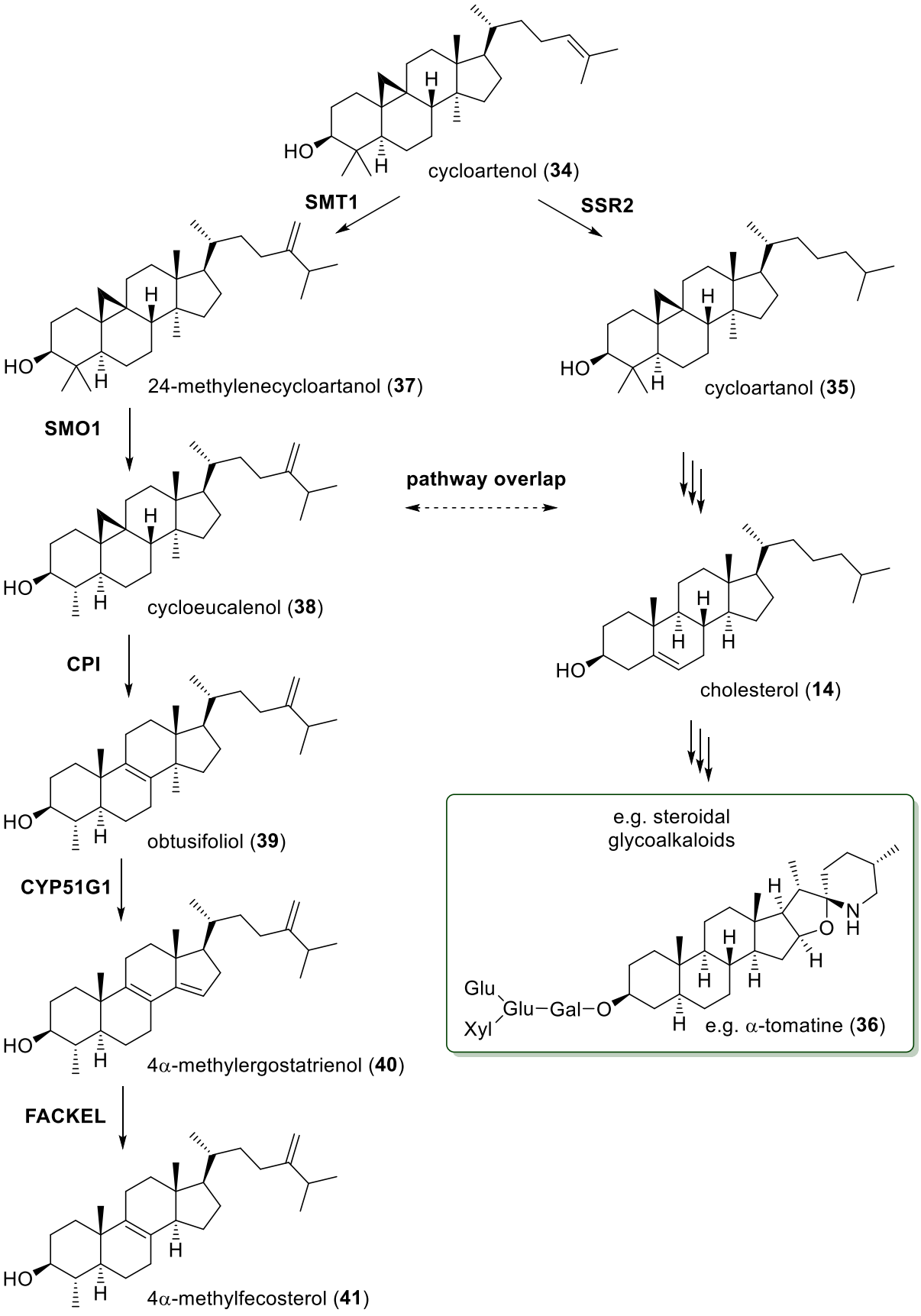
Cyclization of 2,3-oxidosqualene (**29**) marks a branching point between sterol/withanolide biosynthesis and other triterpenoids such as β-amyrin (**34**). In *W. somnifera* oxidosqualene cyclases β-amyrin synthase (βAS), lupeol synthase (LUP) and cycloartenol synthase (CAS) have been characterized. [51] The latter is the most likely candidate to be involved in withanolide biosynthesis, as it follows the path towards phytosterols. [51]

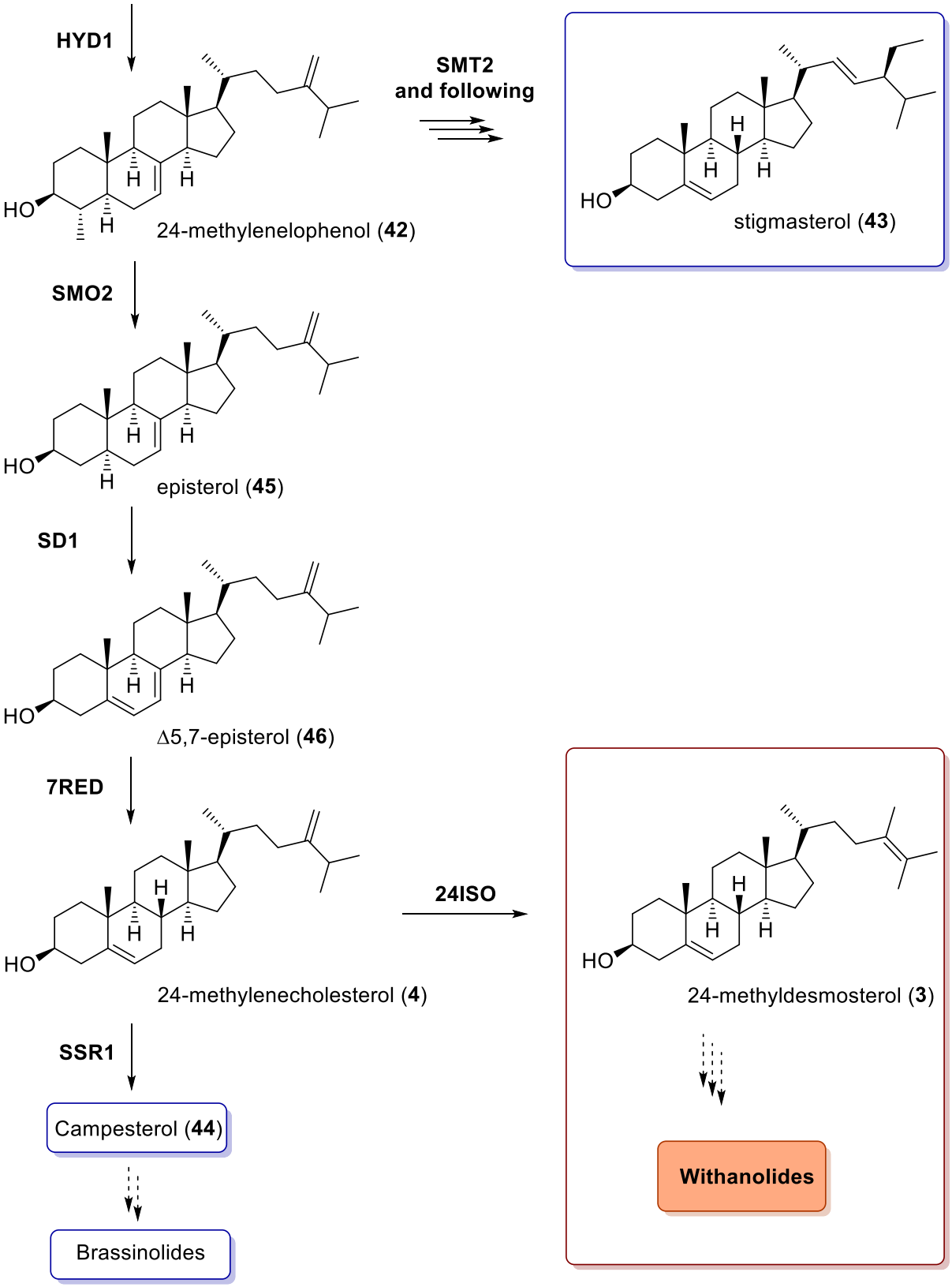
From this point onwards, research regarding withanolide biosynthesis specifically gets scarce. In general, cycloartenol (**6**) can be further processed in two directions. Reduction of the side chain towards cycloartanol (**35**) by a sterol side chain reductase enzyme (SSR2) marks the branching point towards biosynthesis of cholesterol (**15**) and more specialized metabolites such as steroidal glycoalkaloids (e.g. α-tomatine, **36**, Scheme 1.4). [52]

A distinguishing feature of phytosterols is the alkylation at the C-24 position. [53] Consequently in phytosterol biosynthesis cycloartenol (**6**) is methylated to 24-methylenecycloartanol (**37**) by a sterol methyltransferase (SMT1, Scheme 1.4). [52] It has to be noted that plant cholesterol and phytosterol biosynthesis show significant overlap due to promiscuous activity of the enzymes involved. [52] For clarity, only the common phytosterol pathway will be described further.

Further steps in phytosterol biosynthesis include demethylation at C-4 to cycloeucalenol (**38**, SMO1), [54] cleavage of the cyclopropane ring forming obtusifoliol (**39**, CPI), [55] demethylation at C-14 that gives 4α-methylergostatrienol (**40**, CYP51G1), [56] reduction at C-14 forming 4a-methylfecosterol (**41**, FACKEL) [57] and isomerization of the double bond from C-8 to C-7 (HYD1) [58] to reach the intermediate 24-methylenelophenol (**42**, Scheme 1.4). Here, phytosterol biosynthesis branches again. While the following reactions are identical, SMT2 can extend the alkylation at the C-24 position yielding stigmasterol (**43**) instead of campesterol (**44**). [59] In any case, intermediates are subsequently demethylated at C-4 (SMO2), [60] forming episterol (**45**) in case of campesterol (**44**) biosynthesis, followed by an elimination at C-5 (Δ5,7-episterol, **46**, SD1) [61] and reduction of the double bond at C-7 (7RED, Scheme 1.4). [62] When processing 24-methylenelophenol this results in the intermediate 24-methylenecholesterol (**4**). Here another important branching point is reached. While reduction of the side chain at the C-24 position (SSR1) [63] yields the phytosterol campesterol (**44**), withanolide biosynthesis branches by isomerization of the side chain resulting in the intermediate 24-methyldesmosterol (**3**, 24ISO, Scheme 1.4). [64] Activity of the 24ISO was confirmed by expressing a Δ24-isomerase from *Withania somnifera* in a 24-methylenecholesterol (**4**) producing yeast strain and involvement of the 24ISO in withanolide biosynthesis was verified by silencing experiments in *W. somnifera*. [64] Furthermore 24-methylenecholesterol (**4**) was confirmed as a withanolide

precursor by labeling studies. [65] 24-Methyldesmosterol (**3**) therefore marks the last known intermediate in withanolide biosynthesis (Scheme 1.4).





Scheme 1.4 Known steps and branching points in phytosterol and withanolide biosynthesis.

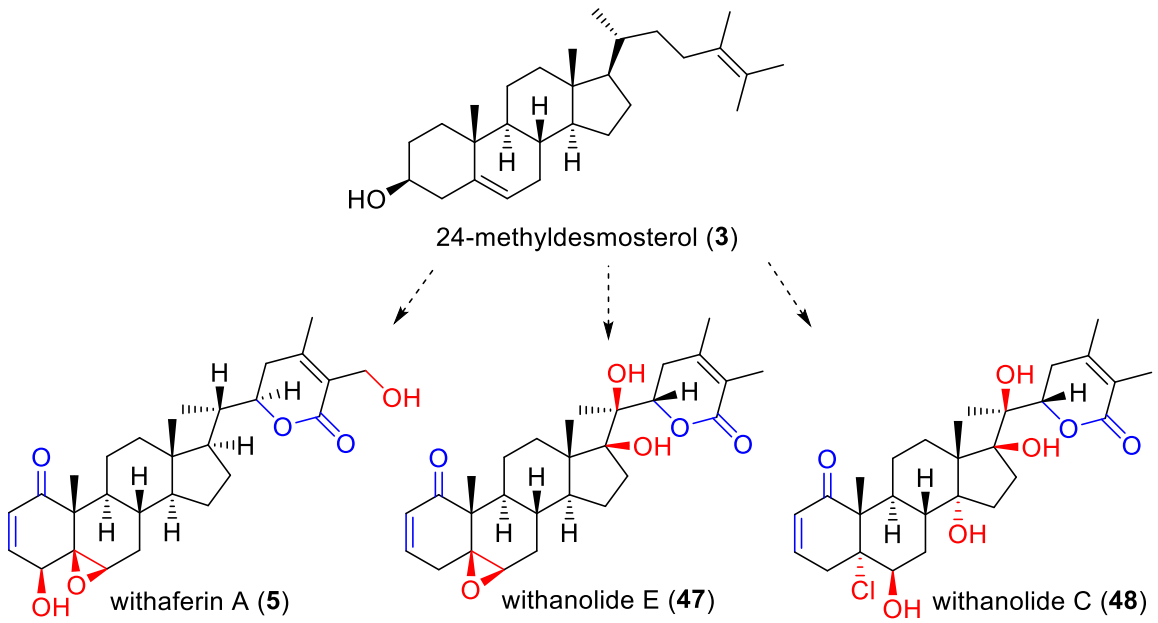
Further steps in withanolide biosynthesis have not been identified to date. Characteristic features of type A withanolides are multiple oxidations in the A and B ring and an oxidation of

the side chain to form a lactone ring. As their biosynthesis remains elusive, the following chapter will summarize isolated intermediates and form a hypothesis towards withanolide formation.

1.3.3 Biosynthetic hypothesis of withanolide biosynthesis

The formation of 24-methyldesmosterol (**3**) marks the last known step in withanolide biosynthesis. Afterwards, the exact order of reactions and enzymes involved are completely unknown. Due to their plethora of subclasses and different functionalizations, it is impossible to draw one general withanolide biosynthesis proposal. For simplicity, only type A withanolides with the example of withaferin A (**5**) will be discussed in the following section.

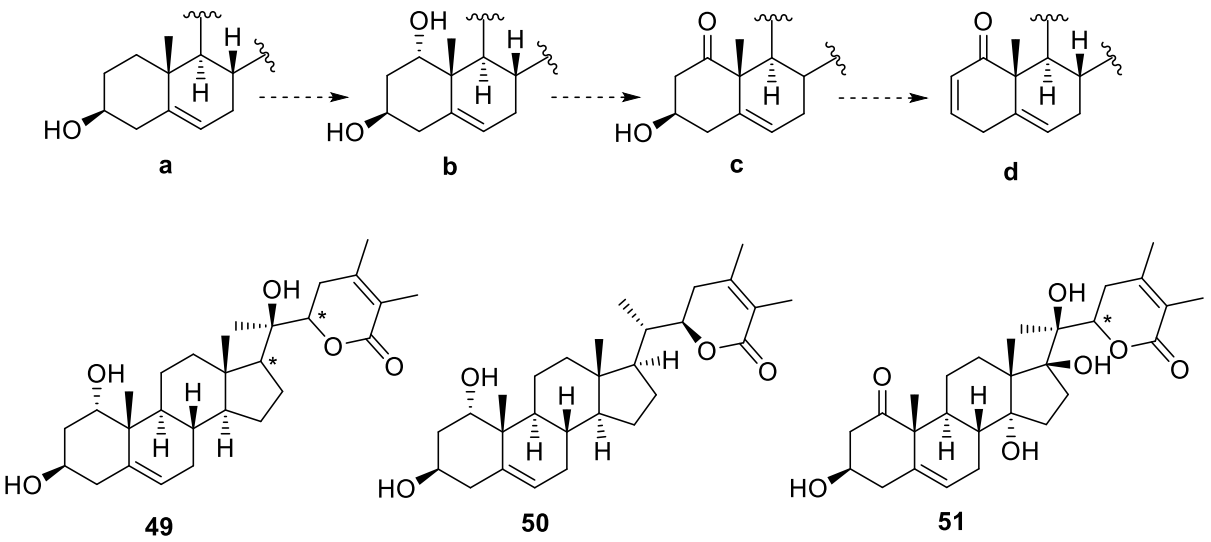
When comparing 24-methyldesmosterol (**3**) with withaferin A (**5**) and type A withanolides in general, modifications can be divided into two groups: general modifications such as the formation of the Michael system in the A ring and the lactonization of the side chain (blue, Scheme 1.5) and additional modifications such as epoxidations and hydroxylations, as demonstrated in comparison of withaferin A (**5**) to withanolide E (**47**, additional modifications in red, Scheme 1.5). Further modifications as for example halogenations are possible (e.g., withanolide C (**48**), Scheme 1.5) [10] but will not be discussed here in detail. Due to the large structural diversity it is possible that enzymes involved present a certain promiscuity to generate the multitude of compounds reported.



Scheme 1.5 Modifications in withanolide biosynthesis after formation of 24-methyldesmosterol (**3**) can be grouped into general (blue) and additional (red) functionalizations. Here shown exemplary for withaferin A (**5**), withanolide E (**47**) and withanolide C (**48**).

In the literature, the formation of the Michael system and the lactonization of the side chain have been discussed and potential intermediates have been isolated.

For the general modification of the A-ring, it was proposed that hydroxylation would occur at the C-1 position (b, Scheme 1.6) and would further be oxidized to the keto group (c, Scheme 1.6). Afterwards elimination of the 3-hydroxy moiety would be performed (d, Scheme 1.6). [66,67] Intermediates that show the hydroxy group at C-1 [66,68] have been isolated from *Withania somnifera* (**49**, Scheme 1.6)[66] and *Dunalia australis* (**50**, Scheme 1.6). [68] Furthermore an intermediate with the keto group at the C-1 position and a hydroxy-group at C-3 has been reported from *Withania coagulans*, however this structure might form spontaneously by addition of water to the double bond (**51**, Scheme 1.6). [67]

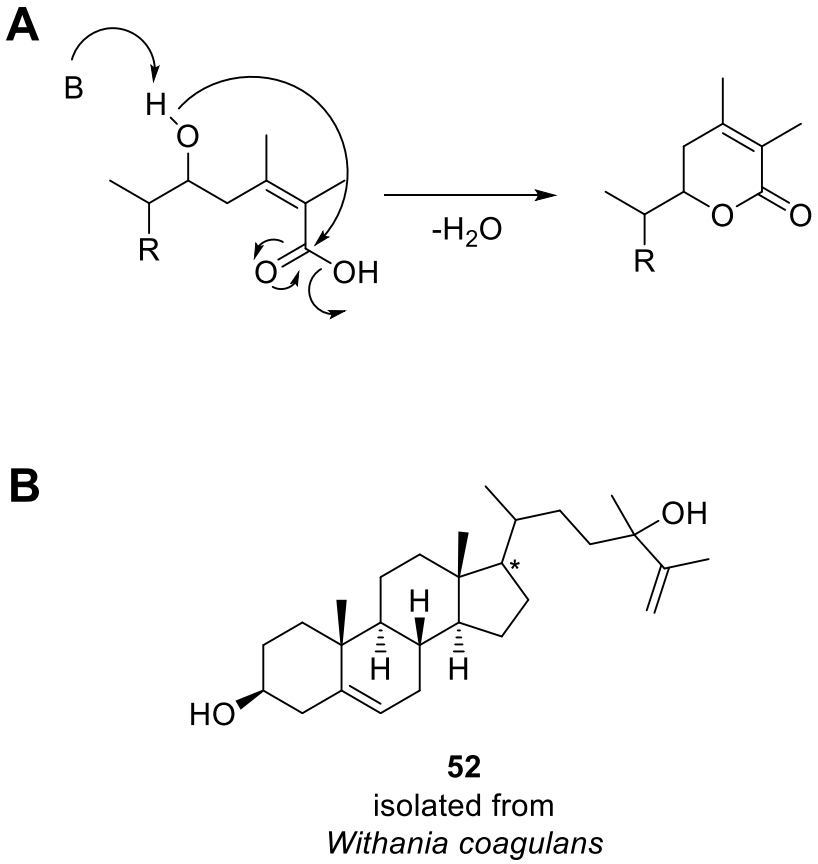


Scheme 1.6 Proposal of the Michael system formation in the withanolide A rings according to the literature with reported intermediates. [66–68]

While the reported intermediates in Scheme 1.6 show different additional modifications in the C-17 to C-20 region, a common feature is the lactone ring that is already present in those compounds. This suggests that lactone formation occurs at some point before the second oxidation step at C-1 (step c, Scheme 1.6).

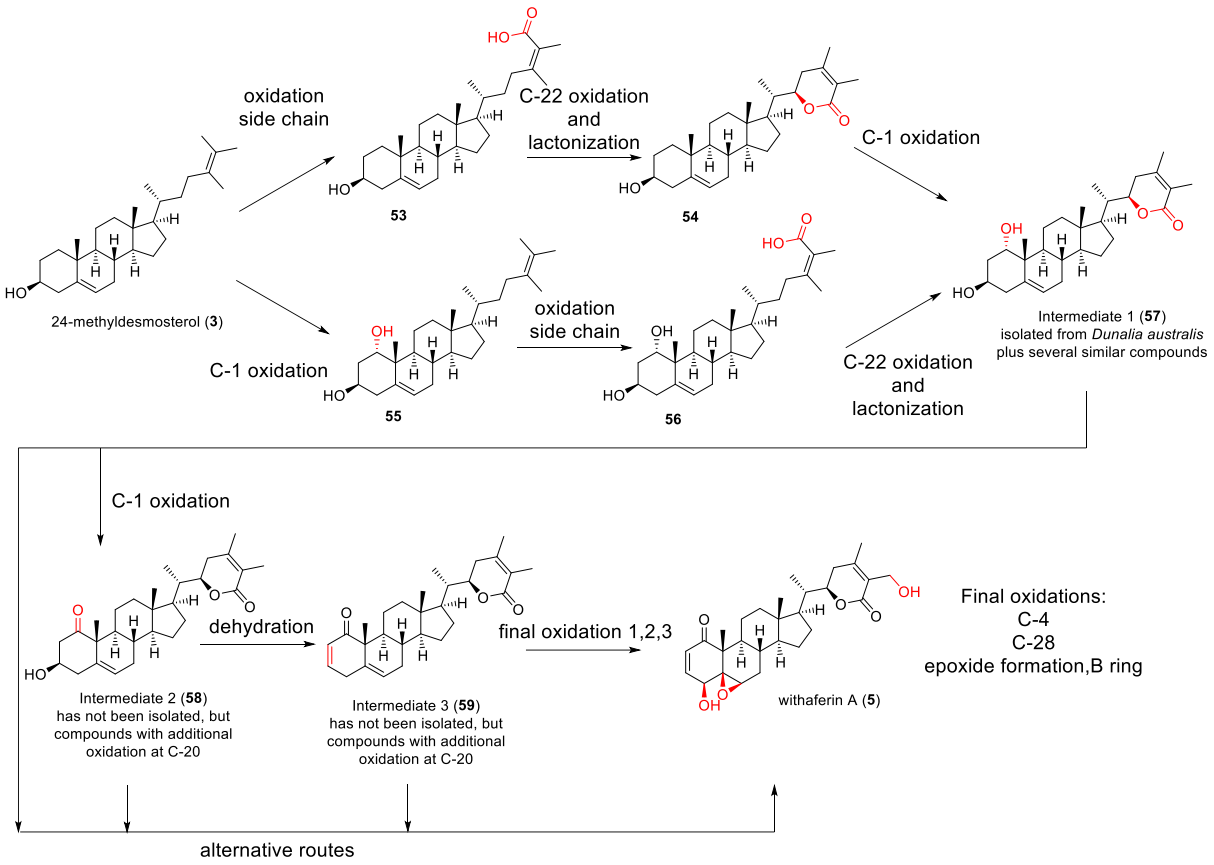
The most likely route for lactone formation would be the hydroxylation of the side chain at C-22 and oxidation to a carboxylic acid at C-26. Afterwards esterification could form the lactone ring (Scheme 1.7A). C-22 hydroxylation of sterol-type compounds is well reported, albeit not for withanolide precursors. In steroidal glycoalkaloid biosynthesis, a P450 from the CYP710A subfamily was reported to catalyze hydroxylation of cholesterol at the C-22 position. [69,70] However, only one compound related to withanolide metabolism with an acyclic oxidized side

chain, was isolated to date. This structure was isolated from *Withania coagulans* and shows hydroxylation at the C-24 position (**52**, Scheme 1.7B). [67]



Scheme 1.7 Proposed mechanism for lactonization of the side chain (A) and compound isolated by Velde *et al.* (B, stereochemistry depicted as far as reported). [67]

A proposal for withaferin A (**5**) biosynthesis, based on isolated intermediates, is depicted in Scheme 1.8. Besides the illustrated pathway, multiple alternative roads can be proposed depending on the time point of final oxidations 1-3. For clarity, alternative routes are not shown in Scheme 1.8, but a more comprehensive scheme can be found in the appendix (Scheme 10.1).



Scheme 1.8 Proposal for the biosynthesis of withaferin A (**5**) based on intermediates and similar structures isolated. [66–68]

During withanolide biosynthesis, intermediates can be branched to glycowithanolides. [71] Sterol glycosyltransferases (SGT) have been studied in the past, and three SGTs from *Withania somnifera* have been identified. [72] Furthermore, silencing experiments in *W. somnifera* confirmed their *in vivo* activity. [73]

In summary, withanolide biosynthesis shares many of the early steps with common phytosterol and triterpene biosynthesis. However, the main functionalizations that distinguish withanolides from phytosterols have not been characterized so far. The high level of oxidation of withanolides suggests oxygenases, most likely cytochrome P450 enzymes (P450), to be involved, as this enzyme class comprises the most versatile catalysts in nature. [74,75] Another class of oxidative enzymes with similar versatility are α-ketoglutarate dependent dioxygenases. [76] However, little evidence of activity of those enzymes in triterpene metabolism has been documented yet. [70] So far, no P450s have been assigned to withanolide biosynthesis either, but tissue specific expression analysis of several P450s in *W. somnifera* showed correlation of expression levels and withanolide contents. [77] Furthermore one P450 was

heterologously expressed in *E. coli* and was shown to hydroxylate withaferin A (**5**). [78] However, the exact processes of withanolide biosynthesis on molecular level remain uncertain. To further understand reactivity of P450 enzymes, a brief overview into this vast topic will be given in the following section.

1.3.4 Cytochrome P450 Monooxygenases

As cytochrome P450 monooxygenases (P450) are hypothesized to catalyze many of the central functionalizations in withanolide biosynthesis, this enzyme class should be further discussed. P450s comprise one of the most prominent classes of enzymes to catalyze a variety of oxidative reactions and are involved in many tailoring reactions of secondary metabolites. Discovered in the 1950-60s, and confirmed with O_2 labelling experiments, they challenged the existing opinion that oxygen in biological processes is derived exclusively from water. [79] P450s are found in all kingdoms of life, albeit not omnipresent.

The main functional roles for P450 enzymes are the metabolization of xenobiotics and the biosynthesis of important molecules such as hormones or secondary metabolites. [80] This is underlined by the fact that in plants, a kingdom well known for their plethora of secondary metabolites, CYP genes can comprise as much as 1% of the overall plant genes. [81] In general, P450s are often bound to the membrane of the endoplasmic reticulum or the mitochondrial membrane. [82]

P450s are divided into families with >40% amino acid identity, and subfamilies with >55% amino acid identity. [79,83] Their low sequence identity exacerbates prediction and engineering of P450 reactions. However, all P450s feature a signature motif of 10 amino acids for coordination of the heme iron. [79] The iron (III) protoporphyrin IX complex (Figure 1.5) in the active site is the center for reaction and is covalently linked to the protein by the sulfur atom of the cysteine ligand, which is conserved and essential for activity (Figure 1.5). [84] Furthermore substrate recognition sites (SRS) with high sequence variation around the active site shape the reactivity according to the substrates and products required. [80,85] Especially the high sequence variation of SRS makes reactivity of P450s hard to predict.

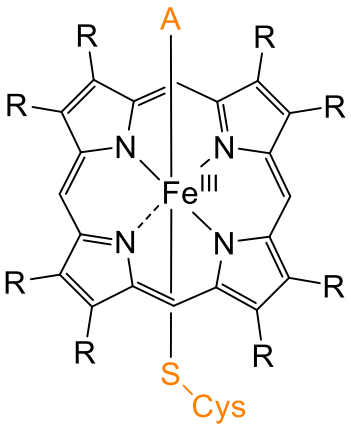


Figure 1.5 The iron (III) protoporphyrin IX complex of cytochrome P450 monooxygenases is covalently linked to a cysteine ligand in the proximal position, while the distal position (A) is responsible for substrate binding.

In a reaction cycle, one O-atom is incorporated into the substrate and the other one is reduced to a water molecule, which defines them as monooxygenases. [79] For this reaction, two electrons are used, provided from NADPH via a reductase. [86] P450s can be classified by their use of different reduction systems, with the most common ones being class I to III (Figure 1.6). [85] Class I P450 systems encompass most bacterial and mitochondrial systems, where electrons are transferred from a NAD(P)H molecule to a ferredoxin via a FAD-containing reductase. [86] The ferredoxin then reduces the P450 (Figure 1.6A). [86] Class II P450s are the most common ones in eukaryotes and can perform extremely diverse reactions. This system is localized in the membranes of the endoplasmic reticulum (ER) and contains two membrane proteins: the P450 itself and a NADPH-dependent cytochrome P450 reductase (CPR, Figure 1.6B), containing the prosthetic groups flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN). [87] Class III is similar to class I, where electrons are transferred from NAD(P)H via a FAD-containing ferredoxin reductase. However, the receiving molecule is a flavodoxin, containing FMN as a redox center instead of the iron sulfur cluster of the ferredoxin. This depicts a similar reduction system as the CPR, but divided onto separate proteins (Figure 1.6C). [88]

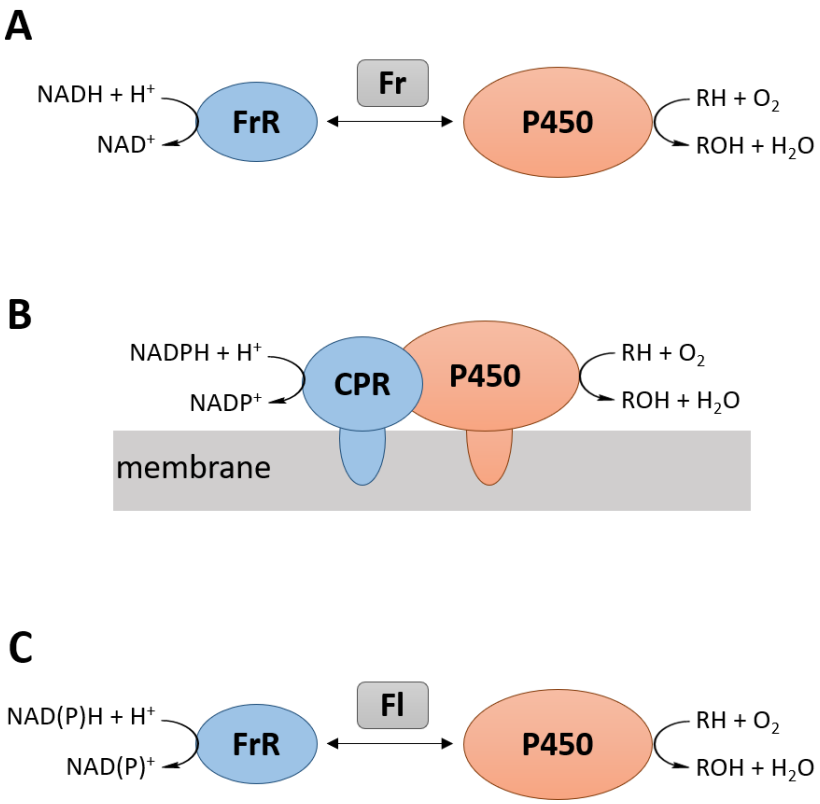
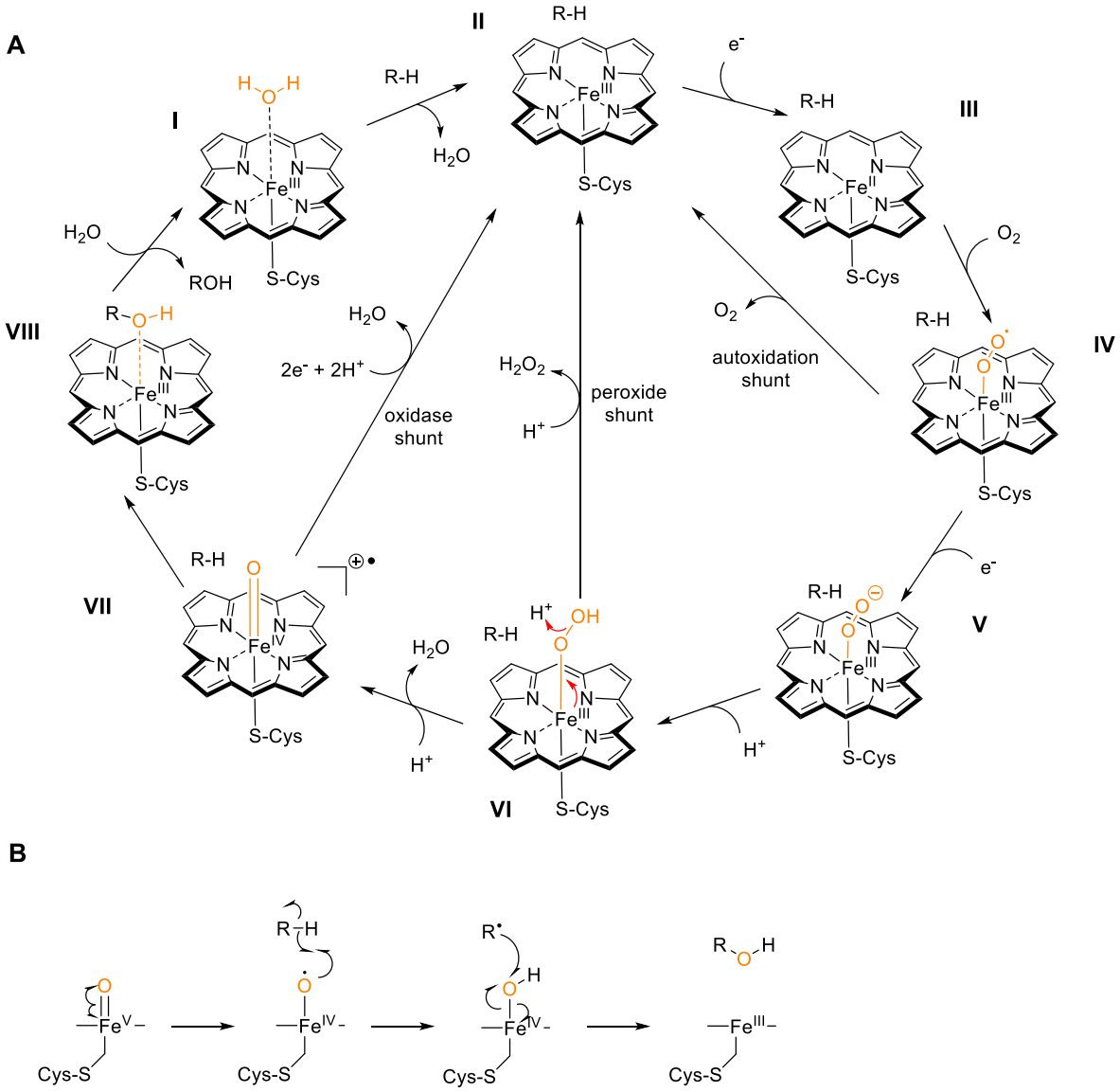


Figure 1.6 Common P450 reductions systems class I (A), class II (B) and class III (C); P450: cytochrome P450 monooxygenase; FrR: ferredoxin reductase, Fr: ferredoxin; CPR: cytochrome P450 reductase; Fl: flavodoxin (adapted from Hannemann *et al.*). [86]

Besides the reduction systems, the putative mechanism of oxidation seems to be identical in P450 enzymes. While intermediates for the initial steps in the reaction cycle were isolated and extensively studied (I to IV, Scheme 1.9A), [79,89–91] the later stages of the reaction cycle cannot be isolated and can only be targeted with few techniques due to their reactivity (V to VIII, Scheme 1.9A). [74,80] Depiction and description of the P450 reaction cycle varies slightly, depending on the focus of the respective publication. Here, the general consensus described in the literature should be described. [74,79,80,91]



Scheme 1.9 General consensus for the reaction cycle of cytochrome P450 oxidases (A) and presumed hydroxylation mechanism (B). [74,79,80]

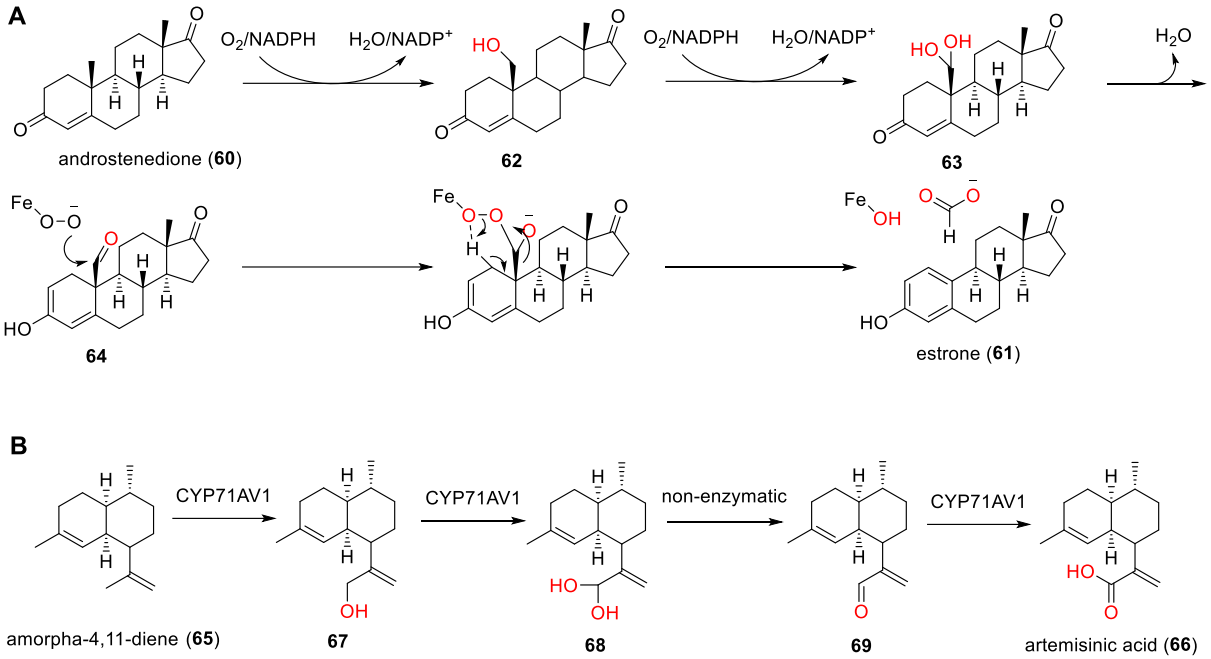
In the resting state, the iron is complexed with water as its distal ligand (Scheme 1.9A, I). [74] Addition of the substrate generates the pentacoordinate high spin ferric state with a vacant coordination site (Scheme 1.9A, II). [74] The first reduction step, performed by the respective reductase, generates a 5-coordinate high spin deoxyferrous state (Scheme 1.9A, III). [92] However, reduction kinetics vary drastically depending on the P450, the reductase and the respective environment, hence generalization of P450 kinetics and influential factors is not possible. [92]

In the following, oxygen can bind to form the ferric superoxide, the last stable intermediate of this reaction cycle (Scheme 1.9A, IV). [80] A second reduction, typically the rate limiting step of the mechanism, generates the ferric peroxide (Scheme 1.9A, V). [74,80] Protonation then yields the hydroperoxide complex (Scheme 1.9A, VI), which, after a second protonation releases

water to generate the proposed oxo-ferryl porphyrin radical (Scheme 1.9A, VII). [74,80] The proximal thiolate ligand serves as a strong internal electron donor to facilitate the O-O bond cleavage. [74] The oxo-ferryl radical is the most likely candidate to perform the hydroxylation, presumably by an oxygen rebound mechanism (Scheme 1.9B). [74] Reaction steps V-VIII have been targeted by many experimental methods, however, the high reactivity of this complex and the resulting low accumulation exacerbate measurements. [80] In addition to the general reaction cycle, P450 reactivity contains at least 3 branch points where side reactions can occur naturally. [93] The most commonly described shunt reactions include autoxidation of the oxy-ferrous state (Scheme 1.9A, IV), hydrogen peroxide dissociation from stage VI, or oxidation of the ferryl-oxo intermediate to water (VII), all resulting in return of the P450 to its resting state (Scheme 1.9A, II). [74,80,93]

Cytochrome P450s are known to not only catalyze simple hydroxylations, but also complex consecutive oxidation reactions. A good example is the aromatase (CYP19) in human sex hormone biosynthesis. It catalyzes the conversion of testosterone to estradiol, and the conversion of androstenedione (**60**) to estrone (**61**). [94] Here, two consecutive hydroxylations at the C-19 position (yielding **62** and **63**), followed by dehydration and spontaneous enolization at C-3, yield the 19-oxo androstenedione (**64**). [95] A third hydroxylation leads to elimination of formic acid, resulting in the aromatization of ring A (**61**, Scheme 1.10A). [96,97]

A popular example for a plant P450 catalyzing multiple oxidations is the CYP71AV1 in artemisinin (**13**) biosynthesis. It performs a three-step oxidation of amorpha-4,11-diene (**65**) to artemisinic acid (**66**) via **67**, **68** and **69** (Scheme 1.10B). This enzyme was used in a semi-synthetic approach to produce artemisinin (**13**). [98]



Scheme 1.10 Mechanism of the CYP19 aromatase (A) and mechanism of CYP71AV1 (B).

Further, cytochrome P450 enzymes have been reported to catalyze steps in phytosterol metabolism, [56] or related triterpene pathways such as steroidal glycoalkaloid [70] or brassinosteroid [99] biosynthesis. Additionally WSCYP93Id, a P450 from *Withania somnifera,* was able to hydroxylate withaferin A in *in vitro* experiments. [78]

Their ability to catalyze sequential oxidation reactions and their reported involvement in phytosterol related metabolism makes them the most likely candidates for functionalizations such as lactone formation in withanolide biosynthesis. Enzyme classes with similar versatility, e.g. α-ketoglutarate dependent dioxygenases, [76] have seldomly been reported in phytosterol or triterpene metabolism. [70]

Due to their numerous occurrences in plant genomes and localization in membranes, investigation of P450s is a difficult task. Further challenges and potential solutions for investigating withanolide biosynthesis will described in the following chapter.

1.4 Challenges and potential solutions for the investigation of withanolide biosynthesis

With their various medicinal properties and complex oxidation pattern, withanolides are appealing to research. However, chemical synthesis of withanolides or potential intermediates is complex due to many chiral centers and, depending on the withanolide, an epoxy ring. Furthermore, precursors for withanolide synthesis are either unavailable, expensive or require

many steps to receive a product. This makes a synthetic approach for the broader use of withanolides economically unappealing. Consequently, production by fermentation has come into focus, which requires the identification of biosynthetic steps involved. However, research on withanolide biosynthesis is aggravated because of many factors.

The investigation of plant biosynthesis in general faces a few obstacles, starting with the determination of relevant pathway genes. Sequencing of plant genomes is often a complicated, enduring task due to their size, compared to sequencing of microbial genomes. [100] Consequently, no withanolide producer has been characterized on genome level, so that analysis of gene candidates has to be conducted with transcriptome data. Furthermore, biosynthetic genes are rarely clustered and the identification of the genes involved often requires extensive testing. Due to their abundancy in the plant genome, the prioritization of P450 candidates for testing is especially complex.

As pathways in multicellular organisms are heavily structured, plant pathway genes are often expressed only in a certain tissue or at a certain developmental stage. For the investigation of plant metabolite biosynthesis, the producing tissue has to be identified first, and then the right time point in plant development in which the gene is expressed has to be caught. However, once certain pathway genes have been identified, co-expression analysis can provide insights on further genes involved.

Expression experiments of plant genes often require complex systems that are able to perform posttranslational modifications or possess a suiting folding machinery to produce either soluble proteins or membrane integrated proteins. [101] Specifically the expression of P450s is often complicated, as P450s are usually membrane attached, requiring a eukaryotic host. Furthermore, a compatible CPR has to be available, which often means a second transformation or a more complex vector system has to be employed.

The obvious choice for expression experiments is a plant host, as it usually provides the machinery necessary for expression of plant genes. A commonly used system is the model plant *Nicotiana benthamiana*, a close relative of tobacco. [102] However, plants are not able to take up DNA like unicellular hosts such as *E. coli* or yeast do. Methods for DNA transfer into plants include biolistic inoculation or protoplast transformations. [103,104] The most popular tool for DNA delivery, however, is the indirect transfer via agroinoculation. As this technique is based on a complex system, *Agrobacterium* mediated gene transfer should be discussed in the following section.

1.4.1 *Agrobacterium tumefaciens* as a tool for DNA transfer

Agroinoculation is one of the most commonly used techniques for DNA transfer into plants, as it is fast and versatile. This technique employs the natural pathogenic properties of the *Agrobacterium* species. In nature Agrobacteria are plant pathogens that cause tumor-like growth in wounded plant tissues to exploit plant metabolism for the production of opines, such as agrocinopine A. [105] *Agrobacterium tumefaciens* induces galls on roots and crowns, while an infection with *Agrobacterim rhizogenes* leads to abnormal root production. Those tumors are in general non-lethal, but affect general plant growth negatively. [106]

The virulent properties of *Agrobacterium* species are derived from a plasmid, in the case of *Agrobacterium tumefacies* this is the tumor-inducing (Ti) plasmid. [106] *Agrobacterium rhizogenes* carries the root-inducing (Ri) plasmid. [106] As plant infections are typically carried out with *A. tumefaciens*, the Ti plasmid will be described in further detail (see Figure 1.7).

The tumor-like growth of plant cells is induced by the expression of several oncogenes which are encoded on the transfer DNA (T-DNA) that is transported into the plant cell and integrated into the plant genome. It carries three oncogenes that establish a new biosynthetic pathway for the production of auxin, cytokinin and opines. [107] This T-DNA is a section on the Ti plasmid, flanked by conserved 25 bp nuclease signaling sequences (left and right border, LB/RB), that can be excised to generate active T-DNA. Furthermore a virulence region is encoded on the Ti plasmid that harbors the activators for this virulence region, the nuclease for T-DNA excision and a machinery for the transfer of the T-DNA (vir and CT). [106,108]

Agrobacterium tumefaciens is able to specifically degrade the opines in tumors they have incited. For this purpose not only synthesis genes, but also genes for the catabolism are encoded on the Ti plasmid (opine catabolism, OC). [109,110] This redirection of plant metabolism to a form that bacteria can utilize gives the bacteria a competitive advantage against other pathogens. By modifying the T-DNA, scientists can transfect DNA of interest into a plant host.

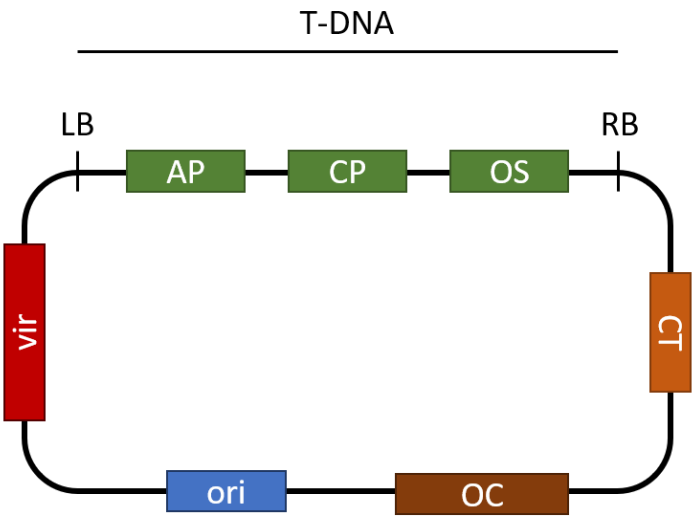


Figure 1.7 Schematic depiction of the native Ti-plasmid. The transfer DNA (T-DNA) region is flanked by the left and right border (LB and RB) and includes oncogenes responsible for the production of auxin (AP) and cytokinin (CP) and genes for opine synthesis (OS). Furthermore, the Ti plasmid includes a virulence region (vir), a region for establishing the conjugative transfer (CT), genes for opine catabolism (OC) and an origin of replication. Figure adapted from Zerbini *et al.* [111]

As the process of plant infection is essential for the understanding of this technique, the different steps will be described in further detail (Figure 1.8). Wounded plant tissue secrets acetosyringone and other phenolic compounds that can be sensed by Agrobacteria, which can attach to the plant tissue with cellulose fibrils. [112] Acetosyringone and other substances further stimulate the upregulation of the virulence region, and a type 4 secretion system (T4SS) is formed. [113–116] The T-DNA is excised from the Ti plasmid by the nuclease encoded in the virulence region and furthermore transferred as a single strand, coated by single strand binding proteins, through the canal of the pilus. [117–119] The integration occurs preferably in regions with high transcriptional activity. [120]

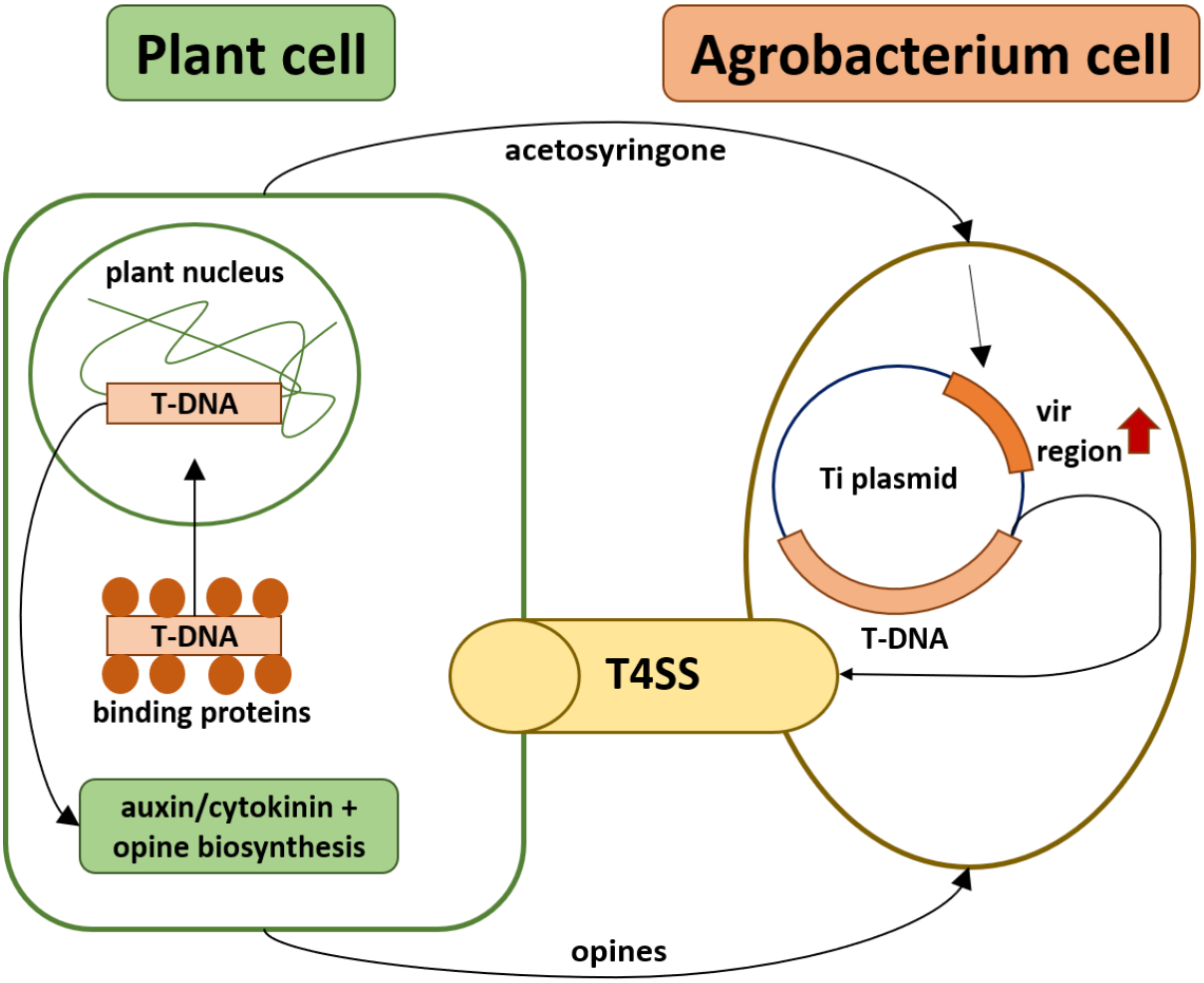


Figure 1.8 Schematic depiction of the agroinfection process (modified after Hwang *et al.* [106]). The transfer T-DNA from the Ti plasmid is transferred via a type 4 secretion system (T4SS) from the *Agrobacterium* cell to the plant cell. In the plant the T-DNA will be integrated into the genome in the plant nucleus, resulting in an upregulation of auxin/cytokinin biosynthesis and production of opines, which are beneficial for the Agrobacteria.

Agrotransformation allows easy transfer of genes of interest, for improved fitness of the plant by genomic integration or to study the function of a certain protein. For the latter, transient expression in model organisms such as the commonly used *Nicotiana benthamiana*, an Australian relative of tobacco, [121] is preferred. In general, transient expression can be achieved via plasmids based on modified Ti vectors. In this case it is essential to avoid plant innate immune responses, e.g. posttranscriptional gene silencing (PTGS), to generate high titers of protein. To prevent PTGS, suppressors are usually co-expressed either in a binary vector system or on the same plasmid. [122] A vector that is commonly used is the pEAQ-HT system, [123] that includes the p19 protein from the tomato bushy stunt virus (TBSV), preventing siRNA association with the RNA-induced silencing complex (RISC). [124]

Besides expression of candidate genes, another technique employed in biosynthetic research is silencing of genes of interest. In plant biosynthetic research transient techniques such as virus

induced gene silencing (VIGS) are preferred over stable genetic knockouts, as VIGS is easier to establish in non-model plant species. [125] Withanolide producing plants are not common model organisms, therefore this technique can provide further insights and should be discussed in the following section.

1.4.2 Virus induced gene silencing – investigating plant biosynthesis by knockdown

For studying biosynthetic pathways, two general approaches can be considered: experiments that try to reconstruct one or more steps of a biosynthetic pathway with the help of a heterologous host, or *in vivo* experiments that target the biosynthetic mechanism in the producing organism e.g., by downregulation of genes of interest.
The downregulation of certain steps in a pathway can be employed as a strategy to confirm a hypothesis, as knockdown of the gene of interest will result in metabolic changes, unless targeted genes are redundant. Furthermore, the analysis of accumulated intermediates can provide insights in the sequence of a biosynthetic route (Figure 1.9).

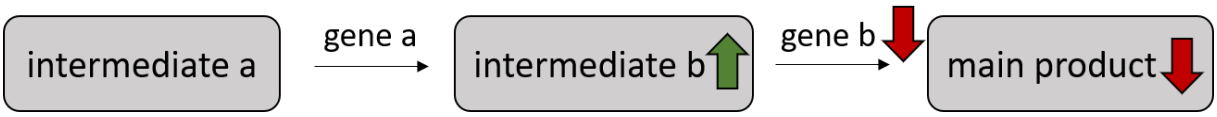


Figure 1.9 Effect of knockdown experiments on metabolites in a biosynthetic pathway.

In plants, knockouts can be achieved by the generation of transgenic lines. However, this process is challenging and time consuming. A technique known for its fast and transient procedure is virus induced gene silencing (VIGS). [126] Here, a recombinant virus is employed to trigger the innate immune response of the plant to eliminate viral RNA. If viral RNA has been exchanged with sequences complementary to the gene of interest, silencing of this gene can be observed instead. This immune response, the post transcriptional gene silencing (PTGS) [127,128] is essential for understanding the VIGS technique and will therefore be described in more detail. In the cell, viral RNA is replicated by an endogenous RNA-dependent RNA polymerase. Those double stranded dsRNA molecules can be recognized by DICER-like proteins that cleave the dsRNA into 21-30 nt long fragments, termed short interfering RNAs (siRNAs). [129–131] The resulting siRNAs can be recognized and bound by Argonaute proteins, [132,133] forming the RNA induced silencing complex (RISC). This complex can further use single stranded siRNAs to target complementary RNAs in the plant cell, which are degraded in this process. [129,134–136] The siRNA sequences will be amplified and transported, leading to a systemic gene silencing in all parts of the plant. [137] As a consequence, translation of viral proteins is prevented, and the plant is protected. [129,134] The general mechanism is summarized in Figure 1.10.

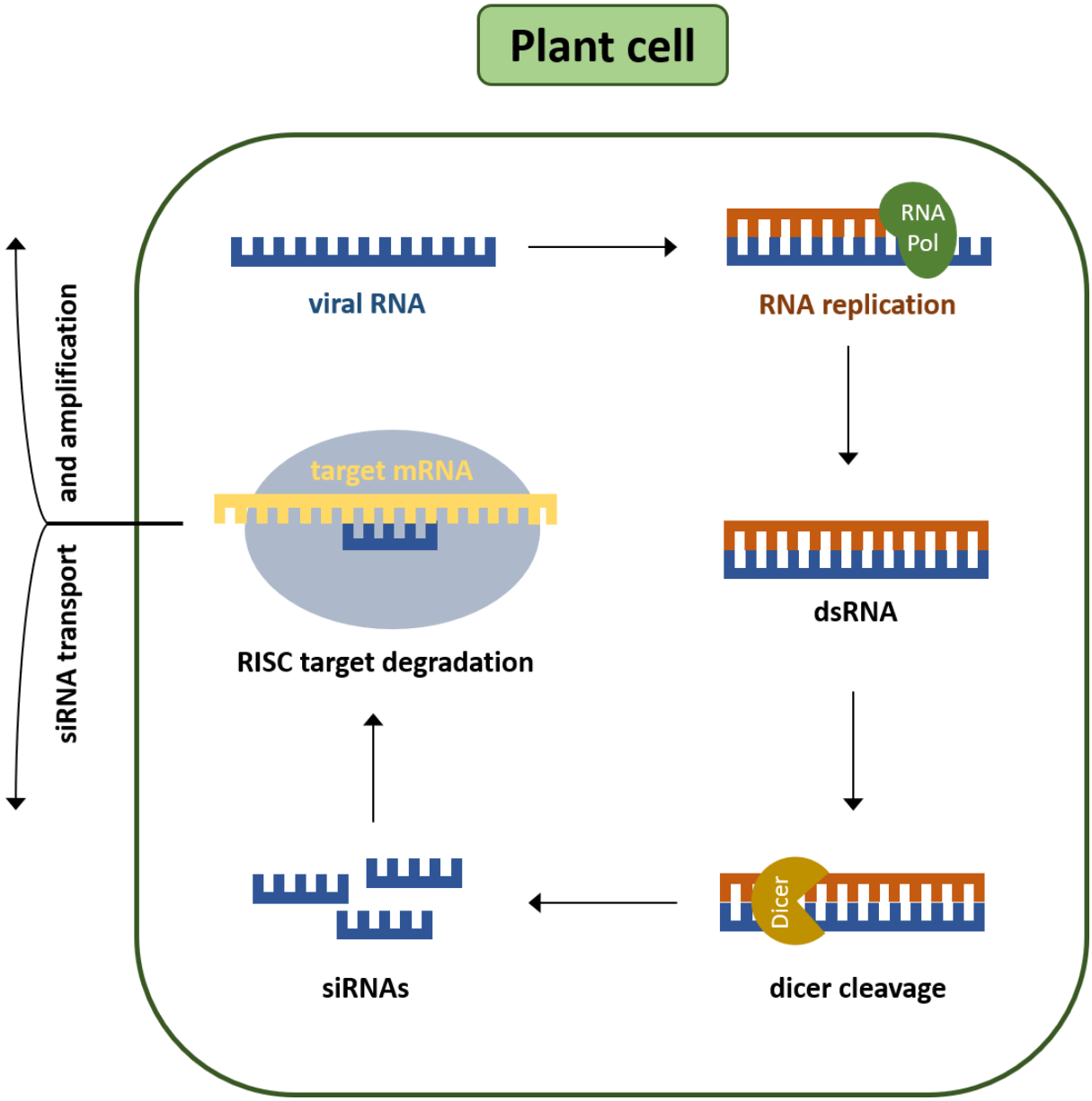


Figure 1.10 General mechanism of VIGS (adapted from Bekele *et al.*) [138].

Several plant viruses have been adapted as VIGS vector systems. Those can be on the basis of both RNA or DNA plant viruses, but many are based on positive strand RNA viruses such as potato virus X (PVX), tobacco mosaic virus (TMV) or tobacco rattle virus (TRV). [126] Single strand DNA viruses are employed as well, but may require a helper virus or satellite DNA for symptom induction. [103,139]

Furthermore, RNA virus-derived VIGS systems with satellite and helper RNAs have been developed which uncouple viral and silencing components. Those systems may result in stronger phenotypes and were called satellite virus induced silencing system (SVISS). [140]

The most popular system to date is the binary pTRV vector system, where the pTRV1 vector contains the virulence genes, while the silencing fragment is provided on the second vector, pTRV2. [141,142]

Several Solanaceae species such as *Capsicum annum*,[143] *Solanum lycopersicum*[141] and *Solanum tuberosum* [144] have been targeted by VIGS, making this the most widely used family for those silencing experiments. The first plant in which VIGS was reported is the well-studied model organism *Nicotiana benthamiana*.[145]

Furthermore, studies have been published in withanolide producers *Withania somnifera* [64,73,146,147] *Physalis floridana* [148] and *Physalis peruviana*, [149] showing that VIGS and consequently investigation of withanolide biosynthesis is possible in those organisms.

The techniques described in this chapter (Agroinfiltration, VIGS) provide a toolbox for investigating withanolide biosynthesis. As withanolide biosynthesis still has many unanswered questions, those techniques should be employed for gaining further insights into withanolide biosynthesis and related pathways in this work. The following section will discuss the aims set for this work.

1.5 Aim of the thesis

Withanolides are plant natural products with various medicinal properties beneficial to humankind. [6,8,9] As yields in plants are low, it is of scientific interest to understand the withanolide biosynthetic pathway to enable heterologous production and possibly manipulation towards derivatives with enhanced properties. Withanolides possess several modifications compared to the precursors, which have not yet been investigated due to the challenging chemistry and enzymology involved. Common features of type A withanolides are the Michael system in the A ring that requires oxidation and elimination, and a lactone ring in the side chain that requires multiple oxidations compared to the precursor. The aim of this thesis is to provide further insights into oxidation and elimination reactions involved in creating those features.

When investigating withanolide biosynthesis, multiple difficulties have to be faced. As, a plethora of withanolide type structures have been identified from multiple plant species, depending on the species, chemotype or cultivation environment, withanolide producers for research should be chosen and evaluated for their main withanolides observed in this environment.

Furthermore, potential intermediates of withanolide biosynthesis are difficult to obtain, as they are not commercially available. However, those compounds are needed as standards or substrates in enzyme assays. Here one approach can be metabolic engineering. This work aims to hijack yeast ergosterol metabolism to channel it towards production of 24-methyldesmosterol (**3**). Furthermore, such a strain could be transformed with candidate genes for *in vivo* analysis.

Withanolides contain many oxidations that are thought to be performed by cytochrome P450 enzymes. Furthermore, the Michael system in the A ring requires elimination of the 3-hydroxy group observed in 24-methyldesmosterol (**3**), potentially catalyzed by a dehydratase. Consequently, suitable P450 and dehydratase candidates should be selected as a base for further investigation of withanolide biosynthesis.

While bioinformatic analysis can provide a rough selection of gene candidates, experimental data is needed to narrow down the field and connect candidates with the withanolide biosynthesis pathway. An approach that enables screening of larger batches of candidates is gene-knockdown by virus induced gene silencing. VIGS for withanolide producers *Withania somnifera* [64,73,146,147] and *Physalis peruviana* [149] has been reported in the literature with varying extent. Therefore, a VIGS system should be established for those plants in this institute and

potential withanolide biosynthesis candidates should be silenced to see whether an effect on withanolide accumulation can be observed.
Besides knockdown experiments, heterologous expression is an option for investigating biosynthesis. Due to the specific requirements of cytochrome P450 enzymes, a eukaryotic host is needed. Here, two systems for expression should be explored. First, expression in yeast with subsequent microsome isolation and enzyme assays with external substrate feeding should be tested. Furthermore, expression of P450 candidates should be conducted in the common plant host *Nicotiana benthamiana*.

In summary, this work aims to provide novel insights into withanolide biosynthesis to pave the way for heterologous production and rational engineering.

2 Identification of main withanolides from *Withania somnifera* and *Physalis peruviana*

2.1 Introduction

Withanolides are plant natural products with a complex oxidation pattern present in numerous genera of the Solanaceae family, for example *Withania*, *Datura*, *Dunalis*, *Iochroma*, *Jaborosa*, *Lycium* and *Physalis*.[14] Their diverse modifications and multitude of producing plants that often contain different subspecies and chemotypes, all with different withanolides, result in more than 300 known representatives of this compound group.[14] For those, various different purification protocols have been published. [150–154]

For this work, *Withania somnifera* and *Physalis peruviana* were selected for further investigation, as they are both used as traditional remedies, are not toxic, and transcriptome data for identification of gene candidates is available. [155–157] As silencing experiments are planned in this work, an overview of main withanolides of both plants is needed, to track effects of silencing of withanolide biosynthesis gene candidates.

Withania somnifera Dunal is the most prominent withanolide producer and a plethora of withanolide-type compounds have been isolated in different chemotypes grown in various countries. Some common examples include withaferin A (**5**), the first withanolide isolated, [158,7,159,160] withanolide A (**9**), [161] withanolide D (**70**), [162] withanolide E (**47**), [163] withanolide S (**71**) [164] and withanone (**72**) [165] (see Figure 2.1).

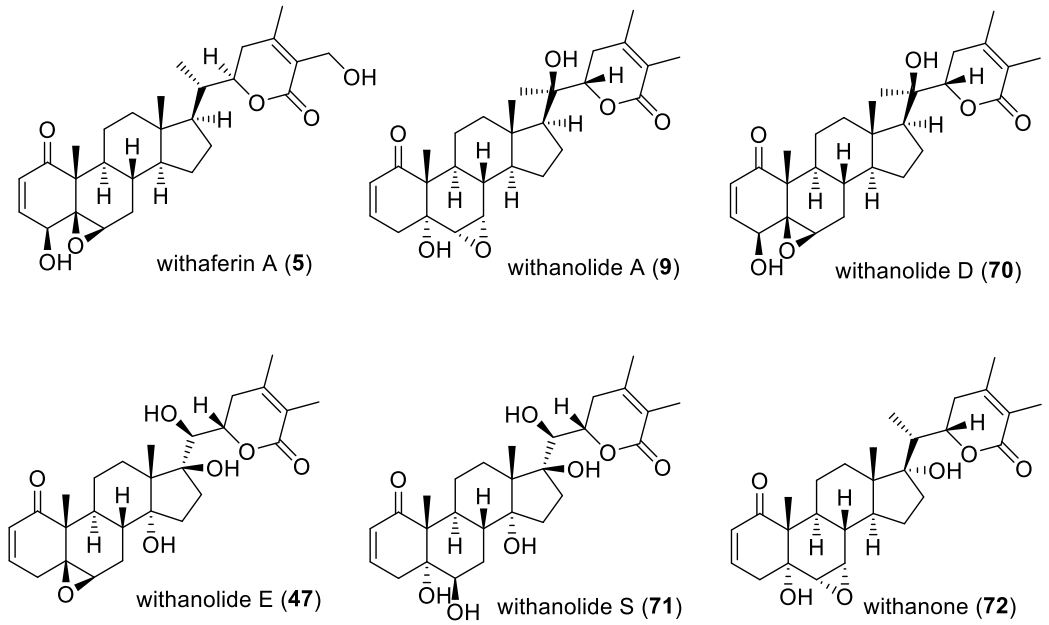


Figure 2.1 Examples for withanolides isolated from *Withania somnifera* Dunal include withaferin A (**5**), [2-5] withanolide A (**9**), [161] withanolide D (**70**), [162] withanolide E (**47**), [163] withanolide S (**71**) [164] and withanone (**72**). [165]

The other target for investigation is *Physalis peruviana*, a plant commonly cultivated for its edible berries. The first withanolide reported from this plant is 4β-hydroxywithanolide E **(73)**. [166,167] Structural diversity of withanolides in *Physalis* species is vast, and subclasses reported in *P. peruviana* grown in different countries include common withanolides (e.g. withanolide C, **48**), [10] perulactones (perulactone H, **74**), [168] withaperuvins (withaperuvin C, **75**),[169] phyperunolides (phyperunolide D, **76**) [170] and physalolactones (physalolactone C, **77**, see Figure 2.2). [171]

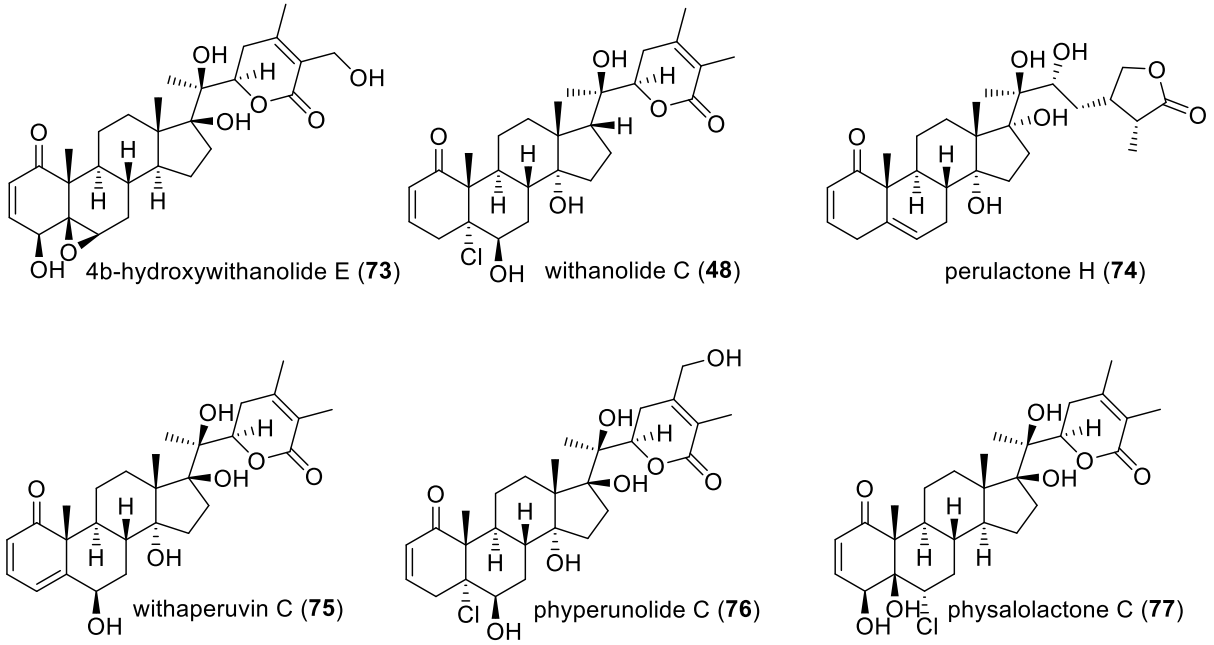


Figure 2.2 Examples for withanolides isolated from *Physalis peruviana* include 4β-hydroxywithanolide E (**73**), [167] withanolide C (**48**), [10] perulactone C (**74**), [168] withaperuvin C (**75**),[169] phyperunolide D (**76**) [170] and physalolactone C (**77**). [171]

For this project, a general purification strategy needed to be established. Once a protocol suitable for purification of main withanolides was found, *W. somnifera* and *P. peruviana* plants were extracted and analyzed for their main withanolides. Here, commonly reported withanolides were identified. Furthermore, a new subclass of truncated withanolides (irinan A (**1**) and B (**2**)) was discovered in *P. peruviana* and results were further published in the Beilstein Journal of Organic Chemistry. [172]

2.2 Establishment of an isolation strategy for withanolides

At first, a purification strategy for main withanolides should be established. Literature suggested various extractions protocols, depending on the subclass of withanolides isolated. In general type A withanolides showed good solubility in chloroform, while glycowithanolides are soluble in butanol. [152,154,173] Therefore, purchased *Withania somnifera* root powder was extracted following common themes in the literature by extraction with methanol and further sub-fractionization for defatting, withanolide and glycowithanolide extraction.

The purification was tested on analytical scale (10 g powder, method see chapter 1.1, purification scheme see Figure 2.3). Root powder was extracted with H_2O/MeOH (3:1) and was further divided into petroleum ether, chloroform and butanol fractions (see chapter 8.2.4). The chloroform crude fraction was further purified via automated flash chromatography (SNAP KP-Sil 10 g, chloroform/methanol 1% - 10%). The resulting fractions were combined in two major subfractions (F1: fractions 3-6, 6 mg; F2: fractions 7-10, 26 mg, see Figure 2.3 and Figure 10.3) and were examined via LCMS and NMR. F1 and F2 contained NMR signals typical for withanolides e.g., signals at 6-7 ppm corresponding to the α,β unsaturated A-ring. [174] Comparison to reference NMR spectra identified F1 and F2 as withanolide A (**9**) [19,161] and withaferin A (**5**) [175,176,11] respectively (see appendix Figure 10.1 and Figure 10.2). However, both samples still contained impurities.

The *n*-butanol fraction was examined via LCMS and seemed to contain glycowithanolides due to the following reasons: Compounds eluted at earlier retention times and were consequently more polar. Furthermore, observed masses showed a difference matching added glucose molecules (+162 per glucose molecule). However, as glycowithanolides are not the focus of this work, no further analysis was conducted.

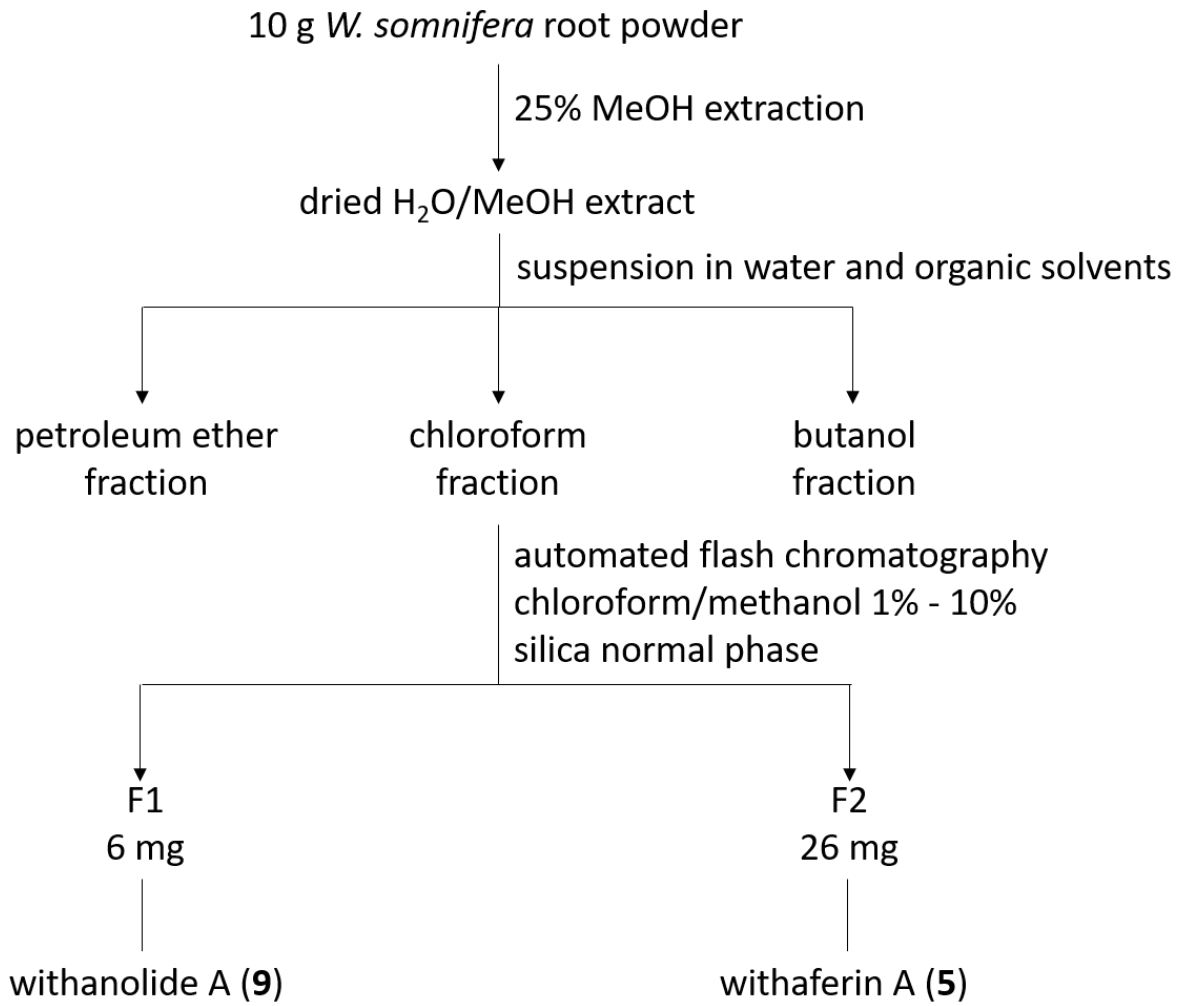


Figure 2.3 Schematic depiction of the isolation process for withanolide A (**9**) and withaferin A (**5**) from *W. somnifera* root powder.

It was concluded that the purification method used was able to extract the main withanolide, withaferin A (**5**), in *W. somnifera* root powder. This matches the expectations, as withaferin A (**5**) is typically the main withanolide observed in *Withania somnifera*.[158,159,7,160] In order to know which metabolites to target in the analysis of silencing samples (see chapter 5), self-cultivated *Withania somnifera* plants were investigated for their main withanolides by Samuel E. Hakim.

