I. Introduction

I.1 Plant secondary metabolism

Plants are capable of synthesizing an overwhelming variety of small organic molecules, called secondary metabolites, usually with very complex and unique carbon skeleton structures and many of them of high interest to the pharmaceutical and chemical industry. For instance, plant derived pharmaceuticals represent a large market value; about 25% of today’s pharmaceuticals contain at least one active ingredient of plant origin (Rischer et al., 2006). By definition, secondary metabolites are not essential for the growth and development of a plant, but rather are required for the interaction of plants with their environment (Kutchan and Dixon, 2005). Many secondary compounds have signalling functions; among them are plant hormones which will be discussed later in more details. They influence the activities of other cells, control their metabolic activities and co-ordinate the development of the whole plant. Other substances flower colours serve to communicate with pollinators (Atsatt, 1996) or protect the plants from feeding by animals or infections by, producing specific phytoalexines after fungi infections that inhibit the spreading of the fungi mycelia within the plant (Atsatt, 1996). Plant secondary metabolites have been a fertile area of chemical investigation for many years, driving the development of both analytical chemistry and of new synthetic reactions and methodologies. The subject is multi-disciplinary with chemists, biochemists and plant scientists all contributing to our current understanding. One overall goal is to genetically characterise the molecular mechanisms driving secondary metabolite biosynthesis in plant cells. Identifying key elements involved in these processes will allow generating novel tools for metabolic engineering of plant cells. Even so, only a little fraction of the enormous biosynthetic potential of plant cells is being exploited. Metabolic engineering of plant cells so far has added little to the problem, since insight into the molecular mechanisms driving plant secondary metabolism is still very limited. Genetic maps of biosynthetic pathways are still far from complete and the regulation of these pathways is hardly understood. However, technological advances in analytical chemistry, in particular in the development of high-field nuclear magnetic resonance spectroscopy and Fourier
transform-ion cyclotron mass spectrometry, have facilitated the elucidation of structures of secondary metabolites that are present even at low levels within a plant.

I.2 Modern methods of gene discovery

Genetic approaches have been successful in unraveling the metabolic network for phenylpropanoid and flavonoid metabolism, as well as in finding transcriptional regulators of structural genes in the genetically tractable plant species maize, petunia, snapdragon and, more recently, Arabidopsis (Davies et al., 2003). The revolution in views on phenylpropanoid metabolism demonstrates the power of genetics and shows that long-term and worldwide research efforts that employed biochemical approaches have failed to correctly describe the basic network of intermediates and enzyme activities in this pathway. Enormous advances have also been made in understanding the basic network of glucosinolate biosynthesis, and its relationship to auxin homeostasis, in Arabidopsis (Memelink, 2005). A powerful tool in phenotype-driven genetics is gain-of-function mutagenesis with a strong constitutive promoter that is carried on an insertion element such as Agrobacterium tumefaciens T-DNA. T-DNA activation tagging causes dominant gain-of-function mutations that are caused by the over-expression of the genes that flank the T-DNA insertion (Memelink, 2005). It allows the recovery of mutant phenotypes that are conferred by functionally redundant genes. In addition, the T-DNA tag allows gene recovery from the primary transformant without a need for genetic crosses. T-DNA activation tagging has been successfully used in the model species Arabidopsis to identify the transcriptional regulator PAP1 (Production of anthocyanin pigment1) (Borevitz et al., 2000). Over-expression of this MYB-type protein up regulates multiple genes that are involved in anthocyanin biosynthesis and gives the transformant a purple color. In non-model plant species, the use of genetic approaches (with the exception of some reverse-genetic RNAi approaches) appears to have come to an almost complete standstill. For example in opium poppy, codeinone reductase (COR), a terminal enzyme in morphine biosynthesis, has been knocked out by RNAi using a chimerical inverted repeat that targets all seven members of the gene family, leading to a significant reduction in the morphine and codeine levels of the transgenic poppy latex (Allen et al.,
Introduction

Surprisingly, none of the morphine-type precursors that are normally present in the latex were detectable. Instead, the transgenic latex accumulated rare alkaloids. These included the upstream precursor reticuline, which is located seven enzymatic steps before the COR-mediated reaction in the biosynthesis pathway, and several methylated derivatives of reticuline. The reasons for the apparent shutdown of the complete morphine-specific branch, resulting in a drastic switch in the alkaloid pattern, are not clear. It has been demonstrated that some genetic approaches, such as T-DNA activation tagging, can be successful with non-model species but this has not led to the expected explosion of similar approaches in a variety of plant species. The picture is far from complete, however, and how the expression and the activities of these transcription factors are regulated remains largely mysterious (Von Endt et al., 2002).

