The ORCA-ome as a key to understanding alkaloid biosynthesis in Catharanthus roseus

Hasnain, G.

Summary

Catharanthus roseus is one of the best studied medicinal plants. It produces an important class of plant secondary metabolites known as terpenoid indole alkaloids (TIAs). It is known to biosynthesize more than 130 TIA. Most of this rich molecular pool has been screened for therapeutics since several decades, and a few of them are marketed (serpentine, ajmalicine and the bisindole alkaloids vinblastine, vincristine and 3`,4`, anhydrovinblastine). Vincristine and vinblastine are effective drugs against many types of cancer (van der Heijden et al., 2004). The amounts of these compounds in the plant are very low and consequently they are difficult to isolate (Scott et al., 1979; De Luca and Laflamme, 2001). Plant tissue culture is considered as an alternative source for natural product synthesis because it is independent of environmental factors, climate changes, pests, and geographical and seasonal constraints. Various approaches involving manipulation of growth conditions have been applied to enhance the production of important alkaloids such as ajmalicine and catharanthine in *C. roseus* cell cultures. These approaches have met with limited success. Probably due to regulatory constraints alkaloid production is very low or even absent in undifferentiated cells, which clearly is an obstacle for commercial scale production (Verpoorte et al., 2000). Metabolic engineering is thought to be a solution to overcome the rate-limiting factors. Cloning of enzyme-encoding genes involved in TIA metabolism is an important prerequisite for metabolic engineering.

The classical strategy to clone genes encoding biosynthetic enzymes is to purify the target enzymes to homogeneity through several chromatographic steps from high-alkaloid-producing plant tissue or cell cultures, and then to determine N-terminal or internal peptide sequences or develop antibodies as tools for isolation of DNA sequences. Genes for many of the enzymes amenable to purification have already been cloned in this way, but the remaining enzymes in the TIA pathway may be recalcitrant to conventional purification strategies. A more recent trend is to avoid biochemical purification and to directly isolate candidate clones by combination of expression pattern analysis and homology-based screening (Hashimoto and Yamada, 2003). This homology-based cloning and functional testing approach has been successful for obtaining several new cDNAs encoding TIA biosynthetic enzymes. A variation of this approach is to look for genes with altered expression in mutant plants affected in regulation of TIA biosynthesis, however such mutants do not exist for *C. roseus*. Luckily two transcription factors called ORCA2 and ORCA3 regulating the methyl-jasmonic acid (MeJA)-responsive expression of at least half of the known genes from the TIA biosynthetic pathway in *C. roseus* have been described. It is quite likely that these transcription factors regulate other unknown genes of the TIA biosynthetic pathway. The aim of the research described in this thesis was to look into the possibility to identify and characterize novel ORCA-regulated genes from the TIA pathway using *C. roseus* cell cultures inducibly overexpressing ORCA2 or ORCA3.

ORCA2 and ORCA3 were expressed in an inducible manner because constitutive overexpression may lead to the activation of regulatory feedback systems and consequently to loss of information and may cause unspecific stress responses leading to massive shifts in the expression levels of irrelevant genes. In addition, with inducible cell lines changes in gene expression between the control and overexpression conditions are measured in the same genetic background, and inducible expression maximizes the likelihood of isolating primary target genes. In **Chapter 2** we report the comprehensive transcript profiling analysis of *C. roseus* cell lines overexpressing ORCA2

or ORCA3 in an inducible manner. By using cDNA-amplified fragment-length polymorphism (cDNA-AFLP) technology the quantitative accumulation patterns of 11,277 transcript tags were determined and analyzed. In total 1874 transcripts were differentially expressed between cell lines and treatments. Since ORCAs are positive regulators of the TIA pathway, for identification of the tags we focused on those differentially expressed genes, which were up-regulated in estradiol-treated ORCA2 and ORCA3 lines alone or in both and expressed at lower levels in estradiol-treated GFP lines and in DMSO-treated ORCA lines. This resulted in the selection of 74 transcript tags, among which 36 transcripts were upregulated both in response to overexpression of ORCA2 and ORCA3 while 22 tags and 16 tags were upregulated specifically in ORCA2 and ORCA3 overexpression lines, respectively. Upon BLAST searching of the NCBI database, 17 tags (22%) gave perfect matches with publically available *C. roseus* sequence entries. Thus, the majority of the 74 selected tags represent previously undescribed *C. roseus* sequence information. Many of these tag sequences are predicted to encode enzymes, but the assignment of these enzymes to specific metabolic steps in the TIA pathway is not straightforward. We thought that metabolic profiling of ORCA2 and ORCA3 overexpressing cell lines may be helpful to assign putative functions to the enzymes towards specific metabolic steps in the TIA pathway.