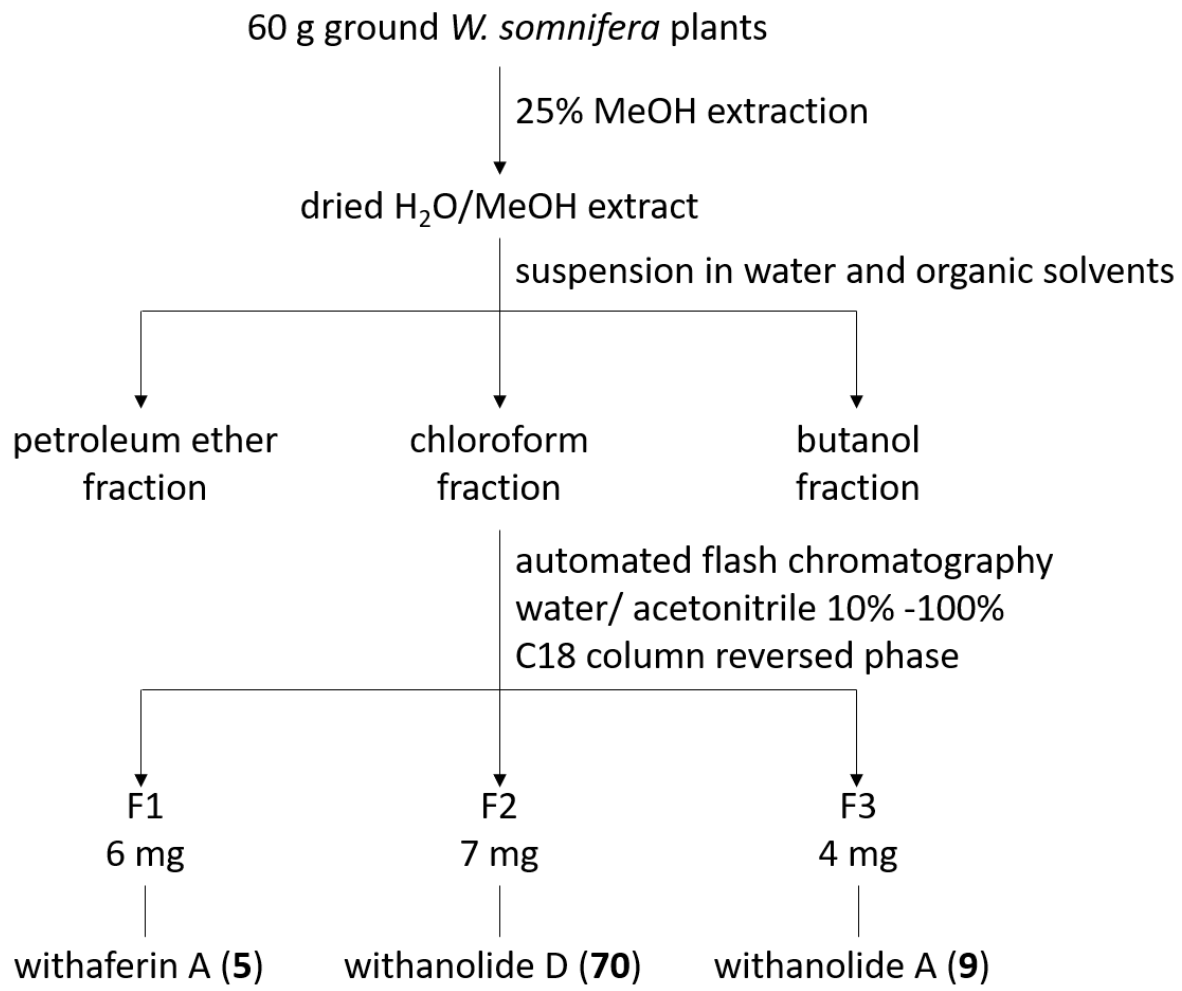


Figure 2.4 Schematic depiction of the isolation process for withaferin A (**5**), withanolide D (**70**) and withanolide A (**9**) from *W. somnifera* performed by Samuel E. Hakim.

Here, a chloroform crude extract was purified using automated flash chromatography (SNAP C_{18}, water/acetonitrile 10% - 100%) and yielded three subfractions (F1-F3, see Figure 2.4). Analysis via LCMS and NMR with subsequent comparison to literature data showed known withanolides withaferin A (**5**), [11,175,176] withanolide D (**78**) [177] and withanolide A (**9**) [19] respectively (see Figure 2.5).

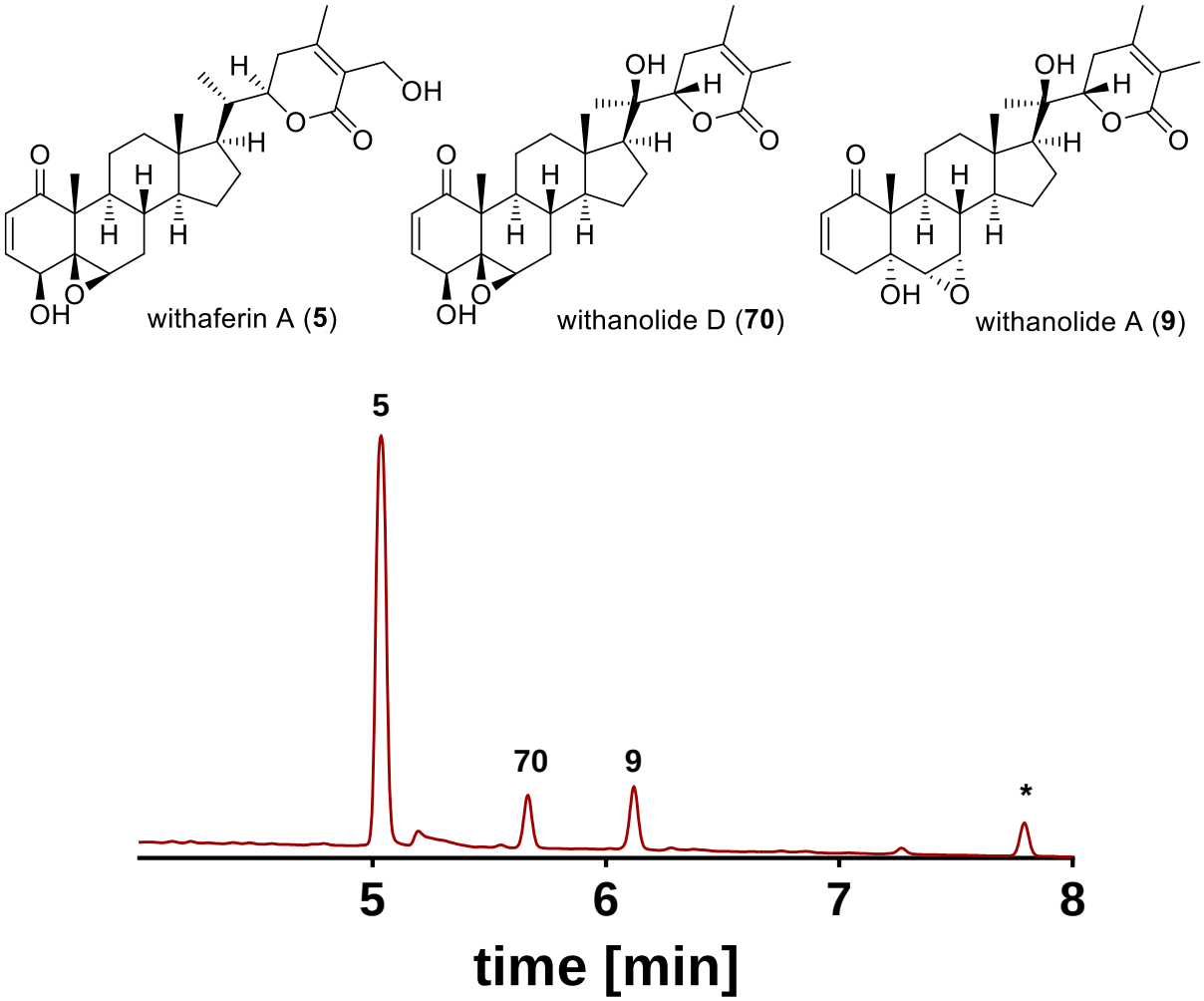


Figure 2.5 Exemplary LCMS chromatogram (UV trace at 220 nm) of *Withania somnifera* MeOH crude extract with main withanolides (**5**,**70** and **9**) identified by Samuel E. Hakim (*: unidentified compound).

2.3 Identification of key withanolides and new compounds from *Physalis peruviana*

Initial silencing experiments were planned with *Physalis peruviana* and therefore an overview of withanolides observed in plants cultivated in this work was needed. The results obtained in this chapter were part of a publication in Beilstein Journal of Organic Chemistry. [172]

Five 9-week old *Physalis peruviana* plants (137 g) were ground in liquid nitrogen and extracted with H_2O:MeOH (3:1). The crude extract was further and divided into fractions soluble in petroleum ether (yield 660 mg), chloroform (386 mg), and *n*-butanol (1.17 g), according to the protocol developed in the previous section (see Figure 2.6). $CHCl_3$ and *n*-BuOH fractions were analyzed via LCMS and contained withanolides.

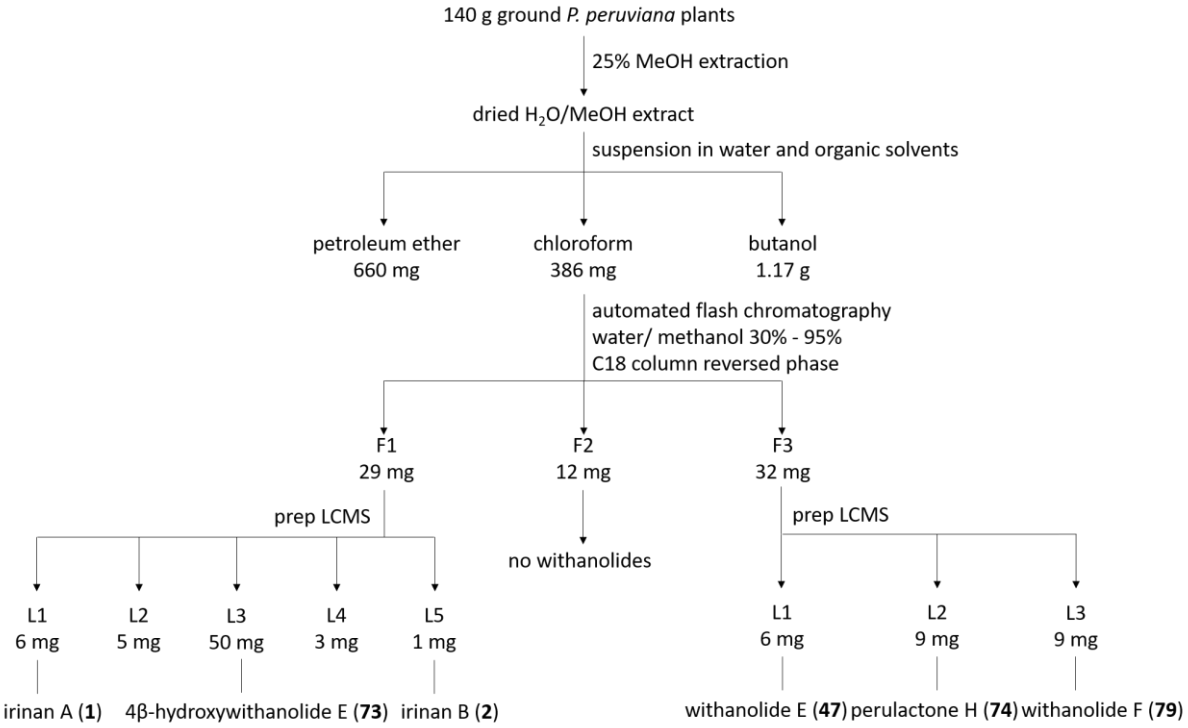


Figure 2.6 Schematic depiction of the isolation process for compounds reported from *P. peruviana*. [172]

The chloroform crude fraction was pre-purified via automated flash chromatography and the resulting fractions were combined in three major subfractions (F1: fractions 1-21, 29 mg; F2: 22-31, 12 mg; F3: 32-45, 32 mg see Figure 10.3, in the appendix).

NMR analysis showed that F2 did not contain withanolides, but F1 and F3 showed some signals typical for withanolides. [174] F1 was purified via preparative LCMS and 5 fractions (L1-L5, Figure 2.6) were collected that contained masses in the expected range of *m/z* 350-500.

The LCMS-prep fractions were analyzed via NMR, with fractions L2 (5 mg) and L4 (3 mg) showing only small amounts of withanolide-type compounds and many impurities. The remaining LCMS fractions, L1 (6 mg), L3 (50 mg) and L5 (1 mg) were further analyzed and the one with the highest yield (L3) was quickly identified as 4β-hydroxywithanolide E (**73**) via ^{1}H-NMR by comparison to reference spectra (see Table 2.1 and Figure 10.4 in the appendix). [166,174]

Withanolides often exhibit signals at 6-7 ppm (**73** at 6.92 and 6.22 ppm, see Table 2.1),[i] that can be attributed to the protons of the α,β unsaturated western fragment. [174] Furthermore, doublet of doublet signals at 4-5ppm (**73** at 4.88 ppm) are characteristic for the eastern lactone ring. [174] Furthermore withanolides often exhibit methyl groups with downfield shifts at C-27 and C-28 (**73** at 1.88 and 1.94 ppm), if no hydroxy group at C-27 is present. [174]

[i] For carbon numbering see Figure 2.7

Fractions L1 and L5 were expected to be withanolides as well, since the characteristic signals for the α,β-unsaturated western fragment were present. For 4β-hydroxywithanolide E (**73**) coupling of proton signals at C-3 (doublet of doublets) indicates that C-4 only possesses one proton. [174]

L1 showed the same coupling as 4β-hydroxywithanolide E (**73**) for protons at C-3, while L5 showed a ddd, indicating that two protons are present at C-4 (see Table 2.1). The signals for the deshielded methyl groups at C-27 and C-28 were missing for L1 and L5. Furthermore, both fractions seemed to be missing the signals of the lactone ring that in type A withanolides are typically observed at 4-5 ppm for C-22. [174] This led to the assumption that the eastern lactone ring was either missing or was structured differently compared to other compounds isolated in this work.

For further structure elucidation ^{13}C, HSQC, HMBC und COSY measurements were analyzed by Dave Biedermann. The ^{13}C NMR spectrum confirmed the assumption of a truncated withanolide by showing only 19 signals. Furthermore, the ^{13}C NMR revealed another oxo-group group besides C-1, as two signals were present >200 ppm. NMR data suggested that the second carbonyl group was located at C-17, where usually the side chain with the lactone ring is located. The presumed structures were confirmed with HSQC, HMBC and COSY measurements by Dave Biedermann and the compounds present in L1 and L5 were termed irinan A (**1**) and B (**2**) respectively (see Figure 2.7).[ii]

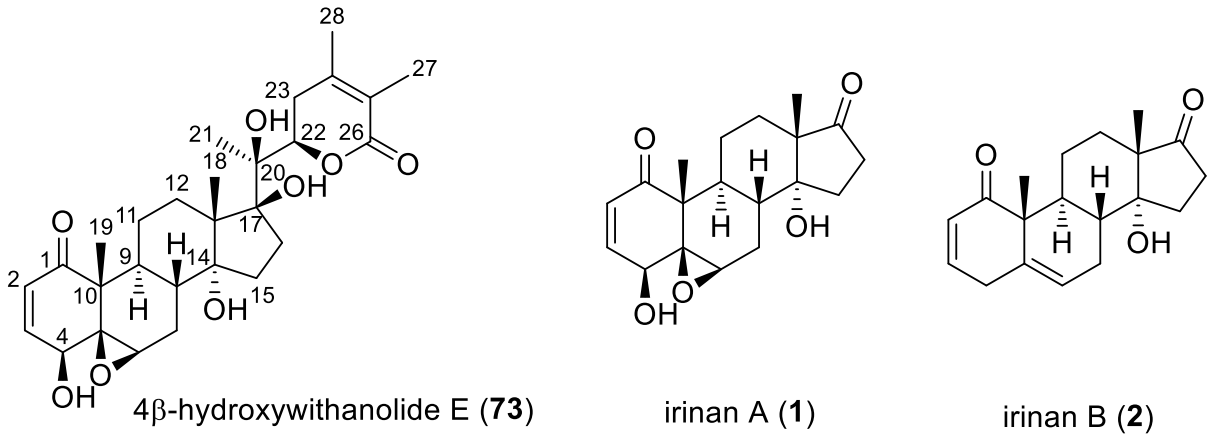


Figure 2.7 Structures of the main withanolide 4β-hydroxywithanolide E (**73**) and truncated withanolides irinan A (**1**) and B (**2**) isolated from *Physalis peruviana*.

[ii] For ^{1}H-NMR see Figure 10.5 and Figure 10.6 in the appendix. For further NMR data see Stein *et al.* [172]

Table 2.1 ^{13}C and ^{1}H NMR data ($CDCl_3$, 500 MHz, 298 K) of irinans A (**1**) and B (**2**) in comparison to the known compound 4β-hydroxywithanolide E (**73**, $CDCl_3$, 400 MHz, 298 K, δ in ppm, *J* in Hz). [172]

	^{13}C			^{1}H		
position	**73**	**1**	**2**	**73**	**1**	**2**
1	201.9	202.0	203.8	-	-	-
2	133.2	132.3	128.1	6.22 (1H, d, 9.9)	6.22 (1H, d, 10.0)	5.91 (1H, ddd, 10.0, 3.1, 1.2)
3	141.4	142.1	145.3	6.92 (1H, dd, 9.9, 6.1)	6.94 (1H, dd, 10.0, 5.8)	6.79 (1H, ddd, 10.0, 5.0, 2.6)
4	70.4	69.9	33.6	3.74 (1H, d, 6.1)	3.79 (1H, dd, 5.8, 2.4)	3.31 (1H, dddddd, 21.3, 2.8, 2.8, 2.8, 2.8, 2.8)[a] 2.88 (1H, dd, 21.2, 4.9)
5	64.2	63.9	135.7	-	-	-
6	63.1	63.1	124.3	3.28 (1H, br s)	3.37 (1H, m)	5.64 (1H, dt, 5.7, 2.0)
7	26.0	24.9	24.1	2.03 (2H, m)	2.11 (1H, dt, 14.2, 3.1) 1.84 (1H, ddd, 14.1, 11.7, 1.4)	2.08 (1H, m) 1.95 (1H, m)
8	34.3	32.6	35.5	1.83 (1H, m)	1.90 (1H, m)	1.88 (1H, m)
9	36.7	38.1	37.1	1.69 (1H, m)	1.51 (1H, m)	2.10 (1H, m)
10	47.9	47.8	50.9	-	-	-
11	21.5	20.4	21.6	1.72 (1H, m) 1.56 1H, (m)	1.91 (1H, m) 1.46 (1H, m)	2.34 (1H, m) 1.52 (1H, m)
12	29.8	24.3	25.0	2.25 (1H, m) 1.28 (1H, m)	1.66 (1H, d, 13.2) 1.55 (1H, m)	1.86 (1H, m) 1.63 (1H, m)
13	54.6	52.6	52.5	-	-	-
14	81.9	80.9	81.0	-	-	-
15	32.5	30.0	29.9	1.66 (1H, m) 1.59 (1H, m)	1.92 (2H, m)	1.96-1.85 (2H, m)
16	38.0	33.1	33.1	2.72 (1H, m) 1.45 (1H, m)	2.44 (1H, ddd, 18.9, 7.6, 4.1) 2.33 (1H, dt, 18.8, 8.8)	2.35-2.46 (2H, m)
17	87.8	218.0	218.5	-	-	-
18	20.4	17.9	18.1	1.07 (3H, s)	1.01 (3H, s)	1.05 (3H, s)
19	16.9	17.8	19.2	1.42 (3H, s)	1.45 (3H, s)	1.27 (3H, s)
20	79.2	-	-	-	-	-
21	19.8	-	-	1.42 (3H, s)	-	-
22	79.7	-	-	4.88 (1H, dd, 11.8, 5.3)	-	-
23	34.4	-	-	2.51 (2H, m)	-	-
24	150.8	-	-	-	-	-
25	121.6	-	-	-	-	-
26	166.0	-	-	-	-	-
27	12.5	-	-	1.88 (3H, s)	-	-
28	20.8	-	-	1.94 (3H, s)	-	-
14-OH				n.d.	1.41 (1H, br s)	1.41 (1H, br s)
4-OH				n.d.	2.57 (1H, d, 2.50)	n.d.

[a] apparent dsext
n.d. not detected

2.3.1 Discussion of Irinanes: A subclass of androstanes in Physalis?

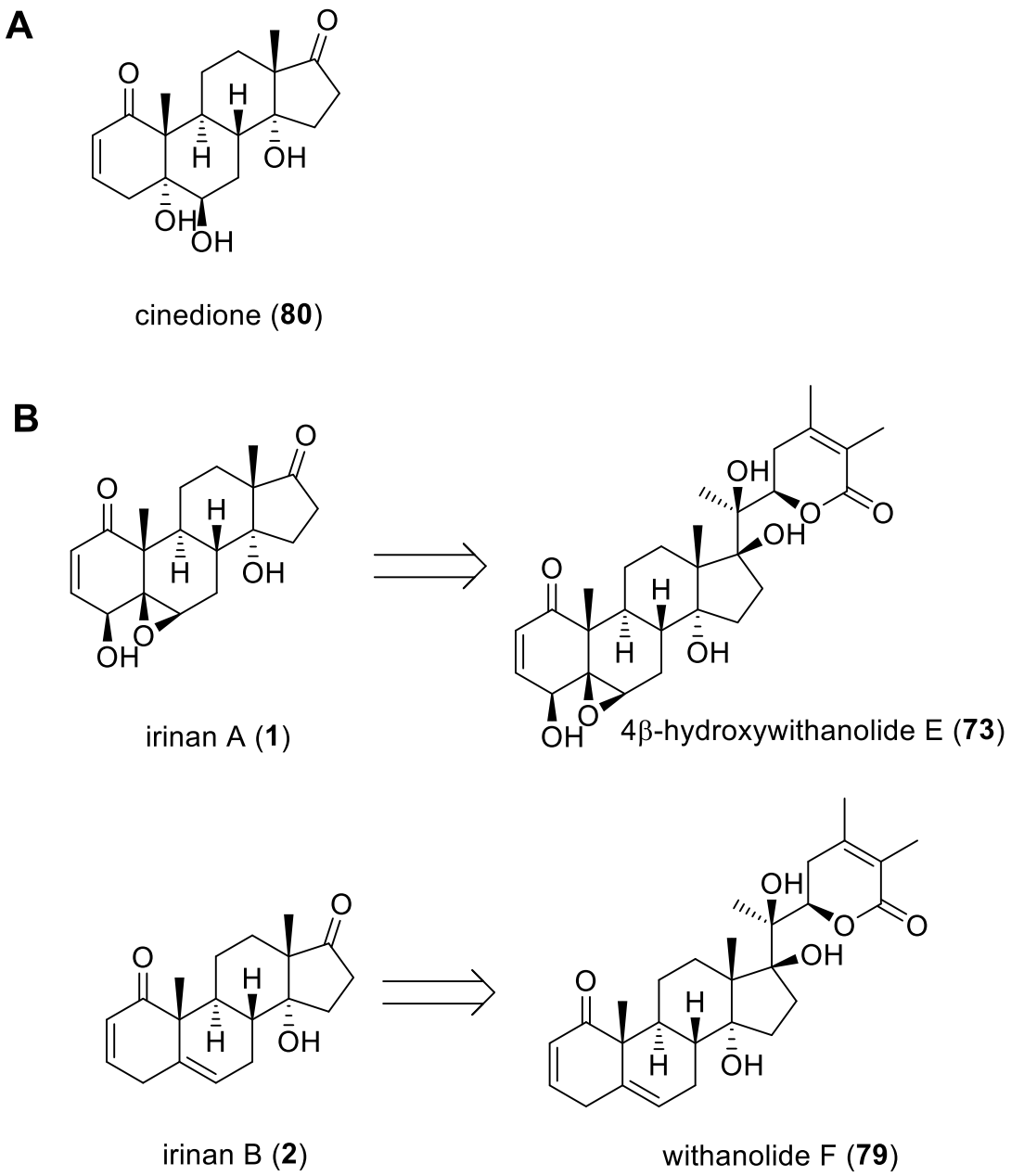
Since the core structure in both irinan A (**1**) and B (**2**) resembles the ergostane-type structure of withanolides, it was assumed that irinans are degradation products of withanolides. Consequently, **1** would be derived from 4β-hydroxywithanolide E (**73**) and the potential precursor for **2** would be withanolide F (**79,** see Scheme 2.1B).

To further prove stereochemistry of irinan A (**1**), 4β-hydroxywithanolide E (**73**) was oxidatively cleaved to **1** with sodium periodate at 70 °C.[iii] The product was analyzed via LCMS and showed matching retention times and masses to irinan A (**1**) samples. Furthermore, 4β-hydroxywithanolide E (**73**) was exposed to various solvents (chloroform, methanol, DMSO, acetonitrile) and pH conditions (pH 3 at 70 °C, pH 0 and pH 11 at 70 °C). [iv] Formation of **1** was observed in none of these tests.

As the cleavage of 4β-hydroxywithanolide E (**73**) to irinan A (**1**) had shown, harsh oxidative conditions were required, while various other chemical conditions could not evoke formation of **1**. Since no oxidizing agents were used during isolation, it is possible that androstane-type compounds occur naturally in *Physalis*. A similar compound, cinedione (**80**, see Scheme 2.1A), has been isolated from *Physalis cinerascens,* [178] suggesting that this cleavage might be characteristic for *Physalis* species. As a potential cleavage of withanolides and accumulation of irinan B (**2**) implied that the precursor withanolide F (**79**) would be present in this plant, all side fractions of the isolation process were screened again. NMR analysis showed trace amounts of withanolide F (**79**) [174] in the residing flash-chromatography fraction F3. [172]

[iii] Experiment performed by Dave Biedermann

[iv] Experiment performed by Dave Biedermann



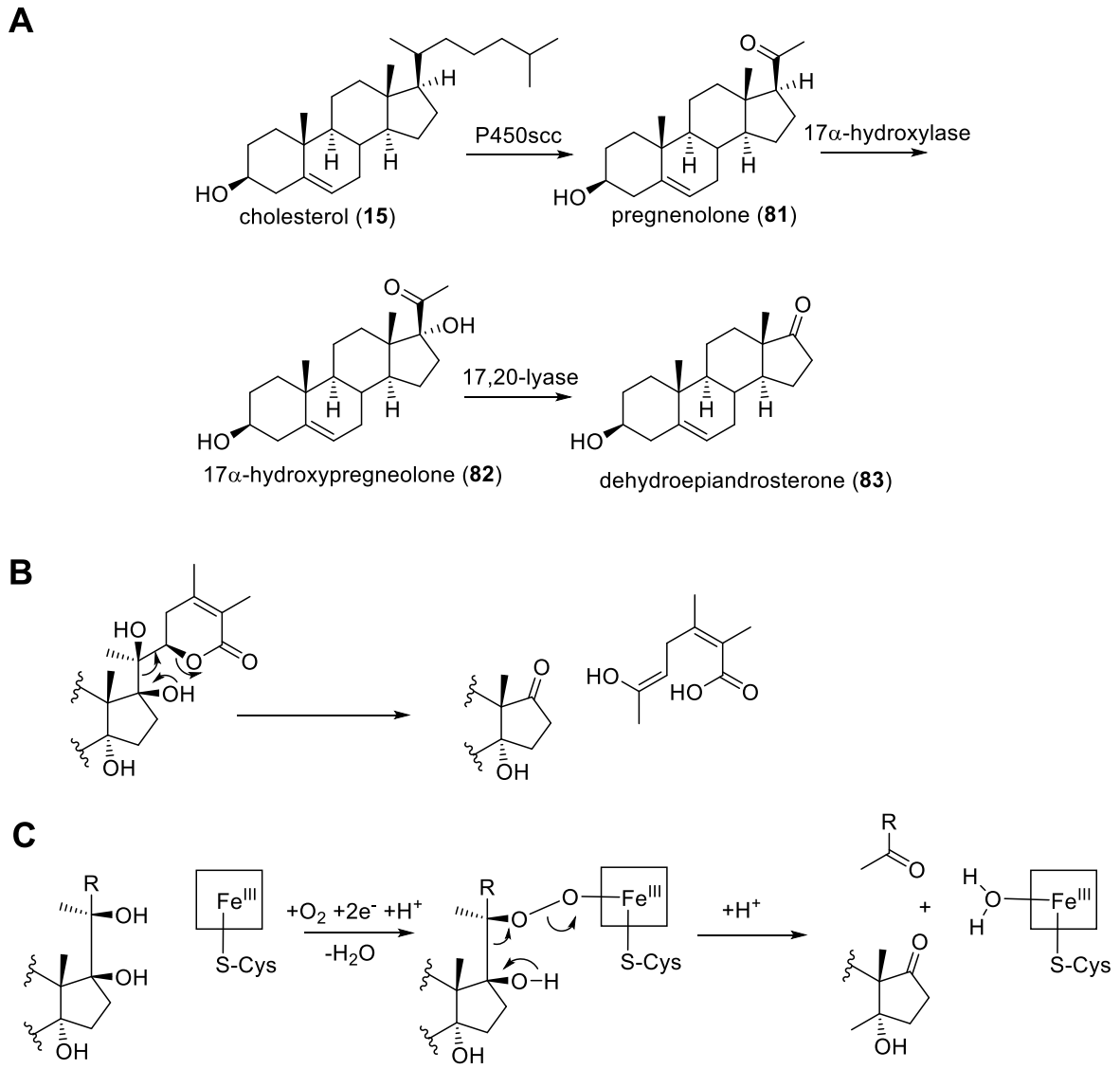
Scheme 2.1 Structure of previously reported cinedione (**80**) from *Physalis cinerascens* (A) [178] and presumed precursors of irinan A (**1**), B (**2**) isolated from *Physalis peruviana* (B).

Irinanes, with their lack of the typical side-chain lactone ring, resemble human sex hormones, called androstanes. [179] Androstane-type compounds have been rarely reported in plants. [178,180-182]

In human androstane biosynthesis cholesterol (**15**) is cleaved to pregnenolone (**81**) by the cholesterol side-chain cleavage P450 (P450scc), which catalyzes the cleavage of the C20-C22 bond via sequential hydroxylation (Scheme 2.2A). [183,184] The cholesterol 17α-hydroxylase/17,20-lyase can then catalyze the hydroxylation to 17α-hydroxypregnenolone (**82**) and subsequent deacetylation at C-17 yielding dehydroepiandrosterone (**83**). The presumed

ferric peroxide mechanism is discussed by Yoshimoto *et al*. [185] However, related enzymes in plants have not been reported to date. A tBLASTn search of *P. peruviana* transcriptome data, [155] conducted by Prof Dr. Jakob Franke concluded that no P450 enzymes similar to human androstane biosynthesis enzymes (UniProtKB accession P05108 and P05093) exist in *P. peruviana*. [172]

Furthermore, the substitution pattern of withanolides differs strongly from the one in cholesterol (**15**), suggesting a different mechanism for irinane biosynthesis. As possible routes for the degradation, two mechanisms come to mind. Either a non-oxidative Grob-fragmentation (Scheme 2.2B) or an oxidative cleavage. The latter could employ a similar mechanism as the P450scc, via a ferric peroxide intermediate (Scheme 2.2C). [186] Several P450s have been reported to perform cleavage of diols. [79,186]



Scheme 2.2 Cholesterol (**15**) is cleaved to androstandes by three P450 enzymes (A); Possible routes for degradation of withanolides to irinanes: Grob-fragmentation (B) or cleavage performed by a P450 (C).

As withanolides possess potent bioactivity, [187,188] the question arose whether loss of the side-chain lactone would affect their antiproliferative activity. Therefore **73**, **1** and **2** were forwarded to Prof. Dr. Mark Brönstrup (Helmholtz Centre for Infection Research) to perform bioactivity tests against four different tumor cell lines. Bioactivity of irinanes may be of interest, as synthesis and potential use as a medicinal product would be facilitated since less stereocenters would be required.

The EC_{50} values measured for **73** [172] corresponded to previously reported values. [170,189,190] Irinan A (**1**) and B (**2**) exhibited 1.3 to 10-fold less activity compared to 4ß-hydroxywithanolide E (**73**). [172] An exception was the activity of irinan B (**2**) in A549 cells, where similar EC_{50} values were measured. [172] It was concluded that irinans exhibit antiproliferative activity, even though reduced compared to 4ß-hydroxywithanolide E (**73**). This is in agreement with the literature, as withanolides with an α,β-unsaturated ketone in ring A and an α,β-unsaturated δ-lactone group were reported to exhibit the highest antitumor activity. [17,191]

2.3.2 Identification of main withanolides in *Physalis peruviana* methanol extracts

The previous section described the isolation of unknown compound from *Physalis peruviana*. However, the main motivation for this chapter was the identification of main withanolides from producing plants, to identify potential markers for silencing experiments. For VIGS experiments *Physalis peruviana* MeOH/H_2O (90:10) crude extract should be used for high throughput analysis of changes in the withanolide profile. Such a crude extract was was examined for the main compounds. Here 4β-hydroxywithanolide E (**73**), irinan B (**2**) and withanolide F (**79**) could be identified in the chromatogram. However, the compound that gave the major peak (*****, Figure 2.8) had not been isolated in previous experiments.

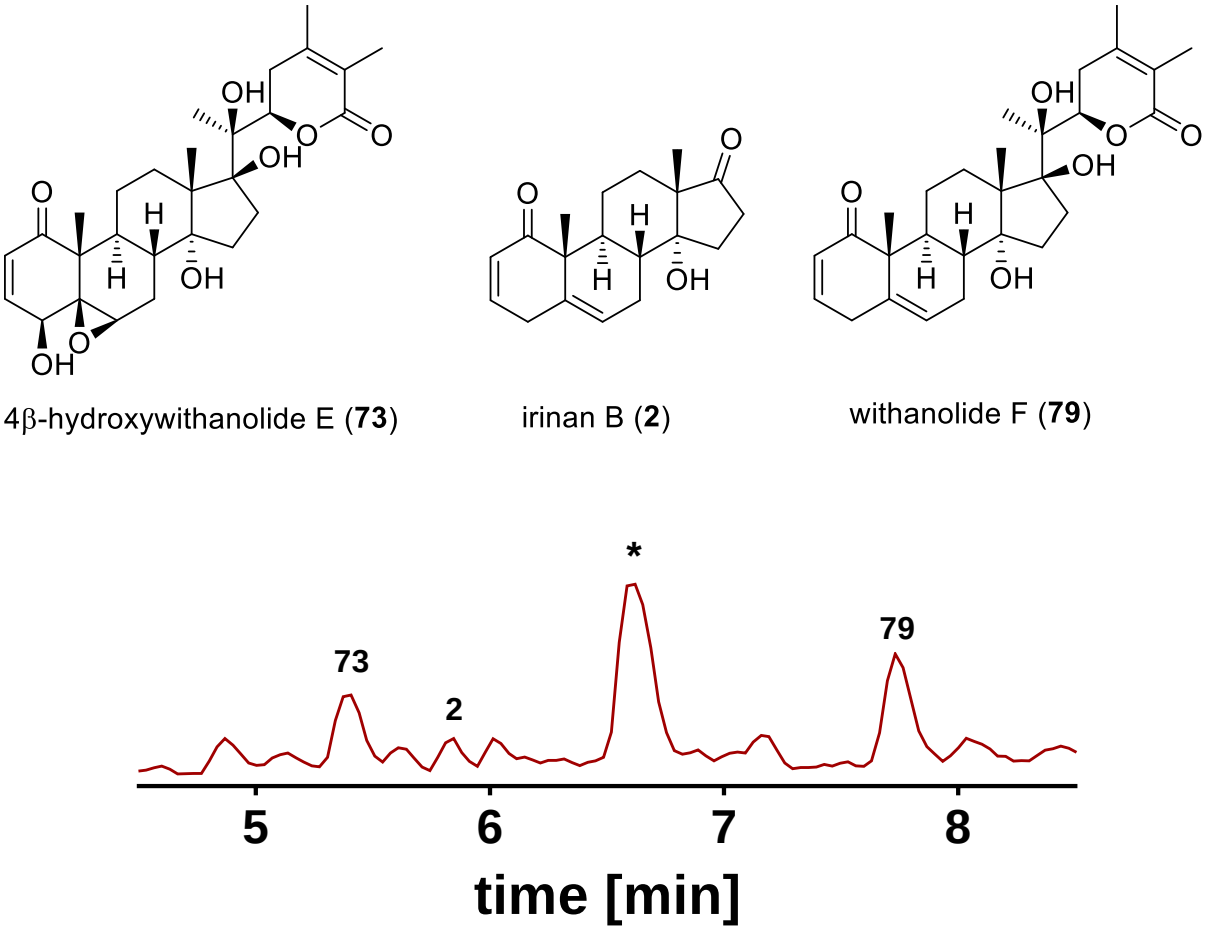


Figure 2.8 Exemplary LCMS chromatogram (TIC) of *Physalis peruviana* MeOH crude exhibits main withanolides (**73**, **2** and **79**) and one unidentified compound (*: unidentified compound).

Consequently, isolation and identification of this compound had to be conducted and was performed within the scope of the bachelor's thesis of Katrin Dumann. Here, 8–10-week-old *P. peruviana* plants (73 g) were processed according to the protocol developed in section 2.2. Ground plant powder was extracted with H_2O:MeOH (3:1) and the concentrated extract was divided into fractions soluble in petroleum ether, chloroform (735 mg), and *n*-butanol (263 mg). The $CHCl_3$ extract was further separated with automated flash chromatography and fractions were combined in three major subfractions (F1: fractions 23-29; F2: 30-31; F3: 32-36).

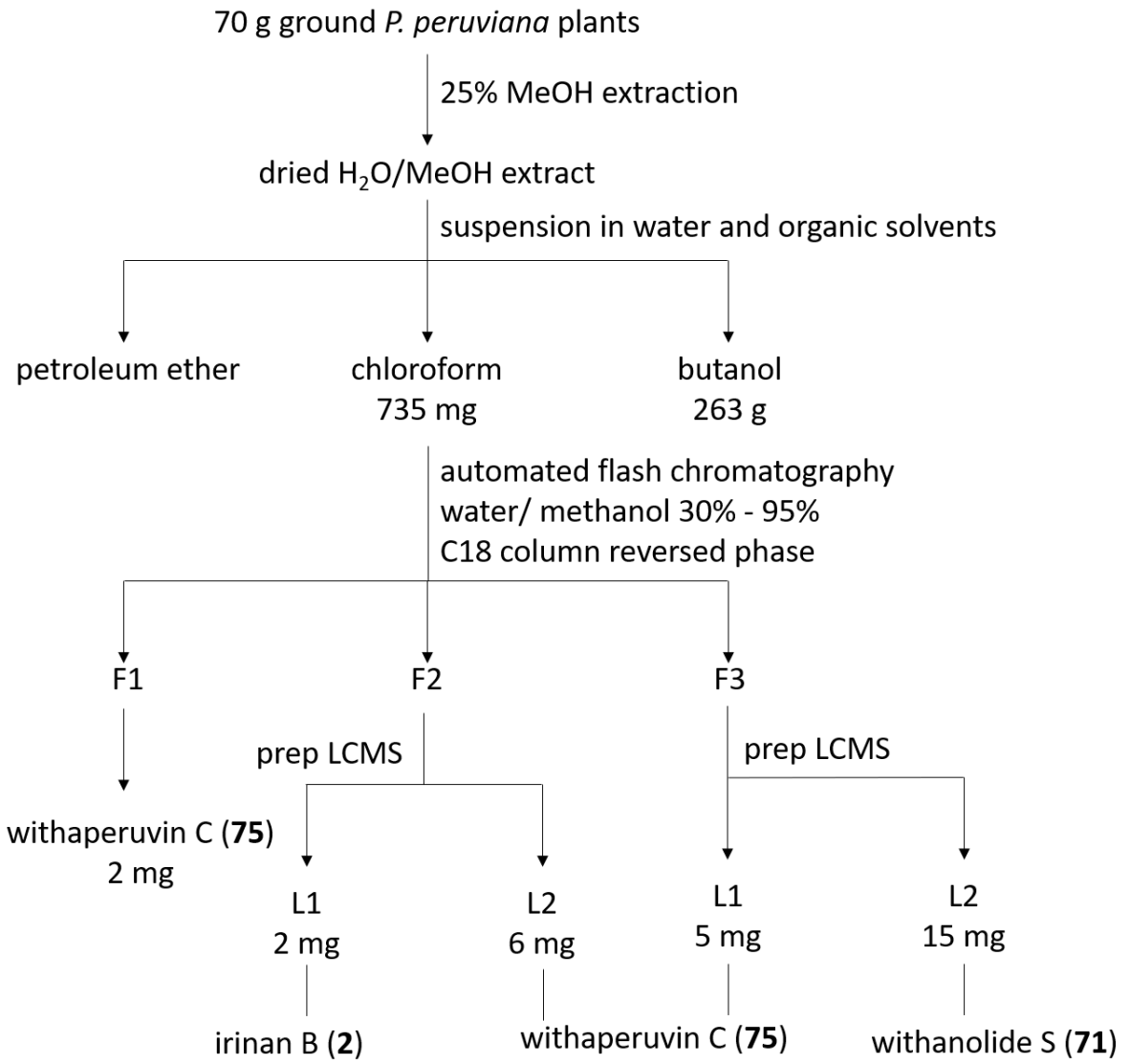


Figure 2.9 Schematic depiction of the isolation process for withaperuvin C (**75**) and withanolide S (**71**) from *P. peruviana* plants.

The subfractions were further analyzed with LCMS and NMR. While F2 contained some irinan B (**2**), it also contained the unidentified compound depicted in Figure 2.8. Separation of this compound with preparative LCMS and comparison to NMR reference data concluded that it is withaperuvin C (**75**, Figure 2.10). [192] Consequently, F1 and F3 were investigated for their withanolide content as well. While F1 contained mostly withaperuvin C (**75**), F3 was separated with preparative LCMS into two subfractions (L1-L2). L1 contained **75**, while L2 was identified as withanolide S (**71**) by comparison to literature NMR data. [6,193]

In conclusion, the two known withanolides (withaperuvin C (**75**) and withanolide S (**71**), Figure 2.10) were identified in locally grown *Physalis peruviana* extracts within the scope of Katrin Dumann's bachelor thesis.

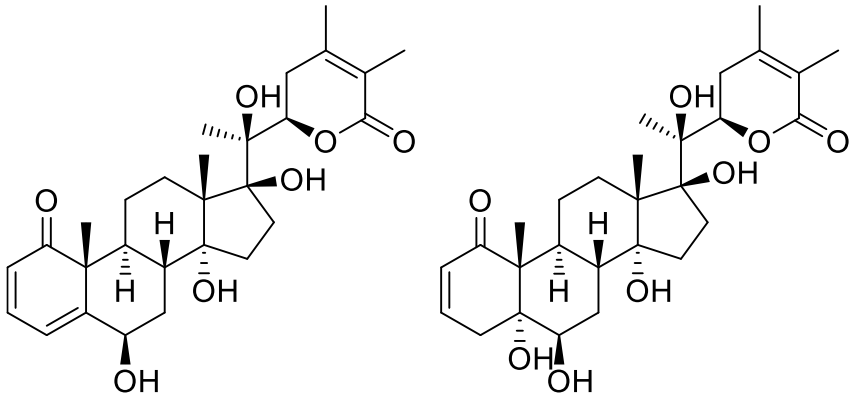


Figure 2.10 Known compounds withaperuvin C (**75**) and withanolide S (**71**) were isolated and identified from *Physalis peruviana* within the scope of Katrin Dumans bachelor thesis.

2.4 Conclusion

For the conduction of further experiments (e.g., silencing and expression), a strategy for isolation of main withanolides was needed to obtain potential reference compounds and substrates. Therefore, a protocol for purification of withanolides from *W. somnifera* and *P. peruviana* was established.

The analysis of *W. somnifera* root powder showed that withanolides are mostly present in the chloroform fraction, compared to petrol ether or butanol fractions, which matched reports from the literature. [152,154,173] The further separation of the chloroform extract was performed via automated flash chromatography. Here, withaferin A (**5**) and withanolide A (**9**) were identified in comparison to literature data. [11,19,161,175,176] However, both fractions still contained impurities, showing that subsequent isolation attempts need to be refined. An analysis of Samuel E. Hakim targeting *W. somnifera* plants cultivated for this project identified withaferin A (**5**), withanolide D (**70**) and withanolide A (**9**) as the main withanolides.

Furthermore, self-cultivated *P. peruviana* were examined for their main withanolides. Here, two sets of experiments were performed that gave an overview about the main withanolides in this plant. At first, a chloroform crude extract was pre-purified using automated flash chromatography and obtained subfractions were further separated by preparative LCMS. NMR analysis of the LCMS-prep fractions resulted in the identification of the main withanolide (4β-hydroxywithanolide E, **73**) and furthermore two new, truncated withanolides with an androstane backbone. Detailed structure elucidation was performed by Dave Biedermann using HSQC, HMBC and COSY NMR. The isolated compounds were termed irinan A (**1**) and irinan

B (**2**) and are assumed to be cleavage products of the withanolides **73** and **79** observed in *P. peruviana*. The isolated compounds were further investigated for their bioactivity by the group of Prof. Mark Brönstrup and found 1.3 to 10-fold less activity compared to 4ß-hydroxywithanolide E (**73**), demonstrating the importance of the side chain lactone for cytotoxic activity. [172] The isolation of irinans has been published in Beilstein Journal of Organic Chemistry. [172]

As VIGS experiments with *Physalis peruviana* were planned, a MeOH crude extract was examined for its main withanolides. Here, previously identified withanolides 4β-hydroxywithanolide E (**73**), irinan B (**2**) and withanolide F (**79**) were detected. In addition, another, yet unidentified compound accumulated. Isolation of this peak and comparison to literature data [192] identified it as withaperuvin C (**75**). [v]

Furthermore, withanolide S (**71**) was detected in the remaining fractions. Both compounds were not detected in previous analyses, demonstrating that *Physalis peruviana* possesses a diverse and fluctuating withanolide profile.

[v] Isolation and identification were performed by Katrin Dumann for her bachelor's thesis under supervision in this work.

3 Attempts towards a 24-methyldesmosterol-producing *Saccharomyces cerevisiae* strain

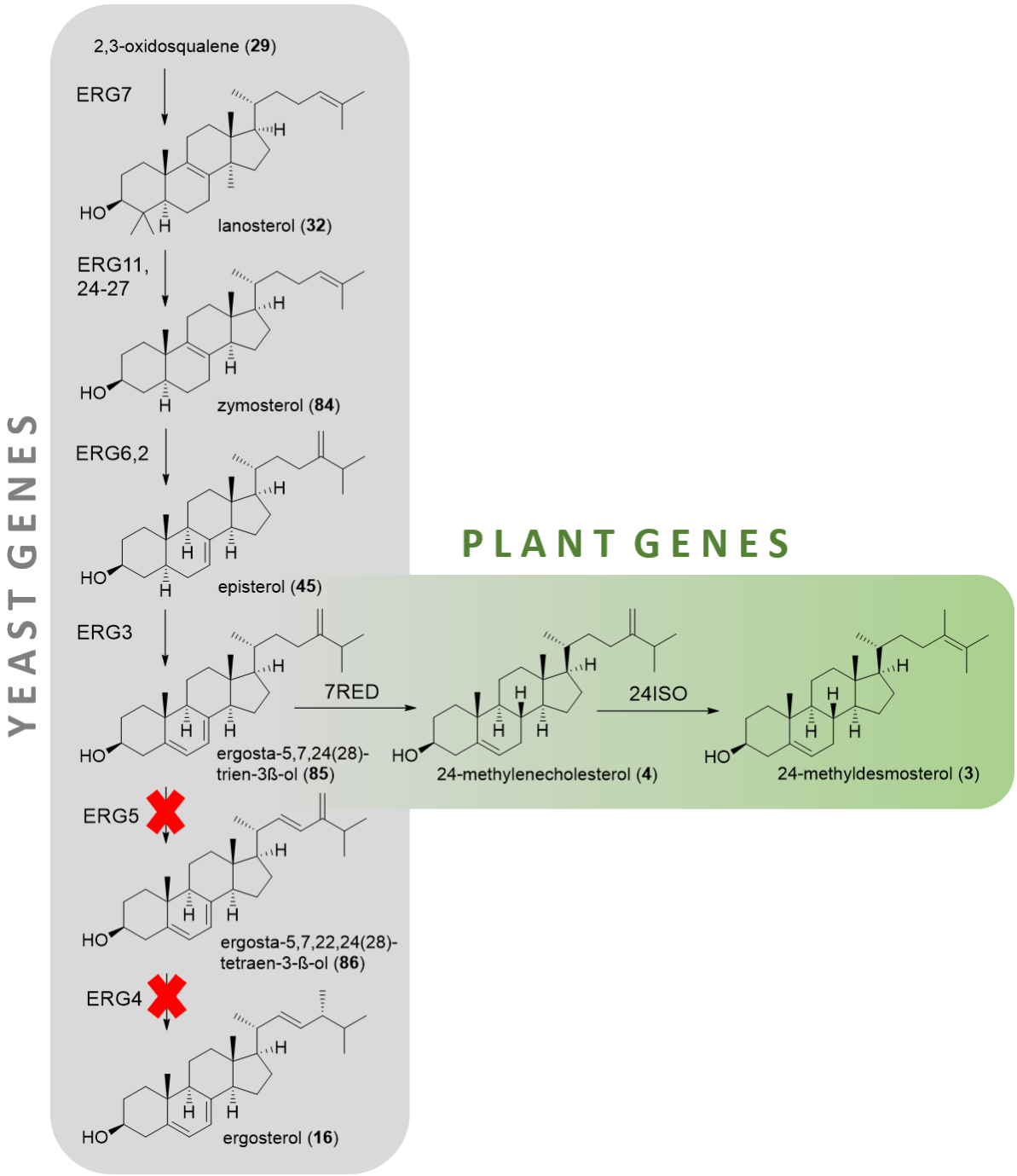
3.1 Introduction

In order to perform enzyme assays in planned experiments (see chapter 6) a substrate has to be accessible. The last known substrate in withanolide biosynthesis is 24-methyldesmosterol (**3**). [64] As this compound is not commercially available, and yeast would not be able to assimilate exogenous sterols under aerobic conditions, [194,195] yeast metabolism should be modified instead. Furthermore, this approach would create a set of *S. cerevisiae* strains which are customizable to the current hypotheses of withanolide biosynthesis, as genes could be added to the yeast strain and compounds of interest could be purified for structural analysis.

In order to generate a 24-methyldesmosterol (**3**) producing strain, ergosterol biosynthesis should be redirected towards the synthesis of **3** by including two plant genes. Sawai *et al.* constructed a 24-methylenecholesterol (**4**) producing strain by expressing a *sterol Δ7-reductase* [62] from plasmid in a *Δerg4/Δerg5* strain. [196] This concept should now be expanded to enable 24-methyldesmosterol (**3**) production.

In general, **3** is structurally close to the main fungal sterol ergosterol (**16**), only differing in the B-ring and side chain, so that modifications at the right intermediate could generate the desired structure. Like other sterols, ergosterol (**16**) is synthesized via condensation of C_5 isoprene units, generated by the mevalonate pathway (see chapter 1.3.1). [197,198] In contrast to phytosterol biosynthesis, 2,3-oxidosqualene (**29**) is cyclized to lanosterol (**32**) in fungi. [199] Scheme 3.1 illustrates key steps in ergosterol biosynthesis, with the intermediates being lanosterol (**32**), zymosterol (**84**), episterol (**45**), ergosta-5,7,24(28)-trien-3ß-ol (**85**) and ergosta-5,7,22,24(28)-tetraen-3-ß-ol (**86**). In general, deletion strains of the late ergosterol biosynthesis genes show compromised phenotypes (e.g. lowered tolerance against osmotic stress) but are viable.[200]

Consequently, knockouts of *ERG4* and *ERG5* would be non-lethal and should increase the availability of ergosta-5,7,24(28)-trien-3ß-ol (**85**), similar to the concept reported in the literature. [196] This intermediate could then be converted by two plant enzymes to the desired substrate **3**. The Δ7-reductase (7RED) was previously reported to be flexible enough to accept ergosta-5,7,24(28)-trien-3ß-ol (**85**) as a substrate and then reduce it to 24-methylenecholesterol (**4**). [196] A second plant enzyme, the Δ24-isomerase (24ISO), has been reported to perform the isomerization of 24-methylenecholesterol (**4**) to 24-methyldesmosterol (**3**). [64] An overview of the planned modifications is illustrated in Scheme 3.1.



Scheme 3.1 The concept for a 24-methyldesmosterol (**3**) producing yeast strain includes knockouts of the last two genes involved (*erg4* and *erg5*) to increase accumulation of ergosta-5,7,24(28)-trien-3ß-ol (**85**). Insertion of plant genes *7RED* and *24ISO* could then enable conversion of (**4**) to the desired substrate (**3**).

Now that the general modifications needed are clarified, an exact strategy for knockouts and gene additions was needed. Sawai *et al.* [196] used BY4742 *Δerg4*, [201] which can be purchased. The knockout of *ERG5* can be achieved using an antibiotic selection marker (Figure 3.1A). For this purpose a nourseothricin cassette (*natMX*) can be amplified from the pAG25 vector [202] and homologous overhangs targeting the *ERG5* locus can be added by PCR. Afterwards, the *7RED* should be integrated into the yeast genome, as this increases genetic stability, compared to the

expression from a vector. [203] For this purpose, the pXP system was selected. [204] Multiple pXP vectors are available that provide different promoters and markers adjoined to the multiple cloning site. Furthermore the respective yeast marker is flanked by *loxP* sites, enabling marker recycling via CreA recombinase. [204] In addition, this publication described the homologous sequences used for integrations targeting multiple common loci in *S. cerevisiae*. [204] After the planned genomic integration of *7RED* (Figure 3.1B), the marker used for selection can be excised by transformation of the CreA plasmid (pBF3060, Figure 3.1C). Lastly, the integration of the *24ISO* at a different site can be conducted using the same marker (Figure 3.1B).

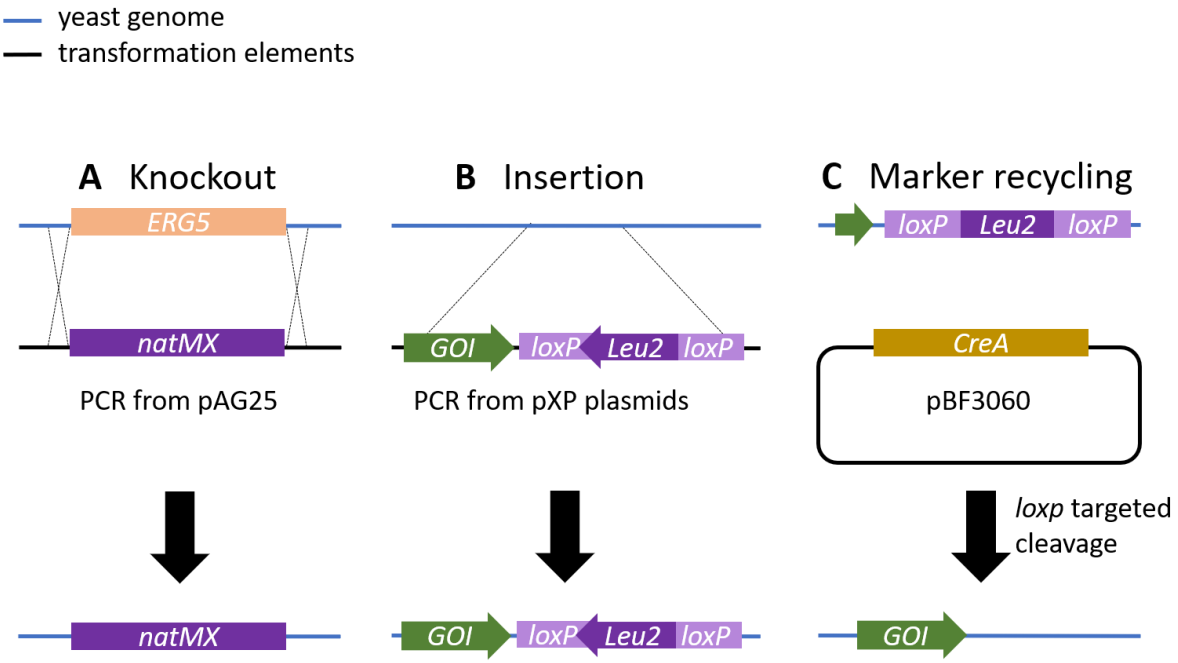


Figure 3.1 Transformations needed for knockout of *ERG5* (A), integration of genes of interest (B) and marker recycling (C) to generate a 24-methyldesmosterol (**3**) producing yeast strain.

3.1.1 Knockout of *ERG5* leads to accumulation of ergosta-5,7,24(28)-trien-3ß-ol

For the generation of a substrate producer strain, knockouts of the two ergosterol (**16**) biosynthesis genes *ERG4* and *ERG5* had to be generated. As single knockouts of haploid yeast strains are readily available in the EUROFAN yeast collection, [201] a BY4742 *Δerg4* strain was purchased (Y14463, EUROSCARF). Furthermore, the knockout of *ERG5* had to be established. Therefore, a nourseothricin resistance cassette (*natMX*) was amplified, using the pAG25 vector [202] as a template, with overhangs specific for the *ERG5* locus. Transformation into the *Δerg4* strain yielded the double knockout strain. Successful transformation was verified by colony PCR with primers spanning the junction of the locus and the insert (Figure 3.2B). The

PCR for the nourseothricin cassette shows bands at ca. 1.1 kb and 1.0 kb, as expected, while for the *ERG5* control no amplification was observed (Figure 3.2A). The generated *Δerg4/Δerg5* yeast strain will further be termed ASY01. It has to be noted that the ASY01 strain was grown on media containing 20 µg/mL ergosterol for subsequent transformations as growth was hampered in this strain. Slowed growth for *Δerg5* strains has been reported before. [205] Even though not lethal, mutations in *ERG4* and *ERG5* can change the composition of the cell membrane, which affects membrane potential, salt tolerance and drug resistance. [200]

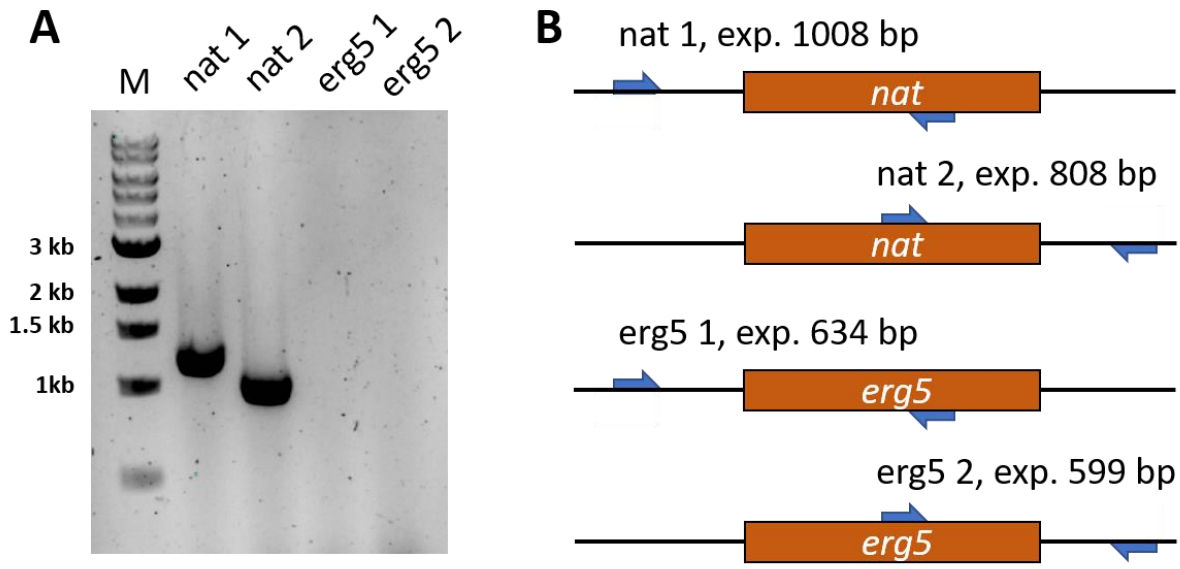


Figure 3.2 Colony PCR of ASY01 showed bands for integration of the *natMX* cassette at the *ERG5* locus (A); Placement of verification primers is locus and insert specific (B); 1% agarose gel; (M): 1 kb ladder; (nat): nourseothricin knockout cassette (nat 1 exp. 1008 bp, nat 2 exp. 808 bp).

Now that the first modification was generated, the question was whether the knockout strains (*Δerg4* and ASY01) exhibited a different sterol profile than a wildtype strain. Therefore, yeast cells were extracted and analyzed via GCMS. For GCMS analysis ASY01 was grown without ergosterol supplementation, to give a representative depiction of the compounds produced in this strain. A comparison of GCMS chromatograms and compounds in their biosynthetic context is depicted in Figure 3.3. For a list of compounds detected in all yeast strains see Table 3.1. In general, compounds were assigned to structures based on three different levels of confidence. [206] The highest confidence level were compounds that were identified via comparison to a standard (e.g. ergosterol, **16**). The second level included compounds that matched the respective reference spectra, but no standard was available to compare retention times. Lastly, the third category included structures that were likely sterol-type compounds due to their mass spectra, but no exact structure or reference spectrum could be assigned.

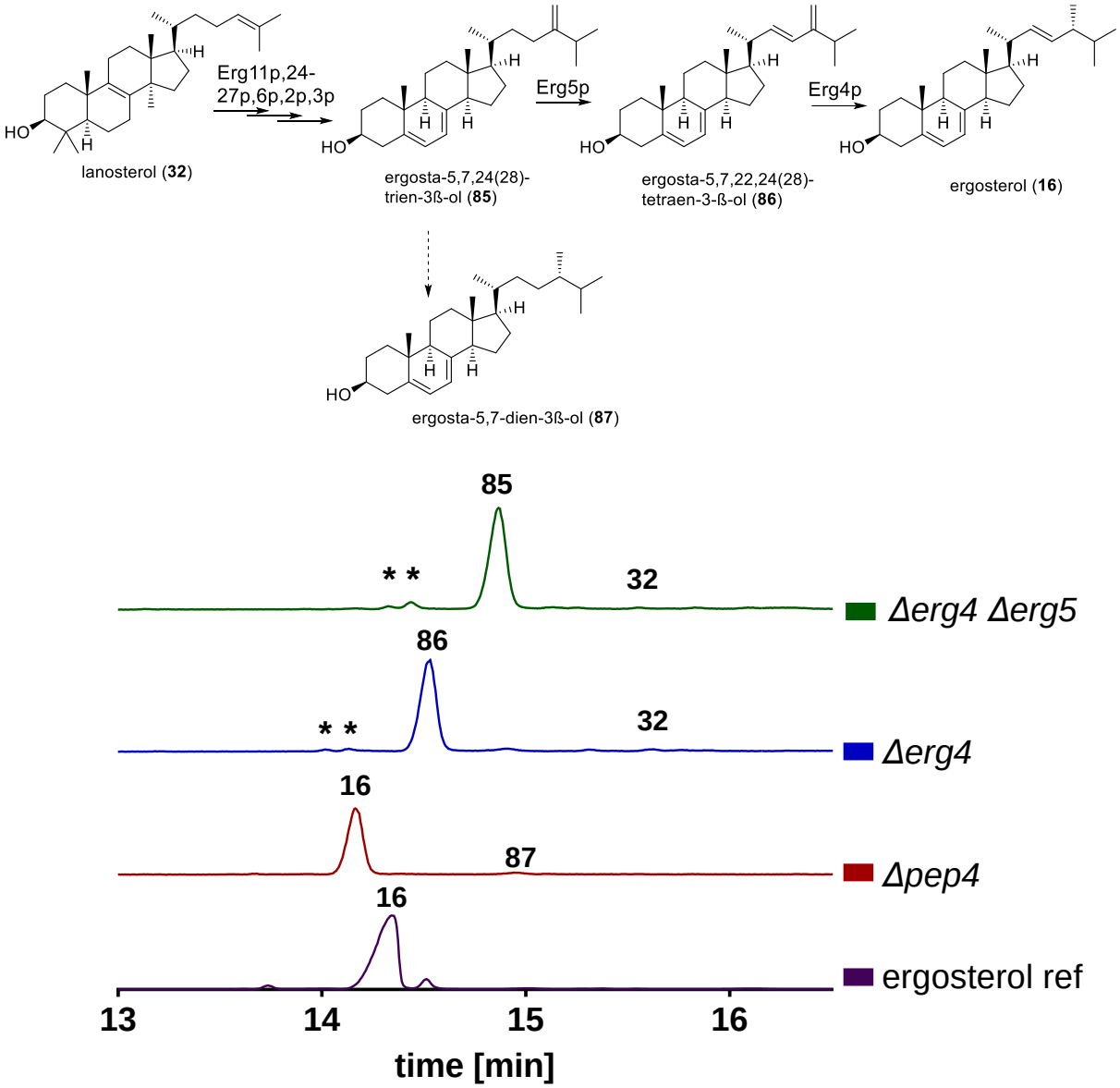


Figure 3.3 GC measurements of *S. cerevisiae* strains Δpep4, Δerg4 and Δerg4/Δerg5 (ASY01) and detected compounds in their biosynthetic context (*: unidentified compound).

For the analysis of *Δerg4* and ASY01 strains the protease-deficient *Δpep4* [207] strain was measured as a control, as it does not contain modifications of the ergosterol (**16**) pathway and was available in the laboratory. The strains *Δpep4*, *Δerg4* and ASY01 were compared to an ergosterol standard (**16**, 14.27 min, mass TMS-ether: 468). [208,209] In the *Δpep4* strain **16** was detected, albeit with slightly different retention times (14.16 min instead of 14.27 min), as measurements were carried out at a different time point. Furthermore trace amounts of a compound at 14.96 min were detected in the *Δpep4* control that exhibited a fragmentation pattern similar to reference spectra of ergosta-5,7-dien-3ß-ol (**87**). [210–212] When comparing the *Δerg4* strain, the peak for **16** is clearly not visible anymore (Figure 3.3). Instead, a new peak at 14.52 min appears, containing a signal at *m/z* 466, which corresponds to the mass of a TMS-

ether of ergosta-5,7,22,24(28)-tetraenol (**86**), [210,213] the product converted by Erg5p and further substrate of Erg4p (Figure 3.3). A comparison to a mass spectrum recorded by Le Fur *et al.* [212] showed that the same fragments were present, albeit with different intensities. Müller *et al.* stated different key fragments for this compound (376, 361 and 251), [210] that matched the observed intensities in this work. Furthermore, the ergosterol precursor lanosterol (**32**) and trace amounts of two unidentified sterol-like compounds were detected (see Table 3.1). It was concluded that ergosterol (**16**) production was lost by the knockout of *ERG4* and the strain accumulates ergosta-5,7,22,24(28)-tetraenol (**86**) instead.

In the ASY01 (*Δerg4/Δerg5*) strain production of **86** was not observed. Instead, a new peak at 14.86 min appears (Figure 3.3). The signals at *m/z* 468, 363, 337, 253 and 211 lead to the assumption that this compound is structurally similar to ergosterol (**16**), presumably ergosta-5,7,24(28)-trien-3ß-ol (**85**), the product of Erg3p and substrate of Erg5p (Figure 3.4). [214] Comparison to reference spectra from the literature supported this assumption. [210,211] Again, trace amounts of lanosterol (**32**) and two unidentified sterol-like compounds were detected (Figure 3.3).

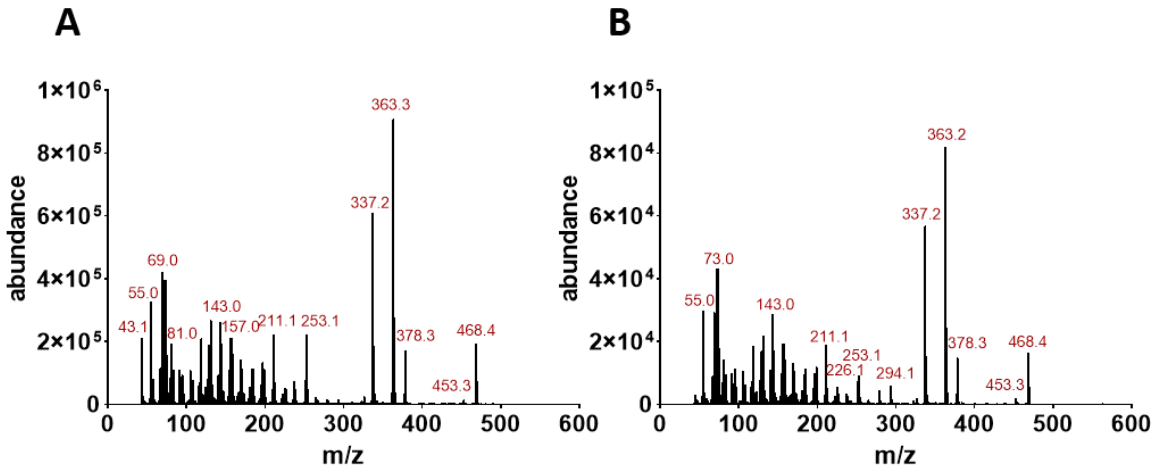


Figure 3.4 Comparison of mass spectra of an ergosterol (**16**) standard at 14.16 min (A) and the peak at 14.86 min accumulated in the ASY01 strain (B) shows that the compound is an isomer of ergosterol (**16**) and comparison to reference spectra concluded it was ergosta-5,7,24(28)-trien-3ß-ol (**85**). [210,211]

It was concluded that knockout of *ERG4* and *ERG5* was successful and ergosterol (**16**) production was lost in favor of a compound that matched reference spectra of the desired precursor ergosta-5,7,24(28)-trien-3ß-ol (**85**). [210,211] It has to be noted that growth of the ASY01 was slower than the reference *Δpep4* [215] strain that contains intact *ERG4* and *ERG5* genes. Similar tendencies in growth behavior have been described previously for *Δerg5* strains. [205] As the next step, the addition of plant genes *7RED* and *24ISO* was planned to evoke conversion towards 24-methylenecholesterol (**4**) and further to 24-methyldesmosterol (**3**).

Table 3.1 Summary of relevant peaks of all chromatograms.

Name	**Retention time [min]**	**Confidence level structure**	**Detected in strain**	**Fragments and intensities measured**	**Fragments and intensities in reference spectra**	**Reference**	**Suggested structure**
unidentified compound	14.01	3	*Δerg4*	464 (6)[a], 374 (20), 359 (11), 339 (7), 325 (9), 266 (7), 263 (11), 251 (100), 249 (34), 235 (11), 223 (14), 211 (9)			
unidentified compound	14.13	3	*Δerg4*	466 (24)[a], 451 (2)[b], 423 (1)[c], 376 (15)[e], 361 (100)[f], 343 (2)[g], 341 (4)[h], 325 (2)[q], 253 (21)[l], 251 (72)[m], 245 (1)[j], 211 (33), 147 (4)			isomer of (**86**)
ergosterol (**16**)	14.30	1	*Δpep4*	468 (21), 453 (2), 378 (19), 363 (100), 337 (68), 253 (25), 251 (8), 211 (42)	468 (42)[a], 453 (5)[b], 378 (32)[e], 363 (100)[f], 343 (29)[g], 341 (7)[h], 337 (100)[i], 253 (48)[l], 251 (18)[m], 227 (10)[p]	208,212,216,217	
24-methylenecholesterol (**4**)	14.32	1	ASY02, ASY03	470 (32), 455 (25), 386 (100), 380 (66), 371 (24), 365 (49), 343 (56), 341 (78), 296 (67), 281 (44), 271 (7), 267 (6), 257 (51), 253 (45), 243 (18), 227 (15), 217 (15), 215 (19), 213 (37), 211 (21)	470[a], 455[b], 386[d], 365[f], 341, 296, 281, 253, 227, 213	64,196,218	
unidentified compound	14.32	3	AS01	466 (15)[a], 451 (19), 376 (100), 361 (40), 343 (9), 299 (15), 292 (14), 279 (24), 277 (32),			isomer of (**86**)

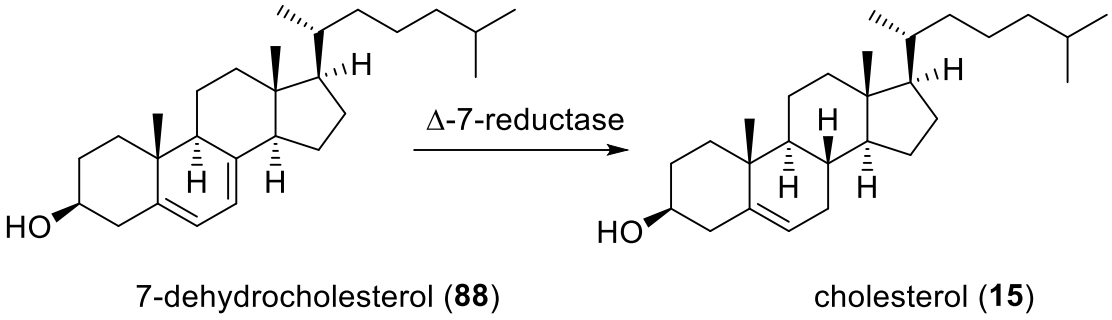
unidentified compound	14.43	3	AS01	468 (24)[a], 453 (5), 378 (7), 363 (100), 337 (77), 279 (8), 253 (13), 251 (12), 211 (18)			side chain isomer of (**16**) or (**85**)
ergosta-5,7,22,24(28)-tetraen-3-ß-ol (**86**)	14.52	2	*Δerg4*	466 (24), 451 (2), 423 (1), 376 (15), 361 (100), 343 (2), 341 (4), 325 (2), 253 (21), 251 (72), 245 (1), 211 (33), 147 (4)	466 (100)[a], 451 (8)[b], 423 (3)[b], 376 (9)[e], 361 (5)[f], 343 (11)[g], 341 (19)[h], 325 (5)[q], 253 (7)[l], 251 (21)[m], 245 (7)[j], 211 (18), 147 (51)	210,212,213	
ergosta-5,7,24(28)-trien-3ß-ol (**85**)	14.86	2	ASY01, ASY02, ASY03	468 (18), 453 (2), 378 (16), 363 (100), 337 (66), 294 (6), 253 (11), 251 (9), 226 (6), 211 (23)	468 (44)[a], 453 (4)[b], 378 (18)[e], 363 (100)[f], 337 (79)[i], 294 (9), 253 (11), 251 (9), 226 (8), 211 (19)	210,211,214,219	
ergosta-5,7-dien-3ß-ol (**87**)	14.96	2	*Δpep4*	470 (14), 380 (20)[e], 365 (100), 352 (10), 340 (19), 339 (70), 266 (10), 253 (26)[m]	470 (43)[a], 455 (5)[b], 380 (23), 365 (100)[f], 337 (17), 253 (12), 211 (20)	210–212	
lanosterol (**32**)	15.56	1	*Δerg4* and following	498 (27), 483 (25), 393 (100), 351(2), 323 (2), 297 (2), 255 (9), 241 (14), 227 (13), 149 (10), 147 (15), 145 (16), 135 (19), 109 (38)	498 (16)[a], 483 (15)[b], 408 (3)[e], 393 (100)[f], 376 (48), 323 (5)[k], 297 (5)[l], 255 (11), 241 (22), 227 (12), 135 (26), 109 (68)	210,212	

Key fragments of TMS-ions described in the literature: a: $[M]^{+\bullet}$, b: $[M-15]^{+}$, c: $[M-43]^{+}$: [M−C(25) to C(27)]$^{+}$, d: $[M-84]^{+\bullet}$: [M−C(23) to C(28)−H]$^{+}$, e: $[M-90]^{+\bullet}$: $[M-(CH_3)_3-SiOH)]^{+\bullet}$, f: $[M-90-15]^{+}$, g: [M−side chain]$^{+}$, h: [M−side chain−2H]$^{+}$, i: $[M-131]^{+}$, j: $[M-90-131]^{+}$, k: [M−90−15−C(23) to C(28)−H]$^{+}$, l: [M−side chain−90]$^{+}$, m: $[M-141]^{+}$: [M−90−C(1) to C(4))]$^{+}$. [210,212]

Confidence levels for structure assignment: 1: identified via standard, 2: matched to reference spectrum, 3: sterol-type compound, but no reference spectrum found.

