Sinapate ester metabolism in Brassica and Arabidopsis

HABILITATIONSSCHRIFT

zur Erlangung des akademischen Grades

doctor rerum naturalium habilitatus (Dr. rer. nat. habil.)

vorgelegt der

Biologisch-Pharmazeutischen Fakultät

der Friedrich-Schiller-Universität Jena

von

Dr. rer. nat. Carsten Milkowski

geb. am 30.07.1963 in Lübz
(Mecklenburg-Vorpommern)

Jena, den 9.April 2009

Gutachter

Prof. Dr. Jonathan Gershenzon
Max-Planck-Institut für Chemische Ökologie
Abteilung Biochemie
Hans-Knöll-Straße 8
07745 Jena

Prof. Dr. Birger Lindberg Møller
Department of Plant Biology
The Royal Veterinary and Agricultural University
Thorvaldsensvej 40
DK-1871 Frederiksberg C, Copenhagen
Denmark

Prof. em. Dr. Thomas Hartmann
Technische Universität Carolo-Wilhelmina zu Braunschweig
Institut für Pharmazeutische Biologie
Mendelssohnstraße 1
38106 Braunschweig

Erteilung der Lehrbefähigung für das Fach „Allgemeine Botanik“

14. Dezember 2009

TABLE OF CONTENTS

1 Introduction

Higher plants display a tremendous metabolic plasticity. This is illustrated by their ability to synthesize myriads of so-called secondary organic compounds ('natural products') that seem to be dispensable for growth and development and are often differentially distributed among limited taxonomic groups (Harborne and Turner 1984). Investigation of these plant products was initiated about 200 years ago by Friedrich Wilhelm Sertürner, who confined the active principle of opium poppy to a single organic substance, known as morphine. Since then, the structures of more than 200,000 plant natural products have been elucidated and assigned to the groups of terpenoids, alkaloids, polyketides or phenylpropanoids and various derived phenolic compounds (Wink 1988; Hartmann 2007).

As functional aspects of plant secondary metabolism were hardly apparent, its products have long been considered to be inert metabolic waste or detoxification products (Peach 1950). This view changed when entomologists recognized secondary metabolites as chemical mediators between plants and insects and established the concept of coevolution (Fraenkel 1959). Eventually, this line of research unravelled the enormous ecological importance of plant secondary compounds.

On the other hand, products of secondary metabolism pathways have also proven functional in internal plant processes. This is ideally illustrated by some terpenoids acting as plant hormones like abscisic acid or gibberellins. Lignin is the second most abundant biopolymer on earth (Buchanan et al. 2000). It provides rigidity and impermeability to secondary thickened cell walls, thus being indispensable for plant structure and defense as well as for the development of xylem elements as an efficient long-distance water transport system. Accordingly, in evolutionary terms, the establishment of the phenylpropanoid metabolic pathway toward lignin was a basic requirement for plants to conquer terrestrial biotopes (Boudet 2007). As products of another branch of the phenylpropanoid pathway, flavonoids are ubiquitously distributed among higher plants and serve a wide array of fundamental functions as revealed by their involvement in attraction of pollinators and seed dispersers, protection against UV-B radiation or the establishment of plant-rhizobium bacteria interactions for nitrogen fixation in legumes. In Arabidopsis, flavonoids have been shown to be involved in auxin transport (Brown et al. 2001). These are only few prominent examples, which indicate that a strict line between primary and secondary metabolism is difficult to be drawn. Hence, in contrast to the traditional meaning of 'less important', the term 'secondary' is now interpreted as an indispensable layer of functionality that is inherent to plant metabolic

networks and contributes significantly to the plasticity of plant metabolism, which is required to afford the sessile life style of a land plant under changing environmental conditions (Hartmann 2007).

Only in recent years, the combined application of the upcoming Omics technologies for gene, protein and metabolite analyses begins to discern the many interactions within the network of secondary metabolism and between secondary and primary metabolism (Böttcher et al. 2008). This paves the way to a better understanding of plant metabolism in its outstanding complexity and will support targeted metabolic engineering approaches to generate plants with altered metabolite contents for food industry or pharmaceutical use (Dixon 2005). The complete transfer of the biosynthetic pathway of the cyanogenic glucoside dhurrin from *Sorghum bicolor* to Arabidopsis making this plant cyanogenic demonstrated both the feasibility and the power of engineering secondary plant metabolism (Tattersall et al. 2001). The most prominent example of such a combinatorial biosynthesis in plants is the so-called 'Golden Rice' (Ye et al. 2000). It was generated by simultaneous expression of a phytoene synthase and a lycopene β-cyclase from *Narcissus pseudonarcissus* as well as a bacterial phytoene desaturase from *Erwinia uredovora* in the rice endosperm to produce β-carotene (provitamin A). Concerning phenylpropanoid metabolism, a fundamental interest in improving the processing efficiency of plant biomass for pulping, forage digestibility and biofuels has produced a wealth of approaches to engineer the amount and composition of lignin in several plant species (reviewed by Vanholme et al. 2008).

The inherent flexibility of plant secondary metabolism is reflected by a remarkable plasticity of plant genomes. Within the Arabidopsis genome 15-20% of genes are predicted to be involved in secondary metabolism (D'Auria and Gershenzon 2005). Accordingly, many enzymes of plant secondary metabolism are organized in gene families that have developed from single or few hypothetical ancestors (Moore and Purugganan 2005). In the last decade, of several enzymes involved in plant secondary metabolism unexpected homologs with functions in primary metabolism have been detected (Ober and Hartmann 1999; Steffens 2000). This has encouraged novel research strategies aimed at understanding the evolution of metabolic diversity.

Our research is focused on the metabolism of soluble sinapate esters. These compounds are a hallmark of Brassicaceae plants (Bouchereau et al. 1991) that enabled us to work with the model plant Arabidopsis. For a long time it has been known that seeds of many Brassicaceae species accumulate considerably high amounts of sinapine, the choline ester of sinapate, as the predominating phenolic compound (Gadamer 1897). Given its antinutritive

impact on the seed protein fraction of the major crop plant *Brassica napus* (oilseed rape; Canola), a low sinapine content has become a major aim of conventional and molecular breeding programmes designed to increase the nutritional value of seeds.

The subsequent chapters describe the progress in gene identification, characterization of enzymes and targeted engineering of sinapate ester metabolism made by our and other laboratories throughout the last decade. With emphasis on *B. napus* and Arabidopsis, the structure and function of the relevant proteins and genes will be discussed and aspects of evolutionary enzyme recruitment will be surveyed.

2 Genes and Enzymes of Sinapate Ester Metabolism

Sinapate (3-(4-hydroxy-3,5-dimethoxyphenyl)prop-2-enoate), the phenolic component of sinapate esters, is produced by the phenylpropanoid pathway (Figure 1). The general part of this metabolism converts the aromatic amino acid L-phenylalanine, which is provided by the plastid-located shikimate pathway, to 4-coumaroyl-CoA. This energy-rich thioester marks an important branchpoint of phenylpropanoid metabolism since it feeds into different types of hydroxycinnamate side-chain reactions (Barz et al. 1985), i.e. extension with formation of additional ring systems (e.g. flavonoids or stilbenes), degradation (e.g. hydroxybenzoates), reduction (e.g. hydroxycinnamyl alcohols fueling lignin biosynthesis), oxidation and lactonization (e.g. coumarins) or conjugation with a wide variety of different primary and secondary compounds to form esters or amides. Within this complex biosynthetic network, the major metabolic path toward sinapate proceeds from 4-coumaroyl-CoA via 3-hydroxylation of coumaroylshikimate (Schoch et al. 2001) and 3-*O*-methylation of caffeoyl-CoA followed by 5-hydroxylation of coniferaldehyde and subsequent 5-*O*-methylation of 5-hydroxy-coniferaldehyde (Meyer et al. 1996; Humphreys et al. 1999). Oxidation of the resulting sinapaldehyde leads to sinapate (Nair et al. 2004). In Brassicaceae plants like Arabidopsis and *B. napus*, the pronounced metabolic flux of sinapate is channelled via sinapoylglucose (1-*O*-sinapoyl-β-D-glucose) to various sinapate esters. Sinapoylmalate (sinapoyl-L-malate), a proven UV-shielding component, accumulates in vacuoles of the leaf epidermal cell layer (Landry et al. 1995; Hause et al. 2002), whereas sinapine (sinapoylcholine) constitutes the major phenolic component in seeds localized to the embryonic tissues (Strack et al. 1983; Shahidi and Naczk 1992). During seed germination, sinapine is hydrolized by an esterase activity (Tzagoloff 1963; Nurmann and Strack 1979). In

the young seedlings, the released sinapate is re-esterified via sinapoylglucose to sinapoylmalate (Lorenzen et al. 1996; Figure 1).

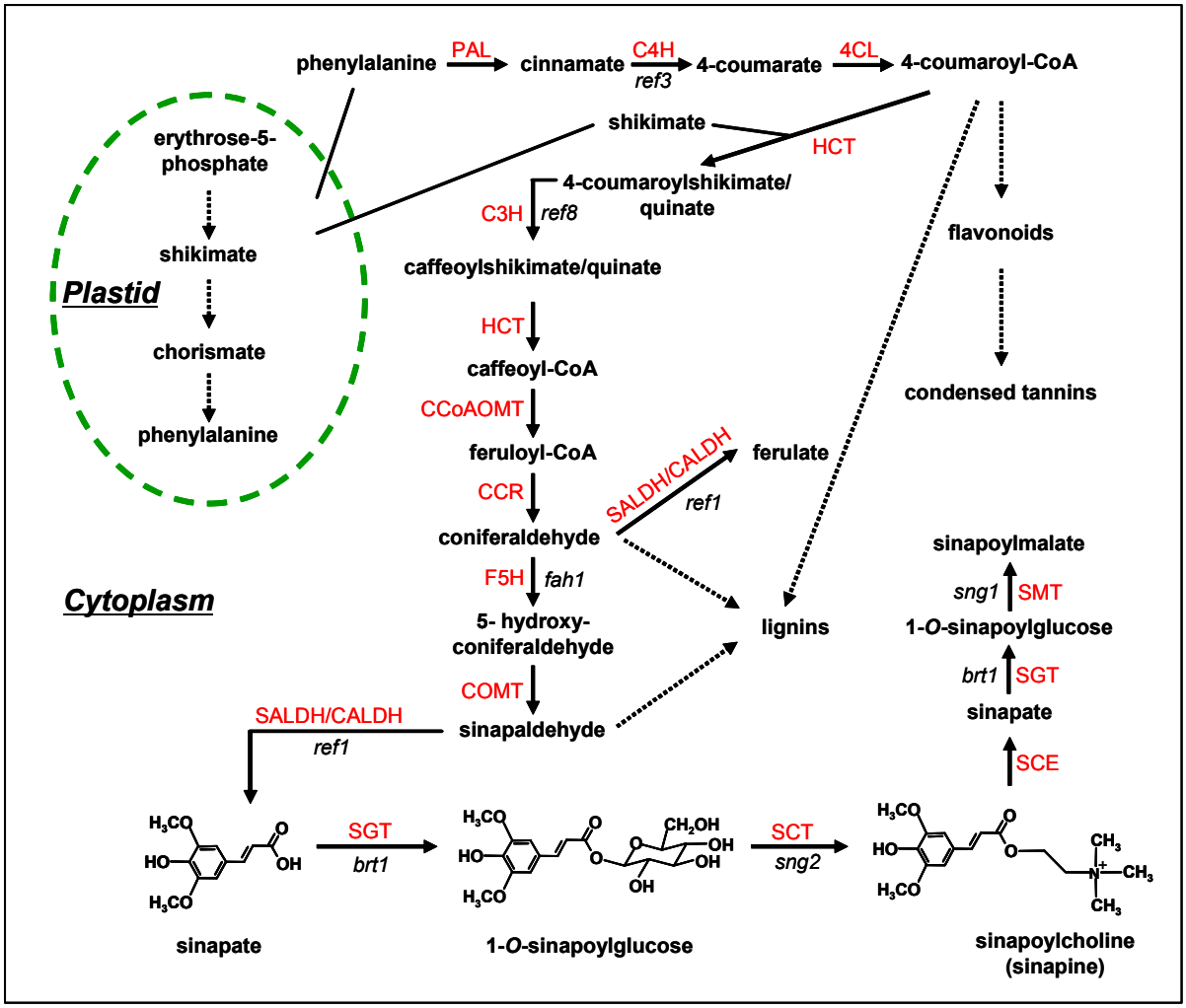


Figure 1 The pathway of sinapate ester biosynthesis in Arabidopsis and Brassica. Aromatic amino acids, including phenylalanine, are synthesized in the plastids via the shikimate pathway. Sinapine accumulates in seeds, and is hydrolized during germination. In seedlings, the liberated sinapate is conjugated via 1-*O*-sinapoylglucose with malate to form sinapoylmalate. Dashed arrows denote multiple enzymatic steps. Lines indicate transport from plastids to the cytoplasm. Enzymes are given in red: PAL, phenylalanine ammonia lyase; C4H, cinnamate-4-hydroxylase; 4CL, 4-coumarate:CoA ligase; HCT, hydroxycinnamoyl-CoA:shikimate/quinate hydroxycinnamoyl transferase; C3H, 4-coumaroyl ester 3-hydroxylase; CCoAOMT, caffeoyl-CoA *O*-methyltransferase; CCR, cinnamoyl-CoA reductase; F5H, ferulate 5-hydroxylase; COMT, caffeate *O*-methyltransferase; SALDH/CALDH, sinapaldehyde/coniferaldehyde dehydrogenase; SGT, UDP-glucose:sinapate glucosyltransferase; SCT, sinapoylglucose:choline sinapoyltransferase; SCE, sinapoylcholine esterase; SMT, sinapoylglucose:L-malate sinapoyltransferase. Names of Arabidopsis mutant alleles are given in italics: *ref*, *reduced epidermal fluorescence*; *brt*, *bright trichomes*; *sng*, *sinapoylglucose accumulator*; *fah*, *ferulate hydroxylase*.

When we started our work, the basic biochemistry of sinapate ester metabolism had been established (Strack et al. 1983; Lorenzen et al. 1996). However, there was a remarkable lack of knowledge about the genes involved and the structures of the encoded enzymes. For the metabolic pathway downstream from sinapate no gene was available. To overcome this discrepancy and pave the way to studies on evolution and metabolic engineering we identified and cloned genes encoding the major enzymes of sinapate ester metabolism, which belong to the families of glucosyltransferases, acyltransferases and esterases.

2.1 Glucosyltransferases

The first commited enzyme in sinapate ester biosynthesis is a glucosyltransferase (UDP-glucose:sinapate glucosyltransferase; SGT, EC 2.4.1.120) that transfers the glucose moiety from UDP-glucose to the carboxyl group of sinapate (Strack 1980; Figure 1). The resulting β-acetal ester 1-*O*-sinapoyl-β-D-glucose (sinapoylglucose) serves as acyl donor in subsequent transacylation reactions thus providing an alternative to CoA-dependent pathways (Mock and Strack 1993).

The pivotal role of sinapoylglucose as sinapoyl donor provoked a fundamental interest in the genes encoding SGT. However, initial attempts to purify SGT from *B. napus* for reverse cloning approaches failed, most likely due to the low abundance of this enzyme in plant tissues (Wang and Ellis 1998). Weak expression has been reported as a general problem associated with plant secondary metabolism UDP-glycosyltransferases (UGTs), rendering purification of these enzymes a challenging task (Hrazdina 1992).

To circumvent these problems, we developed a homology-based cloning strategy that relied on the conserved sequence element designated as **p**lant **s**econdary **p**roduct **g**lycosyltransferase (PSPG) box or UGT signature motif, which mediates recognition and binding of the dinucleotide sugar by the enzyme (Hughes and Hughes 1994). We designed a PCR approach to target the UGT signature sequence in cDNA libraries from developing seeds and young seedlings of *B. napus*, which display high enzymatic SGT activities. Implicating that UGTs, which catalyze the formation of sugar esters instead of *O*-glycosides, would cluster together in a distinct group within the UGT enzyme family, the resulting PCR products were classified by sequence identity to indole-3-acetate-UGT from maize and limonoid-UGT from *Citrus* – the only proven ester-forming UGTs at that time (Szerszen et al. 1994; Kita et al. 2000). Heterologous expression of the isolated full-length cDNAs and enzyme activity

assays with the recombinant protein were used for further characterization. By this approach, we were able to describe the first ester-forming hydroxycinnamate-UGT (HCA-UGT) with a pronounced preference for sinapate as glucose acceptor (Milkowski et al. 2000a). Since the cDNAs isolated from seeds and seedlings were identical, we concluded that SGT activity during seed development and early seedling growth was mediated by the same enzyme. According to the UGT nomenclature (Mackenzie et al. 1997), this glucosyltransferase originally referred to as BnSGT1 was renamed as UGT84A9. By sequence identity with UGT84A9 we predicted four genes from the reference plant Arabidopsis to encode UGTs with SGT activity. Heterologous expression and enzyme assays led us to define At3g21560 (UGT84A2) as the Arabidopsis SGT whereas At4g15480 (UGT84A1), At4g15490 (UGT84A3) and At4g15500 (UGT84A4) were characterized as unspecific ester-forming HCA-UGTs (Milkowski et al. 2000b). Subsequently, our results were confirmed by extensive studies of the full complement of Arabidopsis UGTs (Lim et al. 2001; Li et al. 2001).

The identification of the UGT84A-encoding genes enabled us to initiate a systematic approach aimed at elucidating the physiological role of ester-forming HCA-UGTs *in planta*, beyond their obvious biochemical role as precursors for several HCA esters. Surprisingly, Arabidopsis mutants carrying individual *ugt84A* null alleles were only marginally affected in the gross metabolic flux toward sinapoylmalate in leaves or sinapoylcholine in seeds. Analysis of the phenylpropanoid pattern of *ugt84A2* mutant seeds, however, revealed a slight metabolic redirection toward aromatic choline esters and disinapoylspermidine (Meißner et al. 2008). This is in accordance with the phenotype of the Arabidopsis *brt1* (*bright trichome 1*) mutant impaired in UGT84A2 which still displayed about 70% of the wild-type sinapoylmalate and sinapoylcholine but, most likely, revealed a partial metabolic redirection of sinapate to an unknown polyketide causing increased trichome fluorescence (Sinlapadech et al. 2007). Obviously, the concomitant expression of the four UGT84A enzymes together with their overlapping glucose acceptor specificity confers functional redundancy *in planta* that prevents a distinct mutant phenotype. Our finding that individual overexpression of UGT84A1-A3 led to increased levels of sinapoylglucose in each transgenic line corroborated the assumption of functional redundancy of Arabidopsis UGTs. The absence of related glucose esters like coumaroylglucose or feruloylglucose, which *in vitro* were in the enzymatic product spectrum of the overexpressed UGT84A1-3, indicated a restricted availability of 4-coumarate and ferulate in the Arabidopsis metabolic network (Meißner et al. 2008). Similar results, i.e. only marginally increased product levels or the absence of certain possible reaction products, have been reported for overexpression of an anthocyanin 3'-UGT from

gentiana (Fukuchi-Mizutani et al. 2003) and of the Arabidopsis enzyme UGT71C1 glucosylating the 3-hydroxyl position of caffeate (Lim et al. 2003). Interestingly, we found that the expression of UGT84A1-A4 was strongly induced by enhanced UV-B radiation followed by a transient increase in sinapate ester abundance in Arabidopsis leaves. This suggested a key importance of the ester forming HCA-UGTs in mediating the UV stress response (Meißner et al. 2008). In the transcriptome of *B. napus* seeds, we could not detect homologs of the ester-forming Arabidopsis UGTs 84A1, -A3 and -A4 (Mittasch et al. 2007) although our genomic analyses produced evidence for the occurrence of these genes in the allotetraploid *B. napus* genome (Mittasch 2008). This reveals that biosynthesis of sinapoylglucose in seeds of *B. napus* might be controlled essentially by UGT84A9, whereas in Arabidopsis UGT84A2 and the related glucosyltransferases UGT84A1, -A3 and -A4 are involved (Meißner et al. 2008).

2.2 Acyltransferases

The metabolic pool size of sinapoylglucose in Brassicaceae plants is not only dependent on its biosynthesis but also on its further metabolization to a range of sinapate esters of which sinapoylmalate and sinapoylcholine are the major products in Arabidopsis and *B. napus*. Both of these processes, synthesis and transacylation of sinapoylglucose, are developmentally regulated via differential expression of the genes involved in the sinapate ester pathway (Milkowski et al. 2004; Figure 2). The enzymes channelling sinapoylglucose into the major accumulating sinapate ester compounds have been defined as sinapoylglucose:malate sinapoyltransferase (SMT; EC 2.3.1.92) and sinapoylglucose:choline sinapoyltransferase (SCT; EC 2.3.1.91), respectively (Tkotz and Strack 1980; Strack et al. 1983). By use of the activated glucose ester sinapoylglucose as acyldonor, these enzymes define an alternative to the widespread CoA-thioester dependent acyltransferase reactions. Mature SMT was identified as a vacuolar protein (Hause et al. 2002). Together with the presence of N-terminal signal peptides for vacuolar targeting in related proteins, this indicates that the glucose ester-dependent acyltransfer is most probably confined to the central vacuole of the plant cell and therefore appears spatially separated from the cytosolic path of CoA-thioester-dependent transacylation.

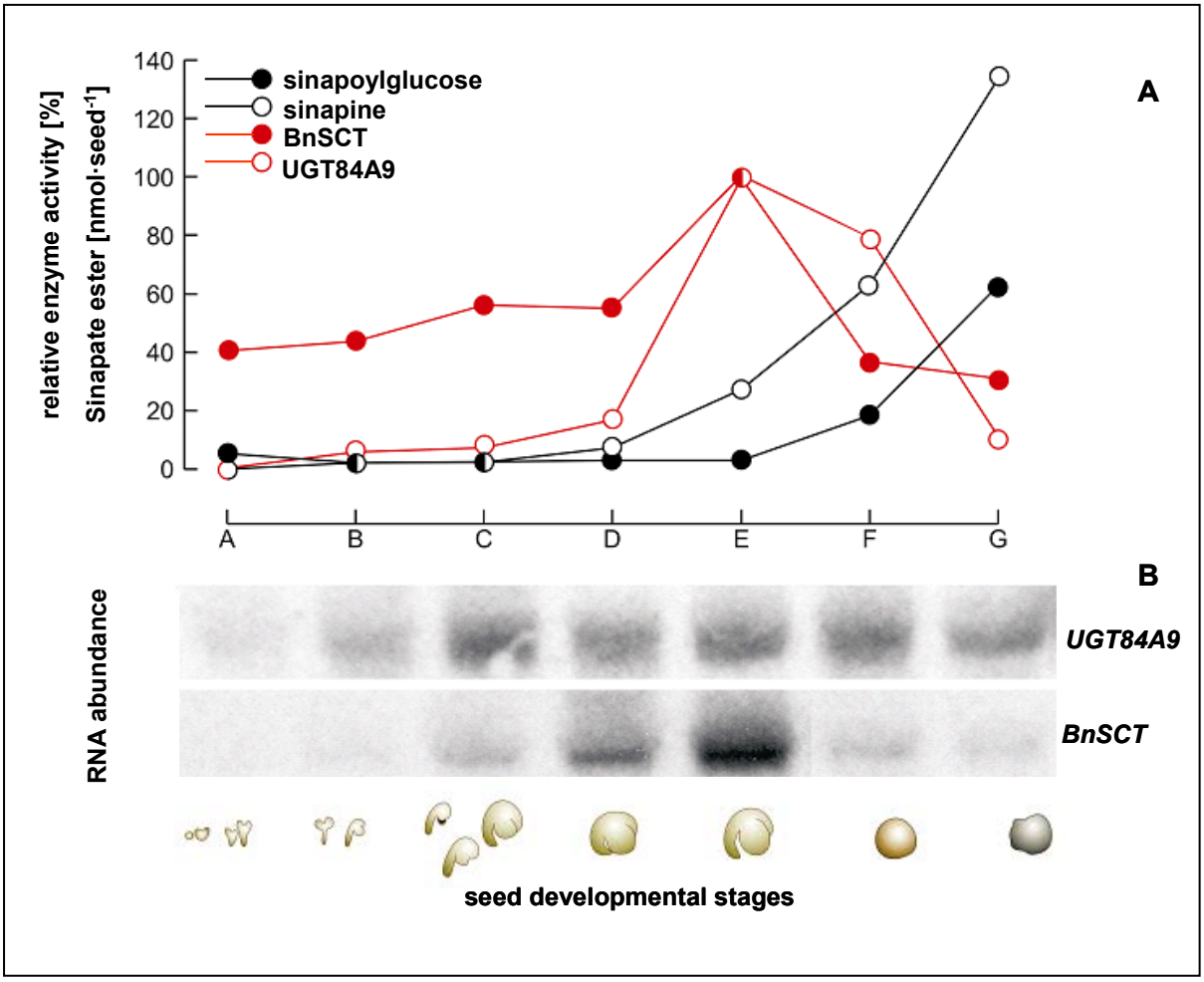


Figure 2 Biosynthesis of sinapoylglucose and sinapine during seed development of *B. napus* (Milkowski et al. 2004; modified). Enzymatic activities of UGT84A9 and BnSCT, depicted in red, were induced during the seed filling phase (seed developmental stages D and E). Sinapine and sinapoylglucose accumulate in dry seeds as shown by the black graphs **(A)**. RNA abundancies assesed by Northern blot analyses reveal regulated expression of *UGT84A9* and *BnSCT* during seed development **(B)**. Seed developmental stages are classified according to embryo and seed morphology: A, globular, torpedo and late torpedo stages; B, early cotyledonary stage to 'walking stick'; C, mid-cotelydonary stage; D, well developed cotyledonary stage; E, well developed embryo; F, seeds with brown seed coat and yellow embryo; G, desiccated seed.

The Arabidopsis genes encoding SMT (*SNG1*, At2g22990) and SCT (*SNG2*, At5g09640) were cloned from mutants of the *sinapoylglucose accumulator* (*sng*) phenotype (Lehfeldt et al. 2000; Shirley et al. 2001). In a parallel approach, we could identify a cDNA encoding the SCT from *B. napus* (Milkowski et al. 2004). Since SCT orthologs were not known, we developed a homology-based cloning strategy by integrating the limited information of only two related sequences from functionally proven ester-dependent acyltransferases – the Arabidopsis SMT (Lehfeldt et al. 2000) and an isobutyroyl transferase

from wild tomato (Li and Steffens 2000). By a stepwise optimization procedure, we developed degenerate primers that allowed amplification of SCT-cDNA fragments from the seed cDNA library of *B. napus*. Combining RACE techniques and genome walking, we eventually isolated a full-length SCT cDNA designated as *BnSCT* from seeds of *B. napus* (Milkowski et al. 2004). Transcripts of *BnSCT* appeared in early seed developmental stages and reached highest abundancies in the phase of seed filling (Figure 2). By expression analysis of *BnSCT* promoter-GUS fusion constructs we could show that *BnSCT* promoter activity was confined to embryo tissues and the aleurone layer (Weier et al. 2007).

2.3 Sinapine esterase

Sinapine, the major phenolic compound of Brassicaceae seeds, is mobilized during seed germination by enzymatic degradation into sinapate and choline (Tzagoloff 1963; Nurmann and Strack 1979). Although initial measurements revealed a broad substrate spectrum, the strong induction of this enzymatic activity upon germination as well as its co-distribution with the seed component sinapine among Brassicaceae species pointed to the existence of a distinct sinapine esterase (SCE; EC 3.1.1.49; Nurmann and Strack 1979). For a long time, however, the nature of this enzyme remained unknown. In the light of the upcoming interest in sinapoyltransferases like SMT or SCT it has been speculated that SCE might represent a hydrolytic relative of these enzymes (Fraser et al. 2005).

Taken together, the previously reported characteristics of SCE did not produce the appropriate information required for the design of promising homology-based cloning strategies. Hence, it became clear that identification of SCE would rely on the conventional reversed strategy of enzyme purification, followed by peptide sequencing and PCR-based cDNA-cloning. For this approach, germinating seeds of *B. napus*, which strongly induce SCE expression, provided a convenient source. Eventually, the combined effort of enzyme purification and subsequent cDNA cloning made by our group led to the identification of *B. napus* sinapine esterase (BnSCE3) as a GDSL lipase-like enzyme (Clauß et al. 2008). The protein identified as BnSCE3 was previously described as a putative unspecific lipase BnLIP2 (Ling et al. 2006). However, this classification was based solely on sequence identity with related lipases without any experimental proof. To produce evidence that BnSCE3 functions as sinapine esterase *in planta*, we generated seed-specific overexpression lines in Arabidopsis and *B. napus*. Transgenic lines differentiated seeds with dramatically reduced sinapate ester

contents due to functional *BnSCE3* expression (Clauß et al. 2008; K. Clauß, personal communication). Thus, the observed loss of sinapine accumulation revealed a role of BnSCE3 as functional sinapine esterase *in planta*. In Arabidopsis, we could identify four *BnSCE3* homologs encoded by a tandemly arranged gene cluster on chromosome I (At1g28640, At1g28650, At1g28660, At1g28670). By expression in *E. coli* and enzyme activity assays we could show that the gene products, except that of At1g28640, displayed SCE activity *in vitro*.

3 Evolution

The outstanding plasticity inherent to plant secondary metabolism is reflected by the development and maintenance of large enzyme families, whose members display distinct or overlapping functions. It is widely accepted that the formation of these enzyme families was brought about by an evolutionary scenario that involves gene duplications followed by subfunctionalization or neofunctionalization of the resulting gene duplicates. Based on the exponentially growing sequence information of whole genomes and advances in unravelling structure-function relationships of enzymes, the evolution of plant metabolic diversity has become one of the challenging topics of modern plant science. The enzymes of sinapate ester metabolism described here, SGT, SCT, SMT and SCE, each belong to distinct clades of large protein families. This allows interesting hypotheses on the evolutionary recruitment of enzymes.

The ester-forming HCA-UGTs including the SGTs from *B. napus* (UGT84A9) and from Arabidopsis (UGT84A2) belong to the class of plant secondary metabolism UGTs. These cluster to family 1 of the glycosyltransferase (GT) enzyme superfamily, which includes about 90 families of GTs (Coutinho and Henrissat 1999). Within family 1 GTs, whose members typically transfer sugars to small-molecule acceptors, the class of plant secondary metabolism UGTs is distinguished by a catalytic mechanism that leads to the inversion of the sugar configuration and by the presence of the PSPG box as conserved binding motif for the dinucleotide-activated sugar donor. Since glucosylation is considered among the most prevalent modifications of secondary metabolites, even small plant genomes encode large UGT families. The Arabidopsis genome harbours 107 open-reading frames for secondary metabolism UGTs, which could be assigned to 14 distinct phylogenetic groups (Ross et al. 2001; Paquette et al. 2003). The ester-forming HCA-UGTs of the UGT84A clade have been classified as members of the phylogenetic group L. Interestingly, other UGTs transferring the

sugar moiety to the acceptor carboxyl groups of different compounds like benzoates or anthranilate were found to cluster consistently in this phylogenetic group L (Lim et al. 2002; Quiel and Benders 2003). In contrast, enzymes catalyzing *O*-glucoside formation were found to be scattered over the other phylogenetic UGT groups. This classification seemed robust enough to allow sequence-based predictions of UGT functions as *O*-glucosylation or ester formation activity. Among the UGTs characterized so far, no enzyme displayed both glucosylation and ester-forming activity with comparable efficiency. To reconstruct the evolutionary events toward differentiation between *O*-glucosylating and ester-forming UGTs would require resolved UGT structures to unravel the molecular mechanisms that direct the formation of esters or *O*-glucosides. Crystal structures for group L UGTs, however, remain to be solved.

The phylogenetic clustering of the sinapoyltransferases SMT and SCT and of sinapine esterase BnSCE3 provided surprising insights into the relation of sinapate ester metabolism with hydrolytic branches of primary metabolism. Sequence analysis indicated that both SMT and SCT shared homology with serine carboxypeptidases (SCPs), a group of hydrolytic enzymes cleaving peptide bonds from the C-terminus of proteins or peptides (Lehfeldt et al. 2000; Shirley et al. 2001). Such unexpected homology had been established before with an isobutyroyl transferase from wild tomato (*Lycopersicon penellii*) involved in the formation of glucose polyesters (2,3,4-tri-*O*-acylglucose; Li and Steffens 2000). Based on these evolutionary relationships, we performed a first phylogenetic analysis of Arabidopsis **s**erine **c**arboxy**p**eptidase **l**ike (SCPL) proteins (Milkowski and Strack 2004). As a result, we identified a distinct cluster within the SCPL protein family that included, besides SMT and SCT, the isobutyroyl transferase from *L. penellii* as another glucose ester-dependent acyltransferase. Functional SCPs like serine carboxypeptidase II from wheat (*Triticum aestivum*; Breddam et al. 1987) and carboxypeptidase Y from yeast (*Saccharomyces cerevisiae*; Svendsen et al. 1982) were found outside this cluster. The functional characterization of the Arabidopsis SCPL protein 24 as a serine carboxypeptidase II (BRS1; Zhou and Li 2005) was in accordance with our hypothesis that acyltransferases and hydrolases form different clades within the SCPL protein family (Figure 3). These observations produced first evidence for a plant-specific group of glucose ester-dependent acyltransferases developed by functional recruitment of a hydrolytic SCPL ancestor enzyme in the course of evolution. A subsequently published phylogenetic analysis largely confirmed our results (Fraser et al. 2005). In summary, it can be concluded that the genome of Arabidopsis encodes a family of 51 SCPL proteins of which 19, including SMT and SCT,

form a separate clade defined by glucose ester-dependent acyltransferases (clade IA SCPL proteins). Recently, other members of this clade were functionally characterized as sinapoylglucose:anthocyanin sinapoyltransferase (SAT; At2g23000) and sinapoylglucose:sinapoylglucose sinapoyltransferase (SST; At2g23010) (Fraser et al. 2007). The specificity for sinapoylglucose as acyl donor shared by all of the investigated clade IA members gives rise to the assumption that the related still uncharacterized enzymes of this group act as sinapoylglucose-dependent acyltransferases as well. All molecular data available support the hypothesis that evolution of these SCPL acyltransferases involved duplication of an ancestor gene encoding a hydrolytic SCPL enzyme. Subsequent neofunctionalization led to a shift from hydrolase to acyltransferase activity, which was followed by subfunctionalization to reach divergence of substrate specificity (Fraser et al. 2007).

For sinapine esterase BnSCE3, we found homology with lipid hydrolases (Clauß et al. 2008). Accordingly, BnSCE3 and its Arabidopsis relatives were classified as members of the SGNH hydrolase subfamily of GDSL-lipases which cleave ester bonds in a wide array of lipids (Akoh et al. 2004). This clustering revealed that BnSCE3 had most likely developed by evolutionary recruitment from ancient lipases. In contrast to the sinapoyltransferases SCT and SMT, the sinapine esterase BnSCE3 obviously kept the catalytic mechanism of the hydrolytic ancestor but experienced a change in substrate specificity. We could show that BnSCE3 was active toward several phenolic esters *in vitro*, although with preference for sinapine. The substrate specificity, however, was not directed toward the aromatic moiety but rather toward the C3 side chain and the nature of the conjugating moiety, predominantly choline (Clauß et al. 2008). In fact, BnSCE3 mediated esterase activity toward phosphatidylcholine pointing to a phospholipase A-like activity (Wang 2001). The characteristic feature of GDSL lipase-like enzymes to accept an extraordinarily broad substrate spectrum for hydrolysis was also demonstrated with thioesterase I (TAPI) from *E. coli* that is – like BnSCE3 – a member of the SGNH hydrolases. TAPI displayed multifunctional activity as protease, thioesterase, arylesterase and lysophospholipase (Lee et al. 1997). At molecular level, this catalytic multifunctionality was explained by a flexible active site allowing several different substrates to bind in optimal conformations for catalysis to occur (Huang et al. 2001).

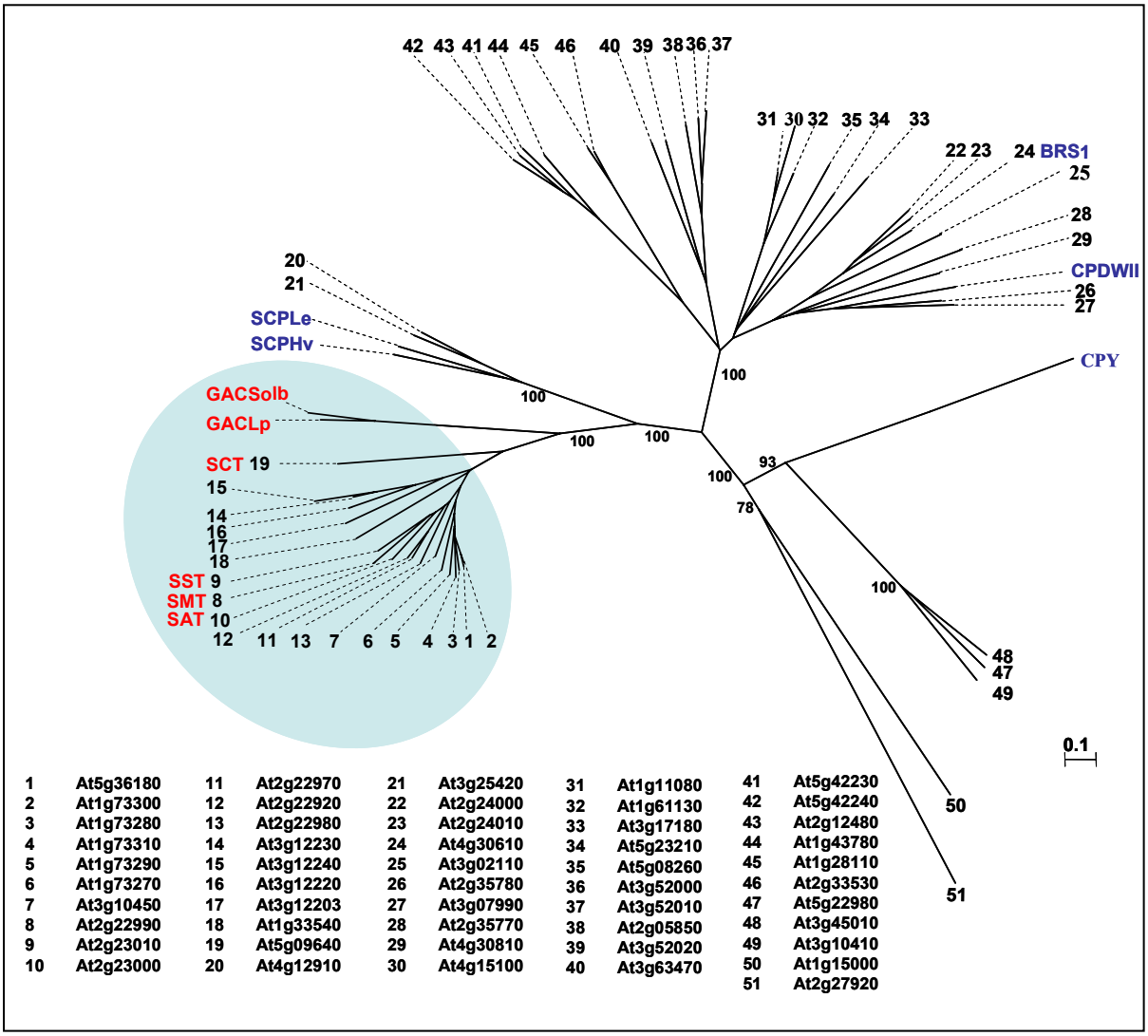


Figure 3 Phylogenetic tree of the 51 Arabidopsis SCPL proteins with functionally characterized acyltransferases labelled in red and carboxypeptidases written in blue (Milkowski and Strack 2004; Fraser et al. 2005; modified). The tree was derived from a multiple alignment of amino acid sequences by neighbour joining algorithm. The scale represents 0.1 fixed mutations per site. Arabidopsis proteins were numbered according to Fraser et al. (2005). Gene names are given below. The clade IA is marked in light blue. BRS1, *bri1* (*brassinosteroid insensitive 1*) *suppressor* (SwissProt accession Q9M099); GACSolb, glucose acyltransferase from *Solanum berthaultii* (GenBank accession AF006078); GACLp, isobutyroyl transferase from *Lycopersicon penellii* (Q9LKY6); SST, sinapoylglucose:sinapoylglucose sinapoyltransferase (Q3EBV8); SMT, sinapoylglucose:malate sinapoyltransferase (Q8RUW5); SCT, sinapoylglucose:choline sinapoyltransferase (Q8VZU3); CPY, serine carboxypeptidase from *Saccharomyces cerevisiae* (P00729); CPDWII, serine carboxypeptidase II from *Triticum aestivum* (P008819); SCPLe, serine carboxypeptidase I from *Solanum lycopersicum* (Q9M513); SCPHv, serine carboxypeptidase I from *Hordeum vulgare* (P07519)

The genes encoding major enzymes of sinapate ester metabolism were frequently found in tandemly arranged clusters. In some cases, enzymes of particular clusters share partly overlapping specificities. This emphasizes the role of gene duplication in the evolution of these enzymes and might support the hypothesis that the branch of sinapate ester metabolism has developed recently in terms of evolution. Thus, an enzymatic activity to form 1,2-di-sinapoylglucose (SST activity) has been proposed for three of the five SCPL acyltransferases clustered on Arabidopsis chromosome II including the Arabidopsis SMT (AtSMT; Fraser et al. 2007). Our work dedicated to a comprehensive *in vitro* enzyme characterization extended the concept of overlapping enzyme functions within this SCPL acyltransferase cluster. In assays lacking the preferred sinapoyl acceptor L-malate, we found that AtSMT catalyzed two reactions - the disproportionation and the hydrolysis of 1-*O*-sinapoylglucose (Stehle et al. 2008a). Given the high level of sequence identity, we assume that other enzymes of this cluster use 1-*O*-sinapoylglucose as substrate and may have kept hydrolytic esterase side activity as well. As a remnant of the ancient precursor, such hydrolytic activity would support the predicted ancestry of SCPL acyltransferases from hydrolases. The close functional relatedness of clustered clade IA SCPL enzymes was also demonstrated by the fact that we were able to engineer another uncharacterized enzyme encoded by At2g22970 to acquire SMT activity by few amino acid replacements. Among the glucosyltransferases and esterases involved in sinapate ester metabolism, we detected tandemly arranged gene clusters with overlapping functions for the Arabidopsis HCA-UGTs 84A1, -3 and -4 on chromosome IV (Milkowski et al. 2000b) and for the BnSCE3 homologs on chromosome I (Clauß et al. 2008).

With regard to the crop species *B. napus*, however, the gene organization reflects a more complex genome evolution coined by additional rounds of polyploidization and hybridization of diploid ancestor genomes to form the allotetraploid. After its separation from the Arabidopsis branch, the genus Brassica experienced a genome triplication (Lagercrantz and Lydiate 1996). Given the genome duplication, brought about by hybridization of the diploid ancestors *B. rapa* (AA, 2n=20) and *B. oleracea* (CC, 2n=18) to form the allotetraploid genome of *B. napus* (AACC, 2n=38), six copies of a single gene locus from Arabidopsis could be expected in oilseed rape. However, in plant evolution genome expansion by polyploidization was commonly counterbalanced by subsequent gene loss. Recent studies suggest that about 35% of the initial gene copies from the triplicated Brassica genome have been lost by deletion mechanisms (Town et al. 2006). This led us to initiate a detailed genomic BAC analysis to identify the *UGT84A9* loci in the allotetraploid genome of *B. napus*.

Using the cultivar ‘Express’ we detected four distinct *UGT84A9* genomic loci, two of which could be traced back to the diploid ancestor genomes of *B. oleracea* (*UGT84A9a* and *UGT84A9c*) and *B. rapa* (*UGT84A9b* and *UGT84A9d*). The four *UGT84A9* genes represent two distinct sequence types. *UGT84A9a* and *UGT84A9b* belong to sequence type 1, whereas *UGT84A9c* and *UGT84A9d* represent sequence type 2. This indicates that *UGT84A9a* and *–b* are orthologous genes, as well as *UGT84A9c* and *–d.* Accordingly, *UGT84A9a/–c* and *UGT84A9b/-d* are considered paralogous gene pairs (Mittasch 2008; Figure 4 A).

During evolution, besides gene loss and sub- or neofunctionalization, a ‘regulatory hypofunctionalization’ became apparent for some paralogous gene copies marked by a strongly reduced expression in all plant organs (Duarte et al. 2006). Our results with *UGT84A9* genes produced evidence for such hypofunctionalization in *B. napus*. In the cultivar “Express”, only sequence type 1, i.e. the genes *UGT84A9a* and *UGT84A9b*, were found to be expressed significantly in all plant organs whereas expression of the according paralogs of sequence type 2, *UGT84A9c* and *–d*, was barely detectable (Figure 4 B). In addition to the evolutionary context, these results are highly relevant for molecular breeding approaches since they define *UGT84A9a* and *UGT84A9b* as the UGT genes involved in sinapine accumulation during seed development.