I.3 Single-cell genomics

Particular emphasis is placed on genomic-based single-cell approaches for an improved understanding of the biosynthesis of specialized metabolites. Alkaloids, such as morphine and codeine, and rubber polymers are found in exudates derived from laticifer cells (Mahlberg, 1993), lignin-like polymers are synthesized in vascular tissue cells (Ye, 2002), UV-absorbing phenylpropanoids are present in epidermal cell layers (Harborn and Williams, 1992), and a variety of metabolite classes that function in the defense against insects and pathogens are localized to secretory trichomes (Wager et al., 2004). Protocols for the isolation of specialized cells usually allow the cell type of interest to be pooled, which increases the amounts of analyses available and thus enables direct interfacing with various post-genomic technologies (Markus, 2005). However, care must be taken to ensure that the collection of a specialized cell type does not bias subsequent analyses. Direct access to the cellular contents of individual cells is facilitated by Pico liter-scale sampling using micro capillaries. Isolation of specialized cell types Glandular secreting trichomes are more easily accessed by mechanical means than are cells or components of cells that are embedded within plant tissues (Heath et al., 2003). Biosynthetically competent secretory cells of peppermint (Mentha x piperita) and basil (Ocimum basilicum), which are responsible for the biosynthesis and secretion of the signature
essential oils of these plants, have been isolated by surface abrasion with glass beads in a complex medium. cDNA libraries were then generated from these cells and randomly selected clones (i.e. expressed sequence tags [ESTs]) (Lange et al., 2000). Laser-capture microdissection (LCM) is a highly effective tool for the isolation of target cells from heterogeneous tissues that has recently been adapted for collecting material from plant sections (Kehr, 2003). The detection of the most abundant proteins and metabolites at the single-cell level has been achieved using a variety of platforms (Tomos et al., 2001), but the potential for multiplexing is limited. Thus, an important challenge in overcoming these current difficulties is the development of more sensitive assays, possibly by advancing the use of nanotechnology in the analysis of biomolecules. Bioanalytical nanotechnology microarrays are used widely to profile thousands of transcripts in parallel and, when interfaced with prior RNA amplification techniques, are powerful tools for single-cell analysis. Microarray analyses might not, however, detect the expression of low abundance transcripts (Czechowski et al., 2004). An alternative transcript-profiling technology, termed massively parallel signature sequencing (MPSS), involves the cloning of DNA molecules onto microbeads and reveals the expression levels of virtually all of the genes that are expressed in a sample by counting the number of individual mRNA molecules produced from each gene. One intriguing possibility is the use of cell-type-specific promoters, which might enable the modulation of essential oil composition in glandular trichomes (Mahmoud et al., 2004) or alkaloid production in laticifers (Frick et al., 2004) by genetic engineering, without causing adverse effects attributable to the ectopic expression of trans-genes.

I.4 Enzymes of plant secondary metabolism

I.4.1 Substrates and products

The ‘classical’ approach to discover enzymes is to start with a given product and to ask what enzymatic reaction is responsible for its formation, and what the substrate is. Substrates can be hypothesized on the basis of biochemical principles and current knowledge of metabolic pathways and types of enzymes. In other words, identification of the enzymatic properties of a newly discovered protein still depends on prior biochemical
knowledge relating to the family of enzymes to which the candidate protein belongs. Thus, many new methyltransferases, terpene synthases, acyl transferases, and glucosyltransferases have been discovered because we already know representative enzymes of these types and can therefore make an informed guess regarding the potential substrates with which to test candidate proteins. New methods of inactivating gene expression, such as the insertion of T-DNAs and transposable elements and RNA interference (RNAi) techniques, could be extremely powerful in combination with metabolic profiling techniques. However, gene identification by such techniques must also be followed by biochemical characterization of the protein (Simkin et al., 2004) because the lack of final product in the mutant does not guarantee that the substrate has been correctly identified. Gene suppression techniques might also lead to misleading results if several similar genes that encode enzymes with different substrates are suppressed by the same construct. In summary, the combination of the new techniques of metabolic and gene expression profiling with classical techniques of enzymatic analysis will allow the identification of the function of the majority of the genes in plant genomes.

The challenge is to develop standardized and automatable methods for enzyme analysis.

