Elucidation of the secoiridoid pathway in Catharanthus roseus

Miettinen, K.

Chapter 5

Summary

Catharanthus roseus (Madagascar periwinkle) is the best studied medicinal plant. It produces the important class of secondary metabolites the monoterpenoid indole alkaloids (MIA) and their precursors the (seco)iridoids, bioactive compounds with a wide spectrum of high-value pharmacological and insect-repellent activities. From the over 130 known *C. roseus* MIAs many have pharmaceutical applications such as the anti-hypertensive drugs serpentine and ajmalicine and the potent antitumor agents bisindole alkaloids vincristine and vinblastine that are widely used to treat several types of cancer such as Hodgkins disease, Kaposi's sarcoma, breast cancer, bladder cancer and testicular cancer. Vinblastine and vincristine are produced by *C. roseus* in extremely low levels, leading to high market prices and poor availability. Because of their complex structures total synthesis is unfeasible (van der Heijden et al., 2004). After decades of research only parts of the MIA biosynthetic pathway are known. Complete knowledge of the biosynthesis is essential for biotechnological production of MIA's and (seco)iridoids.

The aim of this work was elucidation and further description of the iridoid pathway from *C. roseus* from a metabolic engineering perspective. A candidate based approach was taken to find enzymes catalyzing both hypothetical and completely unknown intermediate reactions in iridoid biosynthesis and to find suitable enzymes for biotechnological applications. The candidates were picked based on amino acid and nucleotide sequence homology with known enzymes and the screening was refined by gathering additional information such as expression pattern in different conditions and established tissue localization of proteins and transcripts. Based on the evidence at the onset of these studies the pathway was missing 4-6 enzymes with completely new functionalities. In addition the gene coding for enzyme geraniol synthase (GES) was known from *Ocimum basilicum* (Iijima et al., 2004), *Cinnamomum tenuipilum* (Yang et al., 2005), and *Perilla citriodora* (Ito and Honda, 2007) but *C.roseus GES* (*CrGES*) was not identified. Several oxidoreductases of different classes, cytochrome P450's, glucosyltransferases and a terpene synthase were chosen as candidates according to our pathway model (Fig. 1). These candidates were expressed in *Escherichia coli* and *Saccharomyces cerevisiae* and used for candidate screening with *in vitro* biochemical assays. The pathway was completely reconstituted by transiently expressing the final four candidates and known iridoid biosynthesis genes in *Nicotiana benthamiana* and new metabolites were analyzed by

LC-MS, *CrGES* was validated by *in vitro* biochemical assays of an enzyme produced in *E. coli* and it was further characterized by expression in a farnesyl biosynthesis mutant *S. cerevisiae* strain and by *in situ* hybridization and by transient expression of a GFP-fusion in a *C. roseus* cell culture. Geraniol synthases from *Valeriana officinalis* (*VoGES*) and *Lippia dulcis* (*LdGES*) were characterized *in vitro* and *in planta* to assess biotechnological aspects of monoterpenoid biosynthesis.

Chapter 1 serves as a general overview of MIA and iridoid biosynthesis and present the research strategy for this work.

Chapter 2 presents the cloning and functional characterization of the *C. roseus* cDNA for *CrGES* encoding geraniol synthase (CrGES). The full-length CrGES was over-expressed in *E. coli* and the purified recombinant protein catalyzed the *in vitro* conversion of GPP into geraniol only with a Km value of 58.5 µM for GPP. *In vivo* CrGES activity was evaluated by heterologous expression in *S. cerevisiae* strain K197G mutated in the farnesyl diphosphate synthase gene. Analysis of culture extracts by gas chromatography-mass spectrometry confirmed the excretion of geraniol into the growth medium. Transient transformation of *C. roseus* cells with a Yellow Fluorescent Protein-fusion construct revealed that CrGES is localized in the plastid stroma and the stromules. This could suggest that the availability of its product geraniol to the next enzyme G10H could be facilitated through stromule/ER interactions. In aerial plant organs, RNA *in situ* hybridization showed specific labelling of *CrGES* transcripts in the internal phloem associated parenchyma as observed for other characterized genes involved in the early steps of MIA biosynthesis. Finally, when cultures of *Catharanthus* cells were treated with the alkaloid-inducing hormone methyl jasmonate, an increase in *CrGES* transcript levels was observed. This observation coupled with the tissue-specific expression and the subcellular compartmentalization support the idea that CrGES initiates the monoterpenoid branch of the MIA biosynthetic pathway.