3.1.2 Insertion of Δ7-reductase leads to unstable 24-methylenecholesterol production

After generation of the *Δerg4/Δerg5* double knockout strain, the gene for redirection of substrate flow towards withanolide precursors should be inserted. The Δ7-reductase (7RED) is known to reduce 7-dehydrocholesterol (**88**) to cholesterol (**15**) in higher eukaryotes. [220] Against common belief, cholesterol (**15**) biosynthesis in plants has been reported as well. [221] Consequently, the *7RED* has been identified from different plant species. [62,196]



Scheme 3.2 Δ7-reductase (7RED) converts 7-dehydrocholesterol (**88**) to cholesterol (**15**) in eukaryotes. [62,196]

It was demonstrated in the literature that this enzyme is flexible enough to accept accumulated ergosta-5,7,24(28)-trien-3ß-ol (**85**) from a *Δerg4/Δerg5* strain and reduce it to 24-methylenecholesterol (**4**). [196] Therefore, a *7RED* gene should be integrated into the yeast genome, as this increases expression stability compared to plasmid based expression. [203] For this purpose, the pXP system was employed, as it provides all necessary tools for systematic metabolic engineering in yeast. [204] For the generation of a *7RED* integration cassette, the vector pXP422 was selected, that contains the *TEF1* promoter and a *LEU2-d8* selection marker (Figure 3.5). [204] The yeast selection marker in all pXP plasmids is flanked by *loxP* sites (Figure 3.5), enabling marker recycling via CreA recombinase. [204] As an integration site the *URA3* locus (Figure 3.5) was selected, since insertion at this site was previously reported with the pXP system. [204]

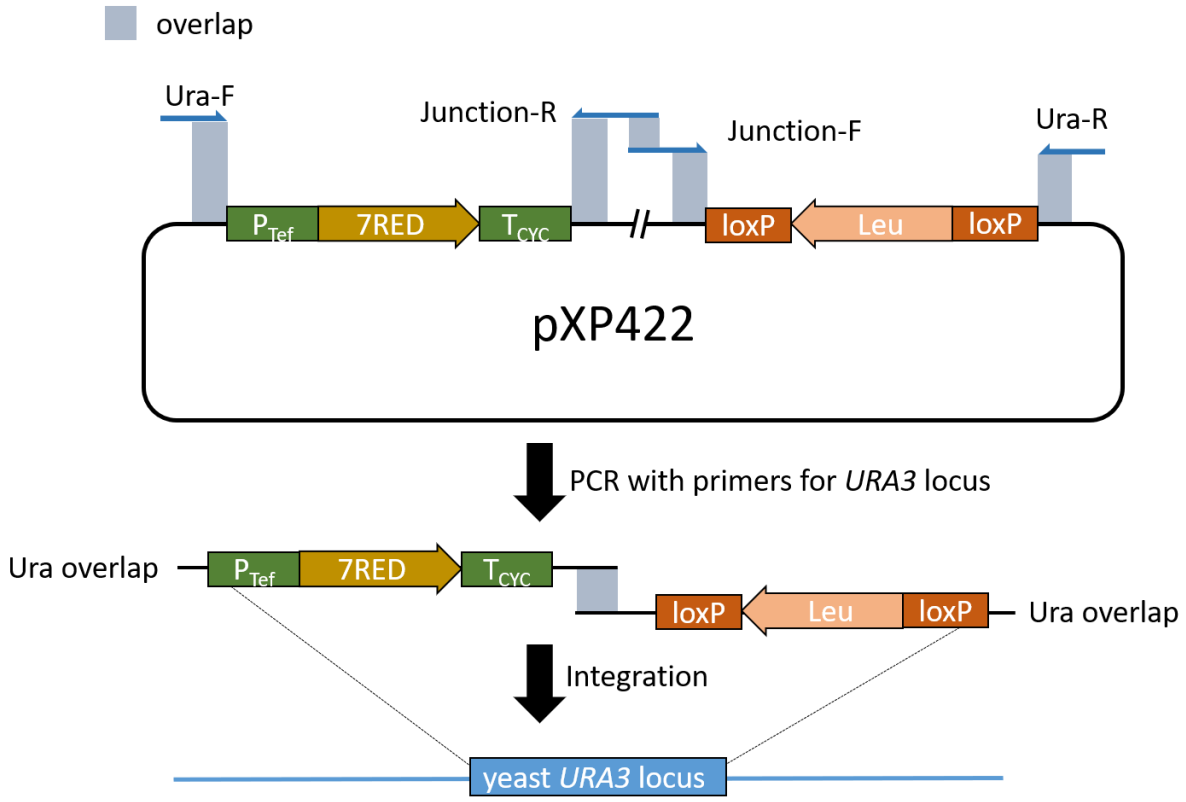


Figure 3.5 Schematic depiction of the *7RED* integration fragments for *URA3* integration, amplified from a pXP422 vector.

Here, two *7RED* genes were tested, as different orthologs might exhibit differences in activity and selectivity. Sawai *et al.* created a yeast strain that produced **4** by adding a plasmid that contained a *S. tuberosum 7RED*. [196] Furthermore the *Physalis peruviana 7RED* was tested, as cDNA of this plant was readily available.

The *Solanum tuberosum 7RED* described in the literature (accession nr.: SGN-U269317, further termed *St7RED*) [196] was amplified from cDNA transcribed from *S. tuberosum* sprout RNA. To determine the *Pp7RED*, a BLASTn search of the *7RED* CDS from *Arabidopsis thaliana* (DWF5; accession nr.: NM_103926) with the *P. peruviana* transcriptome [155] was conducted. Transcript c12488_g1_i1 was the best hit with 74.0% nucleotide identity and a query coverage of 94.6% and will further be termed *Pp7RED*. The insert sequence was obtained from *P. peruviana* aerial tissue cDNA. Both pXP422-7RED plasmids were constructed using Gibson assembly (In-Fusion®, Takara). Successful generation of plasmids was confirmed by colony PCR and DNA sequencing (Eurofins, Ebersberg).

The integration of a *7RED-LEU2d* cassette into the *URA3* locus was conducted using two integration fragments amplified from the respective pXP422-7RED plasmid. Each fragment contained one overhang specific to the URA3 locus and one overhang specific to the respective other fragment (junction site, see Figure 3.5).

ASY01 was transformed with the two overlapping integration fragments generating strains ASY02 (*Pp7RED*) and ASY03 (*St7RED*). The transformation mix was streaked out on selection media containing 20 µg/mL ergosterol, as the ASY01 strain had already shown hampered growth. Transformants were examined by colony PCR.

Colony PCRs of *7RED* strains with insert specific primers were positive. Consequently, GCMS analysis of the strains was conducted. A *Pp7RED* (ASY02) strain showed the ergosta-5,7,24(28)-trien-3ß-ol (**85**) peak at 14.86 min that was previously observed in the ASY01 strain. Furthermore, a new peak at 14.42 min was observed that matched the retention time mass spectrum of a 24-methylenecholesterol (**4**) standard (Figure 3.6). However, when examining the same cryostock with slightly different conditions (different age of cultivation media) accumulation of **4** was not always observed. After attempted marker recycling, **4** was not detected anymore. Similar behavior was observed for the *St7RED* (ASY03) strains. Here, accumulation of **85** and **4** was detected (Figure 3.6). In general, *St7RED* strains analyzed showed stronger accumulation of 24-methylenecholesterol (**4**) compared to the ASY02 strain. However, production of **4** was not observed under all conditions examined even though the same cryostock was used. Here, marker recycling was not attempted.

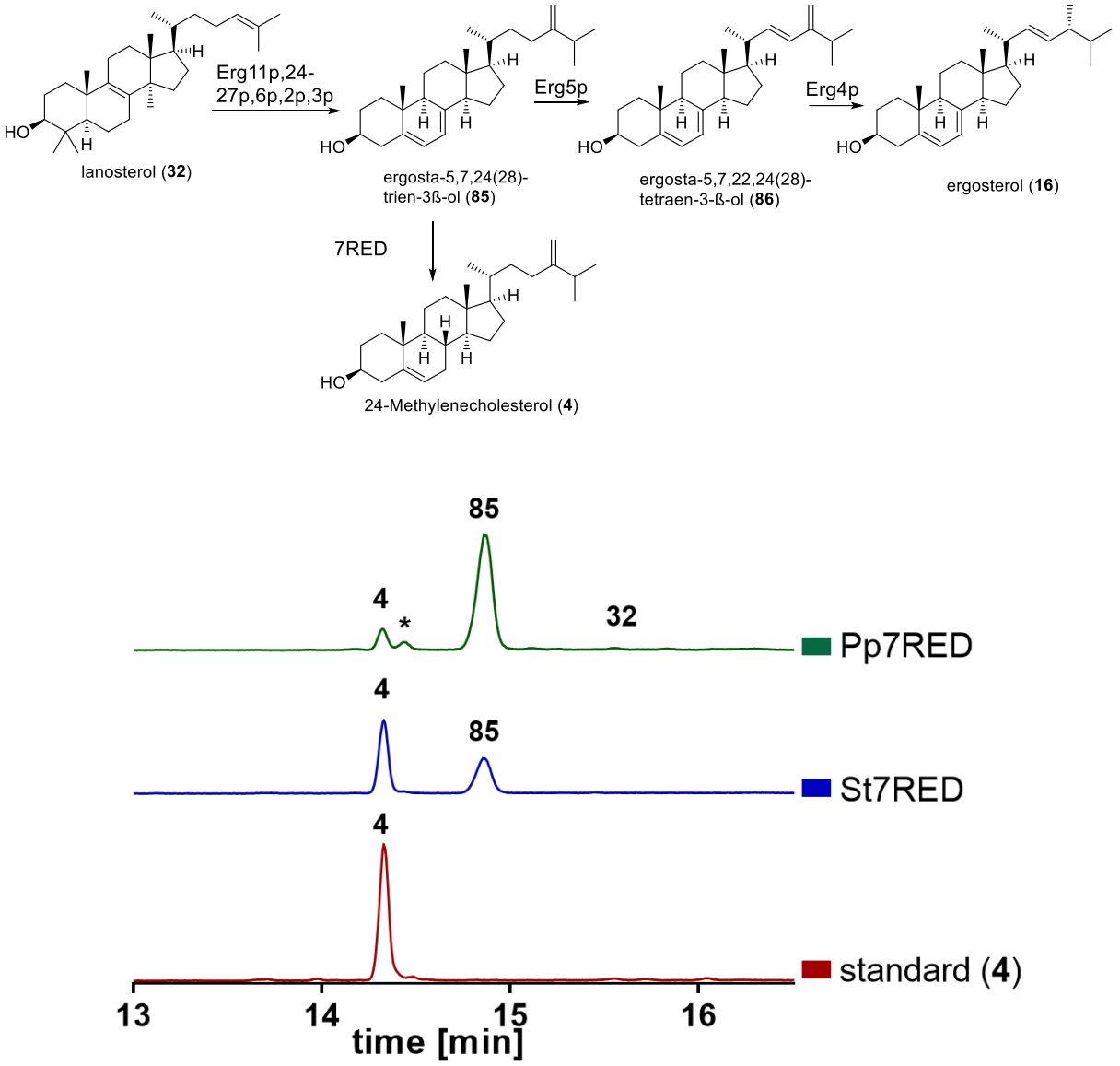


Figure 3.6 GCMS measurements of *S. cerevisiae* strains *Pp7RED* (ASY02) and *St7RED* (ASY03) in comparison to a 24-methylenecholesterol (**4**) standard and other identified compounds (*: unidentified compound).

To find an explanation for inconsistent 24-methylenecholesterol (**4**) production, further colony PCRs were conducted with the strains. While initial colony PCRs, conducted before GCMS analysis, were performed with insert specific primers (Figure 3.7A), now primers spanning the junction of the *7RED* insert and the *URA3* locus were used (Figure 3.7B).

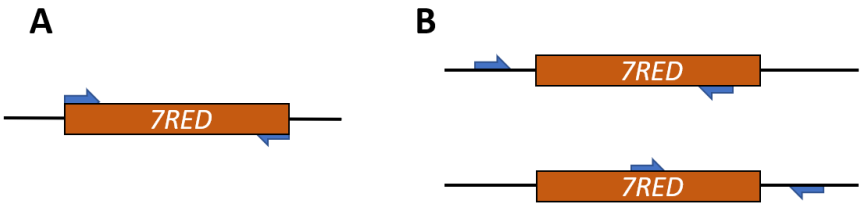


Figure 3.7 Primers used for 7RED colony PCR were insert specific (A) or spanning the junction of the URA3 locus and the 7RED insert (B).

However, the *7RED* was not detected at the *URA3* locus (data not shown). This suggests that the respective *7RED* gene was integrated elsewhere in the genome and exhibited unstable activity. It has to be noted that growth of ASY02 and ASY03 was slower than for the already hampered ASY01 strain and could not be mitigated by addition of ergosterol to the media. Sawai *et al.* however did not report difficulties in their *Δerg4/Δerg5 St7RED* strain. [196]
Even though the production of **4** was unstable under different conditions, the addition of a pYES2-Pp24ISO plasmid to strains ASY02 and ASY03 was tested, to investigate whether conversion towards 24-methyldesmosterol (**3**) could be observed. However, no accumulation of **3** or **4** was detected in those strains.

Multiple attempts of transformation of both *7RED* orthologs did not lead to the desired stable integration and it was concluded that integration of this gene with the techniques used was not a promising approach to generate a 24-methylenecholesterol (**4**) or 24-methyldesmosterol (**3**) producing strain. As the investigation of withanolide biosynthesis candidate genes requires the expression of at least four plant genes (*7RED, 24ISO* and candidate + CPR) in yeast, plasmid-based expression does not seem feasible. The creation of such a strain requires elaborate metabolic engineering, potentially further modifying terpene biosynthesis to provide a sufficient flow of ergosta-5,7,24(28)-trien-3ß-ol (**85**). As metabolic engineering of yeast is not the focus of this project, the production of **3** in *S. cerevisiae* was consigned to Dr. Karan Malhotra. He will be working on establishing a 24-methyldesmosterol (**3**) producing yeast strain system in this group using CRISPR-Cas9.

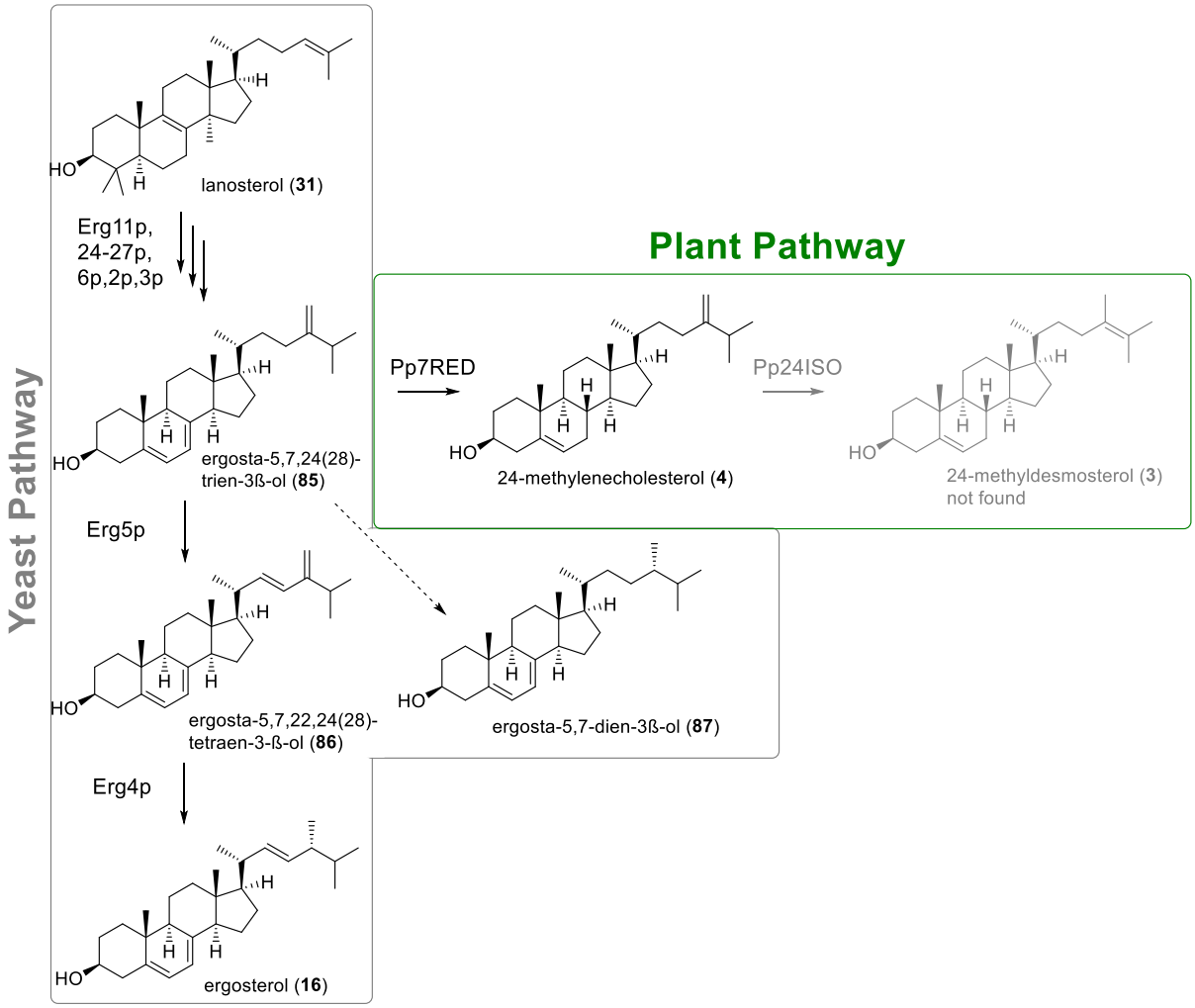
3.1.3 Conclusion, Discussion and Outlook

This part of the project aimed to generate a 24-methyldesmosterol (**3**) producing yeast strain that would enable purification of this substrate or addition of withanolide biosynthesis gene candidates to the yeast strain for *in vivo* assays. Therefore, a similar approach to Sawai *et al*. [196] was taken, as they expressed a St7RED from plasmid in a BY4742 *Δerg4 Δerg5* strain.
For the conversion from 24-methylenecholesterol (**4**) to 24-methyldesmosterol (**3**) an additional gene was needed, the 24ISO. Here, generation of a double knockout (*Δerg4/Δerg5*) strain with subsequent genomic integration of plant genes *7RED* and *24ISO* was planned, as this might increase expression stability compared to plasmid based expression. [203]
A BY4742 *Δerg4* strain was purchased and a further knockout of *ERG5* was performed by integration of a *natMX* cassette. When examining GCMS chromatograms, ergosterol (**16**)

production was lost in the *Δerg4* strain, compared to a *Δpep4* control strain and instead accumulation of ergosta-5,7,22,24(28)-tetraen-3-β-ol (**86**) was observed. Consequently, in the *Δerg4/Δerg5* (ASY01) strain lowered levels of (**86**) and instead accumulation of ergosta-5,7,24(28)-trien-3ß-ol (**85**) was observed. An overview of all sterol-type compounds identified in the yeast strains is depicted in Scheme 3.3 (see also Table 3.1). It was concluded that *ERG5* deletion was successful and insertion of plant genes could be tested.

Insertion of the *7RED* at the *URA3* locus was tested with two orthologs, the *S. tuberosum 7RED* used by Sawai *et al.* [196] and a *P. peruviana 7RED*, as cDNA for cloning was available. After transformation 24-methylenecholesterol (**4**) production was observed, however not under all conditions tested (e.g., different cultivation media). While colony PCRs with insert specific primers were positive, colony PCRs spanning the locus-insert junction could not detect the *7RED* at the *URA3* locus. It was concluded that the *7RED* gene was integrated elsewhere in the genome and exhibited unstable activity. To improve production of **4**, the integration at a different locus could be tested, as transcription levels in different chromatin regions can vary drastically. [222] However, the targeted *URA3* locus is a well-known site for genomic integration with good expression, [223,224] and it is unclear if switching the target site would resolve the unstable genomic integration in this case.

Recently, production of 24-methylenecholesterol (**4**) in *S. cerevisiae* was reported by performing a replacement integration of *7RED* at the *ERG5* locus in a *ERG4* deletion strain. [225] The strain used for production of **4** was a modified WAT11 [226] strain that accumulated higher levels of terpenes. [227] In this study 7RED orthologs from *Oryza sativa* (rice), *Physalis angulata* (related to *P. peruviana*) and *Xenopus laevis* (African clawed toad) were tested. The observed titer of **4** was low but could be increased by adding a second copy of the *Xl7RED* gene upstream of the *ERG4* locus. [225] The authors stated that there is considerable room for improvement of 24-methylenecholesterol (**4**) yield by optimizing precursor flux. [225]



Scheme 3.3 Structures of sterol-type compounds in their biosynthetic context identified in the five *S. cerevisiae* strains described in this work.

It was concluded that accumulation of the precursor ergosta-5,7,24(28)-trien-3ß-ol (**85**) was observed only after deletion of *ERG4* and *ERG5* and that deletions are therefore necessary to create a precursor flow towards **3**. Furthermore, genomic integration of the *7RED* and *24ISO* and therefore the creation of a 24-methyldesmosterol (**3**) producing strain with the methods employed was not productive.

Instead, a CRISPR-Cas9 based approach will be tested by Dr. Karan Malhotra that enables simultaneous metabolic engineering of multiple genes involved in production of **85** and consequently **3**. The *S. cerevisiae* strain planned by Dr. Karan Malhotra will be based on Cas9-containg strain ST7574, derived from CEN.PK113-7D. [228] It was reported that *erg* mutants constructed from BY strains (as in this work) were impaired in the uptake of exogenous amino acids and consequently in cell growth. [229] The use of prototrophic CEN.PK113-7D strain instead of auxotrophic BY strains for *erg* deletions was shown to circumvent the effect of exogenous amino acid supply to cell growth. [230] Precursor flow towards **85** and therefore **3** can

be further engineered by enhancing the yeast mevalonate pathway with 1-2 transformations by CRISPR based techniques, compared to multiple ones as described in this work. Overexpression of mevalonate pathway genes has been successful in various examples to increase production of sterols in yeast. [231–233] Future work will include the overexpression of 9 mevalonate pathway genes [vi] up to squalene biosynthesis.

Preliminary results by Dr. Karan Malhotra suggest that 24-methyldesmosterol (**3**) production by genomic integration of *Pp7RED* and *Pp24ISO* can be observed after deletion of *ERG5* was established. However, titers need to be improved. This is in agreement with results from this thesis, as substantial accumulation of ergosta-5,7,24(28)-trien-3ß-ol (**85**) was only observed once *ERG4* and *ERG5* were deleted. As usable amounts of 24-methyldesmosterol (**3**) were not obtained by production in yeast yet, further experiments will use **3** synthesized by Dave Biedermann.

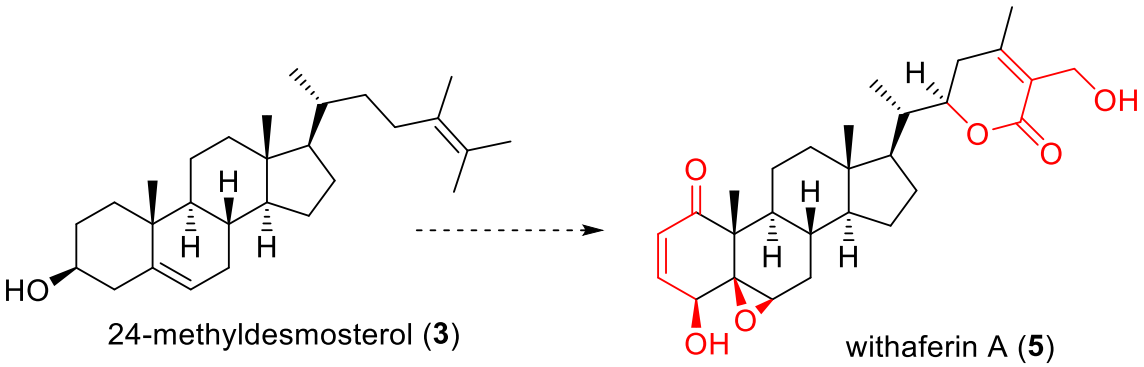
[vi] Overexpressed genes include *ERG8-10*, [234–237] *ERG12-13*, [238,239] *ERG19-20*, [240,241] *IDI1* [242] and *HMG1* [243]

4 Identification of gene candidates for withanolide biosynthesis from *Physalis peruviana* and *Withania somnifera* transcriptome data

4.1 Introduction

This project aims to investigate unknown steps in the withanolide biosynthesis. The last known precursor is 24-methyldesmosterol (**3**). [64] Towards the formation of withanolides such as withaferin A (**5**), multiple oxidations e.g., in the A-ring or the side chain are necessary (see Scheme 4.1). The diverse oxidations in withanolides suggest multiple oxygenases, most likely cytochrome P450s (P450), to be involved in further processing. Another class of oxidative enzymes with similar versatility are α-ketoglutarate dependent dioxygenases. [76] However, little evidence of involvement of those enzymes in triterpene metabolism has been documented yet. [70]

Furthermore, a double bond is present in the A-ring in all commonly occurring withanolides that is not present in the precursor 24-methyldesmosterol (**3**). Therefore, it can be reasoned that an elimination has to occur, most likely with the 3-hydroxy group leaving the A-ring. This reaction could either occur spontaneously once the keto group at C-1 is formed, or it could be performed by a dehydratase. Lastly, some withanolides such as withaferin A (**5**) contain an epoxide group in the B-ring. This reaction is most likely performed by additional P450 enzymes or an epoxidase. However, this particular reaction step will not be further investigated in this work. A summary of the modifications from 24-methyldesmosterol (**3**) towards withaferin A (**5**) is shown in Scheme 4.1 (highlighted in red). For a comprehensive analysis of the proposed biosynthetic route see chapter 1.3.3. This work will concentrate on identifying P450s involved in withanolide biosynthesis and will furthermore approach the elimination reaction in the A-ring.



Scheme 4.1 The withanolide biosynthesis from 24-methyldesmosterol (**3**) to withaferin A (**5**) includes several unresolved reactions (highlighted in red).

A general approach for the elucidation of biosynthetic pathways is the heterologous expression or downregulation of candidate genes to observe accumulation or decrease of products. This work will employ both concepts. An initial screening of gene candidates will be performed with knockdown experiments using virus induced gene silencing (VIGS). Further analysis will be conducted using heterologous expression of candidate genes in *Nicotiana benthamiana* and yeast.

This chapter will describe the identification of gene candidates for withanolide biosynthesis and control genes from *Physalis peruviana*[155] and *Withania somnifera* [156,157] transcriptome data. Table 4.1 shows an overview of transcriptomes used for this analysis that will be further described in the following sections.

Table 4.1 Transcriptome data used in this work.

Organism	**Reference**	**Tissues**	**Sequencing parameters**	**Total reads**	**Assembly**
Physalis peruviana	Fukushima *et al.* [155]	leaves	Illumina, 100 bp paired-end reads	11,611,403	in publication; Trinity
Withania somnifera	Gupta *et al.* [156]	leaf and root from 3 chemotypes	454 pyrosequencing	2,713,592	by Prof. Franke; Trinity and Oases
Withania somnifera	Senthil *et al.* [157]	leaf and root	Illumina, 100 bp paired-end reads	249,036,060	by Prof. Franke; Trinity and Oases

4.2 Identifying cytochrome P450 candidates for withanolide biosynthesis

The main interest in this work is the identification and characterization of enzymes responsible for the biosynthesis of withanolides. Due to their involvement in triterpene pathways in general, [56,70,78] and their ability to perform complex, sequential oxidation reactions, [98] cytochrome P450 enzymes are likely involved in key oxidation steps in withanolide biosynthesis (e.g., lactone ring formation, chapter 1.3.3).

As estimations calculate that P450s constitute up to 1% of the total genes of each plant species, [81] identifying gene candidates of this enzyme family for the pathway of interest is a challenging task. As sequence variability is very high in P450 enzymes, aa identities of more than 40% are considered part of the same cytochrome P450 family and often exhibit similar function.[79,83]

The transcriptomes of two commonly available withanolide producers, the nightshade plants *Physalis peruviana* and *Withania somnifera* were investigated for P450 genes presumably involved in withanolide biosynthesis. The results are described in the following section.

4.2.1 Identification of *Physalis peruviana* P450 gene candidates for withanolide biosynthesis

Initial bioinformatic analysis of potential P450 candidates involved in the withanolide pathway was conducted in the common producer *Physalis peruviana.* A first search for P450 gene candidates was conducted by Prof. Dr. Jakob Franke and supporting analysis was conducted in this work.
The raw reads of *W. somnifera* transcriptomes (pyrosequencing) [156] were assembled by Prof. Franke using Trinity and Oases. [244,245] The assembled *Physalis peruviana* and *Physalis alkekengi* transcriptomes were taken from the regarding publication (Illumina sequencing, 100 bp paired-end reads). [155]

When searching the annotations provided in the publication for the *Physalis peruviana* transcriptome [155] for terms related to "cytochrome P450" (e.g. β-amyrin oxidase, monooxygenase), 200 hits were obtained, highlighting that this enzyme family is strongly represented in this plant. However, this number is far lower than expected, suggesting that either transcriptome data or BLASTx annotations are incomplete. Even though VIGS allows for larger batches of candidate genes, this field had to be narrowed down to enable experimental analysis. Therefore, several indicators for involvement in withanolide biosynthesis were set to generate a candidate pool. For the following analysis only transcripts with an open reading frame (ORF) length of ca. 1.5 kb (±150 bp) were considered, as P450 enzymes typically display a sequence length of ca. 500 amino acids (aa) and shorter or longer ORFs can be considered an assembly artefact.

One indicator for potential involvement in withanolide biosynthesis was the gene expression in leaves, as this is where most withanolides are localized. A study in *Withania somnifera* showed elevated gene expression of phytosterol pathway genes together with increased concentrations of withanolides in the leaves compared to the roots. [246] However, the obtained *Physalis* transcriptome data did not contain information about expression levels assorted to different plant tissues. Therefore, Prof. Dr. Jakob Franke filtered the expression pattern of *Withania*

P450s for transcripts with expression in leaves using a heatmap. Afterwards the closest *Physalis* homologs were determined.

As genomic data from plants is often not available and plant biosynthetic genes are rarely clustered, a tool that is commonly used for identification of genes involved in pathways of interest is the creation of co-expression networks. [247–249] Typically, genes involved in a certain pathway can be used as a guide, to then identify genes with a similar expression pattern that can then be tested for involvement in the respective pathway. [250] Here, co-expression of P450 gene candidates with known phytosterol pathway genes was considered as an indicator, as the early substrates of both withanolides and phytosterols are built by the same pathway (see chapter 1.3.2). Co-expression however, cannot be considered as strictly necessary, since withanolide biosynthesis genes could be expressed independent of phytosterol biosynthesis and furthermore metabolites can be transported within the different tissues of the plant. [251] Furthermore, observation of similar expression patterns does not necessarily imply a regulatory relationship. [252,250]

A co-expression analysis with self-organizing maps (SOM) [253] was conducted by Prof. Dr. Jakob Franke. Transcripts from the *Withania* transcriptome by Gupta *et al*. [156] were grouped by expression pattern. Afterwards 9 phytosterol biosynthesis genes were selected and P450s with similar expression pattern were identified. A co-expression to the selected phytosterol genes was observed for 12 P450s in *Withania*. Figure 4.1 gives a two-dimensional representation of the three-dimensional SOM matrix. The nodes represent genes with a similar expression pattern, with green nodes representing genes with similar expression pattern to neighboring green clusters, but different expression pattern to adjoining white expression clusters. The small red circles show the phytosterol reference genes. 12 cytochrome P450s (stars) are contained within the same nodes as one of the reference genes and therefore exhibit a similar expression pattern.

The *Withania* candidates obtained from the co-expression analysis were screened for orthologs in both *Physalis* species. The presence of homologous sequences in both producers is not strictly specific to withanolides, as orthologous P450s can be involved in other biosynthetic pathways. However, it is an indicator for the quality of the assembly, as identified P450s are not an assembly artefact. A list with 16 *P. peruviana* P450 candidates was provided by Prof. Dr. Jakob Franke.

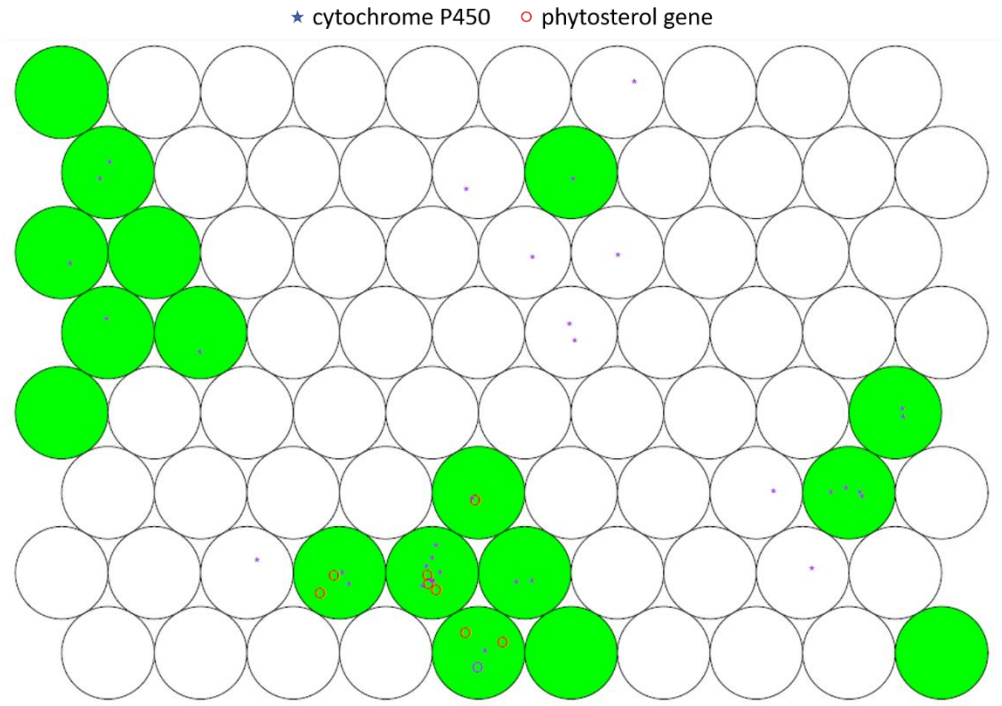


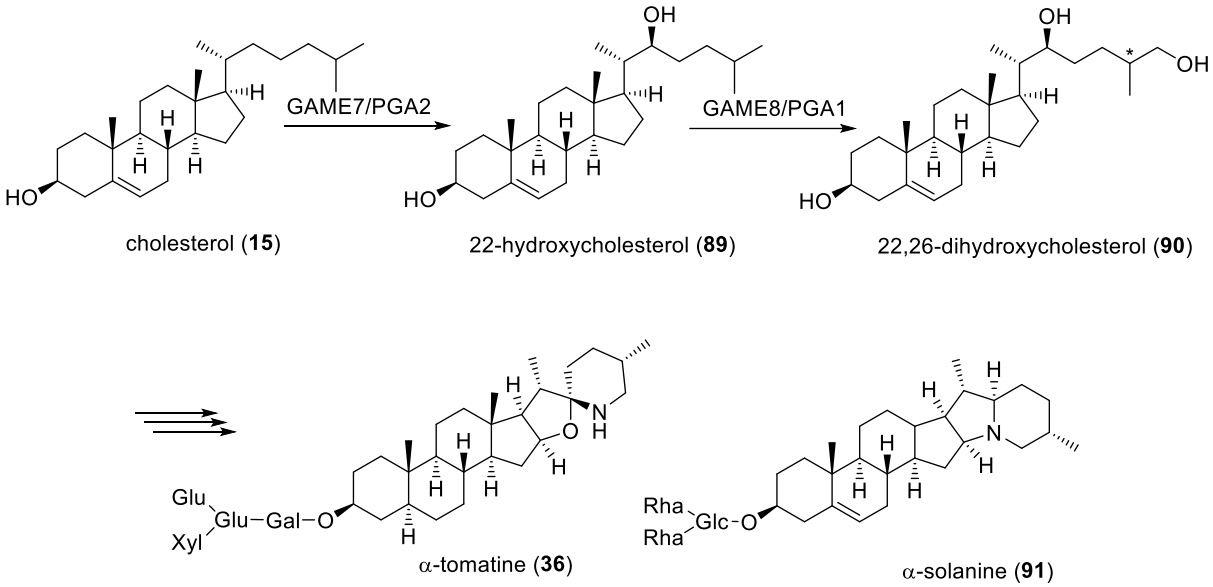
Figure 4.1 An analysis with self-organizing maps showed a similar expression pattern for 12 P450 transcripts (star) to phytosterol reference genes (red circle).

The provided list of 16 candidates was further revised for additional criteria. In phytosterol biosynthesis and related triterpene pathways, multiple P450s were reported. [56,254–256] Consequently, P450s involved in primary metabolism have to be conserved in the plants targeted for this analysis. Therefore, the candidate pool should be revised for P450s with known function. Transcripts with moderate aa identity (30-65%) however, could have evolved to act on withanolide precursors instead of phytosterol intermediates. A tBLASTn[vii] analysis for both *Withania* and *Physalis* transcriptomes was performed, using 39 known P450 genes from plant triterpene biosynthesis as a template (TriForC database). [257] Since aa identity >80% were most likely the respective ortholog to the reference P450, hits for those transcripts were excluded. In *P. peruviana*, two transcripts showed aa identities past that threshold. C33266_g1_i1 exhibited 89% aa identity to a *Solanum lycopersicum* sterol 22-desaturase (AB223043), while c33058_g1_i1 matched multiple obtusifoliol-14-demethylase genes (AF116915, AY091203, AY552551) with 82%-95% aa identity. Both sequences were not included in the candidate list. Most candidates from the candidate pool showed aa identities in the 30-65% range to multiple of the reference genes. One additional transcript caught the attention, as it was not yet included in the list. *P. peruviana* transcript c12144_g1_i2 showed 40-56% aa identity with a query

[vii] due to high sequence variability of cytochrome P450s, BLAST analyses were performed with amino acid query sequence

coverage of 95-98% for 7 of the reference genes (AB223041, AB558145, AB558146, AB558153, DQ335782, LC143440, XM_013608494). The annotations for the matched reference genes contained different functions in varying triterpene pathways, suggesting that a moderately similar enzyme might be involved in withanolide biosynthesis. Consequently, c12144_g1_i2 was included in the candidate pool.

Another group of compounds derived from steroid biosynthesis are steroidal glycoalkaloids (SGAs) in various Solanaceae plants, for example tomato (*Solanum lycopersicum*) or potato (*Solanum tuberosum*). SGA biosynthesis involves two hydroxylations in the cholesterol (**15**) side chain performed by cytochrome P450 monooxygenases, giving 22-hydroxycholesterol (**89**) and 22,26-dihydroxycholesterol (**90**, Scheme 4.2; *S. lycopersicum*: GAME7 and GAME8; *S. tuberosum*: PGA1 and PGA2). [70,258]



Scheme 4.2 Similar as hypothesized for withanolide formation, the biosynthesis of steroidal glycoalkaloids (e.g. α-tomatine (**36**) in tomato and α-solanine (**91**) in potato) involves subsequential hydroxylations of the cholesterol (**15**) side chain by two cytochrome P450 monooxygenases. [70,258]

Since cholesterol (**15**) is structurally close to 24-methyldesmosterol (**3**), the last known withanolide precursor, and this pathway is observed in several nightshade plants, it can be reasoned that a side chain hydroxylase in withanolide producers shares a common ancestor with the SGA pathway. Therefore, the candidate pool was investigated for transcripts similar sequences to the cholesterol side chain oxidases.

A tBLASTn search was conducted with peptide sequences of *Solanum lycopersicum* GAME7 (Solyc07g062520) and GAME8 (SGN-U578058) and *Solanum tuberosum* PGA1 (AB839752)

and PGA2 (AB839753). In the *P. peruviana* candidate pool, 4 transcripts showed aa identities ranging from 36% to 55% to both side chain hydroxylases, however aa identity observed was generally closer to *S. lycopersicum* hydroxylases than to the ones from *Solanum tuberosum* (see Table 4.2).

Table 4.2 Amino acid identities (%) of SGA side chain hydroxylases and *P. peruviana* candidate P450s.

	c17586	c12144	c17629	c8953	GAME7	PGA2	PGA1	GAME8
c17586_g1_i1		36.2	37	40.1	37.0	36.4	35.7	35.5
c12144_g1_i2	36.2		57.2	64.0	60.5	48	44.7	44.9
c17629_g1_i1	37	57.2		72.4	65.8	45.9	46.1	45.9
c8953_g1_i2	40.1	64.0	72.4		85.2	48.2	48.3	49.4
GAME7	37.0	60.5	65.8	85.2		48.5	44.6	45.3
PGA2	36.4	48	45.9	48.2	48.5		51.6	52.1
PGA1	35.7	44.7	46.1	48.3	44.6	51.6		96.6
GAME8	35.5	44.9	45.9	49.4	45.3	52.1	96.6	

In total 17 *Physalis peruviana* P450 gene candidates were determined. The transcript identifiers (ID) of the Fukushima transcriptome data [155] of the gene candidates are listed in Table 4.3.

Table 4.3 *Physalis peruviana* cytochrome P450 gene candidates for withanolide biosynthesis.

P. peruviana ID	**ORF length [bp]**	***W. somnifera* expression pattern**	**Co-expression phytosterol genes**	**Hit TriForC**	**Homology SGA hydroxylases**
c22743_g1_i1	1,449	leaves	✓	✓	
c12180_g1_i2	1,491	leaves	✓		
c12127_g1_i2	1,518	leaves			
c17629_g1_i1	1,560	leaves	✓	✓	✓
c5657_g1_i2	1,449	leaves	✓	✓	
c18183_g1_i1	1,608	leaves			
c29050_g1_i1	1,536	leaves			
c12321_g2_i1	1,533	leaves	✓		
c10311_g2_i1	1,521	leaves			
c903_g1_i1	1,569	leaves			
c23936_g1_i1	1,560	leaves	✓		
c932_g1_i1	1,665	leaves			
c17586_g1_i1	1,536	leaves	✓	✓	✓
c12144_g1_i2	1,581	leaves		✓	✓
c12233_g3_i1	1,353	leaves	✓		
c8953_g1_i2	1,053*	leaves		✓	✓
c8960_g2_i1	1,278*	leaves	✓		

*ORF length did not conform with cutoff length, but candidates were still selected due to other matching criteria

4.2.2 Selecting *Withania somnifera* cytochrome P450 candidates

Preliminary experiments for VIGS in *Physalis peruviana* concluded that silencing in this plant was difficult to achieve, while *Withania somnifera* exhibited high susceptibility to this technique (see chapter 5.2). Therefore, further experiments should be conducted in *W. somnifera*, and *P. peruviana* cytochrome P450 candidates determined in the last subsection should be identified in the new target organism. Furthermore, additional analyses should expand the candidate pool here.

The initial search for *P. peruviana* P450 candidates was supported with *W. somnifera* transcriptome data by Gupta *et al.*[156] Here, different tissues of several chemotypes were analyzed. In total, 2,713,592 reads were obtained by pyrosequencing, divided onto different tissues and chemotypes. A second publication by Senthil *et al.* [157] used paired-end Illumina sequencing and obtained 135,186,223 (root) and 113,849,837 (leaf) reads after removal of low-quality reads. Since the transcriptome data by Senthil *et al.* was more comprehensive and the sequencing technique used is of higher reliability, further analysis was conducted with this transcriptome data. The reads were assembled with Trinity [244] in the publication, however only raw reads were accessible. Therefore assembly with Trinity was repeated and furthermore an additional assembly was conducted by Prof. Dr. Jakob Franke using Oases. [245]

Since all criteria that were mentioned for the determination of *P. peruviana* P450 candidates were valid in *W. somnifera* as well, orthologs of the previously selected *P. peruviana* candidates were identified from the Senthil Trinity data using BLASTn.

Furthermore a second co-expression analysis was performed by Prof. Dr. Jakob Franke with CoExpNetViz. [259] Here four bait genes from phytosterol/withanolide biosynthesis were analyzed against Gupta *Withania* transcriptome data (self-assembled). [156] This analysis yielded 3 transcripts with the expected ORF size of ca. 1.5 kb that were not yet part of the candidate pool.

In addition to the preceding analysis the annotations (BLASTx, BLASTp, pfam) of the Trinity assembly were assessed (key words: cytochrome P450, beta-amyrin oxidase, monooxygenase, hydroxylase) and compared to the pool of gene candidates. Four additional transcripts with suiting annotations were identified and added to the candidate pool.

A tBLASTn search with cholesterol side chain hydroxylases GAME7/PGA2 and GAME8/PGA1 (see chapter 4.2.1) against the *Withania* transcriptome data from Senthil *et*

al. [157] showed amino acid identities ranging from 53% to 74% for four of the selected candidates. Similar to *P. peruviana*, tomato SGA hydroxylases showed higher identities to the candidates than SGA hydroxylases from potato, with one exception: candidate transcript TRINITY_DN4953_c0_g1_i1 showed an aa identity of 90.3% with a coverage of 100% to PGA2 (see Table 4.4).

Table 4.4 Amino acid identities (%) of SGA side chain hydroxylases and *W. somnifera* candidate P450s.

	GAME7	GAME8	PGA1	PGA2	DN3273	DN13186	DN751	DN4953
GAME7		45.7	44.7	48.0	78.9	64.4	62.4	47.7
GAME8	45.7		96.6	52.9	50.6	45.0	45.2	53.6
PGA1	44.7	96.6		52.1	50.2	45.0	45.0	52.8
PGA2	48.0	52.9	52.1		53	45.1	44.1	90.3
DN3273_c0_g7_i1	78.9	50.6	50.2	53		72.5	68.8	52.2
DN13186_c2_g1_i1	64.4	45.0	45.0	45.1	72.5		81.0	46.4
DN751_c0_g1_i8	62.4	45.2	45.0	44.1	68.8	81.0		44.5
DN4953_c0_g1_i1	47.7	53.6	52.8	90.3	52.2	46.4	44.5	

Lastly, all transcripts that did not contain an ORF of ca. 1.5 kb in either the Trinity or Oasis assembly were excluded as those are potentially artefacts of the assembly. In summary, candidates that matched to at least one of the following criteria were selected for VIGS experiments in *Withania*: homology to previously selected *Physalis peruviana* candidates, co-expression to known triterpene pathway genes, annotations (BLASTx, BLASTp, pfam) suggesting activity on withanolide-like structures or homology to SGA hydroxylases. This resulted in a total of 21 candidates that are shown in Table 4.5.

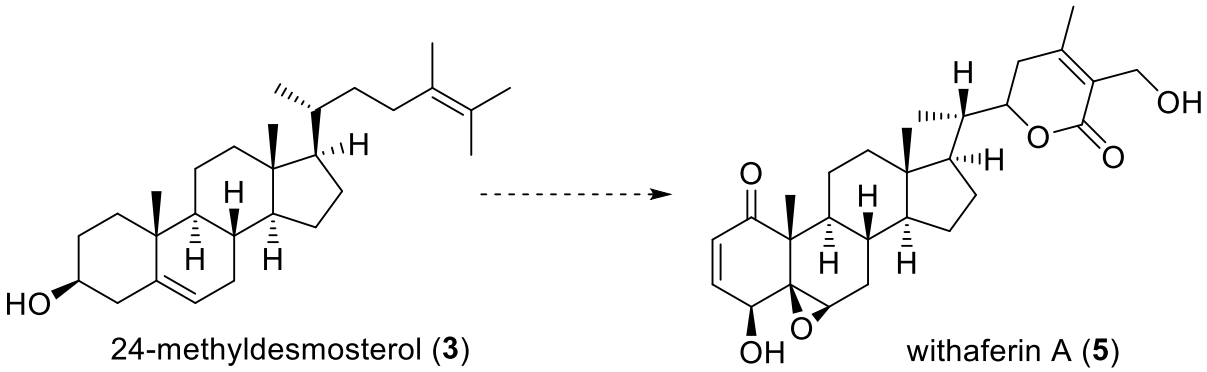
Table 4.5 Candidate genes chosen for VIGS experiments in *Withania somnifera*.

Abbreviation	*W. somnifera* Senthil ID	*P. peruviana* ID	Homology to Pp candidates	CoExpNetViz Hit	Selection from annotation	SGA hydroxylase homology
P450-1	TRINITY_DN5012_c0_g2_i5	c10311_g2_i1	✓			
P450-2	TRINITY_DN3361_c0_g1_i8	c5657_g1_i2	✓	✓		
P450-3*	TRINITY_DN730_c0_g2_i2	c8960_g2_i1	✓	✓		
P450-4	TRINITY_DN730_c0_g1_i6	c22743_g1_i1	✓		✓	
P450-5	TRINITY_DN751_c0_g1_i8	c17629_g1_i1	✓	✓		✓
P450-6	TRINITY_DN26488_c0_g1_i3	c12321_g2_i1	✓			
P450-7	TRINITY_DN16059_c0_g2_i2	c6704_g1_i2			✓	
P450-8	TRINITY_DN4953_c0_g1_i1	c17586_g1_i1			✓	✓
P450-9	TRINITY_DN2831_c0_g1_i3	c22743_g1_i1	✓	✓		
P450-10	TRINITY_DN2147_c0_g1_i7	c17586_g1_i1	✓	✓		
P450-11	TRINITY_DN2102_c0_g1_i19	c12180_g1_i1	✓	✓		
P450-12	TRINITY_DN1039_c0_g1_i9	c33058_g1_i1			✓	
P450-13*	TRINITY_DN3273_c0_g7_i1	c17629_g1_i1	✓	✓		✓
P450-14	TRINITY_DN21326_c0_g3_i3	c5136_g1_i1		✓		
P450-15	TRINITY_DN18544_c0_g1_i3	c23936_g1_i1		✓		
P450-16	TRINITY_DN13951_c0_g1_i2	c2263_g1_i1		✓		
P450-17	TRINITY_DN16765_c1_g3_i1	c932_g1_i1	✓			
P450-18	TRINITY_DN14593_c0_g1_i1	c903_g1_i1	✓			
P450-19*	TRINITY_DN13186_c2_g1_i1	c17629_g1_i1	✓		✓	✓
P450-20	TRINITY_DN9661_c0_g1_i57	c23936_g1_i1	✓			
P450-21	TRINITY_DN406_c0_g1_i16	c29050_g1_i1	✓			

* ORFs for analysis taken from Oases Assembly, as Trinity assembly showed truncated ORF for this transcript

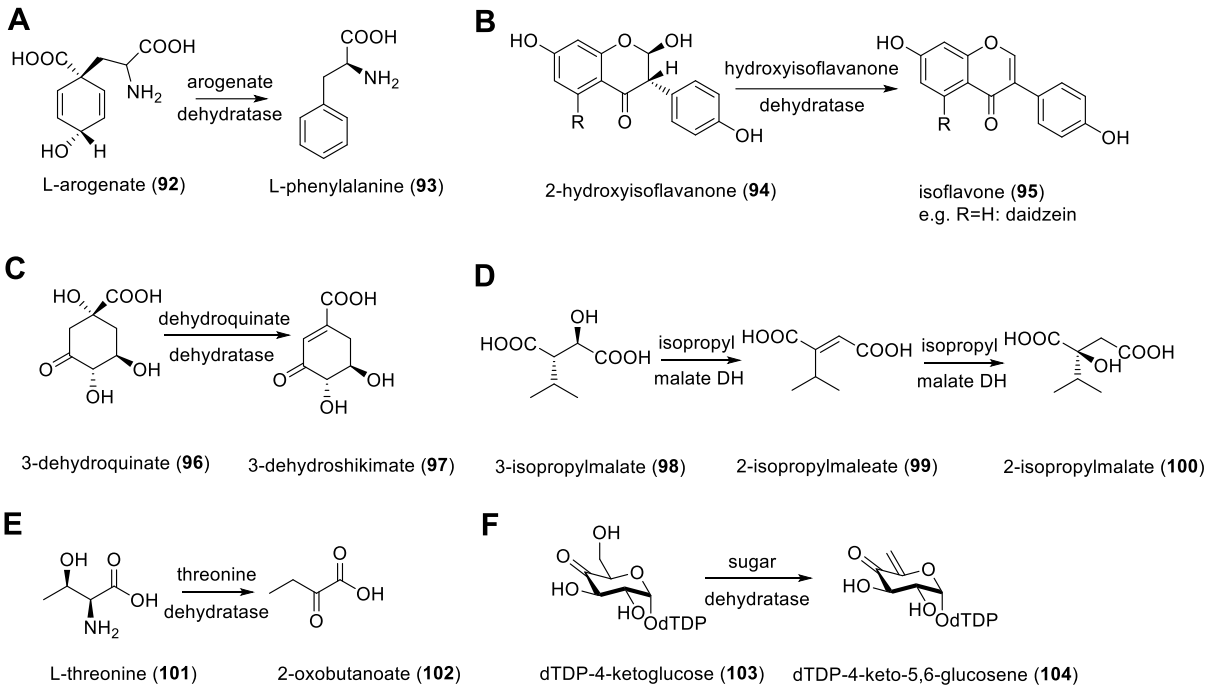
4.3 Selecting *Withania somnifera* dehydratase candidates

Most withanolides possess a double bond in the A-ring. When examining the last known precursor, 24-methyldesmosterol (**3**), the double bond is not present (see Scheme 4.3). Instead, **3** possesses a 3-hydroxy group that is derived from cyclization of 2,3-oxidosqualene (**29**) in the sterol/triterpene biosynthetic pathway (see chapter 1.3.2). Therefore, it can be reasoned that an elimination has to occur with the 3-hydroxy group leaving the A-ring. This reaction is most likely performed by a dehydratase.



Scheme 4.3 Withaferin A (**5**) contains a double bond in the A-ring that is not present in the precursor.

In order to determine candidate genes for the dehydratase reaction, *W. somnifera* transcriptome data from Senthil *et al.*[157] was searched for BLASTp annotations containing dehydratase function. In total, 6 different classes of dehydratases can be found in the transcriptome data: arogenate dehydratase (**92** to **93**, Scheme 4.4A),[260] hydroxyisoflavanone dehydratase (**94** to **95**, B),[261] dehydroquinate dehydratase (**96** to **97**, C),[262] isopropyl malate dehydratase (**98** to **99** to **100**, D),[263] threonine dehydratase (**101** to **102**, E),[264] sugar dehydratase (**103** to **104**, F).[265] For the typical reactions catalyzed by the respective enzyme class see Scheme 4.4. Of the aforementioned annotation classes only the first 3 classes (A-C, Scheme 4.4) contain ring structures that are roughly resembling withanolides and substrates exhibited similar polarity, suggesting that an evolution of those enzymes towards withanolide biosynthesis would seem plausible. Together, transcripts with annotations for arogenate dehydratase, hydroxyisoflavanone dehydratase and dehydroquinate dehydratase yielded 16 candidates for the elimination reaction.



Scheme 4.4 Reactions catalyzed by dehydratase classes present in *W. somnifera* transcriptome data.[157] Arogenate dehydratase (A),[260] hydroxyisoflavanone dehydratase (B),[261] dehydroquinate dehydratase (C),[262] isopropyl malate dehydratase (D),[263] threonine dehydratase (E)[264] and sugar dehydratase (F).[265]

Furthermore, the ORF size of each DH candidate was compared to the CDS length that is typical for its respective annotation. For reference see Table 4.6.

Table 4.6 Typical CDS length for classes of reference dehydratases

Dehydratase	**Accession**	**CDS length [bp]**	**Organism**
2-hydroxyisoflavanone dehydratase	Q5NUF3	960	*Glycine max* [261]
arogenate dehydratase	Q9SA96	1,176	*Arabidopsis thaliana* [260]
3-dehydroquinate dehydratase	Q9SQT8	1,812	*Arabidopsis thaliana* [262]

This eliminated two candidates, as their ORFs with less than 600 bp were too short. However, one of the candidates eliminated was a truncated sequence of another candidate (DH3), suggesting that this could be an artefact of transcriptome assembly. As the extent of the dehydratase gene candidate pool was rather small and did not need to be narrowed down, further analysis, e.g. co-expression analysis, was not conducted. In summary, 14 dehydratase candidates for the elimination reaction in the A-ring were selected for further evaluation (Table 4.7).

Table 4.7 Dehydratase candidate genes for VIGS experiments

Abbreviation	Senthil Trinity ID	ORF [bp]	Annotation
DH 1	TRINITY_DN9817_c0_g1_i6	981	2-hydroxyisoflavanone dehydratase
DH 2	TRINITY_DN980_c0_g1_i2	1,290	arogenate dehydratase [A]
DH 3	TRINITY_DN22085_c0_g1_i50	1,551	3-dehydroquinate dehydratase [B]
DH 4	TRINITY_DN14466_c0_g1_i32	1,557	3-dehydroquinate dehydratase [B]
DH 5	TRINITY_DN6879_c0_g1_i1	957	2-hydroxyisoflavanone dehydratase
DH 6	TRINITY_DN2092_c0_g1_i5	1,191	arogenate dehydratase [A]
DH 7	TRINITY_DN980_c0_g1_i8	1,290	arogenate dehydratase [A]
DH 8	TRINITY_DN980_c0_g1_i21	1,380	arogenate dehydratase [A]
DH 9	TRINITY_DN3290_c0_g1_i1	954	2-hydroxyisoflavanone dehydratase
DH 10	TRINITY_DN59359_c0_g1_i1	948	2-hydroxyisoflavanone dehydratase
DH 11	TRINITY_DN9502_c0_g1_i1	987	2-hydroxyisoflavanone dehydratase
DH 12	TRINITY_DN5853_c0_g1_i2	963	2-hydroxyisoflavanone dehydratase
DH 13	TRINITY_DN5853_c0_g1_i1	939	2-hydroxyisoflavanone dehydratase
DH 14	TRINITY_DN2092_c0_g1_i24	1,191	arogenate dehydratase [A]

A: full annotation reads “arogenate dehydratase/prephenate dehydratase 6, chloroplastic”; B: full annotation reads “bifunctional 3-dehydroquinate dehydratase/shikimate dehydrogenase, chloroplastic”.

4.4 Conclusion

In summary, the transcriptome data of withanolide producers *Physalis peruviana* and *Withania somnifera* was assessed to determine candidates for the oxidation and elimination reactions performed in the late withanolide biosynthesis.

Due to their involvement in the triterpene and phytosterol pathway, [56,70,78] enzymes from the cytochrome P450 superfamily seem as the most plausible candidates to perform sequential oxidations in withanolide biosynthesis. P450 gene candidates were selected according to the following criteria: orthologs in both producing plants, expression pattern, co-expression to genes involved in phytosterol biosynthesis, homology to known phytosterol P450s or annotations that function on structures similar to withanolides. In total, 21 P450 candidates were selected for further silencing experiments in *Withania somnifera*.

For the identification of dehydratase gene candidates, *W. somnifera* transcriptome data was assessed for annotations that would be able to perform elimination reactions on structures similar to withanolides. Here, 14 candidates were selected. The determined gene candidates will be further evaluated by silencing and heterologous expression experiments described in the following chapters.

5 Screening of withanolide biosynthesis candidate genes with virus induced gene silencing

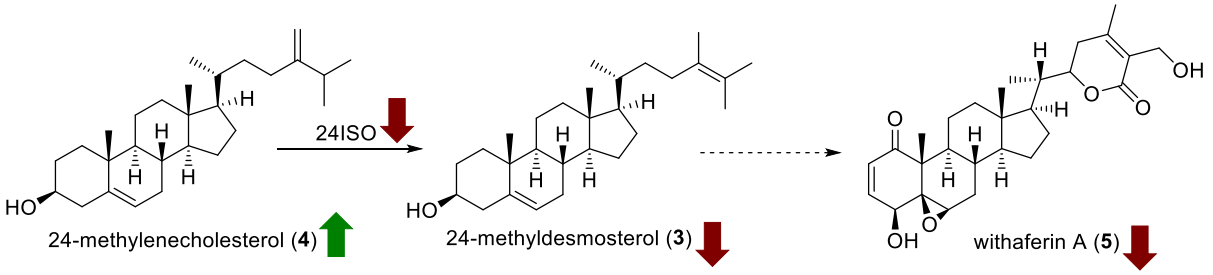
5.1 Introduction

When investigating natural product biosynthesis, a general approach is the up- or downregulation of candidate enzymes to observe changes in product accumulation. An example for the latter is virus induced gene silencing (VIGS), a technique that can be used in plants which exploits RNA-mediated antiviral defense mechanisms.

For this purpose, small fragments of the respective target gene, in this case putative withanolide biosynthetic genes, are cloned into the pTRV2 vector that is based on the tobacco rattle virus.[142] Together with the pTRV1 vector, carrying the genes for virulence, pTRV2 - gene of interest (GOI) constructs can be infiltrated into the plant using *Agrobacterium* mediated gene transfer. Expression of the small fragments in the plant then results in mRNA silencing. [127,130] An advantage of this method is that silencing effects on biosynthesis can be monitored without knowing the intermediates, as the decrease of main product levels can function as an indicator for silencing. The identification of accumulated intermediates might furthermore support or abolish the hypothesis of a biosynthetic route. For a detailed description of mechanisms involved see the chapter 1.3.3.

It has to be noted that VIGS is not a knockout on genome level, but rather a transcriptional downregulation and therefore product formation, albeit inhibited, can still be observed. This broadens the field of applications to essential genes, where a knockout would be lethal to the plant. Furthermore, silencing experiments with VIGS can be achieved in a relatively short time frame. Stable genetic transformations however are challenging to accomplish for non-model plant species. [125]

When performing silencing of genes that are responsible for the formation of key structures in withanolide biosynthesis (e.g. the lactone ring), this would result in a decrease of main withanolides of the respective plant. Silencing of the *Withania somnifera 24-isomerase* was conducted by Knoch *et al.* [64] They observed a decrease of withaferin A (**5**) and 24-methyldesmosterol (**3**) in addition to accumulation of 24-methylenecholesterol (**4**), the substrate of the 24ISO (Scheme 5.1).



Scheme 5.1 Knockdown of 24ISO results in a decrease of withanolide levels (red arrows) and consequently in an increase of the precursor 24-methylenecholesterol (**4**).

The application of virus induced gene silencing has been well studied. [125,126,266] However, achieving silencing in non-model plants is often a difficult task that requires conditions such as plant cultivation and agroinfection geared to the respective plant species.

In the case of *Physalis peruviana*, one example for the application of VIGS has been reported, [149] but multiple examples are present for its relative *Physalis floridana*.[148,267,268] For *Withania somnifera* VIGS of withanolide biosynthesis related genes has been reported before.[64,73,146,147] However, silencing attempts in *W. somnifera* have focused on glycosyltransferases, responsible for attaching sugar moieties to form glycowithanolides, and on genes that were known to be involved in withanolide biosynthesis prior to their branching point from phytosterols. [73,146,147] This work aims to gain insights into transformations that further functionalize withanolides from their last known intermediate 24-methyldesmosterol (**3**).

In order to investigate withanolide biosynthesis, withanolide producers *Physalis peruviana* and *Withania somnifera* were examined for their compatibility with VIGS experiments. Chapter 5.2 describes the conditions determined in this work for effective silencing, using a visual marker, the *phytoene desaturase* (*PDS*). [269] As VIGS functioned more reliably in *Withania somnifera*, subsequent silencing of candidate cytochrome P450 and dehydratase genes was conducted in this plant (see chapter 5.3.2 and 5.4.2).

5.2 Establishing a virus induced gene silencing system in a withanolide producing plant

5.2.1 Construction of a phytoene desaturase control for VIGS in *Physalis peruviana* and *Withania somnifera*

In order to establish VIGS experiments, a positive control is needed that indicates successful silencing with the tested conditions. A common control is the phytoene desaturase (PDS), an enzyme that catalyzes the dehydrogenation of 15-cis-phytoene in the carotenoid biosynthesis. [269] Since carotenoids are essential in protecting plants from photobleaching, [270] a bleached phenotype can be observed in infected areas. Silencing of the *PDS* has been reported for both *Physalis peruviana* [149] and *Withania somnifera*.[73,271]

In order to establish VIGS for this project, *P. peruviana* and *W. somnifera* PDS sequences needed to be obtained. Information about the respective sequences, however, was either not published or not available at the time of this analysis. Therefore, to identify the PDS transcripts in *P. peruviana* and *W. somnifera* the coding DNA sequence (CDS) of a *Physalis pubescens* PDS (Accession Nr: JX255734) [148] was used as a template for BLASTn[viii] analysis of the *P. peruviana* [155] and *W. somnifera* [157] (Trinity assembly) transcriptomes.

The BLAST analysis revealed one transcript, c12986_g1_i1 as the putative *P. peruviana* PDS with a nucleotide identity of 99.4% and a query coverage of 100%. This transcript will further be referred to as *PpPDS*. For *W. somnifera* the transcript TRINITY_DN5499_c0_g1_i7 was the best hit with an identity of 99.6% and a query coverage of 100%. This transcript will further be termed as *WsPDS*. A DNA sequence identity of >99% in both cases is in consensus with the literature, as the highly conserved PDS typically exhibits >95% identity. [272,273]

VIGS needs a 150-500 bp fragment in order to generate a silencing response. [274] However, it is crucial to select segments that are specific for the gene in order to avoid off-targeting and consequential unpredictable metabolic effects. The siFi21 prediction tool [275,276] compares a target sequence to a custom database, in this case the respective plant transcriptome and detects potential off-targeting sites (red, Figure 5.1). Furthermore, segments with higher probability for RNAi are highlighted (red line, Figure 5.1). An exemplary diagram of the *WsPDS* gene analyzed against *W. somnifera* transcriptome [157] is given in Figure 5.1.

[viii] As phytoene desaturase is a highly conserved enzyme, a nucleotide BLAST was conducted

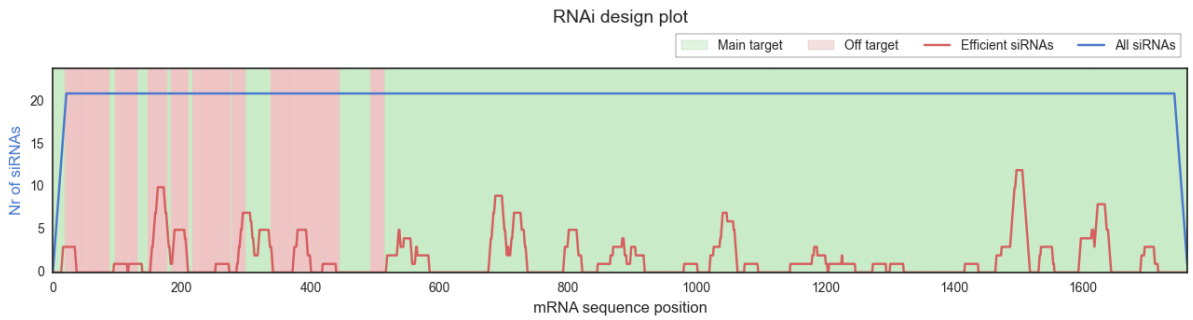


Figure 5.1 An exemplary RNAi design plot of *WsPDS* compared to the *W. somnifera* transcriptome [157] highlights sequence areas with potential off-targeting (red) and higher probability for RNAi (red line). [275,276]

Since PDS silencing of *Physalis peruviana* was described in the literature the target fragment used by Osorio-Guarín *et al.* (440 bp, further termed PDSO) [149] was tested in addition to the site selected with siFi21 for this work (default settings, 347 bp, further termed PDSS).

An analysis for a PDS target site in *Withania somnifera* was conducted with siFi21 and compared to the target site described in the literature. [271] A protocol for *WsPDS* silencing by Bomzan *et al.* suggested a target fragment ranging from 350-830 bp in the ORF. [271] However, the RNAi design tool showed potential off-targeting between 0-500 bp in the ORF. Therefore, the fragment selected for this work was shifted slightly downstream to avoid off targeting (600-930 bp in ORF). The target sites for PpPDS and WsPDS are illustrated in Figure 5.2.

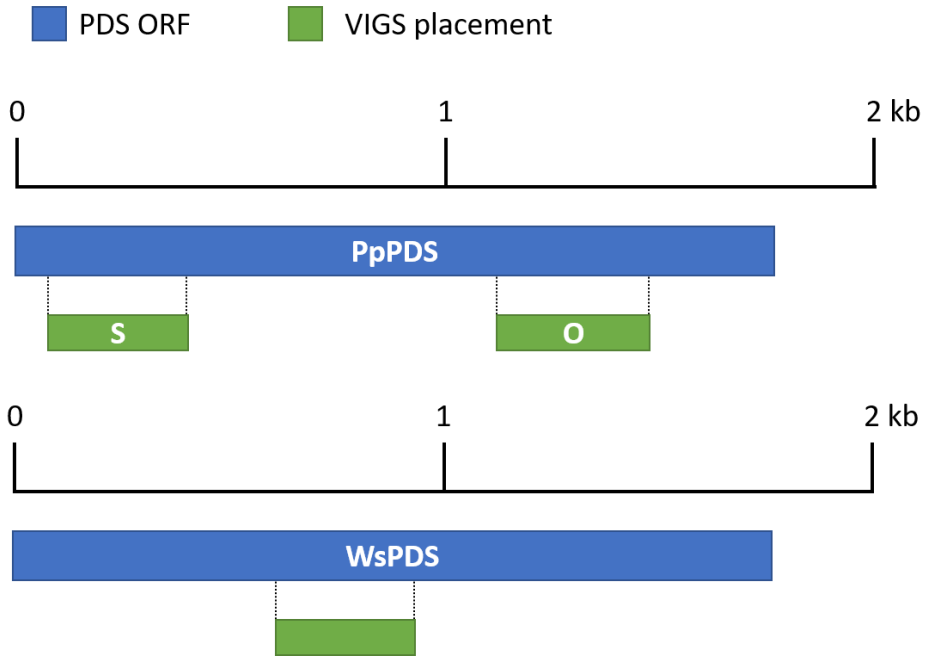


Figure 5.2 Target sites (green) for virus induced gene silencing of *phytoene desaturase* (*PDS*, blue) in *P. peruviana* and *W. somnifera* (S: target site for PpPDS selected in this work; O: target site selected by Osorio-Guarín *et al.*). [149]

For the construction of vectors (pTRV2-GOI) the Gibson assembly technique was applied (In-Fusion®, Takara). As a template for PCR of GOI inserts, cDNA generated from aerial tissue RNA of self-cultivated *W. somnifera* and *P. peruviana* plants was used. The obtained plasmids were confirmed by colony PCR and DNA sequencing (Eurofins, Ebersberg). Subsequently, *A. tumefaciens* strains GV3101, AGL-1, LBA4404 and EHA105 were transformed with pTRV1 and pTRV2-PDS plasmids using electroporation. The resulting transformants were confirmed by colony PCR.

5.2.2 Determination of VIGS conditions for *Physalis peruviana* and *Withania somnifera*

Physalis peruviana

Initial experiments were conducted with *Physalis peruviana* and showed that finding the right set of conditions can be a challenging task. Many factors such as light exposure, OD_{600} of the *A. tumefaciens* strains, stress and temperature can have a big impact on the success of VIGS experiments and are individual for each plant species. [149,277–279]

Furthermore, PDS bleaching has to be distinguished from yellowing of leaves due to stress response. Bleaching due to VIGS spreads systemically and is therefore typically observed in young leaves, exhibiting white coloring of affected areas. [272,280] In contrast, stress response was often observed in leaves used for infiltration due to mechanical stress, exhibiting a yellow coloring, and often resulted in abscission of the affected leaves. Real-time quantitative PCR (qPCR) of such yellow leaves, performed in this work, showed no decrease in PDS expression levels (data not shown). A comparison of leaf coloring of infiltrated *Physalis peruviana* plants is shown in Figure 5.3.

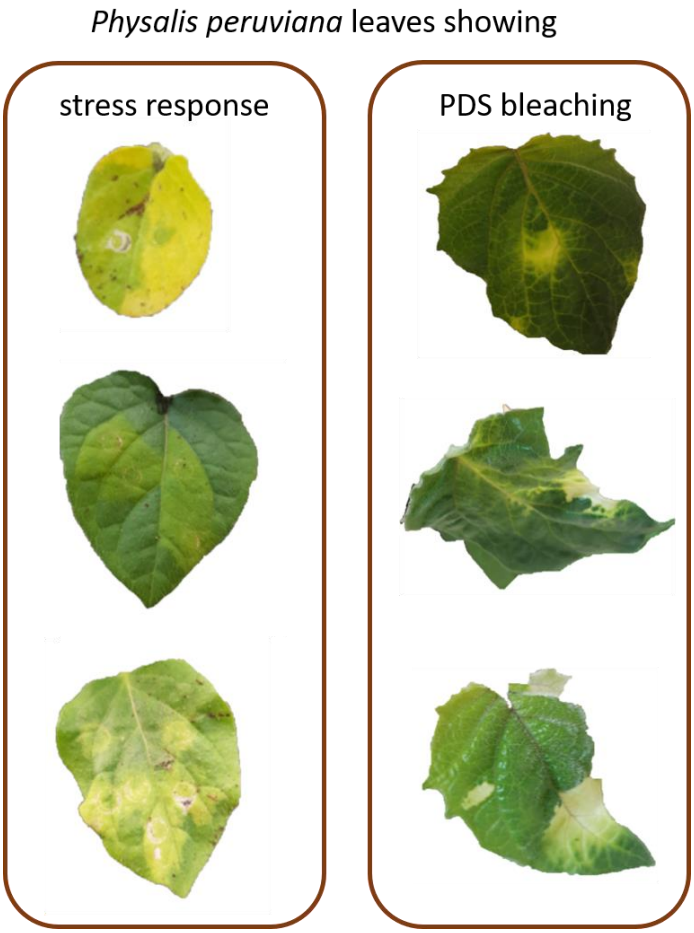


Figure 5.3 Systemic bleaching after PDS infiltration shows a faded white phenotype in young leaves (right) compared to yellowing of infiltrated leaves due to stress response (left) in *P. peruviana*.

Osorio-Guarín *et al.* [149] investigated silencing of a *P. peruviana PDS* using *A. tumefaciens* strain GV3101 at OD_{600} 1.0 and 1.5. They reported that the silencing phenotype spread uniformly across the entire leaf, when using an OD_{600} of 1.0, while the infiltration with OD_{600} of 1.5 resulted in a more mosaic-like phenotype. All plants were cultivated after infection at 25 °C, 16/8 h photoperiod and silencing effects were observed 7 days post infection (dpi) (Table 5.1, condition O).

An experiment with similar conditions, was conducted with both PpPDSO (taken from Osorio-Guarín *et al.*) [149] and PpPDSS (designed in this work) target sequences for OD_{600} of 0.1, 1.0 and 1.5, and plants were monitored for 28 dpi. Both PDS and buffer control samples showed yellowing due to stress response, but no systemic bleaching, suggesting that optimization of silencing conditions was necessary. The process of agroinfection involves both *Agrobacterium*-related and plant-related factors. Multiple factors concerning plant cultivation [148,149,277] are described in the literature and therefore optimization of VIGS in this work concentrated on those conditions first.

Osorio-Guarín *et al.* noted that the duration and intensity of the photoperiod is important for the appearance of the bleached phenotype of PDS silencing. [149] A photoperiod of 12 h light/12 h dark did not support the bleaching, while a 16 h light/8 h dark schedule resulted in fast bleaching. This effect has also been observed for PDS silencing in *Gossypium barbadense*. [277] Zhang *et al*. noticed that incubation in darkness for 48 h after infiltration has positive effects on silencing in *Physalis floridana*. [148] An incubation in darkness is presumed to improve the entry of Agrobacteria into plant cells. [281,282] Another factor influencing the effects of PDS silencing is the temperature. In general it is noted that more intense photobleaching can be observed at temperatures below 25 °C. [278,279]

Therefore, different available plant cultivation conditions post infiltration were tested. All following experiments were conducted for PsPDSO and PpPDSA at OD_{600} 1.0, as this was reported to be most effective in *P. peruviana*. [149] Experiments were conducted for all three cultivation conditions (A-C, Table 5.1) and plants were incubated in darkness for 48 h post infiltration as described in the literature. Furthermore *A. tumefaciens* cultures were regularly tested by colony PCR to exclude the possibility of plasmid loss.

Table 5.1 Conditions tested for *Physalis peruviana* cultivation after agroinfiltration with PsPDSO and PpPDSA at OD_{600} 1.0.

Condition	Temperature [°C]	Humidity [%]	Light intensity [µmol/s]	Photoperiod [h]	Result
O	25	60	not reported in the literature [149]	16/8	no photobleaching
A	21-24	69	500	11/13	no photobleaching
B	30	60	221	16/8	no photobleaching
C	22	not controlled	150	16/8	no photobleaching

A change of plant growth conditions while maintaining agroinfiltration conditions did not result in the desired photobleaching. Instead yellowing in infiltrated leaves as a part of stress response was observed in all plant cultivation locations. Therefore, the process of agroinfection was targeted for optimization. All following experiments were conducted for plant growth conditions A-C with subsequent cultivation in the dark (48 h) after infiltration.

At first, the mechanical barrier that was faced during *P. peruviana* agroinfiltration was addressed. *Physalis peruviana* leaves are very sturdy, so that it was technically difficult to infiltrate larger volumes of *Agrobacterium* suspension into the leaf. Therefore, addition of surfactants to the infiltration buffer was tested (0.2% (v/v) Tween 20), as literature reported an

increase in silencing efficiency due to decreased surface tension in the suspension and improved cell attachment for bacterial invasion. [283–286] However, the addition of surfactants did not lead to a facilitated infiltration or a PDS-silenced phenotype.

Second, the infecting *A. tumefaciens* strain was addressed, as various strains are available that exhibit different properties regarding their pathogeny. All previous experiments used *A. tumefaciens* strain GV3101 as this strain is very efficient for infection of Solanaceae [287] and both *W. somnifera* and *P. peruviana* PDS silencing have been reported with this strain. [149,271] In order to increase infection efficiency in *P. peruviana*, three additional strains were tested.[ix]

The *A. tumefaciens* strain AGL-1 is recombinase deficient, which facilitates infection with larger plasmids such as pTRV2-GOI, [288] while LBA4404 is widely used and many examples for application in Solanaceae can be found. [289] Lastly, EHA105 shows high virulence and has been reported to transfer T-DNA more efficiently than GV3101 in Solanaceae. [290,287] All plants were observed for 30 dpi, however no bleaching apart from yellowing due to stress response was observed.

Lastly, an additional infection method was tested. Previous experiments employed the commonly used leaf infiltration technique (see Figure 5.4). In addition to leaf infiltration, the stem of *P. peruviana* plants was injured using a canula and *Agrobacterium* suspension was injected in multiple spots in the stem. It was reasoned that by injuring the stem, virus would occasionally reach the phloem and would therefore lead to a more systemic spread of silencing effects. This technique was previously reported to improve systemic spread of VIGS effects. [291,272]

[ix] *Agrobacterium* strains were kindly provided by the group of Prof. Dr. Witte, Institute of Plant Nutrition, Leibniz University Hanover

Figure 5.4 Leaf infiltration technique demonstrated on *Nicotiana benthamiana* plants.[x]

Agroinfiltration was conducted with both EHA105 and GV3101 strains, as those were the ones most commonly used for Solanaceae. For GV3101 no significant bleaching was observed. For the EHA105 infiltrated plants, bleaching was observed in 40% of PDSS and 20% of PDSO plants in condition C, 21 dpi (see Table 5.2). However, bleaching was observed only in small areas of the leaf and never in a whole section of the plant (see Figure 5.5). The examination of PDS expression levels in bleached *P. peruviana* plants was not further pursued, as PDS silencing in *Withania somnifera* showed stronger bleaching and was therefore further investigated.

Table 5.2 Final conditions tested for *Physalis peruviana* agroinfection with PsPDSO and PpPDSA at OD_{600} 1.0.