I.4.2 Evolution and function

The chemical diversity of plants is the result of ongoing evolutionary processes. Recent advances in the molecular biology of plants, particularly in the area of large-scale genomics (Borvitz and Ecker, 2004), are revealing how enzymes of natural product biosynthesis arise through mutation and gene duplication, leading to the continued elaboration of new chemical structures that will be selected for if the impart an adaptive advantage on the plant (Pichersky and Gang, 2000). Structural biology provides an important tool set for the detailed structure–function characterization of proteins at the atomic level (Kim et al., 2003; Eisenberg, 2005). The level of functional understanding derived from experimentally determined structures or from realistic models that are constructed from homologous protein folds (Goldsmith-Fischman and Honig, 2003) can lead to a more complete appreciation of complex biosynthetic pathways. Such information can elucidate the mechanisms of individual biosynthetic reactions (Naismith,
and provides new views at atomic resolution of the temporal and spatial architecture of multi-protein complexes that are vital to metabolic flux and channeling, and this results in a practical and rational basis for engineering useful metabolic traits (Whittle and Shanklin, 200; Singh et al., 2005) into medicinally and agriculturally important plants. The fact that Arabidopsis has more than one hundred small molecule glucosyltransferases (GSTs) (Bowles et al., 2004) suggests that these enzymes to be relatively promiscuous, at least in vitro, presents a challenge in relating in vitro biochemistry to in vivo function and poses fascinating questions as to why such complex gene families have evolved (Kliebenstein et al., 2005). Studies are currently underway to generate transgenic A.thaliana lines that specifically vary in each of these steps to verify their epistatic and phenotypic effect on both glucosinolate structure and herbivores.

I.5 Multiple levels of regulation of plant secondary pathways

Secondary metabolism is an integral part of the developmental program of plants, and accumulation of secondary metabolites often marks the onset of developmental stages. Although this association between plant differentiation and secondary metabolism has long been known, a picture of the molecular mechanisms that connect these two programs is starting to emerge, particularly in the latest findings related to the regulation of flavonoid biosynthesis in Arabidopsis. It emphasizes the role played by a common set of transcription factors that control both this pathway and specific aspects of cellular differentiation, and discusses the importance of WD40 proteins in coordinating flavonoid regulation in Arabidopsis and other plant species (Broun, 2005). In this respect we have much to learn about the transcription factors, and we need to know more about the post-translational events that control their activity, and their sub-cellular localization and turnover.

I.5.1 Metabolon formation and metabolic channeling in secondary metabolism

Classically, biological membranes such as the endoplasmic reticulum (ER) have been considered as homogenous fluid structures that are composed of lipid bilayers,
which serve as a two-dimensional solvent phase for fully or partly embedded membrane proteins (Singer and Nicolson, 1972). The organization of cooperating enzymes into macromolecular complexes is central feature of cellular metabolism. A major advantage of such spatial organization is the transfer of biosynthetic intermediates between catalytic sites without diffusion into the bulk phase of the cell. This so-called “metabolic channeling” offers unique opportunities for enhancing and regulating cellular biochemistry (Winkel, 2004). The formation of plant secondary metabolites is complex and dynamic process that involves multiple sub-cellular compartments such as the cytosol, endoplasmic reticulum, and the vacuole. The metabolic activities of a plant are highly coordinated at the whole-plant, organ, tissue, cellular, organelle and molecular levels. At the cellular level, channeling of substrates to their target enzymes is facilitated by the compartmentalization of the cell into different organelles and sub-structures. This serves to co-localize and optimize the concentrations of enzymes and their substrates. A limited number of key genes encode the enzymes that are responsible for the synthesis of the pivotal backbone structures that constitute the hallmarks of the different classes of natural products, and progress has been made in the identification of these genes (Kuchan 1995; Bak et al., 2003). The subsequent decoration of the backbone structures generates the huge diversity of plant secondary products. The large majority of these decoration processes are mediated by a limited number of enzyme classes, such as glycosyl-, methyl- and acyltransferases, which are all encoded by multi gene families. The positioning of enzymes that have broad substrate specificity downstream of the conserved early pivotal enzymes of plant secondary metabolism opens the possibility of producing new secondary compounds without major re-structuring of the enzyme complement. Metabolic channeling and metabolon formation provide the key to resolving and avoiding potential negative interference in plant natural product formation either by narrowing substrate specificity as a result of conformational changes upon binding or because binding into the metabolon prevents access of unwanted substrates (Jørgensen et al., 2005). A single glycosyl-, methyl- or acyltransferase might possess the ability to bind to different metabolons. In this manner, the possibility of combinatorial defined substrate specificity might explain how the desired substrate specificity is achieved with a minimum number of enzymes (Jørgensen et al., 2005). The latest findings stress the
importance of specific types of transport molecules, and genes encoding various types of multi drug resistance, and ABC transporter proteins. ABC transporters constitute a large protein family that is found in a range of organisms from bacteria to humans. Because of intensive studies on the roles of ABC transporters in multi drug resistance in animal cancer cells, it had long been believed that they exhibit broad substrate specificity (Yazaki 2005). Recent studies have demonstrated that the function of ABC transporters is not restricted to detoxification processes (Martinoa at all). Furthermore, they have been found to be involved in many specific biological activities, such as cell signaling, that have strict substrate specificity, as well as in other divergent physiological functions (Klein et al., 2003; Pighin et al., 2004). It was suggested recently that the ABC transporter family might have evolved according to the need to transport specific substrates in organism, and not as drug efflux pumps (Sheps et al., 2004).