Chapter 3 describes the study of metabolite accumulation in ORCA2 and ORCA3 overexpressing cell lines. Overexpression of ORCA2 or ORCA3 alone did not affect metabolite composition, but when overexpression of ORCAs was combined with precursor feeding (loganin and tryptamine) we were able to measure metabolite accumulation and to discriminate between different samples. Precursor feeding led to accumulation of high levels of strictosidine, the central precursor for all TIAs, independent of ORCA overexpression. Accumulation of downstream TIAs was differentially affected by ORCA2 and ORCA3 overexpression. TIA accumulation patterns allowed us to predict in which set of cDNA-AFLP tags genes encoding the corresponding enzymes should be present. The accumulation of catharanthine was observed only in ORCA3 overexpressing cell lines indicating that gene(s) corresponding to enzyme(s) involved in the catharanthine biosynthetic pathway must be present in the data set of ORCA3-specific tags. In contrast, ajmalicine concentration was increased when ORCA2 or ORCA3 were overexpressed, hence the gene involved in ajmalicine biosynthesis must be present in the tag set which was regulated by both ORCA2 and ORCA3. The literature describes the involvement of an NADPH-dependent oxidoreductase in the conversion of cathenamine to ajmalicine (Stoeckigt et al., 1983). The dataset with the genes up-regulated by both ORCA2 and ORCA3 contains the tag CR-75 which upon TBLASTX searching showed homology to NADPH-dependent oxidoreductase enzymes.

To prove the hypothesis that the enzyme corresponding to tag CR-75 functions as cathenamine reductase (CR) we isolated the full-length CR-75 cDNA and its characterization is reported in **Chapter 4**. CR-75 transcripts were found to accumulate in response to overexpression of either ORCA2 or ORCA3 and in response to MeJA or yeast elicitor. All of the known TIA pathway genes are responsive to MeJA (Chapter 2), and therefore the expression pattern of the CR-75 transcript was consistent with a role of the encoded enzyme in TIA biosynthesis. We wanted to provide biochemical proof that CR-75 is working as cathenamine reductase. To this end the CR-75

protein was produced in *E.coli* with a strep-tag at the N-terminal end and a His tag at the C terminus. The substrate cathenamine is highly unstable and therefore cannot be obtained in pure form since it exists in an equilibrium with 4,21-dehydrogeissoschizine. To obtain a cathenamine preparation we coupled tryptamine and secologanin enzymatically using strictosidine synthase. The formed strictosidine was then incubated with crude protein extract from a *C. roseus* cell culture, which contains a high level of strictosidine β-D-glucosidase enzyme activity. Products of this reaction were used as a source of substrate for the CR enzyme assay. We observed the formation of ajmalicine in an NADPH-dependent manner, proving that indeed CR-75 is working as cathenamine reductase.

Chapter 5 reports the isolation and characterization of alcohol dehydrogenase genes from *C. roseus* which may function as 10-hydroxygeraniol oxidoreductases (10HGO). During our genome-wide transcriptome analysis of *C .roseus* cells inducibly overexpressing ORCA2 or ORCA3 we found a cDNA-AFLP tag (CR-36) which was up-regulated by both ORCA2 and ORCA3. Homology search with TBLASTX showed that CR-36 has 64 % identity to the cDNA of a putative Cr10HGO (Accession no AY352047) at the amino acid level. The same tag (called CRG258 in that study) was also isolated from another cDNA-AFLP study of *C. roseus* cells. It was found to be upregulated in response to MeJA and the expression profile was similar to the known TIA pathway genes (Rischer et al., 2006). The regulation of CR-36 by ORCA2 and ORCA3, the responsiveness of a same tag (CRG258) to MeJA and its similarity to a cDNA annotated as encoding a putative 10HGO led us to hypothesize that the CR-36 protein has 10HGO activity. During our search for the full-length CR-36 cDNA we accidently isolated another full-length cDNA (CR-36-1) which showed 96 % nt identity with the CR-36 tag. After isolation of the full-length cDNA corresponding to the CR-36 tag we had three closely related genes in hand, the putative 10HGO, CR-36-1 and CR-36. Interestingly only CR-36 transcripts were regulated by ORCA2 and ORCA3 and were also found to accumulate in response to MeJA. For the conversion of 10-hydroxygeraniol to 10-oxogeranial involvement of a alcohol dehydrogenase enzyme has been described in *R. serpentina* cells (Ikeda et al., 1991). A similar enzymatic step is also expected in *C. roseus.* The proteins encoded by CR-36-1 and CR-36, and the putative 10HGO belong to the alcohol dehydrogenase family and all the catalytically important amino acid residues were conserved in all three proteins. All three recombinant proteins were able to accept 10-hydroxygeraniol as a substrate to reduce $NADP^+$. Determination of the precise substrate specificities of these enzymes and of the identities of the formed products requires further studies.