Strain	Agrosuspension injection in stem	Plant cultivation condition	Result photobleaching
EHA105	yes	A	no photobleaching
GV3101	yes	A	no photobleaching
EHA105	yes	B	no photobleaching
GV3101	yes	B	no photobleaching
EHA105	yes	C*	2/5 replicates for PDSS, 1/5 replicates for PDSO
GV3101	yes	C	no photobleaching

* Plant cultivation conditions C involves 22 °C (humidity not controlled), 150 µmol/s light intensity and 16/8 h photoperiod

[x] Photo taken by Prof. Dr. Jakob Franke

Figure 5.5 Systemic bleaching was observed in newly emerging *P. peruviana* leaves after PDS silencing when agrosuspension was injected into the stem.

It was concluded that VIGS in *P. peruviana* was possible, but required very specific conditions. The plant seemingly presents a very robust immune system, as infiltrated leaves were often abscised and systemic bleaching was only achieved when the *Agrobacterium* suspension was directly injected into the phloem.

For VIGS in *Physalis peruviana* not many examples are reported in the literature. The only known example for PDS silencing of this plant was described by Osorio-Guarin *et al*. [149] Here, photobleaching was observed in 50% of replicates and a maximal reduction of transcript levels by 88% was observed. Those results could not be replicated when using similar silencing conditions. However, changing the *Agrobacterium* strain used from GV3101 to EHA105, injecting agrosuspension into the stem in addition to leaf infiltration and lowering plant cultivation temperatures resulted in systemic bleaching in up to 40% of the replicates.

Since the efficiency in both the number of silenced replicates as well as the degree of bleaching in the individual plant were quite low, it was decided that this plant was not suited for high-throughput screening of gene candidates for withanolide biosynthesis. Therefore, another withanolide producer, the Indian *Withania somnifera* was examined as a silencing system, that will be described in the following section.

Withania somnifera

As an alternative VIGS system, the common withanolide producer *Withania somnifera* was examined. Silencing of the PDS in *Withania somnifera* has been previously reported.[73,147,271] Literature suggested agroinfection with strain GV3101 with an OD_{600} of 1.3. [271] Plant growth

conditions after agroinfiltration were described as 70% humidity and a 16/8 h light-dark cycle at 20-22 °C.[271]

Therefore, infiltrations were carried out with GV3101 and EHA105 with OD_{600} 1.3. Plants were incubated in condition C (Table 5.1), as it was similar to conditions described in the literature. Strain EHA105 was included in this experiment, as it had been successful in *P. peruviana* PDS silencing.

For *W. somnifera*, bleaching was observed in newly emerging leaves 14 dpi in 50% (5/10) of replicates, when infiltrated with the EHA105 strain (see Figure 5.6). Leaf samples were harvested 21 dpi. Subsequent repetitions of WsPDS infiltration experiments with the same conditions exhibited bleaching in up to 100% of replicates. The frequency of PDS silencing reported here exceeds the observed frequency of 49% in the literature. [147] Plants infiltrated with GV3101 exhibited bleaching in 30% of replicates (3/10), however, bleaching affected smaller areas of the plant compared to plants infiltrated with EHA105. Therefore, strain EHA105 was chosen for further experiments.

Figure 5.6 Photobleaching of new leaves after PDS infiltration was observed in *W. somnifera* 14 dpi when infiltrated with *A. tumefaciens* strain EHA105.

As bleaching was observed in *W. somnifera* after agroinfection with WsPDS, those phenotypic changes were validated by expression analysis. Green and white leaves of infiltrated plants were harvested separately and RNA was extracted for real-time quantitative PCR (RT-qPCR). Relative quantification was performed with the $\Delta\Delta C_t$ method, using elongation factor 1α for normalization. [xi] Average expression of PDS decreased by 43% in green leaves and 71% in white leaves respectively (see Figure 5.7).

[xi] *Withania somnifera* transcript TRINITY_DN6100_c0_g1_i26 was identified by BLASTn search, using *Solanum lycopersicum* elongation factor 1α as reference (Accession Nr.: NM_001247106); 92% identity at 99% coverage

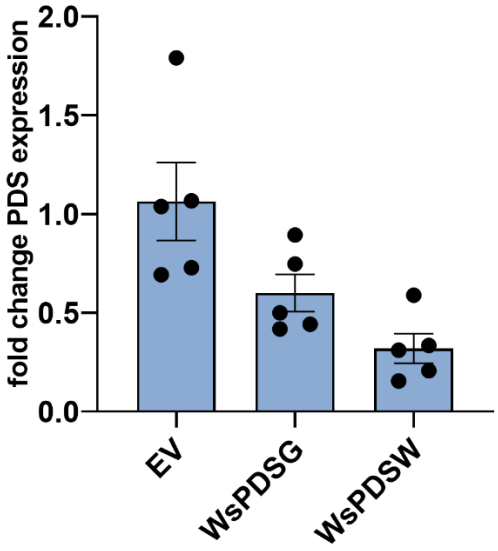


Figure 5.7 Fold changes detected in WsPDS expression compared to the EV mean for green (G) and white (W) leaves. Error bars represent SEM (n=5).

W. somnifera PDS silencing described in the literature led to a decrease of PDS mRNA levels ranging from 49% to 94%.[73,147] Therefore, it was concluded that silencing of PDS was in accordance with results from the literature and decrease of expression could be observed independently of the phenotype. However, knockdown of expression was stronger in tissues that exhibited the desired bleaching. Since VIGS in *Withania somnifera* exhibited a higher infection frequency and stronger photobleaching than *P. peruviana* (Figure 5.8), this plant was selected as a system for further screening of gene candidates for withanolide biosynthesis.

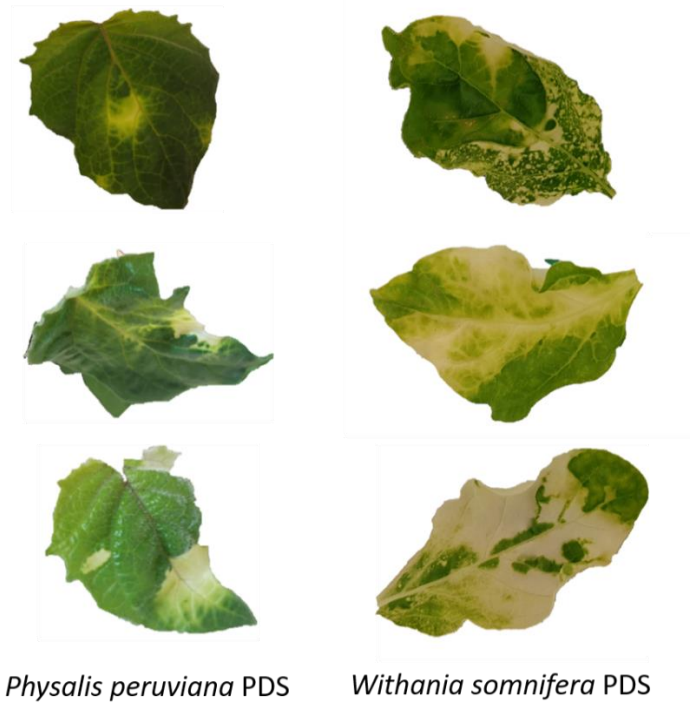


Figure 5.8 Comparison of *P. peruviana* and *W. somnifera* PDS silenced leaves show a stronger bleached phenotype for *W. somnifera* leaves.

5.2.3 Applying VIGS to withanolide biosynthesis by targeting the *24-isomerase* gene

Construction and transformation of a 24ISO plasmid for VIGS

While a PDS control can prove the general applicability of VIGS in a plant, a second control was needed that was directly linked to withanolide metabolism to confirm that silencing affected withanolide levels and the extraction and analysis protocol applied would be able to detect such changes.

The last known gene in withanolide biosynthesis, the 24ISO, was set as a secondary control, as silencing of this gene would result in a decrease of withaferin A (**5**) levels and potentially in accumulation of 24-methylenecholesterol (**4**, see Scheme 5.1 in the introduction of this chapter). VIGS of this gene has been reported before. [64]

The sequence of *W. somnifera* 24ISO was described in the literature. [64] While Knoch *et al.* generated a 150 bp silencing fragment from three joined 50 bp sequences, this work attempted cloning of one coherent silencing fragment for VIGS. In order to generate suitable silencing fragments and to avoid off targeting, the siFi21 software was used. [276] An illustration of the selected target sites is depicted in Figure 5.9.

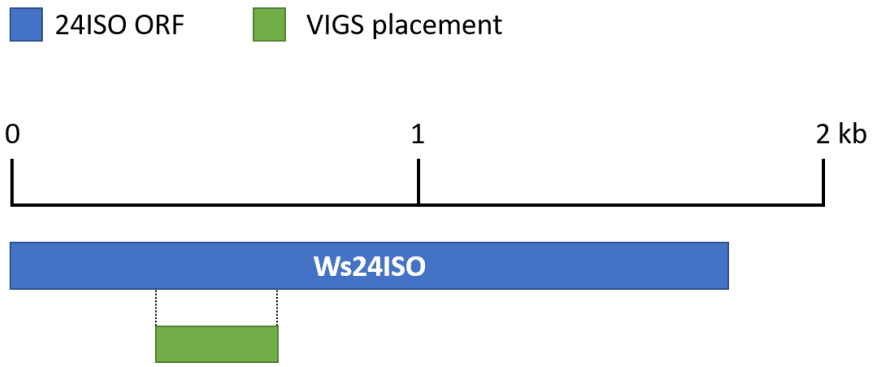


Figure 5.9 Target site (green) for virus induced gene silencing of *Withania somnifera* 24ISO (PDS, blue) used in this work.

The pTRV2-24ISO constructs were obtained using the same technique as for the pTRV2-PDS plasmids. cDNA generated from aerial tissues RNA of *W. somnifera* and *P. peruviana* plants was used as a template for Gibson assembly (In-Fusion®, Takara). The obtained plasmid was confirmed by colony PCR and DNA sequencing (Eurofins, Ebersberg). PDS silencing was most efficient with *A. tumefaciens* strain EHA105. Therefore, the pTRV2-Ws24ISO was transformed into this strain using electroporation. The resulting transformants were confirmed by colony PCR.

Analysis of withaferin A levels after agroinfection

As the VIGS positive control 24ISO targeted withanolide biosynthesis, an overview of main withanolides expected in *Withania sominfera* plants grown in this project is needed.

The plants used for this work accumulated withaferin A (**5**), withanolide D (**70**) and withanolide A (**9**) as their main withanolides (see chapter 2.2). However, the concentrations for **9** and **70** were far lower compared to **5** and exhibited higher fluctuations, exacerbating statistical analysis of VIGS results. Therefore, the subsequent analysis of VIGS effects was based on withaferin A (**5**) levels.

Agroinfection of *W. somnifera* with 24ISO was conducted with the conditions determined in chapter 5.2.2. Leaf samples were extracted and analyzed via LCMS with emodin (**105**) as an internal standard for relative quantification (see chapter 8.2.4).

Figure 5.10 shows an exemplary chromatogram of a *W. somnifera* leaf extract infiltrated with empty vector (EV). Withaferin A (**5**) elutes at 5.1 min, withanolide D (**70**) at 5.7 min, withanolide A (**9**) at 6.2 min and emodin (**105**) at 7.9 min.

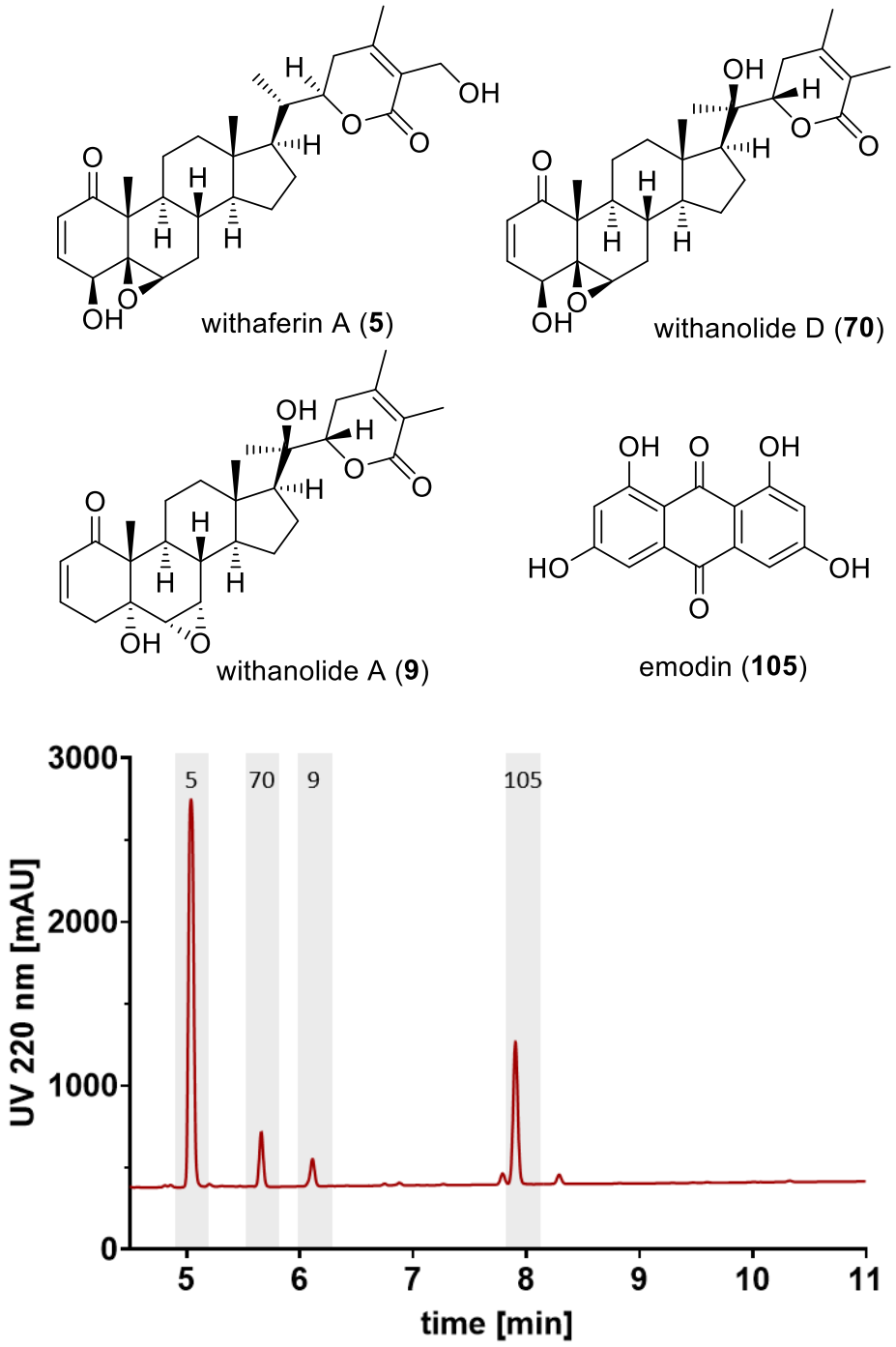


Figure 5.10 The exemplary UV trace of a chromatogram of *Withania somnifera* leaf extract infiltrated with empty vector shows the three main withanolides withaferin A (**5**), withanolide D (**70**) and withanolide A (**9**).

VIGS experiments in the following chapter were conducted in two independent iterations. The second repetition had to be cultivated in a different plant growth chamber due to construction work, albeit with the same settings. Thus, 24ISO infiltration was repeated in the second growth chamber as well to ensure reproducibility in a different location (blue, Figure 5.11).

An average decrease of 31% and 34% of withaferin A (**5**) content was observed for 24ISO plants compared to the empty vector control. One replicate in particular showed a decrease by 90% (see Figure 5.11A). In order to confirm that those metabolic effects were due to decreased expression, quantitative PCR (RT-qPCR) was performed with the remaining leaf harvest of experiment 1. Comparative quantification was carried out with the $\Delta\Delta C_t$ method, [292] using β-tubulin [xii] as a reference gene. The average expression of the 24ISO was decreased by 23%, with the outlier in withaferin A (**5**) levels showing the lowest expression of 24ISO (54% decrease, Figure 5.11B).

Samples were furthermore investigated for accumulation of 24-methylenecholesterol (**4**), the substrate of the 24ISO.[xiii] However, no significant accumulation of this compound was detected. Since **4** is a precursor in phytosterol biosynthesis [293,294] it can be reasoned that substrate accumulated by silencing of the 24ISO would be utilized by the phytosterol pathway.

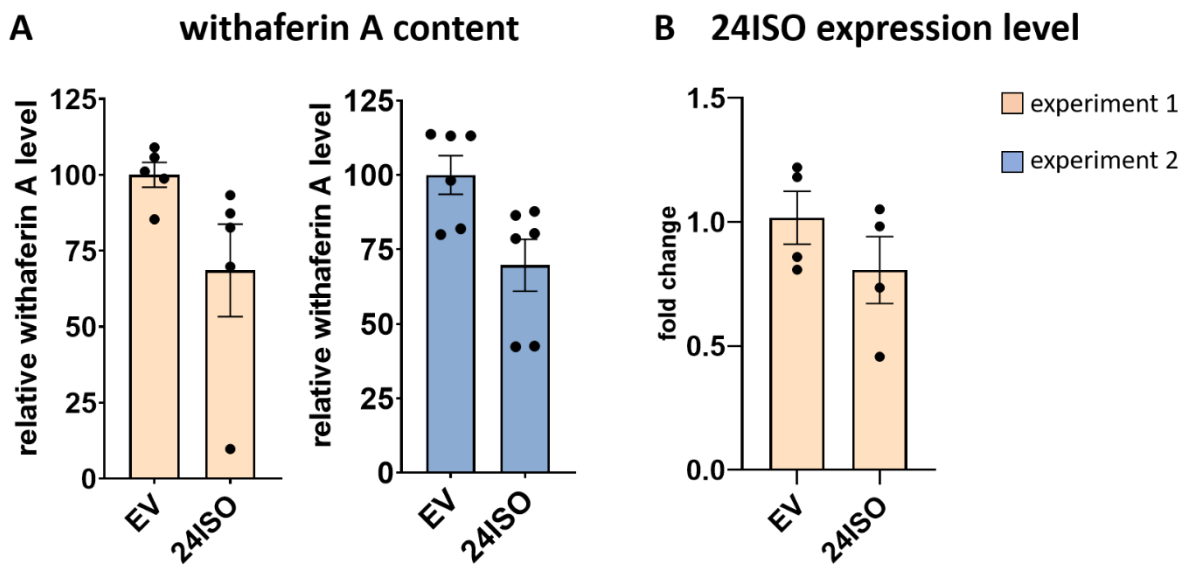


Figure 5.11 Virus induced gene silencing of the 24ISO showed decrease in withaferin A (**5**) content (A) and in 24ISO expression levels (B) compared to the EV-control. Error bars represent SEM (n=4-6).

VIGS results for the 24ISO can be compared to the results of Knoch *et al*. [64] Their research described a decrease by ca. 60% in withaferin A (**5**) content and ca. 80% of 24ISO gene expression when silencing the 24ISO. Conditions for agroinfiltration were similar as far as is known.

[xii] *Withania somnifera* transcript TRINITY_DN2481_c0_g1_i21 was identified by BLASTn search, using *Lycopersicon esculentum* β-tubulin (Accession Nr.: DQ205342); 84% identity at 98% coverage

[xiii] Analysis conducted by Samuel Edward Hakim with an LCMS gradient suited for (xMC)

In general, it was concluded that silencing of the 24ISO evoked a decrease of withaferin A (**5**) content, albeit lower than described in the literature. This can be due to different target sequences used. Knoch *et al.* used a different target fragment for 24ISO silencing that contained three distinct target sequences within the 24ISO CDS. Targeting of multiple locations could potentially amplify the observed silencing effect. Furthermore, natural fluctuations in withanolide content in *W. somnifera* were observed that could possibly override silencing effects. Lastly, it has to be noted that harvest of leaf tissues is untargeted, as no phenotypic changes (e.g. photobleaching) are present. During extraction tissues with different degrees of withaferin A (**5**) decrease can be processed together, leading to a "dilution" of observed silencing effects. In order to direct harvest towards VIGS affected tissues and therefore observe silencing response without mitigation of unaffected tissues the following chapter describes P450-PDS co-infiltration in addition to P450-only infiltrations.

5.3 Virus induced gene silencing of cytochrome P450 candidates

One of the most prominent questions of withanolide biosynthesis is the biosynthetic origin of the multiple oxidations observed in withanolides (see chapter 1.3.3). Chapter 4 described the selection of 21 cytochrome P450 candidates that could potentially perform oxidation of the side chain with subsequent lactonization and furthermore oxidation of the A-ring.
Withanolide biosynthesis involves multiple steps from the last known intermediate, 24-methyldesmosterol (**3**), to main withanolides such as withaferin A (**5**). Until now no enzymes or substrates have been reported for any of those missing steps and potential intermediates are not commercially available. Therefore *in vitro* assays with P450 candidates would be tedious and not very promising unless information can be gathered about enzymes or substrates involved. A primary screening of P450 candidates, using knockdown experiments, aimed to gain insights into withanolide biosynthesis and narrow down the P450 candidate pool.

5.3.1 Construction of WsP450- plasmids for VIGS

As described in chapter 5.2.1, VIGS fragment selection needs to consider potential off-targeting. An analysis with the siFi21 prediction tool [275,276] was conducted for all P450 candidates and unique target sequences were selected. An overview of VIGS target sites is given in Figure 5.12.
The pTRV2-P450 plasmids were constructed using Gibson assembly (In-Fusion®, Takara). As a template for PCR cDNA from *Withania somnifera* was used. Since most withanolides are

found predominantly in leaves (e.g. withaferin A, **5**),[7] but *de novo* production in roots is possible as well, [295] RNA was extracted from root and aerial tissues and resulting cDNAs were pooled (chapter 8.2.2). Cloning experiments were confirmed with colony PCR and DNA sequencing (Eurofins, Ebersberg).

In total, the desired fragments of 17 out of 21 P450 candidates could be cloned. The results are shown in Table 5.3. Repeated attempts to integrate the missing fragments into pTRV2 did not show positive colonies after transformation, although the same aliquot of cut vector successful in other assemblies was used. Since the particular P450 candidates were selected due to only one indicator respectively (in case of P450-2 two) and the candidate pool already contained a large number of P450s, cloning of those candidates was considered as not essential for VIGS testing and silencing experiments were conducted with the 17 obtained constructs.

Table 5.3 Cloning results for pTRV2-P450 constructs.

Candidate	Cloning result
P450-1	verified
P450-2	no positive colonies
P450-3	verified
P450-4	verified
P450-5	verified
P450-6	verified
P450-7	verified
P450-8	verified
P450-9	verified
P450-10	verified
P450-11	verified
P450-12	verified
P450-13	verified
P450-14	verified
P450-15	verified
P450-16	no positive colonies
P450-17	verified
P450-18	no PCR product
P450-19	verified
P450-20	no positive colonies
P450-21	verified

The obtained pTRV2-P450 plasmids were transformed into *A. tumefaciens* EHA105 by electroporation. Successful transformation was confirmed by colony PCR. The results for VIGS with those constructs will be described in the following chapter.

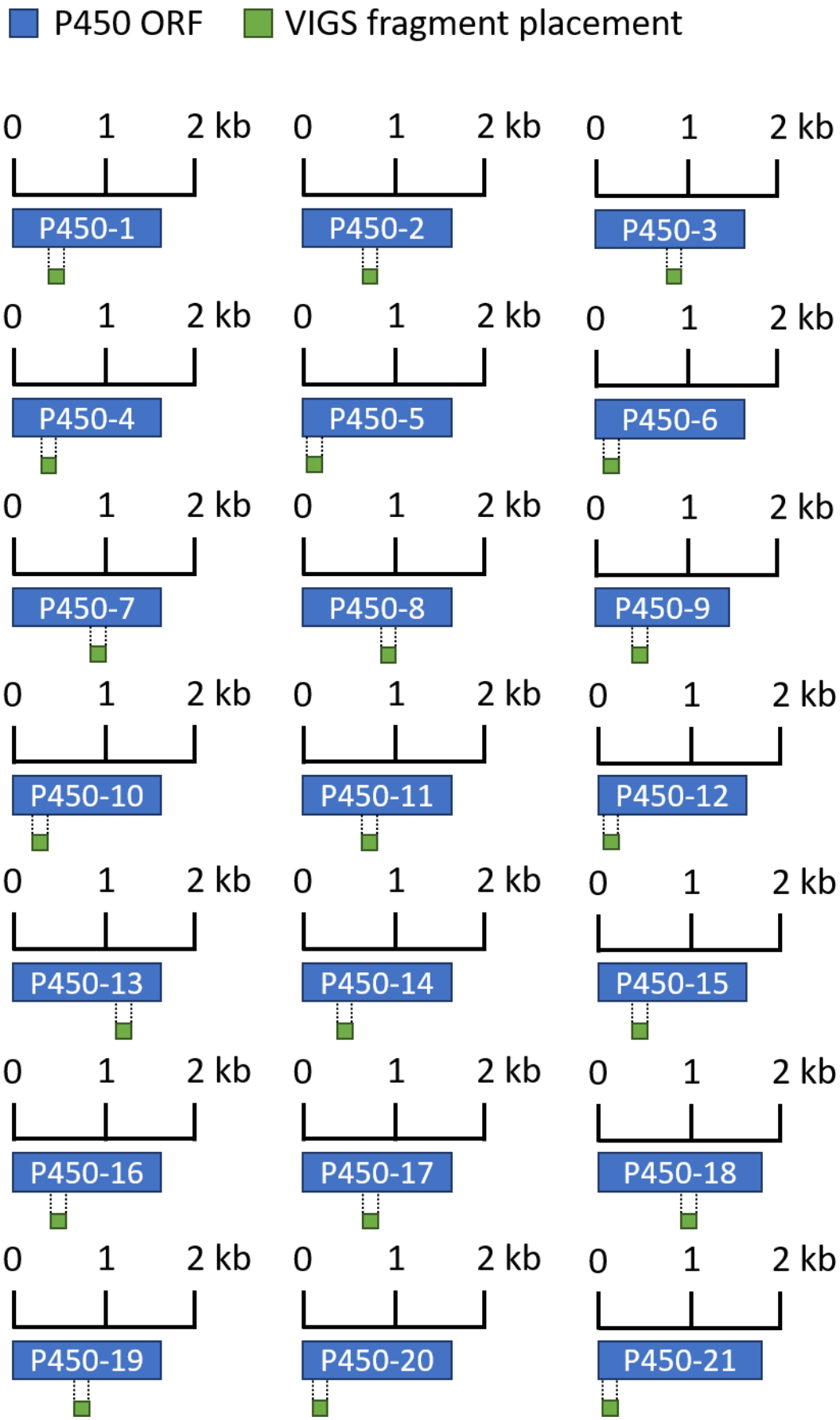


Figure 5.12 VIGS target sites (green) for silencing of P450-candidates (blue) for withanolide biosynthesis.

5.3.2 WsP450 agroinfiltration resulted in significant withaferin A decrease for five candidates

As silencing of the 24-isomerase demonstrated that VIGS in *W. somnifera* could evoke decrease of withaferin A (**5**) levels, candidate genes for the oxidations involved in withanolide biosynthesis should be investigated. Fragments of cytochrome P450 genes were cloned into pTRV2 and transformed into *Agrobacterium tumefaciens* EHA105 (see previous chapters). The respective pTRV2-GOI strain was infiltrated together with pTRV1 to elicit silencing and samples were harvested 28 dpi. Extraction was conducted based on fresh weight to preserve the option for RNA extraction for RT-qPCR analysis.

VIGS samples were extracted and LCMS chromatograms were analyzed for their withaferin A (**5**) content, using internal standard emodin (**105**) for normalization. An exemplary chromatogram is depicted in chapter 5.2.3, Figure 5.10. Besides withaferin A (**5**), the main withanolides present in *Withania somnifera* are withanolide D (**70**) and withanolide A (**9**).[xiv] The detected amounts of **8** and **70**, however, are far lower and exhibit higher fluctuations and therefore exacerbate statistical analysis. Thus, analysis of VIGS experiments was conducted on the base of withaferin A (**5**) levels. VIGS experiments for P450 silencing was conducted in 2 independent iterations to investigate reproducibility of the experiment.

Figure 5.13 gives an overview of relative withaferin A (**5**) contents after agroinfection for both experiments (experiment 1: orange, experiment 2: blue), normalized with the internal standard emodin (**105**). All infiltrations were analyzed for a significant decrease of withaferin A (**5**) levels compared to an empty vector (EV) control (Welch's t-test, P-value ≤ 0.05). A significant decrease of withaferin A (**5**) levels in at least one of two experiments was detected for five P450 candidates.

P450-7 showed a significant decrease of withaferin A (**5**) content of 25% in experiment 1. Iteration 2 showed a decline of 28%, albeit not statistically significant due to broad dispersion of observed withaferin A (**5**) levels.[xv] For P450-8 a decrease of 39% was observed in experiment 1, with one replicate showing a strong decrease of 76%. This result was not repeated in experiment 2. P450-14 showed a reproducible significant decrease of **5** in both infiltrations of 12% and 21%, respectively. The same was observed for P450-15, with a decrease of 18% and 21%. Lastly, P450-17 showed a decrease of 15% in iteration 1, and a downwards trend of

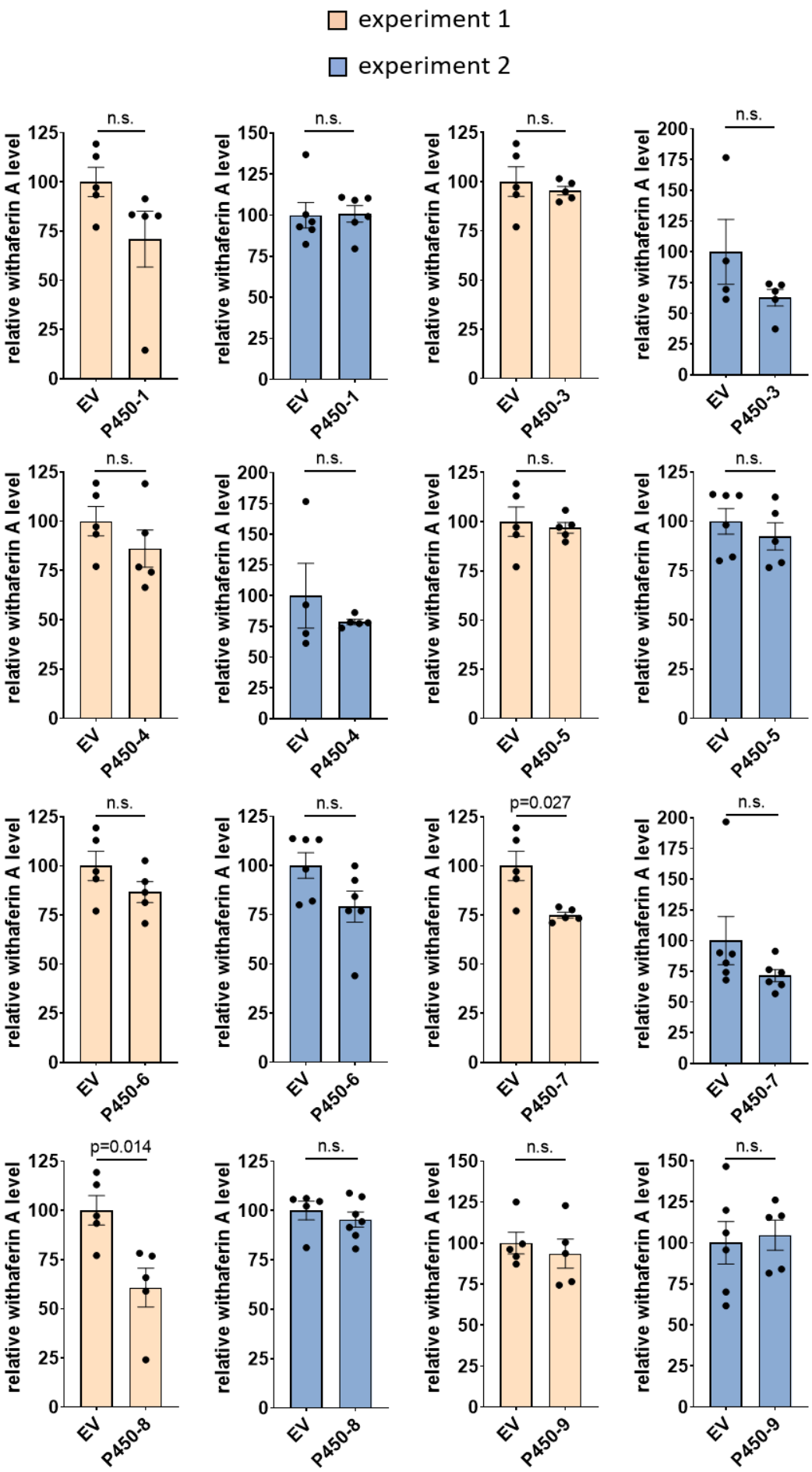
[xiv] Analysis conducted by Samuel Edward Hakim

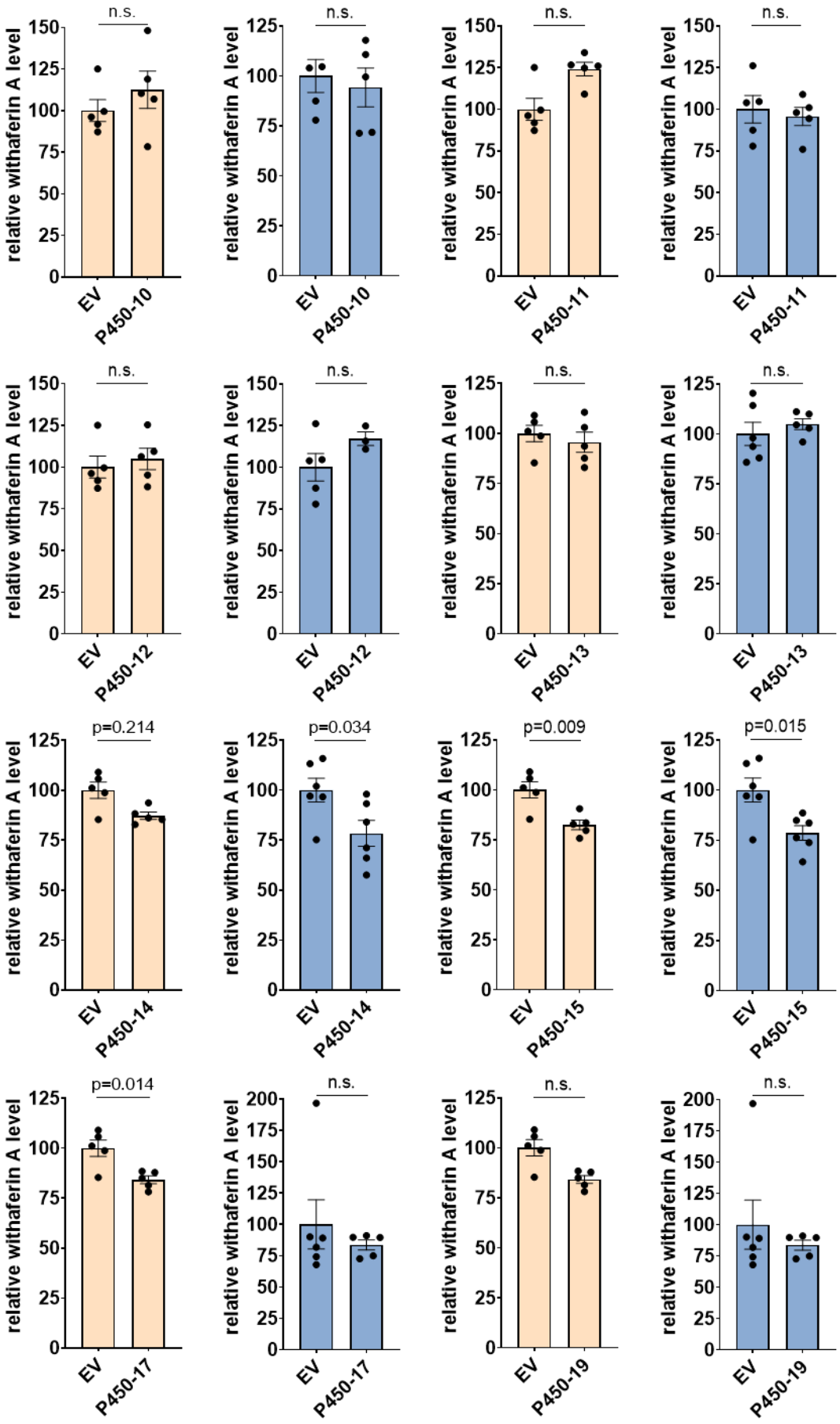
[xv] When excluding the outlier in the EV control, a decrease of 11% in withaferin A levels was observed for P450-7, experiment 2.

16% in experiment 2, that was not statistically significant due to fluctuations in withaferin A (**5**) content. Furthermore, LCMS chromatograms were investigated for accumulation of new compounds that could indicate which intermediates are present after silencing. However, no accumulation of new masses was detected. This can be due to low solubility of the intermediate in the chosen extraction method or low concentrations in general. Furthermore, it is possible that accumulated intermediates might be converted by other enzymes in the complex withanolide pathway.

In addition to P450-8, P450-1 and P450-6 showed one replicate each with a drastic withaferin A (**5**) decrease. As levels of **5** exhibited strong fluctuations and such strong decreases were not observed for P450-1 and P450-6 again, those candidates were not included in the pool for further testing.

In this section, a total of five P450 candidates were determined that showed significant decrease in withaferin A (**5**) levels after silencing, that were therefore chosen for further *in vitro* assays. As silencing effects do not spread evenly throughout the plant and harvest of leaves is not guided by a phenotype, as it would be the case for PDS-bleached plants, an additional approach was tested. To identify strongly affected leaves and improve the selection of tissues for extraction, the following chapter describes VIGS experiments conducted with a GOI+PDS co-infiltration.


experiment 1
experiment 2
relative withaferin A level
n.s.
p=0.027
p=0.014
EV
P450-1
P450-3
P450-4
P450-5
P450-6
P450-7
P450-8
P450-9


n.s.
relative withaferin A level
EV
P450-10
n.s.
relative withaferin A level
EV
P450-10
n.s.
relative withaferin A level
EV
P450-11
n.s.
relative withaferin A level
EV
P450-11
n.s.
relative withaferin A level
EV
P450-12
n.s.
relative withaferin A level
EV
P450-12
n.s.
relative withaferin A level
EV
P450-13
n.s.
relative withaferin A level
EV
P450-13
p=0.214
relative withaferin A level
EV
P450-14
p=0.034
relative withaferin A level
EV
P450-14
p=0.009
relative withaferin A level
EV
P450-15
p=0.015
relative withaferin A level
EV
P450-15
p=0.014
relative withaferin A level
EV
P450-17
n.s.
relative withaferin A level
EV
P450-17
n.s.
relative withaferin A level
EV
P450-19
n.s.
relative withaferin A level
EV
P450-19

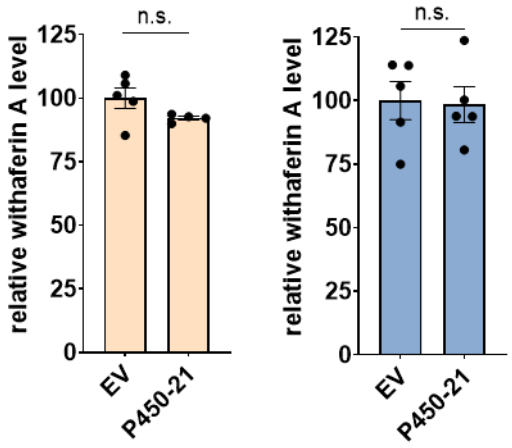


Figure 5.13 Relative withaferin A (**5**) levels of P450 candidates (P450) compared against empty vector control (EV) reveal significant decrease for 5 candidates (Welch's t-test, significant difference: $P \leq 0.05$). Error bars represent SEM (n=4-6).

Silencing of P450 genes with PDS guiding

The previous chapter described virus induced gene silencing of all P450 candidates for withanolide biosynthesis. Here, the harvest of leaf samples was not guided by a phenotype as it is the case for photobleaching after PDS infiltration. As observed in 5.2.2 a viral infection and therefore silencing effects do not spread evenly throughout the plant. This exacerbates selection of leaves for analysis, as silencing effects can be "diluted" by blending strongly affected with less affected tissues.

In order to direct harvest of leaf samples towards strongly silenced tissues the co-infiltration of EHA105 strains pTRV1, pTRV2-P450- and pTRV2-PDS was investigated. VIGS with a GOI-ChlH fusion construct to aid harvest has been reported previously. [296,297] Application of this method allowed to pinpoint tissues subjected to VIGS and facilitated analysis of experimental data. [297] Preliminary results from fellow PhD student Ling Chuang indicate that a PDS-GOI fusion functions similarly and aids selection of affected tissues. Here a slightly different approach was tested, as instead of fusion constructs, both targets were infiltrated from individual strains. This technique enables a fast mix-and-match principle and does not require additional cloning. In general, a systemic spread of silencing effects is achieved due to viral movement and RNAi within the plant. Thus, the presence of one or two viral vectors functioning as siRNA templates should be irrelevant. It was reasoned that PDS and P450 targeting siRNAs should spread together throughout the plant and therefore bleached tissues would indicate areas with generally high VIGS activity.

PDS co-infiltration was conducted in the greenhouse, formerly described as growth condition A (21-24 °C, 69% humidity, 11/13 h photoperiod), since it was the growth condition closest to condition C (22°C, no controlled humidity, 16 h light period) that showed to be the most effective for VIGS. The growth chamber described as condition C in previous chapters was not available due to construction work.
In this experiment, the positive control 24ISO and the five candidates that showed a significant decrease in withaferin A (**5**) levels in the previous chapter were co-infiltrated with PDS and harvested 28 dpi (P450-7, P450-8, P450-14, P450-15, P450-17). Extraction was performed as described previously, normalizing withaferin A (**5**) levels to the internal standard emodin (**105**). Most plants showed a varying degree of white patches on some leaves. Green and white leaves of each replicate were extracted separately, and the results are shown in Figure 5.14 (leaf color marked for each sample). As not all plants showed bleaching, the replicate number for white leaves was lower than for green leaves.

When comparing withaferin A (**5**) levels in white and green tissues, a stronger decrease in white leaves, indicating higher viral activity, was not observed. Instead, strong fluctuations in withaferin A (**5**) content limited reliability of statements as statistical power was compromised. Since the tested candidates had shown a significant decrease of withaferin A levels in previous experiments, it was assumed that PDS silencing might interfere with P450-silencing and therefore no decrease of withaferin A (**5**) levels was observed. A significant overall decrease compared to the EV control was detected only for P450-17 (20%, Welch's t-test, $P \leq 0.05$), but P450-8 and P450-14 showed a downwards trend in average levels of **5** as well. The positive control 24ISO did not show the expected decrease of withaferin A (**5**) levels.

This experiment was further repeated in the same location (greenhouse) with a higher number of replicates (data not shown). However, withaferin A (**5**) levels varied strongly throughout the entire iteration of the experiment, limiting validity of statistical statements. It has to be noted that a sudden increase in outside temperature was observed simultaneously with the second infiltration, raising the inside temperatures in the greenhouse >25°C. As previous results were not reproducible in this environment, it was concluded that successful conduction of VIGS requires temperatures below 25 °C.

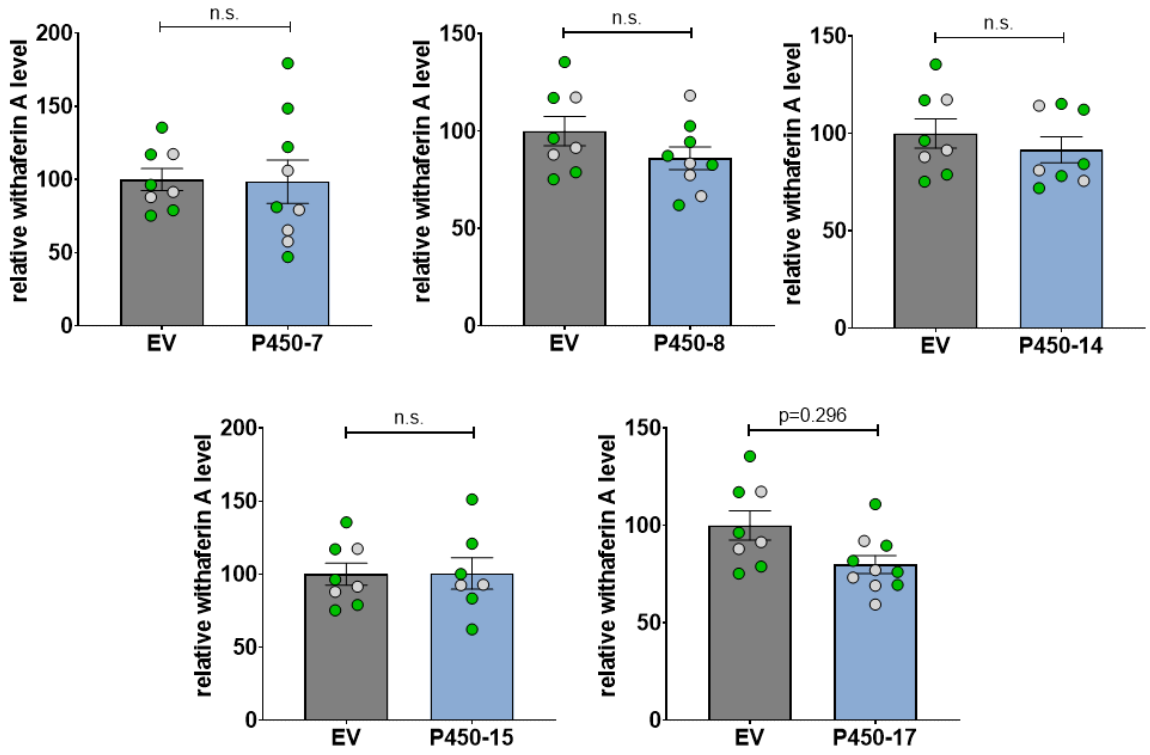


Figure 5.14 Co-infiltration of PDS and P450 candidates does not exhibit a stronger decrease of withaferin A (**5**) levels in bleached leaf areas (grey dots) compared to green leaf areas (green dots). Relative withaferin A (**5**) levels were compared to an empty vector (EV) control (Welch's t-test, significant difference: $P \leq 0.05$). Error bars represent SEM (n=7-10).

Conclusion and discussion of P450 VIGS experiments

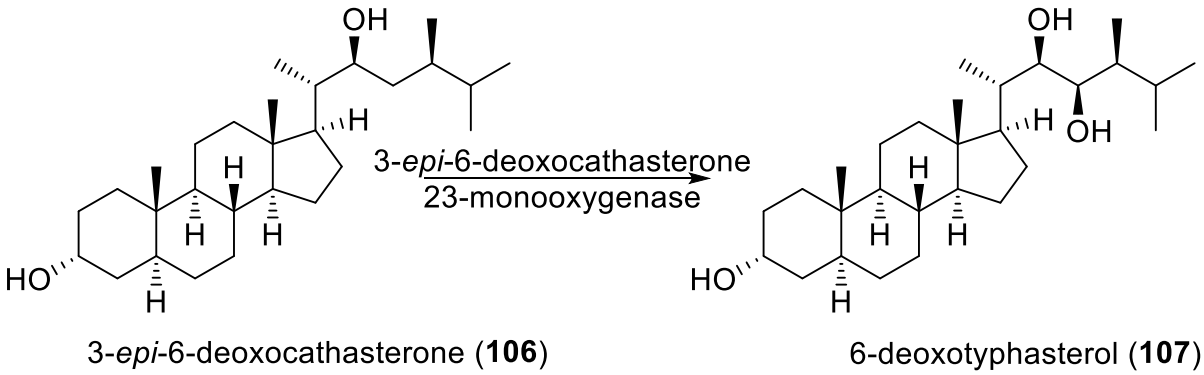
VIGS experiments with P450 gene candidates generated multiple insights. First, the negative influence of elevated temperatures on VIGS experiments was underlined. Chapter 5.2.2 described silencing attempts of the *Physalis peruviana* phytoene desaturase for multiple plant growth conditions. The expected bleaching of leaves was observed in condition C at 22 °C, but never at condition B (30 °C). PDS infiltration of *Withania somnifera* was thus conducted in condition C and showed the desired phenotype. In this chapter, VIGS of P450 gene candidates for the withanolide biosynthesis was conducted in three different locations due to construction work. Experiments that were conducted in a greenhouse without temperature regulation showed strong fluctuations in withaferin A (**5**) levels, once temperatures raised above 25 °C.

It was concluded that VIGS in *Withania somnifera* is strongly dependent on plant cultivation temperature, and effects of silencing cannot be observed past a certain threshold temperature. In the literature VIGS in *W. somnifera* has been reported for temperatures below 24°C. [147,271] However, it is unclear whether higher temperatures were not successful or not tested.

Furthermore, PDS co-infiltration was tested, as it was reasoned that silencing effects for both targets would simultaneously spread in the same tissues within the plant and leaves with strong bleaching would indicate strong silencing activity in general. This approach was based on a protocol by Palmer and O'Connor, that used GOI-ChlH fusions to target harvest of infected tissues. [296,297] However, no correlation between photobleached tissues and stronger withaferin A decrease was observed, indicating that PDS and GOI siRNAs either spread independently within the plant or GOI silencing was not effective.

Lastly, the main motivation for P450 silencing in *W. somnifera* was to narrow down the candidate pool of cytochrome P450 gene candidates for withanolide biosynthesis. VIGS experiments determined five candidates that showed significant decrease of withaferin A (**5**) levels, four of those showed this decrease in more than one independent repetition of the experiment.

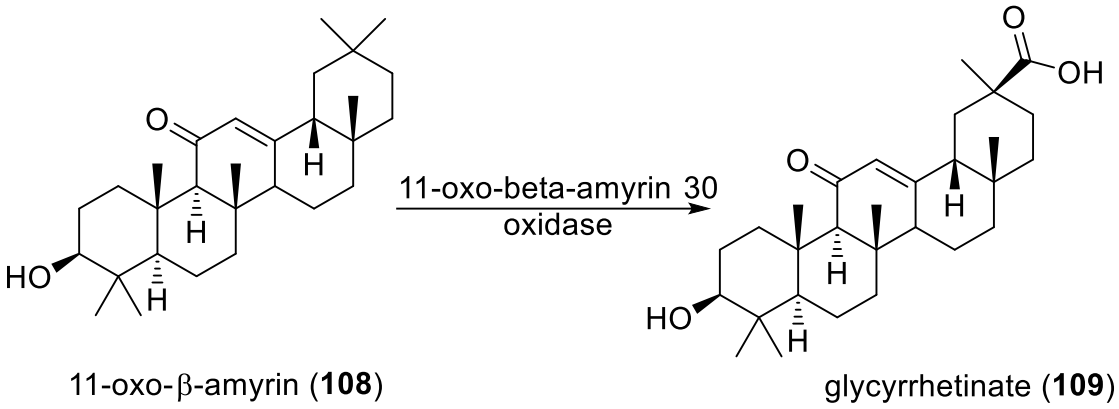
P450-7 showed a significant decrease of 25% in experiment 1. This candidate was selected for VIGS experiments based on its BLASTx annotation (see chapter 4.2.2). The annotation states a 3-*epi*-6-deoxocathasterone 23-monooxygenase. This enzyme catalyzes the hydroxylation at the C-23 position of 3-*epi*-6-deoxocathasterone (**106**) to form 6-deoxotyphasterol (**107**) in the brassinosteroid biosynthesis (see Scheme 5.2). [298] Due to the structural similarity of brassinosteroids and withanolides, it seems reasonable that P450-7 might have evolved to hydroxylate the withanolide side chain.



Scheme 5.2 The 3-*epi*-6-deoxocathasterone 23-monooxygenase catalyzes a C-23 hydroxylation in brassinosteroid biosynthesis. [298]

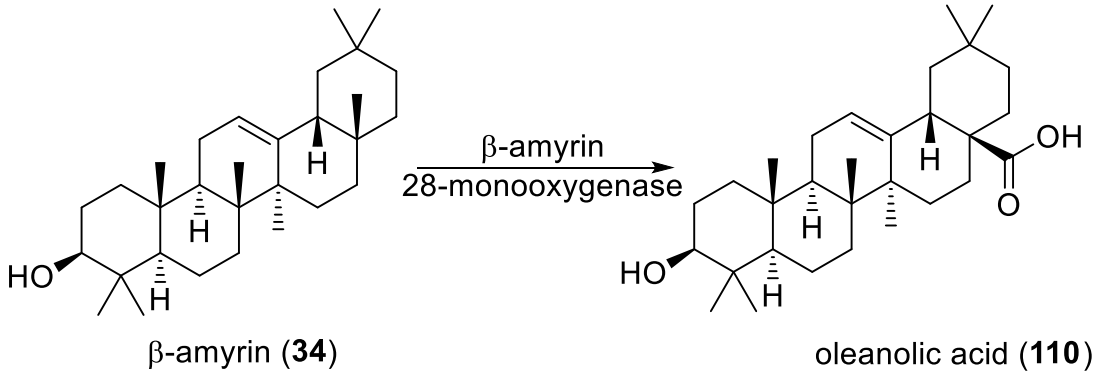
P450-8 showed a 39% decrease in withaferin A (**5**) levels in experiment 1. The candidate was selected due to BLASTx annotation and homology to steroidal glycoalkaloid hydroxylases. The annotation for P450-8 was 11-oxo-β-amyrin 30-oxidase, an enzyme that usually catalyzes three sequential oxidation steps at C-30 of 11-oxo-β-amyrin (**108**) towards glycyrrhetinate (**109**,

Scheme 5.3).[299] Furthermore a 54% aa identity to SGA side chain hydroxylases was observed. This candidate could have evolved to accept withanolides as a substrate and catalyze subsequential oxidation of the withanolide precursor side chain to form the lactone ring.



Scheme 5.3 11-Oxo-β-amyrin 30-oxidase catalyzes three sequential oxidation steps at C-30 of 11-oxo-β-amyrin (**108**) towards glycyrrhetinate(**109**). [299]

P450-14 showed a significant decrease of withaferin A (**5**) levels in both infiltrations with 12% and 21% respectively. The annotated function states β-amyrin 28-monooxygenase that is known to catalyze the carboxylation of β-amyrin (**34**) at the C-28 position to form oleanolic acid (**110**, Scheme 5.4). [254,300,301] As the function of the described enzyme is similar to the annotation of P450-8, this candidate might also be involved in subsequential oxidation of the withanolide side chain with consequential lactone formation. P450-14 was selected based on co-expression analysis performed with CoExpNetViz [259] using phytosterol and withanolide biosynthesis genes.

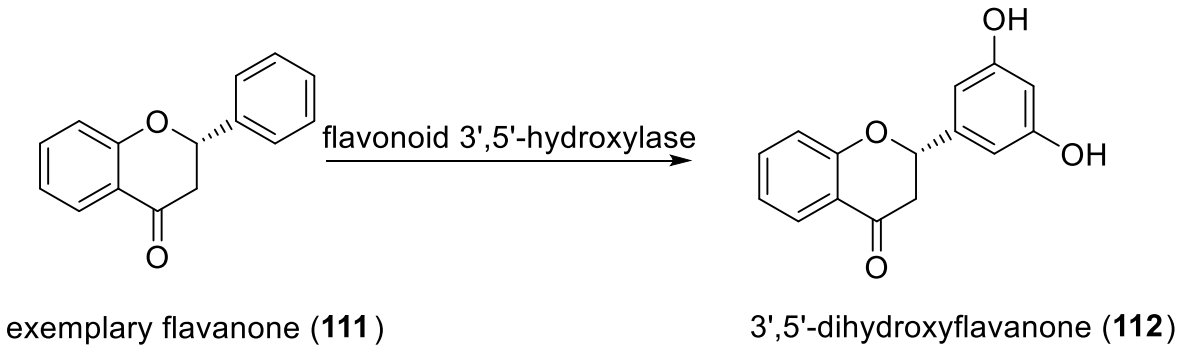


Scheme 5.4 β-amyrin 28-monooxygenase catalyzes carboxylation of β-amyrin (**34**) at the C-28 position to form oleanolic acid (**110**).[254,300,301]

P450-15 showed a significant decrease of withaferin A (**5**) levels in both iterations of agroinfiltration with 17% and 21% respectively. Here the annotation was cytochrome P450 94C1, an enzyme that has been reported to be involved in the oxidation of plant hormone

jasmonoyl-L-isoleucine. [302–304] In vitro experiments have furthermore demonstrated a function as in-chain fatty acid hydroxylase. [302] Based on this annotation, no hypothesis can be formed regarding an involvement in withanolide biosynthesis. This candidate was selected based on the CoExpNetViz [259] analysis as well.

P450-17 showed a significant decrease in withaferin A (**5**) levels in iteration 1 and furthermore in a P450-PDS co-infiltration experiment. The annotation states a flavonoid 3',5'-hydroxylase, which is known to catalyze multiple 3',5'-hydroxylations in flavonoids such as anthocyanins (**111** to **112**, Scheme 5.5). [305] Here, again no hypothesis can be formed regarding an involvement in withanolide biosynthesis based on this annotation. This candidate was selected based on homology to a previously determined *P. peruviana* candidate.



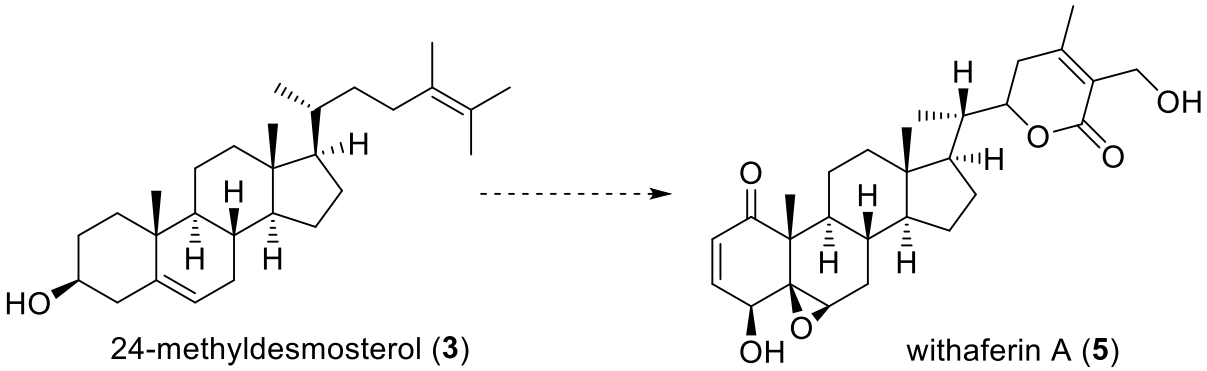
exemplary flavanone (**111**) 3',5'-dihydroxyflavanone (**112**)

Scheme 5.5 Flavonoid 3',5'-hydroxylase catalyzes 3',5'-hydroxylations with various flavonoid substrates. [305]

LCMS chromatograms were further investigated for accumulation of new compounds, but no new peaks were detected. In summary five cytochrome P450 candidates were determined that will be further examined for *in vitro* activity on 24-methyldesmosterol (**3**), the last known withanolide intermediate.

5.4 Virus induced gene silencing of dehydratase candidates

As a virus induced gene silencing system was established in *Withania somnifera*, reactions other than the multiple oxidation steps in withanolide biosynthesis can be investigated. In contrast to the last known intermediate 24-methyldesmosterol (**3**), most withanolides exhibit a double bond in the A-ring (e.g. **5**, Scheme 5.6). Therefore, a dehydratase is needed, that could perform an elimination reaction with the 3-hydroxy group leaving the A-ring. Suitable gene candidates were identified in chapter 4.3 and should now be subject of experimental analysis.



Scheme 5.6 Withaferin A (**5**) contains a double bond in the A-ring.

5.4.1 Construction of dehydratase plasmids for virus induced gene silencing

The 14 dehydratase (DH) candidates determined in chapter 4.3 were analyzed for off-targeting using siFi21. [275,276] This eliminated 2 candidates from the pool, as no fragment with sufficient length for VIGS was available that did not present potential off-targeting and promiscuous silencing could potentially result in unpredictable metabolic responses. For the remaining 12 candidates a 150-300 bp fragment for VIGS was selected. An illustration of VIGS fragment placement can be found in Figure 5.15.

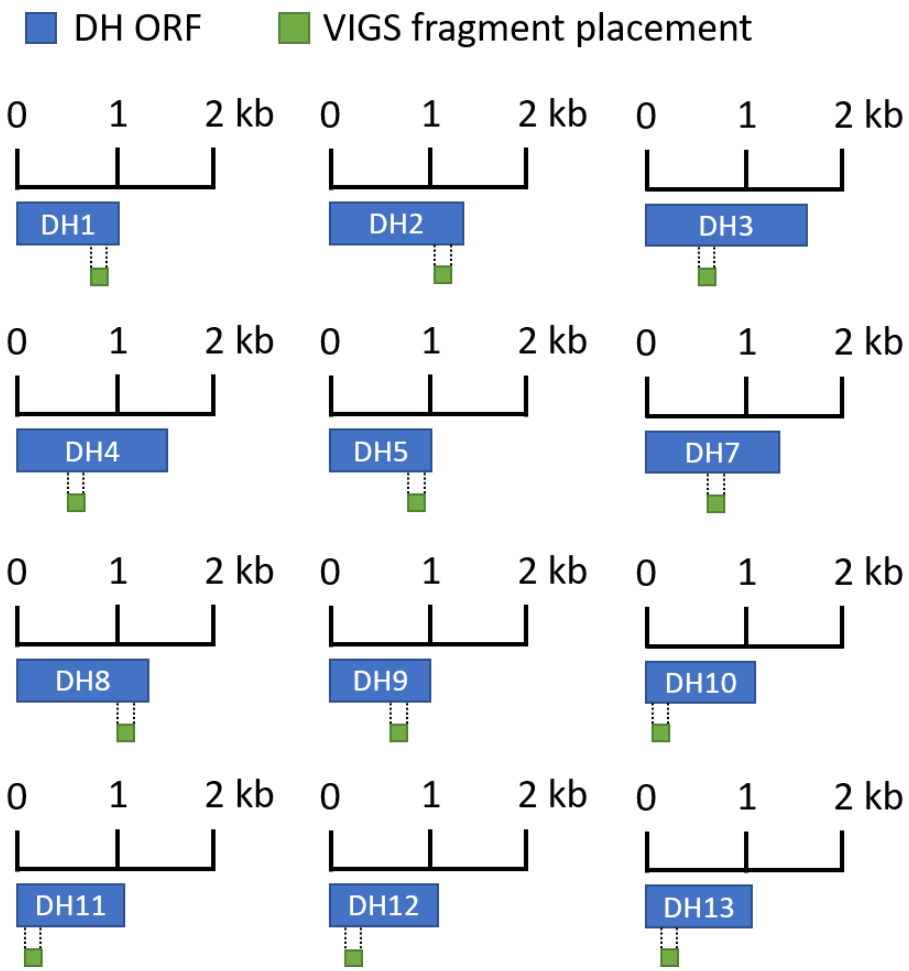


Figure 5.15 VIGS fragment placement for dehydratase candidates.

Cloning and agrotransformation was conducted together with Lara Tenge as part of her bachelor's thesis. In summary, 6 of the 12 fragments were successfully cloned into pTRV2 using Gibson assembly (In-Fusion®, Takara) as previously described (see chapter 8.2.2). An overview of the obtained plasmids is given in Table 5.4. All plasmids were transformed into *A. tumefaciens* EHA105 and transformants were confirmed by colony PCR and DNA sequencing (Eurofins, Ebersberg).

Table 5.4 Cloning results of pTRV-DH constructs.

Plasmid	Cloning result
pTRV2-DH1	verified
pTRV2-DH2	verified
pTRV2-DH3	verified
pTRV2-DH4	no positive colonies
pTRV2-DH5	verified
pTRV2-DH6	no unique target sequence
pTRV2-DH7	verified
pTRV2-DH8	no positive colonies
pTRV2-DH9	no positive colonies
pTRV2-DH10	no positive colonies
pTRV2-DH11	verified
pTRV2-DH12	no positive colonies
pTRV2-DH13	no positive colonies
pTRV2-DH14	no unique target sequence

5.4.2 WsDH agroinfiltration resulted in withaferin A decrease for two candidates

While chapter 5.2.3 and 5.3 showed that VIGS could elicit decrease of withaferin A (**5**) levels, this technique should now be employed for investigation of dehydratase candidates. Infiltration of dehydratase candidates was performed together with Lara Tenge as part of her bachelor's thesis. Since experiments were conducted simultaneously to P450-PDS co-infiltrations, dehydratase candidates were co-infiltrated with PDS into *W. somnifera* as well.

PDS co-infiltration in chapter 5.3.2 showed no elevated silencing levels in photobleached tissues compared to green leaves. Nonetheless, for DH silencing only leaves with some degree of bleaching were analyzed to ensure comparability of GOI and EV samples. Extraction of VIGS samples followed the previously described protocol (see chapter 8.2.4).
Figure 5.16 shows the relative withaferin A (**5**) levels, normalized with internal standard emodin, for each candidate compared to an empty vector (EV) control. A significant decrease of 41% was detected for DH11, and while the decrease for DH7 was not statistically significant due to strong dispersion of withaferin A (**5**) levels, a decreasing trend was observed as well (Welch's t-test, P-value ≤ 0.05). Furthermore, LCMS data was investigated for accumulation of new compounds. However, no putative intermediates were detected. A second independent infiltration was performed without PDS co-infiltration in growth chamber 2, but no significant

changes in withaferin A (**5**) levels were detected compared to an EV control. As reproducibility was not given in this set of infiltrations, *in vitro* assays were conducted with all dehydratase candidates.

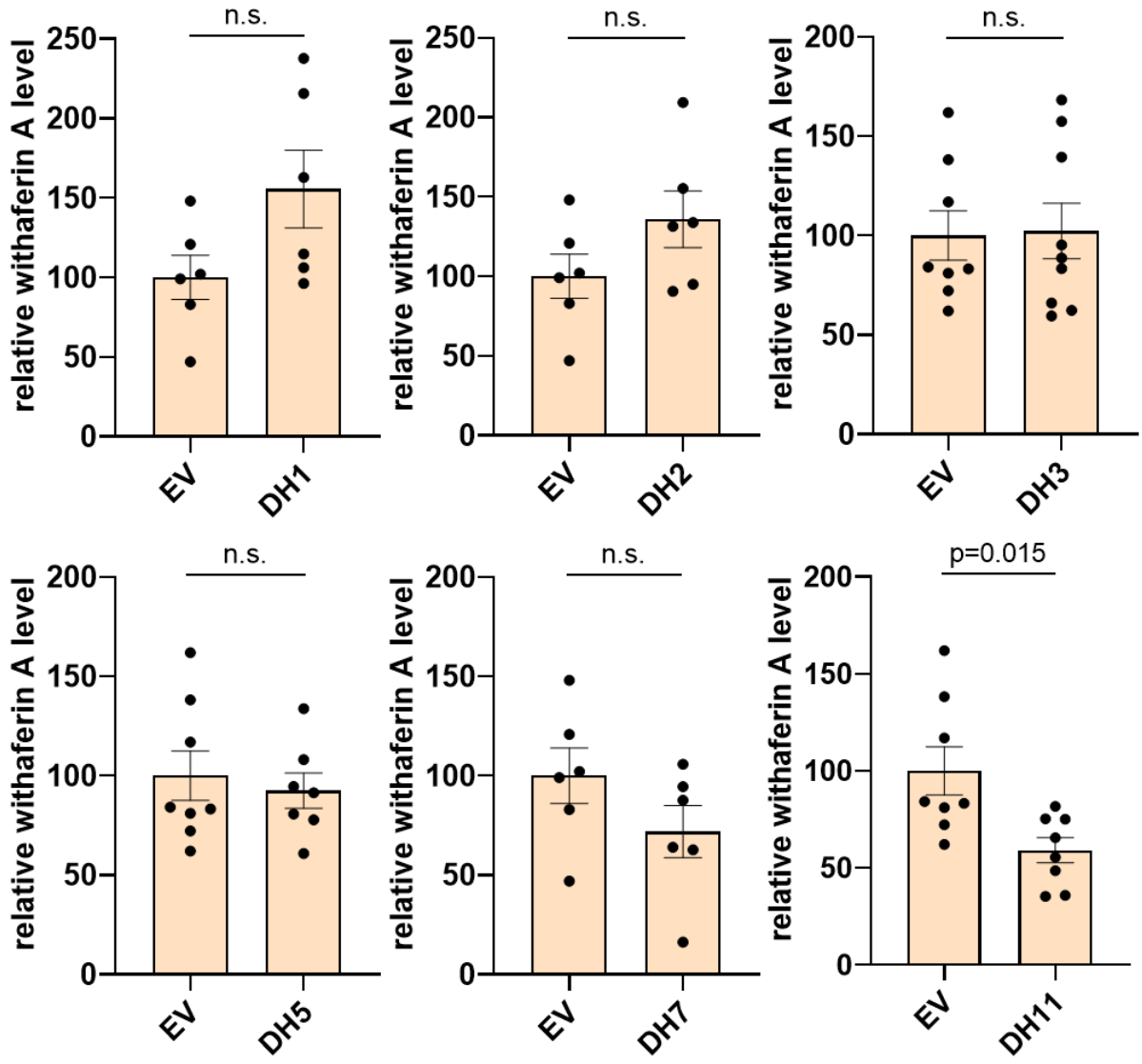


Figure 5.16 Co-infiltration of PDS and dehydratase candidates (DH) shows decreased withaferin A levels for DH7 and DH11 compared to an empty vector control (EV) in bleached leaves. Error bars represent SEM (n=6-9).

5.5 Conclusion

Virus induced gene silencing experiments are the first step in investigation of potential gene candidates for withanolide biosynthesis and will be summarized briefly.

Silencing experiments were tested in two non-model plants, *P. peruviana* and *W. somnifera*, using the positive control phytoene desaturase (PDS). For *P. peruviana* multiple conditions were tested for plant cultivation and agroinfection and systemic spread of photobleaching was only observed after injection of agrosuspension to the stem. The conditions described in the literature could not evoke a photobleaching phenotype. [149] *P. peruviana* seems to have a robust immune system, as infected leaves were quickly abscised and systemic bleaching was hardly observed. *W. somnifera* on the other hand was more susceptible to VIGS, as the first condition tested, that had been reported in the literature, [271] showed the desired bleached phenotype. *W. somnifera* PDS silencing showed a bleached phenotype in 50%-100% of replicates, while PDS silencing in *P. peruviana* exhibited a frequency of up to 40%. As *W. somnifera* showed the desired phenotype in a larger proportion of replicates and individual replicates showed stronger symptoms, this plant was selected for further experiments.

Consequently, the VIGS system was tested for the application to the withanolide biosynthesis, by targeting the 24ISO, the branching point towards withanolides. Here, an average withaferin A (**5**) decrease of 31%-34% was observed and expression levels of the 24ISO decreased by 23% on average. A similar experiment has been reported previously, using a different fragment for VIGS and resulted in ca. 60% decrease in withaferin A (**5**) content and ca. 80% of 24ISO gene expression. [64] It was concluded that silencing of the 24ISO evoked a decrease of withaferin A content, however intensity was lower than described in the literature, potentially due to the different target sequences used.

Furthermore, involvement of cytochrome P450 candidates in withanolide biosynthesis was investigated with VIGS experiments. In chapter 4.2.2, 21 P450 candidates were selected that showed indicators for participation in withanolide biosynthesis. Target silencing sequences for 17 out of 21 candidates could be cloned into pTRV2. Agroinfiltration with those candidates showed significant decrease of the main withanolide withaferin A for five candidates: P450-7, P450-8, P450-14, P450-15 and P450-17. In addition, co-infiltration of those five candidates with PDS was tested in order to direct harvest of silenced tissues towards leaves with stronger VIGS activity. This approach was an adapted version to a protocol using GOI-ChlH fusion products for targeting VIGS harvest. [296,297] However, bleached tissues did not show a stronger

decrease in withaferin A levels compared to green leaves. Therefore it was concluded that GOI silencing was either not effective or PDS and GOI siRNAs spread independently within the plant. Additionally, LCMS chromatograms were analyzed for accumulation of intermediates, but no new compounds were detected.

Lastly, 6 of 12 previously determined dehydratase candidates (chapter 4.3) were cloned and tested in VIGS experiments. A significant decrease of withaferin A (**5**) content was observed for DH11, and further DH7 showed a downwards trend, albeit not statistically significant. Repetition of this experiment did not show a clear silencing trend for any candidate. Therefore all 6 candidates should be tested in *in vitro* experiments again.

In summary, VIGS is a valuable technique that, once established, allows relatively fast screening of many candidates. However, processing of large quantities of candidates and replicates is labor intensive and natural fluctuations of metabolite levels often exacerbate statistical analysis of VIGS effects. Furthermore, the detection of accumulated substrates is strongly dependent on the degree of silencing, the respective target compound, the biosynthetic pathway investigated (e.g. alternative metabolization of intermediates) and the methods employed. Therefore, any candidate that was selected in this chapter needs to be further investigated with *in vitro* experiments to allow further conclusions on their activity.

6 Conduction of enzyme assays with yeast and plant hosts

6.1 Introduction

While silencing experiments were conducted in chapter 5 and provided a narrowed pool of gene candidates for withanolide biosynthesis, enzyme assays of said gene candidates should be performed in this chapter. Now as this work focuses on the multiple oxidation reactions in withanolide biosynthesis that are assumed to be performed by cytochrome P450 enzymes, the chosen expression system has to be able to cater to this protein class.

P450s are typically membrane bound, and depend on a heme group being incorporated at the active site. [306] Furthermore a cytochrome P450 reductase (CPR) is needed to transfer electrons to the P450 in its reaction cycle. [307] As prokaryotic hosts such as *Escherichia coli* do not provide the required membrane structure, folding machinery or reductase partners, [308–311] the expression in such hosts can only be achieved by employing several strategies such as N-terminal sequence cleavage in combination with chaperone co-expression. [310–312] The observed yields of active enzyme are often quite low and require optimization. An alternative is the expression of P450 gene candidates in eukaryotic hosts such as insect cells, yeast or plants as they provide the necessary structures and machinery. For those hosts numerous examples of plant *P450* expression can be found in the literature. [313,69,314,315]

As optimizing membrane protein expression in *E. coli* is a complicated process and expression of plant genes in the yeast *Saccharomyces cerevisiae* [313,316–318] or the plant host *Nicotiana benthamiana* [319–321] generally have more examples for application, the latter were chosen for expression experiments.

Besides the selection of a host system, another issue is the available substrate. The last known intermediate in withanolide biosynthesis is 24-methyldesmosterol (**3**). In chapter 5, a withaferin A (**5**) decrease was observed in *Withania somnifera* after individual silencing of five different P450 candidates. However, no accumulation of intermediates was observed and therefore the exact substrates are unknown. Therefore 24-methyldesmosterol (**3**) represents the only verified and available substrate and *in vitro* assays will be tested with this compound.

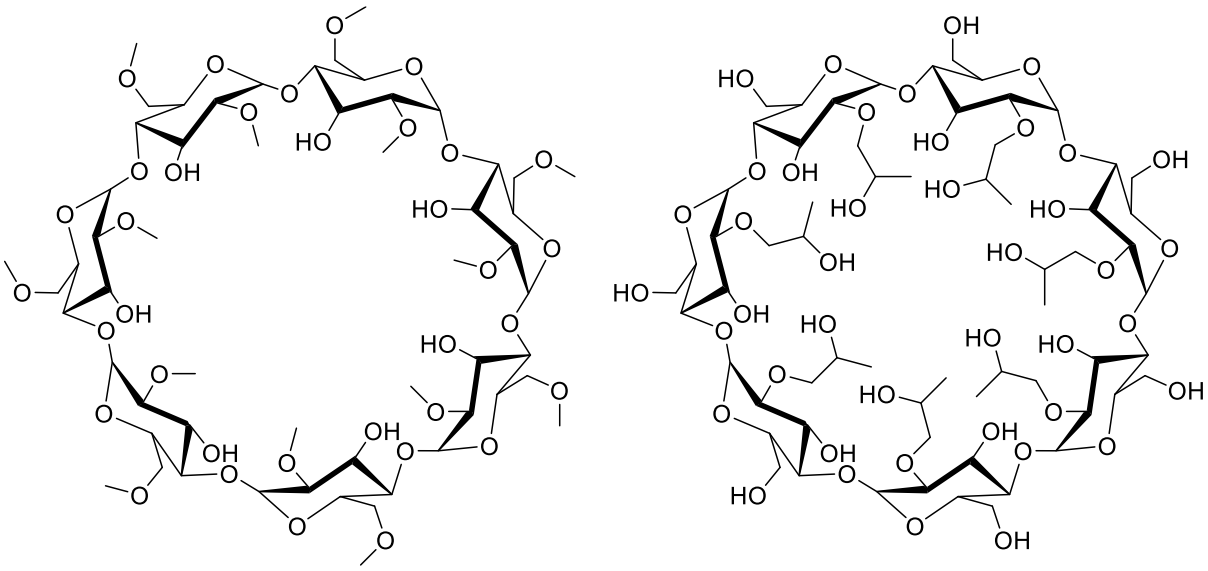
6.2 Solubilization of hydrophobic substrates

Chapter 3 described the attempt of generating a 24-methyldesmosterol (**3**) producing strain. However, no production of **3** was observed in this work and a yeast strain generated by Dr. Karan Malhotra with a CRISPR based approach was not available at the time of this experiments. Instead, **3** for this work was obtained via synthesis by Dave Biedermann. In addition to the availability, the solubility of **3** provides another issue. As **3** is strongly hydrophobic, it can only be dissolved in organic solvents such as ethyl acetate or 100% methanol. Since organic solvents interfere with enzyme folding and activity, an alternative has to be found to join substrate and enzyme in the same medium.
In the literature multiple strategies can be found that describe solubilization of hydrophobic substrates in aqueous solutions. Examples include solubilization with Triton-X, [322,323] Tween, [324,325] solution with acetone [326] or addition of cyclodextrins. [327–329] Due to the simplicity of protocols described, the addition of cyclodextrins was tested and furthermore solubilization with Tween was considered in this work.

