

Biosynthesis and metabolic engineering of glucosinolates

Review Article

M. D. Mikkelsen^{1,2}, **B. L. Petersen**^{1,2}, **C. E. Olsen^{2,3}, and B. A. Halkier**^{1,2}

1Plant Biochemistry Laboratory, Department of Plant Biology, The Royal Veterinary and Agricultural University, 2Center for Molecular Plant Physiology (PlaCe), The Royal Veterinary and Agricultural University, and ³Department of Chemistry, The Royal Veterinary and Agricultural University,

Frederiksberg C, Copenhagen, Denmark

Accepted December 21, 2001

Summary. Glucosinolates are amino acid-derived natural plant products found throughout the Capparales order. Glucosinolates and their degradation products have a wide range of biological activities, *e.g.* in plant defense as deterrents against insect and fungi. The conversion of amino acids to aldoximes is a key step in glucosinolate biosynthesis. This step is catalyzed by cytochromes P450 from the CYP79 family. The post-aldoxime enzymes in the glucosinolate pathway have high substrate-specificity for the functional group and low substrate-specificity for the side chain. Therefore, we have been able to metabolically engineer new glucosinolate profiles into Arabidopsis by altering the levels of endogenous CYP79s and by introducing exogenous CYP79s. The approach has great potential for design of metabolically engineered plants with improved pest resistance and increased nutritional value.

Keywords: Glucosinolates – Metabolic engineering – CYP79s – Aldoximes – *Arabidopsis* – CYP83s

Introduction

Glucosinolates are natural plant products characterized by having a thioglucose moiety, a sulfonated oxime, and a side chain derived from aliphatic and aromatic amino acids as well as tryptophan (Fig. 1). Glucosinolates cooccur with the endogenous thioglucosidases, myrosinases (for review see Halkier, 1999 and Rask et al., 2000). Following tissue disruption, *e.g.* as a consequence of insect feeding, glucosinolates are hydrolyzed by myrosinases. This leads to the formation of a wide range of biologically active compounds

Fig. 1. The glucosinolate structure

including isothiocyanates, nitriles and thiocyanates. The biological activities of these compounds include inhibition of microbial growth (Mari et al., 1993; Manici et al., 1997), intermediacy in the biosynthesis of indole phytoalexins (Pedras et al., 2000), stimulation of oviposition and feeding by specialist insects (Stadler, 1978; Nielsen 1988) as well as possible deterrent effects on non-crucifer specialist insects and generalist herbivores. In addition, glucosinolates are important as flavor compounds, cancer-preventive agents, and as biopesticides. Brassica vegetables such as *e.g.* cauliflower, brussels sprouts, cabbage, broccoli and especially broccoli sprouts contain high levels of glucosinolates, of which the degradation products have strong anticarcinogenic properties (Fahey et al., 1997; Zhang and Talalay, 1994). There is a strong interest in controlling the level of glucosinolates to improve flavor and nutritional qualities of food crops and to study the physiological role of glucosinolates in plants, *e.g.* plant-insect interaction.

In the biosynthesis of the core glucosinolate structure, the first common step for the precursor amino acids is the formation of the aldoxime structure. This step is common for glucosinolates and the related group of natural plant products, the cyanogenic glucosides (Fig. 2). Cytochromes P450 of the CYP79 family have been shown to catalyze the conversion of both aliphatic and aromatic amino acids to their corresponding aldoximes in the biosynthesis of cyanogenic glucosides (Halkier et al., 1995; Andersen et al., 2000; Nielsen and Møller, 2000). Similarly, cytochromes P450 from the CYP79 family have been shown to catalyze the conversion of both aliphatic and aromatic amino acids as well as tryptophan to their corresponding aldoximes in the biosynthesis of glucosinolates (Wittstock and Halkier, 2000; Hull et al., 2000; Mikkelsen et al., 2000; Hansen et al., 2001a; Reintanz et al., 2001). Formation of the aldoxime is a key regulatory step in glucosinolate biosynthesis and identification of the CYP79 enzymes catalyzing this process has provided important tools to metabolically engineer functional food crops with improved nutritional value and improved insect and pathogen resistance. In the present review, we will focus on various aspects of metabolic engineering of glucosinolate profiles using CYP79 genes.

Fig. 2. Biosynthetic pathways of cyanogenic glucosides and glucosinolates. In the glucosinolate pathway, it is unknown whether the aldoxime is oxidized to an aci-nitro compound or to a nitrile oxide ▶

Structure and taxonomic distribution

Structure

Glucosinolates are derived from few protein amino acids (alanine, leucine, isoleucine, valine, phenylalanine, tyrosine and tryptophan), and a number of chain-elongated amino acids derived from methionine and phenylalanine, but more than 120 different glucosinolates have been identified (for review see Fahey et al., 2001). Particularly methionine is extensively chain-elongated and methionine-derived glucosinolates with as much as nine additional methylene groups have been reported (Kjær and Schuster, 1972). The precursor amino acids alone would give rise to glucosinolates with simple and branched alkyl side chains as well as ω -methylthioalkyl, aryl or heterocyclic side chains. After biosynthesis of the core structure, glucosinolates are subject to secondary modifications. These include oxidations of the side chain sulfur to sulfinyl or sulfonyl compounds and loss of the methylthio group, which leads to the formation of a terminal double bond. Other modifications include hydroxylation, methoxylation, glucosylation, sulfation and some unusual derivatives such as the 3-apiosyloxy-4-hydroxybenzylglucosinolate from *Hesperis matronalis* (Sørensen, 1990). The secondary modifications are not limited to the side chain. The presence of sinapoyl derivatives of the thioglucose moiety has been reported in *Raphanus sativa* (Linscheid et al., 1980), as has isoferuloyl derivatives in *Barbarea* species (Sørensen, 1990). The diversity of glucosinolates is therefore produced by the combination of several protein amino acids, variation in chain-elongation for some of these, and secondary modifications of the glucosinolate side chain and thioglucose moiety.

Taxonomic distribution

Glucosinolates are found in the order Capparales, which includes the economically important oilseed rape, *Brassica napus*, and *e.g.* Brassica cabbages, mustard and the model plant *Arabidopsis*. A comparison of nucleotide sequences of the chloroplast *rbcL* gene and the 18S ribosomal RNA gene supported the existence of an expanded Capparales order as suggested by Dahlgren (Dahlgren, 1975; Rodman et al., 1996; Rodman et al., 1998). This places all glucosinolate-producing species in the Capparales order with the exception of the genus *Drypetes* in the Euphorbiaceae.