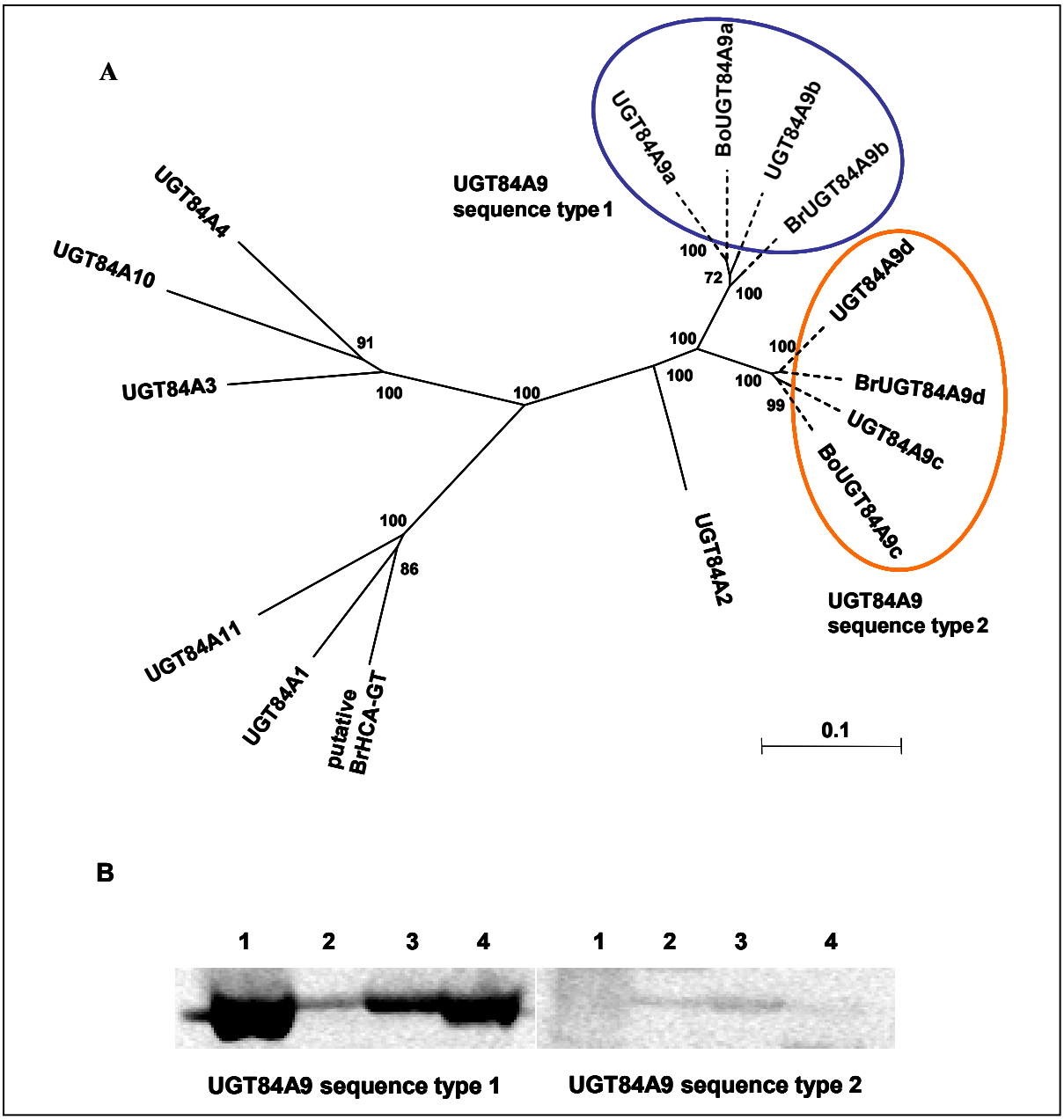


Figure 4 Phylogenetic relationships and expression patterns of *UGT84A9* genes (Mittasch 2008; modified). The phylogenetic tree was derived by neighbour joining algorithm from a multiple alignment of nucleotide sequences (**A**). The scale represents 0.1 fixed mutations per site. Clades comprising the different *UGT84A9* sequence types are highlighted in blue (*UGT84A9* sequence type 1) and orange (*UGT84A9* sequence type 2). Bo, *Brassica oleracea*; Br, *Brassica rapa*. Expression analysis by cDNA-AFLP identified sequence type 1 as the predominantly expressed *UGT84A9* variant in *B. napus* (**B**). RNA abundance was quantified in the following tissues of *B. napus*: seedlings at 2 days after germination (lane 1); flowers (2); seeds at cotyledonary stage (3); seeds at well developed mature embryo stage (4).

The genomic fine structure of the four *UGT84A9* loci revealed the presence of several transposon-derived sequence elements and, on the other hand, only marginal remnants of collinearity with the homologous genomic region of Arabidopsis chromosome III (Figure 5).

This indicates that the *UGT84A9* loci are localized to dynamic genome regions, which underlie increased reorganization. By comparative analyses with a summer type Canola line we produced evidence that the presence of four distinct *UGT84A9* loci is a common feature of *B. napus* cultivars (Mittasch 2008). However, with regard to genomic fine structure and nucleotide sequence, the results from cultivar 'Express' could not be directly transferred to other cultivars.

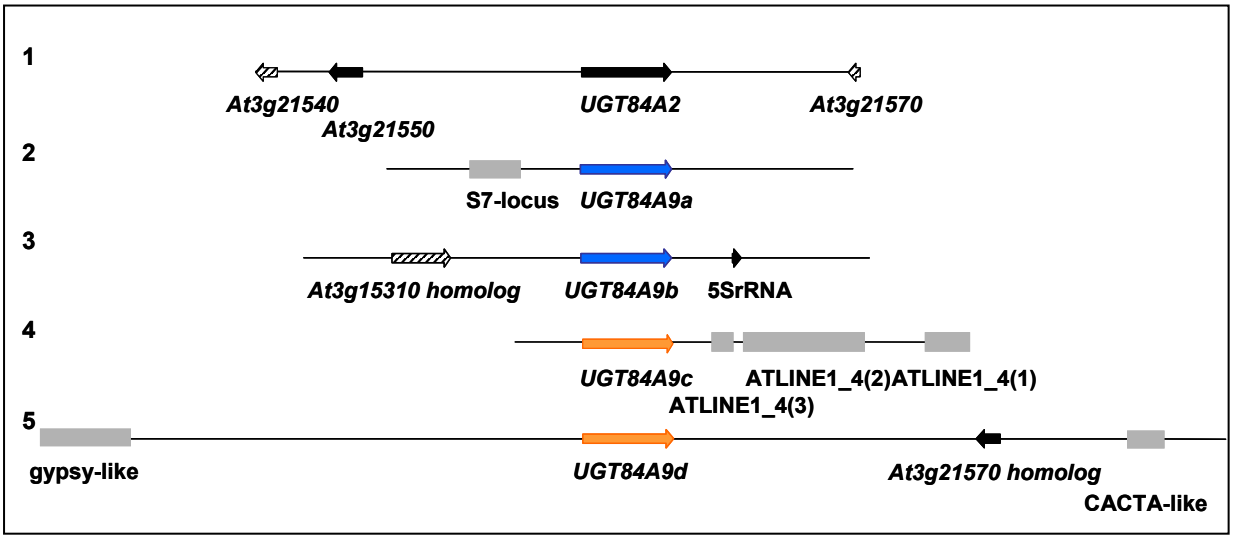


Figure 5 Microstructures of the *UGT84A9* genomic regions from *B. napus* compared to the homologous *UGT84A2* locus from Arabidopsis (Mittasch 2008; modified). Genes are shown as pentagons pointing in the direction of transcription. Hatched pentagons represent partial genes. The *UGT84A9* orthologs of sequence type 1 (*UGT84A9a*, *UGT84A9b*) are shown in blue; those representing sequence type 2 (*UGT84A9c*, *UGT84A9d*) are marked in orange. Transposon-derived elements are shown in gray. 1, genomic *UGT84A2* locus from Arabidopsis chromosome III; 2-5, homologous *UGT84A9* loci isolated from a genomic BAC library of *B. napus*

The shared evolutionary ancestry of the sinapoyltransferases SMT and SCT with serine carboxypeptidases raised the question on molecular mechanisms that governed the functional shift from hydrolase to acyltransferase activity within the SCPL protein family. SCPL proteins belong to the exceptionally large and diverse family of α/β-hydrolase fold enzymes, which carry the α/β-hydrolase fold as a central catalytic domain of unique topology and three-dimensional structure (Ollis et al. 1992). In the active center of these enzymes, the catalytic triad is built by hydrogen bonds between a nucleophile (Ser), an acid (Asp) and a base (His) (Dodson and Wlodawer 1998). Substrate recognition and binding is mediated by a hydrogen bond network (Liao et al. 1992; Endrizzi et al. 1994). It has long been known that some α/β-hydrolase fold enzymes bear the inherent potential for catalyzing transfer reactions.

Accordingly, yeast carboxypeptidase Y was shown to catalyze transpeptidation and transesterification reactions under non-physiological alkaline conditions (Widmer and Johansen 1979). More recently, bi-functional enzymes have been described that mediate transacylation and hydrolysis like indole-3-acetyltransferase from *Zea mays* (Kowalczyk et al. 2003). With regard to evolution, the catalytic versatility of α/β-hydrolases might have converted ancient members of this family into eligible candidates for acquiring the acyltransfer function upon selective pressure.

To unravel molecular mechanisms responsible for the functional shift from hydrolase to acyltransferase activity we choosed the AtSMT enzyme (Lehfeld et al. 2000) for detailed analysis. Applying model structures and site-directed mutagenesis, we could show that the essential functional elements of hydrolytic SCPs – catalytic triad, oxyanion hole and hydrogen bond network for substrate recognition – were adopted by AtSMT to carry out the acyltransferase reaction (Stehle et al. 2006). This led us to conclude that AtSMT applied the same basic chemistry as the hydrolytic SCPs (Figure 6). Accordingly, the alkoxide ion derived from the catalytic Ser by proton abstraction mediates the nucleophilic attack onto the carbonyl carbon of the acyl donor sinapoylglucose. This leads to the liberation of the glucose moiety and formation of an acylenzyme intermediate (sinapoylated enzyme). The latter is cleaved by L-malate as nucleophile resulting in the formation of the transacylation product sinapoylmalate and regeneration of the active site seryl residue. These conversions proceed via a negatively charged tetrahedral transition state that is stabilized by the oxyanion hole. The model structure developed for AtSMT supports the proposed reaction mechanism (Figure 7). Besides catalytic triad, hydrogen bond network for substrate recognition and oxyanion hole, it emphasizes the amino acid residue Arg322 as crucial for substrate binding and positioning of the acyl acceptor L-malate.

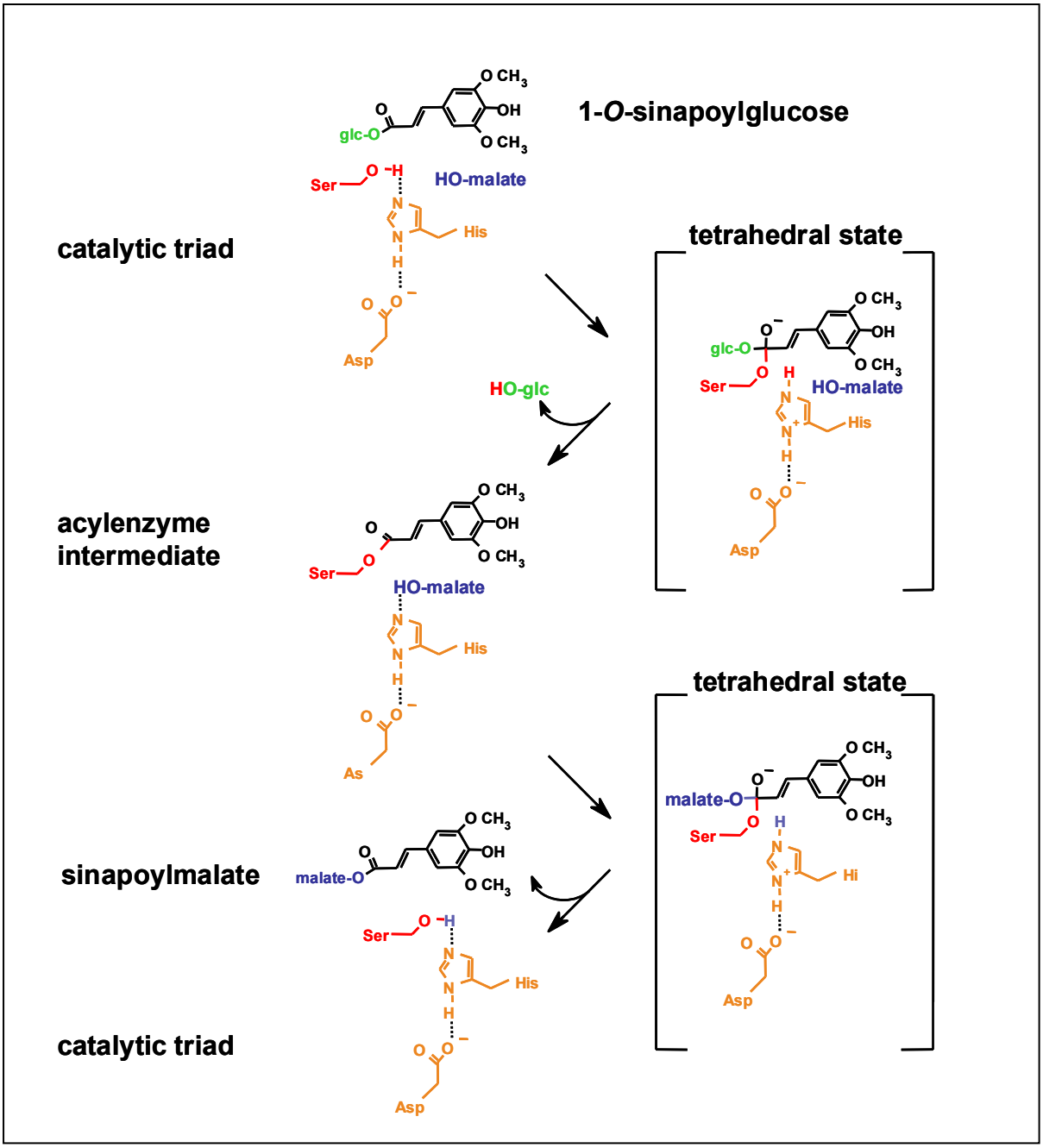


Figure 6 Proposed reaction scheme of AtSMT involving an acylenzyme intermediate (Stehle 2008; modified). Members of the AtSMT catalytic triad are marked in red (Ser) and orange (His, Asp). Of the acyl donor, 1-*O*-sinapoylglucose, the sinapoyl moiety is labelled in black and the glucose moiety in green. The acyl acceptor L-malate is shown in blue.

By sequence comparison between SCPL acyltransferases and proven SCPs we identified two highly conserved substitutions that were localized to the characteristic serine peptidase motif surrounding the catalytic seryl residue (Ser*). Accordingly, this pentapeptide motif present in hydrolytic SCPs as Gly-**Glu**-Ser*-Tyr-**Ala** (Derewenda et al. 1992) was found to be changed to Gly-**Asp**-Ser*-Tyr-**Ser** in the vast majority of predicted SCPL acyltransferases. Interestingly, the Glu/Asp position preceding the catalytic Ser forms part of

the hydrogen bond network for substrate recognition in SCPL proteins (Figure 7). As a functional implication, our Arabidopsis SMT structure model suggests that the smaller Asp side chain at this position affords the necessary space to accommodate the glucose moiety of the acyl donor (Stehle et al. 2006). Therefore, the substitution of Glu by Asp might have been essential in the evolution of substrate specificity of SCPL acyltransferases toward 1-*O*-glucose esters as acyl donors.

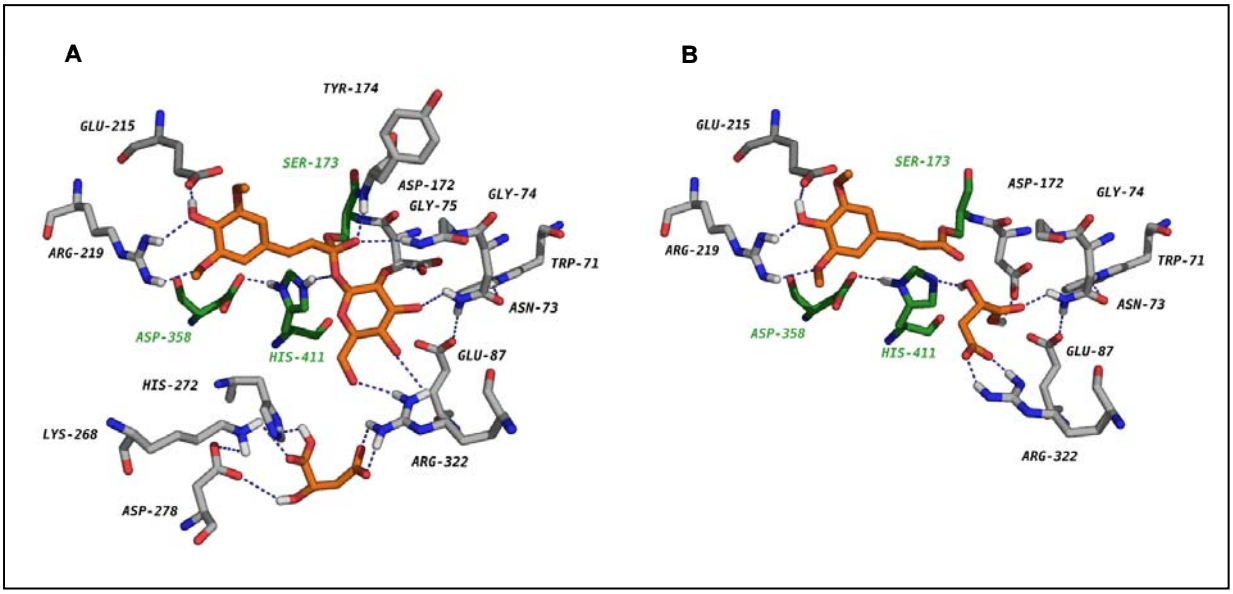


Figure 7 The active site of AtSMT as revealed by the model structure (Stehle et al. 2006; modified). In the tetrahedral transition state the negative charge of the carbonyl carbon is stabilized by the backbone amino groups of Gly75 and Tyr174 contributing to the oxyanion hole. Substrate recognition is facilitated by the hydrogen bond network (Trp71, Asn73, Gly74, Asp172) (**A**). Formation of the acylenzyme is accompanied by a conformational change of Arg322 that enables L-malate to be activated as a nucleophile due to proton abstraction from the hydroxyl group by the catalytic His411 (**B**). Amino acid residues of the catalytic triad (Ser173, His411, Asp358) are marked in green. The substrates sinapoylglucose and L-malate are highlighted in orange. Hydrogen bonds are visualized by dashed lines.

As one critical point in understanding the functional shift from hydrolase to acyltransferase activity, the molecular mechanism that favours transesterification over hydrolytic ester cleavage of the acyldonor needs explanation. Without crystal structures available, we used kinetic studies to elucidate the mechanism of the glucose ester-dependent acyltransfer catalyzed by AtSMT (Stehle et al. 2008b). Applying recombinant AtSMT that was purified from an optimized yeast expression system, we produced kinetic data that were consistent with a random sequential bi-bi mechanism (Schellenberger 1989). This mechanism requires the presence of both sinapoylglucose and L-malate in an enzyme-donor-acceptor

complex before transacylation starts. In accordance with the AtSMT model structure, we propose the formation of a very short-lived acylenzyme that would be accompanied by a conformational change of Arg322 moving L-malate to a position, which favours the proton abstraction by the catalytic His411 and thereby the nucleophilic attack onto the acylenzyme (Figure 7). This scenario would emphasize the facilitation of substrate proximity as major mechanism to prevent ester hydrolysis instead of active water exclusion by shielding the active site (Stehle et al. 2006, 2008b). Our results with recombinant AtSMT were in accordance with previously reported kinetic data on SMT from *Raphanus sativus* (Gräwe et al. 1992). However, with regard to the whole group of glucose ester-dependent acyltransferases, the kinetic measurements reveal some contradiction. For SCT from *B. napus* and Arabidopsis, a double displacement ping-pong mechanism for acyl transfer was reported, which is similar to that applied by the homologous SCPs for peptide hydrolysis (Vogt et al. 1993; Shirley and Chapple 2003). This would require a hydrophobic shield at the active center to prevent attack of water onto the acylenzyme. Since this assumption could not be supported by our SCT structure model (Stehle et al. 2006), further studies of the reaction kinetics of SCT remain essential.

As sinapoylated compounds have been frequently reported from plant metabolomes, several other putative functions for sinapoylglucose-dependent acyltransferases of the SCPL type can be hypothesized. These involve sinapoylation of several flavonoid glycosides and gentiobiose in seeds of *B. napus* (Baumert et al. 2005) as well as the sinapoyltransfer to a wide range of anthocyanins present in Brassicaceae plants (e.g. Hrazdina et al. 1977; Honda et al. 2005) or in wild carrot (*Daucus carota*) where sinapoylglucose was identified as acyl donor (Gläßgen and Seitz 1992). On the other hand, several acyltransferases of the BAHD family (St Pierre and De Luca 2000) have been identified as anthocyanin acyl transferases in plant species like *Gentiana triflora* (Fujiwara et al. 1998), *Perilla frutescens* (Yonekura-Sakakibara et al. 2000) and *Arabidopsis thaliana* (Luo et al. 2007). The identified anthocyanin acyltransferases of the BAHD family accept the CoA-thioesters of 4-coumarate, ferulate and caffeate, but not sinapoyl-CoA (Fujiwara et al. 1998; Luo et al. 2007). Accordingly, anthocyanin sinapoylation in plants may be catalyzed by SCPL acyltransferases and would therefore be localized to vacuoles whereas the acylation of anthocyanins with other hydroxycinnamates can be mediated alternatively by BAHD acyltransferases in the cytosol or SCPL acyltransferases in the vacuole (Fraser et al. 2007). Further insight into possible functions of SCPL acyltransferases could be gained from investigation of plant secondary metabolism pathways known to involve glucose ester-dependent acyltransfer reactions. These

include the biosynthesis of chlorogenic acid in *Ipomoea batatas* (Villegas and Kojima 1986) or of 2-*O*-acetyl-3-*O*-(p-coumaroyl)-*meso*-tartrate in *Spinacia oleraceae* (Strack et al. 1987). Other metabolic pathways to be considered involve the conversion of 1-*O*-hydroxcinnamoylglucoses to celosianin I (4-coumaroylamaranthin) and celosianin II (feruloylamaranthin) in *Chenopodium rubrum* (Bokern and Strack 1988; Bokern et al. 1991) and the formation of 4-coumarate- and caffeate esters of glucaric acid in *Cestrum elegans* (Strack et al. 1988). The above mentioned biochemical pathways and the related plant species may provide promising sources to isolate other members of the β-acetal ester-dependent acyltransferases, thereby extending the knowledge on function and evolution of these interesting enzymes of plant secondary metabolism.

4 Metabolic engineering

The accumulation of sinapine is a prominent trait of various Brassicaceae seeds. In Arabidopsis and *B. napus* this phenolic compound accumulates up to 2% (w/w) of the seed dry mass. Within the seed, the distribution of sinapine is almost restricted to the embryo (Bell 1993; Velasco and Möllers 1998). The subcellular localization of sinapine has not been elucidated, so far. Besides the predominating sinapine and varying amounts of its metabolic precursor, sinapoylglucose, the presence of related phenolic seed compounds in minor concentrations has been described. For Arabidopsis, the latter comprise about 20 phenolic choline esters including substituted hydroxycinnamoylcholines and hydroxybenzoylcholines (Böttcher et al. 2008). In *B. napus*, 14 different sinapate esters, mainly sinapoylated glucose, gentiobiose and kaempferol glycosides, have been detected (Baumert et al. 2005). Together with the relatively high fiber content, the sinapate esters contribute to antinutritive characteristics of the *B. napus* seed protein fraction as revealed by bitter taste, astringency and low digestibility (Kozlowska et al. 1990; Shahidi and Naczk 1992). Therefore, reduction of the sinapate ester quantity in seeds to establish the 'low sinapine trait' is a major goal in oilseed rape breeding. Based on the identification of the major genes for biosynthesis (*UGT84A9*; *BnSCT*) and degradation (*BnSCE3*), we developed strategies of metabolic engineering to reduce the amount of sinapate esters in seeds of *B. napus*. These strategies included suppression of sinapate ester biosynthesis, metabolic diversion of precursors and induced degradation of accumulating compounds.

4.1 Suppression of biosynthesis

We initiated dsRNAi-mediated strategies to silence *UGT84A9* and *BnSCT* in developing seeds. This led to transgenic *B. napus* lines that were homozygous for a single copy insertion of the *UGT84A9*-RNAi cassette and showed up to 76% reduction of the total sinapate ester content accompanied by a decrease in sinapine of about 70% relative to control plants (Hüsken et al. 2005a). Comprehensive analysis of transgenic seeds revealed that both the amount of sinapoylglucose and that of the minor sinapate esters were drastically reduced to trace amounts below the HPLC detection limit (Figure 8). An unusual cyclic spermidine amide, however, was not affected in the *UGT84A9*-suppressing seeds (Baumert et al. 2005). Interestingly, there was no indication of a possible compensative metabolic redirection of sinapoyl moieties into other conjugates of the soluble fraction. This net loss of sinapoyl moieties might point to a hypothetical negative metabolic feedback effect on sinapate biosynthesis triggered by the reduced sink pathway toward sinapate esters in the *UGT84A9*-suppressing seeds. To unravel the complex metabolic response in *UGT84A9*-suppressing seeds, a comprehensive metabolome analysis complemented by quantification of metabolic fluxes will be required. In summary, our results corroborate the pivotal role of sinapoylglucose as universal precursor of sinapate esters in seeds of *B. napus* and confirm the importance of *UGT84A9a* and *UGT84A9b* as major target genes for molecular breeding approaches toward low sinapine lines. We found that the suppression effect was stably inherited as revealed by the low seed sinapate ester phenotype still maintained in the seventh transgenic generation. A concomitant maintenance of hairpin-RNA mediated silencing through successive transgenic generations has been reported for cotton as another important crop plant (Townsend and Llewellyn 2007). Local and temporal specificity of silencing could be obtained by inducible expression of the dsRNAi constructs. In *B. napus*, we used the seed-specific napin promoter and found the post-transcriptional silencing of *UGT84A9* restricted to seed development and early seedling growth (Mittasch 2008). Similar results with inducible or organ-specific promoters have been reported for tobacco, Arabidopsis, *B. napus* and cotton (Chen et al. 2003; Byzova et al. 2004; Sunilkumar et al. 2006).

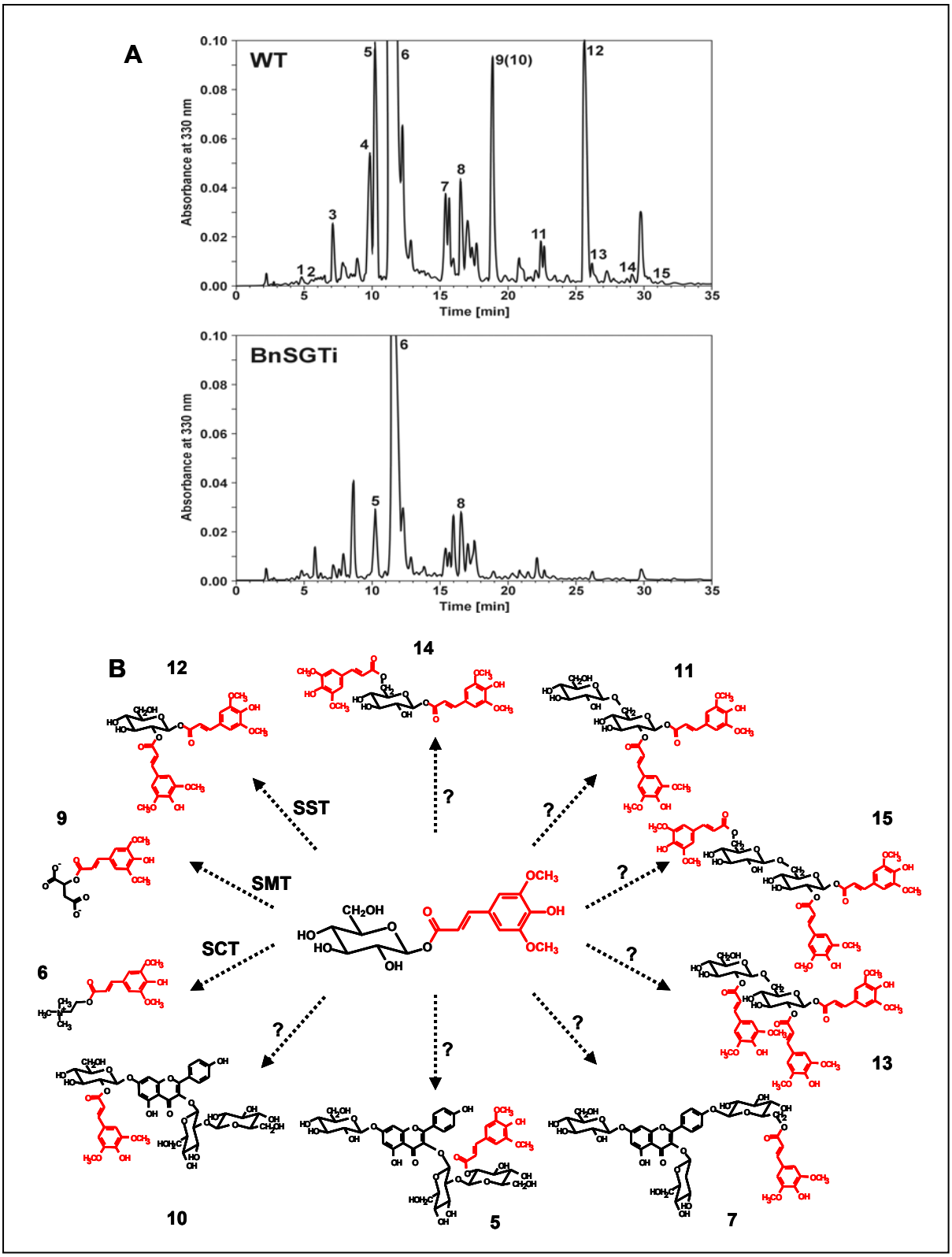


Figure 8 Effect of *UGT84A9* suppression on sinapate ester content of *B. napus* seeds (Baumert et al. 2005; modified). **A**) HPLC traces of methanolic seed extracts from wild type (WT) and *UGT84A9*-suppressing RNAi plants (BnSGTi). **B**) Structures of sinapate esters from *B. napus* seeds emphasizing the pivotal role of sinapoylglucose as metabolic precursor. Arrows indicate acyltransferase activities. HPLC peaks and corresponding structures are assigned by identical numbers. SCT, sinapoylglucose:choline sinapoyltransferase; SMT, sinapoylglucose:malate sinapoyltransferase; SST, sinapoylglucose:sinapoylglucose sinapoyltransferase; ?, putative SCPL acyltransferase. Sinapoyl moieties are marked in red.

Interestingly, the physiological importance of UGT84A9 as a key enzyme of sinapate ester biosynthesis in *B. napus* could not be confirmed for the homologous Arabidopsis glucosyltransferase UGT84A2. By analysis of a *ugt84A2* null mutant we found only moderately decreased levels of soluble sinapate esters, most likely due to a pronounced functional redundancy brought about by the other members of the UGT84A clade in Arabidopsis (Meißner et al. 2008). Accordingly, the UGT84A2-deficient mutant *brt1* showed a weak *reduced epidermal fluorescence* (*ref*) phenotype, which was indicative of a scarcely decreased sinapoylmalate content (Ruegger and Chapple 2001; Sinlapadech et al. 2007).

In Arabidopsis, suppression of SCT, the final enzyme in sinapine biosynthesis, resulted in a reduction of seed sinapine content down to 48% of the wild type (Weier et al. 2007). However, this decrease in sinapine content was counterbalanced by a shift to higher amounts of its metabolic precursor, sinapoylglucose, thus resembling the *sinapoylglucose accumulator* phenotype of the Arabidopsis *sng2* mutant impaired in the gene At5g09640 encoding SCT (Shirley et al. 2001). With *BnSCT*-suppressing *B. napus* seeds we obtained similar results (D. Weier, personal communication). Since the accumulation of sinapoylglucose kept the total seed amount of anti-nutritive sinapoyl moieties nearly constant, silencing of *BnSCT* is not recommended as a promising strategy to improve the nutritional value of *B. napus* seeds.

A recent approach designed to silence the genes encoding C3H, C4H, COMT and SCT in *B. napus* revealed that suppression of the upper part of phenylpropanoid biosynthesis (C3H, C4H, COMT) was not effective in reducing the sinapine content whereas the results of *BnSCT* silencing agreed with our findings (Bhinu et al. 2009). As the most effective strategy to decrease sinapine and total phenolics amounts in seeds, the authors present a seed specific concomitant down-regulation of F5H and SCT resulting in a reduction of the sinapine content by about 90% and of total soluble phenolics by about 75%. As expected from the Arabidopsis *sng2* mutant, the down-regulation of *BnSCT* was accompanied by an increase in the seed level of free choline that was not channelled into sinapine. Since choline is an important feed additive this effect could potentially contribute to an increased nutritive value of *B. napus* seed products in addition to the low sinapine content (Zeisel 2000). Another promising strategy to suppress sinapine biosynthesis became obvious from the Arabidopsis mutant *ref 1* (Nair et al. 2004). The *REF1* gene encodes a bifunctional cinnamaldehyde dehydrogenase (SALDH/CALDH) converting sinapaldehyde and coniferaldehyde to the related hydroxycinnamates sinapate and ferulate, respectively. Silencing of *REF1* should, therefore, provide the possibility to suppress the metabolic flux toward sinapate esters without affecting

lignin biosynthesis or other related pathways. Accordingly, Arabidopsis *ref1* mutant plants displayed strongly reduced sinapate ester levels but without any detectable change in lignification pattern, growth phenotype or fertility compared to the wild-type (Nair et al. 2004). The remaining low sinapine level in *ref1* mutant seeds, however, points to the existence of alternative biosynthesis pathways toward sinapine. Therefore, the efficiency of *REF1* suppression for generating low sinapine seeds in *B. napus* remains to be proven.

4.2 Metabolic diversion

To reinforce the effect of *UGT84A9* suppression, the biosynthesis of the phytoalexin resveratrol (*trans*-3,5,4'-trihydroxystilbene) was established in seeds of *B. napus* by expression of the gene encoding stilbene synthase (*VST1*) from *Vitis vinifera* (Hüsken et al. 2005b). Stilbene synthase catalyzes the formation of resveratrol from one molecule of 4-coumaroyl-CoA and three molecules of malonyl-CoA, thus competing with chalcone synthase, the first committed enzyme of the flavonoid pathway, for substrates (Rupprich and Kindl 1978). In *B. napus* seeds, mediating both *UGT84A9* suppression and *VST1* expression, we found the sinapate ester level decreased below that of the best-performing *UGT84A9*-suppressing line. This indicates that the transgene-induced synthesis of resveratrol promoted suppression of sinapate ester biosynthesis by decreasing the metabolic flux toward sinapate (Hüsken et al. 2005b). Conversely, the suppression of *UGT84A9* led to an increase in the yield of piceid, the 3-*O*-β-D-glucoside of resveratrol, known for its health-promoting effects in humans. By these experiments we demonstrated the generation of *B. napus* plants with increased performance for two economically important traits due to a double transgenic strategy based on metabolic crosstalk in the phenylpropanoid pathway.

Currently, metabolic diversion was described to reduce the sinapine content of seeds by coupling to an increased flux of choline into glycine betaine. In Arabidopsis seeds it has been shown that expression of a bacterial choline oxidase (COX, EC 1.1.3.17) that converts choline into betaine reduced the sinapine accumulation (Huang et al. 2008). In addition, since betaine is a proven protective compound against various abiotic stress factors, its biosynthesis in seeds could potentially support germination and increase the vigour of young seedlings under adverse environmental conditions. However, in *B. napus* seeds the endogenous supply of choline was a limiting factor that prevented accumulation of betaine to physiologically relevant amounts in *COX* expressing lines (Huang et al. 2000). A recent study on *COX*

expression in an Arabidopsis T-DNA insertion mutant lacking SCT activity revealed increased betaine levels in seeds due to increased endogenous choline supply as a consequence of blocked sinapine synthesis (Huang et al. 2008). However, even in the presence of high COX activity, the seeds kept a certain level of free choline. This unknown mechanism of choline homeostasis might prevent *COX* expressing seeds to accumulate betaine to amounts required for increased abiotic stress tolerance.

In Arabidopsis, the perturbation of phenylpropanoid metabolism at different enzymatic steps revealed insights into metabolic links that should be considered for future engineering strategies. Thus, the perturbation of *O*-glucosylation and *O*-methylation has been related to decreasing amounts of soluble sinapate esters. Accordingly, overexpression of the genes encoding the monolignol-specific glucosyltransferases UGT72E2 and UGT72E3 led to a substantial loss in sinapoylmalate accompanied by the production of massive levels of coniferyl and sinapyl alcohol 4-*O*-glucosides (Lim et al. 2005; Lanot et al. 2006, 2008). Likewise, the inactivation of the Arabidopsis *O*-methyltransferase AtOMT1 by insertion mutagenesis caused a significant reduction of the sinpoylmalate content in leaves and stems of mature plants as well as in seedlings (Goujon et al. 2003). This decrease was related to the appearance of some unusual metabolites, 5-OH-feruloylmalate and 5-OH-feruloylglucose, in the metabolome. Among others, these studies point to the plasticity of phenylpropanoid metabolism allowing the plants to cope with perturbations brought about by modulation of enzyme activities.

4.3 Induced degradation

Based on the identification of BnSCE3 as sinapine esterase, we expressed the corresponding cDNA under control of the napin promoter during seed development with the aim to synchronize biosynthesis with hydrolysis of sinapine. In Arabidopsis, this approach produced nearly sinapine-free seeds. Remarkably, there was no increase of another sinapoylated compound in the soluble methanolic seed fraction (Clauß et al. 2008). This was surprising since sinapine hydrolysis by BnSCE3 liberates sinapate, which could be assumed to be metabolically redirected toward related compounds. Therefore, the dramatic decrease of the total amount of sinapate esters in *BnSCE3*-expressing seeds supported the hypothesis that free sinapate could act as a negative metabolic feedback signal for suppression of the metabolic flux toward its formation.

In summary, our results indicate metabolic engineering as efficient technology to generate low sinapine *B. napus* lines (Figure 9). Combination of different strategies bear the potential to produce seeds with the major phenolic sinapine converted to a trace component.

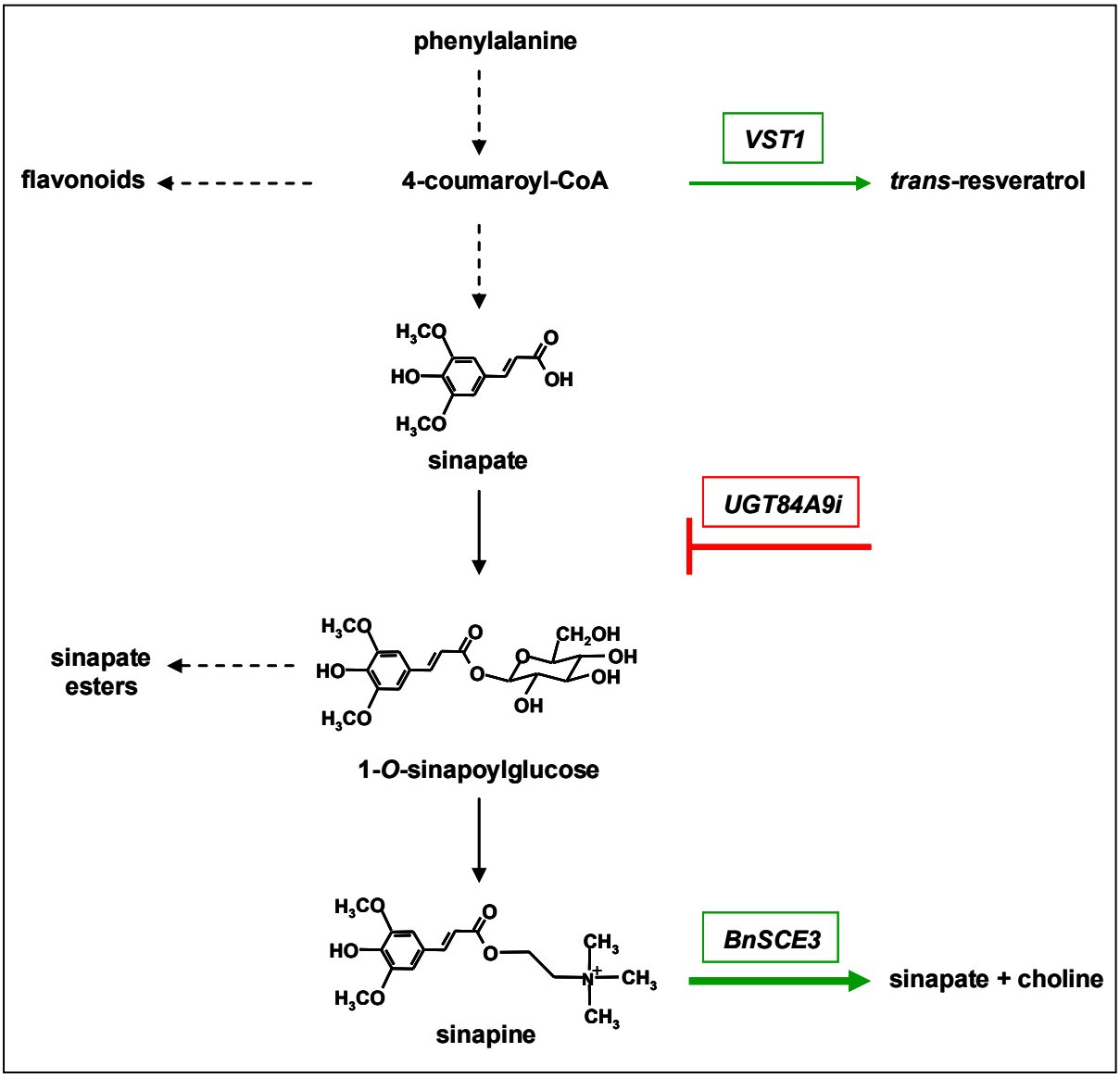


Figure 9 Metabolic engineering strategies to decrease the sinapate ester content in *B. napus* seeds. Suppression of biosynthesis by silencing of *UGT84A9* (*UGT84A9*i) caused a reduction of total sinapate ester content of about 75%. Metabolic diversion by overexpression of *VST1* from *Vitis vinifera* encoding stilbene synthase led to a decrease of about 50%. Induced degradation by overexpression of sinapine esterase (BnSCE3) provoked a nearly complete loss of sinapine. Related sinapate esters were present in trace amounts. Sinapate as a cleavage product was not detectable. Combined *UGT84A9* silencing and *VST1* overexpression led to a decrease in sinapate ester content of about 80%. Green arrows indicate seed-specific overexpression, suppression is marked by a red line. Line strength symbolizes the efficiency of the strategy in sinapate ester decrease. Dashed arrows indicate multiple enzymatic reactions.

4.4 Economical Relevance

According to recent estimates, *B. napus* ranks amongst the three major oilseed crops produced worldwide providing between 10% and 13% of the world's supply of vegetable oil (www.gmo-safety.eu; Hajduch et al. 2006). The success of oilseed rape is tightly related with the valuable oil provided by the seeds that is considered one of healthiest vegetable oils because of its high percentage of unsaturated fatty acids. The development of 'double zero' (00) lines with significantly reduced levels of the antinutritive compounds erucic acid and glucosinolates paved the way to a broad use in human diet and led to a steady increase of cultivation areas (Downey 1976).

As the second major storage compound, seeds of *B. napus* accumulate proteins to about 25% (w/w), which are deposited in modified vacuoles called protein bodies (Fenwick 1982). As major components, these seed storage proteins include the 12 S globulins cruciferins and the 2 S albumins napins. The seed protein fraction of *B. napus* is characterized by a well-balanced composition with high levels of essential sulphur-containing amino acids (Ohlson 1978). Accordingly, this protein fraction was shown to exhibit a high biological value in animals and humans (Campbell et al. 1981; Bos et al. 2007). However, this positive effect was counterbalanced by a poor digestibility due to a relatively high percentage of antinutritive components in the seed protein fraction. These are provided by the seed coat tissue (hulls) that delivers undigestible fibres and complex proanthocyanidins as well as by sinapate esters with sinapine as the major compound that accumulates in the embryo to amounts of up to 2% of seed dry matter (Bell 1993). Being oxidized during seed oil processing, sinapate esters may form complexes with proteins, thus causing reduced digestibility of the meal (Naczk et al. 1998; Rawel et al. 2000). Moreover, these phenolic esters contribute to the bitter taste and astringency of rapeseed products. As a result, the meal is considered as a by-product of the *B. napus* seed discounted about 35% relative to soybean meal.