In **Chapter 6,** characterization of two full-length cDNA clones encoding putative TIA pathway enzymes is described. One TIA pathway candidate gene is represented by tag CR-93, which was up-regulated by both ORCA2 and ORCA3. Upon TBLASTX search CR-93 showed close homology to the TIA pathway gene polyneuridine aldehyde esterase (PNAE). Polyneuridine aldehyde esterase catalyses the conversion of polyneuridine aldehyde to epi-vellosimine during the formation of ajmaline and sarpagine in the medicinal plant *R. serpentina* (Dogru et al., 2000). Ajmaline /sarpagine type alkaloids are not synthesized in *C. roseus* but a similar enzymatic reaction may be possible in the biosynthesis of other TIAs. cDNA library screening resulted in a 1128 bp sequence having 100 % nt identity with the CR-93 cDNA-AFLP tag. Northern blot analysis confirmed

the regulation of CR-93 by both ORCA2 and ORCA3 and also showed that the expression of CR-93 was responsive to MeJA. We checked the possibilities that recombinant CR-93 protein may function as a MeJA-cleaving esterase or as a geranyl acetate esterase but did not find enzymatic activity of CR-93 towards MeJA or geranyl acetate. It is possible that CR-93 protein has a specific function in TIA biosynthesis similar to its close homolog, PNAE from *R. serpentina,* in ajmaline biosynthesis.

The transcript tag CR-65 was up-regulated specifically by ORCA2 alone. Upon TBLASTX searching of the DNA database the encoded protein was found to be related to the carboxylesterase gene superfamily. Metabolite profiling of ORCA2 and ORCA3 overexpressing lines resulted in the identification of an ORCA3-specific sub-pathway, but we did not identify an ORCA2-specific sub-pathway. The characterization of an enzyme-encoding gene specifically regulated by ORCA2 indicated the existence of an ORCA2-regulated TIA sub-pathway. Therefore we isolated the full-length cDNA for CR-65. Northern blot analysis validated the CR-65 expression profile observed in cDNA-AFLP analysis. Similarity search in the NCBI database revealed that the protein encoded by the CR-65 cDNA belongs to the carboxylesterase protein superfamily with the highest level of similarity with 2-hydroxyisoflavanone dehydratase from *Glycyrrhiza echinata.*

In conclusion, the strategy for discovery of new genes encoding TIA biosynthesis enzymes using the ORCA transcription factors as tools was successful. We have demonstrated that two of the isolated genes encode enzymes involved in TIA biosynthesis. It is very well possible that the large majority of the missing enzymes is represented by the tags in our collection. In fact, in assigning functions to the orphan enzymes in our collection, the main limiting factor is the lack of classical biochemical information about metabolic intermediates and possible reaction mechanisms in certain parts of TIA metabolism. Another bottleneck in further studies is the fact that many TIA intermediates are not commercially available and are difficult to make via organic synthesis.

References

De Luca V, Laflamme P (2001) The expanding universe of alkaloid biosynthesis. Curr. Opin. Plant Biol. **4:** 225-233

Dogru E, Warzecha H, Seibel F, Haebel S, Lottspeich F, Stöckigt J (2000) The gene encoding polyneuridine aldehyde esterase of monoterpenoid indole alkaloid biosynthesis in plants is an ortholog of the α/β hydrolase super family. Eur. J. Biochem. **267:** 1397-1406

Hashimoto T, Yamada Y (2003) New genes in alkaloid metabolism and transport. Curr. Opin. Biotech. **14:** 163-168

Ikeda H, Esaki N, Nakai S, Hashimoto K, Uesato S, Soda K, Fujita T (1991) Acyclic Monoterpene Primary Alcohol:$NADP^+$ oxidoreductase of *Rauwolfia serpentina* cells: The key enzyme in biosynthesis of monoterpene alcohols. J. Biochem. **109:** 341-347

Rischer H, Oresic M, Seppanen-Laakso T, Katajamaa M, Lammertyn F, Ardiles-Diaz W, Van Montagu MCE, Inze D, Oksman-Caldentey KM, Goossens A (2006) Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* cells. Proc. Natl. Acad. Sci. USA **103:** 5614-5619

Scott AI, Mizukami H, Lee SL (1979) Characterization of a 5-methyltryptophan resistant strain of *Catharanthus roseus* cultured cells. Phytochemistry **18:** 795-798

Stoeckigt J, Hemscheidt T, Hoefle G, Heinstein P, Formacek V (1983) Steric course of hydrogen

transfer during enzymatic formation of 3.alpha.-heteroyohimbine alkaloids. Biochemistry **22:** 3448-3452

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

Verpoorte R, van der Heijden R, Memelink J (2000) Engineering the plant cell factory for secondary metabolite production. Transgenic Res. **9:** 323-343