Bioconversion and combinatorial biosynthesis of selected terpenoids and lignans

Julsing, Mattijs Kamiel

Summary

In this thesis several opportunities for the bioconversion and combinatorial biosynthesis of plant secondary metabolites are described. The focus is on two specific groups of natural products, terpenoids and lignans. The use of a host organism for the bioconversion of natural products is known already for a long time. Feeding of cells or cell derived extracts with the putative substrate is often used for the production of metabolites, but can be applied for identification and characterization of genes and corresponding enzymes as well by cloning and expression of a gene in a microorganism. Combinatorial biosynthesis is a novel strategy in the research of natural products, that can be defined as the approach to combine genes from different organisms for the production of natural products. Since multiple gene expression systems have recently been developed, complete biosynthetic pathways (multi-step reactions) can in principle be reconstructed in microorganisms. Next to that also new gene combinations, which would probably never occur in nature by evolution and selection, can be constructed. This offers new possibilities to develop (leads for) pharmaceutically interesting compounds. This thesis aims to investigate this on three possible levels: first the use of an endogenous bacterial biosynthetic pathway, second the investigation of the expression of a plant gene in *Escherichia coli* and the mechanism of action of the gene product (enzyme), and third the bioconversion of plant compounds using non-related recombinant enzymes.

Chapter 2 gives an elaborate overview of the literature that has appeared so far on the use of combinatorial biosynthesis strategies in relation to the production and development of secondary metabolites from medicinal plants. Some natural product groups, like alkaloids, terpenoids, and flavonoids, are already under investigation for several years and lots of progress has been made since the start of this research. The major focuses of this review are fundamentals of genetic work, used expression systems and latest progress in the field of combinatorial biosynthesis.

The heterologous production of plant secondary metabolites implies the presence or possibility for the reconstruction of biosynthetic pathway in the host cell. The presence of an endogenous pathway in the host organism can be a good reason to choose an organism for the heterologous production of a plant compound. Since terpenoids are formed by most of the living species, the first common steps of the biosynthetic pathway are present in all organisms. Therefore, the biosynthesis of precursors can be performed in the host cell and specific plant genes can be introduced to hook on with the available pathway. The emission of relative high levels of isoprene by *Bacillus subtilis* suggested the high onset of the biosynthesis of the general building block for all terpenoids, isopentenyl diphosphate. Based on homology with known enzymes of the deoxyxylulose phosphate pathway to isoprene, putative enzymes and corresponding genes for the genome of *B. subtilis* were depicted. By measuring the isoprene emission of knock-out mutants for the genes putatively involved in isoprene biosynthesis, we identified most of the genes of the biosynthetic pathway to isoprene in *B. subtilis*. This work is described in **Chapter 3**. The results make it possible in the near future to investigate the production of isoprenoids in more detail and to optimize *B. subtilis* as a host for the production of terpenoids.

The first specific step in the biosynthesis of artemisinin is the formation of the sesquiterpenoid skeleton amorpha-4,11-diene out of the general 15 carbon precursor

farnesyl diphosphate. This bioconversion is catalysed by the enzyme amorphadiene synthase. **Chapter 4** describes an attempt to gain more information about the mechanism of action of this enzyme. Therefore, the gene encoding amorphadiene synthase was cloned into a pET-vector. Based on a primary alignment and a computational model comparing amorphadiene synthase with 5-epi-aristolochene synthase from *Nicotiana tabacum*, three amino acids of the enzyme (F514, Q518, F525) were assumed to be possibly involved in the specific 1,6-cyclization of farnesyl diphosphate in amorpha-4,11-diene, whereas 5-epi-aristolochene is a product from a 2,7-ring closure. Theoretically this could explain the differences in the formation of the products. Mutant proteins were obtained by substitutions of the depicted amino acids. The activity of the purified mutant proteins produced in *E. coli* was investigated and products were identified by GC-MS. Alanine substitutions Q518A, F514AQ518A, and F525A caused inactivation of the protein resulting in no detectable levels of amorpha-4,11-diene, whereas the alanine substitution F514A resulted in a active, but less effective, protein. Substitutions towards the corresponding residue from 5-epi-aristolochene synthase did not result in a complete loss of the activity for the production of amorpha-4,11-diene, but they did not result in the production of 5-epi-atistolochene or other sesquiterpenoids either, as might be expected. Since there were only small differences observed in the active site residues of the two investigated sesquiterpene synthases, it was concluded that variations in the product specificity are caused by several subtle differences in the active site. The elucidation of the crystal structure of amorphadiene synthase may reveal these differences.