Seed meal from *B. napus* is currently used in the feed industry. However, this market will be saturated by an amount of 1.8 – 2.5 million tons per year in Germany (Luck and Borcherding 1995). Accordingly, the yield of *B. napus* produced in Germany in 2007 was estimated to cause a surplus of about 2 million tons of meal that could not be brought into the feed market (Krause et al. 2007). This illustrates that the opening and exploitation of new markets for the seed protein fraction will be of pivotal importance for the future development of oilseed rape industry. Hence, the production of high quality protein products from *B. napus* that could compete with soybean in the field of human food supplements would mark a

breakthrough in oilseed rape commercialization. For this reason, the breeding of *B. napus* lines with lower seed coat related antinutritive compounds and strongly reduced sinapate ester content is of utmost importance.

By conventional plant breeding, lines with a yellow seed coat have been developed and are currently being used to generate high yielding cultivars (Rahman 2001). It was shown that the yellow seed coat correlated with a reduced amount and more digestible hulls with lower fibre content. The high sinapate ester content of the embryo, however, remains to be a challenge. So far, conventional plant breeding could not produce *B. napus* lines with the low sinapate ester trait. Hence, the transgenic approaches reported by us and others provided the best-performing lines with sinapate ester contents ranging between 20% and 30% of the non-transgenic parent lines (Hüsken et al. 2005a, b; Bhinu et al. 2009). As these low sinapate ester levels meet the requirements for human food supplements, the developed lines could be an alternative for the Canadian market that is dominated by transgenic Canola cultivars. In the long term, the combination of the low sinapine and yellow seed coat traits in an elite genetic background would be the most efficient strategy to increase the market value of *B. napus* protein products.

5 Summary and Outlook

During the past decade, plant natural product research has experienced the transition from descriptive biochemistry to engineering of metabolic paths and networks. This development was brought about by the rapidly growing sequence information on plant genomes that allowed the identification of genes involved in secondary metabolism. Large enzyme families have been recognized and many of their members have been characterized with regard to catalytic activity and phylogenetic ancestry. This initiated research efforts aimed at elucidating the evolution of the pronounced metabolic diversity in plants. After a phase, marked by gene hunting and *in vitro* characterization of enzymes as major research activities, it has recently become a challenging task to understand the physiological role of these enzymes and to develop engineering strategies to change plant metabolic networks in a predictive manner.

In this context, our research was focused on the sinapate ester metabolism in *B. napus* and Arabidopsis, emphasizing genetic and evolutionary aspects as well as engineering strategies to generate transgenic plants with reduced seed levels of the antinutritive

compounds, mainly sinapine and related phenolic esters. We identified and cloned the genes encoding the two final enzymes of sinapine biosynthesis, UDP-glucose:sinapate glucosyltransferase (BnSGT, UGT84A9) and sinapoylglucose:choline sinapoyltransferase (BnSCT), as well as the enzyme sinapine esterase (BnSCE3) from *B. napus*.

UGT84A9 clusters to the phylogenetic group L of secondary metabolism glucosyltransferases. Based on sequence identity and catalytic activity, UGT84A2 was identified as Arabidopsis SGT whereas the related UGTs 84A1, -A3 and -A4 were shown to catalyze glucose ester formation with a broad range of hydroxycinnamates. The allotetraploid genome of *B. napus* harbours four *UGT84A9* loci of which *UGT84A9a* and *UGT84A9b* account for the bulk of sinapate ester production. By seed-specific silencing of *UGT84A9* we were able to generate stable homozygous *B. napus* lines with sinapate ester levels reduced down to 24% of control plants. These results define the genes *UGT84A9a* and *UGT84A9b* as major molecular targets for breeding low sinapine oilseed rape.

For SCT, we characterized two genomic loci in *B. napus*, *BnSCT1* and *BnSCT2*. Both genes were induced on the level of transcription during seed development with a peak in the seed filling phase. Silencing of *BnSCT* genes led to reduced amounts of sinapine that were counterbalanced by increased levels of sinapoylglucose. Therefore, suppression of SCT alone is not appropriate to generate *B. napus* lines with low sinapate ester content in seeds. Like the Arabidopsis homologue SNG2, BnSCT enzymes belong to the family of serine carboxypeptidase-like enzymes. With the related Arabidopsis SMT as model enzyme we provided evidence that SCPL acyltransferases employ the same basic chemistry as the hydrolytic SCPs to mediate 1-*O*-glucose ester-dependent acyltransfer reactions. We propose the change to a sequential reaction mechanism and the facilitation of substrate proximity as driving forces for the shift from hydrolase to acyltransferase activity in the course of evolution. For recognition of the acyl donor 1-*O*-sinapoylglucose the highly conserved substitution of Glu by Asp at the sequence position preceding the catalytic Ser might be crucial.

We identified the sinapine esterase BnSCE3 as a member of the GDSL lipase-like enzyme family. The evolution of sinapine esterase activity was most likely brought about by extension of the substrate spectrum of a lipase-like ancestor, whose hydrolytic capacity was preserved. Seed-specific overexpression of *BnSCE3* led to the loss of sinapine accumulation. This indicated the physiological role of BnSCE3 as being involved in sinapine degradation. Seed-specific overexpression of *BnSCE3* might be an efficient strategy to produce low sinapine oilseed rape.

With the identification of genes encoding BnSGT, BnSCT and BnSCE, our work of the past decade laid the foundation for the development of targeted molecular breeding strategies to generate low sinapine oilseed rape. Accordingly, approaches have been started in our research group and in those of our collaborators that include the search for mutations in *UGT84A9a* and *UGT84A9b* genes by TILLING technology, the combined suppression of *UGT84A9* and overexpression of a bacterial *COX* gene as well as the seed-specific overexpression of *BnSCE3*. Of these, the best-performing strategy will be transferred to yellow seeded lines available in Canola. By this strategy, the two crucial traits for high quality seed protein, 'low sinapine' and 'low fibre content', will be combined in one elite genetic background. Based on recently described Arabidopsis mutants, the potential of suppressing the *REF1* homolog to specifically decrease the sinapate ester content should be tested in *B. napus*.

To cope with the future challenges in plant breeding, metabolic engineering must traverse empirical science to predictive modification. This will require an understanding of the fundamental biochemistry and of the rigid and plastic points of the metabolic network for which the characterization of metabolomes, transcriptomes and proteomes will be an indispensable prerequisite. Metabolic flux analyses, the characterization of transporters and transcription factors will extend our understanding of seed-specific sinapate ester biosynthesis and support novel engineering strategies in the near future.

6 References

Akoh CC, Lee GC, Liaw YC, Huang T-H, Shaw J-F (2004) GDSL family of serine esterases/lipases. Progr Lipid Res 43:534–552

Barz W, Köster J, Weltring K-M, Strack D (1985) Recent advances in the metabolism and degradation of phenolic compounds in plants and animals. In: Annual Proceedings of the Phytochemical Society of Europe. The Biochemistry of Plant Phenolics (Van Sumere, C.F., Lea, P.J., eds.), Vol. 25, pp. 307-347. Clarendon Press, Oxford

Baumert A, Milkowski C, Schmidt J, Nimtz M, Wray V, Strack D (2005) Formation of a complex pattern of sinapate esters in *Brassica napus* seeds, catalysed by enzymes of a serine carboxypeptidase-like acyltransferase family. Phytochemistry 66:1334-1345

Bell JM (1993) Factors affecting the nutritional value of canola meal: a review. Can J Anim Sci 73:679–697

Bhinu VS, Schäfer UA, Li R, Huang J, Hannoufa A (2009) Targeted modulation of sinapine biosynthesis pathway for seed quality improvement in *Brassica napus*. Transgenic Res 18:31-44

Bokern M, Heuer S, Strack D (1991) Hydroxycinnamic acid transferases in the biosynthesis of acylated betacyanins: purification and characterization from cell cultures of *Chenopodium rubrum* and occurrence in some other members of the *Caryophyllales*. Bot Acta 105:146-151

Bokern M, Strack D (1988) Synthesis of hydroxycinnamic acid esters of betacyanins via 1-*O*-acylglucosides of hydroxycinnamic acids by protein preparations from cell suspension cultures of *Chenopodium rubrum* and petals of *Lampranthus sociorum*. Planta 174:101-105

Bos C, Airinei G, Mariotti F, Benamouzig R, Berot S, Evrard J, Tome D, Gaudichon C (2007) The poor digestibility of rapeseed protein is balanced by its very high metabolic utilization in humans. J Nutr 137:594-600

Böttcher C, von Roepenack-Lahaye E, Schmidt J, Schmotz C, Neumann S, Scheel D, Clemens S (2008) Metabolome analysis of biosynthetic mutants reveals a diversity of metabolic changes and allows identification of a large number of new compounds in Arabidopsis. Plant Physiol 147:2107-2120

Bouchereau AJ, Hamelin J, Lamour I, Renard M, Larher F (1991) Distribution of sinapine and related compounds in seeds of Brassica and allied genera. Phytochemistry 30:1873–1881

Boudet A-M (2007) Evolution and current status of research in phenolic compounds. Phytochemistry 68:2722-2735

Breddam K, Soerensen SB, Svendsen I (1987) Primary structure and enzymatic properties of carboxypeptidase II from wheat bran. Carlsberg Res Commun 52:297-311

Brown DE, Rashotte AM, Murphy AS, Normanly J, Tague BW, Peer WA, Taiz L, Muday GK (2001) Flavonoids act as negative regulators of auxin transport in vivo in Arabidopsis. Plant Physiol 126:524-535

Buchanan BB, Gruissem W, Jones RL (2000) Biochemistry and Molecular Biology of Plants. American Society of Plant Biologists, Rockville, MD

Byzova M, Verduyn C, De Brouwer D, De Block M (2004) Transforming petals into sepliod organs in Arabidopsis and oilseed rape: implementation of the hairpin RNA-mediated gene silencing technology in an organ-specific manner. Planta 218:379-387

Campbell LD; Eggum BO; Jacobsen I (1981) Biological value, amino acid availability and true metabolizable energy of low-glucosinolate rapeseed meal (canola) determined with rats and / or roosters. Nutr Rep Int 24:791-797

Chen S, Hofius D, Sonnewald U, Börnke F (2003) Temporal and spatial control of gene silencing in transgenic plants by inducible expression of double stranded RNA. Plant J 36:731-740

Clauß K, Baumert A, Nimtz M, Milkowski C, Strack D (2008) Role of a GDSL lipase-like protein as sinapine esterase in Brassicaceae. Plant J 53:802–813

Coutinho PM Henrissat, B (1999) Carbohydrate-active enzymes server at URL: http://afmb.cnrs-mrs.fr/CAZY/

D'Auria JC, Gershenzon J (2005) The secondary metabolism of *Arabidopsis thaliana*: growing like a weed. Curr Opin Plant Biol 8:308-316

Derewenda U, Brzozowski AM, Lawson DM, Derewenda ZS (1992) Catalysis at the interface: the anatomy of the conformational change in a triglyceride lipase. Biochemistry 31:1532–1541

Dixon RA (2005) Engineering of plant natural product pathways. Curr Opin Plant Biol 8:329-336

Dodson G, Wlodawer A (1998) Catalytic triads and their relatives. Trends Biochem Sci 23:347-352

Downey RK (1976) Tailoring rapeseed and other oilseed crops to the market. Chem Ind 1:401-406

Duarte JM, Cui L, Wall PK, Zhang Q, Zhang X, Leebens-Mack J, Ma H, Altmann N, de Pamphilis CW (2006) Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of Arabidopsis. Mol Biol Evol 23:469-478

Endrizzi JA, Breddam K, Remington SJ (1994) 2.8-A structure of yeast serine carboxypeptidase. Biochemistry 33:11106-11120

Fenwick GR (1982) The assessment of a new protein source – Rapeseed. Proc Nutr Soc 41:277-288

Fraenkel G (1959) The Raison d'eˆtre of secondary plant substances. Science 129:1466–1470

Fraser CM, Rider LW, Chapple C (2005) An expression and bioinformatics analysis of the *Arabidopsis* serine carboxypeptidase-like gene family. Plant Physiol 138:1136–1148

Fraser CM, Thompson MG, Shirley AM, Ralph J, Schoenherr JA, Sinlapadech T, Hall MC, Chapple C (2007) Related Arabidopsis serine carboxypeptidase-like sinapoylglucose acyltransferases display distinct but overlapping substrate specificities. Plant Physiol 144:1986–1999

Fujiwara H, Tanaka Y, Yonekura-Sakakibara K, Fukuchi-Mizutani M, Nakao M, Fukui Y, Yamaguchi M, Ashikari T, Kusumi T (1998) cDNA cloning, gene expression and subcellular localization of anthocyanin 5-aromatic acyltransferase from *Gentiana triflora*. Plant J 16:421-431

Fukuchi-Mizutani M, Okuhara H, Fukui Y, Nakao M, Katsumoto Y, Yonekura-Sakakibara K, Kusumi T, Hase T, Tanaka Y (2003) Biochemical and molecular characterisation of a novel UDP-glucose: anthocyanin 3'-*O*-glucosyltransferase, a key enzyme for blue anthocyanin biosynthesis, from gentian. Plant Physiol 132:1652-1663

Gadamer J (1897) Über die Bestandteile des schwarzen und weissen Senfsamens. Arch Pharm 235:44-114

Gläßgen WE, Seitz HU (1992) Acylation of anthocyanins with hydroxycinnamic acids via 1-*O*-acylglucosides by protein preparations from cell cultures of *Daucus carota* L. Planta 186:582-585

Goujon T, Sibout R, Pollet B, Maba B, Nussaume L, Bechtold N, Lu F, Ralph J, Mila I, Barrière Y, Lapierre C, Jouanin L (2003) A new *Arabidopsis thaliana* mutant deficient in the expression of *O*-methyltransferase impacts lignins and sinapoyl esters. Plant Mol Biol 51:973-989

Gräwe W, Bachhuber P, Mock H-P, Strack D (1992) Purification and characterization of sinapoylglucose:malate sinapoyltransferase from *Raphanus sativus* L. Planta 187:236-241

Hajduch M, Casteel JE, Hurrelmeyer KE, Song Z, Agrawal GK, Thelen JJ (2006) Proteomic analysis of seed filling in *Brassica napus*. Developmental characterization of metabolic isozymes using high-resolution two-dimensional gel electrophoresis. Plant Physiol 141:32-46

Harborne JB, Turner BL (1984) Plant Chemosystematics. Academic Press, London

Hartmann T (2007) From waste products to ecochemicals: Fifty years research of plant secondary metabolism. Phytochemistry 68:2831-2846

Hause B, Meyer K, Viitanen PV, Chapple CCS, Strack D (2002) Immunolocalization of 1-*O*-sinapoylglucose:malate sinapoyltransferase in *Arabidopsis thaliana*. Planta 215:26-32

Honda T, Tatsuzawa F, Kobayashi N, Kasai H, Nagumo S, Shigihara A, Saito N (2005) Acylated anthocyanins from the violet-blue flowers of *Orychophragonus violaceus*. Phytochemistry 66:1844-1851

Hrazdina G (1992) Compartmentation in aromatic metabolism. in Recent Advances in Phytochemistry, Vol. 24, Phenolic Metabolism in Plants, (Stafford HA, Ibrahim RK, eds.), pp. 1-23, Plenum Press, New York

Hrazdina G, Iredala H, Mattick LR (1977) Anthocyanin composition of *Brassica oleracea* cv. Red Danish. Phytochemistry 16:297–299

Huang J, Hirji R, Adam L, Rozwadowski KL, Hammerlindl JK, Keller WA, Selvaraj G (2000) Genetic engineering of glycinebetaine production toward enhancing stress tolerance in plants: metabolic limitations. Plant Physiol 122:747-756

Huang J, Rozwadowski K, Bhinu VS, Schäfer U, Hannoufa A (2008) Manipulation of sinapine, choline and betaine accumulation in Arabidopsis seed: towards improving the nutritional value of the meal and enhancing the seedling performance under environmental stresses in oilseed crops. Plant Physiol Biochem 46:647-654

Huang YT, Liaw YC, Gorbatyuk VY, Huang TH. (2001) Backbone dynamics of *Escherichia coli* thioesterase/protease I: evidence of a flexible active-site environment for a serine protease. J Mol Biol 307:1075-1090

Hughes J, Hughes MA (1994) Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava. DNA Sequence 5:41-49

Humphreys JM, Hemm MR, Chapple C (1999) New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multi-functional cytochrome P450-dependent monooxygenase. Proc Natl Acad Sci USA 96:10045-10050

Hüsken A, Baumert A, Strack D, Becker HC, Möllers C, Milkowski C (2005a) Reduction of sinapate ester content in transgenic oilseed rape (*Brassica napus*) by dsRNAi-based suppression of *BnSGT1* gene expression. Mol Breed 16:127–138

Hüsken A, Baumert A, Milkowski C, Becker HC, Strack D, Möllers C (2005b) Resveratrol glucoside (Piceid) synthesis in seeds of transgenic oilseed rape (*Brassica napus* L.). Theor Appl Genet 111:1553–1562

Kita M, Hirata Y, Moriguchi T, Endo-Inagaki T, Matsumoto R, et al. (2000) Molecular cloning and characterization of a novel gene encoding limonoid UDP-glucosyltransferase in Citrus. FEBS Lett 469:173–178

Kowalczyk S, Jakubowska A, Zielinska E, Bandurski RS (2003) Bifunctional indole-3-acetyl transferase catalyses synthesis and hydrolysis of indole-3-acetyl-myo-inositol in immature endosperm of *Zea mays*. Physiol Plantarum 119:165–174

Kozlowska H, Naczk M, Shahidi F, Zadernowski R (1990) Phenolic acids and tannins in rapeseed and canola. In: Shahidi F. (ed.), Canola and Rapeseed. Production, Chemistry, Nutrition and Processing Technology. Van Nostrand Reinhold, New York, pp. 193–210

Krause J-P, Kroll J, Rawel HM (2007) Verarbeitung von Rapssaat – Eigenschaften und Gewinnung von Proteinen. UFOP-Schriften 32:11-101

Lagercrantz U, Lydiate DJ (1996) Comparative genome mapping in *Brassica*. Genetics 144:1903–1910

Landry LG, Chapple CCS, Last RL (1995) Arabidopsis mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. Plant Physiol 109:1159-1166

Lanot A, Hodge D, Jackson RG, George GL, Elias L, Lim EK, Vaistij FE, Bowles DJ (2006) The glucosyltransferase UGT72E2 is responsible for monolignol 4-*O*-glucoside production in *Arabidopsis thaliana*. Plant J 48:286-295

Lanot A, Hodge D, Lim EK, Vaistij FE, Bowles DJ (2008) Redirection of flux through the phenylpropanoid pathway by increased glucosylation of soluble intermediates. Planta 228:609-616

Lee YL, Chen JC, Shaw JF (1997) The thioesterase I of *Escherichia coli* has arylesterase activity and shows stereospecificity for protease substrates. Biochem Biophys Res Commun 231:452-456

Lehfeldt C, Shirley AM, Meyer K, Ruegger MO, Cusumano JC, Viitanen PV, Strack D, Chapple C (2000) Cloning of the *SNG1* gene of *Arabidopsis* reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. Plant Cell 12:1295–1306

Li AX, Steffens JC (2000) An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. Proc Natl Acad Sci USA 97:6902–6907

Li J, Derewenda U, Dauter Z, Smith S, Derewenda ZS (2000) Crystal structure of the *Escherichia coli* thioesterase II, a homolog of the human Nef binding enzyme. Nat Struct Biol 7:555-559

Li Y, Baldauf S, Lim EK, Bowles D (2001) Phylogenetic analysis of the UDP-glucosyltransferase multigene family of *Arabidopsis thaliana*. J Biol Chem 276:4338-4343

Liao D-I, Breddam K, Sweet RM, Bullock T, Remington SJ (1992) Refined atomic model of wheat serine carboxypeptidaseII at 2.2-A resolution. Biochemistry 31:9796-9812

Lim E-K, Higgins GS, Li Y, Bowles DJ (2003) Regioselectivity of glucosylation of caffeic acid by a UDP-glucose:glucosyltransferase is maintained *in planta*. Biochem J 373:987-992

Lim E-K, Doucet CJ, Li Y, Elias L, Worrall D, Spencer SP, Ross J, Bowles DJ (2002) The activity of *Arabidopsis* glycosyltransferases towards salicylic acid, 4-hydroxybenzoic acid, and other benzoates. J Biol Chem 277:586–592

Lim E-K, Jackson RG, Bowles DJ (2005) Identification and characterisation of Arabidopsis glycosyltransferases capable of glucosylating coniferyl aldehyde and sinapyl aldehyde. FEBS Lett 579:2802-2806

Lim E-K, Li Y, Parr A, Jackson R, Ashford DA, Bowles DJ (2001) Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in Arabidopsis. J Biol Chem 276:4344-4349

Ling H, Zuo K, Zhao J, Qin J, Qiu C, Sun X, Tang K (2006) Isolation and characterization of a homologous to lipase gene from *Brassica napus*. Russ J Plant Physiol 53:366-372

Lorenzen M, Racicot V, Strack D, Chapple C (1996) Sinapic acid ester metabolism in wild type and a sinapoylglucose-accumulating mutant of Arabidopsis. Plant Physiol 112:1625-1630

Luck T, Borcherding A (1995) Evaluierung der technischen Verwertungsmöglichkeiten für die Nebenprodukte der Ölerzeugung aus Raps. BMFT-Verbundprojekt Kraftstoffe aus Raps, Fkz 10618 A

Luo J, Nishiyama Y, Fuell C, Taguchi G, Elliott K, Hill L, Tanaka Y, Kitayama M, Yamazaki M, Bailey P, Parr A, Michael AJ, Saito K, Martin C (2007) Convergent evolution in the BAHD family of acyl transferases: identification and characterization of anthocyanin acyl transferases from *Arabidopsis thaliana*. Plant J 50:678–695

Mackenzie P, Owens I, Burchell B, Bock K, Bairoch A, Belanger A, Fournel-Gigleux S, Green M, Hum D, Iyanagi T et al. (1997) The UDP glycosyltransferase superfamily: recommended nomenclature update based on evolutionary divergence. Pharmacogenetics 7:255-269

Meißner D, Albert A, Böttcher C, Strack D, Milkowski C (2008) The role of UDP-glucose:hydroxycinnamate glucosyltransferases in phenylpropanoid metabolism and the response to UV-B radiation in *Arabidopsis thaliana*. Planta 228:663-674

Meyer K, Cusumano JC, Somerville C, Chapple CCS (1996) Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monooxygenases. Proc Natl Acad Sci USA 93:6869–6874

Milkowski C, Strack D (2004) Serine carboxypeptidase-like acyltransferases. Phytochemistry 65:517–524

Milkowski C, Baumert A, Schmidt D, Nehlin L, Strack D (2004) Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. Plant J 38:80–92

Milkowski C, Baumert A, Strack D (2000a) Cloning and heterologous expression of a rape cDNA encoding UDP-glucose:sinapate glucosyltransferase. Planta 211:883–886

Milkowski C, Baumert A, Strack D (2000b) Identification of four *Arabidopsis* genes encoding hydroxycinnamate glucosyltransferases. FEBS Lett 486:183–184

Mittasch J (2008) Hydroxycinnamate glucosyltransferase genes in *Brassica napus*, encoding key enzymes in sinapate ester metabolism. PhD thesis, Martin-Luther-Universität, Halle

Mittasch J, Strack D, Milkowski C (2007) Secondary product glycosyltransferases in seeds of *Brassica napus*. Planta 225:515–522

Mock H-P, Strack D (1993) Energetics of the uridine 5'-diphosphoglucose: hydroxycinnamic acid acyl-glucosyltransferase reaction. Phytochemistry 32:575-579

Moore RC, Purugganan MD (2005) The evolutionary dynamics of plant duplicate genes. Curr Opin Plant Biol 8:122-128

Naczk M, Aramowicz A, Sullivan A, Shahidi F (1998) Current research developments on polyphenolics of rapeseeds/canola: a review. Food Chem 62:489-502

Nair RB, Bastress KL, Ruegger MO, Denault JW, Chapple C (2004) The *Arabidopsis thaliana REDUCED EPIDERMAL FLUORESCENCE1* gene encodes an aldehyde dehydrogenase involved in ferulic acid and sinapic acid biosynthesis. Plant Cell 16:544-554

Nurmann G, Strack D (1979) Sinapine esterase. Part I. Characterization of sinapine esterase from cotyledons of *Raphanus sativus*. Z Naturforsch 34c:715-720

Ober D, Hartmann T (1999) Homospermidine synthase, the first pathway-specific enzyme of pyrrolizidine alkaloid biosynthesis, evolved from deoxyhypusine synthase. Proc Natl Acad Sci USA 96:14777-14782

Ohlson R (1978) Functional properties of rapeseed oil and protein product. In Proc 5th Int Rapeseed Congr, Malmö, Sweden, 152-167

Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J, Sussman JL, Verschueren KHG, Goldman A (1992) The α/β hydrolase fold. Protein Eng 5:197–211

Paquette S, Møller BL, Bak S (2003) On the origin of family 1 plant glycosyltransferases. Phytochemistry 62:399-413.

Peach K (1950) Biochemie und Physiologie der Sekundären Pflanzenstoffe. Springer, Berlin Göttingen Heidelberg

Quiel JA, Benders J (2003) Glucose conjugation of anthranilate by the Arabidopsis UGT74F2 glucosyltransferase is required for tryptophan mutant blue fluorescence. J Biol Chem 278:6275-6281

Rahman MH (2001) Production of yellow-seeded *Brassica napus* through interspecific crosses. Plant Breed 120:463-472

Rawel H, Rohn S, Kroll J (2000) Reaction of selected secondary plant metabolites (glucosinolates and phenols) with food proteins and enzymes – Influence on physicochemical protein properties, enzyme activity and proteolytic degradation. Recent research developments. Phytochemistry 4:115-142

Ross J, Li Y, Lim E-K, Bowles DJ (2001) Higher plant glycosyltransferases. Genome Biol 2: 3004.1-3004.6

Ruegger M, Chapple C (2001) Mutations that reduce sinapoylmalate accumulation in *Arabidopsis thaliana* define loci with diverse roles in phenylpropanoid metabolism. Genetics 159:1741-1749

Rupprich N, Kindl H (1978) Stilbene synthase and stilbene carboxylate synthases. I. Enzymatic synthesis of 3,5,4,-trihydroxystilbene from p-coumaroyl-CoA and malonyl-CoA. Hoppe Seyler's Z Physiol Chem 359:165-175

Schellenberger A (1989) Enzymkatalyse – Einführung in die Chemie, Biochemie und Technologie der Enzyme. VEB Gustav-Fischer-Verlag, Jena

Schoch G, Goepfert S, Morant M, Hehn A, Meyer D, Ullmann P, Werck-Reichhart D (2001) CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. J Biol Chem 276:36566-36574

Shahidi F, Naczk M (1992) An overview of the phenolics of canola and rapeseed: chemical, sensory and nutritional significance. J Am Oil Chem Soc 69:917-924

Shirley AM, Chapple C (2003) Biochemical characterization of sinapoylglucose:choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism. J Biol Chem 278:19870–19877

Shirley AM, McMichael CM, Chapple C (2001) The *sng2* mutant of *Arabidopsis* is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose:choline sinapoyltransferase. Plant J 28:83–94

Sinlapadech T, Stout J, Ruegger MO, Deak M, Chapple C (2007) The hyper-fluorescent trichome phenotype of the *brt1* mutant of *Arabidopsis* is the result of a defect in a sinapic acid:UDPG glucosyltransferase. Plant J 49:655–668

Steffens JC (2000) Acyltransferases in protease's clothing. Plant Cell 12:1253–1255

Stehle F (2008) Struktur-Funktionsbeziehungen der Serin-Carboxypeptidase-ähnlichen Sinapoylglucose:Malat-Sinapoyltransferase. PhD thesis, Martin-Luther-Universität, Halle

Stehle F, Brandt W, Milkowski C, Strack D (2006) Structure determinants and substrate recognition of serine carboxypeptidase-like acyltransferases from plant secondary metabolism. FEBS Lett 580:6366–6374

Stehle F, Brandt W, Schmidt J, Milkowski C, Strack D (2008a) Activities of Arabidopsis sinapoylglucose:malate sinapoyltransferase shed light on functional diversification of serine carboxypeptidase-like acyltransferases. Phytochemistry 69:1826-1831

Stehle F, Stubbs MT, Strack D, Milkowski C (2008b) Heterologous expression of a serine carboxypeptidase-like acyltransferase and characterization of the kinetic mechanism. FEBS J 275:775-787

St Pierre B, De Luca V (2000) Evolution of acyltransferase genes: origin and diversification of the BAHD superfamily of acyltransferases involved in secondary metabolism. In Recent Advances in Phytochemistry 34. Evolution of Metabolic Pathways. Edited by Romeo JT, Ibrahim R, Varin L, De Luca V. Elsevier Science Ltd:285-315

Strack D (1980) Enzymatic synthesis of 1-sinapoylglucose from free sinapic acid and UDP-glucose by a cell-free system from *Raphanus sativus* seedlings. Z Naturforsch 35c:204-208

Strack D, Gross W, Heilemann J, Keller H, Ohm S (1988) Enzymic synthesis of hydroxycinnamic acid esters of glucaric acid and hydroaromatic acids from the respective 1-*O*-hydroxycinnamoylglucoside and hydroxycinnamoyl-coenzyme A thioester as acyldonors with a protein preparation from *Cestrum elegans* leaves. Z Naturforsch 43c:32-36

Strack D, Heilemann J, Boehnert B, Grotjahn L, Wray V (1987) Accumulation and enzymatic synthesis of 2-acetyl-3-(*p*-coumaroyl)-*meso*-tartaric acid in spinach cotyledons. Phytochemistry 26:107-111

Strack D, Knogge W, Dahlbender B (1983) Enzymatic synthesis of sinapine from 1-*O*-ß-D-glucose and choline by a cell-free system from developing seeds of red radish (*Raphanus sativus* L. var. *sativus*). Z Naturforsch 38c:21-27

Sunilkumar G, Campbell LM, Puckhaber L, Stipanovic RD, Rathore KS (2006) Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol. Proc Natl Acad Sci USA 103:18054-18059

Svendsen I, Martin BM, Viswanatha T, Johansen JT (1982) Amino acid sequence of carboxypeptidase Y. Peptides from enzymatic cleavages. Carlsberg Res Commun 47:15-27

Szerszen JB, Szczyglowski K, Bandurski RS (1994) *Iaglu*, a gene from *Zea mays* involved in conjugation of growth hormone indole-3-acetic acid. Science 265:1699-1701

Tattersall DB, Bak S, Jones PR, Olsen CE, Nielsen JK, Hansen ML, Hoej PB, Moeller BL (2001) Resistance to an herbivore through engineered cyanogenic glucoside synthesis. Science 239:1826-1828

Tkotz N, Strack D (1980) Enzymatic synthesis of sinapoyl-L-malate from 1-sinapoylglucose and L-malate by a protein preparation from *Raphanus sativus* cotelydons. Z Naturforsch 35c:835-837

Town CD, Cheung F, Maiti R, Crabtree J, Haas BJ, Wortmann JR, Hine EE, Althoff R, Arbogast TS, Tallon LJ, Vigouroux M, Trick M, Bancroft I (2006) Comparative genomics of *Brassica oleracea* and *Arabidopsis thaliana* reveal gene loss, fragmentation, and dispersal after polyploidy. Plant Cell 18:1348-1359

Townsend BJ, Llewellyn DJ (2007) Reduced terpene levels in cottonseed add food to fiber. Trends Biotechnol 25:239-241

Tzagoloff A (1963) Metabolism of Sinapine in Mustard Plants. I. Degradation of Sinapine into Sinapic Acid & Choline. Plant Physiol 38:202-206

Vanholme R, Morreel K, Ralph J, Boerjan W (2008) Lignin engineering. Curr Opin Plant Biol 11:278-285

Velasco L, Möllers C (1998) Nondestructive assessment of sinapic acid esters in Brassica species: II. Evaluation of germplasm and identification of phenotypes with reduced levels. Crop Sci 38:1650–1654

Villegas RJA, Kojima M (1986) Purification and characterization of hydroxycinnamoyl D-glucose quinate hydroxycinnamoyl transferase in the root of sweet potato, *Ipomoea batatas* LAM. J Biol Chem 261:8729-8733

Vogt T, Aebershold R, Ellis B (1993) Purification and characterization of sinapine synthase from seeds of *Brassica napus*. Arch Biochem Biophys 300:622–628

Vogt T, Jones P (2000) Glycosyltransferases in plant natural product synthesis: characterisation of a supergene family. Trends Plant Sci 5:380-386

Wang SX, Ellis BE (1998) Enzymology of UDP-glucose:sinapic acid glucosyltransferase from *Brassica napus*. Phytochemistry 49:307-318

Wang X (2001) Plant phospholipases. Annu Rev Plant Physiol Biol 52:211–231

Weier D, Mittasch J, Strack D, Milkowski C (2007) The genes *BnSCT1*and *BnSCT2* from *Brassica napus* encoding the final enzyme of sinapine biosynthesis: molecular characterization and suppression. Planta 227:375–385

Widmer F, Johansen JT (1979) Enzymatic peptide synthesis: carboxypeptidase catalyzed formation of peptide bonds. Carlsberg Res Commun 44:37–46

Wink M (1988) Plant breeding: importance of plant secondary metabolites for protection against pathogens and herbivores. Theor Appl Genet 75:225-233

Ye X, Al-Babili S, Kloti A, Zhang J, Lucca P, Beyer P, Potrykus I (2000) Engineering provitamin A (ß-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science 287:303-305

Yonekura-Sakakibara K, Tanaka Y, Fukuchi-Mizutani M, Fujiwara H, Fukui Y, Ashikari T, Murakami Y, Yamaguchi M, Kusumi T (2000) Molecular and biochemical characterization of a novel hydroxycinnamoyl-CoA:anthocyanin 3-*O*-glucoside-6((-*O*-acyltransferase from *Perilla frutescens*. Plant Cell Physiol 41:495-502

Zeisel SH (2000) Choline: an essential nutrient for humans. Nutrition 16:669-671

Zhou A, Li J (2005) Arabidopsis BRS1 is a secreted and active serine carboxypeptidase. J Biol Chem 280:35554-35561

7 List of included publications

1. Milkowski C, Baumert A, Strack D (2000a) Cloning and heterologous expression of a rape cDNA encoding UDP-glucose:sinapate glucosyltransferase. Planta 211:883-886

2. Milkowski C, Baumert A, Strack D (2000b) Identification of four Arabidopsis genes encoding hydroxycinnamate glucosyltransferases. FEBS Lett 486:183-184

3. Milkowski C, Baumert A, Schmidt D, Nehlin L, Strack D (2004) Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. Plant J 38:80-92

4. Baumert A, Milkowski C, Schmidt J, Nimtz M, Wray V, Strack D (2005) Formation of a complex pattern of sinapate esters in *Brassica napus* seeds, catalysed by enzymes of a serine carboxypeptidase-like acyltransferase family. Phytochemistry 66:1334-1345

5. Hüsken A, Baumert A, Strack D, Becker HC, Möllers C, Milkowski C (2005a) Reduction of sinapate ester content in transgenic oilseed rape (*Brassica napus*) by dsRNAi-based suppression of *BnSGT1* gene expression. Mol Breed 16:127–138

6. Hüsken A, Baumert A, Milkowski C, Becker HC, Strack D, Möllers C (2005b) Resveratrol glucoside (Piceid) synthesis in seeds of transgenic oilseed rape (*Brassica napus* L.). Theor Appl Genet 111:1553–1562

7. Mittasch J, Strack D, Milkowski C (2007) Secondary product glycosyltransferases in seeds of *Brassica napus*. Planta 225:515-522

8. Weier D, Mittasch J, Strack D, Milkowski C (2008) The genes *BnSCT1* and *BnSCT2* from *Brassica napus* encoding the final enzyme of sinapine biosynthesis - molecular characterization and suppression. Planta 227:375-385

9. Clauß K, Baumert A, Nimtz M, Milkowski C, Strack D (2008) Role of a GDSL lipase-like protein as sinapine esterase in Brassicaceae. Plant J 53:802-813

10. Meißner D, Albert A, Böttcher C, Strack D, Milkowski C (2008) The role of UDP-glucose:hydroxycinnamate glucosyltransferases in phenylpropanoid metabolism and the response to UV-B radiation in *Arabidopsis thaliana*. Planta 228:663-674

11. Stehle F, Brandt W, Milkowski C, Strack D (2006) Structure determinants and substrate recognition of serine carboxypeptidase-like acyltransferases from plant secondary metabolism. FEBS Lett 580:6366-6374

12. Stehle F, Brandt W, Schmidt J, Milkowski C, Strack D (2008a) Activities of Arabidopsis sinapoylglucose:malate sinapoyltransferase shed light on functional diversification of serine carboxypeptidase-like acyltransferases. Phytochemistry 69:1826-1831

13. Stehle F, Stubbs MT, Strack D, Milkowski C (2008b) Heterologous expression of a serine carboxypeptidase-like acyltransferase and characterization of the reaction mechanism. FEBS J 275:775-787

8 Statement on personal contributions

The work on sinapate ester metabolism was performed at the Leibniz Institute of Plant Biochemistry, where I initially held a postdoc position, three years later followed by a group leader position at the Department of Secondary Metabolism headed by Prof. Dieter Strack.

Construction of cDNA libraries from *B. napus* seeds and seedlings, cloning and heterologous expression of *UGT84A9*-cDNA were performed by myself (Milkowski et al. 2000a). The identification of UGT84A2 as SGT and of UGT84A1, -A3 and -A4 as ester-forming HCA-GTs of Arabidopsis were done by myself with support from Dr. A. Baumert (Leibniz Institute of Plant Biochemistry, Department of Secondary Metabolism) (Milkowski et al. 2000b). Design and construction of the first generation seed-specific dsRNAi-suppression cassette to silence *UGT84A9* in *B. napus* were accomplished by myself (Hüsken et al. 2005a). Homology-based cloning of *BnSCT* was designed and initiated by myself and finished by the former PhD student Dr. Diana Weier under my scientific guidance. First

genomic Southern Blot analyses to estimate gene copy numbers of *UGT84A9* and *BnSCT* in the allotetraploid *B. napus* and its diploid ancestors were accomplished by myself as well as the Northern Blot analyses to study developmental regulation of *UGT84A9* and *BnSCT* (Milkowski et al. 2004). Seed-specific silencing of SCT in Arabidopsis and *BnSCT* promoter studies were carried out under my scientific guidance (Weier et al 2008). I was involved in the experimental design to combine *UGT84A9* suppression with resveratrol biosynthesis by *VST1* expression in seeds of *B. napus* (Hüsken et al. 2005b). The concept of identifying alternative UGTs potentially involved in seed sinapate ester biosynthesis by EST cloning and homology-based cloning strategies was developed by myself and accomplished by the former PhD student Dr. Juliane Mittasch under my scientific guidance (Mittasch et al. 2007). Based on my work on ester-forming UGTs, the project to elucidate the biological function of UGT84A members in Arabidopsis was initiated by myself and performed by the former PhD student Dr. Dirk Meißner (Meißner et al. 2008). To the project aimed at identification of sinapine esterase from *B. napus*, my impact included set up and scientific guidance of the cDNA cloning strategy, based on peptide sequences of the purified enzyme. The metabolic engineering approach of induced sinapine degradation by seed-specific *BnSCE3* expression was designed by myself and carried out by the PhD student Kathleen Clauß (Clauß et al. 2008).

The second line of research aimed at elucidating functional aspects of SCPL acyltransferase evolution was initiated by Prof. Dieter Strack and myself based on emerging sequence information that revealed homology of these enzymes with hydrolases. The first phylogenetic analysis of the Arabidopsis SCPL protein family that led to the identification of a putative acyltransferase cluster and to the identification of critical amino acid positions and substitutions was done by myself. Given the high level of sequence identity with well-characterized serine carboxypeptidases, the structure modelling approach was initiated by myself and Prof. Dieter Strack and essentially performed by Dr. Wolfgang Brandt (Leibniz Institute of Plant Biochemistry, Department of Bioorganic Chemistry) and Felix Stehle as part of his PhD work (Stehle et al. 2006). I developed the basic concept of an optimized host-vector system for heterologous *AtSMT* expression that was refined and realized by the former PhD student Dr. Felix Stehle (Stehle et al. 2008b). Site-directed mutagenesis of AtSMT was done under my scientific guidance by Felix Stehle as well as the kinetic characterization of AtSMT catalysis for which theoretical support by Prof. Dieter Strack and Dr. Stephan König (Martin Luther University Halle, Institute of Biochemistry and Biotechnology) was a substantial benefit (Stehle et al. 2008a, b).

Danksagung

Die vorliegende Arbeit wurde am Leibniz-Institut für Pflanzenbiochemie in der Abteilung Sekundärstoffwechsel angefertigt. Mein besonderer Dank gilt dem Leiter dieser Abteilung Prof. Dr. D. Strack, der mein Interesse für den pflanzlichen Stoffwechsel geweckt hat und mir schon früh die wissenschaftliche Freiheit für ein selbständiges und eigenverantwortliches Arbeiten gewährte. Seinem Vertrauen verdanke ich auch die Möglichkeit, die Arbeitsgruppe „Phenylpropanstoffwechsel“ aufzubauen und zu leiten.

Ganz herzlich möchte ich mich bei den ehemaligen und aktuellen Doktoranden bedanken, die mit Enthusiasmus und Ausdauer ihre individuellen Projekte verfolgt haben und mir in vielen anregenden Diskussionen interessante Aspekte der Arbeiten erschlossen. Es sind dies (in alphabetischer Reihenfolge): Dr. Dirk Meißner, Dr. Juliane Mittasch, Dr. Felix Stehle und Dr. Diana Weier sowie Kathleen Clauß und Jessica Leuchte. Claudia Horn und Sylvia Vetter verdanke ich die sachkundige Durchführung vieler praktischer Arbeiten. Darüber hinaus danke ich allen Mitarbeitern der Abteilung Sekundärstoffwechsel für die angenehme Arbeitsatmosphäre und stets gewährte Hilfsbereitschaft.

Herr Dr. Alfred Baumert hat durch seine großen Erfahrungen in der Analytik pflanzlicher Naturstoffe und in der Proteinreinigung wesentlich zum Gelingen der einzelnen Projekte beigetragen. Ihm verdanke ich außerdem viele interessante Diskussionen über chemische und allgemeine Aspekte unserer wissenschaftlichen Arbeiten. Dr. Lilian Nehlin (Gregor-Mendel-Institut für Molekulare Pflanzenbiologie, Wien) hat geholfen, Arabidopsis als Forschungsobjekt zu etablieren und ihre Sachkenntnis und Begeisterung für die Biologie pflanzlicher Samen eingebracht.

Mein besonderer Dank gilt den Kooperationspartnern, die durch ihr Spezialwissen die hier vorgestellten Arbeiten substanziell bereichert haben. Dies sind: Dr. C. Böttcher, Dr. J. Schmidt und Dr. W. Brandt (Leibniz-Institut für Pflanzenbiochemie, Halle), Prof. M. T. Stubbs (Universtät Halle), Dr. M. Nimtz (Helmholtz-Zentrum für Infektionsforschung, Braunschweig), Dr. W. Heller (Helmholtz-Zentrum München), Dr. R. Snowdon (Universität Gießen), Dr. A. Hüsken (Universität Göttingen, Biologische Bundesanstalt für Land- und Forstwirtschaft, Braunschweig), Prof. B. Weisshaar (Universität Bielefeld), Prof. G. Selveraj (University of Saskatchewan, Saskatoon, Canada).

Vielen Dank an die Projektpartner aus der Wirtschaft, Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (Dr. G. Leckband, Dr. M. Frauen), Deutsche Saatveredelung AG (Dr. D. Stelling), KWS Saat AG (Dr. F. Breuer), Saatenunion Resistenzlabor GmbH (Dr. J. Orsini), die den Bezug zur angewandten Forschung herstellten und ihre Erfahrungen und Ressourcen aus der Rapszüchtung einbrachten.

Für die kritische Durchsicht dieser Arbeit und wertvolle Hinweise danke ich Prof. D. Strack, Prof. C. Wasternack, Dr. B. Hause, Dr. T. Vogt, Dr. J. Mittasch und Dr. F. Stehle.

Ich danke der Deutschen Forschungsgemeinschaft und dem Bundesministerium für Bildung und Forschung für die großzügige finanzielle Unterstützung der hier vorgestellten Arbeiten.

Ehrenwörtliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe. Andere als die angegebenen Quellen und Hilfsmittel wurden nicht benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen wurden als solche kenntlich gemacht.
Die Arbeit wurde nur an der Biologisch-Pharmazeutischen Fakultät der Friedrich-Schiller-Universität Jena vorgelegt. Frühere oder andere laufende Habilitationsverfahren gibt es nicht.