Among the glucosinolate-containing families in the Capparales order there is a wide variation in the distribution of glucosinolates. Benzylglucosinolate has been found in most glucosinolate-producing plants, whereas methyl glucosinolate rarely occurs outside Capparaceae and is close to being a family-specific glucosinolate (Underhill, 1980). Indole glucosinolates are found in Brassicaceae, Capparaceae, Tropaeolaceae, Bataceae and Resedaceae (Rodman, 1991; Griffiths et al., 2001). The chain elongated glucosinolates are restricted to Brassicaceae, Capparaceae and Resedaceae (Ettlinger and Kjær, 1968). An extensive variation in glucosinolate profiles is seen, especially in Brassicaceae where *e.g. B. napus* has approximately 30 glucosinolates (Sørensen, 1990) and *Arabidopsis* has 34 different glucosinolates in 39 different ecotypes (Kliebenstein et al., 2001a). Large differences in the glucosinolate profiles in the leaves and seeds of the various ecotypes have provided a powerful tool for QTL-mapping of chain-elongating and secondary modifying activities (Mithen and Campos, 1996; Kliebenstein et al., 2001b; Kroyman et al., 2001).

Glucosinolate biosynthesis

Chain elongation

The various extended methionine and phenylalanine precursor amino acids have all been subjected to the chain elongation reaction. This reaction is thought to occur by a mechanism similar to the formation of leucine from acetate and valine (Strassman and Ceci, 1963). First, the amino acid is deaminated to produce an α -keto acid, which is then extended by condensation with acetyl-CoA. Following a rearrangement of the hydroxy group the dicarboxyacid is decarboxylated leading to the formation of a chain elongated α -keto acid. Finally, the chain-elongated amino acid is produced in a transamination reaction. This scheme was proposed solely based on evidence from *in vivo* administration of 14C-labeled amino acids and [2-14C]acetate followed by isolation of 14C-labeled glucosinolates (Matsuo and Yamazaki, 1964; Chisholm and Wetter, 1964). This scheme was further supported by 15N feeding studies in *Eruca sativa* (Graser et al., 2000) and *Arabidopsis* (Graser et al., 2001), which also showed that more than 50% of the nitrogen lost in the initial deamination reaction was reincorporated in the final transamination reaction.

Synthetic lines of *B. napus* derived by interspecies crosses with the parental lines, *Brassica rapa* and *Brassica oleracea*, have been used by Magrath et al. to examine the regulation of glucosinolate biosynthesis (Magrath et al., 1994). The synthetic lines contained a number of glucosinolates not found in natural *B. napus*. The lines were used to map two loci, *Gsl-elong-A* and *Gsl-elong-C* that were shown to be involved in the production of butyl- and pentylglucosinolates, and *Gsl-pro* that was shown to regulate production of propylglucosinolates (Magrath et al., 1994). In a similar strategy, recombinant inbred lines between the *Arabidopsis* ecotypes Landsberg *erecta* and Colombia were used to identify the *Gsl-elong-Ar* locus, which was found to regulate the production of butylglucosinolates as opposed to propylglucosinolates (Magrath et al., 1994; Mithen and Campos, 1996). Recently, fine scale mapping of the QTL involved in methionine chainelongation in *Arabidopsis* enabled the identification of two members of an isopropylmalate synthase-like gene family (de Quiros et al., 2000). One of the genes, *MAM-1*, encodes a methylthioalkylmalate synthase and controls whether methionine is chain-elongated with one or two methylene groups (Kroyman et al., 2001). The function of the other gene, termed *MAM-L* for *MAM*-like, is unknown as heterologous expression in *Escherichia coli* did not

result in functional enzyme and/or the correct substrate has not been identified. However, the high sequence homology to *MAM-1* suggests that it is very likely to be catalyzing a similar reaction. Glucosinolate analysis of *MAM-L* knockout lines is likely to answer this question. It is an open question whether every chain-elongation cycle requires a separate MAM-activity or if the substrate specificity enables each MAM-enzyme to metabolize a series of similar substrates. Furthermore, it is puzzling that mutants and ecotypes deficient in the elongation of trihomomethionine to tetrahomomethionine are able to produce long-chain, *i.e.*, penta- and hexahomomethionine. A pathway partially independent from that forming the short-chain methionine-derived glucosinolates has been suggested (Kroyman et al., 2001).

Aldoxime formation

The first common step for the precursor amino acids is the formation of the aldoxime structure by the action of cytochromes P450 from the CYP79 family (Fig. 3). Based on the knowledge from the cyanogenic glucoside pathway and biochemical characterizations, we have used a functional genomics approach to identify aldoxime-producing enzymes in the glucosinolate pathway. In parallel, other groups have identified several of the CYP79 genes in mutant screens (see below). In the genome of *Arabidopsis* ecotype Colombia, seven CYP79s are present: *CYP79A2*, *CYP79B2*, *CYP79B3*, *CYP79C1*, *CYP79C2*, *CYP79F1* and *CYP79F2*. We have successfully expressed and characterized several of these genes. Recombinant CYP79A2 was able to convert L- $[{}^{14}C]$ phenylalanine to the (E) - and (Z) -isomers of phenylacetaldoxime. *CYP79B2* was identified as an EST with homology to the *Sorghum bicolor CYP79A1* (Mikkelsen et al., 2000), and in a screen to find an *Arabidopsis* cDNA with the ability to confer resistance to 5-fluoroindole to yeast (Hull et al., 2000). Both *CYP79B2* (Mikkelsen et al., 2000; Hull et al., 2000) and the homologue *CYP79B3* (Hull et al., 2000) were heterologously expressed in *E. coli* and both enzymes were shown to convert tryptophan to indole-3-acetaldoxime.