Plant cytochrome p450 enzymes (CYP) are often involved in the derivatization of the terpenoid skeletons synthesized by terpene synthases. Several plant cytochrome P450 genes have already been isolated and identified, but the systematic, complete search for genes encoding these specific enzymes is a laborious job. **Chapter 5** describes a study to explore the use of human cytochrome P450 produced in *E. coli* for the metabolization of several selected mono- and sesquiterpenoids. The sesquiterpenoid cedrol was selected as the model compound. Cedrol was converted by CYP3A4 in two metabolites, which were isolated and identified as 2-hydroxycedrol and 4-hydroxycedrol respectively. The fact that several other selected mono- and sesquiterpenoids were metabolized by CYP3A4 emphasizes the broad substrate specificity of the enzyme. Two other investigated cytochrome P450 enzymes, CYP1A2 and CYP2C9, did not show any activity towards the tested terpenoids. It was concluded that recombinant human cytochromes can be useful tools in the production of new natural products and for *in vitro* metabolisation studies.

The metabolization of lignans by recombinant human cytochrome P450 monooxygenases was investigated as well. Again, CYP3A4 was the only enzyme showing activity to the substrates offered. The lignan podophyllotoxin is the natural source of the semi-synthetic anticancer drugs teniposide and etoposide. In **Chapter 6** the bioconversion of deoxypodophyllotoxin by *E. coli* cells containing recombinant CYP3A4 is described. Deoxypodophyllotoxin is the natural precursor of podophyllotoxin and is more easily available from natural sources. Deoxypodophyllotoxin was metabolized for more than 90 % and only one metabolite was observed, which was identified as epipodophyllotoxin. These results show the possibility to use a human cytochrome P450 to mimic the bioconversion in

a plant cell. We conclude that deoxypodophyllotoxin can be used in principle as an alternative source for the semi-synthesis of the cytostatic drugs etoposide and teniposide.

In **Chapter 7** the bioconversion of deoxypodophyllotoxin by CYP3A4 was further investigated. To explain the experimental findings in relation to the conformation of the (active site of the) enzyme and the chemical structure of the lignans used, we subjected the lignan structures to a computational model together with the known crystal structure of CYP3A4 using the Distance Constrained Docking algorithm. Modeling of deoxypodophyllotoxin into CYP3A4 clearly demonstrated the stereoselectivity of this bioconversion for the formation of epipodophyllotoxin instead of podophyllotoxin. Next to that, we applied the modeling procedure to find a possible explanation for the enzyme inhibition caused by the methylenedioxy moiety of lignan structures. For lignans containing a methylenedioxy moiety inhibition of CYP3A4 has been described before. Therefore, inhibition of CYP3A4 by deoxypodophyllotoxin or the bioconversion product epipodophyllotoxin was feasible and should make the bioconversion of deoxypodophyllotoxin to produce epipodophyllotoxin more complex to apply. In order to investigate the influence of the methylenedioxy group on the bioconversion process, the effects of five lignans with and without the methylenedioxy group were compared. The results confirmed the inhibiting effect of the methylenedioxy group of lignans for CYP3A4.

In conclusion, it can be stated that bioconversion and strategies of combinatorial biosynthesis are strong tools in natural product research, both for the elucidation of biosynthetic pathways as for the heterologous production and derivatization of plant secondary metabolites. The knowledge about the regulation of biosynthetic pathways and the mechanisms of the enzymes involved can certainly improve the efficiency of the production of metabolites. This could make heterologous production systems efficient alternatives for the production of pharmaceutical lead compounds in the near future.