Jena, den 16.04.2009

Dr. Carsten Milkowski

Planta (2000) 211: 883–886

Planta
© Springer-Verlag 2000

Rapid communication

Cloning and heterologous expression of a rape cDNA encoding UDP-glucose:sinapate glucosyltransferase

Carsten Milkowski, Alfred Baumert, Dieter Strack

Leibniz-Institut für Pflanzenbiochemie, Abteilung Sekundärstoffwechsel, Weinberg 3, 06120 Halle (Saale), Germany

Received: 14 July 2000 / Accepted: 8 August 2000

Abstract. A cDNA encoding a UDP-glucose:sinapate glucosyltransferase (SGT) that catalyzes the formation of 1-*O*-sinapoylglucose, was isolated from cDNA libraries constructed from immature seeds and young seedlings of rape (*Brassica napus* L.). The open reading frame encoded a protein of 497 amino acids with a calculated molecular mass of 55,970 Da and an isoelectric point of 6.36. The enzyme, functionally expressed in *Escherichia coli*, exhibited broad substrate specificity, glucosylating sinapate, cinnamate, ferulate, 4-coumarate and caffeate. Indole-3-acetate, 4-hydroxybenzoate and salicylate were not conjugated. The amino acid sequence of the SGT exhibited a distinct sequence identity to putative indole-3-acetate glucosyltransferases from *Arabidopsis thaliana* and a limonoid glucosyltransferase from *Citrus unshiu*, indicating that SGT belongs to a distinct subgroup of glucosyltransferases that catalyze the formation of 1-*O*-acylglucosides (β-acetal esters).

Key words: *Brassica* – Brassicaceae – Glucosyltransferase – Hydroxycinnamates – Phenylpropanoids – Sinapate

The enzyme UDP-glucose:sinapate glucosyltransferase (SGT; EC 2.4.1.120) catalyzes the transfer of glucose from UDP-glucose to sinapate and some other hydroxycinnamates, including 4-coumarate, caffeate and ferulate (Strack 1980; Nurmann and Strack 1981; Wang and Ellis 1998). The enzyme is active in ripening seeds and young seedlings in members of the Brassicaceae. The enzymic product, 1-*O*-sinapoylglucose (β-acetal ester), is the energy-rich acyl donor (Leznicki and Bandurski 1988; Mock and Strack 1993) in reactions leading to sinapine (sinapoylcholine) in the seeds (Strack et al. 1983) and sinapoylmalate in the seedlings (Tkotz and Strack 1980; Mock et al. 1992; Bouchereau et al. 1992; Lorenzen et al. 1996). The sinapate esters in members of the Brassicaceae play an important role in UV protection, as shown with *Arabidopsis* (Landry et al. 1995). Figure 1 shows a scheme of the sinapate metabolism emphasizing the pivotal role of SGT. This enzyme has also been detected outside the Brassicaceae, e.g. in cultures of wild carrot (Halaweish and Dougall 1990).

The 1-*O*-acylglucosides of hydroxycinnamates and other phenolic acids are frequently described as being alternative precursors of phenolic esters (Strack and Mock 1993), besides the respective coenzyme A esters. The pivotal role of 1-*O*-acylglucosides in biosynthetic reactions has also been proven in other metabolic areas, e.g. in biosyntheses of indole-3-acetylinositol and indole-3-acetylglycerol in maize (Kesky and Bandurski 1990; Kowalezyk and Bandurski 1990) or glucose polyesters (isobutyrate esters) in wild tomato (Li et al. 1999).

The first gene in sinapate metabolism, encoding 1-*O*-sinapoylglucose-malate sinapoyltransferase (SMT; EC 2.3.1.92), has recently been cloned from *Arabidopsis* (Lehfeldt et al. 2000). In this communication we describe cloning and expression of a rape glucosyltransferase that catalyzes the formation of 1-*O*-sinapoylglucose. The enzyme is most likely a UDP-glucose:sinapate glucosyltransferase (SGT; EC 2.4.1.120).

Plant material. Rape (*Brassica napus* L. cv. Express) seeds were obtained from the Norddeutsche Pflanzenzucht (Holtsee, Germany). Immature pods were harvested from field-grown plants and the immature (green) seeds removed from the siliques. Cotyledons were excised from 3-d-old seedlings, germinated on filter paper moistened with tap water. The plant material was frozen in liquid nitrogen immediately after harvest and stored at −80 °C.

Construction of the cDNA library and screening for glucosyltransferases. Total RNA from immature seeds and cotyledons was

Abbreviations: GT = glucosyltransferase; PCR = polymerase chain reaction; SGT = UDP-glucose:sinapate glucosyltransferase

The cDNA sequence has been submitted to the GenBank/EBJ data bank with accession number AF287143

Correspondence to: D. Strack;
E-mail: dstrack@ipb-halle.de; Fax: +49-345-5582106

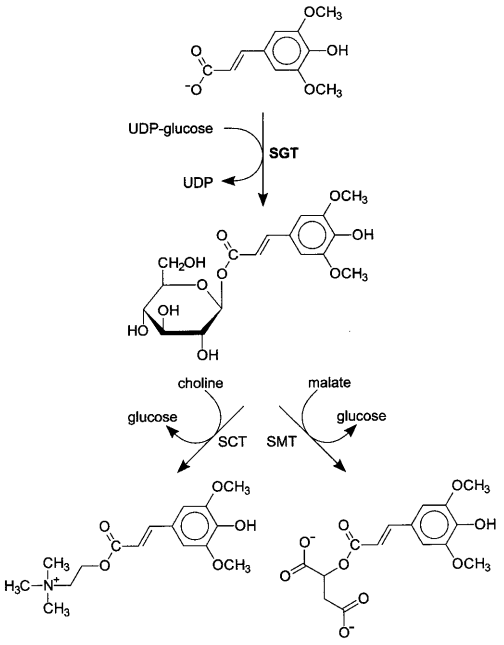


Fig. 1. Scheme of sinapate metabolism emphasizing the pivotal role of UDP-glucose:sinapate glucosyltransferase (*SGT*; EC 2.4.1.120) in the biosynthesis of sinapine (sinapoylcholine) and sinapoylmalate, catalyzed by 1-*O*-sinapoylglucose:choline sinapoyltransferase (*SCT*; EC 2.3.1.91) and 1-*O*-sinapoylglucose:malate sinapoyltransferase (*SMT*; EC 2.3.1.92), respectively

prepared with the Rneasy Plant Mini Kit (Qiagen) and used to purify poly(A)$^+$RNA by selective binding to oligo-dT-Oligotex-beads (Qiagen). The poly(A)$^+$RNA (7.5 µg) from each tissue was used to construct a λ-ZAP cDNA library as described previously (Vogt et al. 1999). Glucosyltransferase-specific sequences were amplified from the cDNA libraries by polymerase chain reaction (PCR) using the degenerate forward primer 5′-ACNCA-YTGYGGNTGGAAC-3′ derived from the 'PSPG box' of plant glucosyltransferases (Hughes and Hughes 1994) and the T7 primer (5′-GTAATACGACTCACTATAGGGC-3′). Screening of the cDNA libraries (40,000 plaque-forming units each) was performed by hybridization with the labeled PCR product mixture. From positive clones, the corresponding recombinant pBluescript phagemids were excised and used for infection of *E. coli* XL-1 blue. Cloned cDNA inserts were subjected to DNA sequence analysis. A similarity search of amino acids and nucleotide sequences of SGT was performed using the different options of the 'Blast'-algorithm (Altschul et al. 1990; Gish and States 1993). Sequence analysis and multiple sequence alignment were performed with the software package DNASTAR (Lasergene, Madison, Wis., USA).

Expression of SGT cDNA in Escherichia coli. The *sgt* coding sequence was amplified from the cloned cDNA fragment by PCR using the forward primer GT1 (5′-ATGGAACTATCATCTTC-TCC-3′) and the T7 primer. The PCR product was inserted into the *Sma*I-linearized vector pQE-32 (Qiagen). The *E. coli* strain M 15 (pREP4) was transformed by the ligation mixture.

Enzyme assay. Isopropylthiogalactoside (IPTG)-induced *E. coli* cells (80 ml) harboring the SGT cDNA were collected, suspended in lysis buffer (0.1 M Mes buffer, pH 6.0), disrupted by sonification and centrifuged. Aliquots of the supernatants were incubated for 30 min at 30 °C with 150 µl assay mixture (4 mM UDP-glucose and 2 mM sinapate in 0.1 M Mes buffer, pH 6.0). The reaction was terminated by adding 10 µl trifluoroacetic acid and the enzymic products analyzed by HPLC. 1-*O*-Sinapoylglucose was identified by liquid chromatography-mass spectrometry (LC-MS) and photodiode array (PDA)-HPLC analyses including authentic 1-*O*-sinapoylglucose. Other enzymic products were tentatively identified by their chromatographic behavior as compared to 1-*O*-sinapoylglucose and rapid hydrolysis by alkaline treatment (adding a few drops of 1 N aqueous NaOH to the assay mixtures). The products were quantified by external standardization with authentic compounds, purchased from Sigma. In HPLC analyses, 20-µl aliquots of the assays mixtures were injected onto a Nucleosil C_{18} column (5 µm; 250 mm long, 4 mm i.d.; Macherey-Nagel) and the assay components separated by using a 15-min linear gradient at 1 ml min^{-1} from 1.5% aqueous H_3PO_4 to 40% aqueous acetonitrile. Compounds were detected photometrically using a photodiode array (PDA) detector (maxplot, 210–500 nm).

Liquid chromatography-mass spectrometry (LC-MS). The negative ion electrospray (ES) mass spectrum and the ES-MS/MS measurement was obtained from a Finnigan MAT (San Jose, USA) TSQ 7000 instrument (electrospray voltage 4.0 kV; heating capillary temperature 220 °C; sheath gas nitrogen) coupled with an UltraPlus MicroLC system equipped with a C_{18} reversed phase column (4 µm; 100 mm long, 1 mm i.d.; SEPSERV; Micro-Tech, Berlin, Germany). The following HPLC gradient elution system was used: a flow rate of 70 µl min^{-1} within 10 min from H_2O-methanol (90:10) (each containing 0.2% acetic acid) to 50:50 and then on hold for a further 10 min. The collision-induced dissociation (CID) mass spectrum of m/z 385 ([M-H]$^-$, sinapoylglucose) was obtained during the HPLC run under the following conditions: collision energy, 20 eV; collision gas, argon; collision pressure, 1.8×10^{-3} Torr. All mass spectra were averaged and background subtracted. ES-CIDMS, m/z (rel. int.): 325 (4), 223 ([M-H-162]$^-$, 26), 205 ([M-H-glucose]$^-$, 100), 190 (6).

Amplification by PCR of rape cDNA sequences using the 'PSPG box'-derived primers in combination with the

```
1    GAA AAA CAG AGC ACA CAG AAG AGA ACC CCT TTT TGA AGG
40   ATG GAA CTA TCA TCT TCT CCT TTA CCT CCT CAT GTT ATG
79   CTT GTA TCC TTC CCA GGA CAA GGC CAC GTT AAT CCA CTT
118  CTT CGT CTC GGC AAG CTC TTA GCT TCG AAG GGT TTA CTC
157  GTC ACT TTT GTC ACC ACA GAA TCA TGG GGC AAA AAG ATG
196  CGA ACC GCC AAC AAG ATT CAA GAC CGA GCC CTC AAA CCT
235  ATC GGT AAA GGT TAT CTC CGG TTC GAT TTC TTC AAC GAT
274  GGG CTC CCT GAA GAC GAC GAT GCA AGC AGA ACC AAC TTA
313  ACC ATC CTC CGA CCA CAA CTC GAG CTG GTC GGA CAA CAA
352  GAG ATC AAA AAC CTC GTG AAA CGT TAC AAG GAA GTG ATG
391  AAA CAG CCC GTG ACG TGT CTC ATC AAC AAC CCT TTC GTC
430  TCT TGG GTC TGT GAC GTA GCC GAA GAT CTT CAA ATC CCC
469  TGT GCT GTT CTC TGG GTC CAG TCT TGT GCT TGC CTA GCT
508  TCT TAT TAT TAT TAC AAC CAC AAG CTT GTC GAC TTC CCG
547  ACC GAA ACA GAT CCC AAG ATC GAT GTC CAG ATC CCG TGC
586  ATG CCT GTC TTG AAA CAC GAC GAG ATC CCT TCT TTC ATC
625  CAT CCT TTT TCA CCT TAT TCG GGT TTA AGA GAA GTG ATC
664  ATT GAT CAG ATC AAA CGT CTT CAC AAG CCT TTC GCT GTT
703  CTC ATC GAT ACT TTC TAC TCC TTG GAG AAA GAT ATC ATC
742  GAC CAC ATG ACA AAC CTC TCT CGC ACC GGC TTT GTC AGA
781  CCG CTC GGA CCG CTT TAC AAA ATG GCC AAA ACG TTG ATT
820  TGT GAT GAC ATC AAA GGA GAT ATG TCT GAG ACG AGG GAT
859  GAC TGC ATG GAG TGG TTA GAC TCG CAG CCT GTT TCG TCC
898  GTT GTT TAC ATC TCA TTT GGT ACC GTG GCT TAC GTG ACA
937  CAA GAA CAG ATC AGC GAG ATT GCG TTA GGC GTT TTA AAC
976  GCT GAC GTT TCG TTC TTG TGG GTG ATA AGA CAA CAA GAA
1015 CTA GGT GTA AAC AAA GAG CGA CAT GTT CTG CCT GAA GAA
1054 CTC AAA GGG AAA GGT AAA GTC ATT GAA TGG TGT TCA CAA
1093 GAG AAA GTC TTA GCT CAT CCT TCT GTG GTT TGT TTC GTG
1132 ACT CAT TGT GGA TGG AAC TCA ACG ATG GAA GCT TTG TCT
1171 AGT GGA GTC CCA ACG GTC TGT TTT CCT CAG TGG GGA GAT
1210 CAA GTC ACC GAC GCT GCT TAC ATG ATC GAC GTG TTC AAG
1249 ACG GGA GTG AGG CTT AGC CGT GGA GAG ACG GAG GAG AGG
1288 GTG GTG CCT AGG GAG GAA GTA GCG GAG AGG CTG AGA GAA
1327 GTT ACG AAA GGA GAG AAA GCG ACG GAG CTG AAG AAG AAT
1366 GCT TTA AAA TGG AAG GAG GAG GCG GAA GCG GCC GTG GCT
1405 CGC GGT GGT TCG TCG GAT CGG AAT CTT GAT GAG TTT GTG
1444 GAA AAG TTG GGC GTG AAA CCT GTG GCT AAA CAG AAC GGA
1483 AGT CTC AAT CAA AAC GGA AGT ATT CAA AAA CTT TTA TTG
1522 CAA AAG TCA TAA ATG TGT TTT GTA TTA TCT GTT GTT TGT
1561 TTT CTG CAA ATC CTG ATA ATT GTG TTT GTT TGT ATG CTC
1600 CCT CCG TTC TTT GTT CTC GTT TTT TTT TTG TAT GGG TTT
1639 GGG GAG ACT AGT TAT CCT TCT TGA ATA AAT GTT AGA TCA
1678 CGA CTT GTG TTC TTA AGA TTA AAA AAA AAA AAA AAA AA
```

Fig. 2. Nucleotide sequence of the SGT cDNA [GenBank accession number AF287143 ("SGT1")]. Start and stop codons are in bold

T7 primer produced DNA fragments ranging from about 400 to 700 bp in length. [The 'PSPG box' is the proposed binding site of UDP-glucose in glucosyltransferases (GTs) (Bairoch 1991; Hughes and Hughes 1994).] To confirm the abundance of GT-specific sequences within this PCR product mixture, all fragments obtained were ligated into the plasmid pGEM-T easy (Qiagen), and randomly selected clones were subjected to DNA sequence analysis. From six PCR products sequenced, five displayed highest identity to GTs of plant metabolism (data not shown).

Hybridization screening of the rape cDNA libraries with the labeled PCR product mixture as a probe resulted in the isolation of a cDNA fragment that revealed an open reading frame encoding a protein of 497 amino acids, that can most likely be ascribed to the SGT. The complete nucleotide sequence of the cDNA is shown in Fig. 2. Alignment of the deduced amino acid sequence and selected GTs of plant origin gave highest sequence identity to GTs catalyzing the formation of β-acetal esters (Fig. 3), i.e. 64, 63 and 61% identity to putative indole-3-acetate GTs (Bevan et al. 1998) as well as 54% to a limonoid GT (Kita et al. 2000). The SGT molecular mass (Da) and the isoelectric point (pI) were calculated to be 55,970 and 6.36, respectively. The SGT sequence was highly abundant among the clones selected from the cDNA library of immature seeds (5 from the total of 9 clones analyzed harbored this sequence), whereas among the clones isolated from cotyledons this sequence was found in 1 of the investigated 11 clones.

To determine whether the isolated cDNA encodes SGT, it was expressed in *E. coli*. Assaying soluble protein with UDP-glucose and sinapate gave the expected enzymic product, i.e. 1-*O*-sinapoylglucose, which was identified by LC-MS analysis and chromatographic comparison (PDA-HPLC) with a standard compound (not shown). Mass spectrometry gave the expected molecular mass of 386 (m/z 385; $[M\text{-}H]^-$) and indicative fragments, e.g. m/z 223 (sinapate) and 205 [M-H-glucose]$^-$. Protein encoded by the cDNA sense constructs showed the expected enzyme activity, whereas neither heat-denatured protein (5 min at 95 °C) nor that from *E. coli* harboring cDNA antisense constructs showed any activity.

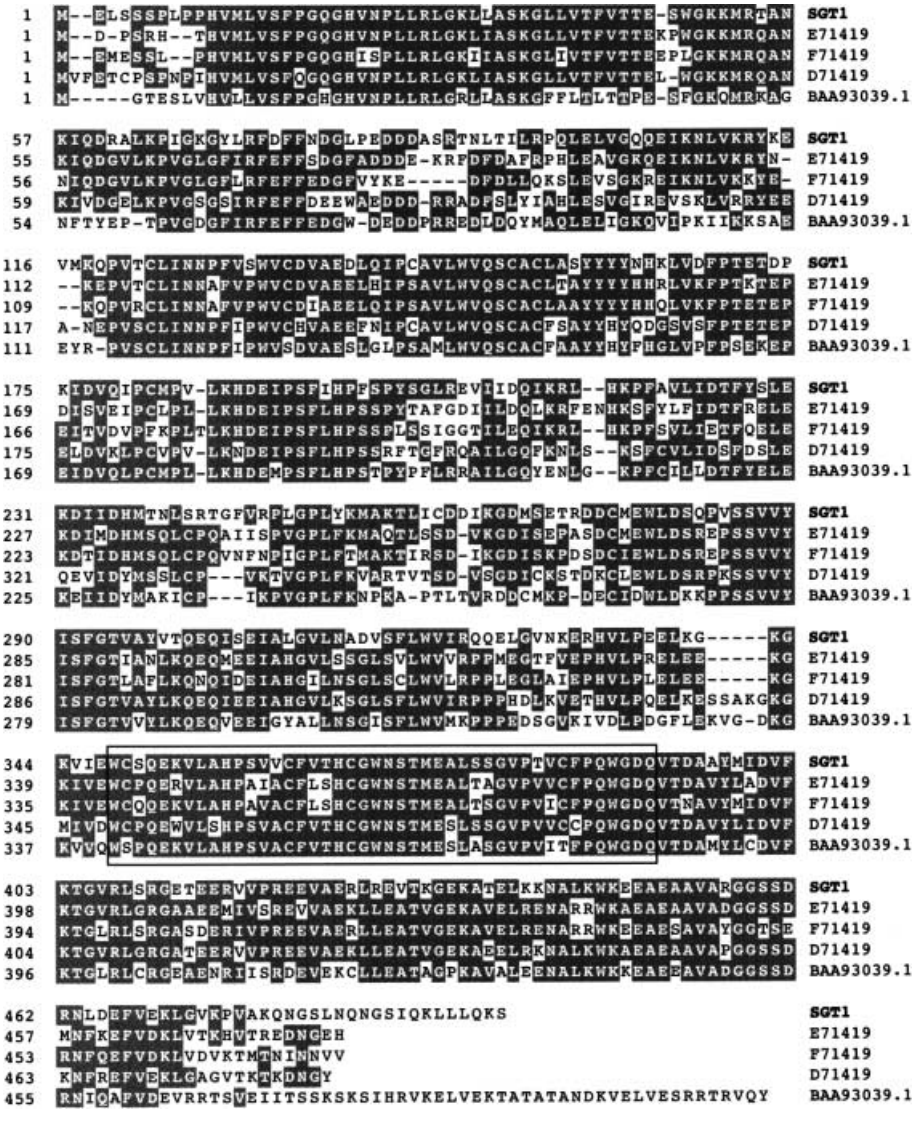


Fig. 3. Alignment of deduced amino acid sequences (single-letter code) of the SGT ("SGT1") and selected GTs of plant origin. *Black shading* shows amino acid identities. In addition, the 'PSPG box', the proposed binding site of UDP-glucose (Bairoch 1991; Hughes and Hughes 1994), is framed. GenBank accession numbers are: E71419, F71419, D71419 ("probable indole-3-actetate beta-glucosyltransferases" from *Arabidopsis thaliana*; Bevan et al. 1998); BAA93039.1 ("limonoid UDP-glucosyltransferase" from *Citrus unshiu*; Kita et al. 2000)

Table 1. Acceptor specificity of the recombinant SGT

Substrate[a]	Relative activity (%)
Sinapate	100[b]
Cinnamate	20
Ferulate	17
4-Coumarate	13
Caffeate	10
Indole-3-acetate	n.d.[c]
4-Hydroxybenzoate	n.d.[c]
Salicylate	n.d.[c]

[a] 2 mM at 4 mM UDP-glucose
[b] 297 pkat (mg protein)$^{-1}$
[c] Not detected

Testing some other potential acceptors (Table 1), the recombinant SGT exhibited highest activity towards sinapate [297 pkat (mg protein)$^{-1}$, 100%] followed by cinnamate (20%), ferulate (17%), 4-coumarate (13%) and caffeate (10%). Indole-3-acetate, 4-hydroxybenzoate or salicylate were not conjugated.

In summary, our results indicate that SGT belongs to a distinct subgroup of GTs catalyzing the formation of 1-*O*-acylglucosides (β-acetal esters). Detailed analyses of the purified SGT and the expression of the corresponding gene in planta are in progress.

The authors thank T. Vogt [Leibniz-Institut für Pflanzenbiochemie (IPB), Halle, Germany] for valuable discussions and T. Kutchan (IBP) for critical reading of the manuscript. We also thank J. Schmidt (IPB) for MS analysis. This work is part of the research project "NAPUS 2000 – Gesunde Lebensmittel aus transgener Rapssaat", supported by the Bundesministerium für Bildung und Forschung.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–410

Bairoch A (1991) PROSITE: a dictionary of sites and patterns in proteins. Nucleic Acid Res 19: 2241–2245

Bevan M, et al. (1998) Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. Nature 391: 485–488

Bouchereau A, Hamelin J, Renard M, Larher F (1992) Structural changes in sinapic acid conjugates during development of rape. Plant Physiol Biochem 30: 467–475

Gish W, States DJ (1993) Identification of protein coding regions by database similarity search. Nat Genet 3: 266–272

Halaweish FT, Dougall DK (1990) Sinapoyl glucose synthesis by extracts of wild carrot cell cultures. Plant Sci 71: 179–184

Hughes J, Hughes MA (1994) Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava. DNA Sequence 5: 41–49

Kesky JM, Bandurski RS (1990) Partial purification and characterization of indol-3-ylacetylglucose:myo-inositol indol-3-ylacetyltransferase (indoleacetic acid-inositol synthase). Plant Physiol 94: 1598–1604

Kita M, Hirata Y, Moriguchi T, Endo-Inagaki T, Matsumoto R, Hasegawa S, Suhayda CG, Omura M (2000) Molecular cloning and characterization of a novel gene encoding limonoid UDP-glucosyltransferase in *Citrus*. FEBS Lett 469: 173–178

Kowalezyk S, Bandurski RS (1990) Isomerization of 1-*O*-indol-3-ylacetyl-β-D-glucose. Enzymatic hydrolysis of 1-*O*-, 4-*O*-, and 6-*O*-indol-3-ylacetyl-β-D-glucose and synthesis of indole-3-acetyl glycerol by a hormone metabolizing complex. Plant Physiol 94: 4–12

Landry LG, Chapple CCS, Last R (1995) *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. Plant Physiol 109: 1159–1166

Lehfeldt C, Shirley AM, Meyer K, Ruegger MO, Cusumano JC, Viitanen PV, Strack D, Chapple C (2000) Cloning of the *SNG1* gene of *Arabidopsis* reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. Plant Cell 12: 1295–1306

Leznicki AJ, Bandurski RS (1988) Enzymic synthesis of indole-3-acetyl-1-*O*-β-D-glucose. II. Metabolic characterization of the enzyme. Plant Physiol 88: 1481–1485

Li AX, Eanetta N, Changas GS, Steffens JC (1999) Glucose polyester biosynthesis. Purification and characterization of a glucose acyltransferase. Plant Physiol 121: 453–460

Lorenzen M, Racicot V, Strack D, Chapple C (1996) Sinapic acid ester metabolism in wild type and a sinapoylglucose-accumulating mutant of *Arabidopsis*. Plant Physiol 112: 1625–1630

Mock H-P, Strack D (1993) Energetics of the uridine 5′-diphosphoglucose: hydroxycinnamic acid acyl-glucosyltransferase reaction. Phytochemistry 32: 575–579

Mock H-P, Vogt T, Strack D (1992) Sinapoylglucose: malate sinapoyltransferase activity in *Arabidopsis thaliana* and *Brassica rapa*. Z Naturforsch Teil C 47: 680–682

Nurmann G, Strack D (1981) Formation of 1-sinapoylglucose by UDP-glucose: sinapic acid glucosyltransferase from cotyledons of *Raphanus sativus*. Z Pflanzenphysiol 102: 11–17

Strack D (1980) Enzymatic synthesis of 1-sinapoylglucose from free sinapic acid and UDP-glucose by a cell-free system from *Raphanus sativus* seedlings. Z Naturforsch Teil C 35: 204–208

Strack D, Mock HP (1993) Hydroxycinnamic acids and lignins. In: Lea PJ (ed) Enzymes in secondary metabolism, vol 9. Methods in plant biochemistry. Dey PM, Harborne JB (eds). Academic Press, London, San Diego, pp 45–97

Strack D, Knogge W, Dahlbender B (1983) Enzymatic synthesis of sinapine from 1-*O*-sinapoyl-β-D-glucose and choline by a cell-free system from developing seeds of red radish (*Raphanus sativus* L. var. *sativus*). Z Naturforsch Teil C 38: 21–27

Tkotz N, Strack D (1980) Enzymatic synthesis of sinapoyl-L-malate from 1-sinapoylglucose and L-malate by a protein preparation from *Raphanus sativus* cotyledons. Z Naturforsch Teil C 35: 835–837

Vogt T, Grimm R, Strack D (1999) Cloning and expression of a cDNA encoding betanidin 5-*O*-glucosyltransferase, a betanidin- and flavonoid-specific enzyme with high homology to inducible glucosyltransferases from the Solanaceae. Plant J 19: 509–519

Wang SX, Ellis BE (1998) Enzymology of UDP-glucose: sinapic acid glucosyltransferase from *Brassica napus*. Phytochemistry 49: 307–318

Note added in proof. Recent data base search showed highest identity of SGT1 (81% on amino acid level) to a new entry (Accession No. BAB00006; "indole-3-acetate beta-glucosyltransferase-like protein" from *Arabidopsis*).

Correspondence

Identification of four *Arabidopsis* genes encoding hydroxycinnamate glucosyltransferases

Carsten Milkowski, Alfred Baumert, Dieter Strack*

First published online 23 November 2000

Members of the Brassicaceae accumulate sinapate esters as major phenylpropanoid secondary metabolites, which are fluorescent UV-protective compounds, as shown with *Arabidopsis thaliana* [1]. A pivotal enzyme in sinapate ester biosynthesis is UDP-glucose:sinapate glucosyltransferase (SGT, EC 2.4.1.120) which catalyzes the transfer of glucose from UDP-glucose to sinapate and some other hydroxycinnamates (HCAs), including 4-coumarate, caffeate and ferulate [2]. 1-*O*-Sinapoylglucose is the immediate acyl donor in reactions leading to sinapine (sinapoylcholine) in seeds [3] and sinapoylmalate plants [4,5].

The first gene in sinapate ester biosynthesis has been cloned from *Arabidopsis* [6] and encodes 1-*O*-sinapoylglucose:malate sinapoyltransferase (SMT, EC 2.3.1.92). The second gene involved in this metabolism has very recently been cloned from rape (*Brassica napus*) [7]. It encodes UDP-glucose:SGT. Here we report the cloning and identification of four *Arabidopsis* glucosyltransferase (GT) genes encoding cinnamate and HCA GTs. These genes have previously been putatively assumed to encode indole-3-acetate (IAA) GTs ([8]; Y. Nakamura, NCBI protein database accession number BAB00006). We unambiguously demonstrate, however, that the heterologously expressed *Arabidopsis* enzymes do not catalyze the synthesis of the IAA glucose ester, but rather specifically cinnamate and HCA glucose esters. This is a prominent example of misidentification of new genes. It shows again that the assignment of genes without functional expression carries a high risk of error.

As the corresponding nucleotide sequences retrieved from the databases reveal no introns, the *Arabidopsis* genes encoding the proteins BAB00006, D71419, E71419 and F71419 were amplified by polymerase chain reaction (PCR) with genomic *Arabidopsis* DNA as template. Genomic DNA was isolated according to Brandstädter et al. [9] from rosette leaves of 5-week-old *Arabidopsis* plants (*Arabidopsis thaliana* Heynh. ecotype Columbia) grown in the green house. The following specific primers were used: BAB00006: 5′-ATG GAG CTA GAA TCT TCT CC-3′ (forward) and 5′-TTA AAA GCT TTT GAT TGA TCC-3′ (reverse); E71419: 5′-ATG GAC CCG TCT CGT CAT ACT CAT G-3′ (forward) and 5′-CTA GTG TTC TCC GTT GTC TTC TCT CG-3′ (reverse); F71419: 5′-ATG GAG ATG GAA TCG TCG TTA CCT C-3′ (forward) and 5′-TTA CAC GAC ATT ATT AAT GTT TGT C-3′ (reverse); D71419: 5′-ATG GTG TTC GAA ACT TGT CCA TCT CC-3′ (forward) and 5′-CTA GTA TCC ATT ATC TTT AGT CTT CG-3′ (reverse). The PCR products were inserted into the *Sma*I-linearized vector pQE-32 (Qiagen). Sense and antisense constructs were used to transform the *Escherichia coli* strain M15 (pREP4).

E. coli M15 (pREP4) cells harboring the *Arabidopsis* genes on the pQE-32 vector were grown to logarithmic phase. Expression of the recombinant *Arabidopsis* proteins was then induced by incubating the cells for another 20 h at 30°C in the presence of 1 mM isopropylthiogalactoside (IPTG).

The enzyme assays were performed as described [7]. In short: IPTG-induced *E. coli* cells harboring the *Arabidopsis* GT genes were collected, suspended in lysis buffer (0.1 M Mes buffer, pH 6.0), disrupted by sonification and centrifuged. Aliquots of the supernatants were incubated for 30 min at 30°C with 4 mM UDP-glucose and various acceptor molecules (cinnamate, HCAs, benzoate, hydroxybenzoates or IAA, 2 mM each) in the presence of 0.1 M Mes buffer (pH 6.0). The reactions were terminated by adding trifluoroacetic acid. Enzymatic products were identified by liquid chromatographic-mass spectrometry (LC-MS) and photodiode array-high performance liquid chromatography (PDA-HPLC) analyses. Quantification was achieved by external standardization with cinnamate and HCAs purchased from Sigma. Further details on PDA-HPLC and LC-MS are described elsewhere [7].

In a recent paper [7] we presented results on cloning and heterologous expression of a rape cDNA encoding UDP-glucose:sinapate GT (SGT1). The rape SGT1 sequence showed 64, 63 and 61% sequence identities to putative *Arabidopsis* IAA GTs [7]. In addition, a recent data base search showed the highest identity of rape SGT1 (81% on amino acid level) to a new entry assigned also to an *Arabidopsis* IAA GT-like protein (Y. Nakamura, NCBI protein database accession number BAB00006). These results prompted us to clone and functionally express the four *Arabidopsis* genes in *E. coli*, expecting to find rape SGT1-like enzymes.

We amplified the four respective genes by PCR with genomic DNA isolated from rosette leaves of 5-week-old *Arabidopsis* plants. These genes were introduced into *E. coli* cells as sense and antisense constructs and their expression induced by the addition of IPTG. Crude *E. coli* protein extracts were assayed for GT activities. Testing potential acceptors, it was found that none of the recombinant enzymes was active towards IAA (Table 1). Thus, the earlier putative assignments of these genes are incorrect. Instead, the recombinant enzymes catalyzed specifically the formation of cinnamate glucose ester (1-*O*-cinnamoylglucose) and HCA glucose esters (1-*O*-4-coumaroyl-, -caffeoyl-, -feruloyl- and -sinapoylglucose) with markedly different acceptor specificities. Possible glucose esterification of benzoate, 4-hydroxybenzoate and salicylate was not observed.

Identification of the products, cinnamoyl- and HCA-glucose esters, was achieved by chromatographic comparison (PDA-HPLC) with standard compounds available from a previous study [7] and by LC-MS. The latter gave the expected molecular mass for cinnamoylglucose of 310, for 4-coumaroylglucose of 326, for caffeoylglucose of 342; for feruloylglucose of 356, and for sinapoylglucose of 386. Neither heat-denatured proteins (5 min at 95°C) nor that from *E. coli*

0014-5793/00/$20.00 © 2000 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Table 1
Acceptor specificity of the recombinant *Arabidopsis* GTs in 0.1 M Mes buffer (pH 6.0)

Substrates[a]	Relative activities (%)			
	AtSGT1 (BAB00006)[b]	AtHCAGT1 (F71419)[b]	AtHCAGT2 (D71419)[b]	AtHCAGT3 (E71419)[b]
Sinapate	100[c]	100[d]	57	32
Ferulate	7	77	77	94
Cinnamate	–[g]	24	24	76
4-Coumarate	–	66	100[e]	100[f]
Caffeate	–	23	72	45
Benzoate	–	–	–	–
4-Hydroxybenzoate	–	–	–	–
Salicylate	–	–	–	–
Indole-3-acetate (IAA)	–	–	–	–

[a]2 mM at 4 mM UDP-glucose.
[b]NCBI protein database accession numbers.
[c]35.5 μkat kg^{-1} protein.
[d]17.9 μkat kg^{-1} protein.
[e]173 μkat kg^{-1} protein.
[f]87.6 μkat kg^{-1} protein.
[g]–, no product detected by the HPLC system applied.

harboring cDNA antisense constructs showed any GT activities.

One protein (accession number BAB00006 with 81% identity and 87% similarity to rape SGT1) exhibited unexpectedly high specificity towards sinapate. Only ferulate was also esterified, but with a relative activity of only 7% compared to sinapate. Cinnamate, 4-coumarate and caffeate were not accepted at all by this protein (based on the HPLC system applied). In contrast, the other enzymes showed broad acceptor specificities. F71419 (63% identity and 76% similarity to rape SGT1) catalyzed the esterification of sinapate (100%), ferulate (77%), 4-coumarate (66%), cinnamate (24%) and caffeate (23%). E71419 (64% identity and 76% similarity to rape SGT1) and D71419 (61% identity and 74% similarity to rape SGT1) converted cinnamate and all the HCAs tested, however, with highest activities towards 4-coumarate (Table 1). In conclusion, BAB00006 can be classified as UDP-glucose:SGT (AtSGT), E71419, F7141 and D71419 as UDP-glucose:HCA GTs (AtHCAGTs). We propose that the four *Arabidopsis* genes and the previously identified rape SGT1 gene [7] as well as a limonoid GT from *Citrus unshiu* [10] belong to a distinct subgroup of GTs catalyzing the formation of 1-*O*-acylglucosides (β-acetal esters).

We thank Jürgen Schmidt (IPB) for MS analysis. We also thank Toni Kutchan (IPB) for critical reading of the manuscript. This work is part of the research project 'NAPUS 2000 – Gesunde Lebensmittel aus transgener Rapssaat', supported by the Bundesministerium für Bildung und Forschung.

References

[1] Landry, L.G., Chapple, C.C.S. and Last, R. (1995) Plant Physiol. 109, 1159–1166.
[2] Strack, D. (1980) Z. Naturforsch. 35c, 204–208.
[3] Strack, D., Knogge, W. and Dahlbender, G. (1983) Z. Naturforsch. 38c, 21–27.
[4] Tkotz, N. and Strack, D. (1980) Z. Naturforsch. 35c, 835–837.
[5] Mock, H.-P., Vogt, T. and Strack, D. (1992) Z. Naturforsch. 47c, 680–682.
[6] Lehfeldt, C., Shirley, A.M., Meyer, K., Ruegger, M.O., Cusumano, J.C., Viitanen, P.V., Strack, D. and Chapple, C. (2000) Plant Cell 12, 1295–1306.
[7] Milkowski, C., Baumert, A. and Strack, D. (2000) Planta 211, 883–886.
[8] Bevan, M. et al. (1998) Nature 391, 485–488.
[9] Brandstädter, J., Roßbach, C. and Theres, K. (1994) Planta 192, 69–74.
[10] Kita, M., Hirata, Y., Moriguchi, T., Endo-Inagaki, T., Matsumoto, R., Hasegawa, S., Suhayda, C.G. and Omura, M. (2000) FEBS Lett. 469, 173–178.

*Corresponding author. Fax: (49)-345-5582 106.
E-mail: dstrack@ipb-halle.de

Abteilung Sekundärstoffwechsel, Institut für Pflanzenbiochemie, Weinberg 3, 06120 Halle (Saale), Germany

PII: S0014-5793(00)02270-5

The Plant Journal (2004) **38**, 80–92 doi: 10.1111/j.1365-313X.2004.02036.x

Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation

Carsten Milkowski, Alfred Baumert, Diana Schmidt, Lilian Nehlin and Dieter Strack*
Department of Secondary Metabolism, Leibniz Institute of Plant Biochemistry, Weinberg 3, D-06120 Halle (Saale), Germany

Received 10 October 2003; revised 9 December 2003; accepted 11 December 2003.
*For correspondence (fax +49 345 5582 1509; e-mail dstrack@ipb-halle.de).

Summary

Members of the Brassicaceae family accumulate specific sinapate esters, i.e. sinapoylcholine (sinapine), which is considered as a major antinutritive compound in seeds of important crop plants like *Brassica napus*, and sinapoylmalate, which is implicated in UV-B tolerance in leaves. We have studied the molecular regulation of the sinapate ester metabolism in *B. napus*, and we describe expression of genes, some properties of the encoded proteins and profiles of the metabolites and enzyme activities. The cloned cDNAs encoding the key enzymes of sinapine biosynthesis, UDP-glucose (UDP-Glc):*B. napus* sinapate glucosyltransferase (BnSGT1) and sinapoylglucose:*B. napus* choline sinapoyltransferase (BnSCT), were functionally expressed. BnSGT1 belongs to a subgroup of plant GTs catalysing the formation of 1-*O*-hydroxycinnamoyl-β-D-glucoses. BnSCT is another member of serine carboxypeptidase-like (SCPL) family of acyltransferases. The *B. napus* genome contains at least two *SGT* and *SCT* genes, each derived from its progenitors *B. oleracea* and *B. rapa*. BnSGT1 and BnSCT activities are subjected to pronounced transcriptional regulation. *BnSGT1* transcript level increases throughout early stages of seed development until the early cotyledonary stage, and stays constant in later stages. The highest level of *BnSGT1* transcripts is reached in 2-day-old seedlings followed by a dramatic decrease. In contrast, expression of *BnSCT* is restricted to developing seeds. Regulation of gene expression at the transcript level seems to be responsible for changes of BnSGT1 and BnSCT activities during seed and seedling development of *B. napus*. Together with sinapine esterase (SCE) and sinapoylglucose:malate sinapoyltransferase (SMT), activities of BnSGT1 and BnSCT show a close correlation with the accumulation kinetics of the corresponding metabolites.

Keywords: sinapate ester metabolism, *Brassica napus*, glucosyltransferases, serine carboxypeptidase-like acyltransferases, molecular evolution.

Introduction

In members of the Brassicaceae family, sinapoylglucose, sinapoylcholine (sinapine) and sinapoylmalate are components of a pathway (Figure 1) catalysed by four enzymes, a glucosyltransferase (sinapate glucosyltransferase (SGT)), two hydroxycinnamoyltransferases (choline sinapoyltransferase (SCT) and malate sinapoyltransferase (SMT)) and a sinapine esterase (SCE). Sinapate, supplied by the phenylalanine/hydroxycinnamate pathway during early seed development (Bopp and Lüdicke, 1980), is conjugated via sinapoylglucose with choline. During seed germination, an esterase-catalysed hydrolysis liberates sinapate (Strack *et al.*, 1980), which is conjugated a second time via sinapoylglucose with malate in the seedling.

It is documented with *Arabidopsis thaliana*, *Raphanus sativus* and some *Brassica* species that the glucose ester (β-acetal ester) of sinapate (1-*O*-sinapoyl-β-D-glucose) is the sinapoyl donor in the biosynthesis of sinapine, catalysed by SCT (Regenbrecht and Strack, 1985; Shirley *et al.*,

© 2004 Blackwell Publishing Ltd

2001; Strack *et al.*, 1983, 1990). This was also shown for the formation of sinapoylmalate in the leaves, catalysed by SMT (Lorenzen *et al.*, 1996; Mock *et al.*, 1992; Tkotz and Strack, 1980). The energy-rich sinapoyl donor (Mock and Strack, 1993) is formed via SGT activity (Strack, 1980; Wang and Ellis, 1998).

In *A. thaliana*, sinapoylmalate and other sinapate esters have been shown to function as UV-protecting compounds in leaf tissues (Booij-James *et al.*, 2000; Landry *et al.*, 1995; Li *et al.*, 1993; Sheahan, 1996). The role of sinapine, however, has not been elucidated so far. Nevertheless, it has been shown that sinapine contributes to the supply of choline for the synthesis of phosphatidylcholine in *R. sativus* seedlings (Strack, 1981).

The metabolism of sinapate esters in members of the Brassicaceae family, especially *Brassica napus*, has been of scientific interest for many years. Previous studies tended to be mainly descriptive because of lack of information about the genes encoding the corresponding enzymes. Attempts to purify these enzymes from plant tissues for protein sequence determination failed (Vogt *et al.*, 1993; Wang and Ellis, 1998), thus preventing isolation of the corresponding cDNAs by the methods of reverse genetics. However, recent advances in the characterisation of plant enzymes from secondary metabolism, especially GTs (Hughes and Hughes, 1994; Lim *et al.*, 2001) and acyltransferases (Lehfeldt *et al.*, 2000; Li and Steffens, 2000), paved the way for homology-based cloning techniques. By use of the highly conserved plant secondary product GT (PSPG) motif, found in plant GTs (Hughes and Hughes, 1994), a cDNA encoding the *B. napus* SGT (BnSGT1) could be isolated (Milkowski *et al.*, 2000a). A similar approach led to the identification of related enzymes from *A. thaliana* (Lim *et al.*, 2001; Milkowski *et al.*, 2000b).

With regard to acyltransferases accepting 1-*O*-acyl-β-D-glucoses, first molecular insights came from investigations of acyltransferases from wild tomato (*Lycopersicon pennellii*; Li and Steffens, 2000) and *A. thaliana* mutants defective in *SMT* and *SCT* genes (Lehfeldt *et al.*, 2000; Shirley *et al.*, 2001). It was shown that these enzymes constitute a new class of serine carboxypeptidase-like (SCPL) proteins (SCPL acyltransferases), co-opted to accept acylglucoses as substrates instead of polypeptides.

The present work documents the close correlation between the metabolism of sinapate esters, the enzyme activities involved and steady state transcript levels of the corresponding genes, *BnSGT1* and *BnSCT*, during *B. napus* seed and seedling development. Cloned cDNAs of these genes were functionally expressed and some properties of the recombinant proteins were determined. Sequence analysis classifies BnSCT as a new SCPL acyltransferase. The preceding enzyme in sinapine and sinapoylmalate biosynthesis, BnSGT1, belongs to a distinct subgroup of GTs catalysing the formation of 1-*O*-acylglucoses.

Results and discussion

Cloning and sequence analysis of BnSGT1 *and* BnSCT

Cloning of a cDNA-encoding BnSGT1 has been previously published (Milkowski *et al.*, 2000a). For the BnSGT1 protein, a molecular mass of 55.97 kDa and an isoelectric point (p*I*) of 6.1 were calculated. *BnSCT* cDNA was isolated by screening a cDNA library from immature *B. napus* seeds with a SCPL protein-specific probe generated by RT-PCR with seed-specific RNA. Full-length cDNA was generated by 5′ rapid amplification of cDNA ends (RACE). The cDNA sequence of *BnSCT* carries a reading frame of 466 amino acid residues (Figure 2). The deduced protein displays characteristics of SCPL proteins. It harbours the typical amino acid residues (S, D, H) forming the serine-based catalytic triad (Bech and Breddam, 1989; Hayashi *et al.*, 1973, 1975; Liao and Remington, 1990; Liao *et al.*, 1992) at conserved positions. According to MEROPS peptidase database (http://merops.sanger.ac.uk), BnSCT can be classified as a member of the family S10 non-peptidase homologues.