Fig. 3. An overview of the endogenous and exogenous CYP79s used in metabolically engineering *Arabidopsis* plants. The generated aldoximes can be precursors for several biosynthetic pathways as indicated

Indole-3-acetaldoxime (IAOX) is a potential intermediate in glucosinolate biosynthesis as well as in the biosynthesis of the plant hormone indole-3-acetic acid (IAA) (for review see Bartel, 1997; Normanly and Bartel, 1999). Accordingly, CYP79B2 and CYP79B3 may act in IAA biosynthesis as well as in the biosynthesis of indole glucosinolates. CYP79B-specific PCR primers have been used to investigate the presence of CYP79Bs in nonglucosinolate producing species. No CYP79Bs was detected in these plant species, whereas the primers detected CYP79Bs in all the glucosinolateproducing species that were tested (Bak et al., 1998). Furthermore, only five families within the Capparales order are known to accumulate tryptophanderived glucosinolates (Rodman, 1991; Griffiths et al., 2001). Therefore, it is unlikely that CYP79-dependent IAOX formation is a general mechanism for IAA biosynthesis, although it may represent a biosynthetic pathway specific for species containing indole glucosinolates. Several other pathways for tryptophan-dependent IAOX formation have been suggested (for review see Bartel, 1997; Normanly and Bartel, 1999). Although IAOX constitute a branching point between indole glucosinolates and IAOX-derived IAA, the role of IAOX in IAA biosynthesis is not yet fully understood.

CYP79F1 was identified in a database search (Hansen et al., 2001a) and as a recessive knockout mutant characterized by having a bushy phenotype (*bus*; Reintanz et al., 2001; *supershoot*; Tantikanjana et al., 2001). Heterologous expression of CYP79F1 in *E. coli* showed that the recombinant enzyme was able to convert homo-, dihomo-, trihomo-, tetrahomo-, pentahomo- and hexahomo-methionine to the respective (E) - and (Z) -isomers of the corresponding aldoximes (Hansen et al., 2001a; Hansen, Chen, Glawischnig and Halkier, unpublished data). *CYP79F2* is 88% identical to *CYP79F1* on the amino acid level. When heterologously expressed in *Saccharomyces cerevisiae* CYP79F2 has been shown to convert long chain methionine-derived amino acids penta-, and hexahomomethionine, to the corresponding aldoximes (Chen, Glawischnig, Naur and Halkier, unpublished data). This shows that CYP79F1 and CYP79F2 are partly redundant. The mechanism for the regulation at the branching point between continuation of chain-elongation of precursor amino acids and entrance to the biosynthetic pathway of the core glucosinolate structure, where CYP79Fs catalyze the first committed step, is unknown. In CYP79F1 knockouts, the long-chain methionine-derived glucosinolates accumulate to higher levels than in the wild type (Reintanz et al., 2001; Hansen et al., 2001a) as does the short-chain methionine precursor amino acids (Hansen et al., 2001a). This suggests that a combination of substrate affinity and fluxes between pools is a determining factor for the final glucosinolate profile. Further studies are required to understand the regulatory network behind the accumulation of chain-elongation of aliphatic glucosinolates.

The CYP79 enzymes presently characterized, are responsible for aldoxime production for the biosynthesis of all indole-, phenylalanine- and all chain-elongated methionine-derived glucosinolates found in *Arabidopsis*. Additional glucosinolates found in *Arabidopsis* include the homophenylalanine-derived glucosinolate, phenylethylglucosinolate, but also

trace amounts of the tyrosine-derived glucosinolate, *p*-hydroxybenzylglucosinolate (Chen et al., 2001), as well as the methionine-derived methylsulfinylethylglucosinolate (Kliebenstein et al., 2001a). The substrate specificity of two closely related CYP79s from *Arabidopsis*, CYP79C1 and CYP79C2, has not yet been determined. The transcripts are present at very low levels, which suggests that CYP79C1 and CYP79C2 are responsible for aldoxime formation of low abundant glucosinolates such as *e.g.* the tyrosine-, methionine- or homophenylalanine-derived glucosinolates. Alternatively, *in vivo* activities of the characterized CYP79s could differ from the *in vitro* activities, in which case *e.g.* CYP79F1 may be able to convert methionine to its aldoxime and CYP79A2 may be able to convert homophenylalanine and possibly tyrosine to their aldoximes, albeit at very low efficiencies.

Cytochromes P450 are generally known to be located on the cytosolic side of the ER membrane. CYP79B2 was suggested to be targeted to the chloroplast when the computer prediction program ChloroP was used on the N-terminal 22 amino acids (Hull et al., 2000; Celenza, personal communication). When we apply the two prediction programs Psort and TargetP v. 1.0 (http://psort.nibb.ac.jp and http://www.cbsdtu.dk, respectively) neither CYP79B2 nor CYP79B3 are targeted to the chloroplast (Mikkelsen and Halkier, unpublished data). This lack of consistency indicates that the predictions are highly questionable. However, translational fusions between the amino-termini of both CYP79B2 and CYP79B3 with GUS have been shown to target GUS to the plastids (Celenza, 2001). In addition, the ATR2 *Arabidopsis* NADPH-cytochrome P450 reductase has been predicted to be targeted to the chloroplast (Urban et al., 1997). However, according to ChloroP, neither ATR1, ATR2 nor ATR3 are predicted to be targeted to the chloroplast (Mikkelsen and Halkier, unpublished data). TargetP predicted that ATR2 is localized in the chloroplast, but this was at the second lowest level of confidence indicating that the prediction is highly speculative. Furthermore, Psort predicted that ATR2 was localized on the ER membrane. Further studies are needed to conclusively determine the subcellular localization of CYP79B2 and CYP79B3.

Formation of thiohydroximic acids

The aldoximes produced by the CYP79s are converted to glucosinolates by a number of reactions. The first step is the oxidation of the aldoxime. CYP83A1 and CYP83B1 from *Arabidopsis* have been shown to be able to catalyze the oxidation of various aldoximes (Bak et al., 2001; Hansen et al., 2001b; Bak and Feyereisen, 2001). It has been proposed that the aldoxime is first oxidized to produce an *aci*-nitro compound or a nitrile oxide compound, which functions as the acceptor for the thiol donor (Ettlinger and Kjær, 1968). Although the nature of the product formed by CYP83B1 has not been identified, it does efficiently conjugate with S-donors to form an *S*-alkylthiohydroximate (Bak et al., 2001; Hansen et al., 2001b). In most species, cysteine has been shown to be the most effective sulfur donor, though methionine has been reported to be

equally effective in other species (Wetter and Chisholm, 1968). CYP83B1 has been shown to have high affinity for tryptophan- and phenylalanine-derived aldoximes (Bak and Feyereisen, 2001), although it does metabolize the aliphatic aldoximes with very low affinity (Naur, Bak, Halkier, in preparation). Similarly, CYP83A1 has high affinity for the aliphatic aldoximes although it does metabolize the aromatic aldoximes and IAOX with low affinity (Bak and Feyereisen, 2001; Naur, Bak, Halkier, in preparation). The biochemical data indicate that CYP83A1 and CYP83B1 are not redundant in the plant under normal physiological conditions.