Chapter 3 discloses the recombinant expression and characterization of two geraniol synthases (GES). GESs from *Valeriana officinalis* (*VoGES*) and *Lippia dulcis* (*LdGES*), were isolated from *E. coli* and were shown to turn geranyl diphosphate (GPP) into geraniol as a specific product *in vitro* with K_m values of 32 µM and 51 µM for GPP, respectively. The *in*

planta enzymatic activity and sub-cellular localization of VoGES and LdGES were characterized in stably transformed tobacco plants and using transient expression in *N. benthamiana*. Transgenic tobacco expressing VoGES or LdGES accumulate geraniol, oxidized geraniol compounds like geranial, geranic acid and hexose conjugates of these compounds to similar levels. Geraniol emission of leaves was lower than that of flowers, which could be related to higher levels of competing geraniol-conjugating activities in leaves. GFP-fusions of the two GES proteins show that VoGES resides (as expected) predominantly in the plastids, while LdGES import into to the plastid is clearly impaired compared to that of VoGES, resulting in both cytosolic and plastidic localization. Geraniol production by VoGES and LdGES in *N. benthamiana* was nonetheless very similar. Expression of a truncated version of *VoGES* or *LdGES* (cytosolic targeting) resulted in the accumulation of 30% less geraniol glycosides than with the plastid targeted VoGES and LdGES, suggesting that the substrate geranyl diphosphate is readily available, both in the plastids as well as in the cytosol. The potential use of GES in the engineering of the MIA pathway in heterologous hosts is discussed.

Chapter 4 reports the discovery of the last five missing steps of the (seco)iridoid biosynthesis pathway from *C. roseus* (Fig. 1) using an integrated approach combining gene expression and coexpression analysis, tissue-specific proteomics and biochemical assays. Five new biosynthesis genes were cloned and the corresponding enzymes characterized. One of the enzymes, the alternative terpene cyclase iridoid synthase (IS) was recently published while this work was in progress (Geu-Flores et al., 2012). The novel enzyme 8-hydroxygeraniol oxidoreductase (8-HGO) catalyzes the oxidation of 8-hydroxygeraniol both at position 1 and 8 to ultimately yield 8-oxogeranial. Another novel enzyme, iridoid oxidase (IO, CYP76A26) was found to turn *cis-trans*-nepetalactol (also known as iridodial hemiacetal form) and *cis-trans*-iridodials into 7-deoxyloganetic acid. Also iridotrial was utilized by the enzyme resulting in the same product. 7-deoxyloganetic acid was in turn used by 7-deoxyloganetic acid glucosyltransferase (7DLGT, UGT709C2) to produce 7-deoxyloganic acid. This compound was hydroxylated by a fourth enzyme, 7-deoxyloganic acid hydroxylase (7DLH, CYP72A224), into loganic acid. All of the enzymes were found to be expressed in vascular tissue of *C. roseus* aerial organs, notably the internal phloem associated parenchyma (IPAP) cells, by in situ hybridization. The novel biosynthesis genes showed a very similar expression

pattern in whole plants and in tissues cultures in multiple conditions. The subcellular localization of GFP fusions of the CYP450 enzymes (IO, 7DLH) was found to be in the ER as expected. 8-HGO and 7DLGT N- and C-terminal fusion proteins were found localize in the cytosol. Expression of the eight genes encoding this pathway together with two genes boosting precursor formation and two downstream alkaloid biosynthesis genes in an alternative plant host allowed the heterologous production of the complex MIA strictosidine.