Computational analysis of the N-terminal sequence suggests a signal peptide that is probably cleaved after residue D19 to form a protein of a calculated molecular mass of 50.27 kDa and a p*I* of 6.64. Two potential N-glycosylation sites were identified by sequence analysis with prediction software (Gupta *et al.*, 2002). These features suggest that BnSCT is synthesised as a precursor protein that is targeted to the endoplasmic reticulum where the signal peptide is removed. As has been concluded for the related *A. thaliana* SMT (Hause *et al.*, 2002), the BnSCT is probably glycosylated in the Golgi apparatus and is subsequently routed to the vacuole. Experiments to immunolocalise BnSCT in *B. napus* embryos are in progress.

Genomic organisation of BnSGT1 *and* BnSCT *genes*

We carried out Southern blot analyses to estimate gene copy number of *BnSGT1* and *BnSCT*, and to characterise the association of these genes to the diploid progenitors of *B. napus* (Figure 3). With the complete *BnSGT1* cDNA as probe, two major bands were detected in restriction patterns of genomic DNA from *B. napus* (Figure 3a). Restriction was carried out with enzymes (*Nco*I, *Spe*I) for which no recognition sites are present in the *BnSGT1* reading frame. Analysis of *B. oleracea* produced only one major band corresponding in size to the larger *B. napus* fragment, and the genomic DNA from *B. rapa* showed a single major band corresponding to the smaller *B. napus* fragment. Moreover, the presence of minor bands suggests that the *Brassica* genomes investigated contain several *SGT1*-related genes.

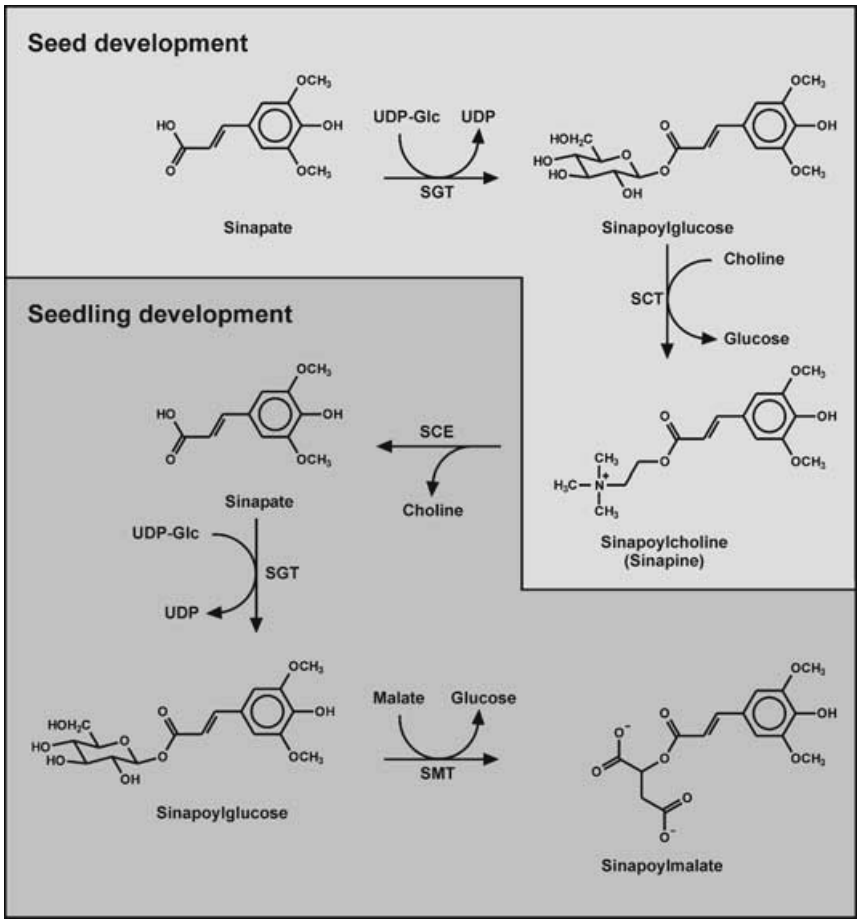


Figure 1. Biosynthetic scheme of the three major sinapate esters in seeds and seedlings of members of the Brassicaceae.
The enzymes involved are SGT (EC 2.4.1.120), SCT (EC 2.3.1.91), SCE (EC 3.1.1.49), and SMT (EC 2.3.1.92).

Southern blot analysis of *BnSCT* was performed using the genomic *BnSCT* fragment as the probe (Figure 3b). Genomic DNA from *B. napus* and its progenitors was digested with endonuclease *Xho*I for which the *BnSCT1* fragment carries a single recognition site near the 5′ end. Analysis of *B. napus* DNA detected two bands; the larger one corresponds to the single signal found in the *B. oleracea* genome. As with *B. oleracea*, a single *BnSCT* fragment was also detected in genomic DNA from *B. rapa*. The *BnSCT* fragment from *B. rapa* was shown to be larger than the corresponding *B. napus* fragment. This could possibly be explained by re-arrangements in the genomes of *B. napus* or *B. rapa*.

These results reflect the fact that the genome of the amphidiploid *B. napus* ($n = 19$) contains the genomes of *B. oleracea* ($n = 9$) and *B. rapa* ($n = 10$; Schenck and Röbbelen, 1982). The two *SGT* and *SCT* genes most likely derive from its progenitors.

An important component of metabolic regulation is the number of genes coding for enzymes that catalyse the same reaction. In *A. thaliana*, besides the sinapate-specific AtSGT1, three other ester-forming GTs (HCAGT1–3) that accept hydroxycinnamates without pronounced preference have been found (Lim *et al.*, 2001; Milkowski *et al.*, 2000b). The genomic Southern analyses showed additional signals originating from related genes (Figure 3a), and therefore we assume that the genomic organisation of genes encoding ester-forming GTs in *B. napus* corresponds to that of *A. thaliana*. This assumption is supported by results from measurements of SGT activities in crude protein extracts. We found significant changes in SGT substrate specificity towards hydroxycinnamates in protein extracts from seeds and seedlings at different developmental stages (not shown).

Expression of BnSGT1 and BnSCT in Escherichia coli

To show that the cloned cDNAs from *B. napus* encode the expected functional enzymes, cDNA fragments carrying the information for the mature BnSGT1 and BnSCT proteins were cloned into expression vector pET28a, which adds a C-terminal His tag to the expressed proteins. Expression was performed in the *E. coli* strain BL21-CodonPlus™ (DE3)-RP, which compensates for rarely used codons in the bacteria. BnSGT1 protein could not be detected on SDS–PAGE after Coomassie staining. Western blot analysis with anti-His antibodies, however, produced signals with both the insoluble and soluble protein fractions.

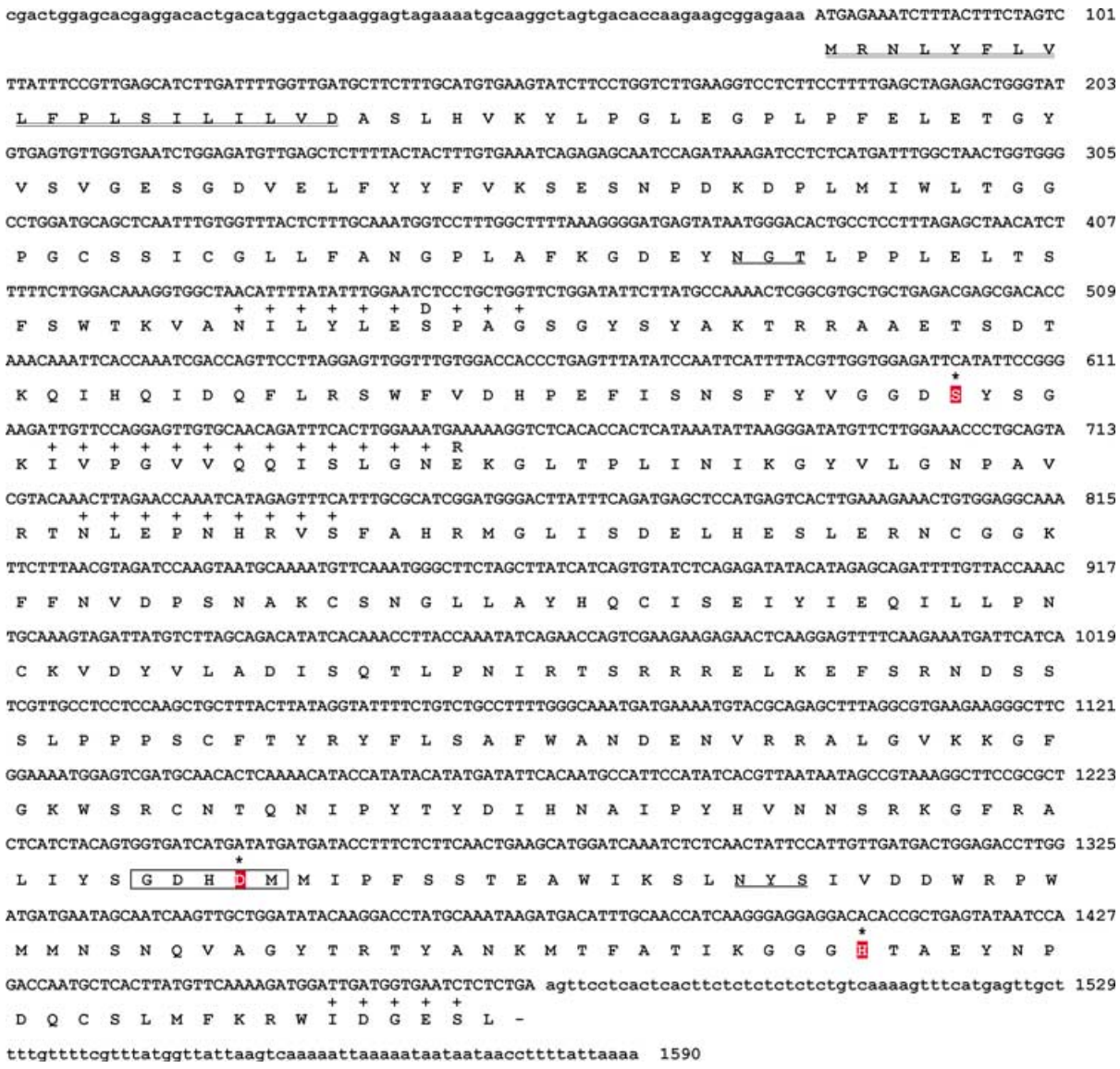


Figure 2. *BnSCT* cDNA and its deduced amino acid sequence.
The predicted N-terminal leader sequence is double underlined. Amino acid residues forming the catalytic triad of SCPs (S, D, H) are marked by asterisks and red shading. Peptide sequences determined by Vogt *et al.* (1993) are indicated (+); amino acid residues that do not fit are designated (D, R). Sites of potential *N*-glycosylation are underlined. The boxed amino acid motif was used to derive degenerate primers for PCR. The binding site of the gene-specific primer used for RACE experiments to amplify the 5′-part of the sequence reached from G(1093) to G(1117).

Comparison of the staining intensities indicated a high portion of the recombinant protein present in the insoluble fraction (Figure S1).

SGT activity was proven by incubating protein extracts with sinapate and UDP-Glc. The product could be identified as sinapoylglucose as described by Milkowski *et al.* (2000a). Heat-denatured protein and extracts from bacteria harbouring the empty vector pET28a failed to produce sinapoylglucose in enzyme assays.

Kinetic analysis suggested that the mechanism of BnSGT1 catalysis fits a random bi–bi model, as proposed for the native *B. napus* enzyme (Wang and Ellis, 1998). As we were unable to purify the recombinant enzyme in amounts necessary for this analysis, the kinetic data came from crude *E. coli* protein extracts.

The calculated apparent K_m values of the BnSGT1 for UDP-Glc and the hydroxycinnamates were found to be independent of the concentration of the second substrate (not documented). The mean apparent K_m value for UDP-Glc in the presence of different sinapate concentrations in the range of 0.05–2 mM was about 1 mM. In the presence of different UDP-Glc concentrations in the range of 0.5–4 mM, the apparent K_m for sinapate was about 0.72 mM, for ferulate 1.7 mM, for caffeate 2.3 mM and for 4-coumarate

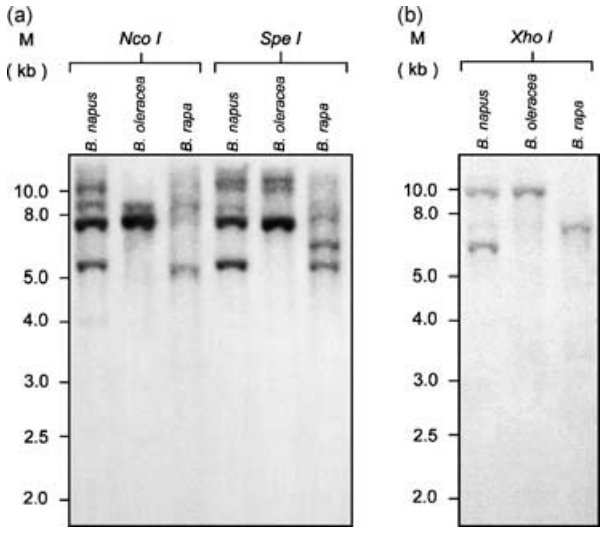


Figure 3. Genomic Southern analysis of *SGT* (a) and *SCT* (b) genes.
DNA (5 μg) was isolated from seedlings of *B. napus*, *B. oleracea* and *B. rapa*, digested with the restriction enzymes as indicated and subjected to Southern analysis with *BnSGT* cDNA (a) and *BnSCT* cDNA (b) as probes.

3.6 mM. BnSGT1 activity showed by far the highest specificity towards sinapate, determined from the reaction efficiencies (V_{max}/K_m). Compared to sinapate (100%), the efficiencies for ferulate, caffeate and 4-coumarate were only 4, 2 and 1%, respectively.

SDS–PAGE analysis of proteins extracted from *E. coli* cells carrying the pET28a-*BnSCT* plasmid showed a strongly induced band in the insoluble protein fraction of induced cells. This band gave a signal in Western blot analysis with anti-His antibodies, indicating that it contains the recombinant BnSCT protein (Figure S2). The protein size corresponds to the calculated mass for the non-glycosylated mature protein. With soluble protein, no conversion of sinapoylglucose and choline to sinapine could be observed. This revealed that in *E. coli* expressing *BnSCT* cDNA, the protein is exclusively present in inclusion bodies. To test functionality of the BnSCT, protein was isolated from purified inclusion bodies and subjected to refolding procedures. When the refolded protein was incubated with sinapoylglucose and choline, a product was found that co-chromatographed on HPLC with authentic sinapine giving identical online UV spectral data. Unambiguous identification was achieved by mass spectrometry. Selected reaction monitoring (SRM), using the reaction *m/z* 310 ($[M]^+$)→*m/z* 251 ($[M^-(CH_3)_3N]^+$), gave identical results in comparison with authentic sinapine. No products were detected in assays with heat-denatured protein or protein subjected to refolding procedures that came from cells transformed with the empty vector.

Accumulation of BnSGT1 *and* BnSCT *transcripts*

To study the regulation of sinapate ester metabolism in *B. napus*, we investigated the expression of the *BnSGT1* and *BnSCT* genes during seed and seedling development (Figure 4). The transcript levels of *BnSGT1* and *BnSCT* seem to determine changes in the corresponding enzyme activities in the developmental regulation of this sinapate ester metabolism (compare Figures 4 and 5). At early embryogenesis stages, i.e. globular, torpedo or late torpedo stage, there was no detectable expression of *BnSCT*, whereas the *BnSGT1* transcript could be found in small quantity. Expression of *BnSGT1* increased until

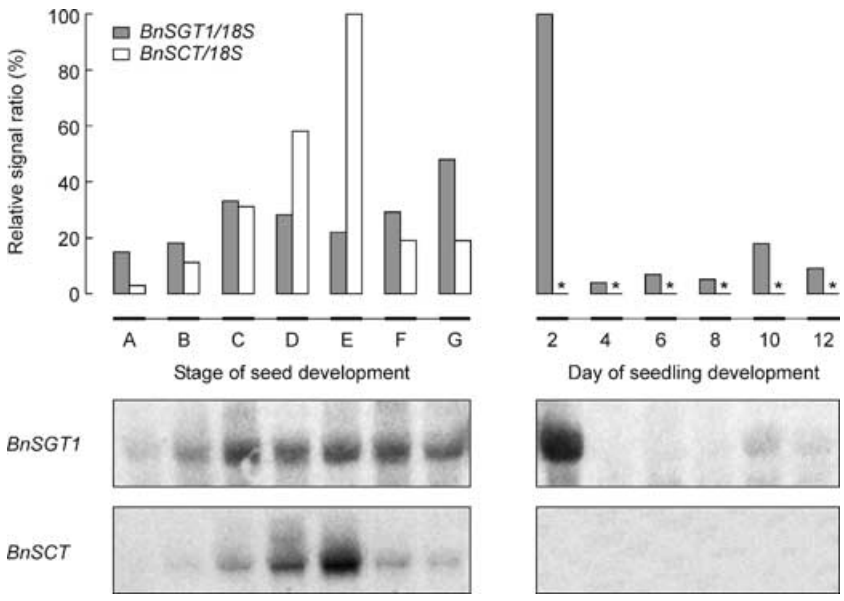


Figure 4. Transcript levels of *BnSGT1* and *BnSCT* genes in developing *B. napus* seeds and seedlings.
Blots of *B. napus* total RNA (10 μg) from various stages of seed and seedling development were hybridised with *BnSGT1* and *BnSCT* probes. All bar values are normalised to equal loading by calculation of specific probe/18S RNA probe signal ratio (100% of *BnSGT1*/*18S*, 2.7; 100% of *BnSCT*/*18S*, 0.5). Undetectable transcript levels are denoted by asterisks. Stages of seed development over *c.* 2 months: A, globular, torpedo and late torpedo stages; B, early cotyledonary stages to 'walking stick'; C, mid-cotyledonary stage, cotyledons expanding in size; D, well-developed cotyledonary stage, the green embryo reached maximal size, seed coat still soft; E, well-developed embryos, desiccation started, seed coat hard, embryos still green; F, seeds with brown seed coat, embryo yellow; G, desiccated seeds, seed coat black.

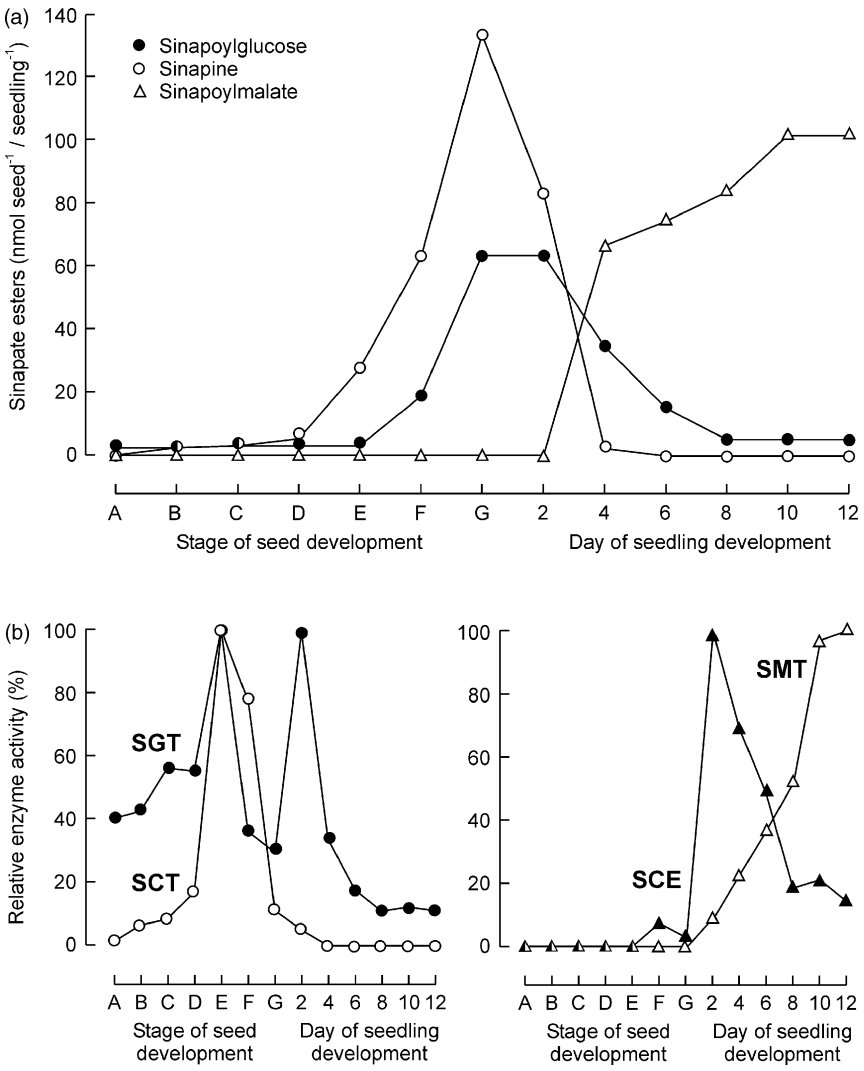


Figure 5. Metabolite accumulation and enzyme activities in developing *B. napus* seeds and seedlings.
Accumulation pattern of the *B. napus* sinapate esters (a) and activities of the involved enzymes (b) in seeds (embryos) and seedlings (cotyledons). For abbreviations of enzyme names and description of seed developmental stages, see Figures 1 and 4, respectively. Enzyme activities are given as per cent of maximal reaction velocities in pkat per milligram protein (100%): SGT, 127; SCT, 185; SCE, 511; and SMT, 55.

developmental stage C and did not significantly change during the following stages of seed development. It is remarkable that during late stages, a high abundance of *BnSGT1* transcripts was maintained. Dramatic changes in *BnSGT1* transcript level were observed at early seedling stages. Two days after sowing, another twofold increase in *BnSGT1* transcripts was detected. Only 2 days later, i.e. 4 days after sowing, the amount of *BnSGT1* transcript had decreased below the detection limit. Minor amounts of *BnSGT1* transcripts were found during late stages of seedling development. *BnSGT1* transcripts were detected neither in adult leaves nor in flower buds and flowers by RNA gel blot analysis (not shown).

The gene encoding BnSCT was found to be subjected to a pronounced transcriptional regulation during seed development. Its expression was restricted to seed development. Nearly undetectable in the earliest stage investigated, the amount of *BnSCT* transcripts increased rapidly to reach their highest levels in stage E. As seed development proceeded, a dramatic and rapid decrease of the *BnSCT* transcript level was observed. Stages F and G were shown to contain still about 20% of the amount of *BnSCT* transcript found in the preceding stage E. By RNA gel blot analysis, we could not detect any *BnSCT* transcripts in the seedlings, adult leaves, flower buds or flowers (not shown).

Sinapate ester accumulation and enzyme activity levels

The major *B. napus* sinapate esters were quantified during seed and seedling development, and activities of the enzymes involved in their biosyntheses were determined (Figure 5). At early stages of embryogenesis, i.e. globular, torpedo or late torpedo (stage A), only sinapoylglucose was detected in trace amounts (Figure 5a). As the cotyledons differentiated, increased synthesis of sinapine and sinapoylglucose was observed, together with rapid expansion of the cotyledons, a developmental feature of *B. napus*. Sinapine appeared at early cotyledonary stage (stage B),

and rapid accumulation was observed after the embryo had entered the well-developed phase (stage D). Sinapine accumulation continued during the following stages of seed development preceding accumulation of sinapoylglucose, which obviously underlies a high turnover rate as a precursor of sinapine biosynthesis. Sinapoylglucose levels exceeded the threshold value only in late developmental stage (stage F).

In members of the Brassicaceae family, zygotic embryogenesis is well described. Using *B. napus* as a model system for studying seed-specific events has the advantage that the endosperm has a transient role and is re-absorbed during the later phases of seed development (Goldberg *et al.*, 1989). Therefore, most metabolic activities occurring during seed development can be exclusively assigned to the embryo itself. The embryo-specific pattern of sinapate ester accumulation could be verified by HPLC analyses of extracts from dissected embryos (details not documented). On the basis of individual embryos, the results were analogous to those obtained with the intact seeds. However, when the amount of sinapate esters was calculated on the basis of the embryo DW, the results were different. Sinapine was first detected in minute amounts at very early stages of cotyledonary differentiation, and sinapoylglucose accumulation started at 'walking stick' stage. Sinapine was detected only in the embryo, whereas sinapoylglucose was present in the embryo, as well as in trace amounts in the seed coat of early- to mid-cotyledonary stages. The most rapid accumulation was observed in embryos between developmental stages C and D, thereafter remaining constant. This is in contrast to the kinetics shown in Figure 5(a), which is most likely because of parallel increase in embryo DW and the amount of the respective sinapate esters.

Early seedling development was characterised by a dramatic decrease in the amount of sinapine, followed by accumulation of sinapoylmalate. Thus, sinapate, liberated from sinapine by SCE activity, was rapidly conjugated with glucose by SGT activity and consumed in SMT-catalysed sinapoylmalate formation. Four days after sowing, sinapine was barely detectable and it completely disappeared thereafter. All the enzyme activities involved, SGT (BnSGT1), SCT (BnSCT), SCE and SMT, are closely correlated with the accumulation pattern of these sinapate esters (Figure 5b). Highest SGT and SCT activities are found in well-developed green embryos at the beginning of desiccation (stage E).

The final stage of embryogenesis involves metabolic inactivation of many pathways and acquisition of desiccation tolerance. As desiccation of the seed proceeded, the activity of BnSGT1 rapidly decreased, whereas that of BnSCT was at a relatively high level at the penultimate stage of seed development (stage F). This difference may reflect different sensitivities of the enzymes to desiccation. During embryogenesis, the sinapate ester metabolism, paralleled by mRNA abundance of BnSGT1 and BnSCT, resembles the biosynthesis and accumulation of most storage compounds, i.e. proteins and lipids (Crouch and Sussex, 1981; Goldberg *et al.*, 1989).

SCE, the specific sinapine esterase, also showed a pronounced developmental regulation. Its activity was detectable only in trace amounts during late seed development, but it rapidly increased after sowing; the highest activity was observed in 2-day-old seedlings. Thereafter, it returned to a low level.

SMT, the second enzyme besides the *B. napus* seed-specific SCT involved in transacylating sinapoylglucose, was not detected in seeds, but appeared in 2-day-old seedlings when SGT reached a second peak of activity. Thereafter, SGT activity declined, whereas the activity of SMT continuously increased until day 10 of seedling development. Taken together, these results show that transcriptional control of BnSGT1 and BnSCT activities regulates the pathway linking sinapine to sinapoylmalate via sinapoylglucose (Bouchereau *et al.*, 1992).

Relationship of BnSGT1 to other GTs

B. napus sinapate glucosyltransferase (BnSGT1) was shown to share high sequence identity with GTs from *A. thaliana*, which catalyse the formation of β-acetal esters of hydroxycinnamates. The closest relative was identified as *A. thaliana* SGT1 (Lim *et al.*, 2001; Milkowski *et al.*, 2000b). An unrooted similarity tree constructed from sequences of plant GTs with experimentally proven substrate specificities shows BnSGT1 clustering in a group composed of ester-forming GTs (Figure 6). This group is clearly separated from the branches of flavonoid GTs and those involved in anthocyanin biosynthesis. The branch of ester-forming GTs harbours, besides BnSGT1, not only the hydroxycinnamate GTs (HCAGT1-3) from closely related *A. thaliana*, but also the limonoid GT from *Citrus*. However, the *A. thaliana* enzymes that catalyse the formation of a sinapate glucoside (sinapate 4-*O*-GT, UGT72E3) and a caffeate glucoside (caffeate 3-*O*-GT, UGT71C1; Lim *et al.*, 2003) belong to a different branch. This resembles the extensive analysis given by Li *et al.* (2001) that groups all known plant GTs with the capacity to form ester bonds in one branch, designated as group L. In Figure 6, the position of ester-forming GTs with acceptor specificity for indole acetate (IAAGT) and salicylate (SAGT) indicates a distant relationship to the hydroxycinnamate GTs. This may reflect specific substrate–enzyme interactions. The ability to catalyse formation of hydroxycinnamoylglucoses might be conserved in specific sequence features. Rapidly growing sequence information produced by genome projects together with progress in elucidating enzyme functions and structural analysis by crystallography will lay the appropriate basis for a dissection of these sequence motifs in the near future.

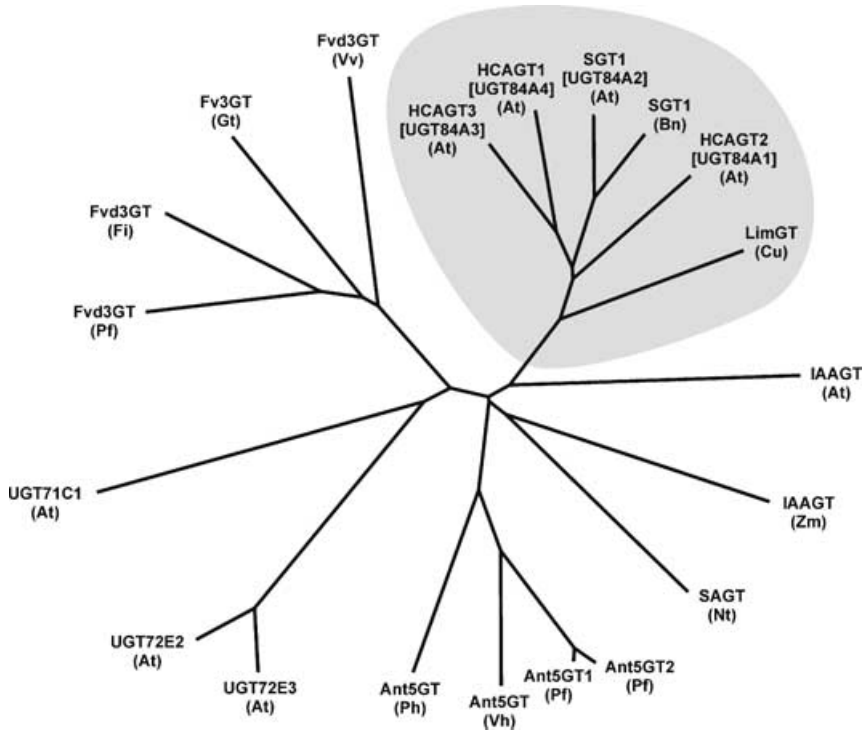


Figure 6. Unrooted similarity tree of plant GTs related to *B. napus* SGT1.
SGT1 (Bn), *Brassica napus* SGT1 (GenBank Accession # AAF98390); SGT1 [UGT84A2] (At), *Arabidopsis thaliana* SGT1 (BAB02351); HCAGT1 [UGT84A4], HCAGT2 [UGT84A1], HCAGT3 [UGT84A3] (At), *Arabidopsis thaliana* hydroxycinnamate-GT1, -2, -3 (F71419, D71419, E71419); LimGT (Cu), *Citrus unshiu* limonoid-GT (BAA93039); IAAGT (At), *Arabidopsis thaliana* indole 3-acetate-GT (AAB87119); IAAGT (Zm), *Zea mays* indole 3-acetate-GT (AAA59054); SAGT (Nt), *Nicotiana tabacum* salicylic acid-GT (AAF61647); UGT71C1 (At), *Arabidopsis thaliana* caffeate 3-*O*-GT and flavonoid GT (AAC35226); UGT72E2 (At), *Arabidopsis thaliana* sinapyl alcohol 4-*O*-GT (BAA97275); UGT72E3 (At), *Arabidopsis thaliana* sinapate 4-*O*-GT (AAC26233); Ant5GT1, Ant5GT2 (Pf), *Perilla frutescens* anthocyanin 5-*O*-GT (BAA36421, BAA36422); Ant5GT (Vh), *Verbena hybrida* anthocyanin 5-*O*-GT (BAA36423); Ant5GT (Ph), *Petunia hybrida* anthocyanin 5-*O*-GT (BAA89009); Fvd3GT (Pf), *Perilla frutescens* flavonoid 3-*O*-GT (BAA19659); Fv3GT (Gt), *Gentiana triflora* flavonol 3-*O*-GT (Q96493); Fvd3GT (Vv), *Vitis vinifera* flavonoid 3-*O*-GT (BAB41026); and Fvd3GT (Fi), *Forsythia intermedia* flavonoid 3-*O*-GT (AAD21086).

Relationship of BnSCT to other SCPL proteins

Multiple alignment of the deduced amino acid sequences of BnSCT and SCPL proteins with known functions (Figure 7) shows highest identities of BnSCT with the *A. thaliana* SCT (84%) and other glucose ester-dependent acyltransferases (GAC) (50% with SMT from *A. thaliana* and 44% with GAC from tomato). The homology of these acyltransferases and hydrolases of the SCP-type is illustrated by significant overall sequence identities of BnSCT with proven peptidases (e.g. 37% with carboxypeptidase 1 (CP1) from barley; 36% with SCP1 from rice). These acyltransferases and SCPs may have developed as a result of divergent evolution. Within plant transacylases, the SCPL acyltransferases form a distinct group. The alignment defines highly conserved blocks of amino acid residues. These blocks are preferentially located around the amino acid residues (S, D, H) that form the catalytic triad of SCPs. It suggests an important functional role for these residues in acyltransferases as well, i.e. the composition of the active site has been conserved between hydrolases and acyltransferases. Inhibitor studies with phenylmethylsulfonyl fluoride (PMSF) revealed the seryl residue as catalytically active in *A. thaliana* SMT (Lehfeldt *et al.*, 2000) and *A. thaliana* SCT (Shirley *et al.*, 2001). Other completely conserved sequence motifs are found within the N-terminal half of the proteins.

The deduced BnSCT protein carries an internal sequence (positions 260–298; Figure 7) that has been proposed to be cleaved out during post-translational processing in the *A. thaliana* homologue (Shirley and Chapple, 2003). This kind of modification, resulting in heterodimer formation, has been shown for CP1 from barley (Doan and Fincher, 1988) and carboxypeptidase 2 (CPDW2) from wheat (Breddam *et al.*, 1987). In the alignment, this region forms the least conserved part, indicating that this part of proteins was not subjected to functional constraints during evolution. It is striking that carboxy pepidase Y (CPY), a proven monomer, and *A. thaliana* SMT lack this amino acid stretch. This raises the question whether SCPL acyltransferases could be systematically subgrouped into enzymes consisting of a single peptide chain and others composed of two peptide chains forming heterodimers. Future work

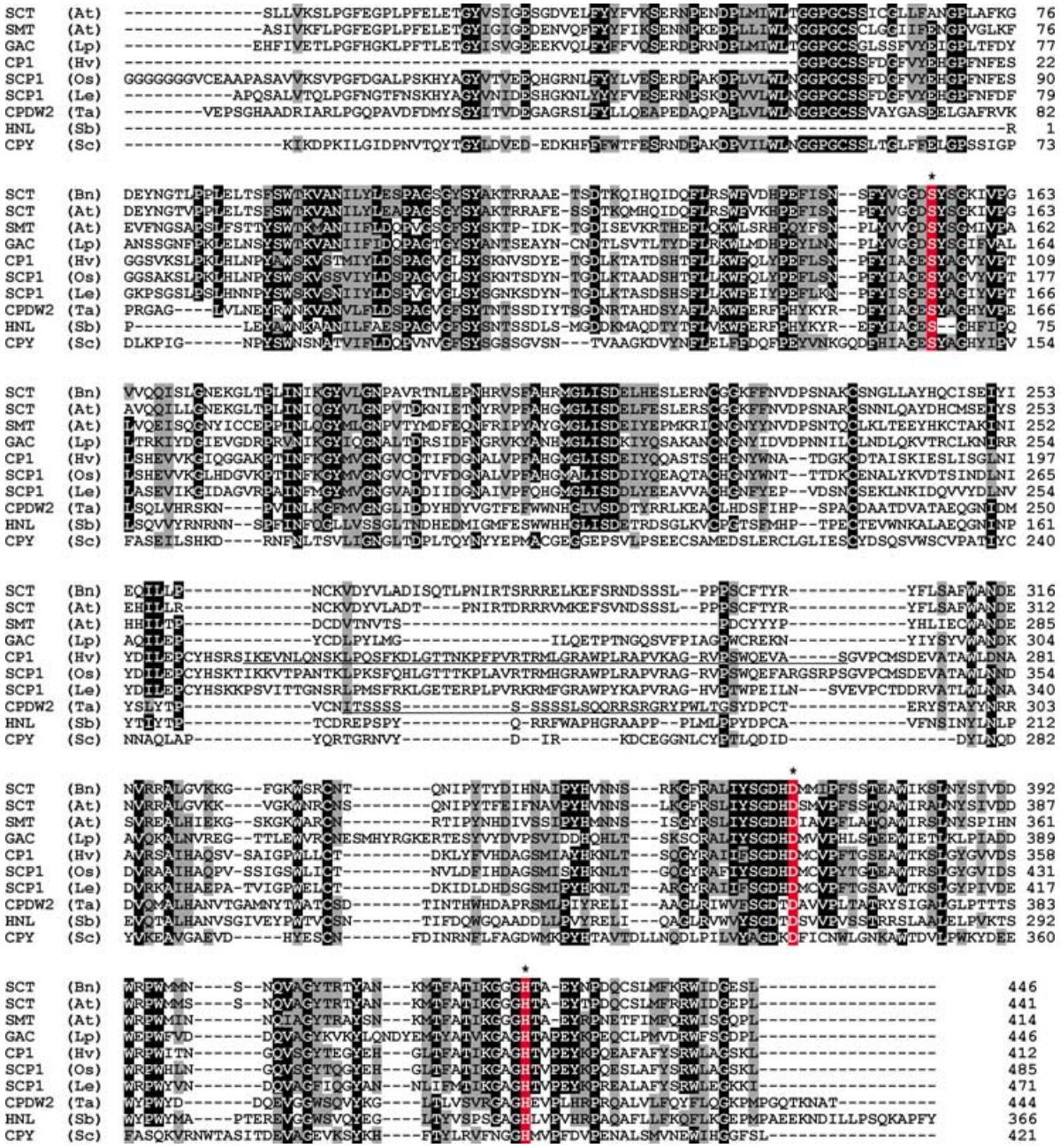


Figure 7. Aligment of the deduced amino acid sequence of the *B. napus* SCT with those of functionally characterised SCPL acyltransferases (SCT and SMT from *A. thaliana* and GAC from *L. penellii*), protein hydrolases (CP, SCP) and a hydroxynitrile lyase (HNL).
N-terminal leader peptides are removed. Fully conserved residues are shaded in black. Grey shading indicates conservation of at least 70%. Amino acid residues forming the catalytic triad of SCPs (S, D, H) are marked by asterisks and red shading. Endopeptides cleaved out during maturation of CP1 (Hv) and CPDW2 (Ta) are underlined. SCT (Bn), *Brassica napus* SCT (GenBank Accession # AY383718) was aligned with SCT (At), *Arabidopsis thaliana* SCT (AY033947); SMT (At), *Arabidopsis thaliana* SMT (AF275313); GAC (Lp), *Lycopersicon penellii* GAC (AF248647); CP1 (Hv), *Hordeum vulgare* Hv (AAA32940); SCP1 (Os), *Oryza sativa* SCP1 (P37890); SCP1 (Le), *Lycopersicon esculentum* SCP1 (AAF44708); CPDW2 (Ta), *Triticum aestivum* CP2 (P08819); HNL (Sb), *Sorghum bicolor* HNL (S53311); CPY (Sc), *Saccharomyces cerevisiae* CPY (1YSC). Sequence identities with BnSCT are: SCT (At), 84%; SMT (At), 50%; GAC (Lp), 44%; CP1 (Hv), 37%; SCP1 (Os), 36%; ACP1 (Le), 35%; CPDW2 (Ta), 27%; HNL (Sb), 27%; and CPY (Sc), 23%.

on this interesting enzyme group will address the fundamental problem of evolutionary history of SCPL acyltransferases, as well as the question of why these enzymes function as acyltransferases instead of hydrolases.

Application

In crop plants like *B. napus*, sinapate esters are antinutritive compounds. They contribute to the bitter taste and

astringency of seed products. Moreover, sinapate esters form complexes with proteins during seed oil processing, thus compromising the use of the valuable seed meal for animal feed (Ismail *et al.*, 1981) and preventing it from being used as human food supplement. Thus, there is a fundamental interest in reducing the amount of sinapate esters in the seeds. Suppressing the expression of the key enzymes of sinapine synthesis, BnSGT1 and BnSCT, by techniques such as dsRNAi should be a valuable step in establishing *B. napus*, currently an important oil crop, as a protein crop as well. Approaches to suppress *A. thaliana SGT1* expression are ongoing in our laboratory. Several transgenic lines with markedly reduced amounts of sinapoylglucose and sinapine have already been identified. Experiments with constructs for suppressing both *SGT* and *SCT* expression in *B. napus* and the prime model plant *A. thaliana* are in progress.

Experimental procedures

Plant material

Seeds of rape (*B. napus* L. cv. Express), red cabbage (*B. oleracea* L. cv. Markola) and turnip (*B. rapa* L. cv. Rex) were obtained from the Norddeutsche Pflanzenzucht (Holtsee, Germany). Seedlings were grown under greenhouse conditions. Leaves, buds, flowers and seeds of *B. napus* were harvested at different developmental stages between June and August 2002 from field-grown plants (near Halle), and were stored at −80°C. For a comparative study of sinapate ester accumulation during embryogensis, zygotic embryos were dissected and frozen in liquid nitrogen, stored at −80°C and were later lyophilised.

Analysis of sinapate esters

The three major sinapate esters (sinapoylglucose, sinapoylcholine and sinapoylmalate) were identified by chromatographic (HPLC) and UV-Vis spectroscopic (DAD-HPLC) comparison, with standard compounds from our own collection of natural products, as well as with products from the corresponding enzyme assays.

Extraction and quantification were achieved by transferring seeds or cotyledons from seedlings (5–10 samples) into 2-ml tubes, which contained 400 μl of a methanol:water mixture (4 : 1) and *c.* 100 zirconia beads (1 mm in diameter). The tubes were vigorously shaken with a Bead Beater (Biospec Products, Bartlesville, OK, USA), the resulting homogenates were centrifuged, aliquots of the supernatants were transferred into HPLC autosampler vials and 10-μl samples were injected. Dissected embryos were analysed likewise. Embryos from very early stages were pooled together (20–50). Analyses of all other stages were performed on single embryos ($n = 5$–18).

The liquid chromatograph was equipped with a 5-μm Nucleosil C18 column (250 mm × 4 mm i.d.; Macherey-Nagel, Düren, Germany). A 20-min linear gradient was applied at a flow rate of 1 ml min^{-1} from 10 to 50% solvent B (acetonitrile) in solvent A (1.5% *o*-phosphoric acid in water). The compounds were photometrically detected at 330 nm and quantified by external standardisation with authentic compounds.

Enzyme preparation and assays

Protein extracts were prepared in triplicate from 90 seeds or cotyledon pairs, or from 100 seedlings, leaves, buds or flowers by treatment with an Ultra-Turrax homogeniser in various volumes of 100 mM potassium phosphate buffer (pH 7.0), depending on sample size. The homogenates were centrifuged (10 min at 20 000 ***g***), and protein of the supernatants was precipitated by adding ammonium sulphate to 80% saturation. After centrifugation, the pellets were re-dissolved in 10 mM of the extraction buffer, and the solutions were desalted using PD-10 Sephadex G-25 columns (Supelco, Bellefonte, PA, USA). Protein contents of the supernatants from the first centrifugation step were determined by the Bradford method (Bradford, 1976).

The enzyme assays contained the following components: *SGT* – 2 mM sinapate (or 4-coumarate, caffeate, ferulate) and 4 mM UDP-Glc in a total volume of 150 μl of 100 mM 2-[*N*-morpholino]ethane-sulfonic acid (MES) buffer (pH 6.0) containing 1 mM 2-mercaptoethanol, 0.5 mM EDTA and 10% glycerol; *SCT* – 2 mM sinapoylglucose and 10 mM choline chloride in a total volume of 100 μl of 100 mM potassium phosphate buffer (pH 7.0) containing 1 mM 2-mercaptoethanol and 0.5 mM EDTA; *SMT* – 2 mM sinapoylglucose and 120 mM L-malate in a total volume of 200 μl of 100 mM MES buffer (pH 6.0) containing 1 mM 2-mercaptoethanol and 0.5 mM EDTA; *SCE* – 0.5 mM sinapine in a total volume of 200 μl of 100 mM tricine buffer (pH 8.5) containing 1 mM 2-mercaptoethanol and 0.5 mM EDTA.