I.5.2 Morphine biosynthesis regulation in *Papaver somniferum*

I.5.2.1 Benzylisoquinoline alkaloids (BAIs) and morphine biosynthesis

Benzylisoquinoline alkaloids are large and diverse groups of natural product containing more than 2500 defined structures found mainly in five plant families, including the Papaveracea (Facchini, 2001). BAIs are all based on the elaboration of a simple skeletal structure. Skeletal representation of known stereochemical conformation and functional group substitutions found in major BAI families (D.K. Liscombe et al., 2005). Opium Poppy (*Papaver somniferum* L.) produces a large number of Benzylisoquinoline alkaloids including morphine and sanguinarine, derived from tyrosine via the Branch -point intermediate (S)-reticuline. Berberine, the first isoquinone alkaloid from an amino acid to have its biosynthetic pathway completely elucidated, which is derived from L-tyrosine via 13 different enzymatic reactions (Sato et al., 2001) Codeine and other morphinan-type alkaloids is centrally involved in the synthesis of many other isoquinoline alkaloid of divergent chemical structure in various plant species (Kutchan, 1996). It is presumed; therefore that morphine biosynthesis proceeds via a common pathway up to (S)-reticuline and that the biosynthetic reactions after this stage constitute the morphine-specific pathway in opium poppy. From morphine-specific
biosynthesis pathway after reticuline, cDNAs of two important genes were cloned i.e. codeinone reductase (Underline et al., 1999) and salutaridinol 7-O-acetyltransferase (SalAT) (Grothe et al., 2001). The cell cultures of opium poppy contained these enzymatic activities, although they did not accumulate morphine itself. From the cultured cells, these enzymes were purified to homogeneity and the internal peptide sequences were used to clone their cDNAs taken from a library. The discovery of (salAT) activity catalyzing the acetylation of salutaridinol is an important step. It was not clarified whether the formation of thebaine, (Figure I.1) the first morphinan alkaloid having the pentacyclic ring system is a protein-dependant or it is a non-enzymatic in vivo reaction at natural conditions of laticifers in poppy plants.

![Proposed reaction sequence](image)

**Figure I.1** Proposed reaction Sequence leading from Salutaridinol via Salutaridinol-7-O-acetate to thebaine (pH 8-9)
The pH value of this novel mechanism in alkaloid biosynthesis is critical for the formation of the phenolate anion that initiates the $S_N^2$ reaction supported by the favored acetate leaving group. In contrast, if the homogeneous Salutaridinol-7-O-acetyltransferase was incubated in the presence of its substrates at pH 7 (Figure I.3), hardly any thebaine was formed but rather a completely different reaction product (dibenz [$d,f]$azonine alkaloid) (Lenz and Zenk, 1995).

![Figure I.2](image)

**Figure I.2** Proposed reaction Sequence leading from Salutaridinol via Salutaridinol-7-O-acetate to [8,9-dihydro-5H-2,12-dimethoxy-1-hydroxy-7-methyl-dibenz[$d,f$]azoninium]acetate (pH 6-7), which was converted chemically to neodihydrothebaine by NaBH$_4$ reduction.