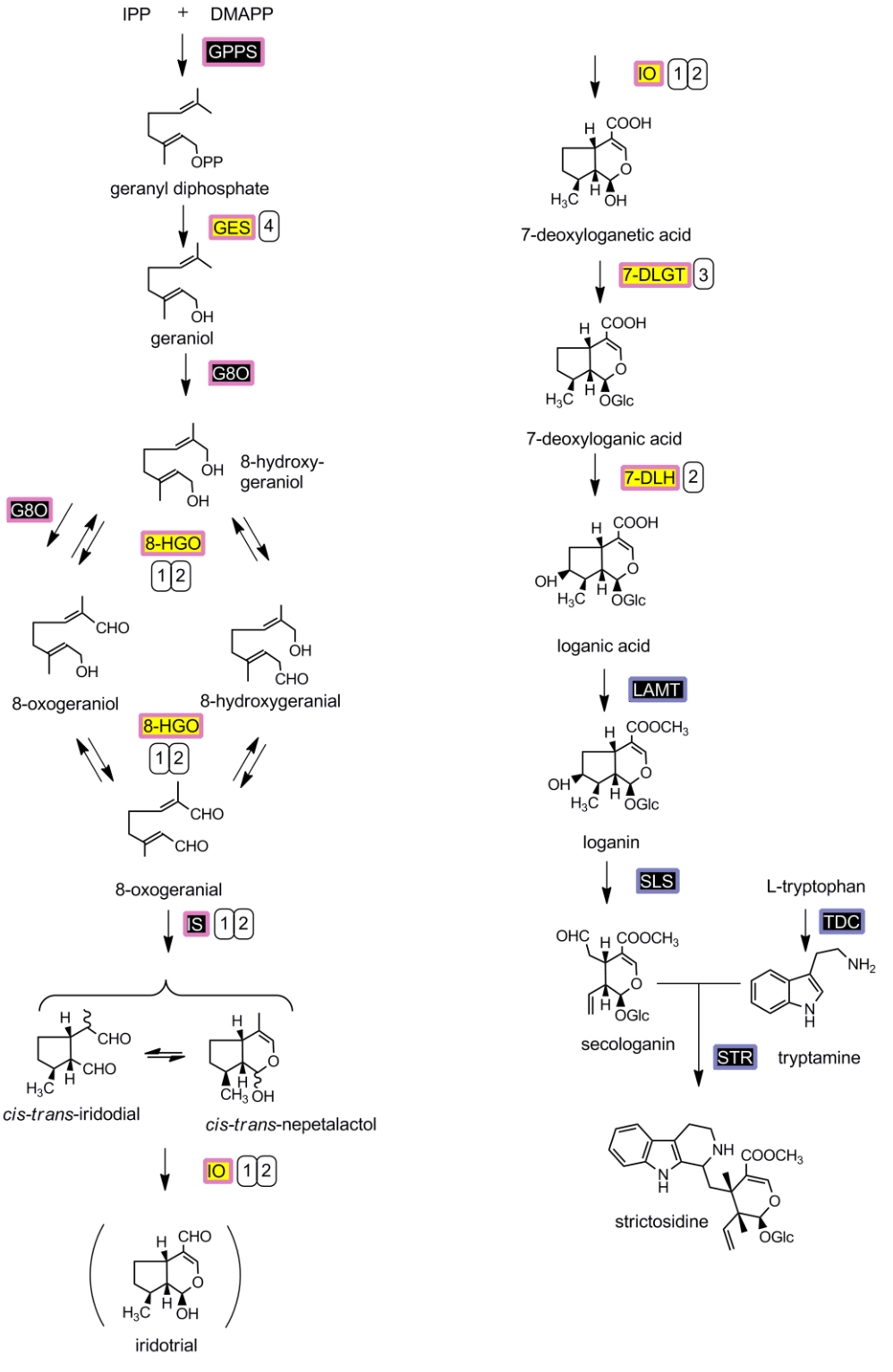
IPP + DMAPP
GPPS
OPP
geranyl diphosphate
GES 4
OH
geraniol
G8O
OH
OH
8-hydroxy-geraniol
G8O
8-HGO
1 2
CHO
OH
8-oxogeraniol
OH
CHO
8-hydroxygeranial
8-HGO
1 2
CHO
CHO
8-oxogeranial
IS 1 2
H
CHO
CHO
H3C
cis-trans-iridodial
H
O
CH3 OH
cis-trans-nepetalactol
IO 1 2
H
CHO
O
H3C
H
OH
iridotrial
IO 1 2
H
COOH
O
H3C
H
OH
7-deoxyloganetic acid
7-DLGT 3
H
COOH
O
H3C
H
OGlc
7-deoxyloganic acid
7-DLH 2
H
COOH
OH
O
H3C
H
OGlc
loganic acid
LAMT
H
COOCH3
OH
O
H3C
H
OGlc
loganin
SLS
L-tryptophan
TDC
OHC
H
COOCH3
O
H
OGlc
secologanin
NH2
N
H
STR
tryptamine
NH
COOCH3
N
H
O
OGlc
strictosidine