Enzyme reaction mixtures containing 25–50 μl protein extracts were incubated in 1.5-ml Eppendorf vials for 15 min at 30°C, and the reactions were terminated by adding 10 μl trifluroroacetic acid (TFA) (SGT, SMT) or 100 μl methanol (SCT, SCE). To determine product formation, the mixtures were centrifuged, aliquots of the supernatants were transferred into HPLC autosampler vials and 10-μl samples were injected. Kinetic data of BnSGT were determined according to Lineweaver and Burk (1934) and Michaelis and Menten (1913).

The liquid chromatograph used for enzyme activity determinations was equipped with a CC 70/4.6 Nucleosil C18 column (120 mm × 3 mm i.d.; Macherey-Nagel, Düren, Germany). A 5-min linear gradient was applied at a flow rate of 1 ml min^{-1} from 10 to 25% solvent B (acetonitrile) in solvent A (1.5% *o*-phosphoric acid in water). The products were photometrically detected at 330 nm (sinapoylglucose, caffeoylglucose and feruloylglucose, as well as sinapate, sinapine and sinapoylmalate) or 315 nm (4-coumaroylglucose) and quantified by external standardisation with standard compounds from our own collection.

cDNA isolation

For amplification of *BnSCT* cDNA, PCR was performed with degenerate forward primer (5′-GGA/T GAC/T CAC/T GAT ATG-3′) derived from a conserved amino acid sequence motif of acyltransferases (GDHDM) and the T7 reverse primer using a cDNA library from immature *B. napus* seeds prepared in the λ Uni-ZAP XR vector system (Stratagene, La Jolla, CA, USA) as template. The resulting 240-bp cDNA fragment was subcloned into plasmid pGEM-Teasy (Promega, Madison, WI, USA). Sequence analysis revealed similarity to SCPL glucose acyltransferases. The fragment was then used to screen the cDNA library from immature seeds. A 5′ incomplete 874-bp fragment was found. In order to isolate a full-length cDNA, 5′ RACE-PCR was performed with gene-specific reverse primer (5′-CCCTT-CTTCACGCCTAAAGCTCTGC-3′) and RNA from immature seeds using the Gene Racer Kit (Invitrogen, Carlsbad, CA, USA). A

RACE-PCR product of 1024 bp was isolated, subcloned in pCR4Blunt-TOPO vector (Invitrogen) and sequenced. Sequence analysis revealed a reading frame starting with ATG, closely related to the N-terminal part of *A. thaliana* SCT (Shirley *et al.*, 2001). RT-PCR was used to amplify a cDNA fragment carrying the whole *BnSCT* reading frame. RNA from immature seeds was reversely transcribed into cDNA using M-MLV RT Rnase H-, Point Mutant (Promega) and oligo(dT) primer. PCR amplification was performed by PCR with specific primers derived from the isolated cDNA fragments covering start and stop codons of *BnSCT* (forward primer, 5′-ATGAGAAATCTTTACTTTCTAGTC-3′; reverse primer, 5′-TCAGAGAGATTCACCATCAATCC-3′) with Platinum *Pfx* polymerase (Invitrogen). The 1401-bp PCR product was cloned into pCR4Blunt-TOPO vector (Invitrogen) and subjected to sequence analysis.

Analysis of nucleic acids

Genomic plant DNA was isolated from young seedlings using a Maxi Kit for plant DNA purification (Qiagen, Hilden, Germany). All plasmid purifications were carried out with a Mini Kit for plasmid isolation (Qiagen). For DNA gel blot analysis, 5 μg of genomic DNA extracted from *B. napus*, *B. oleracea* and *B. rapa* was digested with restriction endonucleases, electrophoretically separated and transferred to Hybond N+ membrane (Amersham, Freiburg, Germany) according to standard protocols (Sambrook *et al.*, 1989). Hybridisation was performed in 7% SDS (w/v), 250 mM NaCl, 1 mM EDTA, 250 mM NaH_2PO_4 (pH 7.0) at 60°C overnight with probes from full-length cDNAs of *BnSGT1* and *BnSCT* ^{32}P-labelled using the Megaprime DNA Labelling System (Amersham). Final washes were carried out in 0.1× SSC, 0.1% SDS at 68°C.

For RNA gel blot analysis, 10 μg of total RNA, extracted from tissues using a kit for plant RNA purification (Qiagen), was electrophoretically separated and transferred to Hybond N+ nylon membrane (Sambrook *et al.*, 1989). Hybridisation was performed as indicated for DNA blot analysis. Loading controls were peformed by hybridising with a ^{32}P-labelled 18S rRNA probe. Blots were exposed to films or to Storm PhospholmagerTM screens for quantification of signals using IMAGEQUANT software. Ratios of signal intensities for specific probes versus 18S rRNA probes were calculated for a given sample to normalise the specific signals for equal loading of RNA.

Sequence analysis of DNA and proteins was done using the software package CLONE MANAGER (Durham, NC, USA). The alignment was carried out by CLUSTALW algorithm using the BLOSUM62 matrix. Potential N-glycosylation sites were identified by prediction software (http://www.cbs.dtu.dk/services/NetNGlyc/). The phylogenetic tree was constructed using the program TREEVIEW (Page, 1996). Leader peptides for vacuolar targeting and cleavage sites were predicted using the algorithm described by Nielsen *et al.* (1997) and available at http://www.cbs.dtu.dk/services/SignalP-2.0/.

Constructs for expression of BnSGT1 *and* BnSCT *in* E. coli

Coding sequences were amplified by PCR with Platinum *Pfx* polymerase (Invitrogen) using the cloned cDNAs as template. The forward primers (*BnSGT1*: 5′-GCTCGGTACCCCATGGAACTATCATCTTCTCC-3′; *BnSCT*: 5′-TATTATCCATGGCTTCTTTGCATGTGAAGTATCTTCC-3′) introduced a *Nco*I restriction site. The reverse primers (*BnSGT1*: 5′-TATTATCCCGGGTGACTTTTGCAATAAAAG-3′; *BnSCT*: 5′-TATTATCCCGGGGAGAGATTCACCATCAATCCATC-3′) incorporated a *Sma*I restriction site before the stop codon. For *BnSCT*, the primer pair amplified a truncated fragment lacking the sequence for the predicted signal peptide. After subcloning into plasmid pGEM-T easy (Promega) and sequencing, cDNAs were ligated as *Nco*I/*Sma*I fragments into pIVEX2.3 (Roche, Mannheim, Germany), fusing both to six His codons at the 3′ terminus. Resulting vectors pIVEX2.3.-*BnSGT1* and pIVEX2.3.-*BnSCT* were cleaved by *Nco*I and *Bam*HI, and cDNAs including the His codons were inserted into similarly cleaved expression plasmid pET28a(+) (Novagen, Madison, WI, USA) to yield pET28a-*BnSGT1* and pET28a-*BnSCT*. For expression analysis, the *E. coli* host BL21-CodonPlus(DE3)-RP (Stratagene) was transformed with the empty pET28a(+) and the recombinant plasmids.

Expression of BnSGT1 *and* BnSCT *in* E. coli *and preparation of* E. coli *extracts*

Overnight cultures of bacteria grown at 37°C in Luria–Bertani (LB) liquid medium containing the appropriate antibiotics were diluted into fresh LB medium, and cultivated at 37°C with vigorous shaking until the early logarithmic growth phase (OD_{600} = 0.4–0.6). Expression of the *B. napus* cDNAs was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM to the cultures grown for another 2 h at 37°C. Cells were harvested by centrifugation at 5000 ***g*** for 10 min at 4°C. Cell pellets were re-suspended in 100 mM NaH_2PO_4 (pH 7.0), 300 mM NaCl, 10% glycerol, and were treated for 30 min with lysozyme (10 μg ml^{-1}) on ice followed by ultrasonication at 0°C. After centrifugation (20 min at 20 800 ***g***, 4°C), the supernatant contained the soluble protein fraction. After removing the soluble protein fraction, the remaining pellet was re-suspended in 100 mM NaH_2PO_4, 10 mM Tris–HCl, 8 M urea, 2.5 mM imidazole (pH 8.0), and was subjected to ultrasonication. The resulting suspension contained the insoluble protein fraction.

Re-folding of recombinant BnSCT

Isolation and purification of inclusion bodies were carried out according to Rudolph *et al.* (1997). Protein folding was initiated using the rapid dilution technique (Rudolph and Lilie, 1996). The denatured inclusion body protein was slowly added to a final concentration of 20 μg ml^{-1} to 20 ml, 100 mM Tris–HCl (pH 8.0), 10% glycerol, 3 mM EDTA, 5 mM reduced glutathione and 1 mM oxidized glutathione. The mixture was then incubated for 4 h at 10°C.

SCT assays contained in 100 μl renaturation solution, 10 mM choline and 2 mM sinapoylglucose solutions. The assays were incubated at 30°C for 2 h, terminated by adding 10 μl of 100% TFA, and product formation was analysed by HPLC.

Immunoblot analysis

Proteins were separated by SDS–PAGE as described by Sambrook *et al.* (1989) and transferred to a nitrocellulose membrane (0.2 μm pore size; Invitrogen) using Xcell II™ Blot Module (Invitrogen). Analysis was carried out with anti-His-Tag monoclonal antibody (Novagen) as primary and anti-Mouse IgG alkaline phosphatase conjugate (Sigma, Taufkirchen, Germany) as secondary antibody using colorimetric signal detection according to the protocols given by the suppliers.

Acknowledgements

We thank Ingrid Otschik and Kerstin Manke for technical assistance, Christine Kaufmann for preparing the figures, Felix Stehle for help in refolding the recombinant BnSCT protein and Jürgen Schmidt for MS analysis of sinapine. The seeds of *B. napus*, *B. oleracea* and *B. rapa* were kindly provided by the Norddeutsche Pflanzenzucht (Holtsee, Germany). We also thank the Swedish Research Council for Engineering Sciences for the postdoctoral fellowship of Lilian Nehlin. This work is part of the research project 'NAPUS 2000 – Gesunde Lebensmittel aus transgener Rapssaat', financially supported by the Bundesminister für Bildung und Forschung.

Supplementary Material

The following material is available from http://www.blackwellpublishing.com/products/journals/suppmat/TPJ/TPJ2036/TPJ2036sm.htm

Figure S1. Analysis of *BnSGT1* expression in *E. coli*.
SDS–PAGE (a) and Western blot analysis (b) of protein (20 μg) from *E. coli*, strain BL21-CodonPlus(DE3)-RP, harbouring pET28a (lanes 1 and 3) and *E. coli* harbouring pET28a-SGT1 (lanes 2 and 4; lanes 1 and 2, insoluble fractions; lanes 3 and 4, soluble fractions).

Figure S2. Analysis of *BnSCT* expression in *E. coli*.
SDS–PAGE (a) and Western blot analysis (b) of protein (20 μg) from *E. coli*, strain BL21-CodonPlus(DE3)-RP, harbouring pET28a (lanes 1 and 3) and *E. coli* harbouring pET28a-SCT (lanes 2 and 4; lanes 1 and 2, insoluble fractions; lanes 3 and 4, soluble fractions).

References

Bech, L.M. and Breddam, K. (1989) Inactivation of carboxypeptidase Y by mutational removal of the putative essential histidyl residue. *J. Biol. Chem.* **262**, 13726–13735.

Booij-James, I.S., Dube, S.K., Jansen, M.A.K., Edelman, M. and Mattoo, A.K. (2000) Ultraviolet-B radiation impacts light-mediated turnover of the photosystem II reaction center heterodimer in *Arabidopsis* mutants altered in phenolic metabolism. *Plant Physiol.* **124**, 1275–1283.

Bopp, M. and Lüdicke, W. (1980) Synthesis of sinapine during seed development of *Sinapis alba*. *Z. Naturforsch.* **35c**, 549–543.

Bouchereau, A., Hamelin, J., Renard, M. and Larher, F. (1992) Structural changes in sinapic acid conjugates during development of rape. *Plant Physiol. Biochem.* **30**, 467–475.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.

Breddam, K., Sorensen, S.B. and Svendsen, I. (1987) Primary structure and enzymatic properties of carboxypeptidase II from wheat bran. *Carlsberg Res. Commun.* **52**, 297–311.

Crouch, M.L. and Sussex, I.M. (1981) Development and storage-protein synthesis in *Brassica napus* L. rape embryos *in vivo* and *in vitro*. *Planta*, **153**, 64–74.

Doan, N.P. and Fincher, G.B. (1988) The A- and B-chains of carboxypeptidase I from germinated barley originate from a single precursor polypeptide. *J. Biol. Chem.* **263**, 11106–11110.

Goldberg, R.B., Barker, S.J. and Perez-Grau, L. (1989) Regulation of gene expression during plant embryogenesis. *Cell*, **56**, 149–160.

Gupta, R., Jung, E. and Brunak, S. (2002) Prediction of N-glycosylation sites in human proteins (in press).

Hause, B., Meyer, K., Viitanen, P.V., Chapple, C. and Strack, D. (2002) Immunolocalization of 1-*O*-sinapoylglucose: malate sinapoyltransferase in *Arabidopsis thaliana*. *Planta*, **215**, 26–32.

Hayashi, R., Moore, S. and Stein, W.H. (1973) Serine as the active center of yeast carboxypeptidase. *J. Biol. Chem.* **248**, 8366–8369.

Hayashi, R., Bai, Y. and Hata, T. (1975) Evidence for an essential histidine in carboxypeptidase Y. Reaction with the chloromethyl ketone derivative of benzyloxycarbonyl-1-phenylalanine. *J. Biol. Chem.* **250**, 5221–5226.

Hughes, J. and Hughes, M.A. (1994) Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. *DNA Seq.* **5**, 41–49.

Ismail, F., Vaisey-Genser, M. and Fyfe, B. (1981) Bitterness and astringency of sinapine and its components. *J. Food Sci.* **46**, 1241–1244.

Landry, L.G., Chapple, C.C.S. and Last, R. (1995) *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. *Plant Physiol.* **109**, 1159–1166.

Lehfeldt, C., Shirley, A.M., Meyer, K., Ruegger, M.O., Cusumano, J.C., Viitanen, P.V., Strack, D. and Chapple, C. (2000) Cloning of the *SNG1* gene of *Arabidopsis* reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. *Plant Cell*, **12**, 1295–1306.

Li, A.X. and Steffens, J.C. (2000) An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. *Proc. Natl. Acad. Sci. USA*, **97**, 6902–6907.

Li, J., Ou-Lee, T.-M., Raba, R., Amundson, R.G. and Last, R.L. (1993) *Arabidopsis* flavonoid mutants are hypersensitive to UV-B irradiation. *Plant Cell*, **5**, 171–179.

Li, Y., Baldauf, S., Lim, E.-K. and Bowles, D.J. (2001) Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. *J. Biol. Chem.* **276**, 4338–4343.

Liao, D.-I. and Remington, S.J. (1990) Structure of wheat serine carboxypeptidase II at a 3.5-Å resolution. *J. Biol. Chem.* **256**, 6528–6531.

Liao, D.-I., Breddam, K., Sweet, R.M., Bullock, T. and Remington, S.J. (1992) Refined atomic model of wheat serine carboxypeptidase II at 2.2-Å resolution. *Biochemistry*, **31**, 9796–9812.

Lim, E.-K., Li, Y., Parr, A., Jackson, R., Ashford, D.A. and Bowles, D.J. (2001) Identification of glucosyltransferase genes involved in sinapate and lignin synthesis in *Arabidopsis*. *J. Biol. Chem.* **276**, 4344–4349.

Lim, E.-K., Higgins, G.S., Li, Y. and Bowles, D.J. (2003) Regioselectivity of glycosylation of caffeic acid by UDP-glucose: glucosyltransferase is maintained *in planta*. *Biochem. J.* **373**, 987–992.

Lineweaver, H. and Burk, D. (1934) The determination of enzyme dissociation constants. *J. Am. Chem. Soc.* **56**, 658–666.

Lorenzen, M., Racicot, V., Strack, D. and Chapple, C. (1996) Sinapic acid ester metabolism in wild type and sinapoylglucose-accumulating mutant of *Arabidopsis*. *Plant Physiol.* **112**, 1625–1630.

Michaelis, L. and Menten, M.L. (1913) Die Kinetik der Invertinwirkung. *Biochem. Z.* **49**, 333–369.

Milkowski, C., Baumert, A. and Strack, D. (2000a) Cloning and heterologous expression of a rape cDNA encoding UDP-glucose: sinapate glucosyltransferase. *Planta*, **211**, 883–886.

Milkowski, C., Baumert, A. and Strack, D. (2000b) Identification of four *Arabidopsis* genes encoding hydroxycinnamate glucosyltransferases. *FEBS Lett.* **486**, 183–184.

Mock, H.-P. and Strack, D. (1993) Energetics of the uridine 5′-diphosphoglucose: hydroxycinnamic acid acyl-glucosyltransferase reaction. *Phytochemistry*, **32**, 575–579.

Mock, H.-P., Vogt, T. and Strack, D. (1992) Sinapoylglucose: malate sinapoyltransferase activity in *Arabidopsis thaliana* and *Brassica napus*. *Z. Naturforsch.* **47c**, 680–682.

Nielsen, H., Engelbrecht, J., Brunak, S. and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**, 1–6.

Page, R.D.M. (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. *Computer Appl. Biosciences*, **12**, 357–358.

Regenbrecht, J. and Strack, D. (1985) Distribution of 1-sinapoylglucose: choline sinapoyltransferase in the Brassicaceae. *Phytochemistry*, **24**, 407–410.

Rudolph, R. and Lilie, H. (1996) *In vitro* folding of inclusion body proteins. *FASEB J.* **10**, 49–56.

Rudolph, R., Böhm, G., Lilie, H. and Jaenicke, R. (1997) Folding proteins. In *Protein Function: a Practical Approach* (Creighton, T.E., ed.). Oxford: IRL Press, pp. 57–99.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, 2nd edn. Cold Spring Harbor, New York: Cold Spring. Harbor Laboratory Press.

Schenck, H.R. and Röbbelen, G. (1982) Somatic hybrids by fusion of protoplasts from *Brassica oleracea* and *B. campestris*. *Z. Pflanzenzücht.* **89**, 278–288.

Sheahan, J.J. (1996) Sinapate esters provide greater UV-B attenuation than flavonoids in *Arabidopsis thaliana* (Brassicaceae). *Am. J. Bot.* **83**, 679–686.

Shirley, A.M. and Chapple, C. (2003) Biochemical characterization of sinapoylglucose: choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism. *J. Biol. Chem.* **278**, 19870–19877.

Shirley, A.M., McMichael, C.M. and Chapple, C. (2001) The *sng2* mutant of *Arabidopsis* is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose: choline sinapoyltransferase. *Plant J.* **28**, 83–94.

Strack, D. (1980) Enzymatic synthesis of 1-sinapoylglucose from free sinapic acid and UDP-glucose by a cell-free system from *Raphanus sativus* seedlings. *Z. Naturforsch.* **35c**, 204–208.

Strack, D. (1981) Sinapine as a supply of choline for the biosynthesis of phosphatidylcholine in *Raphanus sativus* seedlings. *Z. Naturforsch.* **36c**, 215–221.

Strack, D., Nurmann, G. and Sachs, G. (1980) Sinapine esterase. Part II. Specificity and change of sinapine esterase activity during germination of *Raphanus sativus*. *Z. Naturforsch.* **35c**, 963–966.

Strack, D., Knogge, W. and Dahlbender, B. (1983) Enzymatic synthesis of sinapine from 1-*O*-sinapoyl-β-D-glucose and choline by a cell-free system from developing seeds of red radish (*Raphanus sativus* L. var. *sativus*). *Z. Naturforsch.* **38c**, 21–27.

Strack, D., Ellis, B.E., Gräwe, W. and Heilemann, J. (1990) Sinapoylglucose: malate sinapoyltransferase activity in seed and seedlings of rape. *Planta*, **180**, 217–219.

Tkotz, N. and Strack, D. (1980) Enzymatic synthesis of sinapoyl-1-malate from 1-sinapoylglucose and 1-malate by a protein preparation from *Raphanus sativus* cotyledons. *Z. Naturforsch.* **35c**, 835–837.

Vogt, T., Aebersold, R. and Ellis, B. (1993) Purification and characterization of sinapine synthase from seeds of *Brassica napus*. *Arch. Biochem. Biophys.* **300**, 622–628.

Wang, S.X. and Ellis, B.E. (1998) Enzymology of UDP-glucose: sinapic acid glucosyltransferase from *Brassica napus*. *Phytochemistry*, **49**, 307–318.

GenBank Accession number: AY383718 (*BnSCT* cDNA).

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Phytochemistry 66 (2005) 1334–1345
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Formation of a complex pattern of sinapate esters in *Brassica napus* seeds, catalyzed by enzymes of a serine carboxypeptidase-like acyltransferase family?

Alfred Baumert [a], Carsten Milkowski [a], Jürgen Schmidt [a], Manfred Nimtz [b], Victor Wray [b], Dieter Strack [a,*]

[a] *Leibniz-Institut für Pflanzenbiochemie (IPB), Weinberg 3, D-06120 Halle (Saale), Germany*
[b] *Gesellschaft für Biotechnologische Forschung (GBF), Mascheroder Weg 1, D-38124 Braunschweig, Germany*

Received 16 December 2004; received in revised form 15 February 2005
Available online 23 May 2005

Abstract

Members of the Brassicaceae accumulate complex patterns of sinapate esters, as shown in this communication with seeds of oilseed rape (*Brassica napus*). Fifteen seed constituents were isolated and identified by a combination of high-field NMR spectroscopy and high resolution electrospray ionisation mass spectrometry. These include glucose, gentiobiose and kaempferol glycoside esters as well as sinapine (sinapoylcholine), sinapoylmalate and an unusual cyclic spermidine amide. One of the glucose esters (1,6-di-*O*-sinapoylglucose), two gentiobiose esters (1-*O*-caffeoylgentiobiose and 1,2,6′-tri-*O*-sinapoylgentiobiose) and two kaempferol conjugates [4′-(6-*O*-sinapoylglucoside)-3,7-di-*O*-glucoside and 3-*O*-sophoroside-7-*O*-(2-*O*-sinapoylglucoside)] seem to be new plant products. Serine carboxypeptidase-like (SCPL) acyltransferases catalyze the formation of sinapine and sinapoylmalate accepting 1-*O*-β-acetal esters (1-*O*-β-glucose esters) as acyl donors. To address the question whether the formation of other components of the complex pattern of the sinapate esters in *B. napus* seeds is catalyzed via 1-*O*-sinapoyl-β-glucose, we performed a seed-specific dsRNAi-based suppression of the sinapate glucosyltransferase gene (*BnSGT1*) expression. In seeds of *BnSGT1*-suppressing plants the amount of sinapoylglucose decreased below the HPLC detection limit resulting in turn in the disappearance or marked decrease of all the other sinapate esters, indicating that formation of the complex pattern of these esters in *B. napus* seeds is dependent on sinapoylglucose. This gives rise to the assumption that enzymes of an SCPL acyltransferase family catalyze the appropriate transfer reactions to synthesize the accumulating esters.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: *Brassica napus*; Brassicaceae; Oilseed rape; Sinapate esters; Molecular evolution; Serine carboxypeptidase-like acyltransferases

1. Introduction

Seeds of oilseed rape (*Brassica napus* L. var. *napus*) accumulate high amounts of sinapine (sinapoylcholine) among the various phenolics (Kozlowska et al., 1990; Bouchereau et al., 1991; Shahidi and Naczk, 1992; Bell, 1993) that are characteristic of the Brassicaceae family (Hegnauer, 1964). Sinapine is a member of the well-known pathway leading to sinapoylmalate in the seedlings (Linscheid et al., 1980; Bouchereau et al., 1992). A recent study of the molecular regulation of sinapine metabolism in *B. napus* (Milkowski et al., 2004) demonstrated transcriptional regulation of its biosynthesis via 1-*O*-sinapoyl-β-glucose. The role of such acylglucoses as acyl donors in plant secondary metabolism, as an alternative to CoA-dependent pathways, has been demonstrated with a remarkable number of other plants (Strack and Mock, 1993; Steffens, 2000).

* Corresponding author. Tel.: +49 345 5582 1500; fax: +49 345 5582 1509.
E-mail address: dieter.strack@ipb-halle.de (D. Strack).

0031-9422/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2005.02.031

It has been shown (Li and Steffens, 2000; Lehfeldt et al., 2000; Shirley et al., 2001; Milkowski et al., 2004) that 1-*O*-β-acetal ester-dependent acyltransferases constitute a new class of serine carboxypeptidase-like (SCPL) proteins (SCPL acyltransferases). In Arabidopsis these enzymes form a distinct group within a large family of SCPL proteins (Milkowski and Strack, 2004). It is tempting to assume, that there is a corresponding group of SCPL acyltransferases catalyzing the formation of a vast array of structurally diverse sinapate and possibly other hydroxycinnamate esters in Arabidopsis.

The present work documents the formation of a complex pattern of hydroxycinnamate conjugates, mainly sinapate esters, in *B. napus* seeds, composed of glucose, gentiobiose and kaempferol glycoside esters as well as sinapine, sinapoylmalate and an unusual cyclic spermidine amide. We address the question whether the formation of these compounds is dependent on sinapoylglucose indicating that enzymes of a SCPL acyltransferase family would be involved, as hypothesized for Arabidopsis. For this reason we performed a seed-specific dsRNAi-based suppression of the sinapate glucosyltransferase gene (*BnSGT1*) expression. This approach is part of a study (Hüsken et al., 2005) focusing on reducing the amount of antinutritive sinapate esters (Ismail et al., 1981; Kozlowska et al., 1990; Naczk et al., 1998) that compromise the use of the valuable protein-rich seed meal, considered essential for possibly establishing rape as a protein crop.

Seeds of transgenic *B. napus* plants harbouring a dsRNAi construct for suppression of *BnSGT1* expression were analyzed (Hüsken et al., 2005). The *BnSGT1* gene codes for the enzyme that catalyzes the formation of an 1-*O*-β-acetal ester (Milkowski et al., 2000a), here 1-*O*-sinapoyl-β-glucose, the activated substrate of SCPL acyltransferases that transfer the sinapoyl moiety to various acceptors in the formation of the corresponding esters. As expected, the transformed seeds exhibit a significant decrease in the accumulation of 1-*O*-sinapoyl-β-glucose (barely detectable) that in turn results in the disappearance or marked decrease of all the other sinapate esters in *B. napus* seeds. Thus, these esters are synthesized via sinapoylglucose and therefore their formation is most likely catalyzed by a group of enzymes encoded by a SCPL acyltransferase gene family.

2. Results and discussion

2.1. Analysis of *B. napus* constituents

HPLC analyses of methanolic extracts from *B. napus* seeds showed complex patterns of UV-absorbing compounds. Fifteen out of at least 30 detectable compounds were isolated using a three-step protocol. Extracts were fractionated on polyamide SC6 followed by preparative HPLC and final purification on Sephadex LH-20 columns.

Purified compounds were identified using a combination of high-field NMR spectroscopy and high resolution electrospray ionisation mass spectrometry (HR ESI-MS). In all cases the characteristic ^{1}H NMR spin systems of the individual units in each molecule were identified from inspection of the 1D and 2D COSY data. In those cases where the chemical shift data were not sufficient to position substituents of the compounds, this information was determined from responses in 1D ^{1}H NOE difference spectra or from cross peaks in 2D ROESY spectra. HR ESI-MS afforded the molecular formulae and complemented the NMR data.

Three of the identified structures (Fig. 1), i.e. sinapoylglucose (**4**), sinapine (sinapoylcholine, **6**) and sinapoylmalate (**9**), are components of a well-known pathway that is characteristic for members of the Brassicaceae (Linscheid et al., 1980; Bouchereau et al., 1992; Milkowski et al., 2004). The spectroscopic data of these sinapate esters are in accordance with those from compounds isolated from *Raphanus sativus* (Linscheid et al., 1980) and their data are not included in this communication.

Summarizing the structure elucidation of the compounds isolated from *B. napus* seeds, it is noticeable that the molecules essentially fall into three groups, namely glucose, gentiobiose and kaempferol glycoside esters along with a spermidine amide. In most cases the aromatic acid substituent is the ubiquitous *E*-sinapoyl moiety that was recognized in the ^{1}H NMR spectra from its two proton singlet signal at ca. 7 ppm, the singlet of the two aromatic methoxyl groups at ca. 3.9 ppm and olefinic protons at 7.7 and 6.5 ppm with the characteristic *trans* vicinal coupling of ca. 15.8 Hz. Characteristic chemical shifts again identified the presence of the non-symmetrical *E*-4,5-dihydroxy-3-methoxycinnamoyl (5-hydroxyferuloyl) group in **2** and caffeoyl group in **1**. Similarly the universal sugar unit is the β-glucopyranosyl moiety detected from the presence in the ^{1}H spectra of its doublet anomeric proton (^3J ca. 7.8 Hz) and ring protons whose chemical shifts are determined by the individual substitution patterns.

The presence of a single acyl substituent in the glucose ring at the anomeric carbon, C-1 of compounds **1**, **2**, and **4**, was indicated by a doublet (7.8 Hz) for H-1 at ca. 5.6 ppm. Introduction of a second acyl unit at C-2 (**11**, **12**, **13**, **15**) caused a subtle shift to 5.8 ppm of H-1 and the low field shift of the double doublet of H-2 to ca. 5.1 ppm. A downfield shift of H-2 is also evident in the terminal sugar units of gentiobiose (**13**), sophorose (**5**) and glucose (**10**). Substitution at C-6 in **7**, **14** and **15** was evident from the downfield shifts of the two double doublets of H-6A and B to 4.6 and 4.3–4.5 ppm, respectively. Hence the number and simple assessment of chemical shifts provided unambiguous

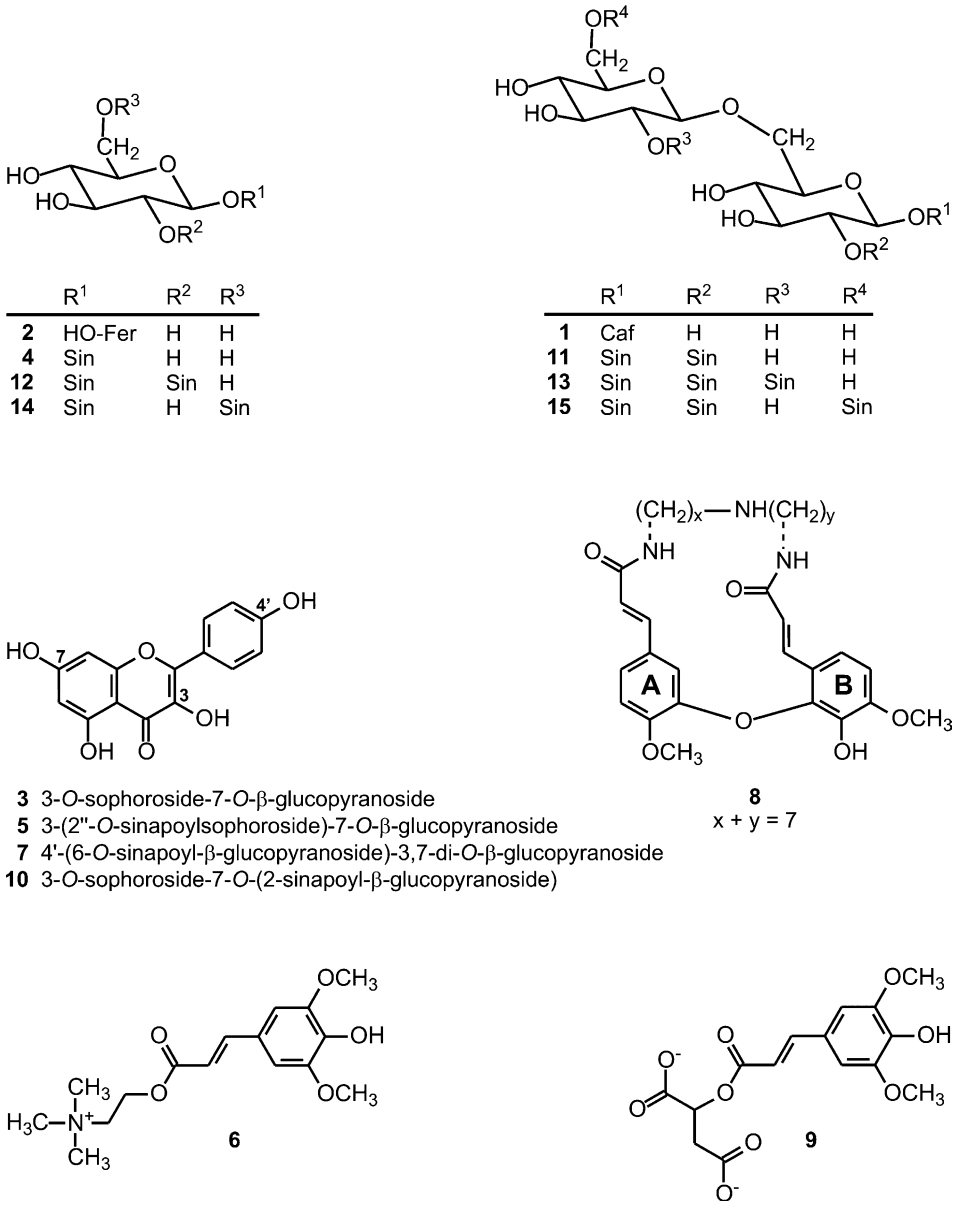


Fig. 1. Structures of compounds isolated from *Brassica napus* seeds; Sin = sinapoyl, HO-Fer = 5-hydroxyferuloyl, Caf = caffeoyl. Compound numbers are given according to the sequence in HPLC elution shown in Fig. 2.

evidence for the structure of the glucose, gentiobiose, and the sugar substituents in the kaempferol derivatives.

The C-7-substituted kaempferol derivatives were characterized by the low field substituent induced ^{1}H shifts of H-6 and H-8, and the AA″BB″ ^{1}H pattern of ring B. The assignments of the sugar units in these systems follow from the 2D COSY spectrum and the position of these units followed from correlations with the anomeric protons in the 1D NOE or 2D ROESY spectra.

The three units in the molecule of the unusual spermidine derivative (**8**), a cyclic spermidine alkaloid, were identified from the 2D COSY spectrum. The relative position of the substituents (methoxyl group and double bond side chains) in both aromatic systems was evident from the long range coupling in this spectrum and from cross peaks in the 2D ROESY spectrum. These substituents are the rare 3-hydroxy-4-methoxycinnamoyl (isoferuloyl) group (Greeaway et al., 1988), and the new 2,3-dihydroxy-4-methoxycinnamoyl (2-hydroxyisoferuloyl) group coupled via their *ortho* and *meta* hydroxyl groups forming a hydroxydiarylether. As two methylene groups of the spermidine suffer acylation and led to low field shifts relative to spermidine of ca. 0.5 ppm, there are two possible assignments; either the central nitrogen atom or the two terminal nitrogens are acylated. Consequently the HMQC spectrum (direct ^{13}C–^{1}H correlation) cannot distinguish between these two alternatives but does provide the two possible assignments. Fortunately, the signal to noise of the

HMBC spectrum (long-range ^{13}C–^{1}H correlation) was sufficient to detect the mutual three-bond correlations between the internal methylene groups H-4 and H-6 and the terminal methylene groups with the carbonyl groups. Thus, the data are only compatible with terminal group acylation. The HR ESI-MS data requires closure of a ring system through the aromatic residues in such a way as to provide a close proximity of H-2 of A with H-7 of B which is evident from their strong NOE interaction in the ROESY spectrum. This is only compatible with the isomeric structures (Fig. 1, compound **8**), arising from the two possible orientations of the spermidine moiety which can not be determined unambiguously.

The structure of this compound, the only amide found among the *B. napus* seed constituents, is described for the first time in a member of the Brassicaceae. It is similar to that of previously isolated compounds from the root bark of *Capparis decidua* (Ahmad et al., 1985, 1987). However, the proposed structures in these publications are questionable and Bienz et al. (2002) pointed to some inconsistencies in these studies and suggested alternative structures that are also different to compound **8**.

Most of the esters identified in the present study are known plant constituents, especially from members of the Brassicaceae. For example, compounds **4**, **6**, **9** and **12** have been found to accumulate in *Raphanus sativus* (Linscheid et al., 1980; Strack et al., 1984), *B. napus* (Bouchereau et al., 1992) and besides **4**, **6** and **9** additionally **2** in *Arabidopsis thaliana* (Lorenzen et al., 1996; Goujon et al., 2003). The gentiobiose di- and trisinapate esters (**11** and **13**) are known to accumulate in broccoli (Plumb et al., 1997; Price et al., 1997) together with gentiobioses esterified with sinapate as well as ferulate. Compound **13** was also found in fruits of *Boreava orientalis* (Sakushima et al., 1994). The kaempferol sinapoylglycoside **5** accumulates in leaves of *B. napus* (Olssen et al., 1998) and *B. oleracea* (Nielsen et al., 1993; Llorach et al., 2003) along with the non-acylated substance (**3**). Some other esters identified are to the best of our knowledge new plant products, i.e. one glucose ester (**14**), two gentiobiose esters (**1**, **15**) and two kaempferol sinapoylglucosides (**7**, **10**).

Changes in compound concentrations were followed during the first eight days of seedling development (Table 1). As shown in one of our previous publications on the sinapate ester metabolism in *B. napus* (Milkowski et al., 2004), the level of sinapine (**6**; 100 nmol seed^{-1}) decreases during seed germination resulting in the accumulation of sinapoylmalate (**9**; 86 nmol seed^{-1}) in the seedling cotyledons. Compounds **4**, **6** and **9** are members of the well-known pathway of sinapate esters in Brassicaceae seedlings (Linscheid et al., 1980; Bouchereau et al., 1992; Milkowski et al., 2004): sinapate is conjugated during seed development via 1-*O*-sinapoyl-β-glucose (**4**) with choline, resulting in the accumulation of sinapine (**6**). During seed germination, a sinapine esterase-catalyzed hydrolysis liberates sinapate that is conjugated a second time via sinapoylglucose with malate, resulting in the accumulation of sinapoylmalate (**9**).

Table 1
Content of *B. napus* compounds given as pmol seed^{-1} and developing seedlings (2-, 4- and 8-day-old), calculated as sinapate (**1**, **2**, **4**, **6**, **8**, **9**, **11**–**15**) or kaempferol equivalents (**3**, **5**, **7**, **10**)

Compound	Day 0 (Seed)	Day 2	Day 4	Day 8
1	320	480	n.d.[a]	n.d.
2	Trace	n.d.	n.d.	n.d.
3	692	130	150	115
4	14,500	39,100	13,000	5200
5	2680	1416	1850	1070
6	100,000	n.d.	n.d.	n.d.
7	560	443	292	n.d.
8	960	240	n.d.	n.d.
9	3600	37,800	73,300	86,000
10	281	n.d.	n.d.	n.d.
11	520	2560	2640	2200
12	3360	2480	400	n.d.
13	640	8900	13,800	10,000
14	100	240	n.d.	n.d.
15	100	80	n.d.	n.d.

[a] Not detected.

With regard to the minor compounds, it is interesting to note that along with sinapoylglucose (**4**) and traces of 5-hydroxyferuloylglucose (**2**), the disinapoylglucoses (**12**, **14**) also decrease in concentration during seedling development, whereas the amounts of two gentiobiose esters (**11**, **13**) slightly increase. From the kaempferol conjugates, the amount of **3** decreases slightly, that of **5** stays more or less constant, while **10** and **7** were not detected anymore in the seedlings after day 2 and 8, respectively. The spermidine alkaloid (**8**) also disappears during seedling development. The fate of these compounds is unknown.

2.2. Suppression of sinapate ester biosynthesis

The acyl donors in acyltransferase-catalyzed ester formation are mostly coenzyme A thioesters (Strack and Mock, 1993). It has been shown, however, with a number of plants that an alternative pathway is facilitated by acyltransferases that accept 1-*O*-β-acetal esters (1-*O*-β-glucose esters) (Strack and Mock, 1993; Milkowski and Strack, 2004). These acyltransferases apparently have been recruited from serine carboxypeptidases (SCPs) and adapted to take over acyltransfer functions (SCPL acyltransferase) (Li and Steffens, 2000; Lehfeldt et al., 2000; Shirley et al., 2001; Milkowski and Strack, 2004). According to the SwissProt database the proteome of Arabidopsis harbours 53 SCPL proteins, of which 21 form a distinct group including the known SCPL acyltransferases (Milkowski and Strack, 2004). In the light of the complex pattern of the *B. napus* sinapate esters and some other

hydroxycinnamate conjugates, we assumed that there is a corresponding SCPL acyltransferase family as in Arabidopsis. We addressed this problem by suppressing sinapoylglucose formation, the putative bottleneck reaction that is a prerequisite for the formation of all the *B. napus* sinapate esters.

Suppression of sinapate ester biosynthesis in seeds of *B. napus* was achieved by a dsRNAi approach designed to silence seed specifically the *BnSGT1* gene. In order to achieve this, *B. napus* was transformed with plasmid pLH–SGT–GUS, a dsRNAi construct that contains a part of the *BnSGT1*-encoding region as inverted repeat under the control of the seed-specific napin promoter (Hüsken et al., 2005). HPLC analysis of methanolic seed extracts revealed a strong reduction of sinapine content by 72% compared to the untransformed plants. For plants (T3 seeds from a homozygous T2 plant) with a single copy insertion it could be shown that the reduction of sinapine is a stable trait that seems not to interfere with other important seed characteristics. Currently this is under thorough investigation in our group using a metabolomics approach.

The loss of about 70% sinapine was accompanied by a strong decrease of the total sinapate ester content of almost 80%. The concentration of sinapoylglucose and most of the other sinapate esters seemed to be below the HPLC detection limit (Fig. 2). Thus suppression of the *BnSGT1* gene expression by the dsRNAi approach indicates that sinapoylglucose is indeed the central precursor not only for sinapine, sinapoylmalate and 1,2-disinapoylglucose biosyntheses (Fig. 3; SCT, SMT and SST activities) but also for transacylation reactions leading to the diverse pattern of the other sinapate esters in *B. napus* seeds (putative SCPL acyltransferase activities). We propose that the sinapate esters identified in this study are produced by distinct acyltransferases of the SCPL-type proteins giving rise to a large protein family of SCPL acyltransferases in *B. napus* as assumed for Arabidopsis (Milkowski and Strack, 2004).

Since we could not achieve stronger reduction of sinapine and some other minor components, it might be possible that *BnSGT1*-related genes are involved in sinapate ester biosynthesis. As was found for Arabidopsis (Milkowski et al., 2000b), there are related genes in *B. napus*

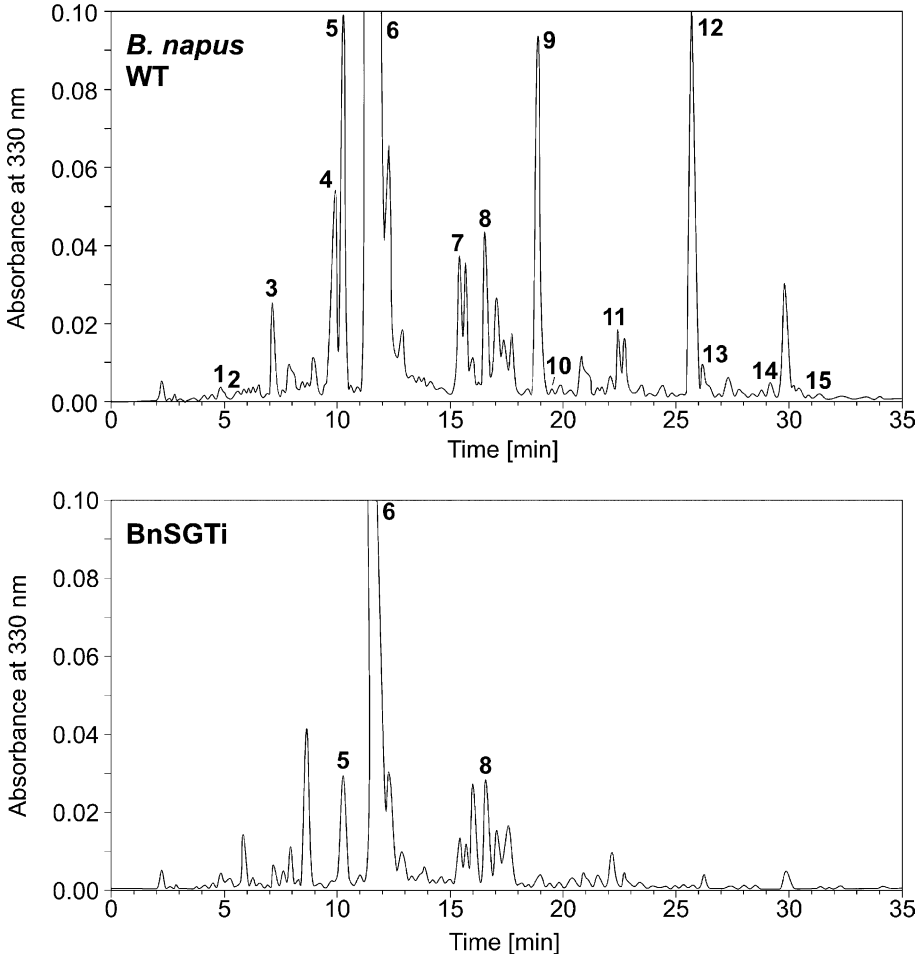


Fig. 2. HPLC traces of methanolic extracts of wild type (**WT**) *Brassica napus* seeds and those carrying the dsRNAi construct pLH–SGT–GUS (**BnSGTi**). Peak numbers correspond to compound numbers in Fig. 1.