### I.5.2.2 Cell-specific localization of morphine biosynthesis key enzymes.

The past decade has provided a lot of new molecular information on the biosynthesis of some of the most complex alkaloids, including the description of all the individual steps in the conversion of tyrosine into berberine and the elucidation of most of the steps in morphine biosynthesis (Qunaroon *et al.*, 2003). The alkaloid (S)-reticuline constitutes an important branch point in alkaloid synthesis because several different subclasses of isoquinoline alkaloids can be formed from this compound, depending on the subsequent regio- or stereo specific transformations that take place (Qunaroon *et al.*, 2003). The cell-specific localization of five key enzymes in the metabolic grid has been
determined in the capsule, stem, and root tissues (Figure I.4). It appears that the early stages of morphine biosynthesis, starting with the decarboxylation of amino acid L-tyrosine occur in the parenchyma cells surrounding laticifers. In contrast, the later stages of morphine biosynthesis occur in the laticifer, which is the storage site of morphine alkaloids thebaine, codeine and morphine. (Kutchan, 2005). This outstanding example of channeling at the cellular level is provided by the different spatial organizations of the key enzymes in the metabolic enzyme grid that results in the synthesis of different morphinans suggest many questions about the involvement of diverse organelles in plant natural product biosynthesis in addition to the simple diffusion transport. Furthermore the interaction of macromolecular biosynthetic enzyme complexes with metabolite transporters might also be a consideration (Kutchan, 2005).
Figure 1.3 Schematic of the biosynthetic pathway leading to tetrahydrobenzylisoquinoline-derived alkaloids in opium poppy. In the capsule and stem, 40OMT, 7OMT and SalAT are found predominantly in phloem parenchyma cells and codeinone reductase is localized to laticifers. These are the site of morphinan alkaloid accumulation, as determined by fluorescence immunocytochemical localization. The berberine bridge enzyme has been localized to vesicles in idioblasts of young shoot and parenchyma cells of the root cortex (by fluorescence immunocytochemical labeling; image reproduced. Major latex proteins are represented by red arrowheads, biosynthetic enzymes by green arrowheads and the co-localization of major latex proteins and biosynthetic enzymes by yellow arrowheads. laticifer; MLP 15, major latex protein 15; Xy, xylem.
I.6 yeast two-hybrid system and Protein-protein interactions detection

All biological processes depend on interactions formed between proteins and the mapping of such interactions on a global scale is providing interesting functional insights. One of the techniques that have proved itself invaluable in the mapping of protein-protein interactions is the yeast two-hybrid system. The yeast two-hybrid screen is a powerful artificial transcription-based assay which provides a rapid and straightforward mechanism to identify proteins that interact in vivo in a yeast model system. A significant issue in the use of two-hybrid system is the degree to which interacting proteins distinguish their biological partner from evolutionarily conserved related proteins and the degree to which observed interactions are specific (Causier and Davies, 2002). A number of variants of the two-hybrid system have been developed for library screening one of them is the Cyto Trap tow-hybrid system (Figure I.4).

Figure I.4 Cyto Trap two-hybrid system: CytoTrap two-hybrid system restores the Ras signal transduction pathway as opposed to transcriptional activation of a reporter gene. The system is able to detect protein-protein interactions in the yeast cytoplasm instead of in the nucleus. Cyto Trap two-hybrid system restores the Ras signal transduction pathways opposed to the transcriptional activation of reporter gene in the traditional systems which rely on transcriptional factors LexA or GAL4. Cyto Trap assay uses an exclusive yeast strain (cdc25H) that harbors a temperature-sensitive mutant of the cdc25 gene, the yeast homologue for hSos. This kind of mutation allows the yeast cells to grow only at room
temperature (25 °C) but not at 37 °C. The Ras rescue depends on the unique design of the bait and target vectors in Cyto Trap system. To use the system target gene or cDNA library should be inserted into the pMyr vector which fuses the target gene to the myristylation factor that anchors the protein to the cell membrane. Bait gene or protein of interest must be cloned into the pSos vector which fuses the bait to hSos. If the bait and target interact, hSos will be recruited to the cellular membrane, and the Ras pathway can be stored, as a result, yeast cells could be assayed for growth at 37 °C. Along the biosynthetic pathway that leads to morphine, salutaridinol is acetylated by salutaridinol 7-O-acetyltransferase to salutaridinol-7-O-acetate. The oxide ring is closed and acetate is eliminated to form thebaine, the first morphinan alkaloid with the complete pentacyclic ring system (Figure III.20). The current study will focus on the molecular characterization and further *in vitro* biochemical analysis of the hypothetical “thebaine synthase” from opium poppy *Papaver somniferum* using the yeast two-hybrid assay.