Since the amount of available 24-methyldesmosterol (**3**) was limited, initial solubility experiments were conducted with other sterol-type compounds available. The obtained substrate solutions were evaluated for two criteria: the concentration achieved and the background generated by the additive, as assays should be analyzed via LCMS for product formation.
Solubilization with Tween 80 was briefly tested, as microsome assays with final concentrations of 0.1% Tween 80 are described in the literature, with higher concentrations exhibiting a sharp decline in activity. [324,325] However, preparation of a stock solution with 211 µM cycloartenol (**6**) and 0.2% Tween 80 in HEPES buffer (pH 7.5, see chapter 8.2.3) did not show complete clearing of the solution. Analysis with LCMS concluded that small amounts of **6** were dissolved, but the chromatogram showed many impurities that would exacerbate identification of newly accumulated compounds. As the addition of cyclodextrins showed more promising results, solubilization with Tween 80 was not further pursued.

Cyclodextrins are cyclic carbohydrates comprised of 6-8 glucose molecules (α-, β- or γ-cyclodextrins, respectively), that form a cylindrical shape. [330] On the inside, a hydrophobic cavity is formed that can accommodate hydrophobic compounds. [331] While unmodified β-cyclodextrin is unable to solubilize substantial amounts of cholesterol (**15**) in water, solubility

of **15** in aqueous solutions can be increased by the addition of substituents on the hydroxyl groups on glucose units of the cyclodextrin. [332] In the case of methyl-β-cyclodextrin (**113**) two hydroxyl-groups per glucose unit are methylated (Figure 6.1).



methyl-β-cyclodextrin (**113**) (2-hydroxypropyl)-β-cyclodextrin (**114**)

Figure 6.1 Structure of methyl-β-cyclodextrin (**113**) and (2-hydroxypropyl)-β-cyclodextrin (**114**).

For this work methyl-β-cyclodextrin (MCD, **113**) and (2-hydroxypropyl)-β-cyclodextrin (HPCD, **114**, Figure 6.1) were tested, as they are commercially available in sufficient amounts for testing. In general, the aim was to generate a concentration of 100-200 µM of substrate in assays. All solubilization experiments mentioned are listed in Table 6.1.

Table 6.1 Substrate concentration determined for different solubilization experiments.

Substrate	Concentration [w/v]	Cyclodextrin	concentration substrate after filtration
cholesterol	20.00%	HPCD	below detection
cholesterol	33.00%	MCD	75 mM
cholesterol	0.23%	MCD	129 µM
cholesterol	0.45%	MCD	266 µM
cholesterol	0.50%	MCD	518 µM
cholesterol	0.60%	MCD	329 µM
24-methyldesmosterol	0.23%	MCD	below detection
24-methyldesmosterol	5.00%	MCD	45 µM
24-methyldesmosterol	10.00%	MCD	92 µM
β-sitosterol	0.45%	MCD	below detection
β-sitosterol	5.00%	MCD	1.13 mM

MCD: methyl-β-cyclodextrin; HPCD: (2-hydroxypropyl)-β-cyclodextrin

At first, solubilization of cholesterol (**15**) was tested by adding 33% (w/v) of MCD (**113**) and 20% (w/v) of HPCD (**114**) to HEPES buffer (pH 7.5) and further adding a surplus (50 mg/mL) of cholesterol (**15**). The suspensions were stirred overnight at room temperature and undissolved cholesterol (**15**) was separated from the suspension by filtration. Afterwards, the aqueous solution was mixed with MeOH for LCMS analysis (see chapter 8.2.4). The MCD (**113**) solution was clear after filtration and yielded 75 mM, while the HPCD (**114**) solution did not show substantial clearing and yielded cholesterol amounts below the detection limit.

As a final concentration of 100-200 µM was desired, the cholesterol-MCD stock solution should be further diluted. However, upon adding further HEPES buffer to the mixture, the dissolved cholesterol precipitated. Several strategies such as heating and stepwise addition of HEPES buffer did not re-dissolve the precipitated cholesterol (**15**).
Therefore, it was concluded that the substrate has to be dissolved in the final concentration to avoid subsequent dilution steps. For planned infiltration of 24-methyldesmosterol (**3**) in *N. benthamiana* this is no problem, as the substrate solution is directly infiltrated into the leaves. For microsome assays however, the substrate solution has to be added to other reaction components such as microsomes and the cofactor NADPH. It was reasoned that microsome assays will be stirred overnight, so that precipitated substrate in the reaction mixture could re-dissolve in the MCD (**113**) complexes and consequently reach the enzyme.
Experiments with low concentrations of MCD (**113**) showed that the observed amount of cholesterol (**15**) was depending on the MCD (**113**) concentration when **15** was added in surplus (Table 6.1). A maximum of 518 µM cholesterol (**15**) was dissolved when adding 0.5% MCD (w/v) to the HEPES buffer. However, solubilization experiments were not entirely consistent, as an independent experiment with 0.6% MCD (**113**) exhibited a cholesterol (**15**) concentration of 329 µM.

Another compound that was tested was β-sitosterol (**14**), as it was a substrate for planned control assays with enzymes with known function. Here 0.45% MCD (**113**) in HEPES buffer led to no substantial solubilization of β-sitosterol (**14**). This shows that solubility of sterol type compounds with cyclodextrins is not equal, as a cholesterol (**15**) solution with the same amount of MCD (**113**) exhibited a concentration of 266 µM. An increase of MCD (**113**) concentration to 5% (w/v) led to solubilization of 1.13 mM β-sitosterol (**14**).

When testing solubilization of 24-methyldesmosterol (**3**), it was noted that this compound exhibited lower solubility with MCD (**113**) than cholesterol (**15**). Initial experiments with **3** in HEPES buffer with 0.23% MCD (**113**) did not detect noteworthy amounts of 24-methyldesmosterol (**3**), while the same amount of MCD (**113**) led to solubilization of 129 µM cholesterol (**15**). An increase of MCD (**113**) concentration to 5% (w/v) while stirring at 40 °C overnight resulted in 45 µM of 24-methyldesmosterol (**3**) in the buffer solution. A second experiment using 10% MCD (**113**) resulted in 92 µM of dissolved **3**.

While dissolving sterol-type compounds (e.g., cholesterol, **15**) in aqueous solutions using cyclodextrins has been reported before, it was confirmed here that this technique works for 24-methyldesmosterol (**3**) and that this compound can therefore be used in enzyme assays. However, the observed concentrations are strongly dependent on the target compound and the concentration of cyclodextrin used. In the following sections, these substrate solutions will be tested for their applicability in microsome and leaf disk assays.

6.3 Microsome Assays

Chapter 5.3 determined P450 candidates of which silencing led to decrease of the main withanolide withaferin A (**5**). Now those enzymes should be further investigated in enzyme assays. As exogenous sterol uptake only occurs under anaerobic conditions in yeast, [194,195] simply transforming a yeast strain with the gene candidates and supplying 24-methyldesmosterol (**3**) was excluded as a method for investigation. Instead, isolated microsomes can be used for *in vitro* assays.

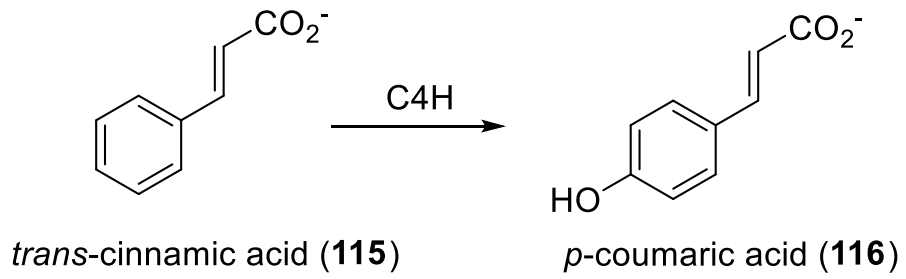
When performing heterologous expression of P450s in yeast the interaction of the P450 with endogenous cytochrome P450 reductases (CPR) is often insufficient. Consequently, product yields can be improved by adding a plant CPR to the host organism. [333] The WAT11 [226] strain contains an *Arabidopsis thaliana* CPR that replaced the endogenous gene in the genome.[xvi] When using this strain, expression of *P450* gene candidates can then be conducted from the pYES2 plasmid (Invitrogen). In order to investigate whether the planned microsome assay system is functioning, positive controls were tested in the following sections.

[xvi] The WAT11 strain was kindly provided by John D'Auria (IPK Gatersleben)

6.3.1 Activity was detected for the positive control cinnamate 4-hydroxylase

Performing microsome assays requires successful production, isolation and activity of enzymes. While successful conduction of those steps can be confirmed by techniques such as western blot or spectral analysis of P450 enzymes, [334] those methods are difficult to establish and provide a common source of error if not routinely performed. Consequently, a P450 positive control is an easier approach to confirm functionality of the techniques used along the processing steps of this system.

The cinnamate 4-hydroxylase (C4H) is a cytochrome P450 enzyme that catalyzes the hydroxylation of *trans*-cinnamic acid (**115**) to *p*-coumaric acid (**116**) in the phenylpropanoid pathway in plants (Scheme 6.1). [316,335] Due to its involvement in plant primary metabolism, C4H genes can be amplified from any plant available in this institute. Furthermore, substrate and product are commercially available, easily detectable by HPLC and water soluble, which excludes enzyme accessibility issues in the assays. Lastly, microsome assays of this P450 have been reported before. [313,316–318] Therefore this enzyme was chosen as a first positive control.



Scheme 6.1 The cinnamate 4-hydroxylase (C4H), used as a control reaction in this work, catalyzes the hydroxylation of *trans*-cinnamic acid (**115**) to *p*-coumaric acid (**116**). [316,335]

The *Arabidopsis thaliana* C4H CDS (Acc. nr.: AM887632.1) was amplified from cDNA obtained from *A. thaliana* plants from the plant collection of the institute of botany. Furthermore, a tBLASTn search was conducted with the AtC4H against the *Withania somnifera* transcriptome [157] to determine the potential C4H of this plant. This yielded transcript k69_Locus_4488_Transcript_3_1 as the best hit (87% aa identity), further termed WsC4H. In this case the Oases assembly by Prof. Dr. Jakob Franke of Senthil *et al.* transcriptome data [157] had to be used, as the provided Trinity assembly gave a truncated sequence (for description of transcriptome assemblies see chapter 4.2). The transcript of interest was amplified using cDNA from *W. somnifera* as a template. DNA assembly of C4H genes into pYES2 was conducted using Gibson assembly (In-Fusion®, Takara). The obtained plasmids were verified by colony PCR and DNA sequencing (Microsynth, Göttingen). Successful transformation into yeast was

confirmed by colony PCR. Enzyme production and microsome isolation was performed as described in chapter 8.2.

Microsome assays in the literature used different buffers and concentrations, suggesting activity of this enzyme over a wider range of conditions. [313,318] Therefore a microsome assay setup with HEPES buffer was tested, as it was readily available (setup see chapter 8.2.3). Analysis of C4H assays was conducted with LCMS UV traces at 345 nm and 270 nm and was compared to a no enzyme (NE) control and *trans*-cinnamic acid (**115**) and *p*-coumaric acid (**116**) standards (depiction of **116** at 345 nm, **115** at 270 nm, Figure 6.2). In both C4H assays a peak is observed that matches the retention time (4.3 min) and UV spectrum of a *p*-coumaric acid (**116**) standard and UV spectra reported in the literature. [318] This peak is not present in the wildtype control. It has to be noted that the peaks for both substrate (**115**) and product (**116**) in the WsC4H sample elute slightly earlier, which is due to the samples being measured at a different time, resulting in a slight shift of retention times.

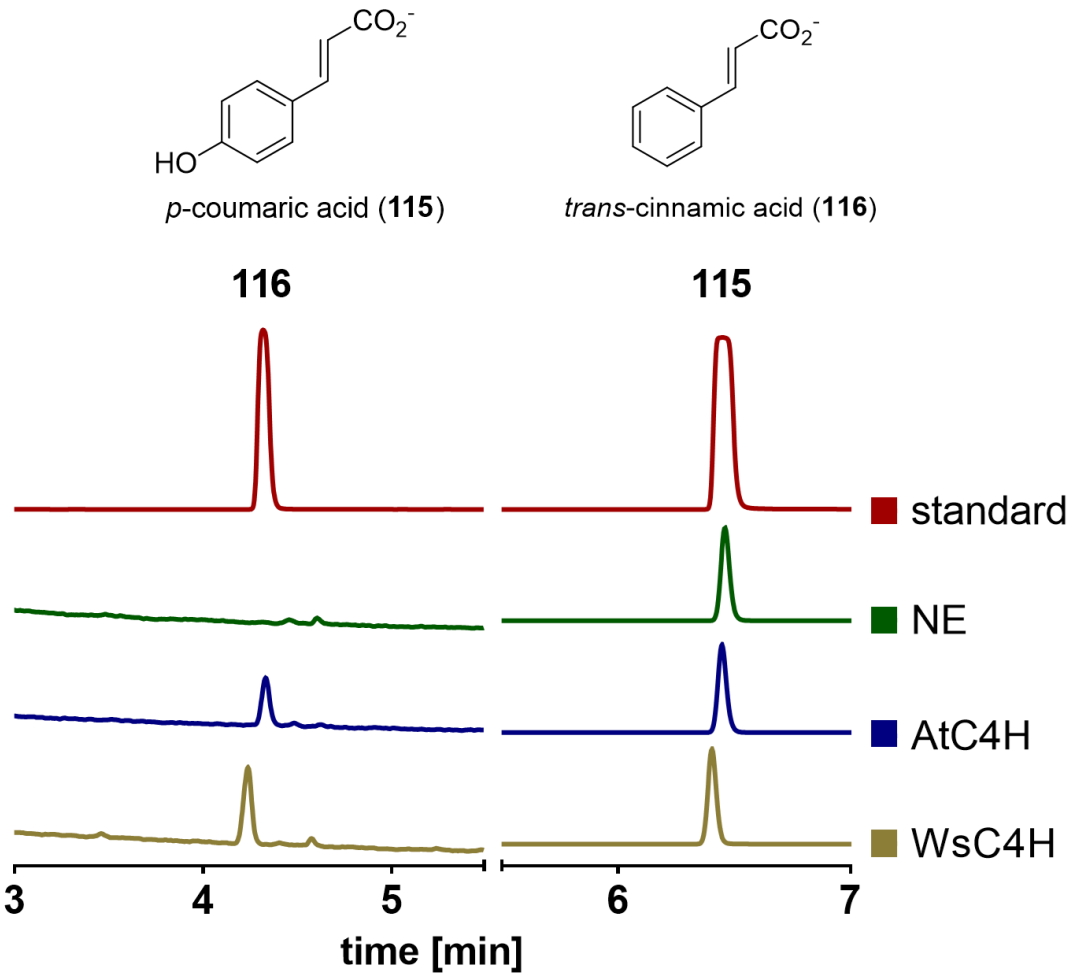
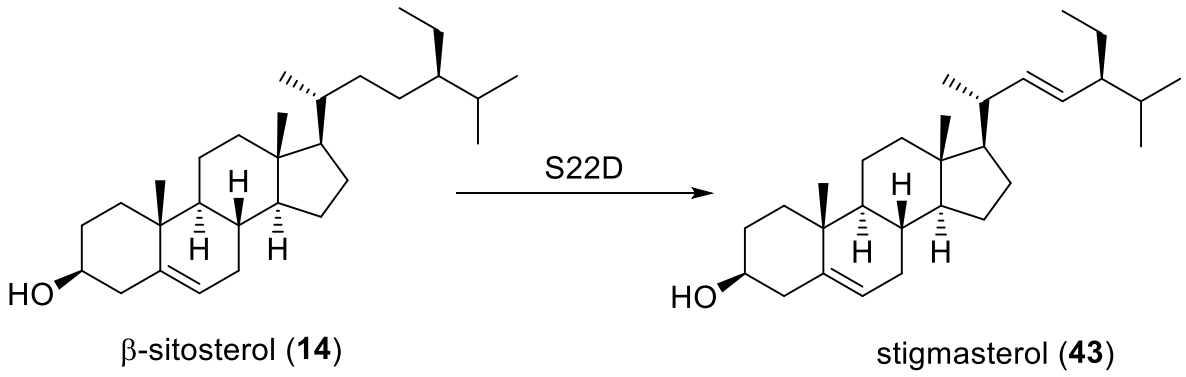


Figure 6.2 LCMS chromatograms of C4H microsome assays show formation of the expected product, *p*-coumaric acid **(116)**, thus proving the general functionality of the microsome assay setup; traces at 3-5.5 min: UV 345 nm; 5.5-7 min: UV 270 nm for visibility of the respective compound.

It was concluded that the general strategy of enzyme production, microsome isolation and the chosen assay setup work for water soluble substrates. However, the strategy of dissolving hydrophobic substrates using cyclodextrins still had to be investigated for its compatibility with microsome assays.

6.3.2 No activity was detected for known enzymes sterol-22-desaturase and cholesterol-22-hydroxylase

The previous section demonstrated that the general process of microsome assays with the methods employed worked. However, the P450 tested (AtC4H) is active on *trans*-cinnamic acid (**115**), a water-soluble compound. As the substrate for assays with P450 candidates in withanolide biosynthesis will be 24-methyldesmosterol (**3**), a hydrophobic compound, this system should now be investigated for compatibility with hydrophobic substrates. Chapter 1.1 described the solution of hydrophobic compounds such as cholesterol (**15**), β-sitosterol (**14**) and 24-methyldesmosterol (**3**) in HEPES buffer by addition of methyl-β-cyclodextrin (**113**). Now this strategy should be employed to test enzymes with known function and hydrophobic substrates. The sterol-22-desaturase (S22D) catalyses C-22 elimination reactions on multiple substrates, for example the conversion of β-sitosterol (**14**) to stigmasterol (**43**, Scheme 6.2). [69] Both substrate and product are commercially available and detection in assays can therefore easily be optimized. However, production of this enzyme has not been reported in yeast, but in insect cells. [69] Therefore, it is unknown whether expression in yeast is possible.

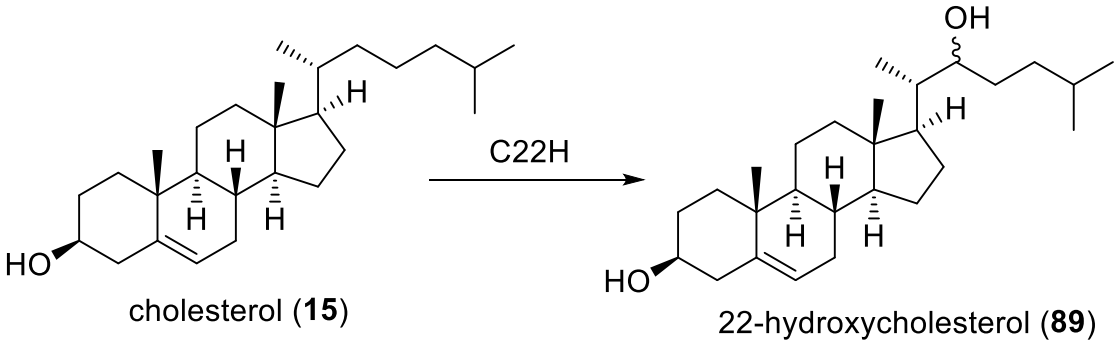


Scheme 6.2 The sterol-22-desaturase converts the conversion from β-sitosterol (**14**) to stigmasterol (**43**). [69]

The *Arabidopsis thaliana sterol-22-desaturase* (*AtS22D*) described in the literature (accession nr.: AB219423.1) [69] was amplified from cDNA obtained from *A. thaliana* plants from the plant collection of the institute of botany. Furthermore, a tBLASTn search was conducted with *Withania somnifera* transcriptome data, [157] with both the Trinity and Oases assembly. The best

hit from the trinity assembly was TRINITY_DN10503_c0_g1_i7, further termed WsS22D1 (60% aa identity, 96% query coverage), while the Oases assembly provided k69_Locus_21786_Transcript_2_1, further termed WsS22D2 (60% aa identity, 96% query coverage). As the predicted sequence obtained from the Oases assembly was 40 aa longer than the sequence obtained from the Trinity assembly, both sequences were selected for cloning. *W. somnifera* inserts were obtained using previously generated cDNA as a template.

Another control enzyme acting on sterol-type substrates is the cholesterol-22-hydroxylase reported in steroidal glycoalkaloid biosynthesis. [70,258] This enzyme catalyzes the hydroxylation of cholesterol (**15**) to 22-hydroxycholesterol (**89**, Scheme 6.3). [70,258] However, *in vitro* assays reported in the literature again used insect cells. [258,314] The *Solanum tuberosum cholesterol-22-hydroxylase* (*StC22H*) described in the literature (accession nr.: AB839753.1) [258] was amplified from cDNA obtained from *S. tuberosum* sprouts.



Scheme 6.3 The cholesterol-22-hydroxylase catalyzes hydroxylation of cholesterol (**15**) to 22-*R*-hydroxycholesterol or 22-*S*-hydroxycholesterol (**89**). [70,258]

DNA assembly of S22D and C22H genes into pYES2 was conducted using Gibson assembly (In-Fusion®, Takara). For WsS22D1, WsS22D2 and StC22H, positive colonies were detected, but not for AtS22D. Multiple attempts of cloning could not obtain the pYES2 AtS22D plasmid. The obtained WsS22D and StC22H plasmids were further verified by DNA sequencing (Microsynth, Göttingen). Transformation into WAT11 was confirmed with colony PCR. Subsequent enzyme production and microsome isolation was performed as described in chapter 8.2.3. The microsome assay setup was similar to the protocol described in the previous section. However, the respective substrate was dissolved in HEPES buffer (pH 7.5) using MCD (**113**).

For the WsS22D assays, β-sitosterol (**14**) was dissolved with 5% MCD (**113**) in the stock solution, yielding 1.13 mM (final concentration in assay 565 µM). Figure 6.3 shows exemplary TIC traces of Ws22D assays, a no enzyme (NE) control and β-sitosterol (**14**) and stigmasterol

(**43**) standards. The substrate β-sitosterol (**14**) can be identified at 4.8 min in all assay samples. The product stigmasterol elutes at 4.4 min (**43**) and even though a peak is visible in the WT and WsS22D1 sample, the mass spectrum does not match the standard. It was concluded that no conversion of **14** to **43** was observed.

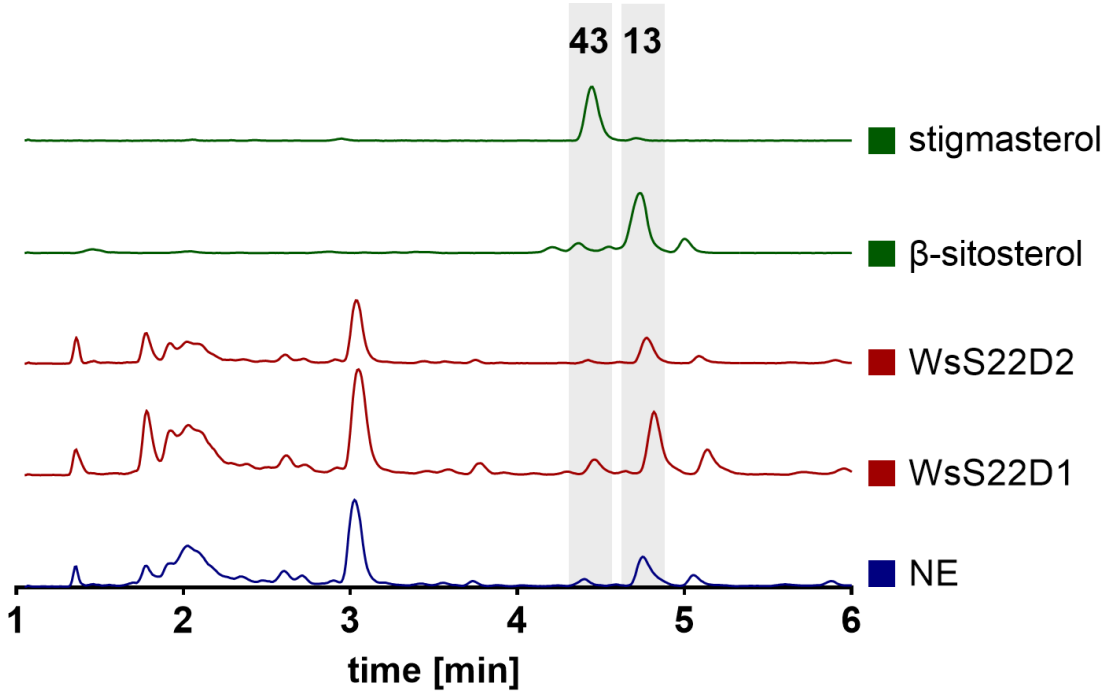


Figure 6.3 LCMS chromatograms (TIC) of S22D microsome assays show no formation of the product stigmasterol (**43**); NE: no enzyme control.

This could be due to multiple reasons. First the sequences determined as WsS22D via BLAST might not be the orthologs of this enzyme in *Withania somnifera*. Alternatively, the enzymes might be inactive.

Before drawing further conclusions, the C22H assays were analyzed, so see whether activity could be observed here. For this experiment cholesterol (**15**) was dissolved with 0.5% MCD (**113**), yielding a concentration of 518 µM in the stock solution (259 µM in the assay). Here, no standard was available for the product 22-hydroxycholesterol (**89**) at the time of assays conduction. Therefore, chromatograms were investigated for masses 403 ($M+H^+$) and 385 ($M-H_2O+H^+$), but no difference to the no enzyme (NE) control was detected. A comparison of TIC traces of C22H and a WT is depicted in Figure 6.4. It was concluded that StC22H did not show activity either.

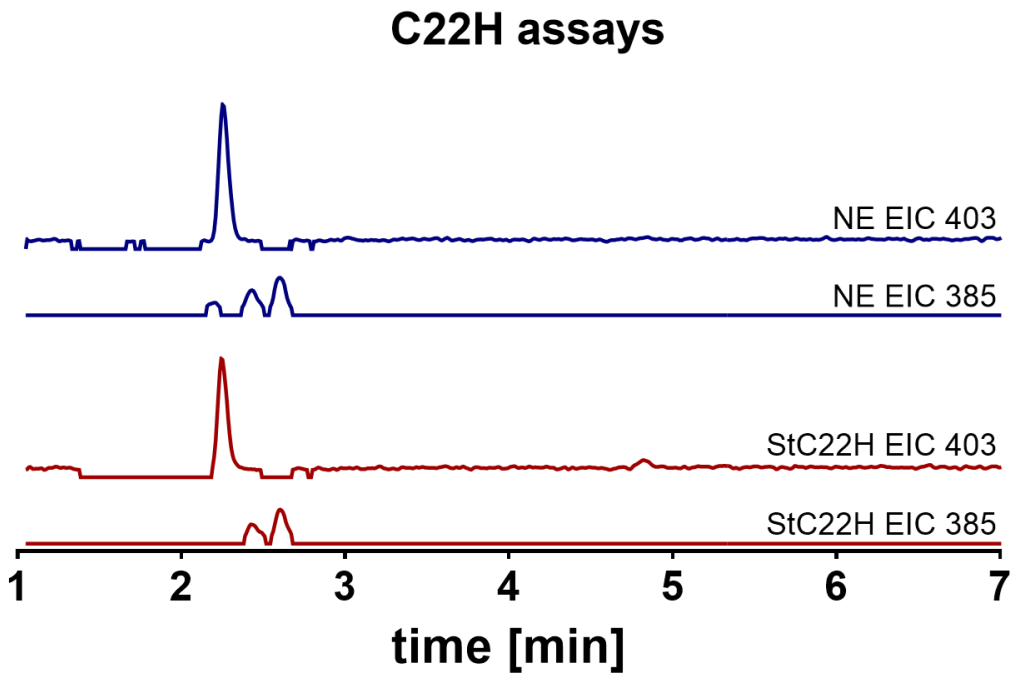


Figure 6.4 LCMS chromatograms (EIC) of a StC22H microsome assay shows no accumulation of new compounds compared to a wild type (WT) control. EIC 403: $[M+H]^+$ of 22-hydroxycholesterol (**89**); . EIC 385: $[M-H_2O+H]^+$ of 22-hydroxycholesterol (**89**).

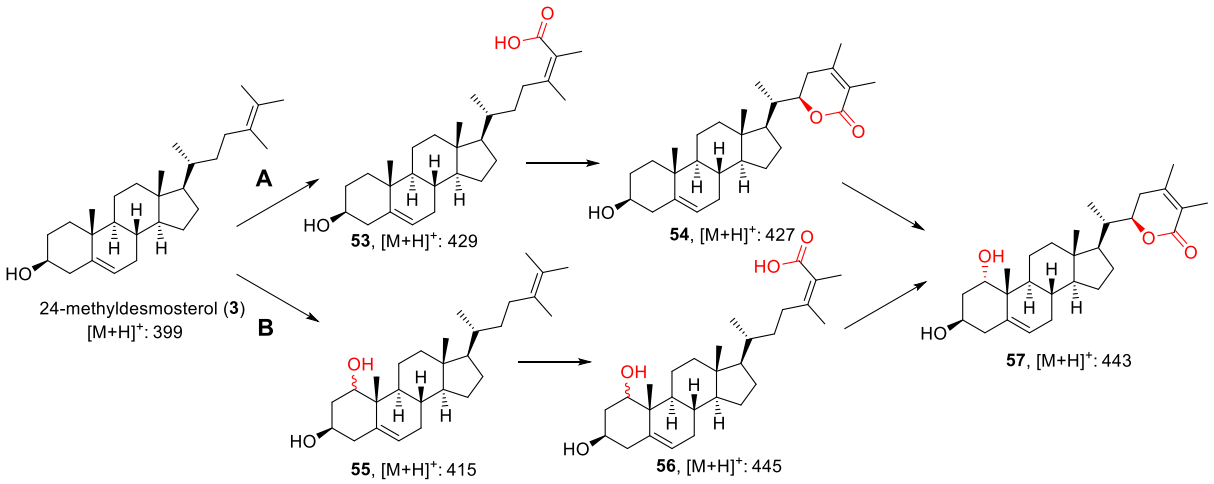
In the second experiment, a sequence was used that had been published in the literature, [258] excluding the possibility of a wrong ortholog. It is possible that both WsS22D and StC22H are inactive in this setup. Experiments in the literature were conducted with insect cell microsomes for both enzymes and it is unclear whether expression in yeast was tested but no activity was observed. [69,258,314] However, experiments with a C4H (see section 6.3.1) showed that the general setup of microsome assays in this work was functioning. Another reason for the lack of product formation in S22D and C22H assays could be the employed strategy for substrate solving. It is unclear whether the substrate is able to leave the cyclodextrin complexes and access the enzyme. Furthermore, initial experiments with cyclodextrin solubilization showed that substrate precipitated upon adding buffer to the substrate solution. As microsome stock solution and the cofactor NADPH have to be added to the substrate solution in order to perform assays, it is possible that the substrate precipitated. Turbidity that would indicate precipitation of the substrate was not observed when preparing the microsome assays, however, substrate concentration might have been too low for visible turbidity.

It was concluded that further investigation of those positive controls was not productive, as the reason for the lack of product formation could not be resolved. Instead, microsome assays with the P450 candidates for withanolide biosynthesis should be tested.

6.3.3 Microsome assays with P450 candidates did not show accumulation of new compounds

As tests with known enzymes sterol-22-desaturase and cholesterol-22-hydroxylase did not result in accumulation of the respective product it was unclear whether the tested enzymes were inactive, or substrate dissolved with cyclodextrins was unable to access the enzyme. Nonetheless, microsome assays should be tested with withanolide biosynthesis P450 candidates. In chapter 5.3 silencing experiments determined five P450 candidates that, when targeted for silencing, evoked decrease of the main withanolide withaferin A (**5**). Here, three of them were targeted for microsome assays. Candidates P450-7, P450-8 and P450-17 were amplified from *W. somnifera* cDNA. DNA assembly with pYES2 was performed using Gibson assembly (In-Fusion®, Takara). All plasmids were verified with colony PCR and DNA sequencing (Microsynth, Göttingen). Yeast transformation was confirmed with colony PCR. Subsequent enzyme production, microsome isolation and assay conduction was performed as described in chapter 8.2.3. As a substrate 24-methyldesmosterol (**3**) was dissolved in HEPES buffer (pH 7.5) using 5% (w/v) MCD (**113**). Here residual precipitated substrate was not filtered, to provide as much **3** as possible in the assay. In total, 68 µM **3** were detected in a substrate only control that was processed identically to the microsome assays.

Chapter 1.3.3 describes the hypothetical route towards withanolides. Here, oxidation of 24-methyldesmosterol (**3**) at either the side chain (**53** then **54** then **57**, pathway A, Scheme 6.4) or the A-ring is proposed (**55** then **56** then **57**, pathway B, Scheme 6.4). However, the exact order of functionalizations is unknown, and it is possible, that the pathway takes a different road. Furthermore, it is possible that reactions that require multiple oxidation steps (e.g., oxidation of the side chain) are performed by multiple enzymes and only intermediates (e.g., hydroxylated or oxo- side chain) can be observed when processed with a single P450. As no oxidized 24-methyldesmosterols are commercially available, enzyme assays cannot be evaluated by comparison to a standard. Therefore, LCMS chromatograms should be analyzed for the respective masses of potential products. In general, oxidized 24-methyldesmosterol would exhibit a sum formula of $C_{28}H_{46}O_2$ for a hydroxylated product (e.g. **55**, $[M+H]^+$: 415), $C_{28}H_{44}O_2$ for a ketone/aldehyde ($[M+H]^+$: 413), $C_{28}H_{44}O_3$ for a carboxylic acid (**53**, $[M+H]^+$: 429) or $C_{28}H_{42}O_3$ for a lactone (**54**, $[M+H]^+$: 427).



Scheme 6.4 Possible routes for oxidation of 24-methyldesmosterol (**3**).

The LCMS gradient used for analysis of P450 microsome assays was adjusted so that a more polar product should be detectable. Figure 6.5 shows TIC and EIC (413, 415) traces of P450-7 and P450-8 in comparison to a no enzyme control (NE). In the conducted assays no difference to the NE control was detected.

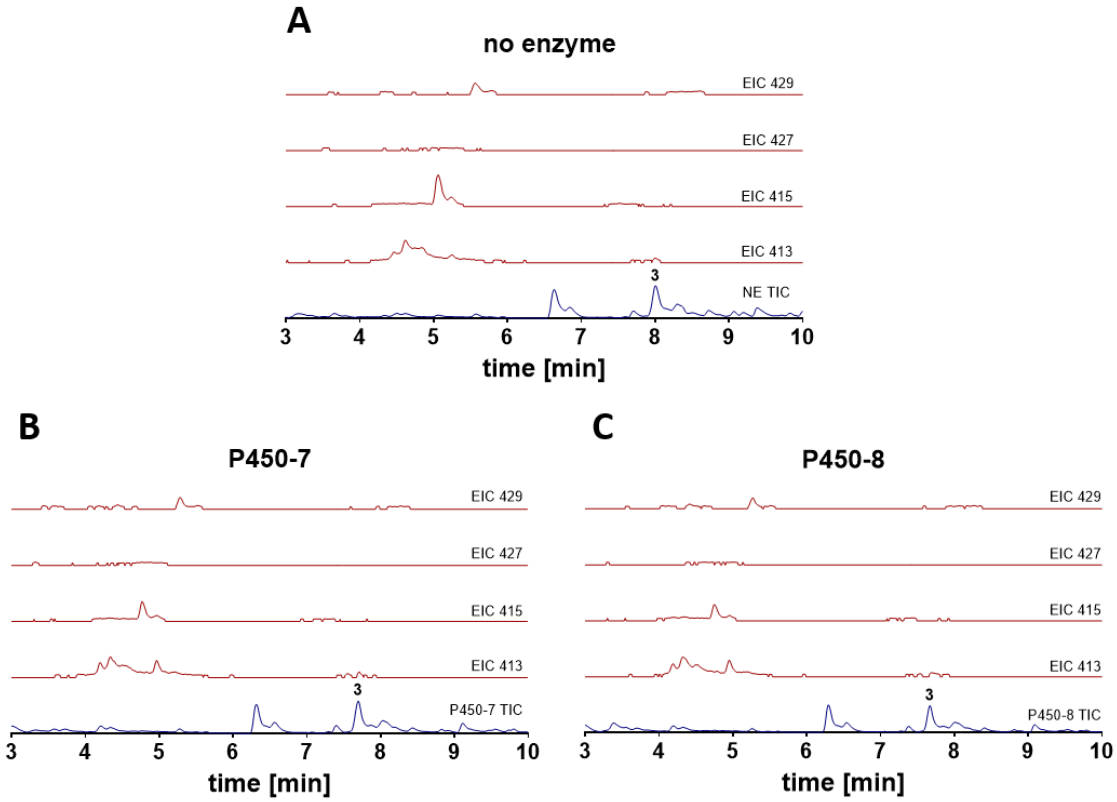


Figure 6.5 Microsome assays with P450-7 and P450-8 did not show accumulation of new compunds compared to a no enzyme (NE) control. LCMS traces of a NE microsome assay with EIC 413, 415, 427 and 429 (A); LCMS traces of a P450-7 microsome assay with EIC 413, 415, 427 and 429 (B); LCMS traces of a P450-8 microsome assay with EIC 413, 415, 427 and 429 (C); All EIC traces were amplified by 100 to overlay with TIC traces.

It was concluded that microsome assays with P450 candidates did not lead to the accumulation of new compounds. The AtC4H control was specifically selected as microsome assays with yeast were reported before. [313,316–318] For sterol-type compounds only one example used yeast microsomes, [325] but did not examine the activity of P450 enzymes. Most publications used other eukaryotic hosts, such as insect cells or mice. [69,258,324,336] It is unclear whether yeast as a host is potentially unable to provide critical unknown factors for the conduction of P450 assays targeting sterol-type compounds.

The other explanation is an accessibility issue of the substrate for the enzyme. While addition of cyclodextrins for microsome assays has been reported before, [336] most protocols for microsome assays with sterol-type substrates use solubilization methods with Tween [324,325] or do not describe any additional solution method, but seem to simply add the substrate to the assay. [69,258] Solubilization with Tween 80 was tested for cycloartenol, but the achieved substrate concentrations were very low, compared to the background created by the additive in LCMS chromatograms. Microsome assays with P450-17 and cycloartenol (**6**) dissolved with 0.1% Tween 80 were tested, but no substantial concentrations of **6** were achieved (see chapter 1.1) and no product formation could be detected (data not shown).

As the reason for the lack of product formation cannot be determined, it was concluded that further attempts at microsome assays for P450 candidates were not productive. Instead, a different expression system, the common plan host *Nicotiana benthamiana* should be tested.

6.4 Leaf Disk Assays

Another option for heterologous expression of P450s is the commonly used plant host *Nicotiana benthamiana*. When investigating plant biosynthetic pathways this host has several advantages compared to other systems. It provides mRNA/protein processing and compartmentalization similar to the plant of interest and necessary co-enzymes (e.g. CPR) and metabolic precursors are present. [337] However, while yeast accumulates mainly ergosterol (**16**), plant hosts such as *N. benthamiana* produce multiple phytosterols, creating a more complex background and exacerbating analysis of sterol-type compounds.

In this chapter *P450* and *DH* candidates were transiently expressed in *N. benthamiana* to investigate whether activity on the respective substrate can be observed. Instead of infiltrating constructs and substrates into whole leaves, a different approach was tested. An unpublished protocol of the O'Connor group (Max Planck Institute for Chemical Ecology, Jena) describes a leaf disk assay, in which sections of GOI-infiltrated leaves are incubated in a substrate

solution. The advantage of this protocol is the decreased need for substrate, as smaller volumes are sufficient for the incubation of leaf disks compared to infiltration into whole leaves.

In order to perform transient expression of genes of interest in *N. benthamiana*, the expression vector pHREAC [338] was selected. This vector is based on the pEAQ-HT [123] plasmid but contains synthetic 5' and 3'UTRs that increase recombinant protein yield. [338]
The pHREAC plasmid enables cloning via Golden Gate assembly (BsaI restriction), a technique that should be tested, as Gibson assembly showed low assembly frequency in this group. To enable comprehensive investigation, all previously selected candidates (chapter 4) should be cloned into a plant expression plasmid. For candidates that contained a BsaI restriction site an alternative cloning into pEAQ-HT was planned (In-Fusion®, Takara). Inserts were generated via PCR from cDNA, obtained from *Withania somnifera* aerial tissue RNA. Assembly of pHREAC-GOI constructs was achieved in 14 out of 15 of tested inserts (see Table 6.2). The obtained plasmids were verified with colony PCR and DNA sequencing (Microsynth, Göttingen). For gene candidates that needed to be cloned into pEAQ-HT, no integration was observed after multiple attempts. As four of the five P450 candidates that showed withaferin A decrease in chapter 5.3 were cloned into pHREAC, troubleshooting of pEAQ-HT cloning was postponed. Obtained plasmids were transformed into *A. tumefaciens* strain GV3101 using electroporation, as this strain is generally used for agroinfection of *N. benthamiana* in this group. Successful agrotransformation was confirmed with colony PCR.

Table 6.2 Cloning result of P450 and DH candidates for transient *Nicotiana benthamiana* expression.

Plasmid name	Cloning result
pHREAC-P450-1	verified
pEAQ-P450-3	no positive colony PCR
pHREAC-P450-4	verified
pEAQ-P450-5	no positive colony PCR
pHREAC-P450-6	verified
pHREAC-P450-7	verified
pHREAC-P450-8	verified
pHREAC-P450-9	no positive colony PCR
pHREAC-P450-10	verified
pHREAC-P450-11	verified
pHREAC-P450-12	verified
pHREAC-P450-13	no PCR product
pEAQ-P450-14	no positive colony PCR
pHREAC-P450-15	verified
pHREAC-P450-17	verified
pEAQ-P450-19	no positive colony PCR
pEAQ-P450-21	no positive colony PCR
pEAQ-24ISO	no positive colony PCR
pHREAC-DH1	verified
pHREAC-DH2	verified
pHREAC-DH3	verified
pHREAC-DH5	verified
pHREAC-DH7	verified
pHREAC-DH11	verified

6.4.1 Analysis of P450 leaf disk assays

Due to limited capacity in terms of processing of leaf disk assays, this work focused on P450 candidates that previously evoked decrease of withaferin A (**5**) levels in VIGS experiments (chapter 5.3). The last known intermediate 24-methyldesmosterol (**3**) was dissolved in HEPES buffer (pH 7.5) using 10% (w/v) methyl-β-cyclodextrin (**113**), resulting in a substrate concentration of 36 µM (see section 1.1). Leaf disks were harvested 3-5 dpi and incubated for 24-48 h in the substrate solution (for details see chapter 8.2.3). After incubation the leaf disks were separated from the substrate solution in order to grind leaf disks with liquid nitrogen. However, remaining substrate solution was not wiped from the leaf disk, so that some of the medium was extracted as well, in case products were transported to the medium. Leaf disk samples were extracted as described in chapter 8.2.4, while buffer samples were stored in case extraction was desired at a later point. The extracts were further processed for LCMS or GCMS analysis. An overview of exemplary GCMS chromatograms is given in Figure 6.6.

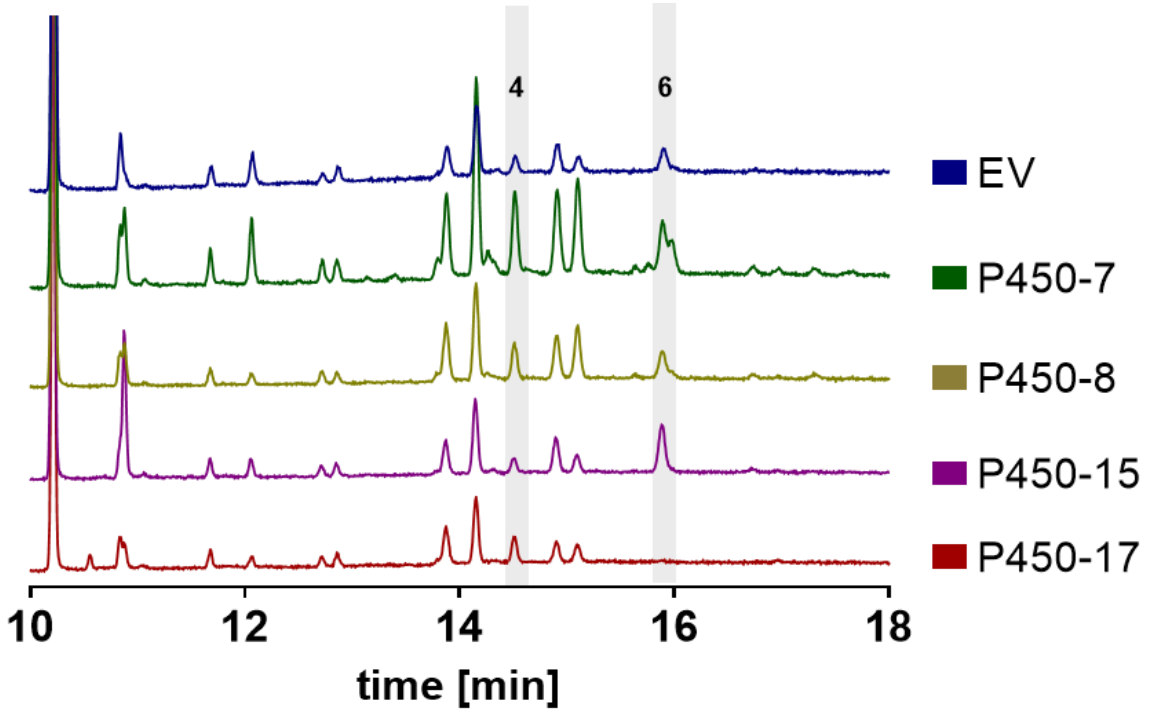


Figure 6.6 Exemplary GCMS (TIC) traces of *N. benthamiana* leaf disk assays show increase of an unidentified compound at 10.87 min for P450-15, accumulation of a new compound at 15.98 min for P450-7 and depletion of cycloartenol (**6**) at 15.90 min for P450-17.

Sterol type compounds can be expected in the time range of 10-18 min. Multiple peaks can be observed in all samples that comprise the natural *N. benthamiana* background. The peaks at 14.61 min and 15.90 min were identified as 24-methyldesmosterol (**3**) and cycloartenol (**6**) by comparison to a standard (see Figure 6.7A).

In P450-15 samples an increase of the peak at 10.87 min can be observed for all replicates including the ones without 24-methyldesmosterol (**3**). A BLASTp search of P450-15 showed up to 90% aa identity to cytochrome 94C1 proteins from various nightshade plants. This enzyme has been reported to be involved in the oxidation of plant hormone jasmonoyl-L-isoleucine. [302–304] *In vitro* experiments have furthermore demonstrated a function as fatty acid hydroxylase. [302] A similar reactivity of P450-15 seems likely, as mass spectra of this peak exhibit multiple fragments in -14 increments, that are characteristic for loss of a CH_2 group (see Figure 10.7 in the appendix). [339] It was concluded that no sterol specific activity was observed for P450-15.

P450-7

Furthermore, slight variations in *N. benthamiana* background peak intensity can be observed, that can be accounted to natural fluctuations and the extraction procedure. In the case of the P450-7 replicate shown in Figure 6.6, multiple small peaks are visible compared to the EV control. It has to be noted that the general signal intensity in this sample is slightly stronger than in other samples, therefore it cannot be excluded that those peaks are part of the natural *N. benthamiana* background. However, P450-7 exhibits a shoulder at the cycloartenol (**6**) peak at 15.90 min that is clearly not present in any other sample. This shoulder was observed in 2 of 3 replicates with 24-methyldesmosterol (**3**) addition and in none of the control samples without the substrate (see Figure 10.8 in the appendix). When examining the mass spectra of the peak at 15.90 min and 15.98 min (Figure 6.7A), the first can be identified as cycloartenol (**6**, compare Figure 6.7B and Figure 6.7D), while the second shows a different mass spectrum (**117**, Figure 6.7C). The mass spectrum of **117** was compared to a 24-methyldesmosterol standard (**3**, Figure 6.7E). Several key fragments were present that were in alignment with key fragments in the **3** standard, suggesting similar core structures. While a mass spectrum of 24-methyldesmosterol (**3**) has been described in the literature, [64] a detailed analysis of fragments observed has not been published. Therefore, key fragments of **117** should be discussed by comparison to key fragments of 24-methyldesmosterol (**3**) isomers.

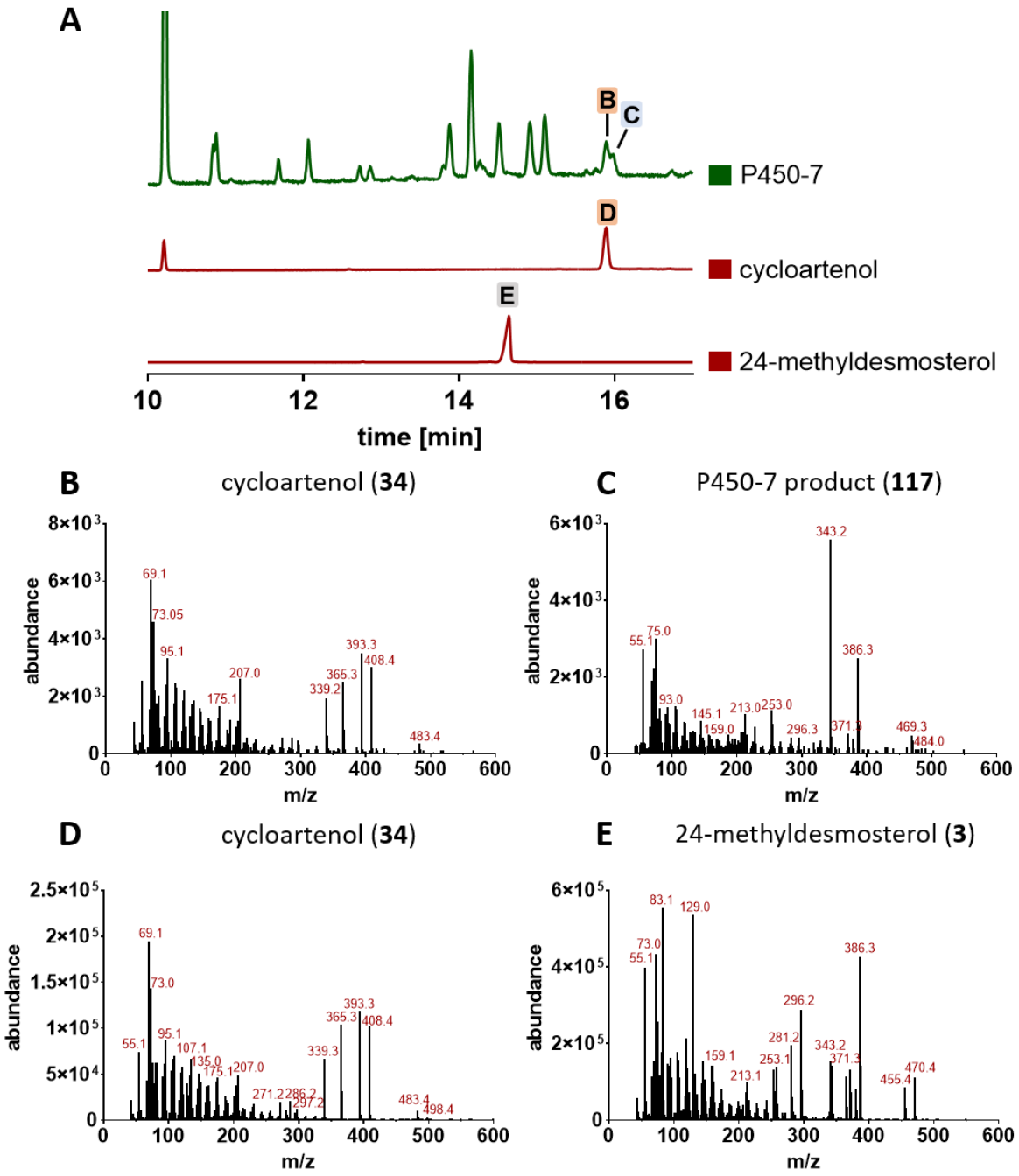
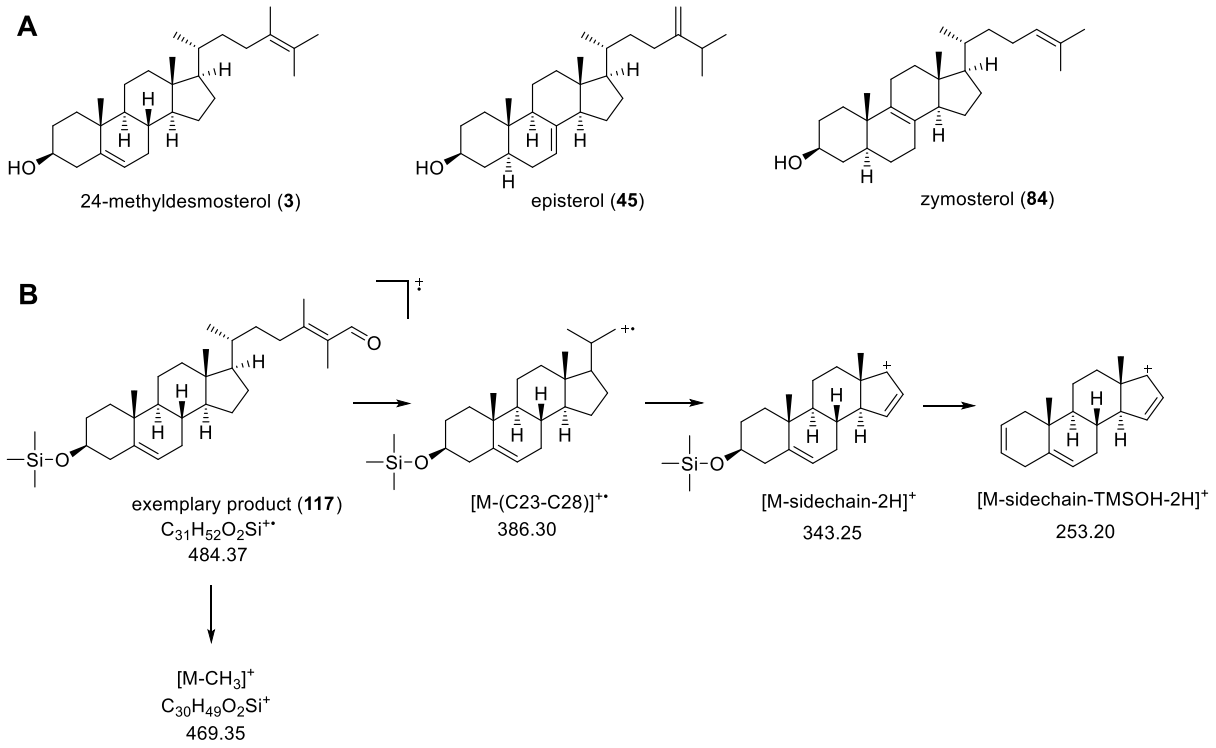


Figure 6.7 GC chromatograms of P450-7 shows accumulation of a new compound compared to reference substances (A); mass spectrum of cycloartenol (**6**) peak of P450-7 sample at 15.90 min (B); mass spectrum of unknown compound of P450-7 sample (**117**) at 15.98 min (C); reference spectrum of cycloartenol (**6**, D); reference spectrum of 24-methyldesmosterol (**3**, E).

The observed signal at *m/z* 484 matches an oxo-24-methyldesmosterol-TMS-ether with the sum formula of $C_{31}H_{52}O_2Si^{+\bullet}$ (Scheme 6.5B). Fragment 469 matches the commonly observed $[M-15]^+$ fragment, that equals loss of a methyl group. [212] Fragment 386 (here $[M-98]^+$) can be observed in fragmentation of episterol (**45**, Scheme 6.5A), a structural isomer of 24-methyldesmosterol (**3**), where it comprises the $[M-(C\text{-}23 \text{ to } C\text{-}28)-H]^+$ fragment. [212] Another

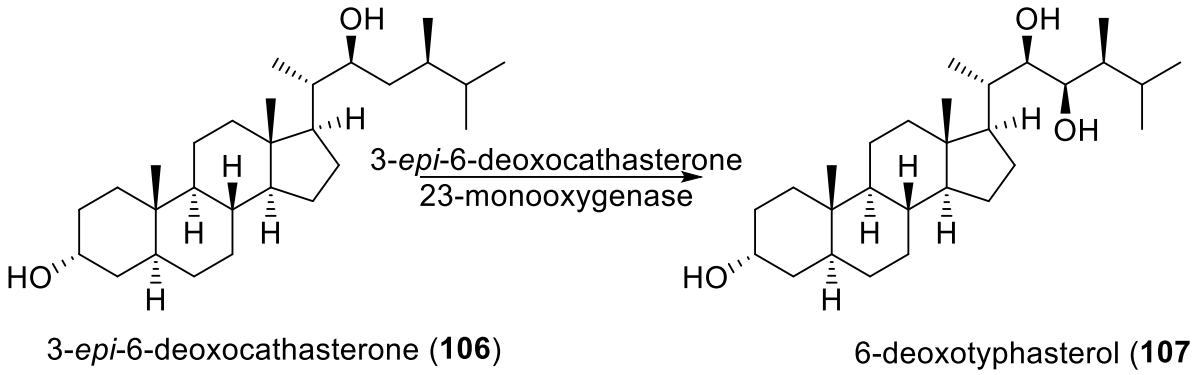
key fragment is 343 that has been described for zymosterol (**84**, Scheme 6.5A) fragmentation in the literature. [212] While **84** does not contain the methyl group at the C-24 position, the core structure without the side chain is an isomer of the 24-methyldesmosterol (**3**) core structure, differing only in the location of the double bond in the B- ring. In the literature the 343 fragment is described as the [M-side chain-2H]$^+$ fragment. [212] Furthermore, the fragment 253 has been described for episterol (**45**), representing the [M−side chain−(TMS-OH) −2H]$^+$ fragment. [212] It was concluded that the observed key fragments of **117** match fragments with a $C_{19}H_{30}O$ core structure similar to 24-methyldesmosterol (**3**), as described in the literature. [212] Consequently, the observed oxidation has likely occurred in the side chain, with the exact position unknown (**117**, Scheme 6.5B). Subsequent assays were analyzed with high resolution LCMS but **117** was not detected. It was concluded that P450-7 might be able to perform the oxidation of the 24-methyldesmosterol (**3**) side chain, but detection of the product has to be improved.



Scheme 6.5 Structures of 24-methyldesmosterol (**3**) and similar compounds episterol (**45**) and zymosterol (**84**, A); potential fragments of an exemplary 24-methyldesmosterol oxidation product (**117**, B).

The BLASTx annotation of P450-7 in the transcriptome data [157] states a 3-*epi*-6-deoxocathasterone 23-monooxygenase. A BLASTp search showed up to 91% aa identity to nightshade 3-*epi*-6-deoxocathasterone 23-monooxygenases (e.g. *Capsicum chinense*, *C. annuum*, *Solanum lycopersicum*). However, those genes were annotated in the respective genome, but function was not proven in enzyme assays. Pairwise alignment with the fully

characterized *A. thaliana* enzyme [340,298] showed 48% aa identity. The function of this enzyme was previously mentioned in chapter 5.3. It catalyzes the hydroxylation of brassinosteroid precursor 3-*epi*-6-deoxocathasterone (**106**) at the C-23 position (see Scheme 6.6). [298] As an aa identity of 48% to the characterized enzyme is not very high, and brassinosteroids and withanolides are structurally similar, it seems reasonable that P450-7 might have evolved to oxidize the withanolide side chain. In conclusion P450-7 exhibited activity that could be relevant for withanolide biosynthesis. However, further characterization of this enzyme and orthologous sequences in other nightshade species (e.g., testing of the substrate spectrum) is needed to come to a conclusion. Future experiments could include expression of this candidate in a 24-methyldesmosterol (**3**) producing yeast strain, generated by Dr. Karan Malhotra, to confirm activity on **3**.



Scheme 6.6 The 3-*epi*-6-deoxocathasterone 23-monooxygenase catalyzes a C-23 hydroxylation in brassinosteroid biosynthesis. [298]

P450-17

Another clear difference is the absence of the cycloartenol (**6**) peak at 15.90 min in the P450-17 sample (Figure 6.8A). Absence or strong decrease of this peak was observed in 5 out of 6 replicates, independent of the presence of 24-methyldesmosterol (**3**). Furthermore, samples analyzed by LCMS showed a new peak at 6.86 min when compared to an EV control (**118**, Figure 6.8B). Further analysis via high resolution LCMS (Orbitrap, see chapter 8.2.4) showed masses matching an oxidized cycloartenol (+16, Figure 6.8C) that is not present in the EV control (Figure 6.8D). It was concluded that P450-17 showed oxidative activity on cycloartenol that is provided by the host organism. Furthermore, microsome assays were tested with P450-17 and cycloartenol (**6**) as a substrate. However, no difference between P450-17 assays and a no enzyme control was detected, underlining, that the microsome assay setup still requires optimization (see Figure 10.9 in the appendix).

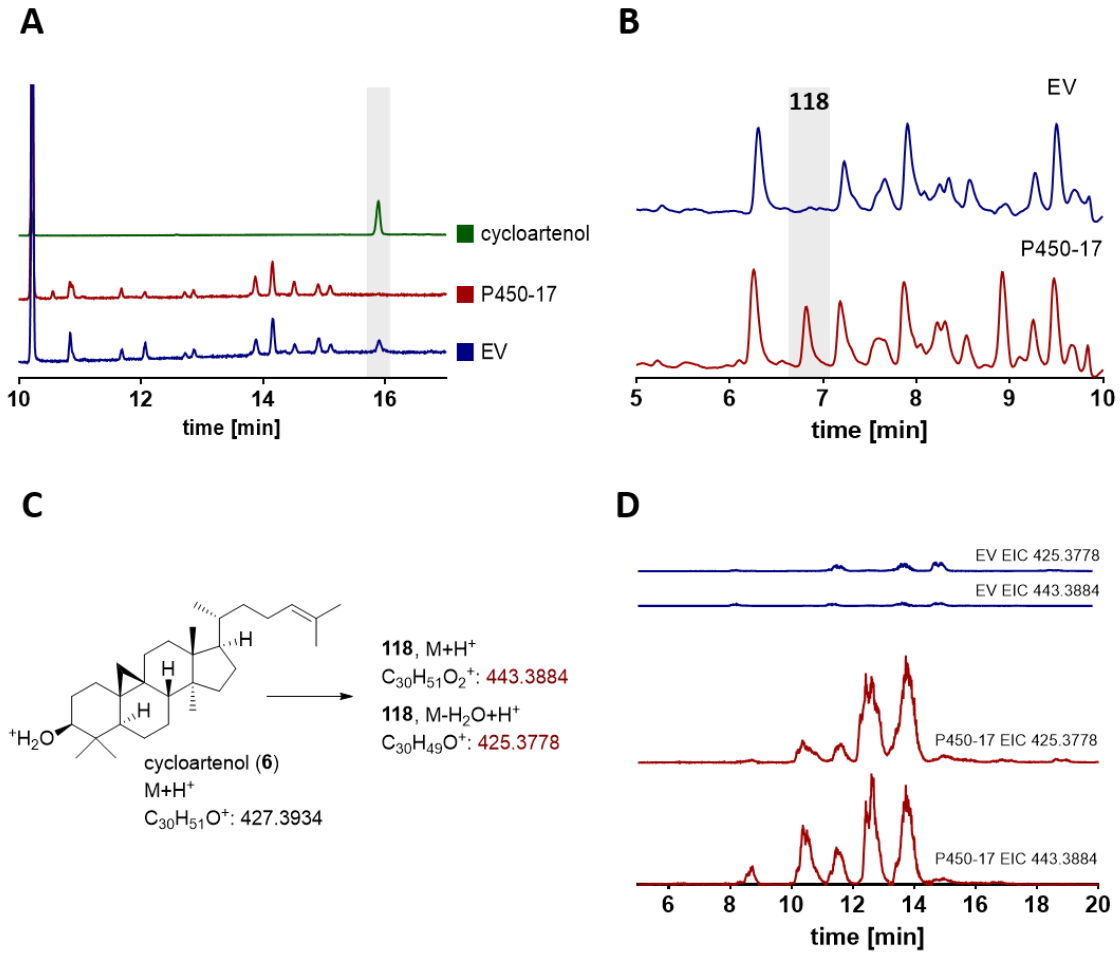


Figure 6.8 P450-17 shows oxidative activity on cycloartenol (**6**). GCMS chromatograms show depletion of cycloartenol (**6**, A); LCMS chromatogram of P450-17 samples show a new peak at 6.86 min (**118**, B); the potential product exhibited a mass corresponding to an oxidized cycloartenol (C); High resolution LCMS EIC traces of the $[M+H]^+$ fragment (443.3884) and $[M-H_2O+H]^+$ fragment (425.3778) showed four distinct peaks in P450-17 samples that are not present in the EV control (D).

The isolation of oxidized cycloartane-type triterpenoids has not been reported from Solanaceae before. However, compounds such as mangiferolic acid (**119**) [341] and cycloartan-24-ene-1α,3β-diol (**120**) [342] were isolated from *Mangifera indica* (Anacardiaceae) and *Commiphora gileadensis* (Burseraceae), respectively (Figure 6.9). Both plants are part of the order of Sapindales, and are therefore not closely related to the Solanaceae family, suggesting that a pathway for production of oxidized cycloartane-type triterpenoids might have developed convergently in both orders.

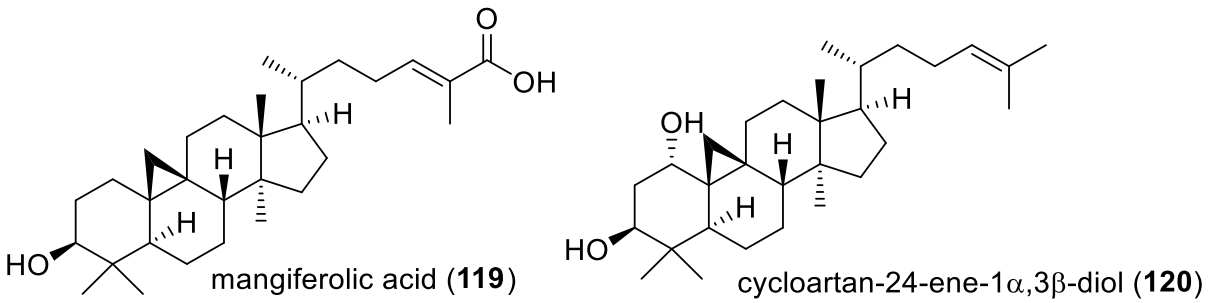
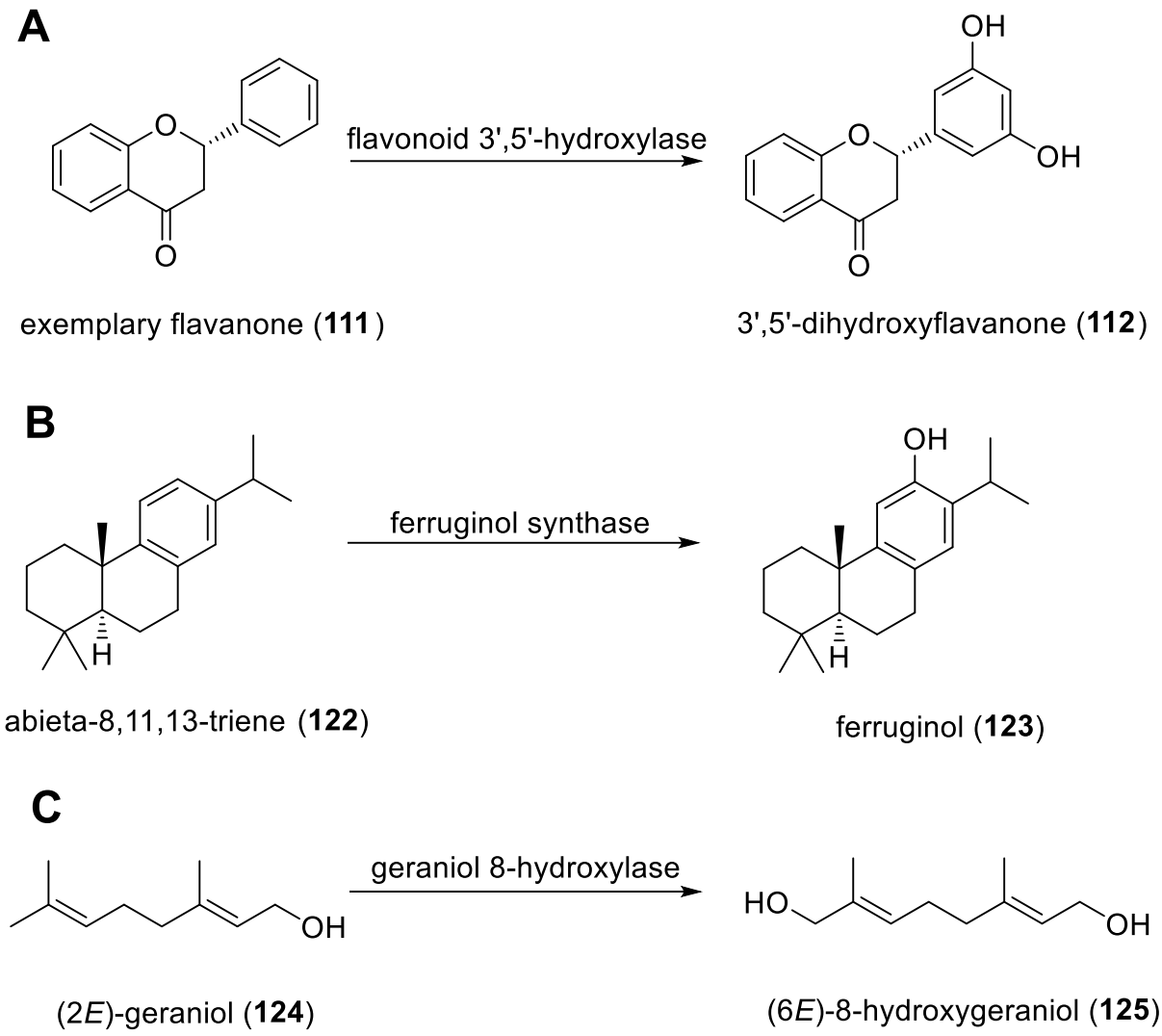


Figure 6.9 Oxidized cycloartane-type triterpenoid mangiferolic acid (**121**) was isolated from *Mangifera indica* [341] and cycloartan-24-ene-1α,3β-diol (**120**) was isolated from and *Commiphora gileadensis*. [342]

Purification of this compound by large scale *N. benthamiana* infiltration is ongoing and will be continued by Samuel E. Hakim. However, until now separation of the P450-17 product from the *N. benthamiana* background has proven difficult. An alternative way of obtaining the desired compound would be the expression of necessary genes in an engineered yeast strain. Dr. Karan Malhotra created a yeast strain that accumulates high amounts of squalene. Expression of a *squalene epoxidase* (*SQE*), *cycloartenol synthase* (*CAS*) and *P450-17* in this strain should generate the product.

As the exact function of P450-17 is unclear, it was investigated whether orthologous sequences exist in other plants that can allow conclusions about the function of this P450. The BLASTx annotation for P450-17 in the *W. somnifera* transcriptome data [157] stated a flavonoid 3',5'-hydroxylase, which is known to catalyze multiple 3',5'-hydroxylations in flavonoids such as anthocyanins (e.g. **111**, to **112**, see Scheme 6.7A). [305] A BLASTp search of P450-17 with the NCBI database [343] yielded hits of up to 84% aa identity to various (hypothetical) proteins in different nightshade species. Examples include a predicted ferruginol synthase (**122** to **123**, Scheme 6.7B) from *Solanum lycopersicum* or a predicted geraniol 8-hydroxylase (**124** to **125**, Scheme 6.7C) from *Nicotiana sylvestris*. All mentioned enzymes catalyze hydroxylations on very different substrates. Therefore, it was concluded that no function can be determined based on homology.



Scheme 6.7 BLASTx annotation and BLASTp hits of P450-17 catalyze hydroxylations on very different substrates. The flavonoid 3',5'-hydroxylase catalyzes 3',5'-hydroxylations with various flavonoid substrates (**111**, A); [305] ferruginol synthase catalyzes hydroxylation of abieta-8,11,13-triene (**122**, B); [344,345] geraniol 8-hydroxylase catalyzes hydroxylation of geraniol (**124**, C). [346,347]

Furthermore, a BLASTp search of P450-17 was conducted by Prof. Dr. Jakob Franke against P450 protein sequences of the database of Dr. David Nelson. [348] Here, the closest homologs were CYP706G5_Solyc390 from *Solanum lycopersicum* with 84% aa identity and Soltu.DM.08G025860.1 from *Solanum tuberosum* with 83% aa identity. In contrast to *Withania somnifera*, genome data is available for *S. lycopersicum* [349] and *S. tuberosum*. [350] Interestingly, P450-17 orthologs in those plants seem to be organized in a gene cluster containing *P450s*, *copper amine oxidases*, and furthermore one *α/β-hydrolase* and one *UDP-D-glucuronate 4-epimerase* (Figure 6.10A). The P450s observed in this gene cluster can be divided into two groups based on their homology to P450-17 (see Figure 6.10B and Table 6.3). As both potato and tomato do not produce withanolides, an importance of this gene cluster beyond withanolide

biosynthesis can be assumed. Future investigation of genes in the *S. tuberosum* cluster, e.g., heterologous expression in *N. benthamiana*, will be conducted by Mahsum Kadah.

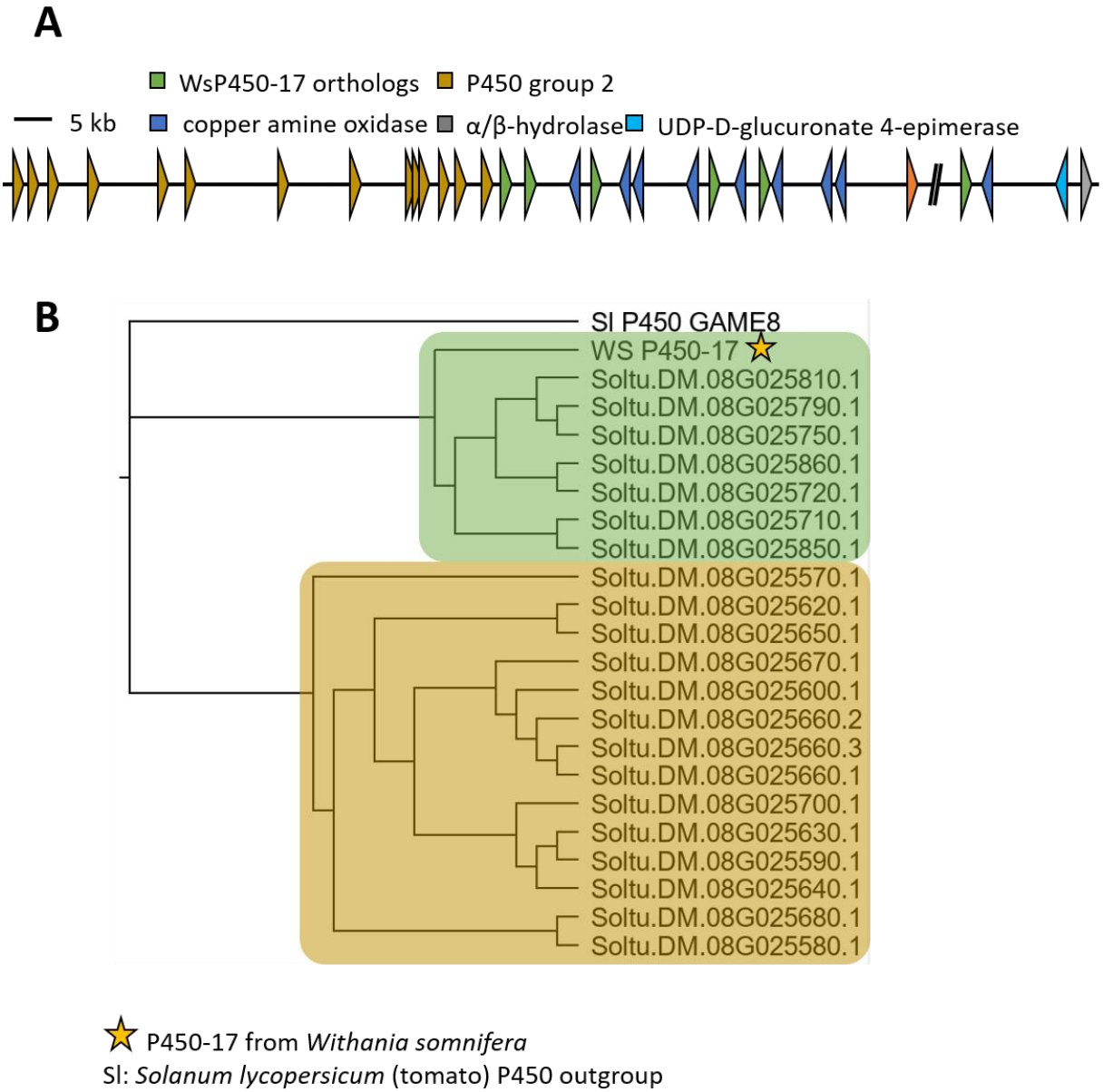


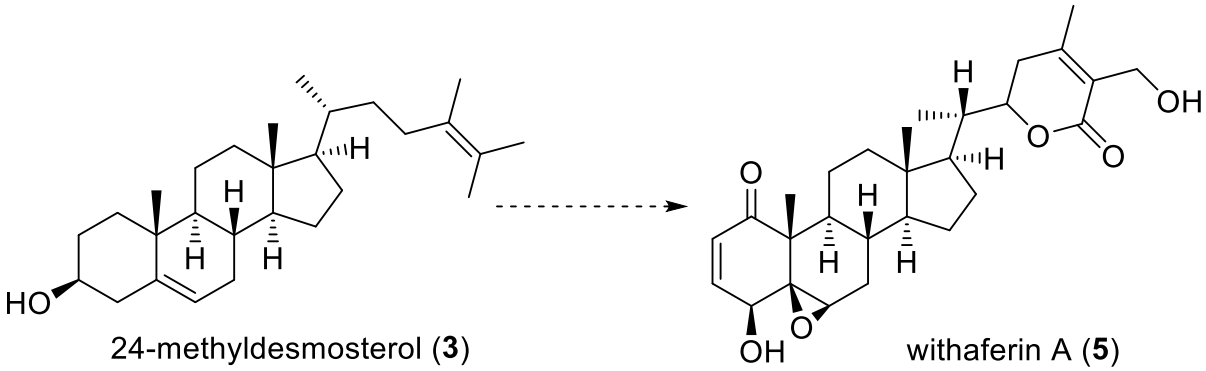
Figure 6.10 P450-17 exhibits orthologous sequences in withanolide non-producers potato and tomato suggesting importance of this P450 beyond withanolide metabolism. Gene cluster containing P450-17 orthologs in *Solanum tuberosum* (A); Phylogenetic tree of P450s in the *Solanum tuberosum* gene cluster shows division into P450-17 orthologs and sequences with less homology to P450-17 (B).