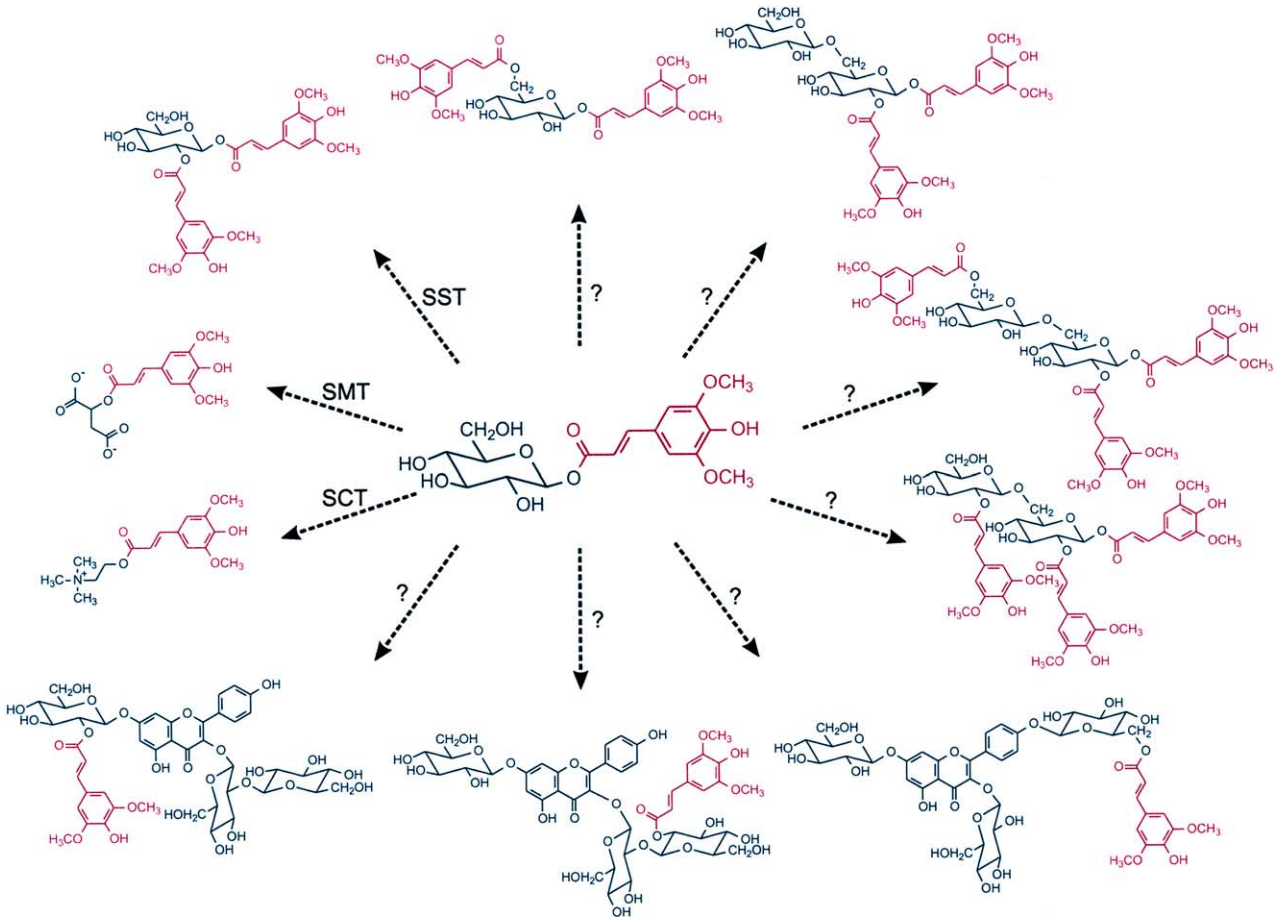


Fig. 3. Proposed *Brassica napus* SCPL acyltransferase family, catalyzing sinapoylglucose-dependent formation of sinapate esters; **SCT** and **SMT**, 1-*O*-sinapoylglucose:choline sinapoyltransferase and 1-*O*-sinapoylglucose:malate sinapoyltransferase from *Brassica napus*; **SST**, assumed to be homologous to the 1-*O*-sinapoylglucose:1-*O*-sinapoylglucose 2-*O*-sinapoyltransferase from *Raphanus sativus* (Dahlbender and Strack, 1986); ?, putative SCPL acyltransferase activities.

encoding enzymes that catalyze (in vitro) the formation of sinapoylglucose, along with other hydroxycinnamoylglucoses (J. Mittasch and C. Milkowski, unpublished). On the other hand, suppression of *BnSGT1* might not be complete. To test suppression strength, experiments are underway to quantify mRNA abundance in developing seeds of T4 plants.

In contrast to hydroxycinnamate esters, up to now the formation of amides seems to be exclusively dependent on coenzyme A activation, see e.g. amide formation with agmatine in barley seedlings (Bird and Smith, 1983) or tyramine and putrescine in tobacco (Negrel and Martin, 1984; Meurer-Grimes et al., 1989). In addition, a spermidine hydroxycinnamoyltransferase, dependent on hydroxycinnamoyl-CoA as acyl donor, has been characterized from tobacco (Negrel et al., 1991). It is most likely that the spermidine amide (**8**) in *B. napus* is also synthesized via coenzyme A thioesters. Thus, in comparison with the control plants, it is not surprising that the amount of this compound does not change significantly in the plants harbouring the dsRNAi construct for suppression of *BnSGT1* expression (Fig. 2).

2.3. Phenolic seed constituents of the progenitors of *B. napus*

We analyzed the phenolic seed constituents of *B. oleracea* and *B. rapa*, the progenitors of *B. napus*. It was shown in a previous publication (Milkowski et al., 2004) that Southern blot analyses of the genes that determine the formation of sinapine, i.e. *BnSGT1* and *BnSCT*, reflect the fact that the genome of the amphidiploid *B. napus* contains the genomes of *B. oleracea* and *B. rapa* (Schenck and Röbbelen, 1982). Thus, it can be assumed that the *B. napus* SCPL acyltransferase gene family most likely derives from the two progenitors. Since we do not know, however, anything about possible gene organization and control of expression of the respective genes, caution is advisable in assuming that the phenolic pattern of *B. napus* seeds is composed

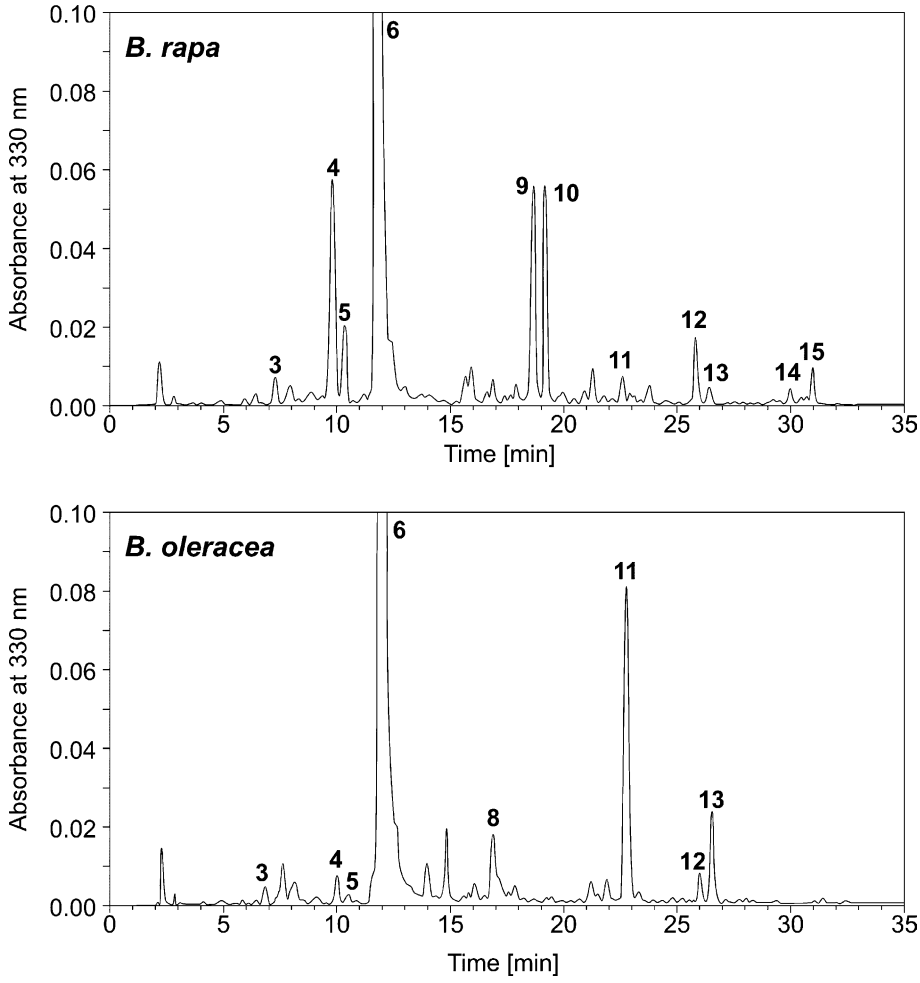


Fig. 4. HPLC traces of methanolic seed extracts from *Brassica oleracea* and *Brassica rapa*. Peak numbers correspond to compound numbers in Fig. 1.

of the compounds detected in the seeds of *B. oleracea* and *B. rapa* (Fig. 4) and the amount of some compounds might be below the HPLC detection limit. Nevertheless, it seems possible that compounds **9**, **10**, **14** and **15** of *B. rapa* add to those found in *B. oleracea* seeds (compare Fig. 4 with the HPLC trace of *B. napus* WT in Fig. 2). Genes involved in the formation of the spermidine amide (**8**) might be derived from *B. oleracea*. It is also interesting to note that formation of sinapoylmalate (**9**) in *B. napus* obviously results from expression of the *B. rapa* SMT gene. It was shown in an earlier publication that red cabbage (*B. oleracea* L. var. *oleracea*) seedlings lack accumulation of sinapoylmalate (Strack et al., 1978), indicating that this plant does not harbour or express the *SMT* gene.

On the genomic level, possible differences in gene organization between A and C genomes combined in the amphidiploid genome of *B. napus* (AACC $n = 19$) and those of the recent cultivars of *B. rapa* (AA, $n = 10$) and *B. oleracea* (CC, $n = 9$) used in this study should be considered (U, 1935). Regarding gene regulation, changes in gene expression patterns are likely to accompany amphidiploid formation to adjust the potential of the novel composite genome to the metabolic need of the plant.

3. Conclusion

In conclusion, we have shown in the present work that the complex pattern of sinapate esters in *B. napus* seeds is dependent on the formation of 1-*O*-sinapoyl-β-glucose. This gives rise to the assumption that enzymes of an SCPL acyltransferase family, known to accept this glucose ester as acyl donor, catalyze the formation of the accumulating esters. With these results we introduce another model system for studying evolution of plant secondary metabolism. As assumed for Arabidopsis (Milkowski and Strack, 2004) and discussed for *B. napus* herein, recruitment of SCPL acyltransferases from hydrolases of the serine carboxypeptidase-type, followed by diversification of the protein structures in plant evolution, leads to a distinct group of enzymes catalyzing an array of structurally diverse hydroxycinnamate esters.

The obvious demand of plant secondary product diversity drives evolution of genetic diversification leading to multigene families.

4. Experimental

4.1. Plant material

Seeds of spring oilseed rape (*Brassica napus* L. var. *napus* cv. Drakkar), forage kale (*B. oleracea* L. var. *medullosa* cv. Markola) and turnip (*B. rapa* L. var. *silvestris* cv. Rex) were purchased from the Norddeutsche Pflanzenzucht, Holtsee, Germany. Seedlings were grown under greenhouse conditions and cotyledons harvested after 2, 4, and 8 days of development. Construction of the dsRNAi suppression plasmid and *Agrobacterium tumefaciens*-mediated transformation and plant regeneration are described elsewhere (Hüsken et al., 2005). Seeds of T2 plants (T3 seeds) were analyzed to evaluate the effect of seed-specific suppression of *BnSGT1* gene expression.

4.2. Extraction and isolation

Seed material (400 g) was first extracted with hexane using an Ultra-Turrax homogenizer. After centrifugation, the defatted pellets were extracted three times with 80% aq. MeOH. The combined and concentrated extracts were chromatographed on Polyamide SC6 (Macherey and Nagel, Düren, Germany), stepwise eluted with H_2O, 50% aq. MeOH and 100% MeOH. Sinapine, the main product in the water fraction, was separated from other seed constituents by formation of the poorly soluble crystalline thiocyanate after addition of solid KSCN prior to further chromatographic steps.

The polyamide eluates were concentrated and further chromatographed by repeated preparative HPLC using a DELTAPAK 300/50 C18 column (Waters, USA) or VP 250/40 Nucleosil 100-10 C18 (Macherey-Nagel, Düren, Germany) using different gradients corresponding to the separation with 1% aq. HOAc (solvent A) and MeCN or MeOH (solvent B) with flow rates of 20 ml min^{-1}. The compounds were photometrically detected (HPLC-DAD, 330 nm). All compounds were finally purified on a Sephadex LH-20 column with 50% aq. MeOH as solvent.

4.3. Analytical HPLC

Fifty seeds or cotyledon pairs at different developmental stages were extracted three times with 80% aq. MeOH and the combined extracts were concentrated to dryness, redissolve in 2 ml 80% aq. MeOH, centrifuged and aliquots injected onto a Nucleosil C18 column (5 μm; 250 × 4 mm i.d.; Macherey-Nagel, Düren, Germany). Separation was achieved using a 40-min linear gradient at 1 ml min^{-1} from 10% to 50% MeCN in 1.5% aq. H_3PO_4. Compounds were photometrically detected (at 330 and 265 nm; maxplot between 210 and 500 nm) by a Waters 2996 photodiode array detector (DAD). Calibration data for quantification were obtained from sinapate for the hydroxycinnamate esters except from kaempferol as standard compound for the kaempferol conjugates.

4.4. NMR

All 1D and 2D [COSY, ROESY (mixing time 500 ms)] ^1H, and 2D ^1H detected ^{13}C–^1H [HMQC, HMBC] NMR spectra were recorded at 300 K on either a Bruker AVANCE DMX 600 or ARX 400 NMR spectrometers locked to the major deuterium signal of the solvent, CD_3OD. Chemical shifts are given in ppm relative to the residual solvent signal at 3.35 ppm and coupling constants in Hz.

4.5. HR ESI-FT-ICR-MS

The high resolution electrospray ionization (HR ESI) mass spectra were obtained from a Bruker Apex 70 e Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR-MS) Bruker Daltonics, Billerica, USA) equipped with an Infinity^{TM} cell, a 7.0 T superconducting magnet (Bruker, Karlsruhe, Germany), an RF-only hexapole ion guide and an external electrospray ion source (Agilent, off axis spray). Nitrogen was used as drying gas at 150 °C. The sample solutions were introduced continuously via a syringe pump with a flow rate of 120 l h^{-1}. All data were acquired with 256k data points and zero filled to 1024k by averaging 32 scans.

4.6. ESI-MS/MS

The positive ion ESI-MS/MS data of the spermidine conjugate (**8**) were obtained from a Finnigan MAT TSQ 7000 instrument (electrospray voltage 4.5 kV; heated capillary temperature 220 °C; sheath gas: nitrogen) coupled with a Surveyor MicroLC system equipped with a RP18 column (5 μm, 1 × 100 mm, Ultrasep). For HPLC a gradient system was used starting from H_2O:MeCN 85:15 (each contained 0.3% HOAc) to 10:90 within 15 min followed by a 10-min isocratic period; flow rate 70 μl min^{-1}. The collision-induced dissociation (CID) mass spectrum of the $[M + H^+]$ of compound 8 (RT = 6.10 min) was performed during the HPLC run using a collision of 35 eV; collision gas: argon, collision pressure: 1.8×10^{-2} Torr.

4.7. 1-O-E-Caffeoyl-β-gentiobiose (1)

UV (HPLC-DAD): λ_{max} (nm) 217, 235, 300sh, 321. ^{1}H NMR: δ = 7.74 [*d*; H-7; *J*(7–8) 15.9], 7.58 [*d*; H-2; *J*(2–6) 2.0], 7.24 [*dd*; H-6; *J* (5–6) 8.4], 6.91 [*d*; H-5], 6.43 [*d*; H-8], 5.61 [*d*; H-1′, *J*(1′–2′) 7.9], 4.85 [*d*; H-1″, *J*(1″–2″) 7.4], 3.98 [*dd*; H-6′A; *J*(5′–6′A) 2.2, *J*(6′A–6′B) 12.1], 3.89 [*dd*; H-6″A; *J*(5″–6″A) 2.0, *J*(6″A-6″B) 12.2]; 3.76 [*dd*; H-6′B; *J*(5′–6′B) 6.1], 3.73 [*dd*; H-6″B; *J*(5″–6″B) 4.6], 3.57 [*dd*; H-2″; *J*(2″–3″) ca. 9], 3.52 [*m*; H-5′], 3.47 [*m*; H-2′, H-5″], 3.56-3.39 [*m*; rest sugar protons]. HR-(+)ESI-MS: 527.1372 $[M + Na]^+$ (calc. for $C_{21}H_{28}O_{14}Na$ 527.1377), HR-(−)ESI-MS: 503.1401 $[M - H]^-$ (calc. for $C_{21}H_{27}O_{14}$ 503.1406).

4.8. 1-O-E-(5-Hydroxyferuloyl)-β-glucopyranose (2)

UV (HPLC-DAD): λ_{max} (nm) 238, 331. ^{1}H NMR: δ = 7.69 [*d*; H-7; *J*(7–8) 15.9], 6.83, 6.80 [*d* × 2; H-2, H-6; *J*(2–6) 1.7], 6.38 [*d*; H-8], 5.61 [*d*; H-1′; *J*(1′–2′) 7.9], 3.92 [*s*; 3-OMe], 3.89 [*dd*; H-6′A; *J* (5′–6′A) 2.0, *J*(6′A–6′B) 12.1], 3.73 [*dd*; H-6′B; *J*(5′–6′B) 4.9], 3.50 [*dd*; H-2′; *J*(2′–3′) 9.4], 3.46 [*m*; H-5′], 3.47–3.40 [*m*; H-3′, H-4′]. HR-(+)ESI-MS: 395.0956 $[M + Na]^+$ (calc. for $C_{16}H_{20}O_{10}Na$ 395.0949), HR-(−)ESI-MS: 371.0986 $[M - H]^-$ (calc. for $C_{16}H_{19}O_{10}$ 371.0984).

4.9. Kaempferol 3-O-sophoroside-7-O-β-glucopyranoside (3)

UV (HPLC-DAD): λ_{max} (nm) 265, 331. ^{1}H NMR: δ = 8.11 ["*d*"; H-2′/6′; *J*(2′–3′) + (2′–5′) 8.9], 6.96 ["*d*"; H-3′/5′], 6.81 [*d*; H-8; *J*(6–8) 2.1], 6.54 [*d*; H-6], sophorose moiety: 5.54 [*d*; H-1; *J* (1–2) 7.4], 4.79 [*d*, H-1′ overlap with water signal], 3.82 [*dd*; H-6′A; *J*(5′–6′A) 2.5, *J* (6′A-6′B) 12.0], 3.78 [overlap; H-2], 3.75 [*dd*; H-6A], 3.73 [*dd*; H-6′B], 3.65 [*dd*; H-3; *J*(2–3) 8.9, *J*(3–4) 8.9], 3.53 [*dd*; H-6B], 3.45-3.40 [overlap; H-2′, H-3′, H-4′], 3.39 [*dd*; H-4], 3.33 [*ddd*; H-5′; *J*(5′–6′B) 6.5, *J*(4′–5′) 9.2], 3.25 [*ddd*; H-5; *J*(4–5) 9.8, *J*(5–6A) 2.2, *J*(5–6B) 5.5], glucose moiety: 5.10 [*d*; H-1; *J*(1–2) 7.5], 3.96 [*dd*; H-6A; *J*(5–6A) 2.2, *J* (6A–6B) 12.1], 3.75 [*dd*; H-6B], 3.57 [*ddd*; H-5; *J*(5–6B) 5.7, *J*(4–5) 9.6], 3.52 [*m*; H-2, H-3], 3.44 [*m*; H-4]. HR-(+)ESI-MS: 795.1976 $[M + Na]^+$ (calc. for $C_{33}H_{40}O_{21}Na$ 795.1954), HR-(−)ESI-MS: 771.1981 $[M - H]^-$ (calc. for $C_{33}H_{39}O_{21}$ 771.1989).

4.10. Kaempferol 3-(2″-O-E-sinapoylsophoroside)-7-O-β-glucopyranoside (5)

UV (HPLC-DAD): λ_{max} (nm) 226, 240sh, 268, 333. ^{1}H NMR: δ = 7.95 ["*d*"; H-2′/6′; *J*(2′–3′) + (2′–5′) 8.9], 6.94 ["*d*"; H-3′/5′], 6.46 [*d*; H-8; *J*(6–8) 2.1], 6.41 [*d*; H-6], sinapoyl moiety: 7.36 [*d*; H-7; *J*(7–8) 15.8], 6.33 [*s*; H-2/6], 6.14 [*d*; H-8], 3.67 [*s*; 3,5-OMe], sophorose moiety: 6.17 [*d*; H-1; *J*(1-2) 7.9], 5.25 [*d*; H-1′, *J*(1′–2′) 7.8], 4.96 [*dd*; H-2′; *J*(2′–3′) 9.5], 3.96 [*dd*; H-6′A; *J*(5′–6′A) small, *J*(6′A–6′B) ∼ 11.5], 3.81 [*dd*; H-6′B], 3.80 [*dd*; H-3′], 3.73 [*dd*; H-3; *J*(2-3) 9.0, *J*(3–4) 9.0], 3.70 [*dd*; H-6A; *J*(5–6A) 2.1, *J*(6A–6B) 12.2], 3.57 [*m*; H-5′], 3.60–3.55 [*m*; H-4′], 3.55 [*dd*; H-2], 3.52 [*dd*; H-6B], 3.35 [*dd*, H-4], 3.29 [*ddd*; H-5; *J*(5–6B) 5.4], glucose moiety: 5.13 [*d*; H-1; *J*(1-2) 7.69], 3.98 [*dd*; H-6A; *J*(5–6A) 2.3, *J*(6A–6B) 12.1], 3.77 [*dd*; H-6B; *J*(5–6B) 5.8], 3.61 [*dd*; H-3; *J*(2–3) 9.1, *J*(3–4) 9.1], 3.60 [*ddd*; H-5; *J*(4–5) 9.6], 3.53 [overlap; H-2], 3.45 [*dd*, H-4]. HR-(−)ESI-MS: 977.2538 $[M - H]^-$ (calc. for $C_{44}H_{49}O_{25}$ 977.2568).

4.11. Kaempferol 4′-(6-O-E-sinapoyl-β-glucopyranoside)-3,7-di-O-β-glucopyranoside (7)

UV (HPLC-DAD): λ_{max} (nm) 223, 240sh, 267, 321. ^{1}H NMR: δ = 8.06 ["*d*"; H-2′/6′; *J*(2′–3′) + (2′–5′) 9.0], 7.18 ["*d*"; H-3′/5′], 6.53, 6.53 [*d* × 2; H-8, H-6; *J*(6–8) 2.2], sinapoyl moiety: 7.57 [*d*; H-7; *J*(7–8) 15.9], 6.86 [*s*; H-2/6], 6.41 [*d*; H-8], 3.89 [*s*; 3,5-OMe], 4′-glucose moiety: 5.09 ["*d*"; H-1; *J*(1–2) 7.7], 4.57 [*dd*, H-6A, *J*(5–6A) 2.5, *J*(6A–6B) 11.9], 4.50 [*dd*; H-6B; *J*(5–6B) 7.3], 3.82 [*m*; H-5], 3.60–3.43 [*m*; H-4], 3.58 [*m*; H-2, H-3], 3-glucose moiety: 5.34 ["*d*"; H-1; *J*(1–2) 7.6], 3.73 [*m*; H-6A], 3.57 [*m*; H-6B], 3.45 [*m*; H-2, H-3], 3.35 [*m*; H-4], 3.25 [*ddd*; H-5; *J*(5–6A) 2.4, *J*(5–6B) 5.5, *J*(4–5) 9.6], 7-glucose moiety: 5.11 ["*d*"; H-1; *J*(1–2) 7.9], 3.94 [*dd*; H-6A; *J*(5–6A) 2.3, *J*(6A–6B) 12.2], 3.74 [*m*; H-6B], 3.60–3.43 [*m*; H-4], 3.56 [*m*; H-2, H-3], 3.55 [*m*; H-5]. HR-(+)ESI-MS: 1001.2516 $[M + Na]^+$ (calc. for $C_{44}H_{50}O_{25}Na$ 1001.2533), HR-(−)ESI-MS: 977.2559 $[M–H]^-$ (calc. for $C_{44}H_{49}O_{25}$ 977.2568).

4.12. Spermidine conjugate (8)

UV (HPLC-DAD): λ_{max} (nm) 236, 305, 310sh. ^{1}H NMR: δ = aromatic moiety A: 7.34 [*d*; H-7; *J*(7–8) 15.7], 7.19 [*dd*; H-6; *J*(2–6) 1.8, *J*(5–6) 8.4], 7.12 [*d*; H-5], 6.78 [*d*; H-2], 6.18 [*d*; H-8], 4.02 [*s*; 4-OMe], aromatic moiety B: 7.64 [*d*; H-7; *J*(7–8) 15.7], 7.31 [*d*; H-6; *J*(5–6) 8.7], 6.99 [*d*; H-5], 6.59 [*d*; H-8], spermidine methylene groups H-2 to H-4 and H-6 to H-9: 3.43 [*m*; H-2] 3.41 [H-9], 2.93 [*m*; H-6], 2.84 [*m*; H-4], 1.88 [*m*; H-3], 1.67 [*m*; H-7, H-8]. ^{13}C NMR (CD_3OD + trace CF_3CO_2H): δ = aromatic moiety A: 129.0 (C-1), 115.1 (C-2), 149.5 (C-3), 152.5 (C-4), 114.1(C-5) 124.9 (C-6), 141.6 (C-7), 119.8 (C-8), 168.8 (C-9), 56.8 (4-OMe), aromatic moiety B: 124.0 (C-1), 143.2 (C-2), 140.8 (C-3), 152.3 (C-4), 110.1 (C-5), 118.3 (C-6), 137.3 (C-7), 119.8 (C-8), 168.9 (C-9), 56.8 (4-OMe), spermidine: 36.8 (C-2), 27.6 (C-3), 45.8 (C-4), 48.5 (C-6), 26.9, 24.1 (C-7, C-8), 39.1 (C-9). HR-(+)ESI-MS: 496.2442 $[M + H]^+$ (calc. for $C_{27}H_{34}N_3O_6$ 496.2442), HR-(−)ESI-MS: 494.2288 $[M - H]^-$ (calc. for $C_{27}H_{32}N_3O_6$ 494.2297). ESI-CIDMS, *m/z* (rel. intensity, %): 496 ($[M + H]^+$, 14),

479 (5), 478 ([M + H − H_2O]$^+$, 5), 425 (7), 408 ([M + H-putrescine]$^+$, 100), 351 ([M + H-spermidine]$^+$, 18), 325 (12), 283 (5), 268 (6), 232 (8), 216 (5), 191 (6), 175 (18), 149 (12).

4.13. Kaempferol 3-O-sophoroside-7-O-(2-O-E-sinapoyl-β-glucopyranoside) (**10**)

UV (HPLC-DAD): λ_{max} (nm) 240, 265, 335. ^{1}H NMR: δ = 8.08 ["*d*"; H-2′/6′; *J*(2′–3′)+(2′–5′) 8.8], 6.92 ["*d*"; H-3′/5′], 6.74 [*d*; H-8; *J*(6–8) 2.1], 6.45 [*d*; H-6], sinapoyl moiety: 7.71 [*d*; H-7; *J*(7–8) 15.9], 6.93 [*s*; H-2/6], 6.46 [*d*; H-8], 3.89 [*s*; 3,5-OMe], sophorose moiety: 5.52 [*d*; H-1; *J*(1-2) 7.5], 4.78 [*d*, H-1′ overlap with water signal], 3.81 [*dd*; H-6′A; *J*(5′–6′A) 2.6, *J*(6′A–6′B) 11.9], 3.76 [overlap; H-2], 3.71 [*dd*; H-6A; *J*(5–6A) 2.1, *J*(6A–6B) 11.9], 3.71 [*dd*; H-6′B; *J*(5–6B) 5.0], 3.64 [*dd*; H-3; *J*(2–3) 9.0, *J*(3–4) 9.0], 3.50 [*dd*; H-6B; *J*(5–6B) 5.6], 3.45–3.40 [overlap; H-2′, H-3′, H-4′], 3.38 [*dd*; H-4], 3.32 [*ddd*; H-5′; *J*(4′–5′) 9.3], 3.23 [*ddd*; H-5; *J*(4–5) 9.7], glucose moiety: 5.38 [*d*; H-1; *J*(1–2) 8.0], 5.16 [*dd*; H-2; *J*(2-3) 9.5, 4.00 [*dd*; H-6A; *J*(5–6A) 2.1, *J*(6A–6B) 12.0], 3.80 [*dd*; H-6B; *J*(5–6B) 5.6], 3.79 [*dd*; H-3], 3.67 [*ddd*; H-5; *J*(4–5) 9.7], 3.57 [*dd*; H-4; *J*(3–4) 9.0]. HR-(+)ESI-MS: 1001.2516 [M + Na]$^+$ (calc. for $C_{44}H_{50}O_{25}Na$ 1001.2533), HR-(−)ESI-MS: 977.2548 [M − H]$^-$ (calc. for $C_{44}H_{49}O_{25}$ 977.2568).

4.14. 1,2-Di-O-E-sinapoyl-β-gentiobiose (**11**)

UV (HPLC-DAD): λ_{max} (nm) 240, 330. ^{1}H NMR: δ = 7.68, 7.67 [*d*×2; H-7^A, H-7^B; *J*(7^A–8^A)/*J*(7^B–8^B)15.8, 15.8], 6.93, 6.89 [*s*×2; H-2^A/6^A, H-2^B/6^B], 6.45, 6.37 [*d*×2; H-8^A, H-8^B], 5.82 [*d*; H-1′; *J*(1′–2′) 8.4], 5.13 [*dd*; H-2′; *J*(2′–3′) 9.5], 4.41 [*d*; H-1″, *J*(1″–2″) 7.8], 4.26 [*dd*; H-6′A; *J*(6′A–5′) 1.8, *J*(6′A–6′B) 11.4], 3.90 [*m*; H-6″A], 3.89 [*m*, H-6′B], 3.90, 3.88 [*s*×2; 3^A/5^A-OCH_3, 3^B/5^B-OCH_3], 3.77 [*dd*; H-3′; *J*(3′–2′) 9.1, *J*(3′-4′) 9.3], 3.73 [*m*; H-5′], 3.72 [*dd*; H-6″B; *J*(6″B–5″) 5.4, *J*(6″B–6″A) 11.8], 3.64 [*dd*; H-4′; *J*(4′–5′) 9.4], 3.41 [*dd*; H-3″; *J*(3″–2″) 9.1, *J*(3″–4″) 9.2], ~3.35 [*m*, H-4″], 3.31 [*m*; H-5″], 3.28 [*dd*, H-2″]. HR-(+)ESI-MS: 777.2187 [M + Na]$^+$ (calc. for $C_{34}H_{42}O_{19}Na$ 777.2213), HR-(−)ESI-MS: 753.2258 [M-H]$^-$ (calc. for $C_{34}H_{41}O_{19}$ 753.2248).

4.15. 1,2-Di-O-E-sinapoyl-β-glucopyranose (**12**)

UV (HPLC-DAD): λ_{max} (nm) 239, 330. ^{1}H NMR: δ = 7.68, 7.67 [*d*×2;H-7^A, H-7^B; *J* (7^A-8^A)/*J*(7^B-8^B) 15.9, 15.8], 6.92, 6.89 [*s*×2; H-2^A/6^A, H-2^B/6^B], 6.45, 6.37 [*d*×2; H-8^A, H-8^B], 5.84 [*d*; H-1′; *J*(1′–2′) 8.3], 5.11 [*dd*; H-2′; *J*(2′–3′) 9.6], 3.94 [*dd*; H-6′A; *J*(6′A–5′) 1.6, *J* (6′A-6′B) 12.3], 3.90, 3.88 [*s*×2; 3^A/5^A-OCH_3, 3^B/5^B-OCH_3], 3.79 [*dd*; H-6′B; *J*(6′B–5′) 4.8], 3.76 [*m*; H-3′], 3.58–3.53 [*m*; H-4′, H-5′]. HR-(+)ESI-MS: 615.1690 [M + Na]$^+$ (calc. for $C_{28}H_{32}O_{14}Na$ 615.1684), HR-(−)ESI-MS: 591.1723 [M − H]$^-$ (calc. for $C_{28}H_{31}O_{14}$ 591.1719).

4.16. 1,2,2′-Tri-O-E-sinapoyl-β-gentiobiose (**13**)

UV (HPLC-DAD): λ_{max} (nm) 239, 327. ^{1}H NMR: δ = 7.79, 7.63, 7.60 [*d*×3; H-7^A, H-7^B, H-7^C; *J*(7^A–8^A)/*J*(7^B–8^B)/*J*(7^C–8^C) 15.9, 15.8, 15.8], 7.02, 6.89, 6.87 [*s*×3; H-2^A/6^A, H-2^C/6^C, H-2^B/6^B], 6.61, 6.38, 6.27 [*d*×3; H-8^A, H-8^C, H-8^B], 5.78 [*d*; H-1′, *J*(1′–2′) 8.3], 5.04 [*dd*; H-2′, *J*(2′-3′) 9.5], 4.86 [*dd*; H-2″, *J*(1″–2″) 8.0, *J*(2″–3″) 9.4], 4.76 [*d*; H-1″], 4.19 [*dd*; H-6′A; *J*(5′–6′) ca. 1, *J*(6′A–6′B) 12.1], 3.93 [*dd*; H-6″A; *J*(5″–6″A) 2.2, *J*(6″A–6″B) 11.8], 3.89, 3.89, 3.86 [*s*×3; 3^B/5^B-OCH_3,3^C/5^C-OCH_3,3^A/5^A-OCH_3], 3.85 [*dd*; H-6′B; *J*(5′–6′B) 6.1] 3.74 [*dd*; H-6″B; *J*(5″–6″B] 5.8], 3.62 [*dd*; H-3″; *J*(3″–4″) 9.1], 3.62 [*ddd*; H-5′], 3.61 [*dd*; H-3′; *J*(3′–4′) 9.2], 3.44 [*dd*; H-4″; *J*(4″–5″) 9.5], 3.41 [*dd*; H-4′; *J*(4′–5′) 9.5], 3.34 [*ddd*; H-5″]. HR-(+)ESI-MS: 983.2794 [M + Na]$^+$ (calc. for $C_{45}H_{52}O_{23}Na$ 983.2792), HR-(−)ESI-MS: 959.2823 [M − H]$^-$ (calc. for $C_{45}H_{51}O_{23}$ 959.2827).

4.17. 1,6-Di-O-E-sinapoyl-β-glucopyranose (14)

UV (HPLC-DAD): λ_{max} (nm) 239, 329. ^{1}H NMR: δ = 7.74, 7.6 [*d*×2; H-7^A, H-7^B; *J*(7^A–8^A)/*J*(7^B–8^B)15.9, 15.9], 6.95, 6.94 [*s*×2; H-2^A/6^A, H-2^B/6^B], 6.45, 6.45 [*d*×2; H-8^A, H-8^B], 5.64 [*d*; H-1′; *J*(1′–2′) 7.5], 4.55 [*dd*; H-6′A; *J*(5′–6′A) 2.1, *J*(6′A–6′B) 12.1], 4.38 [*dd*; H-6′B; *J*(5′–6′B) 5.9], 3.91, 3.91 [*s*×2; 3^A/5^A-OCH_3,3^B/5^B-OCH_3], 3.73 [*ddd*; H-5′; *J*(4′–5′) 9.2], 3.55 [*dd*; H-2′; *J*(2′–3′) ca. 8.9], 3.53-3.45 [*m*; H-3′, H-4′]. HR-(+)ESI-MS: 615.1681 [M + Na]$^+$ (calc. for $C_{28}H_{32}O_{14}Na$ 615.1684), HR-(−)ESI-MS: 591.1719 [M − H]$^-$ (calc. for $C_{28}H_{31}O_{14}$ 591.1719).

4.18. 1,2,6′-Tri-O-E-sinapoylgentiobiose (**15**)

UV (HPLC-DAD): λ_{max} (nm) 239, 328. ^{1}H NMR: δ = 7.65, 7.64, 7.63 [*d*×3; H-7^A, H-7^B, H-7^C; *J*(7^A–8^A)/*J* (7^B–8^B)/*J*(7^C–8^C) 15.9, 15.9, 15.8], 6.93, 6.87, 6.86 [*s*×3; H-2^A/6^A, H-2^B/6^B, H-2^C/6^C], 6.47, 6.41, 6.32 [*d*×3; H-8^A, H-8^B, H-8^C], 5.84 [*d*; H-1′, *J*(1′–2′) 8.4], 5.12 [*dd*; H-2′, *J*(2′–3′) 9.5], 4.62 [*dd*; H-6″A; *J*(5″–6″A) 2.0, *J*(6″A–6″B) 11.9], 4.48 [*d*; H-1″; *J*(1″–2″) 7.8], 4.33 [*dd*; H-6″B; *J*(5″–6″B) 6.0], 4.21 [*dd*; H-6′A; *J*(5′–6′A) 1.6, *J*(6′A–6′B) 11.8], 3.94 [*dd*; H-6′B], 3.89, 3.87, 3.86 [*s*×3; 3^A/5^A-OCH_3, 3^B/5^B-OCH_3, 3^C/5^C-OCH_3], 3.80 [*m*; H-5′], 3.78 [*dd*; H-3′; *J*(3′–4′) 9.6], 3.57 [*dd*; H-4′; *J*(4′–5′) 9], 3.45 [*dd*; H-3″; *J*(2″–3″) 8.9, *J* (3″–4″) 8.9], 3.41 [*dd*; H-4″; *J* (4″–5″) 9.3], 3.30 [*dd*; H-2″]. HR-(+)ESI-MS: 983.2778 [M + Na]$^+$ (calc. for $C_{45}H_{52}O_{23}Na$ 983.2792), HR-(−)ESI- MS: 959.2804 [M − H]$^-$ (calc. for $C_{45}H_{51}O_{23}$ 959.2827).

Acknowledgement

Research on SCPL proteins is supported by the DFG priority program 1152, "Evolution of Metabolic Diversity". This work is also part of the research project "NAPUS 2000 – Healthy Food from Transgenic Rape Seeds", supported by the Bundesministerium für Bildung und Forschung. We thank Ingrid Otschik (IPB), Beate Jaschok-Kentner and Christel Kakoschke (GBF) for technical assistance and Christine Kaufmann (IPB) for preparing the figures.

References

Ahmad, V.U., Arif, S., Amber, Atta-ur-Rahman, Usmanghani, K., Miana, C.A., 1985. A new spermidine alkaloid from *Capparis decidua*. Heterocycles 23, 3015–3020.

Ahmad, V.U., Arif, S., Amber, Atta-ur-Rahman, Fizza, K., 1987. Capparisidine, a new alkaloid from *Capparis decidua*. Liebigs. Ann. Chem. 2, 161–162.

Bell, J.M., 1993. Factors affecting the nutritional value of canola meal: A review. Can. J. Anim. Sci. 73, 679–697.

Bienz, S., Detterbeck, R., Ensch, C., Guggisberg, A., Häusermann, U., Meisterhans, C., Wendt, B., Werner, C., Hesse, M., 2002. Putrescine, spermidine, spermine, and related polyamine alkaloids. In: Cordell, G. (Ed.), The Alkaloids, vol. 58. Academic Press, Cornwall, pp. 83–338.

Bird, C.R., Smith, T.A., 1983. Agmatine coumaroyltransferase from barley seedlings. Phytochemistry 22, 2401–2403.

Bouchereau, A., Hamelin, J., Lamour, I., Renard, M., Larher, F., 1991. Distribution of sinapine and related compounds in seeds of *Brassica* and allied genera. Phytochemistry 30, 1873–1881.

Bouchereau, A., Hamelin, J., Renard, M., Larher, F., 1992. Structural changes in sinapic acid conjugates during seedling development. Plant Physiol. Biochem. 30, 467–475.

Dahlbender, B., Strack, D., 1986. 1-(Hydroxycinnamoyl)-glucose:1-(hydroxycinnamoyl)-glucose hydroxycinnamoyl transferase. Phytochemistry 25, 1043–1046.

Goujon, T., Sibout, R., Pollet, B., Maba, B., Nussaume, L., Bechtold, N., Lu, F., Ralph, J., Mila, I., Barriere, Y., Lapierre, C., Jouanin, L., 2003. A new *Arabidopsis thaliana* mutant deficient in the expression of O-methyltransferase impacts lignins and sinapoyl esters. Plant Mol. Biol. 51, 973–989.

Greeaway, W., Wollenweber, E., Scaysbrook, T., Whatley, F.R., 1988. Novel isoferulate esters identified by gas-chromatography mass-spectrometry in bud exudates of *Populus nigra*. J. Chromatogr. 448, 284–290.

Hegnauer, R., 1964. Chemotaxonomie der Pflanzen, vol. 3. Birkhäuser, Basel, pp. 594–595.

Hüsken, A., Baumert, A., Strack, D., Becker, H.C., Möllers, C., Milkowski, C., 2005. Reduction of sinapate ester content in transgenic oilseed rape by RNA interference-generated glucosyltransferase silencing. Mol. Breed, in press.

Ismail, F., Vaisey-Genser, M., Fyfle, B., 1981. Bitterness and astringency of sinapine and its components. J. Food Sci. 46, 1241–1244.

Kozlowska, H., Naczk, M., Shahidi, M., Zadernowski, R., 1990. Phenolic acids and tannins in rapeseed and canola. In: Shahidi, F. (Ed.), Canola and Rapeseed, Production, Chemistry, Nutrition and Processing Technology. Van Nostrand Reinhold, New York, pp. 103–210.

Lehfeldt, C., Shirley, A.M., Meyer, K., Ruegger, M.O., Cusumano, J.C., Viitanen, P.V., Strack, D., Chapple, C., 2000. Cloning of the SNG1 gene of Arabidopsis reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. Plant Cell 12, 1295–1306.

Li, A.X., Steffens, J.C., 2000. An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. Proc. Natl. Acad. Sci. USA 97, 6902–6907.

Linscheid, M., Wendisch, D., Strack, D., 1980. The structures of sinapic acid esters and their metabolism in *Raphanus sativus*. Z. Naturforsch. 35c, 907–914.

Llorach, R., Gil-Izquierdo, A., Ferreres, F., Tomas-Barberan, F.A., 2003. HPLC-DAD-MS/MS ESI characterization of unusual highly glycosylated acylated flavonoids from cauliflower (*Brassica oleracea* L. var. *botrytis*) agroindustrial byproducts. J. Agric. Food Chem. 51, 3895–3899.

Lorenzen, M., Racicot, V., Strack, D., Chapple, C., 1996. Sinapic acid ester metabolism in wild type and sinapoylglucose-accumulating mutant of *Arabidopsis*. Plant Physiol. 112, 1625–1630.

Meurer-Grimes, B., Berlin, J., Strack, D., 1989. Hydroxycinnamoyl-CoA:putrescine hydroxycinnamoyltransferase in tobacco cell cultures with high and low levels of caffeoylputrescine. Plant Physiol. 89, 488–492.

Milkowski, C., Strack, D., 2004. Serine carboxypeptidase-like acyltransferases. Phytochemistry 65, 517–524.

Milkowski, C., Baumert, A., Strack, D., 2000a. Cloning and heterologous expression of a rape cDNA encoding UDP-glucose:sinapate glucosyltransferase. Planta 211, 883–886.

Milkowski, C., Baumert, A., Strack, D., 2000b. Identification of four Arabidopsis genes encoding hydroxycinnamate glucosyltransferases. FEBS Lett. 485, 183–184.

Milkowski, C., Baumert, A., Schmidt, D., Nehlin, L., Strack, D., 2004. Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. Plant J. 38, 80–92.

Naczk, M., Amarowicz, R., Sullivan, A., Shahidi, F., 1998. Current research developments on polyphenolics of rapeseeds/canola: A review. Food Chem. 62, 489–502.

Negrel, J., Martin, C., 1984. The biosynthesis of feruloyltyramine in *Nicotiana tabacum*. Phytochemistry 23, 2797–2801.