Following the oxidation of the aldoxime by the CYP83s and the conjugation to a sulfur donor, the thiohydroximic acid is formed by cleavage of the Salkylthiohydroximate in a reaction thought to be catalyzed by a C-S lyase (Wallsgrove and Bennett, 1995). No C-S lyase has yet been cloned in the glucosinolate pathway. Generally, the post-aldoxime enzymes have low specificity for the side chain and high specificity for the functional group. However, it remains an open question how many C-S lyases are involved in the pathway, and whether there are *e.g.* two non-redundant C-S lyases as is the case for the aldoxime-metabolizing CYP83s.

Formation of desulfoglucosinolates and glucosinolates

The thiohydroximic acid is further S-glucosylated by a soluble UDPG: thiohydroximate glucosyltransferase (S-GT) to produce a desulphoglucosinolate that subsequently is converted to the glucosinolate by a PAPS: desulfoglucosinolate sulfotransferase. S-GTs have been purified from several species (Guo and Poulton, 1994; Jain et al., 1990a). The S-GT purified from *B. juncea* was highly specific for the thiohydroximic acid moiety and low specificity for the side chain of thiohydroximic acids was seen (GrootWassink et al., 1994). In the study, S-GT activity was shown to be present in all glucosinolate-producing crucifers examined, whereas all non-glucosinolate producing plants examined lacked the S-GT activity. An S-GT from *B. napus* has been cloned (Grootwassink et al., 1997). Based on sequence homology a single candidate S-GT has been identified in the *Arabidopsis* genome (http://www.biobase.dk/p450/gst.shtml, Petersen et al., 2001). This is in agreement with the S-GT enzyme having low specificity for the side chain. Furthermore, S-GTs have high homology to glucosyltransferases involved in the biosynthesis of cyanogenic glucosides. This suggests that the S-GTs have evolved from the family of glucosyltransferases, *e.g.* by gene duplication, and have differentiated into the S-GTs.

Sulfotransferases purified and characterized from *B. juncea*, and *Lepidium sativa* have been shown to catalyze sulfation of a number of different glucosinolates (Glendening and Poulton, 1988; Jain et al., 1990b). The efficiency of the *L. sativa* sulfotransferase was three fold higher when using desulfobenzylglucosinolate than when using desulfoallylglucosinolate and sulfation of other compounds than desulfoglucosinolates was not observed. This indicates limited substrate specificity and suggests that the enzyme may have a preference for specific desulfoglucosinolate side chain structures and that more than one sulfotransferase may be needed to cover the range of glucosinolates found in some species, especially among the Brassicaceae (Jain et al., 1990b). No sulfotranferase in the glucosinolate pathway have yet been cloned. However, with the sequencing of the *Arabidopsis* genome, several putative sulfotransferases have been identified (Marsolais et al., 2000). Accordingly, it will be a major task to identify those specific for glucosinolate biosynthesis.

Secondary modifications

Following the formation of the basic glucosinolate structure a wide range of secondary modifications may take place on the side chain as well as on the glucose moiety. These include oxidation, hydroxylation, methoxylation, desaturation, sulfation and glucosylation. The most powerful tool for dissecting the secondary modifications of glucosinolates have been the use of QTL-mapping in various *Arabidopsis* ecotypes and in synthetic *B. napus* lines. In *B. napus*, two loci, *Gsl-oh-C* and *Gsl-oh-A*, have been shown to regulate the hydroxylation of several alkenylglucosinolates (Giamoustaris and Mithen, 1996; Parkin et al., 1994). Recent fine scale QTL mapping in *Arabidopsis* identified a 2-oxoglutarate-dependent dioxygenase as single candidate gene for this activity (Kliebenstein et al., in submission). Genetic studies have led to the proposal of a model where the initial oxidation of methionine-derived glucosinolates is regulated by alleles of the *Gsl-oxid* loci, and the oxidation to alkenyl glucosinolates is regulated by alleles of the *Gsl-alk* loci (Giamoustaris and Mithen, 1996). In *Arabidopsis*, the same conversion of methylsulfinylalkyl glucosinolates to alkenyl glucosinolates have been shown to be regulated by alleles at the *Gsl-alk* locus (Mithen et al., 1995). Furthermore, the *Gsl-ohp* locus, responsible for the formation of hydroxypropylglucosinolate was located close to the *Gsl-alk* locus (Mithen et al., 1995). Recently, fine scale mapping of these loci identified three genes, *AOP1*, *AOP2* and *AOP3*, which encodes 2-oxoglutarate-dependent dioxygenases (Kliebenstein et al., 2001a, Hall et al., 2001). No function was assigned to AOP1, but recombinant AOP2 was shown to catalyze the conversion of 3-methylsulfinylpropyl- and 4-methylsulfinylbutylglucosinolate to the corresponding alkenylglucosinolates (Kliebenstein et al., 2001b). Furthermore, recombinant AOP3 was shown to catalyze the conversion of 3-methylsulfinylpropyl- to 3 hydroxyalkylpropylglucosinolate (Kroymann et al., 2001).

The aldoxime as an evolutionary branchpoint

The IAOX forms a metabolic branchpoint between indole glucosinolates and IAA, and possibly other indole-derived natural products in the Brassicaceae. The suggested role for CYP79Bs in IAA biosynthesis is supported by the increased auxin phenotype of CYP83B1 knockout plants (Barlier et al., 2000; Bak et al., 2001). Furthermore, *CYP83B1* is induced by IAA (Delarue et al., 1999) and the *CYP83B1* promotor contains four putative auxin-responsive

elements whereas the *CYP83A1* promotor does not (Bak and Feyereisen, 2001). This could be a mechanism to limit IAA biosynthesis during periods of high IAA concentration as *CYP83B1* would be induced and IAOX channeled away from IAA biosynthesis into glucosinolate biosynthesis (Bak et al., 2001).