Table 6.3 Amino acid identities (%) of P450 sequences in the *Solanum tuberosum* gene cluster

	GAME8	P450-17	810.1	790.1	750.1	860.1	720.1	710.1	850.1	620.1	650.1	660.2	680.1	700.1	580.1	570.1	590.1	630.1	670.1	640.1	600.1	660.1	660.3
Sl GAME8		20.8	8.8	19.1	18.9	21.5	12.8	20.2	15.5	18.3	20.9	23	5.2	15.8	18.8	7.2	20.6	19.6	19.6	18.8	19.1	19.1	22.4
WsP450-17	20.8		39.4	75.4	76.6	82.6	60.6	82	58.7	41.4	45.7	57.5	9.8	36.1	46.2	16.6	45	44.6	45.4	46.5	45.4	45.8	56.1
Soltu.DM.08G025810.	8.8	39.4		43.3	43.3	42.7	66.2	42.4	22.9	15.7	19.5	1.5	17	10	20.6	7.1	19.6	19.5	19.6	19.8	19.4	19.3	1.4
Soltu.DM.08G025790.	19.1	75.4	43.3		98.6	83.3	69.1	80.9	62.8	37.1	41.1	44.3	10.5	31.9	42.3	16.4	39.8	40.3	41.4	41.9	40.3	40.3	43.8
Soltu.DM.08G025750.	18.9	76.6	43.3	98.6		84.5	69.1	82.2	65.4	38	42	46.2	10.5	32.8	43.2	16.8	40.6	41.2	42.3	43	41.4	41.2	45.7
Soltu.DM.08G025860.	21.5	82.6	42.7	83.3	84.5		68.2	87.5	63.6	41	45.7	55.3	10.9	35.1	46.4	16.6	43.6	44.2	45.8	46.5	44.4	44.6	54.1
Soltu.DM.08G025720.	12.8	60.6	66.2	69.1	69.1	68.2		62.8	41.3	23.6	27.9	10.7	13.4	17.5	29.2	14.1	27.3	27.7	28.1	28.3	27.3	27.7	12.8
Soltu.DM.08G025710.	20.2	82	42.4	80.9	82.2	87.5	62.8		66.9	41.9	46.4	55.8	11.1	36.1	47.5	16.6	45	45.7	45.7	47.5	45.2	45	54.5
Soltu.DM.08G025850.	15.5	58.7	22.9	62.8	65.4	63.6	41.3	66.9		26.2	35.2	52.2	0	25.5	33.7	33.3	32.2	33.7	32.2	35.2	32.2	31.8	50
Soltu.DM.08G025620.	18.3	41.4	15.7	37.1	38	41	23.6	41.9	26.2		81.8	83.5	23.7	58.4	74.4	27	78.1	77.7	77.5	77.5	80.5	78.4	83.6
Soltu.DM.08G025650.	20.9	45.7	19.5	41.1	42	45.7	27.9	46.4	35.2	81.8		83.5	22.9	58.5	80.9	29.6	86.1	84.2	84.8	84.4	86.3	84.2	83.6
Soltu.DM.08G025660.1	23	57.5	1.5	44.3	46.2	55.3	10.7	55.8	52.2	83.5	83.5		4.8	85	86.4	14.7	84	83.5	86.9	82	91.3	100	100
Soltu.DM.08G025680.	5.2	9.8	17	10.5	10.5	10.9	13.4	11.1	0	23.7	22.9	4.8		4.5	22	11.2	22.5	20.8	22	22	22.7	22.5	4.5
Soltu.DM.08G025700.	15.8	36.1	10	31.9	32.8	35.1	17.5	36.1	25.5	58.4	58.5	85	4.5		56.1	19.3	58.2	59	57.6	58.2	59.6	58.8	85
Soltu.DM.08G025580.	18.8	46.2	20.6	42.3	43.2	46.4	29.2	47.5	33.7	74.4	80.9	86.4	22	56.1		30.4	79.4	81.8	79.2	79.5	81.1	81.4	87.3
Soltu.DM.08G025570.	7.2	16.6	7.1	16.4	16.8	16.6	14.1	16.6	33.3	27	29.6	14.7	11.2	19.3	30.4		29.4	30	28.6	29.2	29.1	29	15.6
Soltu.DM.08G025590.	20.6	45	19.6	39.8	40.6	43.6	27.3	45	32.2	78.1	86.1	84	22.5	58.2	79.4	29.4		88.6	84.1	85.8	86.6	86.4	83.6
Soltu.DM.08G025630.	19.6	44.6	19.5	40.3	41.2	44.2	27.7	45.7	33.7	77.7	84.2	83.5	20.8	59	81.8	30	88.6		83.3	84.6	85.6	84.1	83.2
Soltu.DM.08G025670.	19.6	45.4	19.6	41.4	42.3	45.8	28.1	45.7	32.2	77.5	84.8	86.9	22	57.6	79.2	28.6	84.1	83.3		84.5	87	86	86.4
Soltu.DM.08G025640.	18.8	46.5	19.8	41.9	43	46.5	28.3	47.5	35.2	77.5	84.4	82	22	58.2	79.5	29.2	85.8	84.6	84.5		87	85.4	81.8
Soltu.DM.08G025600.	19.1	45.4	19.4	40.3	41.4	44.4	27.3	45.2	32.2	80.5	86.3	91.3	22.7	59.6	81.1	29.1	86.6	85.6	87	87		88.5	90.9
Soltu.DM.08G025660.	19.1	45.8	19.3	40.3	41.2	44.6	27.7	45	31.8	78.4	84.2	100	22.5	58.8	81.4	29	86.4	84.1	86	85.4	88.5		100
Soltu.DM.08G025660.1	22.4	56.1	1.4	43.8	45.7	54.1	12.8	54.5	50	83.6	83.6	100	4.5	85	87.3	15.6	83.6	83.2	86.4	81.8	90.9	100	

6.4.2 Analysis of DH leaf disk assays

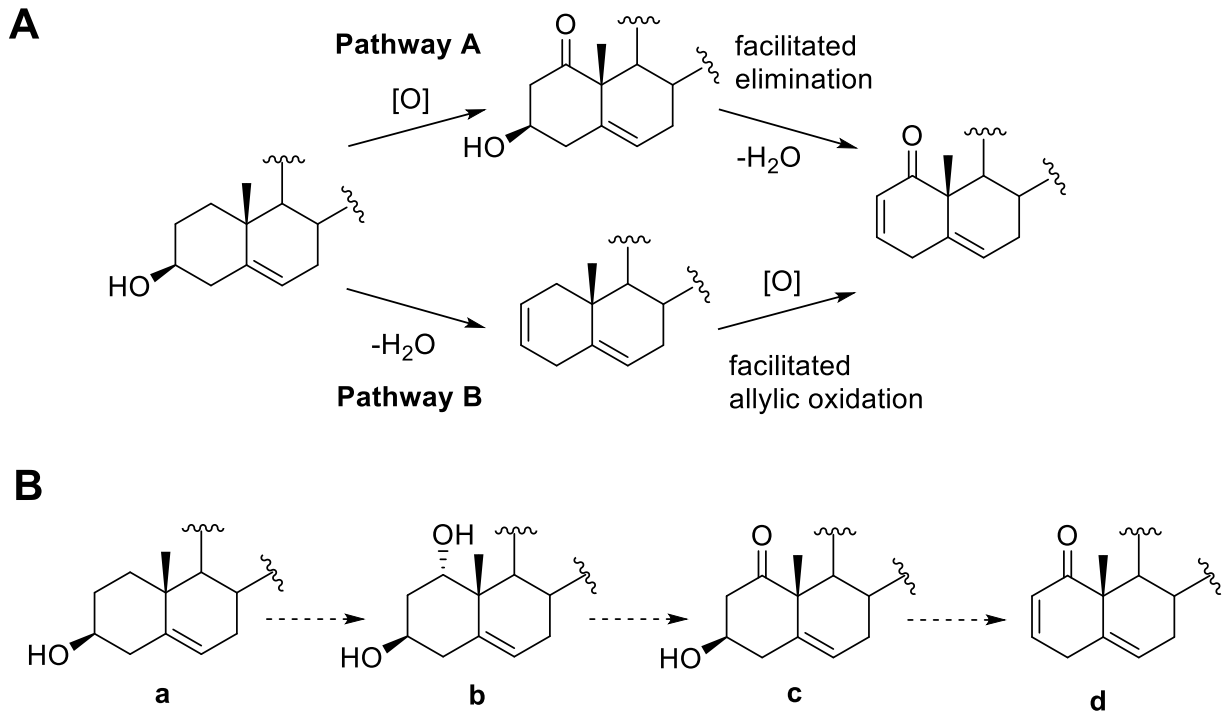
Besides multiple oxidations, withanolides possess a double bond in the A-ring that is not present in the last known precursor 24-methyldesmosterol (**3**, Scheme 6.8). A dehydratase enzyme could perform an elimination reaction with the 3-hydroxy group leaving the A-ring. VIGS experiments targeting DH candidates for this reaction showed decrease of withaferin A (**5**) levels for two candidates (DH7 and DH11). However, no accumulation of intermediates was observed that could give clues on the substrates needed of DH assays.



Scheme 6.8 Withaferin A (**5**) contains a double bond in the A-ring.

In addition to the double bond, most withanolides possess a ketone at C-1, forming a Michael system. As a substrate is needed for DH assays, the order of those functionalizations in the biosynthesis should be discussed briefly.

The first possibility (pathway A, Scheme 6.9A) is that the oxidation occurs first, followed by the elimination of the 3-hydroxy group. The aldol intermediate in pathway A exhibits a more acidic proton in the alpha position, which would facilitate the elimination. The second possibility starts with an elimination followed by the oxidation (pathway B, Scheme 6.9A). An argument for pathway B is that oxidation of unfunctionalized sp^3 carbons is chemically difficult, unless those positions are activated by e.g. adjacent double bonds that can stabilize the active species. [351] Intermediates with a 3-hydroxy and 1-keto group have been isolated (see chapter 1.3.3, compounds **49**, **50** and **51**), therefore creating the consensus for the proposed biosynthesis that is shown in Scheme 6.9B. [66,67]



Scheme 6.9 Possible order of oxidation and elimination (A) and commonly proposed sequence for the A ring functionalization for withanolide biosynthesis (B).

Furthermore, glycowithanolides such as withanoside V (**126**) or withanoside VI (**127**, Figure 6.11) [71] possess a hydroxy groups at C-1 and C-3, presumably presenting a branching point in the functionalization of the A ring.

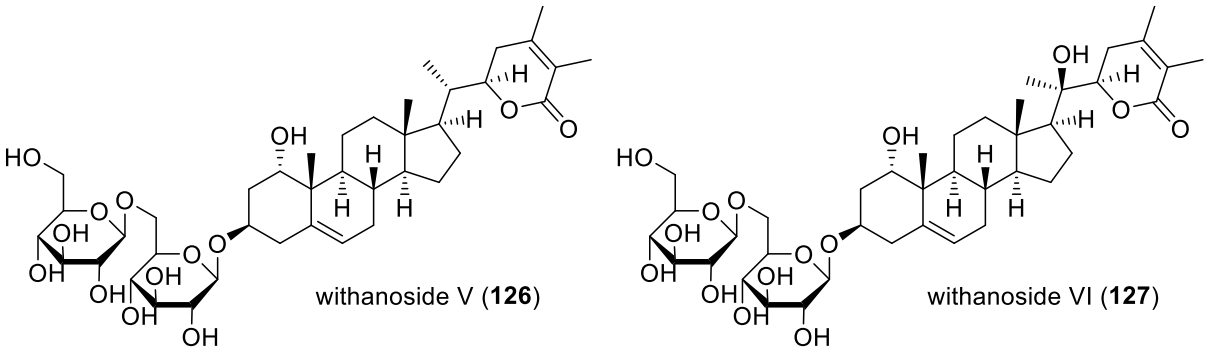
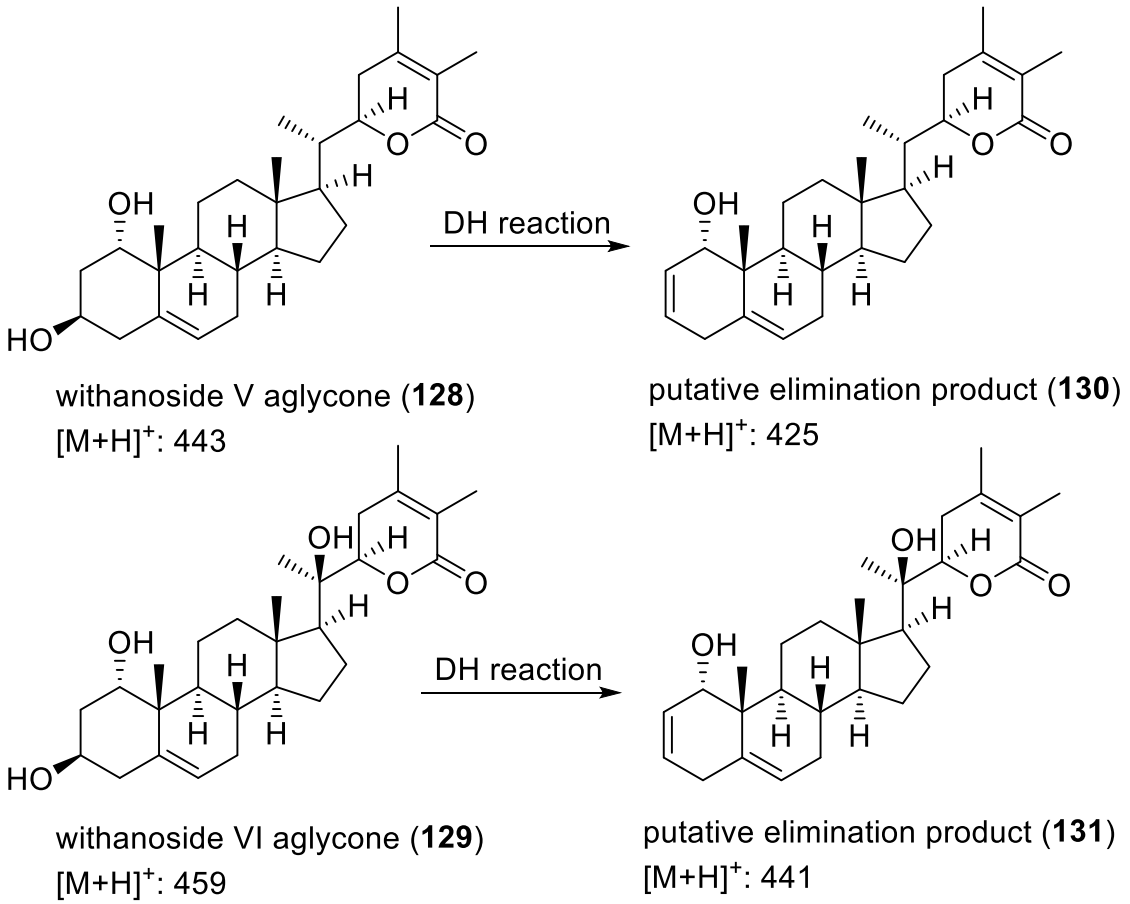


Figure 6.11 Structures of withanoside V and withanoside VI. [71]

Withanoside V (**126**) and VI (**127**) were recently isolated from *W. somnifera* root powder by Samuel E. Hakim in this group. As these compounds were available and dehydratase candidates should be tested in leaf disk assays, the glucose moieties were cleaved off by Samuel E. Hakim. It was reasoned that dehydratases in withanolide biosynthesis might be able to convert the respective aglycone (**128** and **129**) to the unsaturated product. However, it is unclear whether a 1-OH substrate can be accepted by the enzyme or if a keto-group is needed at the C-1 position or whether the pathway takes a different route.

Leaf disk assays were conducted as described in chapter 8.2.3. However, in this case addition of cyclodextrins was not necessary, as withanoside aglycones are soluble in aqueous buffers. As no reference compounds are available, LCMS chromatograms were analyzed for the presumed masses of an elimination reaction. The withanoside aglycones and their elimination products with their respective $[M+H]^+$ ions are depicted in Scheme 6.10.



Scheme 6.10 Potential products of elimination reaction of withanoside aglycones, which might be catalyzed by dehydratase candidates.

Therefore, leaf disk assays were analyzed for masses of the respective elimination product (425 for **130** or 441 for **131**). Figure 6.12 shows exemplary EIC 425 traces of DH leaf disk assays conducted with the withanoside V aglycone (**128**). Here, an additional problem can be observed, as most withanolides exhibit a $[M-H2O+H]^+$ fragment due to in-source fragmentation, exacerbating analysis. Nonetheless, accumulation of an elimination product should be visible, as long as the retention time differs.

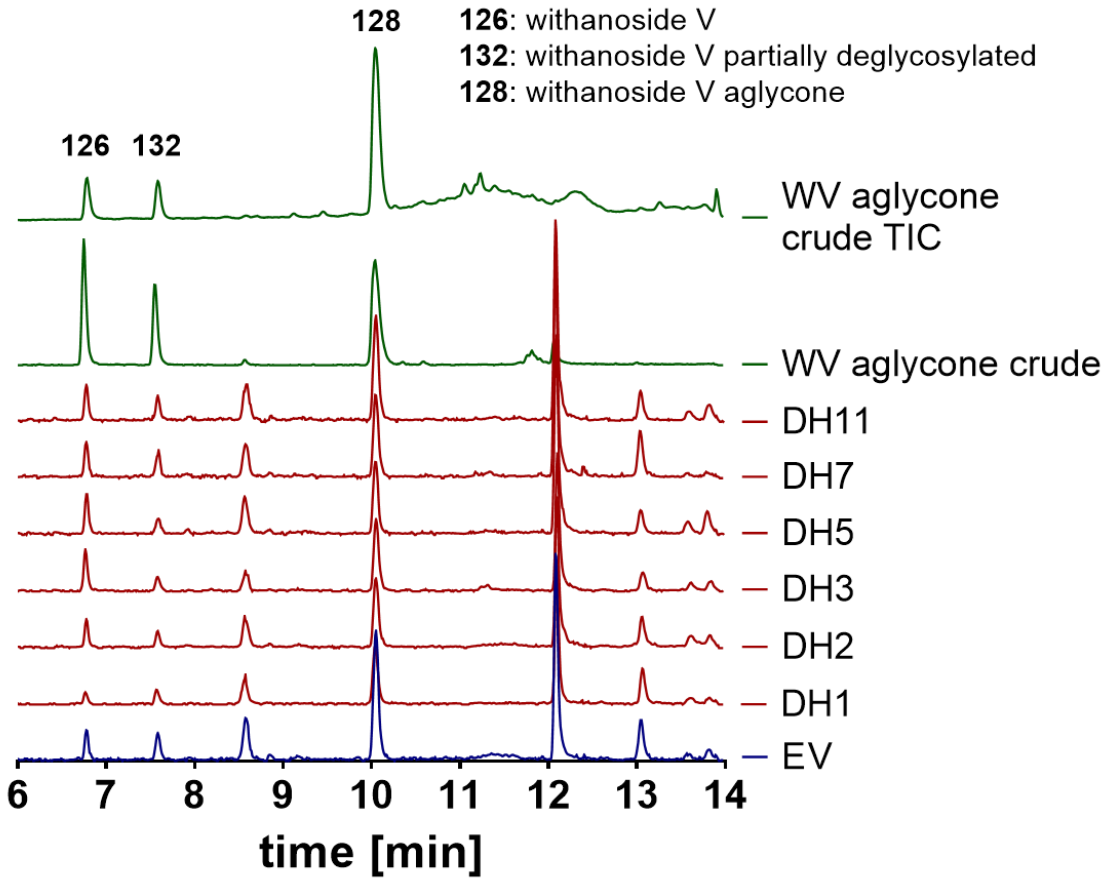


Figure 6.12 LCMS chromatograms of DH leaf disk assays do not show a difference compared to an EV control. Traces depicted are EIC 425, the theoretical elimination product of the withanoside V aglycon substrate. However, multiple peaks with this *m/z* value are visible, as withanolides typically show an $[M-H2O+H]^+$ fragment due to in-source fragmentation. All EIC traces were amplified by 100 to overlay with the TIC trace (**126**: withanoside V, **132**: withanoside V with one glucose cleaved, **128**: withanoside V aglycone).

For the substrate used in leaf disk assays (green, Figure 6.12), multiple peaks can be observed, showing that the cleavage of the glycosyl moiety was incomplete. While the peak at 6.7 min can be identified as withanoside V (**126,** Figure 6.12), the peak at 7.6 min represents withanoside V with one glucose unit cleaved (**132**) and the peak at 10.0 min equals the aglycone (**128**). When comparing the assays to the substrate multiple additional peaks can be observed at 13-14 min that are not present in the standard. All those peaks show signals at *m/z* 443, 425 and 407. However, those peaks are present in the EV control as well and are assumed to be produced by the *N. benthamiana* host. Here, it cannot be excluded that one of those peaks is the expected elimination product due to lack of a standard. Quantification of the aglycone substrate was not informative, as substrate amount in the individual replicates underlies fluctuations, e.g. caused by varying uptake of the substrate into the leaf disk.

In conclusion, no formation of product was detected in DH assays. Future experiments should improve the processing of the aglycone substrate, so that no partially cleaved substrate is interfering with the LCMS analysis of DH assays.

As the product formation in the previous chapter showed (see chapter 6.4.1), the general system of leaf disk assays is at least able to produce active P450 enzymes. Even though DH activity cannot be guaranteed from these results, it seems likely that lack of distinct product formation is not due to an inactive enzyme. Here, three explanations seem reasonable. Either background activity of the host *N. benthamiana* is able to produce the expected product, so that DH samples cannot be distinguished from EV controls, the dehydratase that actually catalyzes the reaction was not tested, or a wrong substrate was used.

As no reference compound is available, it cannot be excluded that one of the multiple peaks with a matching *m/z* value is the expected product. VIGS data showed a decrease of withaferin A (**5**) levels for DH7 and DH11, suggesting that one of those candidates might be involved in the reaction, therefore it seems more likely that instead of a wrong candidate, an unsuitable substrate (withanoside V aglycone, **128**) was used. Here, no further investigation could be conducted due to time constraints. Further experiments should include analysis with high-resolution mass spectrometry to provide further insights (e.g. sum formula) on the multiple peaks with an *m/z* signal at 425 present in the LCMS chromatogram. Furthermore, future experiments should concentrate on testing substrates with different functionalizations in the A-ring (e.g., a keto group at C-1 position) to determine whether substrate selectivity is the reason for the lack of product formation. If no product formation can be observed for any substrate, other dehydratase candidates should be investigated as well.

6.5 Final conclusion of enzyme assays

The conduction of enzyme assays for withanolide biosynthesis presents multiple difficulties. First, the investigation of P450 reactions typically requires a host that is able to provide matching membrane structures, folding machinery and the reaction partner CPR. [306,307] This implies enzyme assays with an eukaryotic host. [308–311] Due to previous experience in this group the hosts *S. cerevisiae* and *N. benthamiana* were chosen for expression experiments.

Another hurdle that has to be faced is the availability of substrates and product standards, as neither are commercially available. Since no accumulation of intermediates was detected in silencing experiments (chapter 5.3.2), the exact substrates for selected P450 candidates are unknown. Therefore, enzyme assays should be conducted with the last known precursor of

withanolide biosynthesis, 24-methyldesmosterol (**3**). [xvii] The main working hypothesis of this thesis was that initial steps of conversion of 24-methyldesmosterol (**3**) towards withanolides would likely include oxidation of the side chain or the A-ring. Therefore, assays were investigated for potential products with a sum formula of $C_{28}H_{46}O_2$ for a hydroxylated product, $C_{28}H_{44}O_2$ for a ketone/aldehyde, $C_{28}H_{44}O_3$ for a carboxylic acid or $C_{28}H_{42}O_3$ for a lactone.

When using 24-methyldesmosterol (**3**) as a substrate, the solubility of this hydrophobic compound is in conflict with the enzymes' need for an aqueous buffer. As the addition of substantial amounts of MeOH or EtOAc to microsome assays or *N. benthamiana* leaves would lead to precipitation of microsomes or death of infiltrated leaves, respectively, a different solution had to be found to dissolve 24-methyldesmosterol (**3**) in aqueous buffers. In this work, the solubilization of various sterol-type compounds with Tween 80 and cyclodextrins was tested. Tests with cycloartenol (**6**) and Tween 80 showed various residing impurities caused by the additive in LCMS chromatograms after extraction of the substrate. Therefore, solubilization with this method was dismissed. Furthermore, the addition of methyl-β-cyclodextrin (MCD, **113**) and (2-hydroxypropyl)-β-cyclodextrin (HPCD, **114**) to substrate stock solutions was tested. While 20% (w/v) HPCD (**114**) was not able to dissolve substantial amounts of cholesterol (**15**), a HEPES Buffer with 33% (w/v) MCD (**113**) exhibited a cholesterol (**15**) concentration of 75 mM. As dilution of the cleared substrate solution caused precipitation of the dissolved cholesterol (**15**), subsequent solubilization attempts had to be directly adjusted to the final concentration used in the assays. Multiple experiments showed that solubility of compounds tested (cholesterol (**15**), sitosterol (**14**) and 24-methyldesmosterol (**3**)) was different and MCD (**113**) concentrations had to be adjusted for each compound. Substrate solutions were subsequently used in microsome assays and leaf disk assays. For 24-methyldesmosterol (**3**), a concentration of 92 µM was achieved when adding 10% MCD (**113**) to the buffer.

As *S. cerevisiae* is unable to take up exogenous sterols under aerobic conditions, [194,195] microsome assays were planned instead of feeding assays. For expression the plasmid pYES2 (Invitrogen) and the *S. cerevisiae* strain WAT11 were selected; the latter contains an *A. thaliana CPR* integrated into the genome. [226] As a positive control the cinnamate 4-hydroxylase (C4H) [316,335] was selected, as yeast microsome assays with this P450 have been reported before. [313,316–318] The substrate, *trans*-cinnamic acid (**115**), was soluble in buffer, circumventing substrate solubilization with cyclodextrins. Here, formation of the product *p*-coumaric acid (**116**) was observed for the *A. thaliana* C4H tested and furthermore for the putative *W.*

[xvii] synthesized by Dave Biedermann

somnifera ortholog. It was concluded that the general system of enzyme production, microsome isolation and enzyme assays was working. Further control reactions were included that had hydrophobic substrates to test compatibility of the cyclodextrin solubilization method with microsome assays. Therefore, the sterol-22-desaturase and cholesterol-22-hydroxylase were selected, which act on sitosterol (**14**) and cholesterol (**15**), respectively. Here, no accumulation of the respective product was detected. Furthermore, microsome assays were conducted with P450-7 and P450-8 using 24-methyldesmosterol (**3**) as a substrate. No accumulation of the presumed products ($C_{28}H_{46}O_2$ or $C_{28}H_{44}O_2$) was detected. As leaf disk assays suggested activity of P450-17 on cycloartenol (**6**), forming a product matching the sum formula of $C_{30}H_{51}O_2$, microsome assays with P450-17 were conducted with **6** instead. In contrast to the leaf disk assays no accumulation of a peak with the mass to charge ratio 443 was observed. It was concluded that the enzymes used in those microsome assays were either inactive or the respective substrate was not able to access the enzyme. Proving either hypothesis is difficult, as there is no direct way to control it. It was concluded that the conduction of microsome assays for investigation of withanolide biosynthesis would require intensive optimization and was therefore dismissed, as the alternative host system *Nicotiana benthamiana* was able to provide more insights.

The second strategy employed in this chapter was the expression of candidates in the common plant host *Nicotiana benthamiana*. Here, a protocol developed by the group of Prof. Sarah O'Connor was employed. Instead of infiltrating substrate into *N. benthamiana* leaves previously infected with the respective GOI overexpression construct and harvesting and processing the whole leaf, small disks of infected *N. benthamiana* plants were incubated in substrate-buffer solution, ultimately consuming less 24-methyldesmosterol (**3**).

Expression of P450 candidates that showed a withaferin A (**5**) decrease in VIGS experiments and furthermore all DH candidates that were tested in VIGS experiments was conducted in pHREAC.[xviii] Experiments with DH candidates and the withanoside V aglycone (**128**) did not yield a potential new product peak that was absent in empty vector control experiments. Here, further work is needed to determine whether the lack of product formation is due to testing the wrong substrates or wrong candidates.

Leaf disk assays with P450 candidates showed that this system is working in general. While P450-8 and P450-15 did not show significant changes linked to sterol-type compounds in

[xviii] P450-7, P450-8, P450-15,P450-17, DH1-3, DH5, DH7 and DH11 were tested; P450-14 could not be cloned.

GCMS chromatograms, P450-7 showed accumulation of a new peak at 15.98 min in 2 of 3 replicates after 24-methyldesmosterol (**3**) feeding. The observed mass spectrum suggested that the side chain was oxidized, as key fragments at lower *m*/*z* values matched the core fragments of a 24-methyldesmosterol (**3**) standard, while a signal at *m/z* 484 likely corresponds to an oxidized 24-methyldesmosterol-TMS-ether with the sum formula of $C_{31}H_{52}O_2Si^{+\bullet}$ (**117**). Independent repetitions of this experiment were analyzed with high resolution LCMS but here no compound matching a sum formula of $C_{28}H_{45}O_2^+$ was observed. BLASTp analysis of P450-7 showed up to 91% aa identity to an enzyme that hydroxylates 3-*epi*-6-deoxocathasterone (**106**), which is structurally close to withanolides. Further investigation, such as the isolation and structure elucidation of the product is needed. Here a possibility could be production of the oxo-methyldesmosterol (**117**) by addition of *P450-7* to a 24-methyldesmosterol (**3**) producing yeast strain generated by Dr. Karan Malhotra.

Another major observation was the activity of P450-17. GCMS analysis of leaf disk assays showed depletion of cycloartenol (**6**) provided by the host plant in 5 out of 6 samples, independent of the presence of the added substrate 24-methyldesmosterol (**3**). Analysis of subsequent leaf disk assays was conducted with high resolution LCMS and showed a mass (443.3884) matching an oxidized cycloartenol with the sum formula $C_{30}H_{51}O_2^+$ (**118**). Due to the lack of standards for such a compound, the product should be isolated from large scale *N. benthamiana* infiltration for structural analysis. The isolation is currently ongoing and will be continued by Samuel E. Hakim. As no oxidized cycloartane-type compounds have been reported in *W. somnifera*, even though the metabolite profile of this plant has been extensively investigated, it can be speculated whether the observed activity is natural for P450-17. It is possible that the enzyme showed promiscuity as cycloartenol (**6**) was easily available in reasonable quantities in the host plant *N. benthamiana*. Another possibility is that oxidized cycloartanes undergo further modifications and are therefore usually not detected in *W. somnifera*.

In order to form a more precise hypothesis about the function of P450-17 BLASTp searches were conducted. The BLASTp hits from the NCBI database [352] catalyze hydroxylations on various substrates, so that no function could be determined based on homology. A second BLASTp search was conducted by Prof. Dr. Jakob Franke against P450 protein sequences of the database of Dr. David Nelson. [348] The closest homologs were CYP706G5_Solyc390 in *Solanum lycopersicum* (84% aa identity) and Soltu.DM.08G025860.1 in *Solanum tuberosum* (83% aa identity). Here, the P450-17 homologous sequences are organized in a gene cluster

containing *P450s* (P450-17 orthologs and other), *copper amine oxidases*, and furthermore one *α/β-hydrolase* and one *UDP-D-glucuronate 4-epimerase*. Such a gene cluster cannot be verified in *Withania somnifera*, as no genome data is known. As both potato and tomato are non-producers of the withanolide class, a function of this gene cluster and P450-17 beyond withanolide biosynthesis can be assumed. Mahsum Kadah will continue analysis of the *S. tuberosum* cluster by heterologously expressing genes in *N. benthamiana*.

In summary the investigation of withanolide biosynthesis with microsome assays did not lead to new results. However, leaf disk assays provided new insights, as GCMS data suggested that P450-7 is able to oxidize the 24-methyldesmosterol (**3**) side chain. Here future work in this group will have to provide further insights to come to a conclusion. Furthermore P450-17 showed unexpected activity, as it is able to oxidize cycloartenol (**6**). Presence of gene clusters containing P450-17 orthologs in tomato and potato suggest an importance beyond withanolide biosynthesis that will be target of further investigations in this group.

7 Overall conclusion and Outlook

This study aimed to gather insights into main functionalizations involved in withanolide biosynthesis. With its plethora of compounds reported, withanolide biosynthesis provides a complex subject for research, and since this project was started in this group with this thesis, several (bio-)chemical and biological methods had to be established in this institute before specific experiments could be started.

Withanolides are produced in many Solanaceae species. For research in this work, the most prominent producer *Withania somnifera* and the popular edible *Physalis peruviana* were selected. To gain an overview of main withanolides produced in the respective species, both plants were extracted and analyzed. In *P. peruviana* the main withanolides were 4β-hydroxywithanolide E (**73**), withaperuvin C (**75**) and withanolide F (**79**) together with two yet unknown compounds. Purification and structure elucidation together with Dave Biedermann came to the conclusion that *P. peruviana* produces truncated androwithanolides, termed irinan A (**1**) and B (**2**), that seem to be cleavage products of common withanolides. Bioactivity assays by the Brönstrup group concluded that albeit antiproliferative activity was present, the lack of the lactone side chain decreased the activity. Isolation, structure elucidation and bioactivity assays of irinans have been published in the Beilstein Journal of Organic Chemistry. [172]

One of the big obstacles faced in withanolide biosynthesis is the lack of commercially available intermediates. This work attempted to generate a yeast strain that produces 24-methyldesmosterol (**3**), the last known intermediate in withanolide biosynthesis. [64] Here, knockout of ergosterol biosynthesis genes *ERG4* and *ERG5* with subsequent genomic integration of plant genes *7RED* and *24ISO* was tested to divert ergosterol (**16**) biosynthesis towards production of 24-methyldesmosterol (**3**). GCMS analysis showed that the double knockout was necessary to accumulate substantial amounts of precursor ergosta-5,7,24(28)-trien-3ß-ol (**85**), a result later confirmed in experiments by Dr. Karan Malhotra. However, it was concluded that genomic integration of the *7RED* gene was not stable, as production of 24-methylenecholesterol (**4**), the first intermediate, was inconsistent and colony PCR did not show integration at the expected locus.

Nonetheless, these results laid the foundation for an improved engineering strategy currently being pursued by Dr. Karan Malhotra using CRISPR-Cas9. For this project, the intermediate 24-methyldesmosterol (**3**) was obtained via synthesis by Dave Biedermann.

Withanolides contain prominent features such as the Michael system in the A ring and the lactonized side chain. Due to their versatility and involvement in phytosterol metabolism,

cytochrome P450 monooxygenases are the most likely candidates to perform most of those reactions. Additionally, a dehydratase is likely to perform the elimination in the A ring. Here, transcriptome data of withanolide producers *W. somnifera* and *P. peruviana* was examined for potential P450 and DH candidates using criteria such as expression pattern in different tissues and homology or co-expression to known phytosterol and withanolide genes. In total, 21 P450 (17 cloned) and 14 DH (6 cloned) candidates were determined for further research.

The experimental focus of this work was on silencing and expression experiments. At first, a virus induced silencing system was established for screening of the selected withanolide biosynthesis gene candidates. Here, the VIGS technique was applied to *Physalis peruviana* and *Withania somnifera* with protocols described in the literature, [149,271] using the visual marker phytoene desaturase. [269] While systemic silencing of *P. peruviana* was not achieved with the methods reported, and instead required substantial fine-tuning, bleaching of *W. somnifera* was observed as expected. It was concluded that *P. peruviana* exhibits a robust immune system and is not susceptible to this technique. Further experiments on withanolide biosynthesis gene candidates were therefore conducted in *W. somnifera* and showed significant decrease of the main withanolide withaferin A (withaferin A) in 5 out of 17 P450 candidates and in 2 out of 6 DH candidates tested.

In another set of experiments those 5 P450 and all DH candidates were tested for their activity in enzyme assays. P450 candidates were initially tested in yeast microsome assays with the hydrophobic substrate 24-methyldesmosterol (**3**) dissolved in buffer by the addition of methyl-β-cyclodextrin (**113**). Functionality of the yeast expression system was tested with a known P450 (C4H) with a water-soluble substrate, where the anticipated reaction was observed. For enzymes with hydrophobic substrates product formation could not be observed neither in the known enzyme controls, nor in the P450 candidates. As the underlying cause such as misfolding, lack of activity or a problem in substrate-enzyme accessibility could not be determined, expression in the plant host *Nicotiana benthamiana* was tested instead. P450 and DH candidates were tested in leaf disk assays, with infiltrated leaves submerged in a substrate solution of 24-methyldesmosterol (**3**, P450) or withanoside aglycones (**128**, DH) respectively. While DH assays did not show product accumulation, suggesting that either the substrate or enzymes tested were wrong, two P450 candidates showed activity. In P450-7 assays a new compound accumulated. Analysis of the mass spectrum concluded that the new compound was an oxo-methyldesmosterol (**117**) with oxidation in the side chain. P450-17 however showed oxidation of the host-derived cycloartenol (**6**). Further bioinformatic investigation of P450-17 showed that orthologs are present in a previously unknown gene cluster in potato and tomato,

both non-producers of withanolides. Further research will be conducted in this group, regarding the adjoining genes and their role in triterpene metabolism in Solanaceae.

This work provides the basis for further research on withanolide biosynthesis in this group and furthermore raised the question of yet unknown phytosterol biosynthetic pathways in Solanaceae (P450-17 gene cluster).

Silencing experiments in *W. somnifera* and expression in *N. benthamiana* narrowed down a large pool of 35 gene candidates to five P450 (P450-7, 8, 14, 15, 17) and two DH (DH7, 11) candidates, that can now undergo further examination. For investigation of DH candidates, different substrates should be tested, as an 1-oxo substrate seems chemically more plausible. [66,67] Samuel E. Hakim is currently working on purifying such a substrate.

P450-17 showed oxidative activity on cycloartenol (**6**) in leaf disk assays. Since oxidized cycloartane-type triterpenoids have not been isolated from Solanaceae before, this topic is of particular interest to understand natural steroid diversity. Furthermore, oxidized cycloartanes were shown to exhibit biological activity that could be of medical interest. [353,354] As P450-17 homologous genes are organized in a gene cluster in potato and tomato, further experiments, e.g. heterologous expression of individual genes in *N. benthamiana,* can provide information whether cycloartenol (**6**), or the oxidation product generated by P450-17 (**118**), is being processed by the enzymes involved as well.

Furthermore, this work provided 5 P450 candidates (P450-7, 8, 14, 15, 17) that should be further investigated for involvement in withanolide biosynthesis. Additional testing can be conducted in a 24-methyldesmosterol (**3**) producing yeast strain with a modified terpene precursor flow. Such a strain will enable production of substrates *in situ*, and addition of gene candidates (e.g., P450-7) might help to reconstruct missing steps in the withanolide biosynthesis. Furthermore, reactivity of P450-17 can be investigated by adding a *squalene epoxidase* and *cycloartenol synthase* to a squalene-overproduction strain. The P450 genes obtained in this work were handed over to Dr. Karan Malhotra and now await further testing in yeast.

In summary, this work laid the technical foundation to study withanolides in this group and furthermore gained new insights into oxidation reactions involved in withanolide biosynthesis and other phytosterol pathways in Solanaceae, providing another building brick in enabling heterologous production of those medicinally active compounds.

8 Experimental

8.1 Material

If not indicated otherwise, chemicals used in this work were purchased in analytical grade (or higher) from *Carl Roth*, *Acros Organics* or *Sigma-Aldrich*. All enzymes used were purchased from *New England Biolabs*, *Invitrogen* or *Applied Biosystems*. All experiments that required kits were conducted according to the manufacturer's protocols with supplied buffers. All buffers and antibiotics were prepared with deionized water. For the preparation of growth media deionized water was further purified with a Barnstead GenPure Pro System from *Thermo Fisher Scientific* (subsequently referred to as ddH_2O).

8.1.1 Antibiotics, growth media and buffers

Antibiotics

Antibiotics were prepared as 1000x concentrated stock solutions. All substances were dissolved with the respective solvents (see Table 8.1) and were afterwards filtered (0.45 µm) and stored at –20 °C.

Table 8.1 Antibiotics used for this work

Antibiotic	Solvent	Stock concentration [mg/mL]	Working concentration [µg/mL]
Carbenicillin (Carb)	H_2O	50	50
Gentamycin (Gent)	H_2O	50	50
Kanamycin (Kan)	H_2O	50	50
Nourseothricin (Nat)	H_2O	50	50
Streptomycin (Strep)	H_2O	50	50
Rifampicin (Rif)	DMSO	50	50

Growth Media

Growth media used in this work were kindly prepared by BMWZ technical staff. All media were prepared with ddH_2O and were sterilized by autoclaving at 121 °C for 15 min (Autoclave 2100 Classic, Prestige Medical). The compositions are listed in Table 8.2.

Table 8.2 Growth media used in this work

Media	Composition
LB	0.5% (w/v) yeast extract 1% (w/v) tryptone 0.5% (w/v) NaCl
SOC	0.5% (w/v) yeast extract 2% (w/v) tryptone 0.06% (w/v) NaCl 0.02% (w/v) KCl 25 mM $MgCl_2$*6 H2O
TB Medium	2.4% (w/v) yeast extract 1.2% (w/v) tryptone 0.4% (w/v) glycerol 10% (v/v) 10X KPI buffer (0.17M KH_2PO_4 and 0.72M K_2HPO_4 added after autoclaving)
YPAD	1% (w/v) yeast extract 2% (w/v) tryptone 2% (w/v) D(+)-glucose monohydrate 0.03% (w/v) adenine
SD-Leu	0.67% (w/v) yeast nitrogen base 2% (w/v) D(+)-glucose monohydrate 0.15% (w/v) amino acid dropout mix without leucin
SD-Leu-Ura	0.67% (w/v) yeast nitrogen base 2% (w/v) D(+)-glucose monohydrate 0.15% (w/v) synthetic complete dropout mixture without leucin and uracil
SM-Ura	0.17% (w/v) yeast nitrogen base 0.5% (w/v) ammonium sulfate 2% (w/v) D(+)-glucose monohydrate or D(+)-galactose 0.077% (w/v) complete supplement mixture minus uracil
Agar	Composition [% (w/v)]
LB Agar	0.5% (w/v) yeast extract 1% (w/v) tryptone 0.5% (w/v) NaCl 1.5% (w/v) agar
YPAD Agar	1% (w/v) yeast extract 2% (w/v) tryptone 2% (w/v) D(+)-glucose monohydrate 0.03% (w/v) adenine 1.5% (w/v) agar
SD-Leu	0.67% (w/v) yeast nitrogen base 2% (w/v) D(+)-glucose monohydrate 0.15% (w/v) amino acid dropout mix without leucin 2% (w/v) agar

SD-Leu-Ura	0.67% (w/v) yeast nitrogen base 2% (w/v) D(+)-glucose monohydrate 0.15% (w/v) synthetic complete dropout mixture without leucin and uracil 2% (w/v) agar
SM-Ura	0.17% (w/v) yeast nitrogen base 0.5% (w/v) ammonium sulfate 2% (w/v) D(+)-glucose monohydrate 0.077% (w/v) complete supplement mixture minus uracil 1.5% (w/v) agar

Buffers

Buffers used in this works were prepared with deionized water and were sterilized with disposable FisherbrandTM PES Bottle Top Filters (pore size: 0.2 µM). The pH was controlled with a FiveEasy Standard pH Meter (Mettler Toledo) and was furthermore adjusted with 2 M HCl or 2 M NaOH. All buffers are listed in Table 8.3

Table 8.3 Buffers and solutions used in this work

Solutions/Buffers	Composition
Yeast transformation master mix	74% (v/v) 50% polyethylene glycol 11% (v/v) 1.0 M lithium acetate solution 15% (v/v) 2 mg/mL boiled salmon sperm DNA
Agroinfiltration buffer	10 mM $MgCl_2$ 10 mM MES 100 µM acetosyringone in DMSO (added from 100 mM stock solution immediately before use)
HEPES Buffer	5 mM (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid pH 7.5
TEK	50 mM Tris, pH 7.4 1 mM EDTA 0.1 M KCl
TES-B	50 mM Tris, pH 7.4 1 mM EDTA 0.6 M sorbitol
TEG	50 mM Tris, pH 7.4 1 mM EDTA 20% (v/v) glycerol
Lysis Buffer A	50 mM Tris-HCl pH 7.3 100 mM NaOH 10% (v/v) glycerol
Lysis Buffer B	50 mM Tris-HCl pH 7.5
Coomassie Dye	25% (v/v) acetic acid 10% (v/v) isopropanol 0.1% (w/v) coomassie

Coomassie Bleach	25% (v/v) acetic acid 10% (v/v) isopropanol
4× Laemmli buffer	10% (v/v) β-mercaptoethanol 0.25% (w/v) bromophenol blue 30% (v/v) glycerol 0.25% xylene cyanol
10× SDS running buffer	25 mM Tris-HCl pH 8.3 192 mM glycine 0.1% (w/v) SDS
50× TAE buffer	2 M Tris acetate pH 8.3 0.05 M EDTA

8.1.2 Strains, vectors and primers

Bacterial strains

E. coli strains used for this work were either obtained from the BMWZ strain collection (Top10, BL21(DE3)) or were purchased from Takara (Stellar). The information about the strains used is listed in Table 8.4.

Table 8.4 *Escherichia coli* strains used for this work

Bacterial strains	Reference	Genotype
Top10	Invitrogen	*E. coli* F$^-$ *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G [355]
BL21(DE3)	Invitrogen	*E. coli* B F$^-$ *ompT gal dcm lon hsdS$_B$* ($r_B^- m_B^-$) λ(DE3[*lacI lacUV5T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K-12}$(λ^S) [356,357]
Stellar	Takara Clontech	*E. coli* F$^-$, *endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, Φ80d lacZΔ M15, Δ (lacZYA - argF) U169, Δ (mrr - hsdRMS - mcrBC), ΔmcrA, λ–* [358]

Yeast Strains

Initial *S. cerevisiae* strains were kindly gifted by the O'Connor Group or purchased from EUROSCARF. All strains used in this work are listed in Table 8.5.

Table 8.5 *Saccharomyces cerevisiae* strains used for this work

Yeast strains (abbreviation)	Reference	Genotype
Δpep4	Ro *et al.*[207,359]	MATa, (*ura3-52, lys2-801, ade2-101, trp1-Δ63, his3-Δ200, leu2-Δ1, Δpep4*)
WAT11	Urban *et al.* [226]	MATα (*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*)
Δerg4	euroscarf	BY4742; MATα*; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; YGL012w::kanMX4*
Δerg4/erg5 (ASY01)	this work	BY4742; MATα*; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; YGL012w::kanMX4 ; YMR015C::natMX*
Δerg4/erg5 Pp7RED (ASY02)	this work	BY4742; MATα*; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; YGL012w::kanMX4 ; YMR015C::natMX; YEL021W::Pp7RED-LEU2*
Δerg4/erg5 St7RED (ASY03)	this work	BY4742; MATα*; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; YGL012w::kanMX4 ; YMR015C::natMX; YEL021W::St7RED-LEU2*
WAT11 pYES2-AtC4H	this work	MATα (*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*)
WAT11 pYES2-WsC4H	this work	MATα (*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*)
WAT11 pYES2-WsS22D1	this work	MATα (*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*)
WAT11 pYES2-WsS22D2	this work	MATα (*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*)
WAT11 pYES2-StC22H	this work	MATα (*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*)
WAT11 pYES2-P450-7	this work	MATα (*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*)
WAT11 pYES2-P450-8	this work	MATα (*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*)
WAT11 pYES2-P450-17	this work	MATα (*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*)

Agrobacterium strains

Initial *Agrobacterium tumefaciens* strains were kindly gifted by the O'Connor (GV3101) or Witte group. For infiltration of genetic material into plants the respective strain was transformed with either pTRV1 or pTRV2-GOI. All information about the strains used is listed in Table 8.6.

Table 8.6 *Agrobacterium tumefaciens* strains used in this work

A. tumefaciens strain (abbreviation)	Reference	Antibiotic resistance
GV3101	Koncz *et al.* [360]	Rif, Gent
AGL-1	Lazo *et al.* [288]	Rif, Carb
LBA4404	Ooms *et al.* [289]	Rif, Strep
EHA105	Hood *et al.* [290]	Rif

Plant Material

In this work, plant material from *Physalis peruviana* and *Withania somnifera* was analyzed. The seeds were initially obtained from FloraSelf or Asklepios Seeds, further experiments were conducted with self-harvested seeds.

All plants were cultivated with min. 11 h light exposure (min. 60 klxh) per day at 21-25 °C, 69% humidity. The seeds were planted in trays, and after germination (14 days) seedlings were transferred to single pots. Infiltration was performed with 4-5 week old plants.

Vectors

All vectors used and constructed in this work are shown in Table 8.7.

Table 8.7 Vectors constructed and used in this work

Vector	Application	Selection	Reference
pESC-Leu2d:AaCPR	yeast expression	Carb	Ro *et al.* [354]
pXP422	yeast genomic integration	Carb	Fang *et al.* [355]
pXP422-Pp7Red	yeast genomic integration	Carb	this work
pXP422-St7Red	yeast genomic integration	Carb	this work
pYES2	microsome assays	Carb	Invitrogen
pYES2-Pp24ISO	microsome assays	Carb	this work
pYES2-Ws24ISO	microsome assays	Carb	this work
pYES2-AtC4H	microsome assays	Carb	this work
pYES2-StPGA2	microsome assays	Carb	this work
pYES2-10503(WsS22D1)	microsome assays	Carb	this work
pYES2-21786 (WsS22D2)	microsome assays	Carb	this work
pYES2-4488 (WsC4H)	microsome assays	Carb	this work
pYES2-16059 (P450-7)	microsome assays	Carb	this work
pYES2-4953 (P450-8)	microsome assays	Carb	this work
pYES2-16175 (P450-17)	microsome assays	Carb	this work
pTRV1	VIGS	Kan	Valentine *et al.* [356]
pTRV2	VIGS	Kan	Liscombe *et al.* [357]
pTRV2-PpPDS	VIGS	Kan	this work
pTRV2-Pp24ISO	VIGS	Kan	this work
pTRV2-WsPDS	VIGS	Kan	this work
pTRV2-Ws24ISO	VIGS	Kan	this work
pTRV2-5012 (P450-1)	VIGS	Kan	this work
pTRV2-730 i2 (P450-3)	VIGS	Kan	this work
pTRV2-730 i6 (P450-4)	VIGS	Kan	this work
pTRV2-751 (P450-5)	VIGS	Kan	this work
pTRV2-26488 (P450-6)	VIGS	Kan	this work
pTRV2-16059 (P450-7)	VIGS	Kan	this work
pTRV2-4953 (P450-8)	VIGS	Kan	this work
pTRV2-2831 (P450-9)	VIGS	Kan	this work
pTRV2-2147 (P450-10)	VIGS	Kan	this work

pTRV2-2102 (P450-11)	VIGS	Kan	this work
pTRV2-1039 (P450-12)	VIGS	Kan	this work
pTRV2-3273 (P450-13)	VIGS	Kan	this work
pTRV2-21326 (P450-14)	VIGS	Kan	this work
pTRV2-18544 (P450-15)	VIGS	Kan	this work
pTRV2-16175 (P450-17)	VIGS	Kan	this work
pTRV2-13186 (P450-19)	VIGS	Kan	this work
pTRV2-406 (P450-21)	VIGS	Kan	this work
pTRV2-19528 (DH1)	VIGS	Kan	this work
pTRV2-11230 (DH2)	VIGS	Kan	this work
pTRV2-5593 (DH3)	VIGS	Kan	this work
pTRV2-20773 (DH5)	VIGS	Kan	this work
pTRV2-1799 (DH7)	VIGS	Kan	this work
pTRV2-24397 (DH11)	VIGS	Kan	this work
pHREAC	*N. benthamiana* expression	Kan	Peyret *et al.*[329]
pHREAC-StPGA2	*N. benthamiana* expression	Kan	this work
pHREAC-5012 (P450-1)	*N. benthamiana* expression	Kan	this work
pHREAC-730 i6 (P450-4)	*N. benthamiana* expression	Kan	this work
pHREAC-26488 (P450-6)	*N. benthamiana* expression	Kan	this work
pHREAC-16059 (P450-7)	*N. benthamiana* expression	Kan	this work
pHREAC-4953 (P450-8)	*N. benthamiana* expression	Kan	this work
pHREAC-2147 (P450-10)	*N. benthamiana* expression	Kan	this work
pHREAC-2102 (P450-11)	*N. benthamiana* expression	Kan	this work
pHREAC-1039 (P450-12)	*N. benthamiana* expression	Kan	this work
pHREAC-18544 (P450-15)	*N. benthamiana* expression	Kan	this work
pHREAC-16175 (P450-17)	*N. benthamiana* expression	Kan	this work
pHREAC-19528 (DH1)	*N. benthamiana* expression	Kan	this work
pHREAC-11230 (DH2)	*N. benthamiana* expression	Kan	this work
pHREAC-5593 (DH3)	*N. benthamiana* expression	Kan	this work
pHREAC-20773 (DH5)	*N. benthamiana* expression	Kan	this work
pHREAC-1799 (DH7)	*N. benthamiana* expression	Kan	this work
pHREAC-24397 (DH11)	*N. benthamiana* expression	Kan	this work

Primers

All oligonucleotides were synthesized by Sigma Aldrich and supplied in lyophilized form. They were dissolved in ddH_2O to a concentration of 10 µM and the stock solutions were stored at −20 °C. The oligonucleotides used in this work are summarized in the appendix, Table 10.1.

8.2 Methods

8.2.1 Biological Methods

Growth and maintenance of *E. coli* strains

All *E. coli* strains were grown on solid or liquid LB medium with appropriate antibiotics. The cultures were incubated at 37 °C, 200 rpm. Furthermore, glycerol stocks (25% glycerol) were stored at –80 °C for long term storage.

Transformation of chemically competent *E. coli* cells

For the transformation 2.5 µL (10 µL for Golden Gate Assembly) of plasmid were added to 50 µL of competent cells. The cells were left on ice for 30 min. The heat shock was performed at 42 °C for 45 s, again with immediate storage on ice for 2 min. Then 250 µL SOC medium was added and the cells were incubated for 1 h at 37 °C, 350 rpm. Of every cell line, 75 µL were streaked out on LB agar plates with appropriate antibiotics. The plates were incubated overnight at 37 °C.

Growth and maintenance of *S. cerevisiae* strains

All *S. cerevisiae* strains were grown on liquid or solid complete medium (YPAD) or appropriate selection media. If necessary for selection, appropriate antibiotics were added. The cultures were incubated at 30 °C, 160 rpm. Furthermore, glycerol stocks (25% glycerol) were stored at –80 °C for long term storage.

Transformation of *S. cerevisiae*

For the transformation of *S. cerevisiae*, 2 mL of overnight cultures were incubated in appropriate medium. The following day, 50 mL of appropriate medium were inoculated with 1 mL of overnight culture and were cultivated for 4-6 h. Yeast cultures were collected by centrifugation at 5,000 × *g*, 5 min, and washed with ddH_2O twice. Afterwards the remaining pellet was solved in transformation master mix (see Table 8.3). For each transformation reaction, 0.1-0.5 µg of DNA was added and the cells were incubated on ice for 3 min. The heat shock was performed at 42 °C for 40 min, again with immediate storage on ice for 2 min. The cells were collected by centrifugation and dissolved in 100 µL ddH_2O. The whole transformation batch was plated on prewarmed plated with appropriate selection media or antibiotics. The plates were incubated at 30 °C until colonies became visible. Colonies were verified by colony PCR (chapter 5.2.2, Polymerase chain reaction).

For the construction of the substrate producing yeast strain multiple primers were used for knockout, verification and integration of genes. All primers are listed in Table 10.1 in the appendix.

Production of compounds or protein in *S. cerevisiae*

For production of compounds of interest or enzymes for microsome isolation the respective yeast strain was streaked from cryogenic stocks on appropriate selection media. Precultures of the respective yeast strain were grown in appropriate media at 30 °C, 160 rpm, overnight. For compound isolation main cultures were grown for 2-4 days at 30 °C, 160 rpm. For protein production (expression from pYES2) main cultures were grown in SM-Ura + 2% Glc for one day. For induction the culture was harvested at 4,000 × *g* (10 min) and cells were resuspended in SM-Ura + 2% Gal. Induced yeast was cultivated for 1 more day at 30 °C, 160 rpm.
Afterwards cells were harvested at 4,000 × *g*, 10 min and further processed for metabolite analysis or microsome isolation.

Growth and maintenance of *A. tumefaciens* strains

All *A. tumefaciens* strains were grown on solid or liquid LB medium with appropriate antibiotics. The cultures were incubated at 28 °C, 160 rpm. Furthermore, glycerol stocks (25% glycerol) were stored at –80 °C for long term storage.

Transformation of *A. tumefaciens* strains

Glycerol stocks of competent *A. tumefaciens* strains stored at −80 °C were thawed on ice. Approximately 150 ng target plasmid were added to the cells and mixed gently. Afterwards, the mixture was transferred into a pre-cooled electroporation cuvette (2 mm gaps, BioRad Laboratories Hercules, CA, USA). The electroporation was performed in a MicroPulse electroporator (BioRad Laboratories) with a voltage of 2.2 kV and 5.5-6 ms pulse (3×). The bacteria were mixed with 150 µL SOC medium and incubated for 2 h at 28 °C in order to recover. Afterwards, 40 µL of the bacterial suspension were spread on a LB agar plate with appropriate antibiotics and incubated for two days at 28 °C. The resulting colonies were further examined by colony PCR.

Agroinfiltration of *P. peruviana* or *W. somnifera*

In order to perform virus induced gene silencing, pTRV1 and pTRV2-GOI were transformed into separate *A. tumefaciens* strains. To facilitate growth strains were streaked out from

cryogenic stocks on LB Agar with appropriate antibiotics and incubated for 2-3 days at 28 °C. Afterwards a single colony was transferred to 2 mL LB agar (with appropriate antibiotics) and incubated at 28 °C, 160 rpm overnight. A 10 mL main culture was inoculated with 200 µL of overnight culture and was incubated over night at 28 °C, 160 rpm.

The cells were harvested by centrifugation at 4000 rpm, 10 min. Afterwards, *A. tumefaciens* was resuspended in infiltration buffer (Table 8.3, OD_{600}=1) with freshly added acetosyringone and incubated at RT for at least 3 h. After incubation, the *A. tumefaciens* strains (pTRV1 and pTRV2-GOI, optionally plus pTRV2-PDS) were mixed in a 1:1 ratio and the bacterial suspension was infiltrated into the leaves of the respective plant.

The infected plants were covered with planting trays for 20 h to generate complete darkness. Afterwards infected plants were incubated for 28 days at 21 °C, 60% humidity and 200 µmol/min light intensity with 16 h of light exposure per day. First signs of PDS bleaching were observed after 7-10 days in *W. somnifera* and 21 days in *P. peruviana*.

8.2.2 Molecular Biology Methods

RNA extraction

For the analysis of *P. peruviana* or *W. somnifera* plants, the biological material was ground in liquid nitrogen and further extracted with the Nucleospin RNA plant Kit (Macherey Nagel) according to the manufacturers' protocol. For cloning experiments, a manual RNA extraction protocol was employed. Due to the length of this protocol, it is attached in the appendix (see section 1.1).

cDNA Synthesis

For RT-qPCR analysis cDNA was synthesized with the High-Capacity RNA-to-cDNA™ Kit (Applied Biosystems). For cloning purposes, the Superscript IV VILO (Invitrogen) was used. Both Kits were used according to the manufacturers' instructions.

Polymerase chain reaction (PCR)

For cloning purpose, the Platinum SuperFi II Master Mix (Invitrogen) was used. After reaction, the amplified DNA fragments were either used directly or purified using the NucleoSpin® Gel and PCR Clean-up kit (Macherey Nagel).

For quick testing of colonies, a colony PCR was performed. For *E. coli* and *A. tumefaciens* testing, 1 µL of culture was mixed into the prepared *OneTaq*® 2x Master Mix (NEB). The initial

annealing time was increased from 30 seconds to 5 min to lyse the cells of the picked colony. For *S. cerevisiae* testing, single colonies were boiled for 5 min in 20 µL 20 mM NaOH. 2 µL of the supernatant were given into the PCR mixture. All reactions were carried out according to the manufacturers' protocol. For *E. coli* and *A. tumefaciens* colony PCRs vector specific primers were used. If those were not available, the cloning primers were used. For yeast colony PCR either vector specific primers were used, or in case of genomic integration, special verification primers were designed. All Primers are listed in Table 10.1 of the appendix.

Agarose gel electrophoresis

In order to separate and analyze DNA samples, agarose gel electrophoresis was performed (BioRad system). DNA samples were mixed with 6× DNA loading Dye and 6× GelRed and 5 µL were loaded on the gel (1% (w/v) or 2% (w/v) agarose). As a molecular marker, the 1 kb DNA ladder or the 100 bp DNA ladder (both NEB) was used. Electrophoresis was carried out at 130 V for 30 min in 0.5× TAE buffer. The gel was photographed using an UVP camera (GelDoc XR+, Bio Rad).

Vector linearization

For restriction digestion, one or two sites per vector were chosen and the reaction was carried out with the respective enzymes (NEB) following the manufacturers' guidelines (pESC CPR: SpeI; pTRV2: BamHI, XhoI; pYES: BamHI, XhoI; pXP422: XhoI, SpeI; pEAQ-HT: AgeI, XhoI). All reactions were performed as 50 µL or 100 µL reactions in CutSmart buffer (NEB) at 37 °C for 3 h or overnight. Per 50 µL reaction volume, 10 U of restriction enzyme were used. In general, 10 µg of vector were digested in one reaction.

The samples were analyzed on a 1% agarose gel and were purified using the gel extraction method or direct PCR cleanup of the NucleoSpin® Gel and PCR Clean-up kit (Machery Nagel) according to the manufacturer's instructions.

DNA Assembly with In-Fusion

In order to generate plasmids, the In-Fusion® system (Takara) was used to integrate the gene of interest into a vector. Primers were designed according to the In-Fusion manual. All primers carried a 15 bp overlap to their insertion site. Vectors were linearized, using either single or double restriction enzyme digestion. For all assembly reactions 1 µL of vector was combined with 3 µL of unpurified PCR product and 1 µL of In-Fusion master mix. The reaction was

carried out at 50 °C for 15 min. The reaction mixture was stored at −20 °C or directly used for transformation in *E. coli*.

DNA Assembly with Golden Gate

The pHREAC vector contains a BsaI cutting site, which enables cloning via Golden Gate Assembly. All potential inserts were examined for BsaI restriction sites. If the insert contained a BsaI site, cloning was instead performed into pEAQ-HT. For all inserts without a BsaI site primers were designed, that contained an 18 bp overlap to the MCS of the pHREAC vector. In the assembly reaction 50 ng of pHREAC vector was combined with at least 100 ng of insert. Furthermore 1 µL of BsaI, 1 µL of T4-Ligase and 2 µL T4-Ligase buffer were added and water was filled up to 20 µL. The reaction was carried out at 37 °C for 2h, followed by an inactivation step at 80 °C for 10 min. The reaction mixture was stored at −20 °C or 10 µL were directly used for transformation in *E. coli*.

Isolation of plasmid DNA from *E. coli*

The isolation of plasmids from overnight culture was performed using the Wizard® Plus SV Miniprep Kit (Promega), following the instructions of the manufacturers' protocol. All elution steps were carried out in nuclease free water. For isolation of plasmids for cloning purposes overnight cultures with TB medium (Table 8.2) were used, as this resulted in higher yields of plasmid.

DNA sequencing

All generated plasmids were sequenced by Eurofins Genomics (Ebersberg) or Microsynth Seqlab GmbH (Göttingen).

8.2.3 Biochemical Methods

Microsome isolation

Microsome isolation was performed according to a protocol from the O'Connor group. After cultivation and induction of the respective yeast strain, cultures were harvested at 4,000 × *g* for 5 min. The cell pellet was resuspended in 10 mL TEK buffer (Table 8.3) and was incubated at room temperature for 5 min. Afterwards cells were again separated from the buffer by centrifugation at 4,000 × *g* for 5 min. Cells were then resuspended in 15-30 mL TES-B buffer (Table 8.3), depending on the quantity of resuspended cells. Lysis was carried out using a

French press at 12,000 psi two times. Cell debris was then separated from the suspension by centrifugation at 10,000 × *g* for 10 min at 4 °C. Afterwards, microsomes were harvested by centrifugation of the supernatant at 100,000 × *g* for 60 min at 4 °C. The pelleted microsomes were slowly resuspended in 0.5-3 mL TEG buffer (Table 8.3), depending on the size of the pellet. Aliquots were stored at −80°C for up to 1 year.

Microsome Assays

As uptake of sterol-like compounds is difficult in yeast, microsome assays should be performed with P450 candidates instead. For this purpose, 50 µL of purified microsomes were incubated with 50 µL 2 mM NADPH tetrasodium salt in HEPES buffer (final conc. 500 µM) and 100 µL 2.5 mM substrate in HEPES buffer (final conc. 1.25 mM) at 30 °C overnight. Hydrophobic substrates were dissolved in HEPES buffer using 5% (w/v) of methyl-β-cyclodextrin (MCD, **113**). For this purpose, **113** was added to the buffer and stirred until it was dissolved. Afterwards the HEPES-MCD buffer was added to a flask containing the respective substrate in surplus. The buffer-substrate suspension was then stirred until substantial clearing was visible, mostly overnight.

After incubation the reaction was quenched with either 40 µL 2 M HCl or chloroform and the reaction mixture was centrifuged at 10,000 × *g* to pellet the microsomes. The supernatant was extracted twice with 300 µL ethylacetate and the combined organic extracts were concentrated. Afterwards the extract was dissolved in 60 µL MeOH and was stored at 4 °C until LCMS analysis.

Leaf Disk Assays

As an alternative expression system *N. benthamiana* was used. Therefore, leaves of *N. benthamiana*, infiltrated with the *A. tumefaciens* GV3101, harboring pHREAC-GOI were harvested 3-5 dpi. A leaf disk (10 mm diameter) was punched out and placed in a 48-well plate with the adaxial epidermis facing up. The leaf disk was then soaked with 400 µL HEPES buffer containing substrate (240 µM for withanoside V aglycone (**130**); 45 µM for 24-methyldesmosterol (**3**), solubilized with 5% (w/v) methyl-β-cyclodextrin, (**113**)). The plates were sealed and the assays were incubated for 2 full day/night cycles in the phytochamber (22 °C, 16/8 h photoperiod, 150 µmol/s light intensity). Afterwards the leaf disks were separated from the buffer and ground with liquid nitrogen. For extraction see chapter 8.2.4.

8.2.4 Chemical Methods

Extraction withaferin A and withanolide A from *Withania somnifera* root powder

Withania somnifera root powder (10 g) was extracted with 50 mL H_2O/MeOH (3:1) for 3 h. After filtration and evaporation of the solvent under reduced pressure, the crude extract was dissolved in equal amounts H_2O and petroleum ether (25 mL each) for degreasing. The aqueous phase was extracted with 25 mL $CHCl_3$ and 25 mL 1-BuOH, two times each. The combined fractions of $CHCl_3$ and 1-BuOH, respectively, were concentrated under reduced pressure. Samples of each fraction were scanned via LCMS for their withanolide content.

For further separation, automated flash chromatography was used (Biotage Isolera Prime One System). Here, the SNAP KP-Sil cartridges (10 g, 50 μm) were used and flow rates were selected according to the manufacturers suggestion (36 mL/min). A gradient of chloroform (solvent A) and methanol (solvent B) was applied. The gradient is listed in Table 8.8. The detection was performed via UV-absorption at 254 nm and 290 nm. The resulting fractions were combined in two subfractions (F1: fractions 3-6, 6 mg; F2: fractions 7-10, 26 mg Figure 10.3) and were concentrated for LCMS and NMR analysis.

Table 8.8 Gradient for the automated flash chromatography.

column volumes	amount solvent A [%]
5	2
1	2
10	2-20
2	20

Extraction of irinans A and B and 4-β-hydroxywithanolide E from *Physalis peruviana* plants

Whole *Physalis peruviana* plants (140 g, 9 weeks old) were frozen and ground with liquid nitrogen. The plant powder was extracted with 500 mL H_2O:MeOH (3:1) at room temperature for 3 h. The suspension was filtrated and solvent was removed under reduced pressure. The crude extract was resuspended in 300 mL H_2O and defatted with 300 mL petroleum ether. Afterwards, the remaining aqueous layer was extracted with 2 × 300 mL $CHCl_3$ followed by 2 × 300 mL *n*-BuOH. This yielded a 660 mg petroleum ether fraction, 386 mg $CHCl_3$ fraction and 1174 mg *n*-BuOH fraction.

The $CHCl_3$ fraction was further separated using reversed phase flash chromatography (Biotage SNAP KP-C18-HS 30 g column) with a H_2O/MeOH gradient. The $CHCl_3$ crude extract was adsorbed onto Celite under reduced pressure for dry loading. For separation, a gradient from 30% to 95% MeOH was used. The resulting fractions were pooled guided by UV detection and

resulted in fraction F1 (subfractions 1-21, 102 mg), F2 (subfractions 22-31, 12 mg) and F3 (subfractions 32-45, 48 mg). For F2 no withanolides were detected in based on LCMS analysis and this fraction was therefore discarded.

Fraction F1 was further separated by preparative LCMS. The fraction was dissolved in MeOH (15 mg/mL) and 100 µL were injected per run. Separation was conducted using a gradient of water and acetonitrile [solvent A: H_2O + 0.05% formic acid; solvent B: acetonitrile + 0.045% formic acid; gradient: 10% to 90% B over 10 min, 20 mL/min]. The post–column flow was split (100:1) with the minority flow (mixed with 1 mL/min with MeOH + 0.045 % formic acid) being branched for in-line analysis by UV, ELSD and MS. The majority flow was collected. The following peaks were identified by NMR: t_R = 5.8 - 6.0 min (irinan A (**1**), 6 mg); 7.2 - 8.0 min (4β-hydroxywithanolide E (**73**), 49 mg); 9.0 - 9.2 min (irinan B (**2**), 1 mg). The collected fractions were evaporated under reduced pressure.

Main fraction F3 was furthermore separated by preparative LCMS as described above, yielding the known compounds withanolide E (**47**, t_R = 7.3 - 7.6 min, 6 mg), perulactone H (**74,** t_R = 7.6 - 7.8 min, 9 mg) and withanolide F (**79,** t_R = 7.8 - 8.5 min, 9 mg) which were identified by NMR.[361]

High throughput extraction of *Withania somnifera* VIGS samples

To analyse infiltrated *W. somnifera* samples, a high throughput method was established. 100 mg fresh weight leaves were frozen in liquid nitrogen and ground with a ball mill. Extraction was carried out in 500 µL of MeOH/H_2O (90:10) with emodin (0.1 mg/mL) as an internal standard for 2 minutes. After filtration, the samples were analysed via LCMS (5 µL injection volume).

Small scale extraction of *Withania somnifera* leaf disk assays

For analysis of leaf disk assays, leaf samples were separated from the buffer and were ground with liquid nitrogen. For GCMS analysis, 500 µL of saponification solution (10% (w/v) KOH in 90% EtOH) was added to the plant powder, followed by an incubation at 70 °C for 1 h. Afterwards 250 µL H_2O was added and the mixture was extracted twice with 500 µL hexane containing 0.01 mg/mL 5-α-cholestane. The combined hexane fractions were evaporated and the dried hexane phase was dissolved in 100 µL silylation solution (50% pyridine, 50% BSTFA with 1% TMCS). Silylation was carried out for 1 h at 70° C, and 1 µL of sample was injected into GCMS (see chapter 8.2.5).

For LCMS analysis plant powder was mixed with 300 µL H_2O and extracted twice with 300 µL ethyl acetate. The combined EtOAc fractions were evaporated and resolved in 50 µL MeOH and 10 µL were injected into LCMS (see chapter 8.2.5).

For high resolution LCMS plant powder was mixed with 500 µL acetonitrile. After coloring of the solvent (green) the plant powder was separated by centrifugation and washed a second time with 500 µl ACN. The combined ACN fractions were pre-purified with Oasis PRiME HLB 6 (30 mg, Waters) solid phase extraction cartridges. The flowthrough was collected and evaporated. The dried extract was resolved in 50 µL MeOH and was handed over to the Witte group for analysis. Here samples were measured on a Vanquish HPLC (ThermoFisher Scientific) coupled to a Q Exactive™ Plus Hybrid Quadrupol-Orbitrap™ mass spectrometer (ThermoFisher Scientific) using a Phenomenex Kinetex column (2.6 µm C8 100 Å 4.6 x 150 mm). The mobile phase consisted of H_2O+5 mM NH_4OAc (solvent A) and MeOH+5 mM NH_4OAc (solvent B) with a gradient of solvent B from 80% to 100% over 25 min at a flow rate of 0.4 mL/min.

Yeast sample preparation for metabolite analysis

In order to detect changes in yeast metabolite profiles GCMS analysis was performed. For this purpose, 5 mL precultures of the respective yeast strain were grown in appropriate media at 30 °C, 160 rpm, overnight. Main cultures (50 mL) were grown for 2-3 days at 30 °C, 160 rpm. Cells were harvested and washed with ddH_2O by centrifugation at 5,000 × *g*, 5 min.

To 100 mg of cell pellet, 100 µL of saponification solution (10% w/v KOH in 50% EtOH) and 50 µL of 5α-cholestane (0.02 mg/mL in hexane, internal standard) was added. After vigorous shaking (10 min, 30 s shaking, 30 s rest), equal amounts of hexane were added to extract the hydrophobic compounds. The samples were shaken vigorously for 3 min and were afterwards incubated at room temperature for 5 min. After separation of the phases, the hexane phase was transferred to a new vial and concentrated in a RVC 2-25 CDplus (Christ, Osterode am Harz; Germany) rotary vacuum concentrator for 90 min at a final pressure of 50 mbar and 1000 rpm. The dry sample was dissolved in 50 µL pyridine and 50 µL BSTFA/TMCS derivatisation solution to generate volatile derivates by silylation of phytosteroids. Samples were incubated for 1 h at 70 °C. If necessary, the samples were diluted with 200 µL pyridine for GC measurements.

Solubilization of hydrophobic substrates for microsome and leaf disk assays

Hydrophobic substrates were dissolved in HEPES buffer (50 mM, pH 7.5) using cyclodextrins (methyl-β-cyclodextrin (**113**) or (2-hydroxypropyl)-β-cyclodextrin (**114**)). In general, a defined amount of substrate was prepared in a flask by adding the respective amount from a stock solution and evaporating the solvent afterwards. The cyclodextrin powder was added to the HEPES buffer (for concentrations see Table 6.1) and were stirred until completely dissolved. Afterwards the HEPES-cyclodextrin buffer was added to the flask containing the respective substrate. The buffer-substrate suspension was then stirred until substantial clearing was visible, mostly overnight. For concentration measurements, the substrate suspension was then filtered over glass wool to remove undissolved substrate. Afterwards the substrate solution was diluted 5× with MeOH for LCMS analysis.

8.2.5 Analytical Methods

Real-time quantitative PCR

For the analysis of expression levels of silenced genes real-time quantitative PCR was performed. When performing RT-qPCR, the reliability of generated data depends on multiple factors, and several controls have to be established to validate the experiments. [362,363]

At first reference genes have to be selected, to normalize expression levels in between experiments. Typically, expression products of ubiquitously expressed genes are used (e.g. "housekeeping genes"). [364,365] Table 8.9 shows the respective reference genes selected for *Physalis peruviana* and *Withania somnifera*.

Table 8.9 Reference genes selected for qPCR.

Name	Accession *S. lycopersicum*	Best Hit *P. peruviana*	AA identity [%]	Best Hit *W. somnifera*	AA identity [%]
Actin	NM_001330119.1	c11400_g1_i1	96	TRINITY_DN3644 _c0_g1_i4	96
EF1a	X14449.1	c12307_g2_i1	100	TRINITY_DN6100 _c0_g1_i20	100
Tubulin	DQ205342.1	c38351_g1_i1	98	TRINITY_DN2481 _c4_g1_i1	98
Ubiquitin	X58253.1	c11766_g1_i3	98	TRINITY_DN11041 _c0_g1_i1	98

As a next step, primer pairs had to be selected for both reference genes and silencing GOIs. In general, a 50 to 150 bp amplicon should be generated, that, in case of silencing GOIs does not include the silencing construct, and furthermore is specific to the target gene. To prevent off targeting, primer pairs for all genes (reference genes and GOIs) were designed with the Primer BLAST tool (standard settings, product size 50-150 bp, custom database with the respective transcriptome). [366,367]

Furthermore, selected primer pairs had to be tested for potential background amplification by running no template control (NTC) and no reverse transcriptase (no RT) control and further for their efficiency (standard curve). Here, a primer pair was selected for each target or reference gene that did not show a second melting point, indicating off targeting, an efficiency of ca 100% (± 10%), and for reference genes a C_T value that was close to C_T values of GOIs. In summary, for *P. peruviana* RT-qPCR a primer pair for the reference gene elongation factor 1α showed the best properties, while analysis for *W. somnifera* was conducted with tubulin as a reference gene. All primers used for RT-qPCR are listed in Table 10.1 in the appendix.