Fig.1 The secoiridoid pathway in C. roseus: Reactions for which the corresponding *C.roseus* gene has been described in literature have a black background, reactions for which the corresponding gene has been newly described in this thesis have a yellow background. IPP: isopentenyl diphosphate, DMAPP: dimethylallyl diphosphate, GPPS: geranyldiphosphate synthase, GES: geraniol synthase, G8O: geraniol-8-oxidase, 8-HGO: 8-hydroxygeraniol oxidoreductase, IS: iridoid synthase, IO: iridoid oxidase, 7DLGT: 7-deoxyloganetic acid glucosyltransferase, 7DLH: 7-deoxyloganic acid hydroxylase, LAMT: loganic acid methyltransferase, SLS secologanin synthase, TDC: tryptophan decarboxylase, STR: strictosidine synthase. Frames indicate mRNA localization in the leaf internal phloem-associated parenchyma (IPAP) (pink) or epidermis (blue). Numbers indicate predicted enzyme classes in the initial gene discovery strategy. 1: oxidoreductase, 2: cytochrome P450, 3:UDP-glucose-glucosyltransferase, 4: terpene synthase.

This confirmed the functionality of all enzymes of the pathway and highlights their utility for synthetic biology programs towards sustainable biotechnological production of valuable (seco)iridoids and alkaloids with pharmaceutical and agricultural applications. Notably, although different segments of the strictosidine pathway are localized in different cell types in *C. roseus*, our results show that the entire pathway can be successfully reconstituted in a single *N. benthamiana* organ. This paves the way for the biotechnological production of valuable iridoids and iridoid-derived compounds, such as the MIAs vincristine and vinblastine, making these important anti-cancer drugs available to more people and at a lower price.

As a conclusion this thesis presents the complete set of iridoid biosynthesis genes and enzymes. The tissue specific expression of the genes confirms the model proposing that the early part of iridoid synthesis occurs in the IPAP cells associated with vascular tissue and the late part in the epidermis. The nearly very similar expression patterns of all the early pathway genes reinforce this model and suggest regulation by common transcription factors. The elucidation of the pathway validates the approach taken in this work of coexpression analysis and tissuespecific proteomics for finding specific biosynthesis genes. The reconstitution of the whole pathway in *N. benthamiana* acts as proof of concept for expressing whole biosynthetic pathways in plants. Each enzyme was found to fulfill its hypothesized function and to be required to complete the pathway. The metabolite analysis from stably transformed tobacco with the GESs and *N. benthamiana* transiently transformed with GESs or the whole pathway and the first MIA biosynthesis genes can help to identify and solve bottlenecks in future metabolic engineering efforts. As formation of the monoterpenoid

moiety of MIAs (MEP pathway and iridoid pathway) is considered to be the rate limiting step in MIA biosynthesis (Morgan and Shanks, 1999), the availability of pathway genes and additional information on their characteristics can bring about new more elaborate metabolic engineering schemes for producing high amounts MIAs and iridoids in heterologous hosts.

References

Geu-Flores F, Sherden NH, Courdavault V, Burlat V, Glenn WS, Wu C, Nims E, Cui Y and O'Connor SE (2012) An alternative route to cyclic terpenes by reductive cyclization in iridoid biosynthesis. Nature **492**: 138-142

Iijima Y, Gang DR, Fridman E, Lewinsohn E, Pichersky E (2004) Characterization of geraniol synthase from the glands of sweet basil. Plant Physiol **134:** 370–379

Ito M, Honda G (2007) Geraniol synthases from *perilla* and their taxonomical significance. Phytochemistry **68:** 446–453

Morgan JA, Shanks JV (2000) Determination of metabolic rate-limitations by precursor feeding in *Catharanthus roseus* hairy root cultures. J Biotechnol **79:** 137–145

Yang T, Li J, Wang HX, Zeng Y (2005) A geraniol-synthase gene from *Cinnamomum tenuipilum.* Phytochemistry **66:** 285–293

van der Heijden R, Jacobs DI, Snoeijer W, Hallard D, Verpoorte R (2004) The *Catharanthus* alkaloids: Pharmacognosy and biotechnology. Curr Med Chem **11:** 607-628