CYP83B1 and CYP83A1 have a high degree of similarity to the aldoximemetabolizing CYP71E1 in the biosynthesis of the cyanogenic glucoside dhurrin in *S. bicolor* (Hansen et al., 2001b). This supports the hypothesis of an evolutionary relationship between the biosynthesis of cyanogenic glucosides and glucosinolates. According to the hypothesis, the glucosinolate pathway may have evolved from the cyanogenic glucoside pathway by a mutation in an aldoxime-metabolizing enzyme (Hansen et al., 2001b). Instead of converting the aldoxime to a β -hydroxynitrile for the cyanogenic glucoside pathway, a reactive intermediate was formed, which was detoxified by conjugation with cysteine, followed by a C-S lyase, a glycosylation and sulfation reaction, all of which are general detoxification reactions (Hansen et al., 2001b). Furthermore, the myrosinase (a β -thioglucosidase) responsible for hydrolyzing the glucosinolates, has high homology to the β -glucosidase in the cyanogenic glucoside pathway, which suggests a possible evolutionary relationship between the hydrolyzing enzymes (Rodman et al., 1998; Hansen et al., 2001b).

Regulation of glucosinolate biosynthesis

Biosynthesis of specific glucosinolates is regulated both by signaling molecules and environmental factors. Indole glucosinolates have been shown to accumulate systemically in plants treated with jasmonic acid, a signal molecule associated with response to wounding, herbivory and infection (Doughty et al, 1995). This suggests a role for indole glucosinolates in wound response (Brader et al., 2001; Mikkelsen and Halkier, unpublished data). Similarly, in *B. napus* biosynthesis of gluconasturtiin (2 phenylethylglucosinolate) is specifically induced by salicylic acid, a signaling compound involved in many plant physiological processes (Kiddle et al., 1994). The accumulation of indole glucosinolates following jasmonic acid treatment and infection with certain pathogens may be regulated at the transcriptional level. In agreement with this, *CYP79B2* and *CYP79B3* are induced by MeJA treatment (Mikkelsen et al., unpublished data) and *CYP79B2* is induced by wounding (Mikkelsen et al., 2000) and by infection with *Pseudomonas syringae* (Hull et al., 2000). Brader et al. showed that *CYP79B3*, but not *CYP79B2*, was induced by treatment with jasmonic acid as well as a cultural filtrate of *Erwinia carotovora* (Brader et al., 2001). This suggests specific but different roles for CYP79B2 and CYP79B3, which may not be physiologically redundant genes.

Metabolic engineering

There is a strong interest in altering levels of specific glucosinolates in crop plants as certain glucosinolates have desirable and others have undesirable

properties. The identification and characterization of the *Arabidopsis* cytochromes P450 belonging to the CYP79 family, have provided an important tool for modulating the profile of glucosinolates derived from protein amino acids. Overexpression, downregulation, and knockouts of endogenous, possibly combined with the introduction of exogenous CYP79s makes it possible to generate custom designed glucosinolate profiles. In the following, several examples of altered, metabolically engineered glucosinolate profiles in *Arabidopsis* will be described.

Metabolic engineering using endogenous CYP79s

CYP79A2 converts phenylalanine to its aldoxime in the biosynthesis of benzylglucosinolate that is barely detectable in the wild type plant (Wittstock and Halkier, 2000). This is reflected in the expression pattern of *CYP79A2* as evidenced by CYP79A2-promotor-GUS fusion (Hansen and Halkier, unpublished data). The expression pattern is very similar to the infection pattern of the bacterial pathogen *Xanthomonas campestris* (Hugouvieux et al., 1998). This suggests that CYP79A2 may function as a 'watchdog' to minimize the chance and the extent of infection by producing benzylglucosinolate at the possible entry points of *X. campestris i.e.* in the hydathodes. *Arabidopsis* transformed with a 35S::CYP79A2 construct accumulated as much as 18nmol/mg dry weight corresponding to approximately 50% of the total glucosinolate content (Wittstock and Halkier, 2000; Fig. 4).

Arabidopsis expressing a 35S::CYP79B2 construct was shown to accumulate up to five-fold more indole glucosinolates than wild type (Mikkelsen et al., 2000). The largest difference was seen for the 'unmodified' indole glucosinolates and less for the methoxylated analogues indicating that the methoxylating activities are rate-limiting. A *CYP79B2* knockout line had the same glucosinolate profile as wild type (Glawischnig et al., unpublished

Fig. 4. Metabolically engineered glucosinolate profiles in transgenic *Arabidopsis* plants transformed with different 35S::CYP79 constructs. Upregulation of endogenous and introduction of exogenous CYP79s provides a powerful tool for alteration of the glucosinolate profiles and introduction of novel glucosinolates. The chain-elongation step might be rate-limiting for the accumulation of glucosinolates derived from chainelongated amino acids

results). This is most likely due to the functional homologue CYP79B3 that appears to be expressed at a higher basal level than CYP79B2 (Mikkelsen et al., unpublished results). Accordingly, a *CYP79B2:CYP79B3* double knockout is required to abolish biosynthesis of indole glucosinolates. However, if CYP79B2 and CYP79B3 are involved in the biosynthesis of IAA such a double knockout may be lethal.

CYP79F1 knockout mutants and 35S::CYP79F1 co-suppressing lines have been shown to have a characteristic bushy phenotype (Hansen et al., 2001a; Reintanz et al., 2001; Tantikanjana et al., 2001; Hansen and Halkier, unpublished results). The co-suppressing line was shown to contain 50 fold more di-homomethionine and 10 fold more tri-homomethionine (Hansen et al., 2001a). An accumulation of homo- to tetrahomomethionine would be expected in this line, as CYP79F1 is the only enzyme able to catalyze this reaction. As a consequence hereof, in both the co-suppressed lines and in the knockout mutants, an increase of the long-chain methionine-derived glucosinolates was seen. In addition, the knockout mutant was shown to contain elevated levels of IAA (Reintanz et al., 2001; Tantikanjana et al., 2001). Since CYP79F1 metabolizes chain-elongated methionine it is possible that the elevated IAA levels found in this plant is an indirect effect due to metabolic stress. *Arabidopsis* overexpressing CYP79F1 was shown to accumulate at the most two-times more of the homo- to tetrahomomethionine derivedglucosinolates (Reintanz et al., 2001; Hansen et al., 2001a). This is significantly less than for *Arabidopsis* overexpressing *e.g. CYP79A2* or *CYP79B2* and most likely reflects that the rate-limiting step in these lines is the chain-elongation pathway.