For conduction of RT-qPCR RNA was isolated from 100 mg leaf samples with the RNA Plant Kit (DNase treatment included, Macherey Nagel) according to manufacturer's instructions. In the following RNA amounts were calculated to generate even amounts of cDNA, which was synthesized with the High-Capacity RNA-to-cDNA™ Kit (Applied Biosystems). The resulting cDNA was diluted 1:10. 2.5 µL of cDNA were pipetted per well and afterwards 7.5 µL of mastermix were added to avoid enclosure of air (master mix: 500 nM target primers and PowerUp SYBR Green Mastermix (Applied Biosystems)). Alternatively, the KiCqStart® SYBR® Green qPCR ReadyMix (Sigma-Aldrich) was tested, but the first setup was preferred due to higher efficiencies in standard curves. Both kits were used according to the manufactures' protocol in terms of concentration and PCR reaction settings (see also

Table 8.10). The reaction was carried out and monitored in a StepOne Real-Time PCR System (Applied Biosystems). The obtained data was exported to MS Excel and contained all information needed for quantification and quality control (e.g., C_t value, C_t mean, melting temperature). Comparative quantification of GOI expression was carried out with the $\Delta\Delta C_t$ method, [292] using the respective reference gene for normalization. For all experiments, 48-well were used, which typically allowed for 4 biological replicates (e.g., 4 GOI, 4 EV samples) and 3 technical replicates (3 for GOI and reference gene each) per plate.

Table 8.10 Fast cycling settings for RT-qPCR with the PowerUp SYBR Green Mastermix.

Step	Temperature	Duration	Cycles
UDG activation	50°C	120 s	hold
Polymerase activation	95°C	120 s	hold
Denaturation	95°C	3 s	40
Anneal/extend	60°C	30 s	
	Ramp rate	Temperature	Time
Melting curve	1.6°C/s	95°C	15 s
	1.6°C/s	60°C	60 s
	0.15°C/s	95°C	15 s

Analytical LCMS

For analytic purposes the Agilent 1260 II Infinity LC System (Santa Clara, CA, USA) was used, equipped with a quaternary pump, autosampler, column heater and diode array detector (DAD). Furthermore, a mass fragmentation detector, the Agilent G6125B single quadrupole MS was coupled. Here, depending on the separation problem, different columns and gradients were used. The columns used for the analytical measurement were reversed phase Agilent Poroshell 120 EC-C18 4.6 x 100 mm, 2.7 µm and Phenomenex Kinetex 2.6 µm C8 100 Å 4.6 x 150 mm.

Analytical LCMS was operated at a flow rate of 1 mL/min. For analytical purposes, different gradients were used on the different columns. An overview is listed in Table 8.11.

Table 8.11 Analytical LCMS methods used in this work.

Name	Column	Solvents	Gradient [% of B]
VIGS samples	C18	A: H_2O + 0.1% formic acid B: Acetonitrile	30%-90% over 10 min
Leaf Disk Assays P450s	C8	A: H_2O + 5mM NH_4OAc B: MeOH + 5mM NH_4OAc	80%-100% over 10 min
Leaf Disk Assays DHs	C18	A: H_2O + 0.1% formic acid B: Acetonitrile	20%-60% over 8 min, increase to 95% over 2 min
Microsome Assay C4H	C18	A: H_2O + 0.1% formic acid B: Acetonitrile	30%-90% over 10 min
Microsome Assay S22D	C8	A: H_2O + 5mM NH_4OAc B: MeOH + 5mM NH_4OAc	90%-100% over 10 min
Microsome Assay S22H	C8	A: H_2O + 5mM NH_4OAc B: MeOH + 5mM NH_4OAc	85%-100% over 10 min
Microsome Assay P450s	C8	A: H_2O + 5mM NH_4OAc B: MeOH + 5mM NH_4OAc	80%-100% over 10 min

For purification of compounds (chapter 1.1) a Waters time–directed autopurification system was used. It contains a Waters 2767 autosampler, the Waters 2545 pump system and a Phenomenex Kinetex Axia column (5 µm, C18, 100 Å, 21.2 x 250 mm) equipped with the Phenomenex Security Guard precolumn (Luna, C5, 300 Å). Here, subfractions were dissolved in MeOH to a concentration of 15 mg/mL. For each run 100 µL were injected. The solvents used for this purification was solvent A: H_2O + 0.05% formic acid; solvent B: acetonitrile + 0.045% formic acid. The gradient was 10% to 90% B over 10 min, at 20 mL/min. The post–column flow was split (100:1) with the minority flow being supplemented with 1 mL/min MeOH + 0.045 % formic acid for in-line analysis by UV, ELSD and MS. The majority flow was collected for structural analysis.

Gas chromatography (GC) analysis

Measurements were performed on a Hewlett Packard HP 6890N GC System (HP Inc. Palo Alto, CA, USA) with an OPTIMA 5MS (internal Ø: 0.25 mm, capillary layer Ø: 0.25 µm, length: 30 m) column (Macherey Nagel, Düren, Germany). A HP 5973 inert mass selective detector was attached to obtain mass spectrometric (MS) data. Analyses were conducted with an injection volume of 1 µL and a 1:5 split. Automated injection was conducted with a G2613A ALS injector (Agilent, Santa Clara, CA, USA). Here, helium was used as the carrier gas with a flow rate of 1.5 mL/min. For recording of mass spectrometric data, the solvent delay was set to 8 min. The temperature program used in this work was developed by Stephan Rohrbach in this group and is shown in Table 8.12.

Table 8.12 Separation program for GCMS analysis

Temperature [°C]	Time [min]	Temperature rise [°C/min]
100	0	30
275	5.8	3
300	14.16	0
300	30	0

Nuclear magnetic resonance analysis (NMR)

NMR spectra were obtained on a Bruker Ascend 400 MHz machine or a DRX 500 MHz machine and were performed by the NMR department of the Institute for Organic Chemistry, Leibniz University Hanover. The chemical shifts were observed relative to $CDCl_3$ (7.26 ppm) and expressed in δ values (ppm) and the coupling constants *J* were quoted in Hz. Analysis was conducted with TopSpin (Version 4.0.6, Bruker).

8.2.6 Bioinformatic Methods

Primer design for RT-qPCR

In order to prevent off targeting, all targets for RT-qPCR were analyzed with the Primer-BLAST [367] program from NCBI. The PCR product size was set at 50-150 bp and all targets were compared to a custom database containing the *Physalis peruviana* transcriptome. [155]

Analysis of Withanolide levels in VIGS plants

For relative quantification of withanolide levels in VIGS plants, the OpenLab CDS ChemStation Edition (Agilent) was used. Therefore, extracted ion chromatograms for the respective main fragment in ES+ mode (471 for withaferin A (**5**), 453 for withanolide D (**70**), 455 for withanolide A (**9**), 271 for emodin(**105**)) were given an integration cutoff (>15% of total area), so that the respective main peak was the only one integrated. The resulting withanolide areas were normalized with the emodin area.

Other software

For data analysis, Microsoft Office Excel (Microsoft, Redmond, WA, USA) was used. Visualization of data was performed with GraphPad Prism (version 8.4.2, GraphPad Software, San Diego, CA, USA).

For further analysis of LCMS data XCMS™ (Version 3.7.1, The Scripps Research Institute) and FreeStyle™ 1.8 SP2 (Version 1.8.63.0, Thermo Fisher Scientific) were used.

GC-FID and MS data were analyzed using OpenChrom Software (version 1.3.0). Initial assignment of unknown peaks according to their mass spectra was performed using the Golm metabolome database. [209] Alternatively, the program MSD ChemStation (version F.01.03.2357, Agilent Technologies) was used.

For the design of cloning experiments or primers, Geneious (version 11.1.3) was used.

For ideal placement of VIGS sequences, and to predict off targeting, the program siFi21 was used with default settings. [275] The transcriptome data of the respective plant [155,157] were used as custom databases.

Software that was used only for the operation of devices, which were readily supplied by the manufacturer, are not indicated separately.

9 References

1. Dixon, R. A. Natural products and plant disease resistance. *Nature* **411**, 843–847 (2001).
2. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803 (2020).
3. Mackowiak, P. A. Brief History of Antipyretic Therapy. *Clin. Infect. Dis.* **31**, S154–S156 (2000).
4. Dhar, N. *et al.* A Decade of Molecular Understanding of Withanolide Biosynthesis and In vitro Studies in Withania somnifera (L.) Dunal: Prospects and Perspectives for Pathway Engineering. *Front. Plant Sci.* **6**, (2015).
5. Damu, A. G. *et al.* Isolation, Structures, and Structure−Cytotoxic Activity Relationships of Withanolides and Physalins from Physalis angulata. *J. Nat. Prod.* **70**, 1146–1152 (2007).
6. Sang-ngern, M. *et al.* Withanolides derived from Physalis peruviana (Poha) with potential anti-inflammatory activity. *Bioorg. Med. Chem. Lett.* **26**, 2755–2759 (2016).
7. Lavie, D., Glotter, E. & Shvo, Y. Constituents of Withania somnifera Dun. III. The Side Chain of Withaferin A*,1. *J. Org. Chem.* **30**, 1774–1778 (1965).
8. Mulabagal, V. *et al.* Withanolide sulfoxide from Aswagandha roots inhibits nuclear transcription factor-kappa-B, cyclooxygenase and tumor cell proliferation. *Phytother. Res. PTR* **23**, 987–992 (2009).
9. Park, E.-J., Sang-Ngern, M., Chang, L. C. & Pezzuto, J. M. Physalactone and 4β-Hydroxywithanolide E Isolated from Physalis peruviana Inhibit LPS-Induced Expression of COX-2 and iNOS Accompanied by Abatement of Akt and STAT1. *J. Nat. Prod.* (2019) doi:10.1021/acs.jnatprod.8b00861.
10. Bessalle, R. & Lavie, D. Withanolide C, A chlorinated withanolide from Withania somnifera. *Phytochemistry* **31**, 3648–3651 (1992).
11. Chatterjee, S. *et al.* Comprehensive metabolic fingerprinting of Withania somnifera leaf and root extracts. *Phytochemistry* **71**, 1085–1094 (2010).
12. Trivedi, M. K., Panda, P., Sethi, K. K. & Jana, S. Metabolite Profiling in Withania somnifera Roots Hydroalcoholic Extract Using LC/MS, GC/MS and NMR Spectroscopy. *Chem. Biodivers.* **14**, e1600280 (2017).
13. Ha, J. W. *et al.* Structural Characterization of Withanolide Glycosides from the Roots of Withania somnifera and Their Potential Biological Activities. *Plants* **11**, 767 (2022).
14. Chen, L.-X., He, H. & Qiu, F. Natural withanolides : an overview. *Nat. Prod. Rep.* **28**, 705–740 (2011).
15. Glotter, E. Withanolides and related ergostane-type steroids. *Nat. Prod. Rep.* **8**, 415–440 (1991).
16. Ray, A. B. & Gupta, M. Withasteroids, a Growing Group of Naturally Occurring Steroidal Lactones. in *Fortschritte der Chemie organischer Naturstoffe / Progress in the Chemistry of Organic Natural Products* (eds. Cárdenas, J. et al.) 1–106 (Springer Vienna, 1994). doi:10.1007/978-3-7091-9281-8_1.
17. Yoshida, M., Hoshi, A., Kuretani, K., Ishiguro, M. & Ikekawa, N. RELATIONSHIP BETWEEN CHEMICAL STRUCTURE AND ANTITUMOR ACTIVITY OF WITHAFERIN A ANALOGUES. *J. Pharmacobiodyn.* **2**, 92–97 (1979).
18. Zhang, H. *et al.* Cytotoxic Withanolide Constituents of Physalis longifolia. *J. Nat. Prod.* **74**, 2532–2544 (2011).
19. Jana, C. K. *et al.* Synthesis of Withanolide A, Biological Evaluation of Its Neuritogenic Properties, and Studies on Secretase Inhibition. *Angew. Chem. Int. Ed.* **50**, 8407–8411 (2011).

20. Ohkubo, M., Hirai, G. & Sodeoka, M. Synthesis of the DFGH ring system of Type B Physalins: Highly Oxygenated, Cage-Shaped Molecules. *Angew. Chem. Int. Ed.* **48**, 3862–3866 (2009).
21. Alberts, A. W. *et al.* Mevinolin: a highly potent competitive inhibitor of hydroxymethylglutaryl-coenzyme A reductase and a cholesterol-lowering agent. *Proc. Natl. Acad. Sci. U. S. A.* **77**, 3957–3961 (1980).
22. Fleming, A. On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of B. influenzæ. *Br. J. Exp. Pathol.* **10**, 226–236 (1929).
23. Martin, W. R., Eades, C. G., Thompson, J. A., Huppler, R. E. & Gilbert, P. E. The effects of morphine- and nalorphine- like drugs in the nondependent and morphine-dependent chronic spinal dog. *J. Pharmacol. Exp. Ther.* **197**, 517–532 (1976).
24. Luo, X.-D. & Shen, C.-C. The chemistry, pharmacology, and clinical applications of qinghaosu (Artemisinin) and its derivatives. *Med. Res. Rev.* **7**, 29–52 (1987).
25. Cheng, A.-X. *et al.* Plant Terpenoids: Biosynthesis and Ecological Functions. *J. Integr. Plant Biol.* **49**, 179–186 (2007).
26. Hartmann, M.-A. Plant sterols and the membrane environment. *Trends Plant Sci.* **3**, 170–175 (1998).
27. Newman, J. D. & Chappell, J. Isoprenoid Biosynthesis in Plants: Carbon Partitioning Within the Cytoplasmic Pathway. *Crit. Rev. Biochem. Mol. Biol.* **34**, 95–106 (1999).
28. Bach, T. J. Some new aspects of isoprenoid biosynthesis in plants—A review. *Lipids* **30**, 191–202 (1995).
29. Eisenreich, W., Bacher, A., Arigoni, D. & Rohdich, F. Biosynthesis of isoprenoids via the non-mevalonate pathway. *Cell. Mol. Life Sci. CMLS* **61**, 1401–1426 (2004).
30. Goldstein, J. L. & Brown, M. S. Regulation of the mevalonate pathway. *Nature* **343**, 425–430 (1990).
31. Lichtenthaler, H. K. The 1-Deoxy-D-Xylulose-5-Phosphate Pathway of Isoprenoid Biosynthesis in Plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 47–65 (1999).
32. Chaurasiya, N. D., Sangwan, N. S., Sabir, F., Misra, L. & Sangwan, R. S. Withanolide biosynthesis recruits both mevalonate and DOXP pathways of isoprenogenesis in Ashwagandha <Emphasis Type="Italic">Withania somnifera</Emphasis> L. (Dunal). *Plant Cell Rep.* **31**, 1889–1897 (2012).
33. Akhtar, N., Gupta, P., Sangwan, N. S., Sangwan, R. S. & Trivedi, P. K. Cloning and functional characterization of 3-hydroxy-3-methylglutaryl coenzyme A reductase gene from Withania somnifera: an important medicinal plant. *Protoplasma* **250**, 613–622 (2013).
34. Gupta, P. *et al.* Cloning and characterization of 2-C-methyl-D-erythritol-4-phosphate pathway genes for isoprenoid biosynthesis from Indian ginseng, Withania somnifera. *Protoplasma* **250**, 285–295 (2013).
35. Oldfield, E. & Lin, F.-Y. Terpene Biosynthesis: Modularity Rules. *Angew. Chem. Int. Ed.* **51**, 1124–1137 (2012).
36. Chaurasiya, N. D., Gupta, V. K. & Sangwan, R. S. Leaf ontogenic phase-related dynamics of withaferin a and withanone biogenesis in ashwagandha (Withania somnifera Dunal.)-an important medicinal herb. *J. Plant Biol.* **50**, 508 (2007).
37. Gupta, P., Akhtar, N., Tewari, S. K., Sangwan, R. S. & Trivedi, P. K. Differential expression of farnesyl diphosphate synthase gene from Withania somnifera in different chemotypes and in response to elicitors. *Plant Growth Regul.* **65**, 93–100 (2011).
38. Poulter, C. D. Biosynthesis of non-head-to-tail terpenes. Formation of 1'-1 and 1'-3 linkages. *Acc. Chem. Res.* **23**, 70–77 (1990).
39. Seo, J.-W. *et al.* Overexpression of squalene synthase in Eleutherococcus senticosus increases phytosterol and triterpene accumulation. *Phytochemistry* **66**, 869–877 (2005).

40. Bhat, W. W. *et al.* Molecular cloning, bacterial expression and promoter analysis of squalene synthase from Withania somnifera (L.) Dunal. *Gene* **499**, 25–36 (2012).
41. Kirschning, Andreas. Skript Terpene. (2018).
42. Jarstfer, M. B., Blagg, B. S. J., Rogers, D. H. & Poulter, C. D. Biosynthesis of Squalene. Evidence for a Tertiary Cyclopropylcarbinyl Cationic Intermediate in the Rearrangement of Presqualene Diphosphate to Squalene. *J. Am. Chem. Soc.* **118**, 13089–13090 (1996).
43. Blagg, B. S. J., Jarstfer, M. B., Rogers, D. H. & Poulter, C. D. Recombinant Squalene Synthase. A Mechanism for the Rearrangement of Presqualene Diphosphate to Squalene. *J. Am. Chem. Soc.* **124**, 8846–8853 (2002).
44. Han, J.-Y., In, J.-G., Kwon, Y.-S. & Choi, Y.-E. Regulation of ginsenoside and phytosterol biosynthesis by RNA interferences of squalene epoxidase gene in Panax ginseng. *Phytochemistry* **71**, 36–46 (2010).
45. Razdan, S. *et al.* Molecular characterization and promoter analysis of squalene epoxidase gene from Withania somnifera (L.) Dunal. *Mol. Biol. Rep.* **40**, 905–916 (2013).
46. Thoma, R. *et al.* Insight into steroid scaffold formation from the structure of human oxidosqualene cyclase. *Nature* **432**, 118–122 (2004).
47. J. Stephenson, M., A. Field, R. & Osbourn, A. The protosteryl and dammarenyl cation dichotomy in polycyclic triterpene biosynthesis revisited: has this 'rule' finally been broken? *Nat. Prod. Rep.* **36**, 1044–1052 (2019).
48. Bloch, K. E. Sterol, Structure and Membrane Function. *Crit. Rev. Biochem.* **14**, 47–92 (1983).
49. Phillips, D. R., Rasbery, J. M., Bartel, B. & Matsuda, S. P. Biosynthetic diversity in plant triterpene cyclization. *Curr. Opin. Plant Biol.* **9**, 305–314 (2006).
50. Plant cholesterol biosynthetic pathway overlaps with phytosterol metabolism | Nature Plants. https://www.nature.com/articles/nplants2016205.
51. Dhar, N. *et al.* Cloning and Functional Characterization of Three Branch Point Oxidosqualene Cyclases From Withania somnifera (L.) Dunal. *J. Biol. Chem.* jbc.M114.571919 (2014) doi:10.1074/jbc.M114.571919.
52. Sonawane, P. D. *et al.* Short-chain dehydrogenase/reductase governs steroidal specialized metabolites structural diversity and toxicity in the genus Solanum. *Proc. Natl. Acad. Sci.* **115**, E5419–E5428 (2018).
53. Benveniste, P. Sterol Biosynthesis. *Annu. Rev. Plant Physiol.* **37**, 275–308 (1986).
54. DARNET, S. & RAHIER, A. Plant sterol biosynthesis: identification of two distinct families of sterol 4alpha-methyl oxidases. *Biochem. J.* **378**, 889–898 (2004).
55. Heintz, R. & Benveniste, P. Plant Sterol Metabolism: ENZYMATIC CLEAVAGE OF THE 9β,19β-CYCLOPROPANE RING OF CYCLOPROPYL STEROLS IN BRAMBLE TISSUE CULTURES. *J. Biol. Chem.* **249**, 4267–4274 (1974).
56. O'Brien, M., Chantha, S.-C., Rahier, A. & Matton, D. P. Lipid Signaling in Plants. Cloning and Expression Analysis of the Obtusifoliol 14α-Demethylase from Solanum chacoense Bitt., a Pollination- and Fertilization-Induced Gene with Both Obtusifoliol and Lanosterol Demethylase Activity. *Plant Physiol.* **139**, 734–749 (2005).
57. Schrick, K. *et al.* FACKEL is a sterol C-14 reductase required for organized cell division and expansion in Arabidopsis embryogenesis. *Genes Dev.* **14**, 1471–1484 (2000).
58. Grebenok, R. J. *et al.* Isolation and characterization of an Arabidopsis thaliana C-8,7 sterol isomerase: functional and structural similarities to mammalian C-8,7 sterol isomerase/emopamil-binding protein. *Plant Mol. Biol.* **38**, 807–815 (1998).
59. Schaller, H., Bouvier-Navé, P. & Benveniste, P. Overexpression of an Arabidopsis cDNA Encoding a Sterol-C241-Methyltransferase in Tobacco Modifies the Ratio of 24-Methyl Cholesterol to Sitosterol and Is Associated with Growth Reduction. *Plant Physiol.* **118**, 461–469 (1998).

60. Darnet, S., Bard, M. & Rahier, A. Functional identification of sterol-4α-methyl oxidase cDNAs from Arabidopsis thaliana by complementation of a yeast erg25 mutant lacking sterol-4α-methyl oxidation1. *FEBS Lett.* **508**, 39–43 (2001).
61. Husselstein, T., Schaller, H., Gachotte, D. & Benveniste, P. Δ7-Sterol-C5-desaturase: molecular characterization and functional expression of wild-type and mutant alleles. *Plant Mol. Biol.* **39**, 891–906 (1999).
62. Lecain, E., Chenivesse, X., Spagnoli, R. & Pompon, D. Cloning by metabolic interference in yeast and enzymatic characterization of Arabidopsis thaliana sterol delta 7-reductase. *J. Biol. Chem.* **271**, 10866–10873 (1996).
63. Youn, J. H. *et al.* Function and molecular regulation of DWARF1 as a C-24 reductase in brassinosteroid biosynthesis in Arabidopsis. *J. Exp. Bot.* **69**, 1873–1886 (2018).
64. Knoch, E. *et al.* Third DWF1 paralog in Solanaceae, sterol Δ24-isomerase, branches withanolide biosynthesis from the general phytosterol pathway. *Proc. Natl. Acad. Sci.* **115**, E8096–E8103 (2018).
65. Lockley, W. J. S., Rees, H. H. & Goodwin, T. W. Biosynthesis of steroidal withanolides in Withania somnifera. *Phytochemistry* **15**, 937–939 (1976).
66. Velde, V. V. & Lavie, D. New withanolides of biogenetic interest from Withania somnifera. *Phytochemistry* **20**, 1359–1364 (1981).
67. Velde, V. V., Lavie, D., Budhiraja, R. D., Sudhir, S. & Garg, K. N. Potential biogenetic precursors of withanolides from withania coagulans. *Phytochemistry* **22**, 2253–2257 (1983).
68. Lischewski, M. *et al.* Withanolide glycosides from dunalia australis*. *Phytochemistry* **31**, 939–942 (1992).
69. Morikawa, T. *et al.* Cytochrome P450 CYP710A Encodes the Sterol C-22 Desaturase in Arabidopsis and Tomato. *Plant Cell* **18**, 1008–1022 (2006).
70. Itkin, M. *et al.* Biosynthesis of Antinutritional Alkaloids in Solanaceous Crops Is Mediated by Clustered Genes. *Science* **341**, 175–179 (2013).
71. Matsuda, H., Murakami, T., Kishi, A. & Yoshikawa, M. Structures of withanosides I, II, III, IV, V, VI, and VII, new withanolide glycosides, from the roots of Indian Withania somnifera DUNAL. and inhibitory activity for tachyphylaxis to clonidine in isolated guinea-pig ileum. *Bioorg. Med. Chem.* **9**, 1499–1507 (2001).
72. Sharma, L. K., Madina, B. R., Chaturvedi, P., Sangwan, R. S. & Tuli, R. Molecular cloning and characterization of one member of 3β-hydroxy sterol glucosyltransferase gene family in Withania somnifera. *Arch. Biochem. Biophys.* **460**, 48–55 (2007).
73. Singh, G. *et al.* Silencing of sterol glycosyltransferases modulates the withanolide biosynthesis and leads to compromised basal immunity of Withania somnifera. *Sci. Rep.* **6**, 25562 (2016).
74. Sono, M., Roach, M. P., Coulter, E. D. & Dawson, J. H. Heme-Containing Oxygenases. *Chem. Rev.* **96**, 2841–2888 (1996).
75. Guengerich, F. P. Common and Uncommon Cytochrome P450 Reactions Related to Metabolism and Chemical Toxicity. *Chem. Res. Toxicol.* **14**, 611–650 (2001).
76. Farrow, S. C. & Facchini, P. J. Functional diversity of 2-oxoglutarate/Fe(II)-dependent dioxygenases in plant metabolism. *Front. Plant Sci.* **5**, (2014).
77. Rana, S. *et al.* Molecular characterization of two A-type P450s, WsCYP98A and WsCYP76A from Withania somnifera (L.) Dunal: expression analysis and withanolide accumulation in response to exogenous elicitations. *BMC Biotechnol.* **14**, 89 (2014).
78. Srivastava, S. *et al.* Light and auxin responsive cytochrome P450s from Withania somnifera Dunal: cloning, expression and molecular modelling of two pairs of homologue genes with differential regulation. *Protoplasma* **252**, 1421–1437 (2015).
79. Montellano, P. R. O. de. *Cytochrome P450: Structure, Mechanism, and Biochemistry*. (Springer Science & Business Media, 2005).

80. Denisov, I. G., Makris, T. M., Sligar, S. G. & Schlichting, I. Structure and Chemistry of Cytochrome P450. *Chem. Rev.* **105**, 2253–2278 (2005).
81. Mizutani, M. & Sato, F. Unusual P450 reactions in plant secondary metabolism. *Arch. Biochem. Biophys.* **507**, 194–203 (2011).
82. Berka, K., Hendrychová, T., Anzenbacher, P. & Otyepka, M. Membrane Position of Ibuprofen Agrees with Suggested Access Path Entrance to Cytochrome P450 2C9 Active Site. *J. Phys. Chem. A* **115**, 11248–11255 (2011).
83. Nelson, D. & Werck-Reichhart, D. A P450-centric view of plant evolution. *Plant J.* **66**, 194–211 (2011).
84. Meunier, B., de Visser, S. P. & Shaik, S. Mechanism of Oxidation Reactions Catalyzed by Cytochrome P450 Enzymes. *Chem. Rev.* **104**, 3947–3980 (2004).
85. Munro, A. W., Girvan, H. M., Mason, A. E., Dunford, A. J. & McLean, K. J. What makes a P450 tick? *Trends Biochem. Sci.* **38**, 140–150 (2013).
86. Hannemann, F., Bichet, A., Ewen, K. M. & Bernhardt, R. Cytochrome P450 systems—biological variations of electron transport chains. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1770**, 330–344 (2007).
87. Hanukoglu, I. Electron Transfer Proteins of Cytochrome P450 Systems. in *Advances in Molecular and Cell Biology* (ed. Bittar, E. E.) vol. 14 29–56 (Elsevier, 1996).
88. Hawkes, D. B., Adams, G. W., Burlingame, A. L., Ortiz de Montellano, P. R. & De Voss, J. J. Cytochrome P450(cin) (CYP176A), isolation, expression, and characterization. *J. Biol. Chem.* **277**, 27725–27732 (2002).
89. Dawson, J. H. & Sono, Masanori. Cytochrome P-450 and chloroperoxidase: thiolate-ligated heme enzymes. Spectroscopic determination of their active-site structures and mechanistic implications of thiolate ligation. *Chem. Rev.* **87**, 1255–1276 (1987).
90. Poulos, T. L., Finzel, B. C. & Howard, A. J. High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* **195**, 687–700 (1987).
91. Poulos, T. L., Cupp-Vickery, J. & Li, H. Structural Studies on Prokaryotic Cytochromes P450. in *Cytochrome P450: Structure, Mechanism, and Biochemistry* (ed. de Montellano, P. R. O.) 125–150 (Springer US, 1995). doi:10.1007/978-1-4757-2391-5_4.
92. Guengerich, F. P. & Johnson, W. W. Kinetics of Ferric Cytochrome P450 Reduction by NADPH−Cytochrome P450 Reductase: Rapid Reduction in the Absence of Substrate and Variations among Cytochrome P450 Systems. *Biochemistry* **36**, 14741–14750 (1997).
93. Bernhardt, R. Cytochrome P450: Structure, function, and generation of reactive oxygen species. in *Reviews of Physiology Biochemistry and Pharmacology, Volume 127: Volume: 127* 137–221 (Springer, 1996). doi:10.1007/BFb0048267.
94. Graham-Lorence, S., Amarneh, B., White, R. E., Peterson, J. A. & Simpson, E. R. A three-dimensional model of aromatase cytochrome P450. *Protein Sci. Publ. Protein Soc.* **4**, 1065–1080 (1995).
95. Sethumadhavan, K. & Bellino, F. L. Human placental estrogen synthetase (aromatase). Effect of environment on the kinetics of protein-protein and substrate-protein interactions and the production of 19-oxygenated androgen intermediates in the purified reconstituted cytochrome P450 enzyme system. *J. Steroid Biochem. Mol. Biol.* **39**, 381–394 (1991).
96. Kao, Y.-C., Korzekwa, K. R., Laughton, C. A. & Chen, S. Evaluation of the mechanism of aromatase cytochrome P450: A site-directed mutagenesis study. *Eur. J. Biochem.* **268**, 243–251 (2001).
97. Sohl, C. D. & Guengerich, F. P. Kinetic Analysis of the Three-step Steroid Aromatase Reaction of Human Cytochrome P450 19A1*. *J. Biol. Chem.* **285**, 17734–17743 (2010).
98. Ro, D.-K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).
99. Fujita, S. *et al.* Arabidopsis CYP90B1 catalyses the early C-22 hydroxylation of C27, C28 and C29 sterols. *Plant J.* **45**, 765–774 (2006).

100. Schatz, M. C., Witkowski, J. & McCombie, W. R. Current challenges in de novo plant genome sequencing and assembly. *Genome Biol.* **13**, 243 (2012).
101. Yesilirmak, F. & Sayers, Z. Heterelogous Expression of Plant Genes. *International Journal of Plant Genomics* https://www.hindawi.com/journals/ijpg/2009/296482/ (2009) doi:10.1155/2009/296482.
102. Goodin, M. M., Zaitlin, D., Naidu, R. A. & Lommel, S. A. Nicotiana benthamiana: Its History and Future as a Model for Plant–Pathogen Interactions. *Mol. Plant-Microbe Interactions®* **21**, 1015–1026 (2008).
103. Purkayastha, A. & Dasgupta, I. Virus-induced gene silencing: A versatile tool for discovery of gene functions in plants. *Plant Physiol. Biochem.* **47**, 967–976 (2009).
104. Rakoczy-trojanowska, M. Alternative methods of plant transformation: a short review. *Cell. Mol. Biol. Lett.* 849–858 (2002).
105. Messens, E., Lenaerts, A., Hedges, R. W. & Montagu, M. V. Agrocinopine A, a phosphorylated opine is secreted from crown gall cells. *EMBO J.* **4**, 571–577 (1985).
106. Hwang, H.-H., Yu, M. & Lai, E.-M. Agrobacterium-mediated plant transformation: biology and applications. *Arab. Book* **15**, (2017).
107. Zambryski, P., Tempe, J. & Schell, J. Transfer and function of T-DNA genes from Agrobacterium Ti and Ri plasmids in plants. *Cell* **56**, 193–201 (1989).
108. Lacroix, B. & Citovsky, V. The roles of bacterial and host plant factors in Agrobacterium-mediated genetic transformation. *Int. J. Dev. Biol.* **57**, 467–481 (2013).
109. Bomhoff, G. *et al.* Octopine and nopaline synthesis and breakdown genetically controlled by a plasmid of Agrobacterium tumefaciens. *Mol. Gen. Genet. MGG* **145**, 177–181 (1976).
110. Montoya, A. L., Chilton, M. D., Gordon, M. P., Sciaky, D. & Nester, E. W. Octopine and nopaline metabolism in Agrobacterium tumefaciens and crown gall tumor cells: role of plasmid genes. *J. Bacteriol.* **129**, 101–107 (1977).
111. Zerbini, F. M., Silva, F. N. da, Urquiza, G. P. C. & Basso, M. F. Chapter 8 - Transgenic Plants. in *Biotechnology and Plant Breeding* (eds. Borem, A. & Fritsche-Neto, R.) 179–199 (Academic Press, 2014). doi:10.1016/B978-0-12-418672-9.00008-8.
112. Matthysse, A. G., Holmes, K. V. & Gurlitz, R. H. Elaboration of cellulose fibrils by Agrobacterium tumefaciens during attachment to carrot cells. *J. Bacteriol.* **145**, 583–595 (1981).
113. Hooykaas, Paul. J. J. & Beijersbergen, A. G. M. The Virulence System of Agrobacterium Tumefaciens. *Annu. Rev. Phytopathol.* **32**, 157–181 (1994).
114. Yuan, Z.-C., Liu, P., Saenkham, P., Kerr, K. & Nester, E. W. Transcriptome profiling and functional analysis of Agrobacterium tumefaciens reveals a general conserved response to acidic conditions (pH 5.5) and a complex acid-mediated signaling involved in Agrobacterium-plant interactions. *J. Bacteriol.* **190**, 494–507 (2008).
115. Wu, C.-F., Lin, J.-S., Shaw, G.-C. & Lai, E.-M. Acid-induced type VI secretion system is regulated by ExoR-ChvG/ChvI signaling cascade in Agrobacterium tumefaciens. *PLoS Pathog.* **8**, e1002938 (2012).
116. Zechner, E. L., Lang, S. & Schildbach, J. F. Assembly and mechanisms of bacterial type IV secretion machines. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 1073–1087 (2012).
117. Lai, E.-M. & Kado, C. I. Processed VirB2 Is the Major Subunit of the Promiscuous Pilus of Agrobacterium tumefaciens. *J. Bacteriol.* **180**, 2711–2717 (1998).
118. Wu, H.-Y., Chen, C.-Y. & Lai, E.-M. Expression and Functional Characterization of the Agrobacterium VirB2 Amino Acid Substitution Variants in T-pilus Biogenesis, Virulence, and Transient Transformation Efficiency. *PLOS ONE* **9**, e101142 (2014).
119. Ballas, N. & Citovsky, V. Nuclear localization signal binding protein from Arabidopsis mediates nuclear import of Agrobacterium VirD2 protein. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10723–10728 (1997).

120. Koncz, C. *et al.* High-frequency T-DNA-mediated gene tagging in plants. *Proc. Natl. Acad. Sci.* **86**, 8467–8471 (1989).
121. Bally, J. *et al.* The Rise and Rise of Nicotiana benthamiana: A Plant for All Reasons. *Annu. Rev. Phytopathol.* **56**, 405–426 (2018).
122. Tyurin, A. A., Suhorukova, A. V., Kabardaeva, K. V. & Goldenkova-Pavlova, I. V. Transient Gene Expression is an Effective Experimental Tool for the Research into the Fine Mechanisms of Plant Gene Function: Advantages, Limitations, and Solutions. *Plants* **9**, 1187 (2020).
123. Sainsbury, F., Thuenemann, E. C. & Lomonossoff, G. P. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol. J.* **7**, 682–693 (2009).
124. Voinnet, O., Rivas, S., Mestre, P. & Baulcombe, D. Retracted: An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J.* **33**, 949–956 (2003).
125. Becker, A. & Lange, M. VIGS – genomics goes functional. *Trends Plant Sci.* **15**, 1–4 (2010).
126. Lange, M., Yellina, A. L., Orashakova, S. & Becker, A. Virus-Induced Gene Silencing (VIGS) in Plants: An Overview of Target Species and the Virus-Derived Vector Systems. in *Virus-Induced Gene Silencing* 1–14 (Humana Press, Totowa, NJ, 2013). doi:10.1007/978-1-62703-278-0_1.
127. Voinnet, O. RNA silencing as a plant immune system against viruses. *Trends Genet.* **17**, 449–459 (2001).
128. Baulcombe, D. Viruses and gene silencing in plants. *Arch. Virol. Suppl.* **15**, 189–201 (1999).
129. Ding, S.-W. & Voinnet, O. Antiviral Immunity Directed by Small RNAs. *Cell* **130**, 413–426 (2007).
130. Baulcombe, D. RNA silencing in plants. *Nature* **431**, 356–363 (2004).
131. Hammond, S. M. Dicing and slicing: The core machinery of the RNA interference pathway. *FEBS Lett.* **579**, 5822–5829 (2005).
132. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nat. Rev. Mol. Cell Biol.* **9**, 22–32 (2008).
133. Peters, L. & Meister, G. Argonaute Proteins: Mediators of RNA Silencing. *Mol. Cell* **26**, 611–623 (2007).
134. Rt, O., Jj, C. & Hb, S. RNAi-associated ssRNA-specific ribonucleases in Tombusvirus P19 mutant-infected plants and evidence for a discrete siRNA-containing effector complex. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 1714–1719 (2007).
135. Silhavy, D. & Burgyán, J. Effects and side-effects of viral RNA silencing suppressors on short RNAs. *Trends Plant Sci.* **9**, 76–83 (2004).
136. Waterhouse, P. M. & Fusaro, A. F. Viruses Face a Double Defense by Plant Small RNAs. *Science* **313**, 54–55 (2006).
137. Kalantidis, K., Schumacher, H. T., Alexiadis, T. & Helm, J. M. RNA silencing movement in plants. *Biol. Cell* **100**, 13–26 (2008).
138. Bekele, D., Tesfaye, K. & Fikre, A. Applications of Virus Induced Gene Silencing (VIGS) in Plant Functional Genomics Studies. *J. Plant Biochem. Physiol.* **07**, (2019).
139. Tao, X. & Zhou, X. A modified viral satellite DNA that suppresses gene expression in plants. *Plant J.* **38**, 850–860 (2004).
140. Gosselé, V., Faché, I., Meulewaeter, F., Cornelissen, M. & Metzlaff, M. SVISS – a novel transient gene silencing system for gene function discovery and validation in tobacco plants. *Plant J.* **32**, 859–866 (2002).
141. Liu, Y., Schiff, M. & Dinesh-Kumar, S. P. Virus-induced gene silencing in tomato. *Plant J.* **31**, 777–786 (2002).

142. Liu, Y., Schiff, M., Marathe, R. & Dinesh-Kumar, S. P. Tobacco Rar1, EDS1 and NPR1/NIM1 like genes are required for N-mediated resistance to tobacco mosaic virus. *Plant J.* **30**, 415–429 (2002).
143. Chung, E. *et al.* A Method of High Frequency Virus-induced Gene Silencing in Chili Pepper (Capsicum annuum L. cv. Bukang). *Mol. Cells* **17**, 377–380.
144. Brigneti, G. *et al.* Virus-induced gene silencing in Solanum species. *Plant J.* **39**, 264–272 (2004).
145. Kumagai, M. H. *et al.* Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. *Proc. Natl. Acad. Sci.* **92**, 1679–1683 (1995).
146. Pal, S. *et al.* RNAi of Sterol Methyl Transferase1 Reveals its Direct Role in Diverting Intermediates Towards Withanolide/ Phytosterol Biosynthesis in Withania somnifera. *Plant Cell Physiol.* doi:10.1093/pcp/pcy237.
147. Singh, A. K. *et al.* Virus-induced gene silencing of Withania somnifera squalene synthase negatively regulates sterol and defence-related genes resulting in reduced withanolides and biotic stress tolerance. *Plant Biotechnol. J.* **13**, 1287–1299 (2015).
148. Zhang, J.-S., Zhao, J., Zhang, S. & He, C. Efficient Gene Silencing Mediated by Tobacco Rattle Virus in an Emerging Model Plant Physalis. *PLOS ONE* **9**, e85534 (2014).
149. Osorio-Guarín, J. A., García-Arias, F. L. & Yockteng, R. Virus-induced gene silencing (VIGS) in cape gooseberry (Physalis peruviana L., Solanaceae). *Univ. Sci.* **24**, 111–133 (2019).
150. Alfonso, D., Bernardinelli, G. & Kapetanidis, I. Withanolides from Iochroma coccineum. *Phytochemistry* **34**, 517–521 (1993).
151. Silva, G. L., Pacciaroni, A., Oberti, J. C., Veleiro, A. S. & Burton, G. A pregnane structurally related to withanolides from Physalis viscosa. *Phytochemistry* **34**, 871–873 (1993).
152. Matsuura, T., Kawai, M., Nakashima, R. & Butsugan, Y. Structures of physalin A and physalin B, 13,14-seco-16,24-cyclo-steroids from Physalis alkekengi var. Francheti. *J. Chem. Soc. C Org.* **0**, 664–670 (1970).
153. Misra, L. *et al.* Withanolides from Withania somnifera roots. *Phytochemistry* **69**, 1000–1004 (2008).
154. Chaurasiya, N. D., Sangwan, N. S., Sabir, F., Misra, L. & Sangwan, R. S. Withanolide biosynthesis recruits both mevalonate and DOXP pathways of isoprenogenesis in Ashwagandha Withania somnifera L. (Dunal). *Plant Cell Rep.* **31**, 1889–1897 (2012).
155. Fukushima, A. *et al.* Comparative Characterization of the Leaf Tissue of Physalis alkekengi and Physalis peruviana Using RNA-seq and Metabolite Profiling. *Front. Plant Sci.* **7**, (2016).
156. Gupta, P. *et al.* De Novo Assembly, Functional Annotation and Comparative Analysis of Withania somnifera Leaf and Root Transcriptomes to Identify Putative Genes Involved in the Withanolides Biosynthesis. *PLOS ONE* **8**, e62714 (2013).
157. Senthil, K. *et al.* Transcriptome analysis reveals in vitro cultured Withania somnifera leaf and root tissues as a promising source for targeted withanolide biosynthesis. *BMC Genomics* **16**, 14 (2015).
158. Yarden, A. & Lavie, D. 567. Constituents of Withania somnifera. Part I. The functional groups of withaferin. *J. Chem. Soc. Resumed* 2925 (1962) doi:10.1039/jr9620002925.
159. Laviee, D. & Shvo, Y. 1371. Constituents of Withania somnifera Dun. Part 1V.l The Xtructure of Withaferin A. 15.
160. Lavie, D., Greenfield, S. & Glotter, E. Constituents of Withania somnifera Dun. Part VI. The stereochemistry of withaferin A. *J. Chem. Soc. C Org.* **0**, 1753–1756 (1966).
161. Subramanian, S. S., Sethi, P. D., Glotter, E., Kirson, I. & Lavie, D. 5,20α(R)-dihydroxy-6α,7α-epoxy-1-oxo-(5α) witha-2,24-dienolide, a new steroidal lactone from withania coagulans. *Phytochemistry* **10**, 685–688 (1971).

162. Lavie, D., Kirson, I. & Glotter, E. Constituents of Withania Somnifera Dun. Part X The Structure of Withanolide D. *Isr. J. Chem.* **6**, 671–678 (1968).
163. Glotter, E., Abraham, A., Günzberg, G. & Kirson, I. Naturally occurring steroidal lactones with a 17α-oriented side chain. Structure of withanolide E and related compounds. *J. Chem. Soc. Perkin 1* **0**, 341–346 (1977).
164. Vitali, G., Conte, L. & Nicoletti, M. Withanolide Composition and in vitro Culture of Italian Withania somnifera. *Planta Med.* **62**, 287–288 (1996).
165. Dhalla, N. S., Sastry, M. S. & Malhotra, C. L. Chemical Studies of the Leaves of Withania somnifera. *J. Pharm. Sci.* **50**, 876–877 (1961).
166. Sakurai, K., Ishii, H., Kobayashi, S. & Iwao, T. Isolation of 4β-Hydroxywithanolide E, a New Withanolide from Physalis peruviana L. *Chem. Pharm. Bull. (Tokyo)* **24**, 1403–1405 (1976).
167. Kirson, I., Abraham, A., Sethi, P. D., Subramanian, S. S. & Glotter, E. 4β-Hydroxywithanolide E, a new natural steroid with a 17α-oriented side-chain. *Phytochemistry* **15**, 340–342 (1976).
168. Fang, S.-T., Li, B. & Liu, J.-K. Two New Withanolides from Physalis peruviana. *Helv. Chim. Acta* **92**, 1304–1308 (2009).
169. MAHENDRA, S. *et al.* STRUCTURES OF WITHAPERUVIN B AND C, WITHANOLIDES OF PHYSALIS PERUVIANA ROOTS. *Struct. WITHAPERUVIN B C WITHANOLIDES PHYSALIS PERUVIANA ROOTS* (1982).
170. Lan, Y.-H. *et al.* New cytotoxic withanolides from Physalis peruviana. *Food Chem.* **116**, 462–469 (2009).
171. Ali, A., Sahai, M., Ray, A. B. & Slatkin, D. J. Physalolactone C, a New Withanolide from Physalis peruviana. *J. Nat. Prod.* **47**, 648–651 (1984).
172. Stein, A., Compera, D., Karge, B., Brönstrup, M. & Franke, J. Isolation and characterisation of irinans, androstane-type withanolides from Physalis peruviana L. *Beilstein J. Org. Chem.* **15**, 2003–2012 (2019).
173. Kumar, S., Singh, R., Gajbhiye, N. & Dhanani, T. Extraction Optimization for Phenolic- and Withanolide-Rich Fractions from Withania somnifera Roots: Identification and Quantification of Withaferin A, 12-Deoxywithastromonolide, and Withanolide A in Plant Materials and Marketed Formulations Using a Reversed-Phase HPLC-Photodiode Array Detection Method. *J. AOAC Int.* (2018) doi:10.5740/jaoacint.18-0081.
174. Ozawa, M. *et al.* Contribution of Cage-Shaped Structure of Physalins to Their Mode of Action in Inhibition of NF-κB Activation. *ACS Med. Chem. Lett.* **4**, 730–735 (2013).
175. Samadi, A. K. *et al.* Withaferin A, a Cytotoxic Steroid from Vassobia breviflora, Induces Apoptosis in Human Head and Neck Squamous Cell Carcinoma. *J. Nat. Prod.* **73**, 1476–1481 (2010).
176. Yousuf, S. K. *et al.* Ring A structural modified derivatives of withaferin A and the evaluation of their cytotoxic potential. *Steroids* **76**, 1213–1222 (2011).
177. Gamoh, K., Hirayama, M. & Ikekawa, N. Stereocontrolled synthesis of withanolide D and related compounds. *J. Chem. Soc. Perkin 1* 449–454 (1984) doi:10.1039/P19840000449.
178. Maldonado, E., Alvarado, V. E., Torres, F. R., Martínez, M. & Pérez-Castorena, A. L. Androstane and Withanolides from Physalis cinerascens. *Planta Med.* **71**, 548–553 (2005).
179. Kicman, A. T. Pharmacology of anabolic steroids. *Br. J. Pharmacol.* **154**, 502–521 (2008).
180. Siddiqui, B. S., Usmani, S. B., Begum, S. & Siddiqui, S. Steroidal alkaloids and an androstane derivative from the bark of Holarrhena pubescens. *Phytochemistry* **33**, 925–928 (1993).
181. Sanogo, R., Germano, M. P., de Tommasi, N., Pizza, C. & Aquino, R. Vernoniosides and an androstane glycoside from Vernonia kotschyana. *Phytochemistry* **47**, 73–78 (1998).

182. Pupo, M. T., Vieira, P. C., Fernandes, J. B., das G.F. da Silva, M. F. & Fo, E. R. Androstane and pregnane 2β,19-hemiketal steroids from Trichilia claussenii. *Phytochemistry* **45**, 1495–1500 (1997).
183. Burstein, S., Middleditch, B. S. & Gut, M. Mass spectrometric study of the enzymatic conversion of cholesterol to (22R)-22-hydroxycholesterol, (20R,22R)-20,22-dihydroxycholesterol, and pregnenolone, and of (22R)-22-hydroxycholesterol to the lgycol and pregnenolone in bovine adrenocortical preparations. Mode of oxygen incorporation. *J. Biol. Chem.* **250**, 9028–9037 (1975).
184. Byon, C.-Y. & Gut, M. Steric considerations regarding the biodegradation of cholesterol to pregnenolone.-exclusion of (22S)-22-hydroxycholesterol and 22-ketocholesterol as intermediates. *Biochem. Biophys. Res. Commun.* **94**, 549–552 (1980).
185. Yoshimoto, F. K., Gonzalez, E., Auchus, R. J. & Guengerich, F. P. Mechanism of 17α,20-Lyase and New Hydroxylation Reactions of Human Cytochrome P450 17A1. *J. Biol. Chem.* **291**, 17143–17164 (2016).
186. Strushkevich, N. *et al.* Structural basis for pregnenolone biosynthesis by the mitochondrial monooxygenase system. *Proc. Natl. Acad. Sci.* **108**, 10139–10143 (2011).
187. Jayaprakasam, B., Zhang, Y., Seeram, N. P. & Nair, M. G. Growth inhibition of human tumor cell lines by withanolides from Withania somnifera leaves. *Life Sci.* **74**, 125–132 (2003).
188. Kaileh, M. *et al.* Withaferin A Strongly Elicits IκB Kinase β Hyperphosphorylation Concomitant with Potent Inhibition of Its Kinase Activity. *J. Biol. Chem.* **282**, 4253–4264 (2007).
189. Yen, C.-Y. *et al.* 4β-Hydroxywithanolide E from Physalis peruviana (golden berry) inhibits growth of human lung cancer cells through DNA damage, apoptosis and G2/M arrest. *BMC Cancer* **10**, 46 (2010).
190. Xu, Y.-M. *et al.* Withanolides from Aeroponically Grown Physalis peruviana and Their Selective Cytotoxicity to Prostate Cancer and Renal Carcinoma Cells. *J. Nat. Prod.* **80**, 1981–1991 (2017).
191. Zhang, H. *et al.* Cytotoxic Withanolide Constituents of Physalis longifolia. *J. Nat. Prod.* **74**, 2532–2544 (2011).
192. Keinan, E., Sahai, M. & Kirson, I. Reductive elimination of vicinal oxygen functions with palladium(0). Applications in the withanolide series. *J. Org. Chem.* **48**, 2550–2555 (1983).
193. Gottlied, H. E. & Kirson, I. 13C NMR spectroscopy of the withanolides and other highly oxygenated C28 steroids. *Org. Magn. Reson.* **16**, 20–25 (1981).
194. Ness, F. *et al.* Sterol Uptake in Saccharomyces cerevisiae Heme Auxotrophic Mutants Is Affected by Ergosterol and Oleate but Not by Palmitoleate or by Sterol Esterification. *J. Bacteriol.* **180**, 1913–1919 (1998).
195. Choudhary, V. & Schneiter, R. Monitoring Sterol Uptake, Acetylation, and Export in Yeast. in *Lipidomics: Volume 2: Methods and Protocols* (ed. Armstrong, D.) 221–232 (Humana Press, 2010). doi:10.1007/978-1-60761-325-1_12.
196. Sawai, S. *et al.* Sterol Side Chain Reductase 2 Is a Key Enzyme in the Biosynthesis of Cholesterol, the Common Precursor of Toxic Steroidal Glycoalkaloids in Potato. *Plant Cell* **26**, 3763–3774 (2014).
197. Miziorko, H. M. Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch. Biochem. Biophys.* **505**, 131–143 (2011).
198. Klug, L. & Daum, G. Yeast lipid metabolism at a glance. *FEMS Yeast Res.* **14**, 369–388 (2014).
199. Hu, Z. *et al.* Recent Advances in Ergosterol Biosynthesis and Regulation Mechanisms in Saccharomyces cerevisiae. *Indian J. Microbiol.* **57**, 270–277 (2017).

200. Kodedová, M. & Sychrová, H. Changes in the Sterol Composition of the Plasma Membrane Affect Membrane Potential, Salt Tolerance and the Activity of Multidrug Resistance Pumps in Saccharomyces cerevisiae. *PLOS ONE* **10**, e0139306 (2015).
201. Kelly, D. E., Lamb, D. C. & Kelly, S. L. Genome-Wide Generation of Yeast Gene Deletion Strains. *Comp. Funct. Genomics* **2**, 236–242 (2001).
202. Goldstein, A. L. & McCusker, J. H. Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast* **15**, 1541–1553 (1999).
203. Da Silva, N. A. & Srikrishnan, S. Introduction and expression of genes for metabolic engineering applications in Saccharomyces cerevisiae. *FEMS Yeast Res.* **12**, 197–214 (2012).
204. Fang, F. *et al.* A vector set for systematic metabolic engineering in Saccharomyces cerevisiae. *Yeast* **28**, 123–136 (2011).
205. Ma, B.-X., Ke, X., Tang, X.-L., Zheng, R.-C. & Zheng, Y.-G. Rate-limiting steps in the Saccharomyces cerevisiae ergosterol pathway: towards improved ergosta-5,7-dien-3β-ol accumulation by metabolic engineering. *World J. Microbiol. Biotechnol.* **34**, 55 (2018).
206. Reisdorph, N. A., Walmsley, S. & Reisdorph, R. A Perspective and Framework for Developing Sample Type Specific Databases for LC/MS-Based Clinical Metabolomics. *Metabolites* **10**, 8 (2020).
207. Ro, D.-K., Ehlting, J. & Douglas, C. J. Cloning, Functional Expression, and Subcellular Localization of Multiple NADPH-Cytochrome P450 Reductases from Hybrid Poplar. *Plant Physiol.* **130**, 1837–1851 (2002).
208. Osumi, T., Taketani, S., Katsuki, H., Kuhara, T. & Matsumoto, I. Ergosterol Biosynthesis in Yeast. *J. Biochem. (Tokyo)* **83**, 681–691 (1978).
209. hummel@mpimp-golm.mpg.de. GMD - the Golm Metabolome Database. http://gmd.mpimp-golm.mpg.de/.
210. Müller, C., Staudacher, V., Krauss, J., Giera, M. & Bracher, F. A convenient cellular assay for the identification of the molecular target of ergosterol biosynthesis inhibitors and quantification of their effects on total ergosterol biosynthesis. *Steroids* **78**, 483–493 (2013).
211. Loeffler, R. S. T. & Hayes, A. L. Effects of Sterol Biosynthesis Inhibitor Fungicides on Growth and Sterol Composition of Ustilago maydis, Botrytis cinerea and Pyrenophora teres. *Pestic. Sci.* **36**, 7–17 (1992).
212. Le Fur, Y., Maume, G., Feuillat, M. & Maume, B. F. Characterization by gas chromatography/mass spectrometry of sterols in saccharomyces cerevisiae during autolysis. *J. Agric. Food Chem.* **47**, 2860–2864 (1999).
213. Dos Santos Dias, A. C. *et al.* Steroids from Marine-Derived Fungi: Evaluation of Antiproliferative and Antimicrobial Activities of Eburicol. *Mar. Drugs* **17**, 372 (2019).
214. Renard, D., Perruchon, J., Giera, M., Müller, J. & Bracher, F. Side chain azasteroids and thiasteroids as sterol methyltransferase inhibitors in ergosterol biosynthesis. *Bioorg. Med. Chem.* **17**, 8123–8137 (2009).
215. Liu, X.-F. *et al.* Expression of GAI gene and disruption of PEP4 gene in an industrial brewer's yeast strain. *Lett. Appl. Microbiol.* **49**, 117–123 (2009).
216. Choon, R. L. T., Sariah, M. & Mariam, M. N. S. Ergosterol from the soilborne fungus Ganoderma boninense. *J. Basic Microbiol.* **52**, 608–612 (2012).
217. Shirane, N. *et al.* Effect on ergosterol biosynthesis of a fungicide, SSF-109, in Botrytis cinerea. *Phytochemistry* **29**, 2513–2520 (1990).
218. Barbier, M. Isolation of 24-methylenecholesterol-derived oxidation products from queen honeybee ovaries (Apis mellifica L.). *J. Chem. Ecol.* **13**, 1681–1687 (1987).
219. Barrero, A. F., Oltra, J. E., Poyatos, J. A., Jiménez, D. & Oliver, E. Phycomysterols and Other Sterols from the Fungus Phycomyces blakesleeanus. *J. Nat. Prod.* **61**, 1491–1496 (1998).

220. Moebius, F. F., Fitzky, B. U., Lee, J. N., Paik, Y.-K. & Glossmann, H. Molecular cloning and expression of the human Δ7-sterol reductase. *Proc. Natl. Acad. Sci.* **95**, 1899–1902 (1998).
221. Sonawane, P. D. *et al.* Plant cholesterol biosynthetic pathway overlaps with phytosterol metabolism. *Nat. Plants* **3**, 1–13 (2016).
222. Pérez-Martín, J. Chromatin and transcription in Saccharomyces cerevisiae. *FEMS Microbiol. Rev.* **23**, 503–523 (1999).
223. Bach, M. L., Lacroute, F. & Botstein, D. Evidence for transcriptional regulation of orotidine-5'-phosphate decarboxylase in yeast by hybridization of mRNA to the yeast structural gene cloned in Escherichia coli.
224. Flagfeldt, D. B., Siewers, V., Huang, L. & Nielsen, J. Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. *Yeast* **26**, 545–551 (2009).
225. Yang, J., Li, C. & Zhang, Y. Engineering of Saccharomyces cerevisiae for 24-Methylene-Cholesterol Production. *Biomolecules* **11**, 1710 (2021).
226. Urban, P., Mignotte, C., Kazmaier, M., Delorme, F. & Pompon, D. Cloning, Yeast Expression, and Characterization of the Coupling of Two Distantly Related Arabidopsis thalianaNADPH-Cytochrome P450 Reductases with P450 CYP73A5 *. *J. Biol. Chem.* **272**, 19176–19186 (1997).
227. Sun, Z. *et al.* Identification of Novel Knockout Targets for Improving Terpenoids Biosynthesis in Saccharomyces cerevisiae. *PLOS ONE* **9**, e112615 (2014).
228. Entian, K.-D. & Kötter, P. 25 Yeast Genetic Strain and Plasmid Collections. in *Methods in Microbiology* (eds. Stansfield, I. & Stark, M. J.) vol. 36 629–666 (Academic Press, 2007).
229. Umebayashi, K. & Nakano, A. Ergosterol is required for targeting of tryptophan permease to the yeast plasma membrane. *J. Cell Biol.* **161**, 1117–1131 (2003).
230. Liu, G., Chen, Y., Færgeman, N. J. & Nielsen, J. Elimination of the last reactions in ergosterol biosynthesis alters the resistance of Saccharomyces cerevisiae to multiple stresses. *FEMS Yeast Res.* **17**, fox063 (2017).
231. Zhang, B., He, X., Tie, C. & Liu, Y. Construction of high ergosterol-producing yeast strains and study on the optimal conditions for culture. *Chin. J. Biotechnol.* **15**, 43–49 (1999).
232. Veen, M., Stahl, U. & Lang, C. Combined overexpression of genes of the ergosterol biosynthetic pathway leads to accumulation of sterols in Saccharomyces cerevisiae. *FEMS Yeast Res.* **4**, 87–95 (2003).
233. Wriessnegger, T. & Pichler, H. Yeast metabolic engineering – Targeting sterol metabolism and terpenoid formation. *Prog. Lipid Res.* **52**, 277–293 (2013).
234. Tsay, Y. H. & Robinson, G. W. Cloning and characterization of ERG8, an essential gene of Saccharomyces cerevisiae that encodes phosphomevalonate kinase. *Mol. Cell. Biol.* **11**, 620–631 (1991).
235. Jennings, S. M., Tsay, Y. H., Fisch, T. M. & Robinson, G. W. Molecular cloning and characterization of the yeast gene for squalene synthetase. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 6038–6042 (1991).
236. Fegueur, M., Richard, L., Charles, A. D. & Karst, F. Isolation and primary structure of the ERG9 gene of Saccharomyces cerevisiae encoding squalene synthetase. *Curr. Genet.* **20**, 365–372 (1991).
237. Hiser, L., Basson, M. E. & Rine, J. ERG10 from Saccharomyces cerevisiae encodes acetoacetyl-CoA thiolase. *J. Biol. Chem.* **269**, 31383–31389 (1994).
238. Oulmouden, A. & Karst, F. Nucleotide sequence of the ERG12 gene of Saccharomyces cerevisiae encoding mevalonate kinase. *Curr. Genet.* **19**, 9–14 (1991).

239. Parks, L. W., Smith, S. J. & Crowley, J. H. Biochemical and physiological effects of sterol alterations in yeast—A review. *Lipids* **30**, 227–230 (1995).
240. Chambon, C. *et al.* Sterol pathway in yeast. Identification and properties of mutant strains defective in mevalonate diphosphate decarboxylase and farnesyl diphosphate synthetase. *Lipids* **26**, 633–636 (1991).
241. Anderson, M. S., Yarger, J. G., Burck, C. L. & Poulter, C. D. Farnesyl diphosphate synthetase. Molecular cloning, sequence, and expression of an essential gene from Saccharomyces cerevisiae. *J. Biol. Chem.* **264**, 19176–19184 (1989).
242. Mayer, M. P., Hahn, F. M., Stillman, D. J. & Poulter, C. D. Disruption and mapping of IDI1, the gene for isopentenyl diphosphate isomerase in Saccharomyces cerevisiae. *Yeast* **8**, 743–748 (1992).
243. Basson, M. E., Thorsness, M. & Rine, J. Saccharomyces cerevisiae contains two functional genes encoding 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *Proc. Natl. Acad. Sci.* **83**, 5563–5567 (1986).
244. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
245. Schulz, M. H., Zerbino, D. R., Vingron, M. & Birney, E. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* **28**, 1086–1092 (2012).
246. Dhar, N. *et al.* Dynamics of withanolide biosynthesis in relation to temporal expression pattern of metabolic genes in <Emphasis Type="Italic">Withania somnifera</Emphasis> (L.) Dunal: a comparative study in two morpho-chemovariants. *Mol. Biol. Rep.* **40**, 7007–7016 (2013).
247. Naoumkina, M. A. *et al.* Genomic and Coexpression Analyses Predict Multiple Genes Involved in Triterpene Saponin Biosynthesis in Medicago truncatula. *Plant Cell* **22**, 850–866 (2010).
248. Fu, F.-F. & Xue, H.-W. Coexpression Analysis Identifies Rice Starch Regulator1, a Rice AP2/EREBP Family Transcription Factor, as a Novel Rice Starch Biosynthesis Regulator. *Plant Physiol.* **154**, 927–938 (2010).
249. Guerin, C. *et al.* Gene coexpression network analysis of oil biosynthesis in an interspecific backcross of oil palm. *Plant J.* **87**, 423–441 (2016).
250. Aoki, K., Ogata, Y. & Shibata, D. Approaches for Extracting Practical Information from Gene Co-expression Networks in Plant Biology. *Plant Cell Physiol.* **48**, 381–390 (2007).
251. Yazaki, K. ABC transporters involved in the transport of plant secondary metabolites. *FEBS Lett.* **580**, 1183–1191 (2006).
252. Stuart, J. M., Segal, E., Koller, D. & Kim, S. K. A Gene-Coexpression Network for Global Discovery of Conserved Genetic Modules. *Science* **302**, 249–255 (2003).
253. Kohonen, T. The self-organizing map. *Proc. IEEE* **78**, 1464–1480 (1990).
254. Han, J.-Y., Kim, M.-J., Ban, Y.-W., Hwang, H.-S. & Choi, Y.-E. The Involvement of β-Amyrin 28-Oxidase (CYP716A52v2) in Oleanane-Type Ginsenoside Biosynthesis in Panax ginseng. *Plant Cell Physiol.* **54**, 2034–2046 (2013).
255. Katsumata, T. *et al.* Arabidopsis CYP85A2 catalyzes lactonization reactions in the biosynthesis of 2-deoxy-7-oxalactone brassinosteroids. *Biosci. Biotechnol. Biochem.* **72**, 2110–2117 (2008).
256. Tamura, K. *et al.* CYP716A179 functions as a triterpene C-28 oxidase in tissue-cultured stolons of Glycyrrhiza uralensis. *Plant Cell Rep.* **36**, 437–445 (2017).
257. TriForC database. http://bioinformatics.psb.ugent.be/triforc/#/enzymes/1//.
258. Umemoto, N. *et al.* Two Cytochrome P450 Monooxygenases Catalyze Early Hydroxylation Steps in the Potato Steroid Glycoalkaloid Biosynthetic Pathway. *Plant Physiol.* **171**, 2458–2467 (2016).

259. Tzfadia, O. *et al.* CoExpNetViz: Comparative Co-Expression Networks Construction and Visualization Tool. *Front. Plant Sci.* **6**, (2016).
260. Cho, M.-H. *et al.* Phenylalanine biosynthesis in Arabidopsis thaliana. Identification and characterization of arogenate dehydratases. *J. Biol. Chem.* **282**, 30827–30835 (2007).
261. Akashi, T., Aoki, T. & Ayabe, S. Molecular and Biochemical Characterization of 2-Hydroxyisoflavanone Dehydratase. Involvement of Carboxylesterase-Like Proteins in Leguminous Isoflavone Biosynthesis. *Plant Physiol.* **137**, 882–891 (2005).
262. Singh, S. A. & Christendat, D. Structure of Arabidopsis dehydroquinate dehydratase-shikimate dehydrogenase and implications for metabolic channeling in the shikimate pathway. *Biochemistry* **45**, 7787–7796 (2006).
263. Cole, F. E., Kalyanpur, M. G. & Stevens, C. M. Absolute configuration of α-isopropylmalate and the mechanism of its conversion to β-isopropylmalate in the biosynthesis of leucine. *Biochemistry* **12**, 3346–3350 (1973).
264. Sharma, R. K. & Mazumder, R. Purification, Properties, and Feedback Control of l-Threonine Dehydratase from Spinach. *J. Biol. Chem.* **245**, 3008–3014 (1970).
265. Kamsteeg, J., Brederode, J. V. & Nigtevecht, G. V. The formation of UDP-L-rhamnose from UDP-D-glucose by an enzyme preparation of red campion (Silene Dioica (L) clairv) leaves. *FEBS Lett.* **91**, 281–284 (1978).
266. Senthil-Kumar, M. & Mysore, K. S. New dimensions for VIGS in plant functional genomics. *Trends Plant Sci.* **16**, 656–665 (2011).
267. Zhang, S., Zhang, J.-S., Zhao, J. & He, C. Distinct subfunctionalization and neofunctionalization of the B-class MADS-box genes in Physalis floridana. *Planta* **241**, 387–402 (2015).
268. Gao, H., Li, J., Wang, L., Zhang, J. & He, C. Transcriptomic variation of the flower–fruit transition in Physalis and Solanum. *Planta* **252**, 28 (2020).
269. Fraser, P. D. *et al.* Expression in Escherichia coli, purification, and reactivation of the recombinant Erwinia uredovora phytoene desaturase. *J. Biol. Chem.* **267**, 19891–19895 (1992).
270. Demmig-Adams, B. & Adams, W. W. The role of xanthophyll cycle carotenoids in the protection of photosynthesis. *Trends Plant Sci.* **1**, 21–26 (1996).
271. Bomzan, D. P., Shilpashree, H. B., Anjali, P., Kumar, S. R. & Nagegowda, D. A. Virus-Induced Gene Silencing for Functional Genomics in Withania somnifera, an Important Indian Medicinal Plant. in *Virus-Induced Gene Silencing in Plants: Methods and Protocols* (eds. Courdavault, V. & Besseau, S.) 139–154 (Springer US, 2020). doi:10.1007/978-1-0716-0751-0_11.
272. Xu, H. *et al.* Tobacco rattle virus-induced PHYTOENE DESATURASE (PDS) and Mg-chelatase H subunit (ChlH) gene silencing in Solanum pseudocapsicum L. *PeerJ* **6**, e4424 (2018).
273. Zhong, X. *et al.* Virus-induced gene silencing for comparative functional studies in Gladiolus hybridus. *Plant Cell Rep.* **33**, 301–312 (2014).
274. Lu, R., Martin-Hernandez, A. M., Peart, J. R., Malcuit, I. & Baulcombe, D. C. Virus-induced gene silencing in plants. *Methods* **30**, 296–303 (2003).
275. Lück, S. *et al.* siRNA-Finder (si-Fi) Software for RNAi-Target Design and Off-Target Prediction. *Front. Plant Sci.* **10**, (2019).
276. si-Fi. *Snowformatics* http://www.snowformatics.com/si-fi.html.
277. Pang, J. *et al.* Development of Agrobacterium-Mediated Virus-Induced Gene Silencing and Performance Evaluation of Four Marker Genes in Gossypium barbadense. *PLOS ONE* **8**, e73211 (2013).
278. Broderick, S. R. & Jones, M. L. An Optimized Protocol to Increase Virus-Induced Gene Silencing Efficiency and Minimize Viral Symptoms in Petunia. *Plant Mol. Biol. Report.* **32**, 219–233 (2014).

279. Sung, Y.-C., Lin, C.-P. & Chen, J.-C. Optimization of virus-induced gene silencing in Catharanthus roseus. *Plant Pathol.* **63**, 1159–1167 (2014).
280. Xu, H. *et al.* Virus-induced Phytoene Desaturase (PDS) Gene Silencing Using Tobacco Rattle Virus in Lilium × formolongi. *Hortic. Plant J.* **5**, 31–38 (2019).
281. Li, Y. *et al.* Establishment of virus-induced gene silencing system and functional analysis of ScbHLH17 in Senecio cruentus. *Plant Physiol. Biochem.* **147**, 272–279 (2020).
282. Kim, J., Park, M., Jeong, E. S., Lee, J. M. & Choi, D. Harnessing anthocyanin-rich fruit: a visible reporter for tracing virus-induced gene silencing in pepper fruit. *Plant Methods* **13**, (2017).
283. Norkunas, K., Harding, R., Dale, J. & Dugdale, B. Improving agroinfiltration-based transient gene expression in Nicotiana benthamiana. *Plant Methods* **14**, 71 (2018).
284. Cheng, M. *et al.* Genetic Transformation of Wheat Mediated by Agrobacterium tumefaciens. *Plant Physiol.* **115**, 971–980 (1997).
285. Curtis, I. S. & Nam, H. G. Transgenic radish (Raphanus sativus L. longipinnatus Bailey) by floral-dip method – plant development and surfactant are important in optimizing transformation efficiency. 9.
286. Zhao, H., Tan, Z., Wen, X. & Wang, Y. An Improved Syringe Agroinfiltration Protocol to Enhance Transformation Efficiency by Combinative Use of 5-Azacytidine, Ascorbate Acid and Tween-20. *Plants* **6**, 9 (2017).
287. Chetty, V. J. *et al.* Evaluation of four Agrobacterium tumefaciens strains for the genetic transformation of tomato (Solanum lycopersicum L.) cultivar Micro-Tom. *Plant Cell Rep.* **32**, 239–247 (2013).
288. Lazo, G. R., Stein, P. A. & Ludwig, R. A. A DNA Transformation–Competent Arabidopsis Genomic Library in Agrobacterium. *Bio/Technology* **9**, 963–967 (1991).
289. Ooms, G. *et al.* Octopine Ti-plasmid deletion mutants of Agrobacterium tumefaciens with emphasis on the right side of the T-region. *Plasmid* **7**, 15–29 (1982).
290. Hood, E. E., Gelvin, S. B., Melchers, L. S. & Hoekema, A. New Agrobacterium helper plasmids for gene transfer to plants. *Transgenic Res.* **2**, 208–218 (1993).
291. Wege, S., Scholz, A., Gleissberg, S. & Becker, A. Highly Efficient Virus-induced Gene Silencing (VIGS) in California Poppy (Eschscholzia californica): An Evaluation of VIGS as a Strategy to Obtain Functional Data from Non-model Plants. *Ann. Bot.* **100**, 641–649 (2007).
292. Heid, C. A., Stevens, J., Livak, K. J. & Williams, P. M. Real Time Quantitative PCR. 10.
293. Klahre, U. *et al.* The Arabidopsis DIMINUTO/DWARF1 gene encodes a protein involved in steroid synthesis. *Plant Cell* **10**, 1677–1690 (1998).
294. Choe, S. *et al.* The Arabidopsis dwarf1 Mutant Is Defective in the Conversion of 24-Methylenecholesterol to Campesterol in Brassinosteroid Biosynthesis1. *Plant Physiol.* **119**, 897–908 (1999).
295. Sangwan, R. S. *et al.* Withanolide A is inherently de novo biosynthesized in roots of the medicinal plant Ashwagandha (Withania somnifera). *Physiol. Plant.* **133**, 278–287 (2008).
296. Palmer, L. & O'Connor, S. E. Virus-Induced Gene Silencing in Nepeta. in *Virus-Induced Gene Silencing in Plants: Methods and Protocols* (eds. Courdavault, V. & Besseau, S.) 111–121 (Springer US, 2020). doi:10.1007/978-1-0716-0751-0_9.
297. Yamamoto, K. *et al.* Improved virus-induced gene silencing allows discovery of a serpentine synthase gene in Catharanthus roseus. *Plant Physiol.* **187**, 846–857 (2021).
298. Ohnishi, T. *et al.* C-23 Hydroxylation by Arabidopsis CYP90C1 and CYP90D1 Reveals a Novel Shortcut in Brassinosteroid Biosynthesis. *Plant Cell* **18**, 3275–3288 (2006).
299. Seki, H. *et al.* Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. *Plant Cell* **23**, 4112–4123 (2011).

300. Carelli, M. *et al.* Medicago truncatula CYP716A12 Is a Multifunctional Oxidase Involved in the Biosynthesis of Hemolytic Saponins. *Plant Cell* **23**, 3070–3081 (2011).
301. Fukushima, E. O. *et al.* CYP716A Subfamily Members are Multifunctional Oxidases in Triterpenoid Biosynthesis. *Plant Cell Physiol.* **52**, 2050–2061 (2011).
302. Kandel, S. *et al.* Characterization of a methyl jasmonate and wounding-responsive cytochrome P450 of Arabidopsis thaliana catalyzing dicarboxylic fatty acid formation in vitro. *FEBS J.* **274**, 5116–5127 (2007).
303. Heitz, T. *et al.* Cytochromes P450 CYP94C1 and CYP94B3 Catalyze Two Successive Oxidation Steps of Plant Hormone Jasmonoyl-isoleucine for Catabolic Turnover. *J. Biol. Chem.* **287**, 6296–6306 (2012).
304. Widemann, E. *et al.* Sequential oxidation of Jasmonoyl-Phenylalanine and Jasmonoyl-Isoleucine by multiple cytochrome P450 of the CYP94 family through newly identified aldehyde intermediates. *Phytochemistry* **117**, 388–399 (2015).
305. Shimada, Y. *et al.* Expression of chimeric P450 genes encoding flavonoid-3′,5′-hydroxylase in transgenic tobacco and petunia plants 1. *FEBS Lett.* **461**, 241–245 (1999).
306. Vazquez-Albacete, D. *et al.* An expression tag toolbox for microbial production of membrane bound plant cytochromes P450. *Biotechnol. Bioeng.* **114**, 751–760 (2017).
307. Sevrioukova, I. F., Li, H., Zhang, H., Peterson, J. A. & Poulos, T. L. Structure of a cytochrome P450–redox partner electron-transfer complex. *Proc. Natl. Acad. Sci.* **96**, 1863–1868 (1999).
308. Barnes, H. J., Arlotto, M. P. & Waterman, M. R. Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in Escherichia coli. *Proc. Natl. Acad. Sci.* **88**, 5597–5601 (1991).
309. Williams, P. A., Cosme, J., Sridhar, V., Johnson, E. F. & McRee, D. E. Microsomal cytochrome P450 2C5: comparison to microbial P450s and unique features. *J. Inorg. Biochem.* **81**, 183–190 (2000).
310. Leonard, E. & Koffas, M. A. G. Engineering of Artificial Plant Cytochrome P450 Enzymes for Synthesis of Isoflavones by Escherichia coli. *Appl. Environ. Microbiol.* **73**, 7246–7251 (2007).
311. Hausjell, J., Halbwirth, H. & Spadiut, O. Recombinant production of eukaryotic cytochrome P450s in microbial cell factories. *Biosci. Rep.* **38**, BSR20171290 (2018).
312. Rawal, S., Yip, S. S. M. & Coulombe, R. A. Cloning, expression and functional characterization of cytochrome P450 3A37 from turkey liver with high aflatoxin B1 epoxidation activity. *Chem. Res. Toxicol.* **23**, 1322–1329 (2010).
313. Ro, D. K., Mah, N., Ellis, B. E. & Douglas, C. J. Functional Characterization and Subcellular Localization of Poplar (Populus trichocarpa × Populus deltoides) Cinnamate 4-Hydroxylase. *Plant Physiol.* **126**, 317–329 (2001).
314. Augustin, M. M. *et al.* Elucidating steroid alkaloid biosynthesis in Veratrum californicum: production of verazine in Sf9 cells. *Plant J.* **82**, 991–1003 (2015).
315. Christ, B. *et al.* Repeated evolution of cytochrome P450-mediated spiroketal steroid biosynthesis in plants. *Nat. Commun.* **10**, 3206 (2019).
316. Urban, P. *et al.* Characterization of recombinant plant cinnamate 4-hydroxylase produced in yeast. *Eur. J. Biochem.* **222**, 843–850 (1994).
317. Hübner, S. *et al.* Functional expression of cinnamate 4-hydroxylase from Ammi majus L. *Phytochemistry* **64**, 445–452 (2003).
318. Kong, J.-Q., Lu, D. & Wang, Z.-B. Molecular Cloning and Yeast Expression of Cinnamate 4-Hydroxylase from Ornithogalum saundersiae Baker. *Molecules* **19**, 1608–1621 (2014).
319. Herpen, T. W. J. M. van *et al.* Nicotiana benthamiana as a Production Platform for Artemisinin Precursors. *PLOS ONE* **5**, e14222 (2010).
320. Sheludko, Y. V., Gerasymenko, I. M. & Warzecha, H. Transient Expression of Human Cytochrome P450s 2D6 and 3A4 in Nicotiana benthamiana Provides a Possibility for

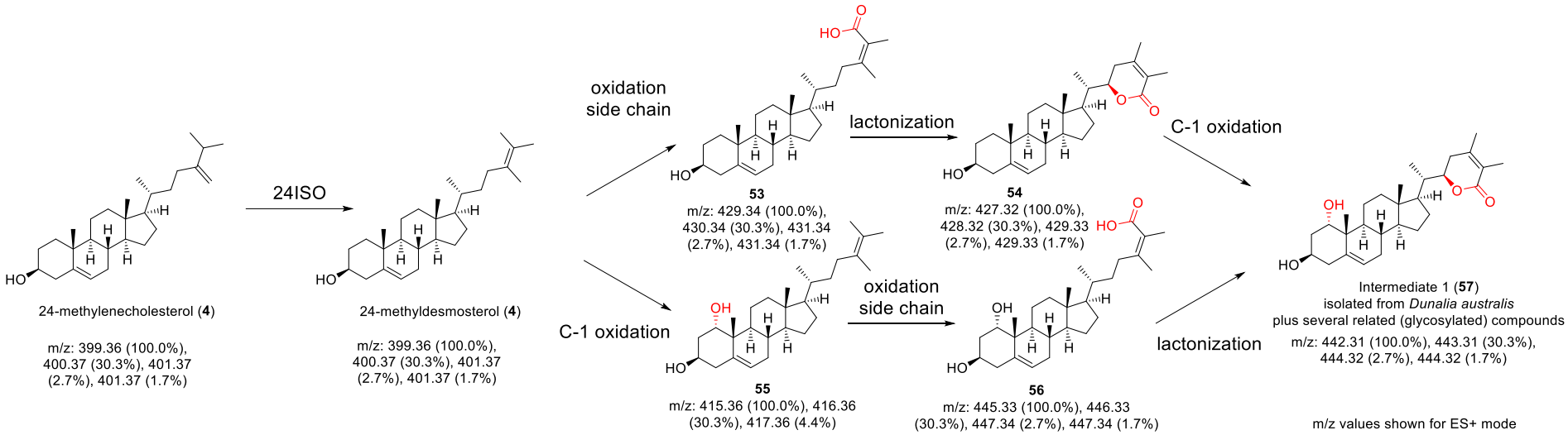
Rapid Substrate Testing and Production of Novel Compounds. *Biotechnol. J.* **13**, 1700696 (2018).
321. Xu, H. *et al.* Pyrethric acid of natural pyrethrin insecticide: complete pathway elucidation and reconstitution in Nicotiana benthamiana. *New Phytol.* **223**, 751–765 (2019).
322. Paik, Y. K., Billheimer, J. T., Magolda, R. L. & Gaylor, J. L. Microsomal enzymes of cholesterol biosynthesis from lanosterol. Solubilization and purification of steroid 8-isomerase. *J. Biol. Chem.* **261**, 6470–6477 (1986).
323. Lin, T. *et al.* A sensitive colorimetric assay for cholesterol based on the peroxidase-like activity of MoS2 nanosheets. *Microchim. Acta* **184**, 1233–1237 (2017).
324. Nes, W. D., Janssen, G. G. & Bergenstrahle, A. Structural requirements for transformation of substrates by the (S)-adenosyl-L-methionine:delta 24(25)-sterol methyl transferase. *J. Biol. Chem.* **266**, 15202–15212 (1991).
325. Bouvier-Navé, P., Husselstein, T. & Benveniste, P. Two families of sterol methyltransferases are involved in the first and the second methylation steps of plant sterol biosynthesis. *Eur. J. Biochem.* **256**, 88–96 (1998).
326. Yehia, H. M., Hassanein, W. A. & Ibraheim, S. M. Purification and characterisation of the extracellular cholesterol oxidase enzyme from Enterococcus hirae. *BMC Microbiol.* **15**, 178 (2015).
327. Li, X., Hylemon, P., Pandak, W. M. & Ren, S. Enzyme activity assay for cholesterol 27-hydroxylase in mitochondria. *J. Lipid Res.* **47**, 1507–1512 (2006).
328. Zidovetzki, R. & Levitan, I. Use of cyclodextrins to manipulate plasma membrane cholesterol content: evidence, misconceptions and control strategies. *Biochim. Biophys. Acta* **1768**, 1311–1324 (2007).
329. López, C. A., Vries, A. H. de & Marrink, S. J. Molecular Mechanism of Cyclodextrin Mediated Cholesterol Extraction. *PLOS Comput. Biol.* **7**, e1002020 (2011).
330. Pitha, J., Anaissie, E. J. & Uekama, K. γ-Cyclodextrin: Testosterone Complex Suitable for Sublingual Administration. *J. Pharm. Sci.* **76**, 788–790 (1987).
331. Irie, T. *et al.* Amorphous Water-Soluble Cyclodextrin Derivatives: 2-Hydroxyethyl, 3-Hydroxypropyl, 2-Hydroxyisobutyl, and Carboxamidomethyl Derivatives of β-Cyclodextrin. *Pharm. Res.* **5**, 713–717 (1988).
332. Somogyi, G., Posta, J., Buris, L. & Varga, M. Cyclodextrin (CD) complexes of cholesterol--their potential use in reducing dietary cholesterol intake. *Pharm.* **61**, 154–156 (2006).
333. Pompon, D., Louerat, B., Bronine, A. & Urban, P. Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol.* **272**, 51–64 (1996).
334. Schenkman, J. B. & Jansson, I. Spectral analyses of cytochromes P450. *Methods Mol. Biol. Clifton NJ* **320**, 11–18 (2006).
335. Kochs, G. & Grisebach, H. Phytoalexin synthesis in soybean: Purification and reconstitution of cytochrome P450 3,9-dihydroxypterocarpan 6a-hydroxylase and separation from cytochrome P450 cinnamate 4-hydroxylase. *Arch. Biochem. Biophys.* **273**, 543–553 (1989).
336. Zou, L., Li, L. & Porter, T. D. 7-Dehydrocholesterol reductase activity is independent of cytochrome P450 reductase. *J. Steroid Biochem. Mol. Biol.* **127**, 435–438 (2011).
337. Stephenson, M. J., Reed, J., Brouwer, B. & Osbourn, A. Transient Expression in Nicotiana Benthamiana Leaves for Triterpene Production at a Preparative Scale. *J. Vis. Exp. JoVE* 58169 (2018) doi:10.3791/58169.
338. Peyret, H., Brown, J. K. M. & Lomonossoff, G. P. Improving plant transient expression through the rational design of synthetic 5′ and 3′ untranslated regions. *Plant Methods* **15**, 108 (2019).
339. Murad, A. M. *et al.* Mass spectrometry characterisation of fatty acids from metabolically engineered soybean seeds. *Anal. Bioanal. Chem.* **406**, 2873–2883 (2014).