Metabolic engineering using exogenous CYP79s

CYP79A1 catalyzes the conversion of tyrosine to its aldoxime in the biosynthesis of the cyanogenic glucoside, dhurrin in *S. bicolor* (Halkier et al., 1995). *Arabidopsis* transformed with a 35S::CYP79A1 construct was shown to accumulate high levels of the tyrosine-derived glucosinolate, *p*-hydroxybenzylglucosinolate (Bak et al., 1999). The breakdown products of *p*-hydroxybenzylglucosinolate are characteristic for the flavor of mustard. Similarly, CYP79D2 catalyzes the conversion of valine and isoleucine to the corresponding aldoximes in the biosynthesis of the cyanogenic glucosides, linamarin and lotoustralin in *Manihot esculenta* Crantz (Cassava; Andersen et al., 2000). *Arabidopsis* transformed with CYP79D2 accumulates the valine- and isoleucine-derived glucosinolates, isopropyl-glucosinolate and 1-methylpropylglucosinolate (Mikkelsen and Halkier, unpublished data). 1 methylpropylglucosinolate has not previously been identified in *Arabidopsis* and only in very few ecotypes have trace amounts of isopropylglucosinolate been found (Kliebenstein et al., 2001a). The breakdown products of isopropylglucosinolate are characteristic for the flavor of capers. This shows that it is possible to specifically alter the concentrations of not only specific endogenous glucosinolates but also to introduce novel glucosinolates not normally present in *Arabidopsis*.

Future prospects for metabolic engineering of glucosinolates

Powerful tools for modifying glucosinolate profiles in glucosinolate-producing plants are now available. Modulating expression levels of endogenous CYP79s have been shown to have large effects on the glucosinolate profiles. Introducing new CYP79s have furthermore enabled us to produce new glucosinolates in *Arabidopsis*, as exemplified by a mustard- and capersflavored *Arabidopsis*. Future metabolic engineering involving the genes from the chain-elongating pathway, the aldoxime-forming CYP79s and the secondarily modifying enzymes, will provide the possibilities of designing crop plants, which are enriched in desirable glucosinolates and free of undesirable glucosinolates. This will ultimately result in crop plants with improved tolerance to pests, improved nutritional values including improved taste and improved cancer-preventing effects.

References

- Andersen MD, Busk PK, Svendsen I, Møller BL (2000) Cytochromes P-450 from cassava (Manihot esculenta Crantz) catalyzing the first steps in the biosynthesis of the cyanogenic glucosides linamarin and lotaustralin. Cloning, functional expression in Pichia pastoris, and substrate specificity of the isolated recombinant enzymes. J Biol Chem 21: 1966–1975
- Bak S, Feyereisen R (2001) The involvement of two P450 enzymes, CYP83B1 and CYP83A1, in auxin homeostasis and glucosinolate biosynthesis. Plant Physiol 127: 108–118
- Bak S, Nielsen HL, Halkier BA (1998) The presence of CYP79 homologues in glucosinolate-producing plants show evolutionary conservation of the enzymes in the conversion of amino acid to aldoxime in the biosynthesis of cyanogenic glucosides and glucosinolates. Plant Mol Biol 38: 725–734
- Bak S, Olsen CE, Petersen BL, Møller BL, Halkier BA (1999) Metabolic engineering of *p*-hydroxybenzylglucosinolate in *Arabidopsis* by expression of cyanogenic CYP79A1 from *Sorghum bicolor*. Plant J 20: 663–671
- Bak S, Tax FE, Feldmann KA, Galbraith DW, Feyereisen R (2001) CYP83B1, a cytochrome P450 at the metabolic branchpoint in auxin and indole glucosinolate biosynthesis in Arabidopsis. Plant Cell 13: 101–111
- Barlier I, Kowalczyk M, Marchant A, Ljung K, Bhalerao R, Bennett M, Sandberg G, Bellini C (2000) The SUR2 gene of Arabidopsis thaliana encodes the cytochrome P450 CYP83B1, a modulator of auxin homeostasis. Proc Natl Acad Sci USA 97: 14819–14824
- Bartel B (1997) Auxin biosynthesis. Annu Rev Plant Physiol Plant Mol Biol 48: 51–66
- Brader G, Tas É, Palva ET (2001) Jasmonate-dependent induction of indole glucosinolates in Arabidopsis by culture filtrates of the nonspecific pathogen *Erwinia carotovora*. Plant Physiol 126: 849–860
- Celenza JL (2001) Metabolism of tyrosine and tryptophan new genes for old pathways. Curr Opin Plant Biol 4: 234–240
- Chen S, Petersen BL, Olsen KE, Schulz A, Halkier BA (2001) Long-distance phloem transport of glucosinolates in Arabidopsis. Plant Physiol 127: 194–201
- Chisholm MD, Wetter LR (1964) Biosynthesis of mustard oil glucosides IV. The administration of methionine- $C¹⁴$ and related compounds to horse radish. Can J Biochem 42: 1033–1040
- Dahlgren R (1975) A system of classification of the angiosperms to be used to demonstrate the distribution of characters. Botaniske Notiser 128: 119–147
- de Quiros HC, Magrath R, McCallum D, Kroymann J, Schnabelrauch D, Mitchell-Olds T, Mithen R (2000) α-Keto acid elongation and glucosinolate biosynthesis in *Arbidopsis thaliana*. Theor Appl Genet 101: 429–437
- Delarue M, Muller P, Bellini C, Delbaraa A (1999) Increased auxin efflux in the IAAoverproducing sur 1 mutant of Arabidopsis thaliana: a mechanism of reducing auxin levels? Physiol Plant 107: 120–127
- Doughty KJ, Kiddle GA, Pye BJ, Wallsgrove RM, Pickett JA (1995) Selective induction of glucosinolates in oilseed rape leaves with methyl jasmonate. Phytochemistry 38: 347–350
- Ettlinger MG, Kjær A (1968) Sulfur compounds in plants. In: Mabry TJ, Alston RE, Runeckles VC (eds) Recent advances in phytochemistry 1. Appleton, Century Crofts New York, pp 59–144
- Fahey JW, Zhang Y, Talalay P (1997) Broccoli sprouts: an exceptionally rich source of inducers of enzymes that protect against chemical carcinogens. Proc Natl Acad Sci USA 94: 10367–10372
- Fahey JW, Zalcmann AT, Talalay P (2001) The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochemistry 56: 5–51
- Giamoustraris A, Mithen R (1996) Genetics of aliphatic glucosinolates. IV. Side-chain modification in *Brassica oleracea*. Theor Appl Genet 93: 1006–1010
- Glendening TM, Poulton JE (1988) Glucosinolate biosynthesis. Sulfation of desulfoglucosinolates by cell-free extracts of cress (*Lepidium sativum* L.) seedlings. Plant Physiol 86: 319–321
- Graser G, Schneider B, Oldham NJ, Gershenzon J (2000) The methionine chain elongation pathway in the biosynthesis of glucosinolates in *Eruca sativa* (Brassicaceae). Arch Biochem Biophys 378: 411–419
- Graser G, Oldham NJ, Brown PD, Temp U, Gershenzon J (2001) The biosynthesis of benzoic acid glucosinolate esters in Arabidopsis thaliana. Phytochemistry 57: 23– 32
- Griffiths DW, Deighton N, Birch ANE, Patrian B, Baur R, Städtler E (2001) Identification of glucosinolates on the leaf surface of plants from the Cruciferae and other closely related species. Phytochemistry 57: 693–700
- Grootwassink JW, Reed DW, Kolenovsky AD (1994) Immunopurification and immunocharacterization of the glucosinolate biosynthesis enzyme thiohydroximate *S*-glucosyltransferase. Plant Physiol 105: 425–433
- Grootwassink JWD, Underhill EW, Hemmingsen SM, Reed DW, Kolenovsky AD (1997) European patent: EP 0 771 878 A1
- Guo L, Poulton JE (1994) Partial purification and characterization of *Arabidopsis thaliana* UDPG-thiohydroximate glucosyltransferase. Phytochemistry 36: 1133–1138
- Halkier BA, Nielsen HL, Koch B, Moller BL (1995) Purification and characterization of recombinant cytochrome P450TYR expressed at high levels in Escherichia coli. Arch Biochem Biophys 322: 369–377
- Halkier BA (1999) Glucosinolates. In: Ikan R (ed) Naturally occurring glycosides. John Wiley & Sons, Inc., New York, pp 193–223
- Hall C, McCallum D, Prescott A, Mithen R (2001) Biochemical genetics of glucosinolate modification in *Arabidopsis* and *Brassica*. Theor Appl Genet 102: 369–374
- Hansen CH, Wittstock U, Olsen CE, Hick AJ, Pickett JA, Halkier BA (2001a) Cytochrome P450 CYP79F1 from *Arabidopsis* catalyzes the conversion of dihomomethionine and trihomomethionine to the corresponding aldoximes in the biosynthesis of aliphatic glucosinolates. J Biol Chem 276: 11078–11085
- Hansen CH, Du L, Naur P, Olsen CE, Axelsen CB, Hick AJ, Pickett JA, Halkier BA (2001b) CYP83B1 is the oxime-metabolizing enzyme in the glucosinolate pathway in *Arabidopsis*. J Biol Chem 276: 24790–24796
- Hugovieux V, Barber CE, Daniels MJ (1998) Entry of *Xanthomonas campestris* pv. *campestris* into hydathodes of *Arabidopsis thaliana* leaves: a system for studying early infection events in bacterial pathogenesis. Amn Phytopathol Soc 11: 537–543
- Hull AK, Rekha V, Celenza JL (2000) *Arabidopsis* cytochrome P450s that catalyze the first step of tryptophan-dependent indole-3-acetic acid biosynthesis. Proc Natl Acad Sci 97: 2379–2384
- Jain JC, GrootWassink JW, Reed DW, Underhill EW (1990a) Persistent co-purification of enzymes catalyzing the sequentioal glycosylation and sulfation steps in glucosinolate biosynthesis. J Plant Physiol 136: 356–361
- Jain JC, GrootWassink JW, Kolenovsky AD, Underhill EW (1990b) Purification and properties of 3'-phosphoadenosine-5'-phosphosulfate: desulfoglucosinolate sulfotransferase from *Brassica juncea* cell cultures. Phytochemistry 29: 1425–1428