340. Kim, G.-T. *et al.* CYP90C1 and CYP90D1 are involved in different steps in the brassinosteroid biosynthesis pathway in Arabidopsis thaliana. *Plant J.* **41**, 710–721 (2005).
341. Escobedo-Martínez, C. *et al.* 1H and 13C NMR characterization of new cycloartane triterpenes from Mangifera indica. *Magn. Reson. Chem.* **50**, 52–57 (2012).
342. Shen, T. *et al.* Cycloartane-Type Triterpenoids from the Resinous Exudates of Commiphora opobalsamum. *J. Nat. Prod.* **71**, 81–86 (2008).
343. BLAST: Basic Local Alignment Search Tool. https://blast.ncbi.nlm.nih.gov/Blast.cgi.
344. Guo, J. *et al.* CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 12108–12113 (2013).
345. Scheler, U. *et al.* Elucidation of the biosynthesis of carnosic acid and its reconstitution in yeast. *Nat. Commun.* **7**, 12942 (2016).
346. Collu, G. *et al.* Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis. *FEBS Lett.* **508**, 215–220 (2001).
347. Sung, P.-H., Huang, F.-C., Do, Y.-Y. & Huang, P.-L. Functional expression of geraniol 10-hydroxylase reveals its dual function in the biosynthesis of terpenoid and phenylpropanoid. *J. Agric. Food Chem.* **59**, 4637–4643 (2011).
348. Nelson, D. R. Cytochrome P450 Homepage - Plants. *Cytochrome P450 Homepage* https://drnelson.uthsc.edu/plants/ (2020).
349. Sato, S. *et al.* The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485**, 635–641 (2012).
350. Pham, G. M. *et al.* Construction of a chromosome-scale long-read reference genome assembly for potato. *GigaScience* **9**, giaa100 (2020).
351. Nakamura, A. & Nakada, M. Allylic Oxidations in Natural Product Synthesis. *Synthesis* **45**, 1421–1451 (2013).
352. BLAST: Basic Local Alignment Search Tool. https://blast.ncbi.nlm.nih.gov/Blast.cgi.
353. Nguyen, H. X. *et al.* Chemical Constituents of Mangifera indica and Their Antiausterity Activity against the PANC-1 Human Pancreatic Cancer Cell Line. *J. Nat. Prod.* **79**, 2053–2059 (2016).
354. Nguyen, H. X., Nguyen, M. T. T., Nguyen, N. T. & Awale, S. Chemical Constituents of Propolis from Vietnamese Trigona minor and Their Antiausterity Activity against the PANC-1 Human Pancreatic Cancer Cell Line. *J. Nat. Prod.* **80**, 2345–2352 (2017).
355. One Shot TOP10 Chemically Competent E. coli - Thermo Fisher Scientific. https://www.thermofisher.com/order/catalog/product/C404003.
356. Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113–130 (1986).
357. One Shot BL21(DE3) Chemically Competent E. coli - Thermo Fisher Scientific. https://www.thermofisher.com/order/catalog/product/C600003.
358. Takara Bio Company. Stellar^{TM} Competent Cells. 2.
359. Nguyen, D. T. *et al.* Biochemical Conservation and Evolution of Germacrene A Oxidase in Asteraceae. *J. Biol. Chem.* **285**, 16588–16598 (2010).
360. Koncz, C. & Schell, J. The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of Agrobacterium binary vector. *Mol. Gen. Genet. MGG* **204**, 383–396 (1986).
361. Ozawa, M. *et al.* Contribution of Cage-Shaped Structure of Physalins to Their Mode of Action in Inhibition of NF-κB Activation. *ACS Med. Chem. Lett.* **4**, 730–735 (2013).
362. Bustin, S. A. *et al.* The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin. Chem.* **55**, 611–622 (2009).
363. Bustin, S. A. Why the need for qPCR publication guidelines?—The case for MIQE. *Methods* **50**, 217–226 (2010).

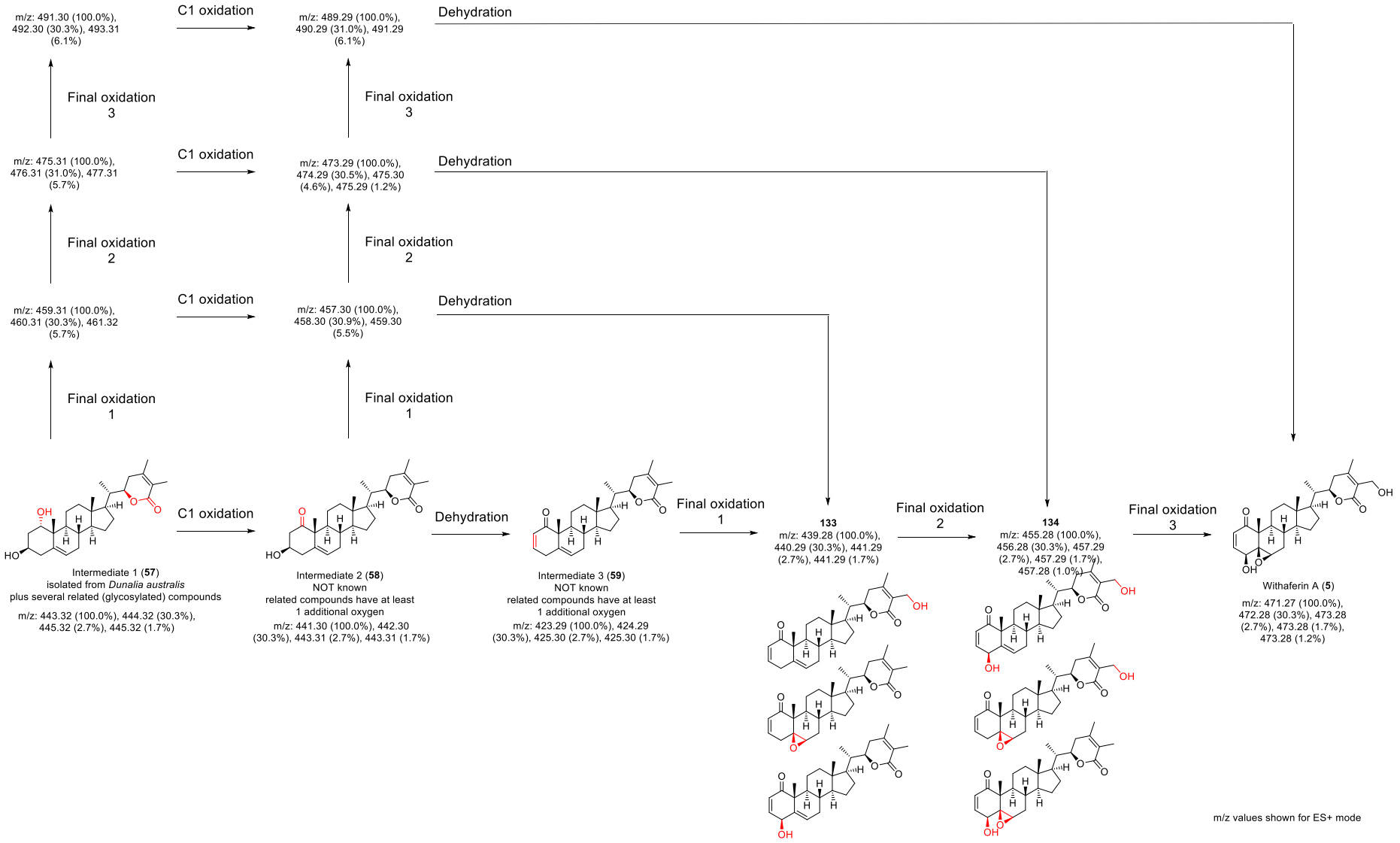
364. Real Time Pcr Handbook Life Technologies. https://www.gene-quantification.de/real-time-pcr-handbook-life-technologies-update-flr.pdf.
365. Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K. & Scheible, W.-R. Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in Arabidopsis | Plant Physiology. http://www.plantphysiol.org/content/139/1/5.short.
366. Primer designing tool. https://www.ncbi.nlm.nih.gov/tools/primer-blast/.
367. Ye, J. *et al.* Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13**, 134 (2012).
368. Meisel, L. *et al.* A rapid and efficient method for purifying high quality total RNA from peaches (Prunus persica) for functional genomics analyses. *Biol. Res.* **38**, 83–88 (2005).

10 Appendix

10.1 Additional figures and schemes



Scheme 10.1 Comprehensive proposal for the biosynthesis of withaferin A (**5**) based on intermediates and similar structures isolated (continued on the next page). [66–68]



Scheme 10.2 Comprehensive proposal for the biosynthesis of withaferin A (**5**) based on intermediates and similar structures isolated. [66–68]

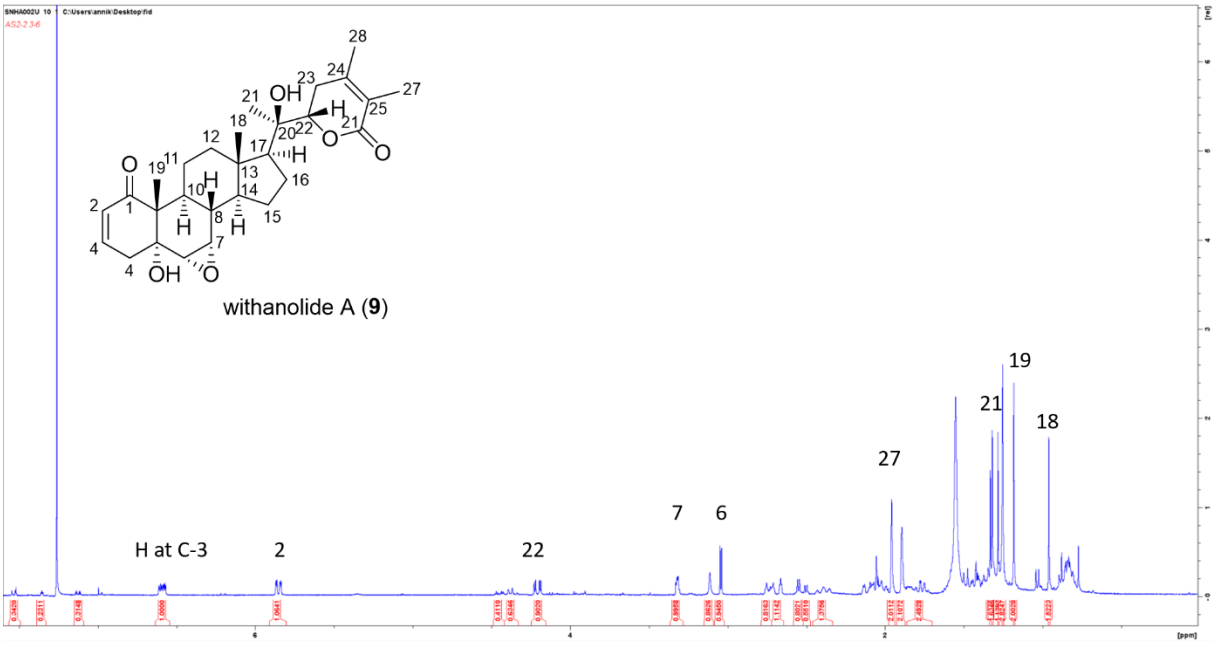


Figure 10.1 ^{1}H-NMR of withanolide A (**9**) identified by comparison to literature data, [11,25] isolated from *Withania somnifera* root powder.

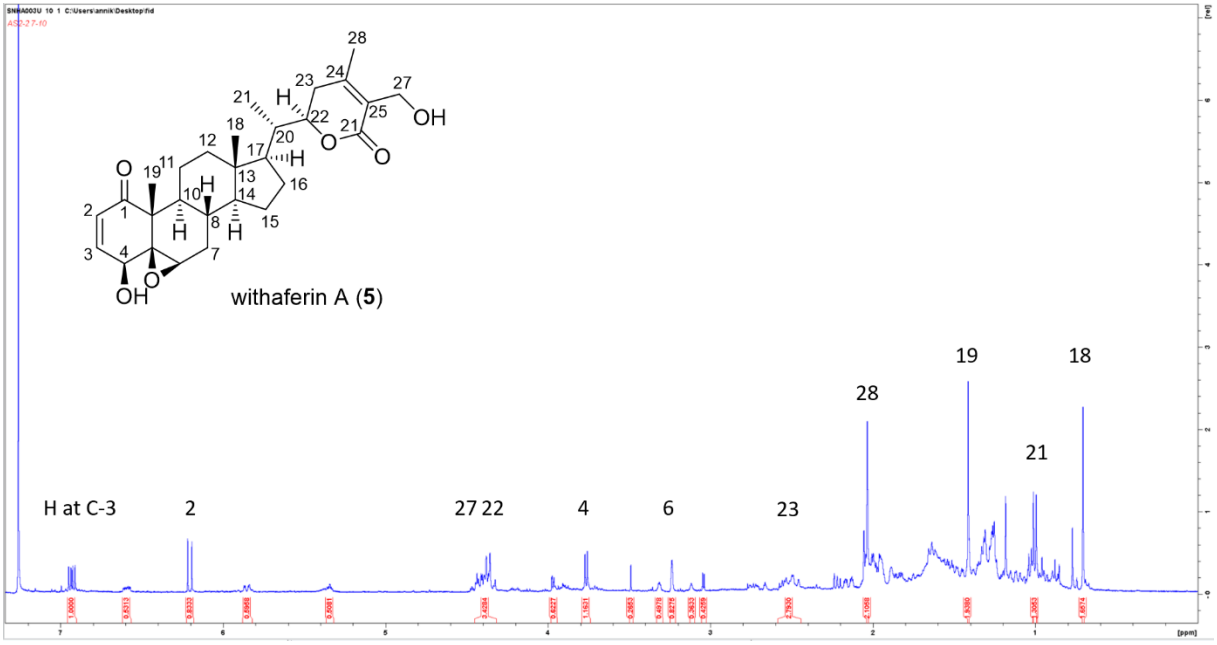


Figure 10.2 ^{1}H-NMR of withaferin A (**5**) identified by comparison to literature data, [26–28] isolated from *Withania somnifera* root powder.

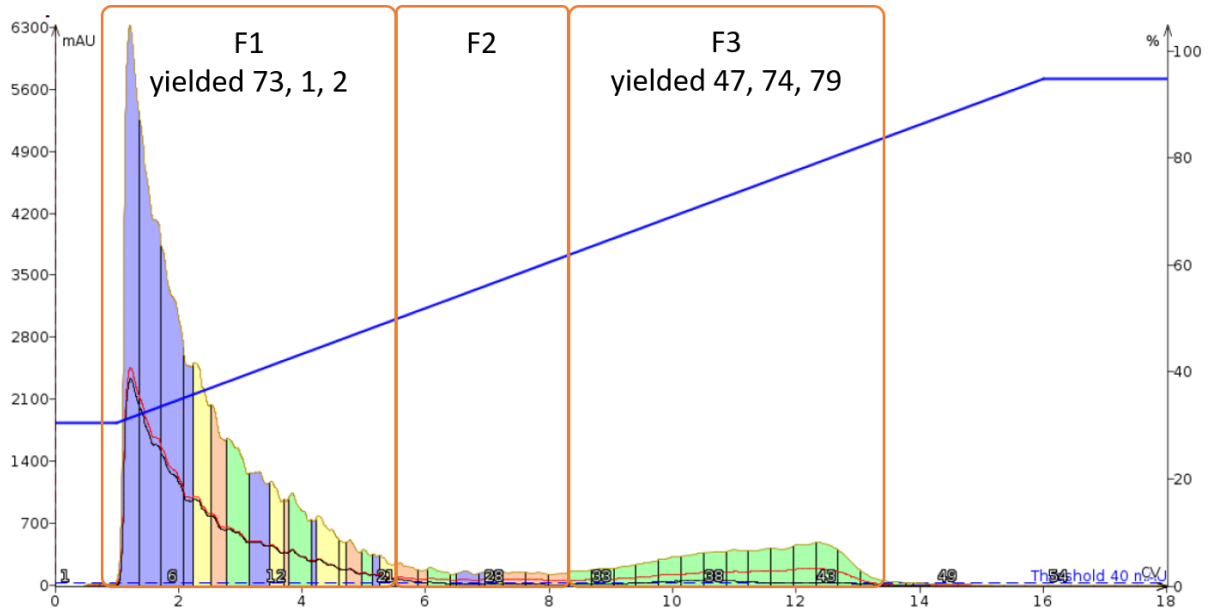


Figure 10.3 The flash chromatography run of $CHCl_3$ crude fraction (UV traces (219 nm and 228 nm)) was divided into three subfractions.

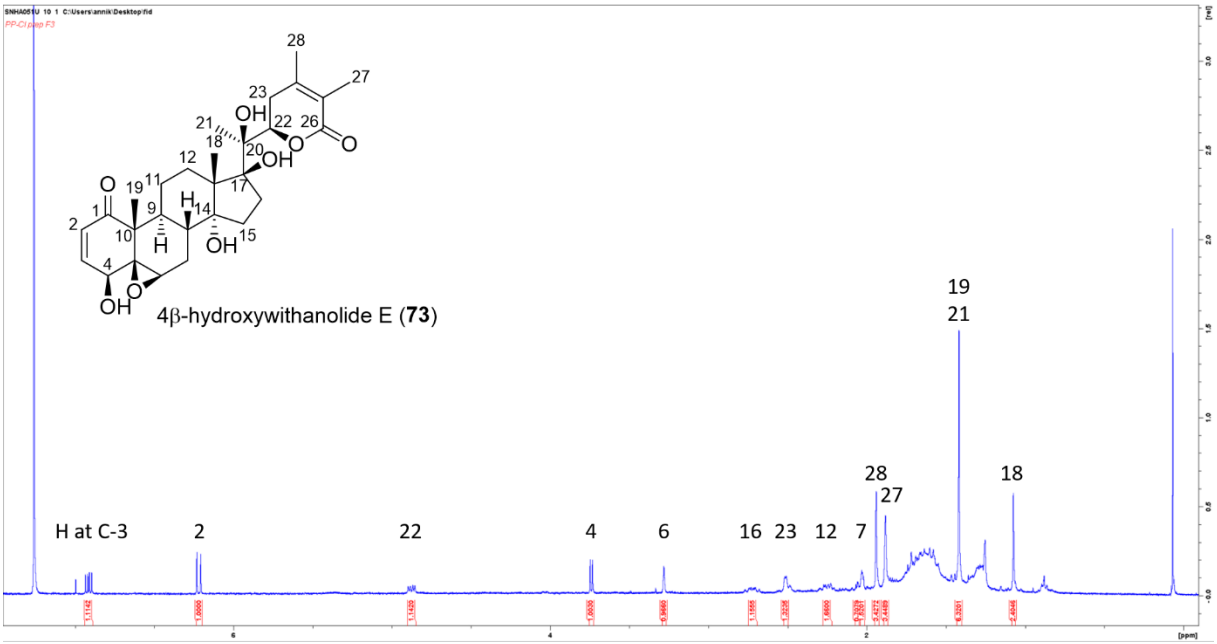


Figure 10.4 ^{1}H-NMR of 4β-hydroxywithanolide E (**73**) identified by comparison to literature data, [174] isolated from *Physalis peruviana*.

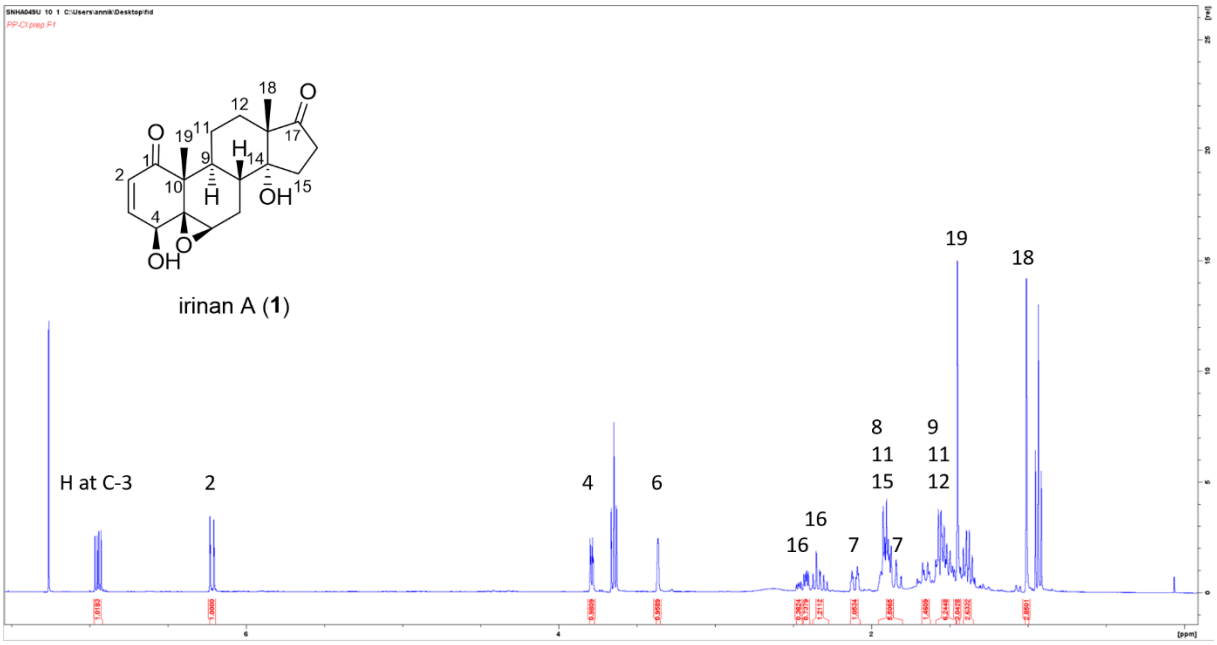


Figure 10.5 ^{1}H-NMR irinan A (**1**) isolated from *Physalis peruviana*.

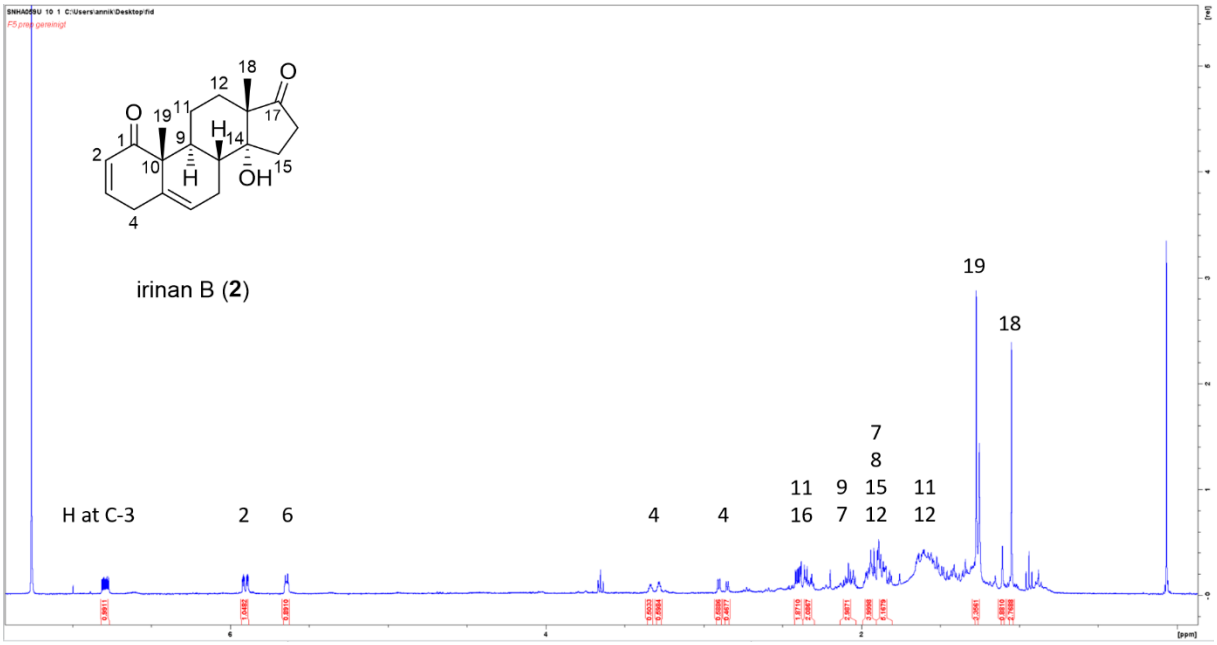


Figure 10.6 ^{1}H-NMR of irinan B (**2**) isolated from *Physalis peruviana*.

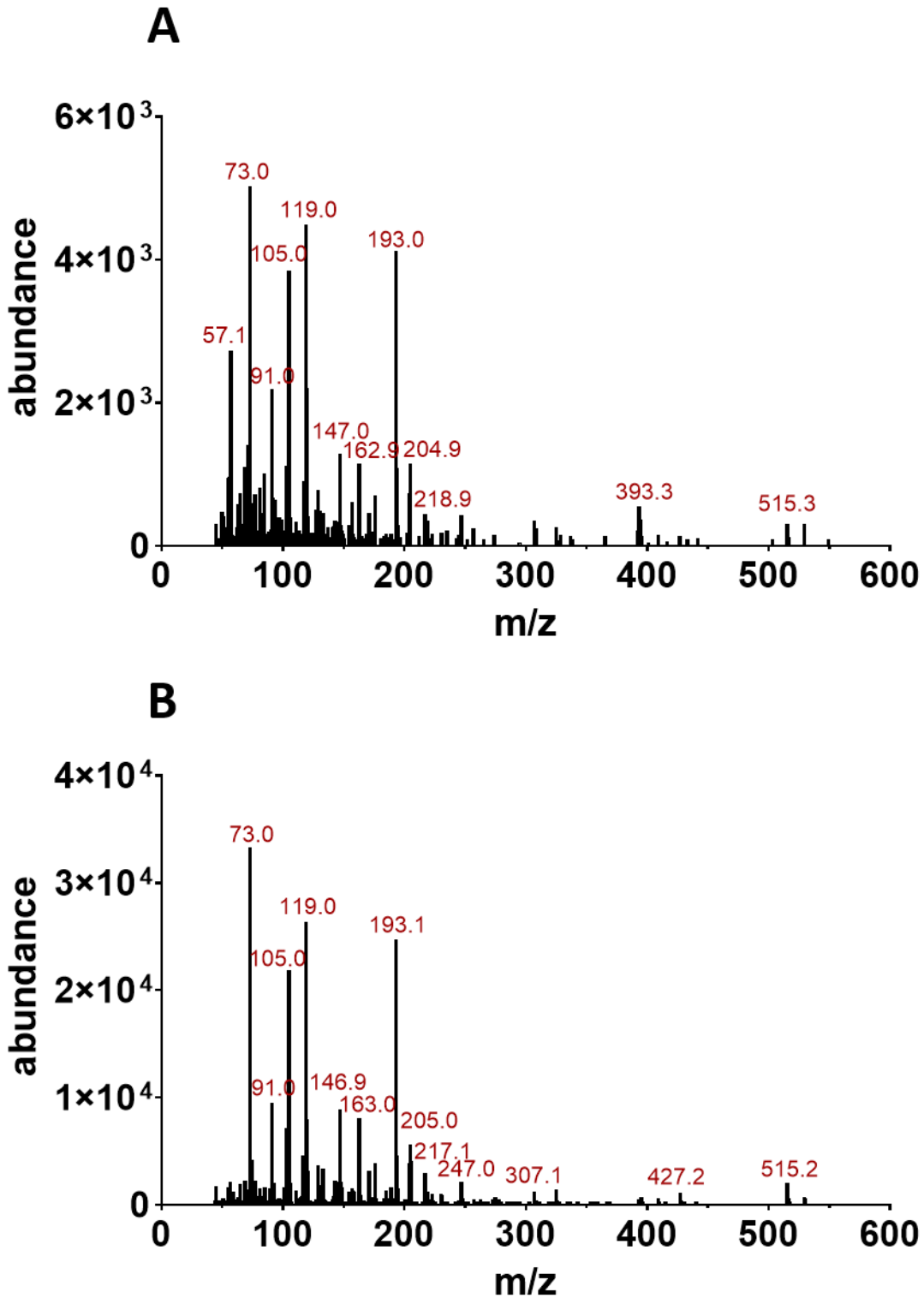


Figure 10.7 Mass spectra of GCMS traces shown in chapter 6.4.1 exhibited peaks at 10.87 min that were increased in the P450-15 leaf disk assay (B) compared to an EV control (A). The −14 increments observed in the spectra are characteristic for sequential loss of CH_2 groups, commonly observed in fatty acid spectra. The peaks at 10.87 min were therefore concluded to be non-sterol compounds and were not considered part of the withanolide biosynthesis.

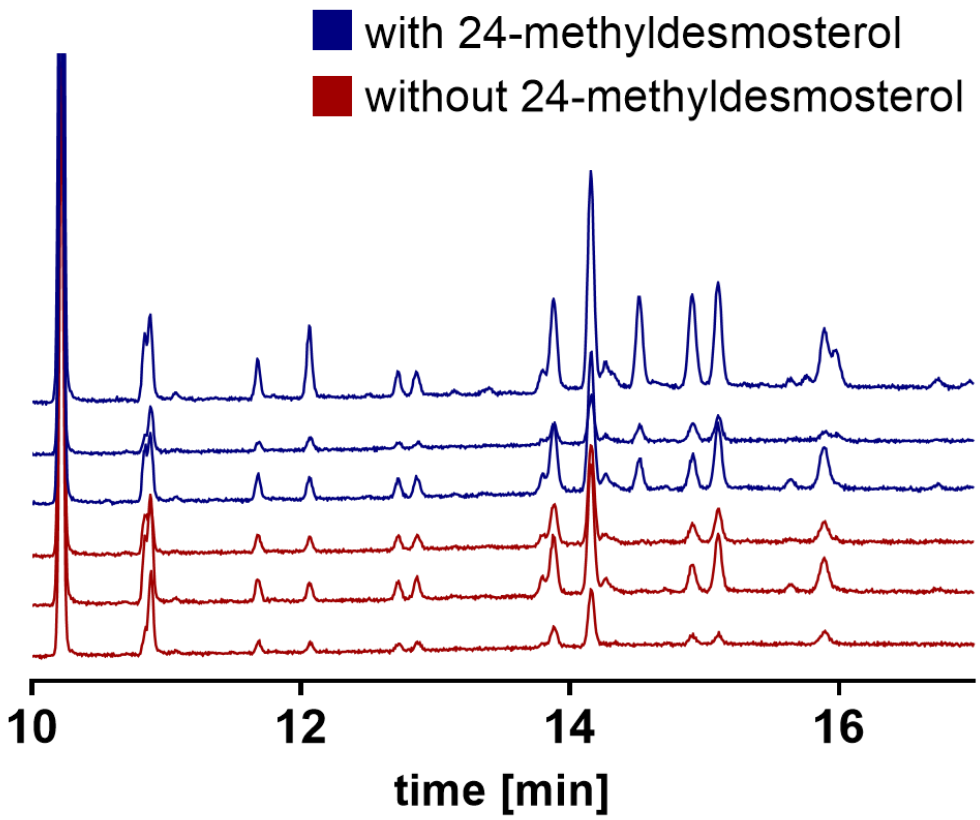


Figure 10.8 GCMS chromatograms of P450-7 *N. benthamiana* leaf disk assays (see chapter 6.4.1) showed a new peak at 15.98 min in leaf disk assays infiltrated with the substrate 24-methyldesmosterol (**3**, blue) compared to samples without the substrate (red).

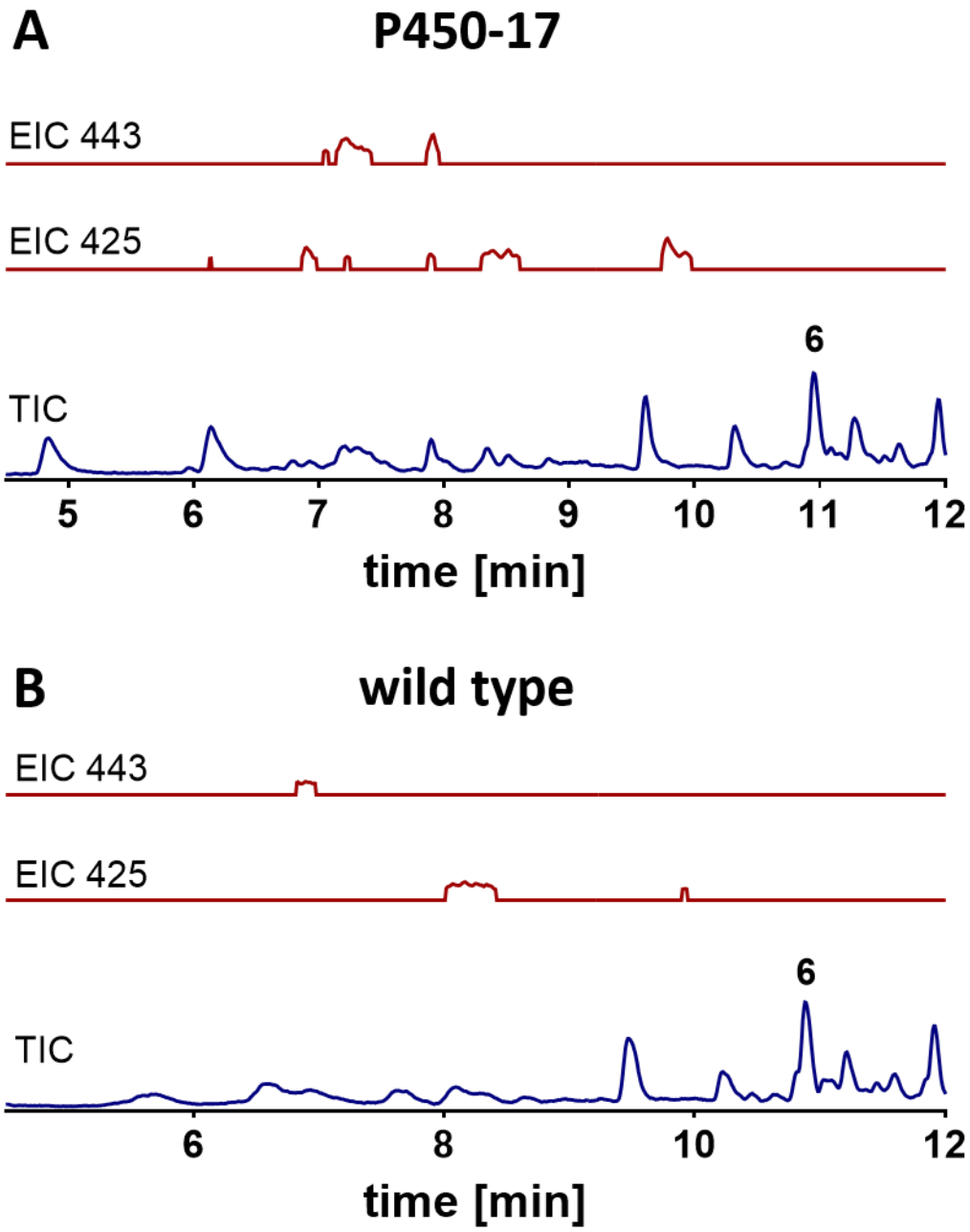


Figure 10.9 Comparison of a P450-17 and wild type (WT) microsome assay for EIC traces of oxidated cycloartenol ($M+H^+$: 443, $M-H_2O+H^+$: 425) show no major accumulation of the suspected product in P450-17; All EIC traces were amplified by 100 to overlay with TIC traces.

10.2 P450 sequences cloned into pHREAC for leaf disk assays

P450-7 (1485 bp)

showed two mutations: at amino acid 234 GAA→GAG and aa 360 AAT→AAC. Both mutations were synonymous.

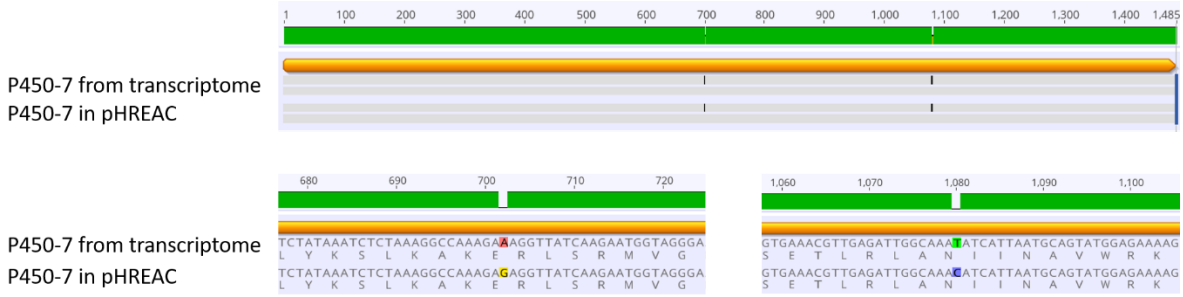


Figure 10.10 Alignment of P450-7 sequence obtained from transcriptome data and P450-17 sequence cloned into pHREAC for leaf disk assays.

atggaaggggaagcagaaagttggatttggatatgtatagcaacaggggtagtagttgtgtg
gttgatttataagatgagaaaaaatcaaaaggaagatgagaatggagttccaagaggtaatt
caggttggccacttttaggagaaacactggatttcattgcttctggatacacttctcgccct
gtcactttcatggagaaacgcaagtctcggtacgggaaagtattcaaaagtaatatactagg
aaaaggaataatagtatcaacggatgcagaagtgaacaaggtgattctacagaacaacgggg
atgtatttataccatgttatcctaaatcaataacagaattacttgggaaaaaatcaatatta
caaatgaatggaccagtacataggagagtgcacggacttattgggagtttcttaaaatcacc
gcaacttaaagctcgtattactcgagacatcgaagcctccgtccggcatttcttgtcaacat
ggctgaaaaaacaacattgtgtctacgtacaagatgaagccaaaaagattgcatttgaggta
ctagtcaagctgatactaagcgtggaccctggggaagaattgaacttgctcaagaaagaatt
tgaagaattcaccaaaggcttgatttgtttgcccattaagcttcctggaaccacactctata
aatctctaaaggccaaagaGaggttatcaagaatggtagggaaattggtagaggagagaaaa
ttgagcatggagaaaagggcagaaaaggggttaccaagtgatgcaattgacgtgctgttagg
gaatgctgatggagcaaagcaacctgtgccgtctgatttcatcagtggcaacttgatagaga
tgatgatccctggagaagagactgtaccaacggcgatcaccttggctgtcaaatttctaagt
gacaaccccgtcgttctagctcgattgctggaggagaatctggaattgaagcagcagaagat
aaatcactgtgaagattatagttggactgattatatgtcattgcccttcactcagaatgtta
tcagtgaaacgttgagattggcaaaCatcattaatgcagtatggagaaaagctctcaaagat
gtcaaaatcaaagggcatttgataccaaaaggatggtgtgttttggcatccttcacttcagt
tcacatggatgaagaaaattacgaaaatccatataattttgatccctccagatgggagaaag
ctggagttgctgtgaccagcaacacatttacaccatttggtggaggacagaggctatgtcct
ggtttagaactttcaaggcttgaaatttccattttcctccaccatcttgtcacttcctacag
atgggtggcagaaaaggatgagattgtctatttcccaacagtgaagatgaagaacaagctgc
ctatcaacatcatgccctttcagcaaaacccactatcaaaccaaacaaaactaacttaa

P450-8 (1551 bp)

no mutations observed

atggccattgtagttttaactgctttggcagtatgctttccgatatccttttttgtgcctgaa
attgctctacttagtatggtggcggcccaaaacagtagaaaaagatctgcggcagcaaggaa
tatatggccgtccatataggtttctttttggaaacctaaaggagatgatggagatgaataaa
atagcgaagtctaaacccatgcctttgaaccacgactacacccctcgacttaatccattgtt
ctatgagctcgccactacatacaagaaacttttcttgttttggctaggaccgatacctcgag
ttaccatcatggatccgaagttaatacgggaagtactgtcaaacaaatctggtgagttcagt
aaaccaagaatcagtgctttcctcaaattattcgtaacgggactggggacttacgatggtga
aaaatgggccaaacacagaaaaattcttaatccagctttccacatggaaaaattgaagctga
tgttaggaacatttgctaactgtacagaagacatgataagcagatggggcaagcttactgga
tcaacaggttcttgtgaattggatatttctcgagaatttcatagtctaacgggagacatgct
atcaaaagcagccttcggcagtaacttcgaagaagggaaattgatattttcacttctggggg
agcaatgtgaactaattttcactgcaaagcttgctattaacgtcttcccatggttaaggttc
gtaccaacgaaaactaatagaagaagattgtacatctataacacagttcgtagttcgctaaa
aggaataatagagaagcgagagaaagaggtacaatcaggaaaagcccacactgacgatctat
tgggtttgttaatgaaatctaatcaagaagaacagcaagagaataagaactcgagcaaagga
atgagtacagaggatgtgatagaggagtgcaactctttctactttgctggtcaagagactac
cgcaactttgttaacatggactgcaattgtattgaccatgcatccagattggcaagagaaag
ccaggaaagaagttcttgaaatcattggaaaagatgagcctaagtttgaccaactgaaccag
ctaaagattgtaactatgatcttgcatgaggttctgaggttgtatccatcaggctctcttgt
tagagaaacaaacaaaaagacaaagcttggagattatacaattccatcaggtgcgcaacttt
tagtgcctctacaaataatccatcgcgatactgaggcatggggagaagacgctctaattttc
aatccagaaaggttttcagaaggggtatccaaagcaaccaaggatctgatgtactttccctt
tggttggggttctcggatatgccttggaatgaatttgccatgattcaagtgaagcttgtta
tggctaaaatcatacagaactactcctttgagctttccccctcctacgctcatggtcctacc
atgccagctcttgttctacaaccacaatatggtgctcctatcatcgttcgaaagctatcatg
a

P450-15 (1476 bp)

no mutations observed

atggctactactatattcttcttgattttggcaattttattagtgctagtaatatcagctct
cttctttctttctagaatcaagccttattgtgagtgcgaaatttgcaaagggtacttgaatt
ctacatggtctttagaattcaagaacctttgtgattggtactctcatttacttaaaaaatca
ccaagtgggacaatacatctgcatgttcttggaaatgttattacagcaaatccagagaacgt
tgagtacatgctcaaaaccaaatttgaaaattttcctaaagggaaacaattttcaacaatct
taggtgatcttttaggccgtggaattttcgccgtggacggcgaaatgtggaaatttcaacgt
aaactggcaagtttagaactcggaagcgtatcgatacgttcgtacgctttcgagatcgttgg
agacgaaattagaaaccggttgattccgtttttggaaagcaacgtacaaaagggcgctgttt
ttgacttacaagatgttttgcgacggttttcgtttgatactatatgtaaattctcttttggt
atggatcctggttgtttgaaactatcattaccagtttcagatcttcaagttgcattcgattt
agcttctaaattgtcagctgaaagagctatgacaatttcacctttggtttggaaaatcaaaa
gagttttgaatattggatcagaaaagaagctaaaagaagcaattaaattggttgatattctt
gctaaagaggtgattaatcataaaaggaaaaatgtttttcttcccaaaacgacctttgtc
taggttcatgaggaatatggatgacgataaactccttagggatattgttgtcagttttcttc
tagctggaagagatactgtagcttcagctttgacaacttttttttggctattaggccaacat
cctcaagtgatagataaaatacgagccgaatctagtagagtcatgggaaaaaagggaaaaac
aacacttgcgaattttgaggaaattcgggaaatgcattatttgactgcggcgttacatgaga
gtataaggctatttccacctgttcaatttgattcgaaattttgtcaagaagatgatactctc
cctgatggtacgttcgtggctaaagggacgagggtaacttatcatccatacgcgatggggag
aatggagagtatttggggtcaagattgtctggaattcaagccagaaagatggttagatgatg
acgggggttttaaggcacaatgtccttttaagtatccggtttttcaaggtgggattagggtt
tgtttgggtaaagatttagctattgtggagatgaagagtgtagctttggctttgattaggca
atttgatatacaagtagtggcaaaagatcagacaccaaagttcatgcccggcttgacagcca
ccgttagaggtggacttcctatcatggttcaagaaagacaacaacaatag

P450-17 (1563 bp)

Amino acid 338 GAG (glutamic acid) → GAC (aspartic acid)

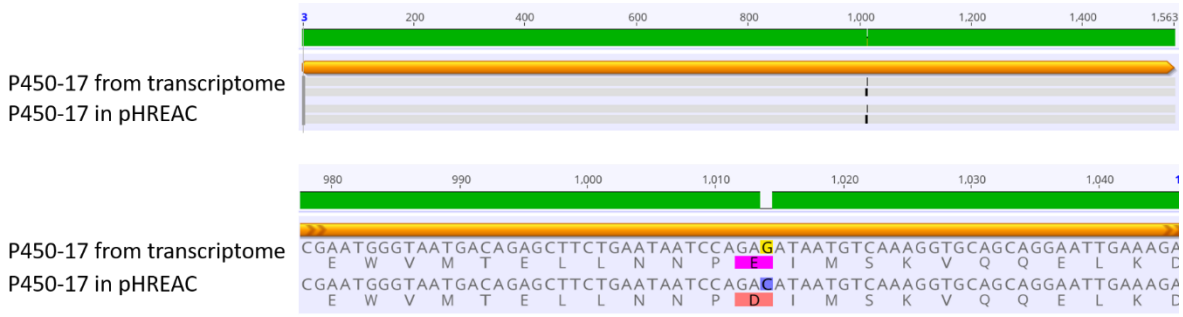


Figure 10.11 Alignment of P450-17 sequence obtained from transcriptome data and P450-17 sequence cloned into pHREAC for leaf disk assays.

atggcttcagcaactttaaacgatctttgtctagcagctctcacccttcaattgtattcct
gactattctatggtacaaactgacctcaaatatatcatccaaggaagaatctaggttgccac
cagggcctcgtggccttccgattgtcggattcataccattcctccgccccaatttgcatcac
cagctcattgagttgtctcaaaagtatggtccaatttacaagttttggttaggacgtaaact
atgtgttgtgttgaattcaccatccttagccaaacaagtggtccgtgaccaagacactgttt
ttgctgaccgcgaccctccaattgcaggattggtggccacatatggtggactagatattgga
ttcgctccctatggctcctactggcgtgagatgcgtaaactgtttgtaagggatatgctaag
caacaggaaccttgaggcctgtcgtagccttcggagacaagaggtcagaaagacaatcagaa
acatgcataccaagattggcagcctcactgatattggtgaattggcttttgtgactcaaatg
aatgtagtcatgagtatgatctttggtagcaattttgttgaggagatggagaagcatagaaa
agatggaactgaattcagggaagtggtgacaagatttcttgaagtgataggaaagccaaaca
tatcagacttcttccctctgcttgcaaggtttgatttacaaggtatacaaaaagagatgaag
gctctagtgaagtcggttgaaagtatattggaccctgccataaacgaacgtatgaagatgct
ttcggacaaaggagatgatgaaatccaggggaacagggagccggattttattcagatcctgt
tagagctaatgaagcagaaagacattggtatatcactcgacttggtgaaaataaaggccatc
ttgttggatatagtgattggtgggactgacactacaatcacgacgatcgaatgggtaatgac
agagcttctgaataatccagaCataatgtcaaaggtgcagcaggaattgaaagatgttgtag
ggatgaatagcatcgttgaagaatctcatgtaccaaatctgcactatcttgatgcagttttg
aaggagacgttacgtttacaccctgcattaccactgcttgtcccgaagcgcccgagccaatc
tgcaatagtaggtggatacacaatacctgagggaactaaagtactcttgaatgtttatgcaa
ttcatagggatcctcaactatggaaaagtccattagaattccagcctgaaaggttcttgaat
cactctacgcaattagactattctggaaatgatatgaagtaccttccatttggatcaggtag
gagaatttgtgcgggacttaatttagcagagaaaatgcttatgtttatattggcttcgttgt
tgcattcctttgattggaaactccctgtgggcaaaaacgttgatcttacagagggacttgga
ctcgtcatcaagaaaagcgagaggctattcgctatccctactcccaggttaccaaattccca
gctatatcagtaa

10.3 List of Figures, Schemes and Tables

10.3.1 List of Figures

10.3.2 List of Schemes

10.3.3 List of Tables

10.4 List of Primers

Table 10.1 Primers used in this work.

Yeast strain Primer	
Name	Sequence *
pAG25 natMX cassette for	attttgctattccaatagacaataaataccttttaacaaaCAGCTGAAGCTTCGTACGC
pAG25 natMX cassette rev	atatgatttattgtctggacaaagttctgtttttccccaGCATAGGCCACTAGTGGATCTG
erg5 locus for	CTTATTGTCTTCAGGCAAAGGGA
erg5 control rev	TTAGCCTTATATTCTTCAAACTTTGGA
erg5 control for	GTTTCCCCTAACTATACCGCA
erg5 locus rev	AAAGTCGCACCTTTAGCAGA
natMX cassette ver for	TTGACGTTGGTGACCTCCAG
natMX cassette ver rev	ACGACGAATCGGACGACG
pXP422 Pp7RED for	atctaagttttaattacaaaaATGGGGGAGTCTCAGTTG
pXP422 Pp7RED rev	gacataactaattacatgacCTAGTAAATTCCGGGTACG
pXP422 St7RED for	atctaagttttaattacaaaaATGGCGGAGTCTCAGTTGG
pXP422 St7RED rev	gacataactaattacatgacCTAATAAATTCCTGGTATGACCCGG
URA3 locus integration for	gttttgaccatcaaagaaggttaatgtggctgtggtttcagggtccataaCCGCGAATCCTTACATCACACC
Junction rev	CCCGGGGATCCTCTAGAGTCGACCGGCCGCAAATTAAAGCCTTCGAGCG
Junction for	CGCTCGAAGGCTTTAATTTGCGGCCGGTCGACTCTAGAGGATCCCCGGG
URA3 locus integration rev	ccaatttttttttttcgtcattatagaaatcattacgaccgagattcccgggAATTCGAGCTCGGTACCCGGG
URA3 locus A ver for	ACCATCAAAGAAGGTTAATGTGGC
URA3 locus B ver rev	TTCTGTATACACCCGCAGAGT
URA3 locus C ver for	CGCATTGGGTCAACAGTATAGA
URA3 locus D ver rev	TTGTTCTTTGGAGTTCAATGCGT
pYES Pp24ISO for	taagcttggtaccgagctcgATGTCAGGGGAGAAGGTG
pYES Pp24ISO rev	gcggccctctagatgcatgcTCAATCCGCGGGCTCATC
VIGS cloning primer	
Name	Sequence
pTRV2 PpPDSO for	gcctccatggggatcTGCTTTTGTGTTTGCCACTC
pTRV2 PpPDSO rev	atgcccgggcctcgaGGTTCACAACCTGGCACAGT
pTRV2 PpPDSS for	gcctccatggggatcTGGAGGTCAAGGTCTTGTTTG
pTRV2 PpPDSS rev	atgcccgggcctcgaATACTTTGCTGTAGACAAACCACC
pTRV2 WsPDS for	gcctccatggggatcGAAGAAACCTGTTCAGCGCG
pTRV2 WsPDS rev	atgcccgggcctcgaAAGCAAGCCGGGAGAATTCA
pTRV2 Ws24IS0 for	gcctccatggggatcCGCTTGCAGTTATTGGCGA
pTRV2 Ws24IS0 rev	atgcccgggcctcgaGACCCCTGAGACCAAGGGA
pTRV2 P450-1 for	gcctccatggggatcTCAGTGAAACGGTGCCAGT
pTRV2 P450-1 rev	atgcccgggcctcgaCCGAAGCCTCCTCGGATTC
pTRV2 P450-2 for	gcctccatggggatcACAGGGGCGTAAAAGTGCT
pTRV2 P450-2 rev	atgcccgggcctcgaAGTGGTTTCATGAGCCGCA
pTRV2 P450-3 for	gcctccatggggatcACAGGGCAATGAAGGCTCAA
pTRV2 P450-3 rev	atgcccgggcctcgaAGCAACAGGATGATGGTGTCA
pTRV2 P450-4 for	gcctccatggggatcTGTTCGGAAGGCCAACCAT
pTRV2 P450-4 rev	atgcccgggcctcgaTTCATCCGTGGCAAGCCAT

pTRV2 P450-5 for	gcctccatggggatcTGCGTGTTGGTTGCTGGAT
pTRV2 P450-5 rev	atgcccgggcctcgaTGGCTTAGACTTGGCTGCA
pTRV2 P450-6 for	gcctccatggggatcTCCCTCCTGGACCGTACAG
pTRV2 P450-6 rev	atgcccgggcctcgaCCTTGGCCGGTCAGAGAAA
pTRV2 P450-7 for	gcctccatggggatcGGGCAGAAAAGGGGTTACCA
pTRV2 P450-7 rev	atgcccgggcctcgaAGGTGATCGCCGTTGGTAC
pTRV2 P450-8 for	gcctccatggggatcGGAAAAGCCCACACTGACG
pTRV2 P450-8 rev	atgcccgggcctcgaCGGTAGTCTCTTGACCAGCA
pTRV2 P450-9 for	gcctccatggggatcACAGGCTTGCTCGTCGATT
pTRV2 P450-9 rev	atgcccgggcctcgaTGCTACTTGATCACCACCA
pTRV2 P450-10 for	gcctccatggggatcGCAATGGGAACAGCAGAGC
pTRV2 P450-10 rev	atgcccgggcctcgaACACTTCCATCCCCAACAA
pTRV2 P450-11 for	gcctccatggggatcAAGCTCAAGCTGTGATGGCT
pTRV2 P450-11 rev	atgcccgggcctcgaTCCTTCCATGGATGTTGGCC
pTRV2 P450-12 for	gcctccatggggatcACCGGTTGTGATGCTGAGG
pTRV2 P450-12 rev	atgcccgggcctcgaACCAGGGCCAAAAGTAGGC
pTRV2 P450-13 for	gcctccatggggatcTGGGAGTCGGAAACCAGATT
pTRV2 P450-13 rev	atgcccgggcctcgaGCACTGGTAAGCAGAGTAGCA
pTRV2 P450-14 for	gcctccatggggatcGCATGGTGGCCAAATTCCG
pTRV2 P450-14 rev	atgcccgggcctcgaAGCAAAGTGACGTTGGGCT
pTRV2 P450-15 for	gcctccatggggatcCGTGGAATTTTCGCCGTGG
pTRV2 P450-15 rev	atgcccgggcctcgaACAGCGCCCTTTTGTACGT
pTRV2 P450-16 for	gcctccatggggatcGTTTGGGCTCCTTACGGCA
pTRV2 P450-16 rev	atgcccgggcctcgaCCCAACGTTTCCCAGTACCA
pTRV2 P450-17 for	gcctccatggggatcCTGCTTGCAAGGTTTGATTTACA
pTRV2 P450-17 rev	atgcccgggcctcgaAAATCCGGCTCCCTGTTCC
pTRV2 P450-18 for	gcctccatggggatcGTGGGCACATTAAAGCCATGA
pTRV2 P450-18 rev	atgcccgggcctcgaCGGTGCTCTCTGATGCTGT
pTRV2 P450-19 for	gcctccatggggatcTGTCATTGAAGTTTCTCGCACC
pTRV2 P450-19 rev	atgcccgggcctcgaCTGTATTGGCCTCCCCTGC
pTRV2 P450-20 for	gcctccatggggatcACTATGCACCAGTGGTGGG
pTRV2 P450-20 rev	atgcccgggcctcgaTCGCTTGCCTTTGGAGTACT
pTRV2 P450-21 for	gcctccatggggatcACCAACACACACCTTCCCA
pTRV2 P450-21 rev	atgcccgggcctcgaCAGGGCCATGACGATTGGA
pTRV2 DH1 for	gcctccatggggatcTCCACTTGCTGAGAAGGCTC
pTRV2 DH1 rev	atgcccgggcctcgaTCCAAATGCCCTTTCCATCC
pTRV2 DH2 for	gcctccatggggatcCCGCTTTCGCATTTCGGAA
pTRV2 DH2 rev	atgcccgggcctcgaCGCATTTTGAGCCCTCACG
pTRV2 DH3 for	gcctccatggggatcTCGCACGGGTTTTCCAAGT
pTRV2 DH3 rev	atgcccgggcctcgaGGTTGCCCAGGAGCTGAAA
pTRV2 DH4 for	gcctccatggggatcTTCACATTGCCAGGTGCCA
pTRV2 DH4 rev	atgcccgggcctcgaGGCTTGTCTAAGGCTGCCA
pTRV2 DH5 for	gcctccatggggatcTCAGGATTGGGTTGTTCGAAGA
pTRV2 DH5 rev	atgcccgggcctcgaTCTGAAAGCAATGACCTTCTCCT
pTRV2 DH7 for	gcctccatggggatcCGGTGAAGTCCAACTCCCC
pTRV2 DH7 rev	atgcccgggcctcgaCTCCAGCGGTATCGTCGAC
pTRV2 DH8 for	gcctccatggggatcAGGAACTAGCGTGCTGTTCA
pTRV2 DH8 rev	atgcccgggcctcgaACATCCGCCATGGAAGCTT

pTRV2 DH9 for	gcctccatggggatcGCCTGATGGACTATGGACGT
pTRV2 DH9 rev	atgcccgggcctcgaCCCAATTCACCTTTCCACCC
pTRV2 DH10 for	gcctccatggggatcTGAGAGATTGATGGGCGAAGG
pTRV2 DH10 rev	atgcccgggcctcgaGTGGAAGTAGATGAAAAGGGGGA
pTRV2 DH11 for	gcctccatggggatcTGAGCGTTTCTACGACGTACA
pTRV2 DH11 rev	atgcccgggcctcgaAGGGAGTTTTTGGCTACGGG
pTRV2 DH12 for	gcctccatggggatcCCACTTTGTGACGATCCGC
pTRV2 DH12 rev	atgcccgggcctcgaACACAAAGGCACGAGGTGA
pTRV2 DH13 for	gcctccatggggatcAAATGCTCCAGTTTGTGACGA
pTRV2 DH13 rev	atgcccgggcctcgaACCACCGTGAATGTAAATCACA
pTRV1 ver for	AGCAGCAACCGACGACTT
pTRV1 ver rev	ACCAACTCCTTCTTCTCAGACT
pTRV2 ver for	TAATGGTTTGGTGGTCAAGGT
pTRV2 ver rev	CAGTCTATACACAGAAACAGATAACAA
RT-qPCR Primer	
Name	Sequence
PDSO qPCR for	CCAGGGCAAAAGCAAATAAA
PDSO qPCR rev	TTTACGCAAGTGGCCAAAC
PDSS qPCR for	GCAATGCTTGGAGGGCAATC
PDSS qPCR rev	CATCTGTCACCCTATCCGGC
WsPDS qPCR for	AAACCGATACTGCTGGAGGC
WsPDS qPCR for	CAGTCTCGTACCAATCCCCA
WsEF1a qPCR for	TCATTGGCCACGTCGATTCT
WsEF1a qPCR rev	GTCAAGCACCCAGGCATACT
Ws24ISO qPCR fwd	CATTTTCGCCGCATAATCAACA
Ws24ISO qPCR rev	CGTGGAGATTTGAGAGACTCC
WsTubulin for	ACCTGCGAAAGTTGGCTGTA
WsTubulin rev	TCCACGAGAGGTTAATGGCG
Microsome Assay Primer	
Name	Sequence
pYES2 24ISO for	taccgagctcggatcATGTCAGATGAGAAGGAGGCCAC
pYES2 24ISO rev	tagatgcatgctcgaTCAATCCGCAGGCTCATCGAC
pYES2 AtC4H for	taccgagctcggatcATGGACCTCCTCTTGCTGG
pYES2 AtC4H rev	tagatgcatgctcgaTTAACAGTTCCTTGGTTTCATAACG
pYES2 WsC4H for	taccgagctcggatcATGGATCTTCTCTTATTAGA
pYES2 WsC4H rev	tagatgcatgctcgaTCAAAAAGATCTTGGTTTCAT
pYES2 AtS22D for	taccgagctcggatcGTTTTCTCTGTTTCTATATTTGCCT
pYES2 AtS22D rev	tagatgcatgctcgaGGAAAAGTTGGGATACTTTGCGA
pYES2 WsS22D1 for	taccgagctcggatcTTGCACATGGCCTCCATTTGG
pYES2 WsS22D1 rev	tagatgcatgctcgaTCATGAAAGAGATGGGAATCGAGTG
pYES2 WsS22D2 for	taccgagctcggatcTTGACCCTTGTTCCTTCTTCCC
pYES2 WsS22D2 rev	tagatgcatgctcgaTCATGAAAGAGATGGGAATCGAGTG
pYES2 StC22H for	taccgagctcggatcATGGGTATTGCAGTTTTCATTGCT
pYES2 StC22H for	tagatgcatgctcgaTCATGATAGCTTTCGAAGGATCA
pYES2 P450-1 for	taccgagctcggatcATGGAATGGGAATGGAGCTATTTG
pYES2 P450-1 rev	tagatgcatgctcgaCTAGATAAGATTGATGAGTGTCTCA
pYES2 P450-7 for	taccgagctcggatcATGGAAGGGGAAGCAGAAAG
pYES2 P450-7 rev	tagatgcatgctcgaTTAAGTTAGTTTTGTTTGGTTTGAT

pYES2 P450-8 for	taccgagctcggatcATGGCCATTGTAGTTTTAACTGC
pYES2 P450-8 rev	tagatgcatgctcgaTCATGATAGCTTTCGAACGATG
pYES2 P450-14 for	taccgagctcggatcTTGTGTGACAACTTATATTTGTTAG
pYES2 P450-14 rev	tagatgcatgctcgaTCATGCCTGTTGGTTGTG
pYES2 P450-15 for	taccgagctcggatcATGGCTACTACTATATTCTTCTTG
pYES2 P450-15 rev	tagatgcatgctcgaCTATTGTTGTTGTCTTTCTTGAACC
pYES2 P450-18 for	taccgagctcggatcATGGCTTCAGCAACTTTAAACG
pYES2 P450-18 rev	tagatgcatgctcgaTTACTGATATAGCTGGGAATTTGG
pYES2 ver for	GCGTGAATGTAAGCGTGAC
pYES2 ver rev	AATATACCTCTATACTTTAACGTC
***N. benthamiana* expression**	
Name	Sequence
pHREAC P450-1 for	caccacaggtctcgaaaaATGGAATGGGAATGGAGCTA
pHREAC P450-1 rev	caccacaggtctcgagcgCTAGATAAGATTGATGAGTGTCTCAAGC
pEAQ-HT P450-3 for	caaattcgcgaccggTTGTCAACTGTTCATGAGTGTG
pEAQ-HT P450-3 rev	agttaaaggcctcgaTCATGCTGAAATCTTCTTGAATTTC
pHREAC P450-4 for	caccacaggtctcgaaaaTTGATTGAGTGGGGAGTAAAAATGG
pHREAC P450-4 rev	caccacaggtctcgagcgTTATGCTGAGAGCTTCTTGA
pEAQ-HT P450-5 for	caaattcgcgaccggATGGAGATTGTTGTTTACTACTTAG
pEAQ-HT P450-5 rev	agttaaaggcctcgaTTACAACTTGTGCAACACAAGAGG
pHREAC P450-6 for	caccacaggtctcgaaaaATGGACATCTGGTTTACCTTCGT
pHREAC P450-6 rev	caccacaggtctcgagcgTCAAACAGAGTTAACTCGGGGG
pHREAC P450-7 for	caccacaggtctcgaaaaATGGAAGGGGAAGCAGAAAGT
pHREAC P450-7 rev	caccacaggtctcgagcgTTAAGTTAGTTTTGTTTGGTTTGA
pHREAC P450-8 for	caccacaggtctcgaaaaATGGCCATTGTAGTTTTAACTGCT
pHEAC P450-8 rev	caccacaggtctcgagcgTCATGATAGCTTTCGAACGATGA
pHREAC P450-9 for	caccacaggtctcgaaaaTTGCCTCCAGGTGACATGG
pHREAC P450-9 rev	caccacaggtctcgagcgCTAAAATAGAACTACAAAATTTGGTTGA
pHREAC P450-10 for	caccacaggtctcgaaaaATGATGATAGCAGTAATAGCT
pHREAC P450-10 rev	caccacaggtctcgagcgTTACAGTGGATGAAGCATGATTTGG
pHREAC P450-11 for	caccacaggtctcgaaaaATGATGCTCTTTCTACTCTTTCT
pHREAC P450-11 rev	caccacaggtctcgagcgCTACAAATAATTTTTAGGAATAAGACA
pHREAC P450-12 for	caccacaggtctcgaaaaATGGAGTTAAGTGACAACAAGT
pHREAC P450-12 rev	caccacaggtctcgagcgTCATTCAGTGGAAAGCTTTCGG
pHREAC P450-13 for	caccacaggtctcgaaaaTGGGAGTCGGAAACCAGATT
pHREAC P450-13 rev	caccacaggtctcgagcgGCACTGGTAAGCAGAGTAGCA
pEAQ P450-14 for	caaattcgcgaccggTTGTGTGACAACTTATATTTGTTAG
pEAQ P450-14 for	agttaaaggcctcgaTCATGCCTGTTGGTTGTG
pHREAC P450-15 for	caccacaggtctcgaaaaATGGCTACTACTATATTCTTCTTGA
pHREAC P450-15 for	caccacaggtctcgagcgCTATTGTTGTTGTCTTTCTTGAACCA
pHREAC P450-17 for	caccacaggtctcgaaaaATGGCTTCAGCAACTTTAAACGA
pHREAC P450-17 rev	caccacaggtctcgagcgTTACTGATATAGCTGGGAATTTGGT
pEAQ-HT Ws24ISO for	caaattcgcgaccggATGTCAGATGAGAAGGAGGCCAC
pEAQ-HT Ws24ISO rev	agttaaaggcctcgaTCAATCCGCAGGCTCATCGAC
pHREAC StPGA2 for	caccacaggtctcgaaaaATGGGTATTGCAGTTTTCATTGCT
pHREAC StPGA2 rev	caccacaggtctcgagcgTCATGATAGCTTTCGAAGGATCA
pHREAC DH1 for	caccacaggtctcgaaaaATGGCTAATTCCAATGACAACAACA
pHREAC DH1 rev	caccacaggtctcgagcgTCACTTATGTTGGATGAAAGAAGCA

pHREAC DH2 for	caccacaggtctcgaaaaATGCATACCAGTGCTCCGTC
pHREAC DH2 rev	caccacaggtctcgagcgCTATTCCCTAGAAGGACACCATGG
pHREAC DH3 for	caccacaggtctcgaaaaATGGAAGGTGAAGGGGCAAT
pHREAC DH3 rev	caccacaggtctcgagcgTCAGTATGTTGCCATTATATTCTTGA
pHREAC DH5 for	caccacaggtctcgaaaaATGGACTCCAGCAGTGACAA
pHREAC DH5 rev	caccacaggtctcgagcgTCACACACAACTCTGTTTGA
pHREAC DH7 for	caccacaggtctcgaaaaATGCATACCATTACTCCCTCATCG
pHREAC DH7 rev	caccacaggtctcgagcgCTACTCGTCTTCCCTCGAAGG
pHREAC DH11 for	caccacaggtctcgaaaaATGTCTTCCGAAGACAATGTGGT
pHREAC DH11 rev	caccacaggtctcgagcgTTACGGGCAATTGATGAATGCA

*Lowercase letters: homology arms to vector/integration site

10.5 Manual RNA extraction using CTAB

This protocol follows a protocol from Meisel *et al.* with some alterations. [368] All consumables, reagents, equipment, & surfaces should be clean from RNA, DNA or any other contaminants and should be treated with RNaseZap or similar products. All water used for purification should be RNAse free.

Preparation of Chemicals

The extraction buffer (EB, Table 10.2) is prepared in water and autoclaved before use. The buffer can be stored at room temperature.

Table 10.2 Components of the extraction buffer for RNA purification.

Chemicals	Final Conc.	Makes 100 mL
CTAB	2% (w/v)	2 g
PVP40	2% (w/v)	2 g
Tris-HCl (1 M, pH 8.0)	100 mM	10 mL
EDTA (0.5 M pH 8.0)	25 mM	5 mL
NaCl (5 M)	2 M	40 mL

On the day of extraction, add 0.05% SPD (spermidine trihydrochloride) and 2% (v/v) BME (β-mercaptoethanol) to extraction buffer (See Step #7).

A 100X SPD stock solution can be prepared by solving 50 mg SPD (spermidine trihydrochloride) in 1 mL water, which should be sterile filtered afterwards (pore size 0.2 µm). The stock solution can be stored at -20 °C for a maximum of 1 month.

Furthermore, the SSTE buffer has to be prepared (Table 10.3). After autoclavation, this buffer can be stored at room temperature.

Table 10.3 Components of the SSTE buffer for RNA purification.

Chemical	Final Conc.	Makes 100 mL
NaCl (5 M)	1 M	20 mL
SDS	0.5%	0.5 g
Tris-HCl (1 M, pH 8.0)	10 mM	1 mL
EDTA (0.5 M pH 8.0)	1 mM	0.2 mL

Further solutions needed are a 10 M LiCl (42.39 g/mol) solution, that needs to be autoclaved before use, 99% and 75% ethanol, that need to be stored at -20°C and a chloroform-isoamyl alcohol (24:1, v/v) mixture.

Extraction Procedure

Clean gloves have to be worn at all times. All surfaces and gloves need to be treated with RNase ZAP regularly. After precipitation (step 14), samples have to be kept on ice.

1. 10 mL EB were heated to 65 °C, then 200 µL of BME were and 100 µL 100X SPD were added.
2. 100 mg of frozen ground tissue were transferred to a pre-cooled 2-mL tube and add 800 µL of warm extraction buffer were added. The samples were mixed immediately by vortexing. If more tissue was available, it was aliquoted to several tubes or stored at -80 °C.
3. All samples were incubated at 65 °C for 10 min, and vortexed vigorously during incubation (8 times). Afterwards all samples were cooled until reaching room temperature.
4. 600 µL of a 24:1 chloroform:isoamyl alcohol mixture were added and samples were shaken vigorously for 2 min.
5. After centrifugation at room temperature 12,000 ×g for 20 min, the aqueous phase was transferred to a 1.5-mL tube. In no clear separation of phases was observed, centrifugation was carried out for longer periods.
6. The aqueous phase was re-extracted equal volume (~550 µL) of 24:1 the chloroform: isoamyl alcohol mixture, and again shaken vigorously for 2 min.
7. Centrifugation and recovery of the supernatant (step 5) were repeated.
8. To precipitate nucleic acids, 0.25 volumes (~135 µL) of 10 M LiCl were added to the ~550 µL aqueous phase and mixed by inverting the tubes. The samples were stored at 4°C overnight.
9. RNA was pelleted by centrifuging at 14,000 ×g for 40 min at 4 °C. The supernatant was discarded.
10. The pellet was resuspended in 500 µL SSTE buffer and was shaken vigorously for ~10 seconds.
11. The sample was extracted with equal volume (ca. 500 µL) of 24:1 chloroform:isoamyl alcohol and shaken vigorously for 2 min.
12. After centrifugation at 14,000 x g for 30 min at 4 °C, the aqueous phase was transferred to a new 1.5 ml reaction tube.
13. Two volumes (~1000 µL) of ice-cold absolute ethanol (100% EtOH) were added and mixed by vigorous shaking for ~10 sec.

14. To precipitate RNA the samples were incubated at -80°C for 90 min and then centrifuged at 14,000 x g for 30 min, 4°C. The supernatant was discarded.
15. The pellets were washed with 1 mL 75% ice cold ethanol and mixed by vigorous shaking for ~10 sec.
16. Centrifugation was carried out at 14,000 ×g for 30 min, 4 °C. The supernatant was discarded.
17. The pellets were dried completely by incubation at room temperature under a fume hood with open lids. Alternatively, residual ethanol was removed under reduced pressure without heating.
18. The RNA was dissolved 30 µL of nuclease-free water.

The quality of the RNA was analyzed with nanodrop and gel electrophorese analysis. Afterwards, a digestion of genomic DNA was performed using the TURBO DNA-free Kit (Invitrogen) according to the manufacturer's instructions. RNA was stored at −80 °C and multiple thawing and freezing was avoided when possible.

Curriculum Vitae **Annika Stein**

Apr 2018 - today	Dr. rer. nat., Natural Product Chemistry, Gottfried Wilhelm Leibniz Universität Hannover
Okt 2015 – Okt 2017	Master of Science, Life Science, Gottfried Wilhelm Leibniz Universität Hannover
Okt 2012- Okt 2015	Bachelor of Science, Life Science, Gottfried Wilhelm Leibniz Universität Hannover
2004 -2011	Abitur, Gymnasium Lutherschule, Hannover

List of Publications

[1] Stein, A., Compera, D., Karge, B., Brönstrup, M. & Franke, J. Isolation and characterisation of irinans, androstane-type withanolides from Physalis peruviana L. *Beilstein J. Org. Chem.* **15**, 2003–2012 (2019).