- Kiddle GA, Doughty KJ, Wallsgrove RM (1994) Salicylic acid-induced accumulation of glucosinolates in oilseed rape (*Brassica napus*) leaves. J Exp Bot 45: 1343–1346
- Kjær A, Schuster A (1972) Glucosinolates in seeds of *Neslia paniculata*. Phytochemistry 11: 3045–3048
- Kliebenstein DJ, Kroymann J, Brown P, Figuth A, Pedersen D, Gershenzon J, Mitchell-Olds T (2001a) Genetic control of natural variation in Arabidopsis glucosinolate accumulation. Plant Physiol 126: 811–825
- Kliebenstein DJ, Lambrix VM, Reichelt M, Gershenzon J, Mitchell-Olds T (2001b) Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutaratedependent dioxygenases control glucosinolate biosynthesis in Arabidopsis. Plant Cell 13: 681–693
- Kroymann J, Textor S, Tokuhisa JG, Falk KL, Bartram S, Gershenzon J, Mitchell-Olds T (2001) A gene controlling variation in Arabidopsis glucosinolate composition is part of the methionine chain elongation pathway. Plant Physiol 127: 1077–1088
- Linscheid M, Wendish D, Strack D (1980) The structures of sinapic acid esters and their metabolism in cotyledons of *Raphanus sativus.* Z Naturforsch 35c: 907
- Magrath R, Bano F, Morgner M, Parkin I, Sharp A, Lister C, Dean C, Turner J, Lydiate D, Mithen R (1994) Genetics of aliphatic glucosinolates. I. Side chain elongation in *Brassica napus* and *Arabidopsis thaliana*. Heredity 72: 290–299
- Manici LM, Lazzeri L, Palmieri S (1997) In vitro fungitoxic activity of some glucosinolates and their enzyme-derived products toward plant pathogenic fungi. J Agric Food Chem 45: 2768–2773
- Mari M, Iori R, Leoni O, Marchi A (1993) In vitro activity of glucosinolate-derived isothiocyanates against postharvest fruit pathogens. Ann Appl Biol 123: 155–164
- Marsolais F, Gidda SK, Boyd J, Varin L (2000) Plant soluble sulfotransferases: structural and functional similarity with mammalian enzymes. In: Romeo JT, Ibrahim R, Varin L, de Luca V (eds) Evolution of metabolic pathways. Elsevier Science Ltd., Amsterdam, pp 433–456
- Matsuo M, Yamazaki M (1964) Biosynthesis of sinigrin. II. Chem Pharm Bull 12: 1388– 1389
- Mikkelsen MD, Hansen CH, Wittstock U, Halkier BA (2000) Cytochrome P450 CYP79B2 from *Arabidopsis* catalyzes the conversion of tryptophan to indole-3 acetaldoxime, a precursoer of indole glucosinolates and indole-3-acetic acid. J Biol Chem 275: 33712–33717
- Mithen R, Clarke J, Lister C, Dean C (1995) Genetics of aliphatic glucosinolates. III. Side chain structure of aliphatic glucosinolates in *Arabidopsis thaliana*. Heredity 74: 210–215
- Mithen R, Campos H (1996) Genetic variation of aliphatic glucosinolates in *Arabidopsis thaliana* and prospects for map-based gene cloning. Entomol Exp Appl 80: 202–205
- Nielsen JK (1988) Crucifer-feeding Chrysomelidae: mechanisms and of host plant finding and acceptance. In: Jolivet P, Petitpierre E, Hsiao TH (eds) Biology of chrysomelidae. Kluwer, Dordrecht, pp 25–40
- Nielsen JS, Møller BL (2000) Cloning and expression of cytochrome P450 enzymes catalyzing the conversion of tyrosine to p-hydroxyphenylacetaldoxime in the biosynthesis of cyanogenic glucosides in Triglochin maritima. Plant Physiol 122: 1311–1321
- Normanly J, Bartel B (1999) Redundancy as a way of life IAA metabolism. Curr Opin Plant Biol 2: 207–213
- Parkin I, Magrath R, Keith D, Sharpe A, Mithen R, Lydiate D (1994) Genetics of aliphatic glucosinolates. II. Hydroxylation of alkenyl glucosinolates in *Brassica napus*. Heredity 72: 594–598
- Pedras MSC, Okanga FI, Zaharia IL, Khan AQ (2000) Phytoalexins from crucifers: synthesis, biosynthesis, and biotransformation. Phytochemistry 53: 161–176
- Petersen BL, Andréason E, Bak S, Agerbirk N, Halkier BA (2001) Characterization of transgenic *Arabidopsis thaliana* with metabolically engineered high levels of *p*-hydroxybenzylglucosinolate. Planta 212: 612–618
- Rask L, Andreasson E, Ekbom B, Eriksson S, Pontoppidan B, Meijer J (2000) Myrosinase: gene family evolution and herbivore defense in Brassicaceae. Plant Mol Biol 42: 93–113
- Reintanz B, Lehnen M, Reichelt M, Gershenzon J, Kowalczyk M, Sandberg G, Godde M, Uhl R, Palme K (2001) *bus*, a bushy Arabidopsis *CYP79F1* knockout mutant with abolished synthesis of short-chain aliphatic glucosinolates. Plant Cell 13: 351–367
- Rodman JE (1991) A taxonomic analysis of glucosinolate producing plants, part 1. Phenetics Syst Bot 16: 598–618
- Rodman JE, Karol KG, Price RA, Sytsma KJ (1996) Molecules, morphology, and Dahlgren's expanded order Capparales. Syst Bot 21: 289–307
- Rodman JE, Soltis PS, Soltis DE, Sytsma KJ, Karol KG (1998) Parallel evolution of glucosinolate biosynthesis inferred from congruent nuclear and plastid gene phylogenies. Am J Bot 85: 997–1006
- Stadler E (1978) Chemoreception of host plant chemicals by oviposition females of *Delia* (Hylemia) *Brassicae*. Entomol Exp Appl 24: 711–720
- Strassman M, Ceci LN (1963) Enzymatic formation of α -isopropylmalic acid, an intermediate in leucine biosynthesis. J Biol Chem 238: 2445–2452
- Sørensen H (1990) Glucosinolates: structure, properties and function. In: Shahidi F (ed) Canola and rapeseed. Production, chemistry, nutrition and processing technology. Van Nostrand Reinhold, New York, pp 149–172
- Tantikanjana T, Yong JWH, Letham DS, Griffith M, Hussain M, Ljung K, Sandberg G, Sundaresan V (2001) Control of axillary bud initiation and shoot architecture in Arabidopsis through the SUPERSHOOT gene. Genes Dev 15: 1577–1588
- Underhill EW (1980) Glucosinolates. In: Bell EA, Chairwood BC (eds) Encyclopedia of plant physiology, vol 8. Springer, New York, pp 493–511
- Urban P, Mignotte C, Kazmaier M, Delorme F, Pompon D (1997) Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with Cytochrome P450 CYP73A5. J Biol Chem 272: 19176–19186
- Wallsgrove RM, Bennett RN (1995) The biosynthesis of glucosinolates in *Brassicas*. In: Wallsgrove RM (ed) Amino acid and their derivatives in higher plants. Cambridge University Press, Cambridge, pp 243–259
- Waser J, Watson WH (1963) Crystallography: electron density in long carbon-carbon bonds*.* Nature (London) 198: 1297–1298
- Wetter LR, Chisholm MD (1968) Sources of sulfur in the thioglucosides of various higher plants. Can J Biochem 46: 931–935
- Wittstock U, Halkier BA (2000) Cytochrome P450 CYP79A2 from *Arabidopsis thaliana* L. Catalyzes the conversion of L-phenylalanine to phenylacetaldoxime in the biosynthesis of benzylglucosinolate. J Biol Chem 275: 14659–14666
- Zhang Y, Talalay P (1994) Anticarcinogenic activities of organic isothiocyanates: chemistry and mechanisms. Cancer Res 54: 1976S–1981S

Authors' address: Barbara Ann Halkier, Assoc. Prof., Ph.D., D.Sc., Plant Biochemistry Laboratory, Department of Plant Biology, The Royal Veterinary and Agricultural University, 40, Thorvaldsensvej, DK-1871 Frederiksberg C, Copenhagen, Denmark, E-mail: bah@kvl.dk