Negrel, J., Javelle, F., Paynot, M., 1991. Separation of putrescine and spermidine hydroxycinnamoyl transferases extracted from tobacco callus. Phytochemistry 30, 1089–1092.

Nielsen, J.K., Olsen, C.E., Petersen, M.K., 1993. Acylated flavonol glycosides from cabbage leaves. Phytochemistry 34, 539–544.

Olssen, L.C., Veit, M., Weissenböck, G., Bornman, J.F., 1998. Differential flavonoid response to enhanced UV-B radiation in *Brassica napus*. Phytochemistry 49, 1021–1028.

Plumb, G.W., Price, K.R., Rhodes, M.J.C., Williamson, G., 1997. Antioxidant properties of the major polyphenolic compounds in broccoli. Free Rad. Res. 27, 429–435.

Price, K.R., Casuscelli, F., Colquhoun, I.J., Rhodes, M.J.C., 1997. Hydroxycinnamic acid esters from broccoli florets. Phytochemistry 45, 1683–1687.

Sakushima, A., Coskun, M., Tanker, M., Tanker, N., 1994. A sinapic acid ester from *Boreava orientalis*. Phytochemistry 35, 1481–1484.

Schenck, H.R., Röbbelen, G., 1982. Somatic hybrids by fusion of protoplasts from *B. oleracea* and *B. campestris*. Z. Pflanzenzücht. 89, 278–288.

Shahidi, F., Naczk, M., 1992. An overview of the phenolics of canola and rapeseed: Chemical, sensory and nutritional significance. J. Am. Oil Chem. Soc. 69, 917–924.

Shirley, A.M., McMichael, C.M., Chapple, C., 2001. The *sng2* mutant of Arabidopsis is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose:choline sinapoyltransferase. Plant J. 28, 83–94.

Steffens, J.C., 2000. Acyltransferases in protease's clothing. Plant Cell 12, 1253–1256.

Strack, D., Mock, H.P., 1993. Hydroxycinnamic acids and lignins. In: Dey, P.M., Harborne, J.B. (Eds.), Enzymes in Secondary Metabolism. Academic Press, London, pp. 45–97.

Strack, D., Tkotz, N., Klug, M., 1978. Phenylpropanoid metabolism in cotyledons of *Raphanus sativus* and the effect of competitive in vivo inhibition of L-phenylalanine ammonia-lyase (PAL) by hydroxylamine derivatives. Z. Pflanzenphysiol. 89, 343–353.

Strack, D., Dahlbender, B., Grotjahn, L., Wray, V., 1984. 1,2-Disinapoylglucose accumulated in cotyledons of dark-grown *Raphanus sativus* seedlings. Phytochemistry 23, 657–659.

U, N., 1935. Genomic analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. Jpn. J. Bot. 7, 389–452.

Molecular Breeding (2005) 16: 127–138
DOI 10.1007/s11032-005-6825-8
© Springer 2005

Reduction of sinapate ester content in transgenic oilseed rape (*Brassica napus*) by dsRNAi-based suppression of *BnSGT1* gene expression

Alexandra Hüsken[1], Alfred Baumert[2], Dieter Strack[2], Heiko C. Becker[1], Christian Möllers[1] and Carsten Milkowski[2,*]
[1]*Institute of Agronomy and Plant Breeding, Georg-August-University, Von-Siebold-Str. 8, D-37075 Göttingen, Germany;* [2]*Department of Secondary Metabolism, Leibniz Institute of Plant Biochemistry, Weinberg 3, D-06120 Halle (Saale), Germany; *Author for correspondence (e-mail: cmilkows@ipb-halle.de; phone: 49-345-5582-1533; fax: 49-345-5582-1509)*

Received 6 December 2004; accepted in revised form 2 May 2005

Key words: Brassica napus, dsRNAi, Seed protein, Sinapate glucosyltransferase suppression, Sinapine

Abstract

Seeds of oilseed rape (*Brassica napus*) accumulate high amounts of antinutritive sinapate esters (SE) with sinapoylcholine (sinapine) as major component, accompanied by sinapoylglucose. These phenolic compounds compromise the use of the protein-rich valuable seed meal. Hence, a substantial reduction of the SE content is considered essential for establishing rape as a protein crop. The present work focuses on the suppression of sinapine synthesis in rape. Therefore, rape (spring cultivar Drakkar) was transformed with a dsRNAi construct designed to silence seed-specifically the *BnSGT1* gene encoding UDP-glucose:sinapate glucosyltransferase (SGT1). This resulted in a substantial decrease of SE content in T2 seeds with a reduction reaching 61%. In T2 seeds a high and significant correlation between the contents of sinapoylglucose and all other sinapate esters has been observed. Among transgenic plants, no significant difference in other important agronomic traits, such as oil, protein, fatty acid and glucosinolate content in comparison to the control plants was observed. Maximal reduction of total SE content by 76% was observed in seeds of one homozygous T2 plant (T3 seeds) carrying the *BnSGT1* suppression cassette as a single copy insert. In conclusion, this study is an initial proof of principle that suppression of sinapoylglucose formation leads to a strong reduction of SE in rape seeds and is thus a promising approach in establishing rape, currently an important oil crop, as a protein crop as well.

Abbreviations: BnSGT1 – *B. napus* UDP-glucose:sinapate glucosyltransferase; dsRNAi – double-stranded RNA interference; SE – sinapate esters

Introduction

Oilseed rape (*Brassica napus*) is the most important oil plant in temperate regions of the world and ranks second amongst oilseed crops produced worldwide. Seeds of *B. napus* accumulate oil to about 40–50% of dry matter. Cultivars strongly reduced in erucic acid and glucosinolates (double-low oilseed rape; 00 quality) give one of the healthiest vegetable oils for human consumption

(Downey and Bell 1990). The meal remaining from oil extraction contains about 40% protein. The protein fraction is known for its well-balanced amino acid composition (Ohlson 1978) making rape seed protein a possible source for food-grade supplements, a growing market which is so far based mainly on soy. Unfortunately, the use of rapeseed in human nutrition is still thwarted by the presence of undesirable phenolic compounds. Due to their concentration that is about 30 times higher than in soybean, these compounds remain a principal antinutritive factor in rape seeds (Sozulski 1979; Ismail et al. 1981; Kozlowska et al. 1990; Shahidi and Naczk 1992).

The predominant phenolic compounds in rape seeds are sinapate esters (SE) with sinapoylcholine (sinapine) as the most common one (Kozlowska et al. 1990; Shahidi and Naczk 1992). They make up 1–2% of the seed dry matter (Bell 1993) and contribute to the bitter taste, astringency and dark colour of rapeseed products. Being oxidized during seed oil processing, SE may form complexes with proteins, thus lowering the digestibility of rapeseed meal (Kozlowska et al. 1990; Shahidi and Naczk 1992; Naczk et al. 1998). This makes the reduction of SE content a substantial requirement for establishing rape as a protein crop.

So far, systematic breeding programmes aimed at developing rape cultivars with low SE content have not been set up. Several studies on the genetic variability of SE content in seeds of *Brassica* ssp. have been carried out (Kerber and Buchloh 1980; Kozlowska et al. 1983; Krähling et al. 1990; Bouchereau et al. 1991; Matthäus 1997; Wang et al. 1998). Analysing 1361 samples of rape breeding lines, Velasco and Möllers (1998) reported on a range in SE content from 5 to 18 mg/g seeds. Zum Felde et al. (2003) found a variation between 3.5 and 13 mg/g seeds in 549 selected seed samples of genotypically divergent winter rape material. These results offer the possibility to reduce the SE content by conventional plant breeding.

On the other hand, since the metabolism of SE is well established, molecular breeding approaches have come into focus and may complement conventional efforts. In members of the Brassicaceae, sinapate is produced via the phenylalanine/hydroxycinnamate pathway (Figure 1). The first commited enzyme in SE biosynthesis is a glucosyltransferase (UDP-glucose:sinapate glucosyltransferase; SGT; EC 2.4.1.120) catalysing the formation of sinapoylglucose (1-*O*-sinapoyl-β-glucose). Sinapoylglucose serves as energy-rich sinapoyl donor in transacylation reactions leading to the synthesis of sinapine (sinapoylcholine) in seeds, catalysed by the enzyme sinapoylglucose:choline sinapoyltransferase (SCT; EC 2.3.1.91). The rape seed fraction of SE contains a range of additional yet unknown SE compounds in minor concentrations most likely also formed via sinapoylglucose. During seed germination sinapine is hydrolysed by a sinapine esterase (SCE; 3.1.1.49). The liberated sinapate is conjugated via sinapoylglucose with malate in the seedling. The resulting sinapoylmalate accumulates in vacuoles of the leaf epidermal cell layer serving as UV screen. So far, the biological role of sinapine accumulation during seed development has not been elucidated, although it has been shown that sinapine contributes to the supply of choline for synthesis of phosphatidylcholine in developing seedlings of *Raphanus sativus* (Strack 1981). In ongoing molecular work on SE in members of the Brassicaceae, we had previously reported the isolation of cDNAs encoding SGT (*BnSGT1*) and SCT (*BnSCT*) from rape (Milkowski et al. 2000, 2004).

The present work is aimed at reducing the SE content in rape seeds by suppression of sinapoylglucose formation, the key metabolite in the biosynthesis of SE. We describe here the regeneration of transgenic rape plants harbouring a dsRNAi construct for suppression of *BnSGT1* expression and changes in composition and concentration of SE in progeny up to the T4 generation.

Materials and methods

Plant material

Spring oilseed rape (*Brassica napus* L. cv. Drakkar) with '00' quality was used. Seeds were surface sterilised by soaking in 5% sodium hypochlorite solution for 30 min. After rinsing three times with sterile distilled water seeds were sown on 1/2 MS-medium (Murashige and Skoog 1962) and germinated in a growth chamber at 20 °C with a day length of 16 h. Plants were cultivated in the greenhouse under a 16 h light regiment with a photon flux density of 200–900 μmol/m^2/s.

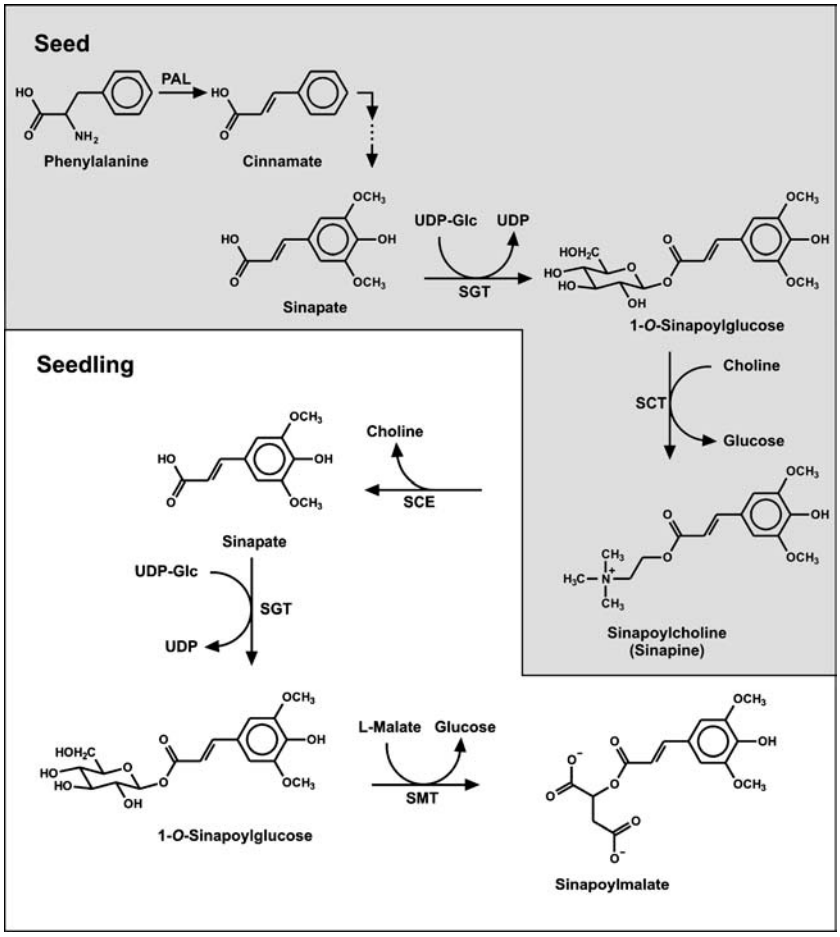


Figure 1. Biosynthetic pathway of the major sinapate esters in seeds and seedlings of rape (according to Milkowski et al. 2004). The enzymes involved are UDP-glucose (UDP-Glc):sinapate glucosyltransferase (SGT; EC 2.4.1.120), 1-*O*-sinapoyl-*β*-glucose:choline sinapoyltransferase (SCT; EC 2.3.1.91), sinapoylcholine esterase (SCE; EC 3.1.1.49), and 1-*O*-sinapoyl-*β*-glucose:malate sinapoyltransferase (SMT; EC 2.3.1.92).

Vector construction

To construct the dsRNAi suppression plasmid pLH-SGT-GUS, a sub-fragment of *BnSGT1*-cDNA (GenBank accession # AAF98390) was amplified by PCR using the following primer pairs:

Nco-SGT-fw (GCTCGGTACCCCATGGAAC-TATCATCTTCTCC) and
Sma-SGT-rev (GTACCCGGGGAGATAACC T TTACCGATAGG);
Nhe-SGT-fw (GTAGCTAGCATGGAACTAT-CATCTTCTCC) and
BH-SGT-rev(GTAGGATCCGAGATAACCTT TACCGATA GG).

The resulting PCR products cover the first 213 bp of the *BnSGT1* reading frame and are flanked by restriction sites for *Nco*I and *Sma*I or *Nhe*I and *Bam*HI allowing the cloning in antisense (as *Sma*I-*Nco*I-fragment) and sense orientation (as *Nhe*I-*Bam*HI fragment) resulting in a suppression cassette consisting of napin590 promoter, the 213 bp *BnSGT1* fragment in antisense orientation, a subfragment of the bacterial *gusA* gene (Chuang and Meyerowitz 2000) as spacer element, the *BnSGT1* fragment in sense orientation and the *nos* terminator (Figure 2). The suppression cassette was constructed in pBluescript from which it was

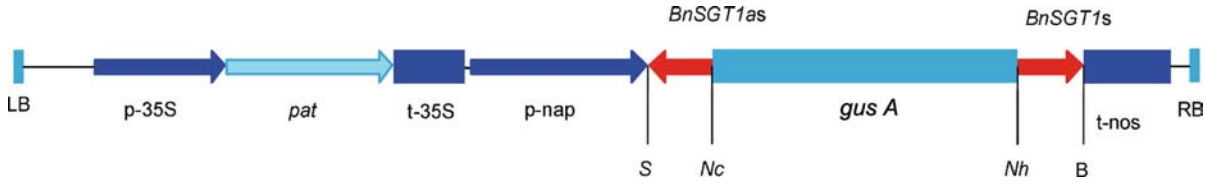


Figure 2. Schematic representation of the dsRNAi construct used in this study. The shown fragment forms the T-DNA part of plasmid pLH-SGT-GUS used for *A. tumefaciens* mediated transformation of rape. Sequences of functional importance are left border (LB) and right border (RB); CaMV 35S promoter (p-35S), *pat* gene encoding phosphinothricin acetyltransferase (*pat*), and CaMV 35S terminator (t-35S). Seed-specific suppression cassette consists of napin promoter (p-nap), a 213 bp subsequence of *BnSGT1* in antisense (as) and sense (s) orientation, the *gusA* gene fragment ('*GUS*) and *nos* transcription terminator (t-nos). Recognition sites are shown for restriction endonucleases *Sma*I (S), *Nco*I (Nc), *Nhe*I (Nh), *Bam*HI (B).

cut as *Spe*I-*Hin*dIII fragment and ligated with the similarly cleaved binary vector pLH7000 (Hausmann and Töpfer 1999) to give the *BnSGT1* suppression vector pLH-SGT-GUS. The construct carries the *pat* gene transcribed from the CaMV 35S promoter as marker for selection of transgenic plants by resistance against phosphinothricin (PPT). All vector construction was done with *E. coli* XL1blue as a host. *Agrobacterium tumefaciens* AGL1 (Hellens et al. 2000) carrying the binary plasmid pLH-SGT-GUS was used for the production of transgenic rape lines.

Rape transformation

Agrobacterium-mediated transformation and plant regeneration were done as described previously (De Block et al. 1989) with slight modification. Hypocotyl segments were inoculated with *A. tumefaciens* carrying the binary plasmid pLH-SGT-GUS. Regeneration of transformed Drakkar hypocotyl segments was achieved in two steps on selective medium with 5 and 10 mg/l PPT (Duchefa, Haarlem, Netherlands), respectively. Three to six weeks after selection, calli with small shoots were regenerated. Shoots were separated and transferred to MS-medium (Murashige and Skoog 1962) for further regeneration and rooting. Putative transgenic plants (T1 plants) propagated *in vitro* were transferred to the greenhouse. The main shoot of each plant was selfed for seed production (T2 seeds). Selected T2 plants were selfed to obtain T3 seeds and selected T3 plants were selfed to obtain T4 seeds.

PAT-ELISA-test

Leaf tissue (100 mg) of putative transgenic plants was used. The PAT-ELISA-test was performed as described by the manufacturer (Steffens, Eberingen, Germany). The absorbance of the reaction mixture was measured at room temperature at a wavelength of 625 nm. The standard curve was made following the manufacturer's instructions.

Selection for phosphinothricin resistance

Surface sterilized T2 seeds (20 seeds per T1 plant) were germinated on 1/2 MS-medium supplemented with 10 mg/l PPT in a growth chamber. The development of herbicide damage symptoms was scored up to 7 days post treatment. As a control, surface sterilised rape seeds (cv. Drakkar) were germinated *in vitro* and plants obtained were propagated.

Southern blot analysis

Genomic plant DNA was extracted from leaf tissue using a kit for plant DNA purification (Qiagen, Hilden, Germany). PCR amplification of the 498 bp *pat* gene fragment was carried out with the primer pair Pat-fw (5′-ATG GGC CCA GAA CGA CGC CC-3′); Pat-rev (5′-GCG TGA TCT CAG ATC TCG GT-3′). For Southern blot analysis, genomic DNA (5 μg) was digested with restriction endonuclease *Eco*RI, electrophoretically separated and transferred to Hybond N+ membrane (Amersham Biosciences, Freiburg, Germany) according to standard protocols (Sambrook et al. 1989). Hybridisation was performed with the *pat* gene as DIG-labelled probe. Labelling was done by PCR using the DIG-Hy-Prime-kit (Roche, Mannheim, Germany). For signal detection, the DIG Nucleic Acid Detection Kit (Roche, Mannheim, Germany) was used.

Analysis of SE

The two major SE from seeds, sinapoylglucose and sinapine, were identified by chromatographic (HPLC) and UV–Vis spectroscopic (DAD-HPLC) comparison with standard compounds. Seed material (20 mg) was extracted with 4 ml of a methanol-water mixture (4:1) in 2 ml-safe-lock tubes by vigorous shaking in the presence of zirconia beads (1 mm in diameter) using a bead beater (Bio Spec Products, Bartlesville, OK, USA). The resulting homogenates were cleared by centrifugation and aliquots of the supernatants transferred into HPLC autosampler vials. Reversed phase HPLC (Waters Separator 2795, Waters 2996 photodiode array detector) was carried out using a 5-μm Nucleosil C_{18} column (250 × 4 mm i.d.; Macherey-Nagel, Düren, Germany). A 20-min linear gradient was applied at a flow rate of 1 ml/min from 10 to 50% solvent B (acetonitrile) in solvent A (1.5% *o*-phosphoric acid in water). SE were photometrically detected at 330 nm and quantified by external standardisation with authentic compounds. The total SE content was calculated as sinapate equivalents.

Analysis of seed quality traits

Seed samples were analysed for oil, protein, glucosinolates and fatty acids by near-infrared reflectance spectroscopy (NIRS) using the *Raps*2001.eqa. (www.vdlufa.de/nirs).

Statistical analysis

All statistical parameters (mean, SD, χ^2) were calculated using the StatGraphics Plus for Windows 3.0 (Statistical Graphics Corp. 1997). For correlation analysis Spearman rank correlation coefficients were used.

Results

Generation of transgenic rape lines

Fifty one PPT-resistant (10 mg/l) Drakkar lines were obtained from 850 explants (Table 1). This number corresponds to a mean regeneration efficiency of 6%. PAT-ELISA test and *pat*-PCR analysis of leaves from putative transgenic plants confirmed the presence of the *pat* marker gene. Thirty of the regenerants were tested positive in both the PAT-ELISA (data not shown) and *pat*-specific PCR (Figure 3). The mean transformation efficiency was 3.6% and varied between the three experiments from 2.8 to 4.3% (Table 1).

Agronomic traits of T1 plants

Thirty transgenic T1 plants (BnSGTi plants) were transferred to the greenhouse. All of the transgenic plants were fully fertile and normal in growth and morphology when compared to untransformed control plants (data not shown). There was no significant difference in other important seed-specific agronomic traits, such as oil content, protein, fatty acids and glucosinolates as measured in T2 seeds (Table 2).

Sinapate ester content in T2 seeds

T2 seeds obtained from all 30 T1 plants were analysed by HPLC for SE accumulation. Seed

Table 1. Transformation of rape with pLH-SGT-GUS: regeneration and transformation efficiency.

Experiment	explants	5 mg/l PPT	10 mg/l PPT	Regeneration efficiency[a] (%)	ELISA[b] and PCR[b] positive	Transformation efficiency[a] (%)
1	250	93	10	4.0	7	2.8
2	300	106	21	7.0	13	4.3
3	300	78	20	6.6	10	3.3
Sum/mean	850	277	51	6.0	30	3.6

[a]Regeneration efficiency is calculated as the percentage of plants growing in the presence of 10 mg/l PPT. Transformation efficiency is expressed as percentage of positive plants in PAT-ELISA and *pat*-PCR assay.
[b]PAT-ELISA-test and *pat*-PCR.

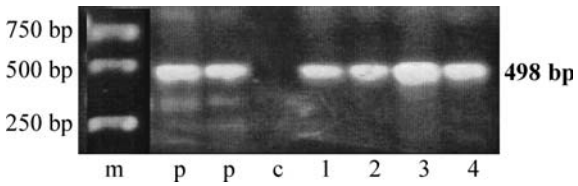


Figure 3. PCR analysis of BnSGTi plants (T1 plants). Gel electrophoresis of the amplified *pat* gene fragment with the expected length of 498 bp. Lanes: m = marker; p = plasmid (pLH-SGT-GUS); c = untransformed control; 1–4 = 1501.6, 1501.10, 1501.24, 1501.26 (T1 plants).

samples showed varying levels for the total SE content (Figure 4). Average concentration of total SE in T2 seeds 7.30 ± 2.44 mg/g. The lowest T2 seed sample contained 2.84 mg/g total SE.

Among the transgenic T2 seeds, a high and significant correlation between sinapoylglucose and all other SE became evident (Table 3; Figure 5), which ranged from 0.76** to 0.92**. Furthermore, contents of sinapine, the sum of non identified SE and the total SE were highly correlated to each other.

Table 4 shows the sinapate ester composition of the three T2 seed samples displaying the lowest SE content. In sample 1501.24, a complete loss of sinapoylglucose was observed whereas the minimum value of the untransformed control samples accounted for 1.21 mg/g. The strongly reduced contents of sinapine (4.02 mg/g) and total reduction of other SE reflected the substantial suppression of sinapate ester biosynthesis in transgenic line 1501.24.

Transgene copy number in low SE lines

The three transgenic plants with the lowest SE content as determined in T2 seeds (lines 1501.6, 1501.24, 1501.26) were characterised for the segregation of PPT-resistance in T2 seedlings. For each line it was found, that 4–5 out of 20 seedlings tested were not resistant to PPTC Table 5). This indicates a 3:1 inheritance of PPT-resistance suggesting that each of the three transformants contains a single T-DNA locus. These results correspond to those determined by genomic Southern blot analysis of T1 plants with the *pat* gene fragment as probe (Table 5).

Twenty T2 plants from line 1501.6 were grouped according to their sinapine and total SE content as measured in T3 seeds. PPT-resistance levels of the

Table 2. T2 seed quality traits (mean ± SD; $n = 30$) of BnSGTi and control plants.

	controls $n = 30$	1501 BnSGTi $n = 30$		Controls $n = 30$	1501 BnSGTi $n = 30$
Oil (%)	44.89 ± 2.37	44.73 ± 2.71	C18:1 (%)	68.53 ± 2.31	68.25 ± 2.81
Protein (%)	24.71 ± 1.97	24.30 ± 2.42	C18:3 (%)	6.70 ± 0.77	6.50 ± 0.72
GSL (μmol/g)	22.35 ± 2.13	21.22 ± 3.54	C22:1 (%)	0.03 ± 0.07	0.00 ± 0.00

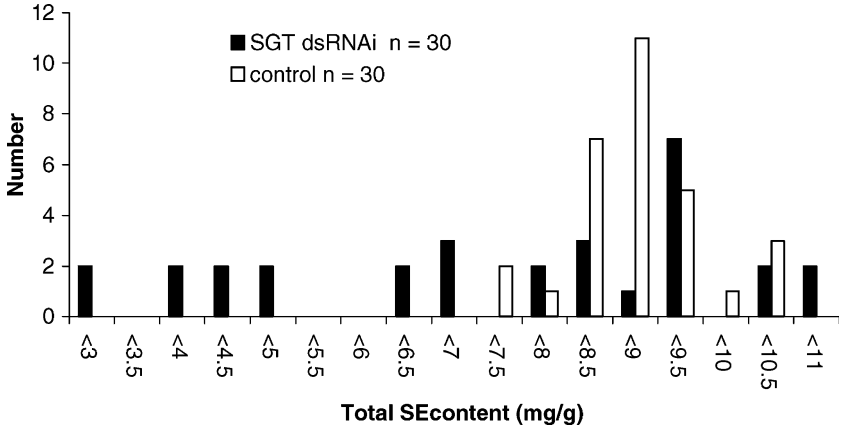


Figure 4. Frequency distribution for the total SE content (mg/g) in T2 seeds of BnSGTi plants and control plants.

Table 3. Spearman rank correlations (r_s; $n = 30$) between sinapate esters in T2 seeds of BnSGTi plants.

	Sinapoyl glucose	Sinapine	Sum of not fully characterised SE[a]
Sinapine	0.76**		
Sum of not fully characterised SE[a]	0.92**	0.76**	
Total SE content[a]	0.91**	0.96**	0.91**

[a]calculated as sinapate equivalents.

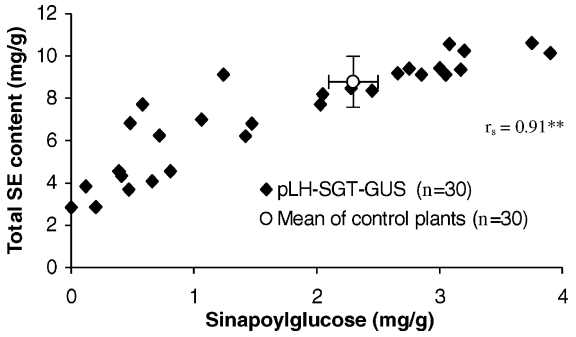


Figure 5. Correlation between sinapoylglucose (mg/g) and total sinapate ester content (mg/g) in T2 seeds ($n = 30$) of BnSGTi plants. For comparison the mean of the untransformed controls (±SD) is also shown.

corresponding T3 seedlings were determined to discriminate between homozygous resistant, hemizygous resistant and homozygous not resistant descendants (data not shown) and correlated to sinapine and total SE contents. Results indicated that the homozygous resistant ones had the lowest sinapine and total SE contents, the hemizygous resistant ones had a medium and the homozygous not-resistant ones had the highest sinapine and total SE content (Figure 6). These groups (low, medium, high sinapine and total SE content, respectively) showed the 1:2:1 segregation as is expected for a single copy insertion of the transgene. Analyses of lines 1501.24 and 1501.26 confirmed the highest sinapine and total SE content for homozygous not-resistants. Hemizygous and homozygous resistant descendants showed severely reduced SE contents but there was no clear evidence for a significant difference between them (data not shown).

Sinapate ester content in the offspring of selected transgenic lines

From each of the three transgenic lines with the lowest SE content and single copy integration (1501.6, 1501.24, 1501.26) 20 T2 seeds were sown together with control seeds ($n = 22$) in the greenhouse. All T2 plants were normal during the vegetative and reproductive growth phases. Mature T3 seeds were harvested and subjected to SE quantification (Table 6). The total SE content in control seeds averaged at 8.40 ± 0.57 mg/g. This value was found to be significantly reduced in transgenic T3 seeds. For line 1501.6 the average concentration of total SE constituted 4.50 ± 1.55 mg/g, line 1501.24 contained 5.42 ± 2.31 mg/g and line 1501.26 accumulated SE to 5.01 ± 1.86 mg/g. The minimum value for total SE content in T3 seeds was detected in line 1501.24.18 with 1.82 mg/g giving rise to a reduction of 76% compared to the lowest control plant value. This particular decrease was accompanied by a marked reduction of sinapine content to 2.66 mg/g corresponding to a loss of 72% from the control level.

The severely suppressed line 1501.24.18 was chosen columns proceeding to T4 generation. T3 seeds ($n = 20$) were sown together with control plants ($n = 13$) in the greenhouse. As with the preceding transgenic generation, no deviations were

Table 4. Total amount (mg/g) of sinapate esters in selected T2 seed samples of BnSGTi and control plants.

Line	Sinapoyl glucose	Sinapine	Sum of non identified SE[a]	Total SE content[a]
Control[b]	1.21	7.44	0.72	7.28
1501.6	0.12	5.30	0.00	3.83
1501.24	0.00	4.02	0.00	2.84
1501.26	0.20	3.71	0.07	2.86

[a]calculated as sinapate equivalents (mg/g).
[b]control plant with the lowest total SE content (from $n = 30$).

Table 5. PPT resistance of T2 seedlings and copy number as determined by Southern blot.

Line	T2 seedlings		Segregation (χ^2)[b]	Copy number[c]
	r[a]	nr[a]		
1501.6	16	4	3:1 (0.26)	1
1501.24	16	4	3:1 (0.26)	1
1501.26	15	5	3:1 (0.00)	1

[a]r/nr: resistant/not resistant (10 mg/l PPT).
[b](χ^2) = 3.84 (5%, 1 DF).
[c]Southern blot.

observable during vegetative and reproductive growth phases of T3 plants in the greenhouse. Mature T4 seeds were collected and analysed for phenolic compounds. In seeds of the control plants, a concentration of 8.65 ± 0.60 mg/g was detected as average value of total SE content (Table 7). The average concentration of total SE in transgenic T4 seeds was 2.68 ± 0.37 mg/g. In one T4 seed sample of line 1501.24.18 the total SE content was reduced to 1.84 mg/g (Table 7). Compared with the lowest control plant this accounts for a reduction of about 75%. For sinapine, a content of 2.68 mg/g was measured representing a reduction by 73% relative to the control (Table 7).

Analysing the PPT-resistance of T4 seedlings (n = 20) from this line, it was observed that the offspring of each T3 plant (n = 20) was completely resistant to PPT, indicating that each T3 plant of line 1501.24.18 is homozygous for the inserted T-DNA (data not shown).

Discussion

Studies with transgenic rape plants revealed that suppression of *BnSGT1* expression is a powerful method for the reduction of SE content in seeds. By introducing a transgene expressing a dsRNA that corresponds to a 213 bp fragment of the 5′-translated region of the *B. napus SGT1* mRNA, we demonstrated the depletion of sinapoylglucose in seeds and a severe reduction of the sinapine content. Our analyses indicated a stable dsRNAi-mediated suppression of SE biosynthesis in the transgenic offspring up to the T4 generation.

Table 6. Absolute (mg/g) and relative (%) sinapine and total SE content in selected T3 seeds of BnSGTi and control plants.

Line	Sinapine	%	Total SE[a] content	%
Control[b]	9.46	100	7.48	100
1501.6.20	3.27	34.6	2.28	30.5
1501.24.18	2.66	28.1	1.82	24.3
1501.26.12	3.33	35.2	2.24	29.9

The differences between control plants and transgenic lines were all significant for $p < 0{,}01$ (students t-test).
[a]Calculated as sinapate equivalents.
[b]Control plant with the lowest total SE content (from n = 22).

Table 7. Total SE (mg/g) and sinapine content (mg/g) in selected T4 seeds of BnSGTi and control plants.

SE	n	mean	SD	min	max
Total					
Control	13	8.65	0.60	7.43	9.93
1501.24.18	20	2.68	0.37	1.84	3.27
Sinapine					
Control	13	8.73	0.60	7.37	9.90
1501.24.18	20	3.92	0.54	2.68	4.77

[a]Calculated as simapate equivalents.

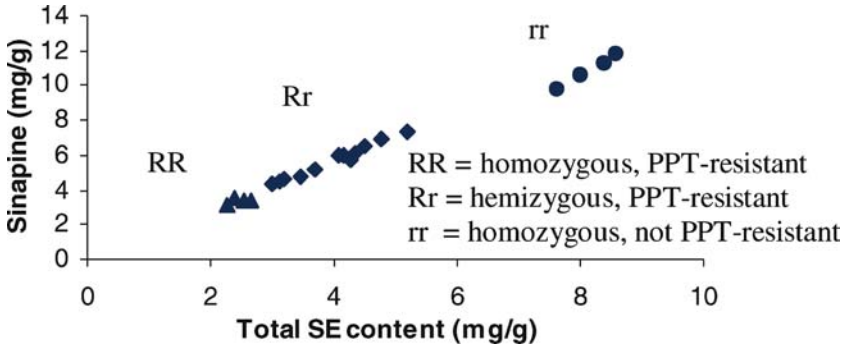


Figure 6. Sinapine (mg/g) and total SE (mg/g) content of T3 seeds (line 1501.6, n = 20).

Efficiency of suppression

The efficiency of the dsRNAi approach calculated as recovery rate of silenced plants in the primary transgenic generation was less than 50%. T2 seeds of only 13 out of 30 T1 plants had an SE content below the levels found in seeds of control plants (Figure 4). This is below the values reported for similar hairpin constructs with the GUS spacer in Arabidopsis which are in the range of 60–90% (Chuang and Meyerowitz 2000; Smith et al. 2000). On the other hand, Liu et al. (2002) found recovery rates of 53 and 62% when silencing fatty acid desaturase genes in seeds of cotton which harbours a more complex allotetraploid genome like *B. napus*. In general, higher silencing frequencies are reached with hairpin constructs carrying an intron in splicing orientation (Smith et al. 2000; Stoutjesdijk et al. 2002). Nevertheless, the maximum strength of suppression is not different when comparing intron containing to conventional hairpin constructs (Stoutjesdijk et al. 2002).

Strength of suppression

Among the rape lines carrying the dsRNAi suppression cassette, one homozygous plant with a single insertion was found to contain only about 24% of SE compared to the untransformed control. The levels of sinapoylglucose and other so far unknown SE were below the detection limit. Regarding sinapine, a strong reduction down to 28% of the level in untransformed control plants was observed. In line 1501.6 the suppression strength seemed to be slightly stronger in homozygous transgenic descendants than in hemizygous ones (Figure 6). Interestingly, a similar result has been described for Arabidopsis transformants carrying a dsRNAi construct for silencing the *FAD2* gene (Stoutjesdijk et al. 2002). However, a similar approach in cotton could not reproduce this effect (Liu et al. 2002). Our results with *B. napus* transgenic lines 1501.24 and 1501.26 displaying no significant differences in SE content between hemizygous and homozygous descendants are in accordance with the bulk of literature data indicating that a single dsRNAi-inducing transgene can mediate a dominant silencing even in polyploids like *Arabidopsis suecica* (Lawrence and Pikaard 2003) or cotton (Liu et al. 2002).

Since no complete reduction of SE was found in the present study, questions remain on how the residual SE biosynthesis is performed in the transgenic. The existence of gene families and the plants corresponding redundancy in gene functions especially regarding plants with large and complex genomes is considered a major obstacle for efficient suppression strategies. Thus, even for Arabidopsis, three additional genes (At4g15480, At4g15490, At4g15500) have been shown to encode glucosyltransferases catalysing the formation of sinapoylglucose *in vitro* (Milkowski et al. 2000; Lim et al. 2001). Sequence identities of about 68% with *AtSGT* probably results in a sufficient mismatch for these genes to escape silencing imposed by an *AtSGT*-specific approach. These genes share sequence identities between 64 and 78% among each other without longer stretches of high identity making it difficult to design a dsRNAi construct for an effective simultaneous suppression. However, as first expression studies in our laboratory failed to detect the transcripts in seeds, a role of these genes in seed-specific sinapine synthesis seems unlikely – at least in the wild type genetic background. In the allotetraploid rape genome, Southern blot analyses indicated the presence of multiple genes related to *BnSGT1* (Milkowski et al. 2004). We have set up homology-based cloning experiments to isolate these *BnSGT1*-related genes. Once isolated, expression analysis will reveal their possible involvement in seed-specific biosynthesis of sinapoylglucose in wild-type and *BnSGT1*-suppressed genetic backgrounds. If this would be the case, a further reduction of seed SE content might be expected by silencing these genes. On the other hand, SE accumulation in seeds could also be fed by vegetative tissues. Transgenic rape plants harbouring antisense constructs to silence the *BnF5H* gene have been reported to suppress sinapine accumulation more efficiently when expressed under control of the constitutive CaMV-35S promoter than under control of the seed-specific napin promoter (Nair et al. 1999). Regarding *BnSGT1*, data on dsRNAi suppression driven by CaMV-35S promoter or the native *BnSGT1* promoter would be helpful to assess a possible influence of maternal tissues. Respective experiments have been started recently in our laboratories.

On the other hand, the activities of the *BnSGT1* and napin promoters may not overlap completely especially during later stages of seed development (Ezcurra 1998; Milkowski et al. 2004). This, however, does not explain residual *BnSGT1* transcript levels given the molecular mechanism of RNAi-mediated silencing which includes amplification and should lead to a continuing silencing effect after RNAi has been triggered early in seed development (Hannon 2002) .

In general, the extent of suppression we observed for sinapine biosynthesis in rape is not unusual. Liu et al. (2002) reported a reduction at metabolite level by about 50% when silencing $\Delta 9$ desaturase in cotton seed. The authors argue that residual enzymatic activity might come from related genes and that a stronger suppression might not be observed because of lethal effects. In rape, a significantly higher suppression of sinapine formation has not been reported, so far. Values of about 20% residual sinapine were reached by diverting choline, a precursor for sinapine synthase (G. Selveraj, personal communication). This raises the interesting question whether a stronger suppression of the sinapine pathway can be tolerated by *B. napus*.

Experiments aimed at the molecular characterisation of plants silenced for *BnSGT1* have been started recently in our laboratory. They will include quantification of *BnSGT1* mRNA and enzyme activity during seed development to seek out whether the residual SE synthesis in these lines is due to remaining *BnSGT1* expression indicating incomplete silencing of the gene. This will help to develop more efficient strategies for suppressing SE biosynthesis in seeds of *B. napus*.

Implications for rapeseed quality improvement

Rapeseed is a significant source of protein supplement in animal nutrition and has the potential for being used in the human food industry. In the present work it was shown that it is possible to achieve a significant reduction of up to 76% in the content of antinutritive SE by suppressing the *BnSGT1* gene expression. Given their antinutritive character it is generally intended to keep the SE content in rape seed meal as low as possible. Hydrothermic treatment could reduce SE levels by up to 95% but it is not economical in bulk processing. Thus, a reduction by about 80% in seed content of SE is considered sufficient for the use of rapeseed meal in animal nutrition – a level that could be reached by suppressing *BnSGT1* expression. There was no indication that other important agronomic traits, like oil, protein, fatty acids and glucosinolates contents of the seeds are affected by decreasing sinapate ester accumulation. A study on segregating doubled haploid winter rapeseed populations confirmed this conclusion by failing to detect significant correlation between contents of SE, oil, protein, fatty acids and glucosinolates (T. zum Felde, pers. communication). Therefore, a drastic suppression of SE accumulation should be a powerful approach to increase protein quality in breeding material of rape.

To be accepted as human food supplement the SE content is required to be below 1 mg/g.

Approaches for further suppression of SE accumulation in rape seeds

So far, only one other gene involved in sinapate metabolism has been used as a molecular target to accomplish a decrease in SE accumulation in rape seeds. Based on sequence information from Arabidopsis (Meyer et al. 1996), three cDNAs were cloned from *B. napus* homologous to ferulate-5-hydroxylase (F5H), the enzyme catalysing the hydroxylation of ferulate to 5-hydroxyferulate (Nair et al. 2000). Recently, it has been shown *in vitro* that the Arabidopsis F5H prefers coniferaldehyde as a substrate (Humphreys et al. 1999). Nevertheless, antisense constructs of the *BnF5H* gene under control of a CaMV-35S promoter were shown to reduce the sinapine content by about 40%. Based on these findings, a crossing of plants suppressing *BnSGT1* and *F5H* appears promising for further reduction of the SE content in seeds. On the other hand, TILLING populations of *B. napus* could be screened for mutations in *BnSGT1*. This would require preliminary information on seed-specific expression of both the *B. oleracea*- and *B. rapa*-derived genes. Since it has been shown that *BnSGT1* is expressed during seedling development and to a lesser extent in adult plants, a possible damaging effect caused by constitutive suppression of this gene has to be examined. An alternative approach to further reduce the SE

content in seeds makes use of the natural variation in rape lines detected by several studies. These strategies should involve transformation of classical low sinapine lines with suppression constructs as well as crossing between classical and low sinapine transgenic material.

Acknowledgements

We thank Rosemarie Clemens and Nicole Ritgen-Homoyounfar for technical assistance and Dr Jose Orsini (Saatenunion Resistenzlabor, Germany) for efficient transformation protocols. This work is part of the research project 'NAPUS 2000-Healthy Food from Transgenic Rape Seeds', financially supported by the Bundesministerium für Bildung und Forschung.

References

Bell J.M. 1993. Factors affecting the nutritional value of canola meal: a review. Can. J. Anim. Sci. 73: 679–697.

Bouchereau A., Hamelin J., Lamour I., Renard M. and Larher F. 1991. Distribution of sinapine and related compounds in seeds of *Brassica* and allied Genera. Phytochemistry 30: 1873–1881.

Chuang C-F. and Meyerowitz E.M. 2000. Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. PNAS 97: 4985–4990.

De Block M., De Brouwer D. and Tenning P. 1989. Transformation of *Brassica napus* and *Brassica oleracea* using *Agrobacterium tumefaciens* and the Expression of the *bar* and *neo* Genes in the Transgenic Plants. Plant Physiol. 91: 694–701.

Downey R.K. and Bell J.M. 1990. New development in canola research. In: Shahidi F. (ed.), Canola and Rapeseed. Production, Chemistry, Nutrition and Processing Technology. Van Nostrand Reinhold, New York, pp. 211–220.

Ezcurra I. 1998. Studies on the Regulation of the Napin *napA* Promoter by Abscisic Acid and the Transclptional Activator ABI3. Doctoral thesis, Swedish University of Agricultural Sciences, Agraria 102.

Hannon G.J. 2002. RNA interference. Nature 418: 244–251.

Hausmann L. and Töpfer R. 1999. Entwicklung von Plasmid-Vektoren. BioEngineering für Rapssorten nach Maß. Vorträge für Pflanzenzüchtung 45: 155–173.

Hellens R.P., Mullineaux P. and Klee H. 2000. A guide to *Agrobacterium* binary Ti vectors. Trends Plant Sci. 5: 446–451.

Humphreys J.M., Hemm M.R. and Chapple C. 1999. New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. PNAS 96: 10045–10050.

Ismail F., Vaisey-Genser M. and Fyfe B. 1981. Bitterness and astringency of sinapine and its components. J. Food Sci. 46: 1241–1244.

Kerber E. and Buchloh G. 1980. Der Sinapingehalt in Cruciferensamen. Angew. Bot. 54: 47–54.

Kozlowska H., Naczk M., Shahidi and Zadernowski R. 1990. Phenolic acids and tannins in rapeseed and canola. In: Shahidi F. (ed.), Canola and Rapeseed. Production, Chemistry, Nutrition and Processing Technology. Van Nostrand Reinhold, New York, pp. 193–210.

Kozlowska H., Rotkiewicz D.A., Zadernowski R., Sosulski F.W. 1983. Phenolic acids in rapeseed and mustard. J. Am. Oil Chem. Soc. 60: 119–123.

Krähling K., Röbbelen G. and Thies W. 1990. Genetic variation of the content of sinapoyl esters in seeds of rape, *B. napus*. Plant Breeding 106: 254–257.

Lawrence R.J. and Pikaard C.S. 2003. Transgene-induced RNA interference: a strategy for overcoming gene redundancy in polyploids to generate loss-of-function mutations. Plant J. 36: 114–121.

Lim E.K., Parr A., Jackson D., Ashford D.A. and Bowles D.J. 2001. Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in *Arabidopsis*. J. Biol. Chem. 276: 4344–4349.

Liu Q., Singh P.S. and Green A.G. 2002. High-stearic and high-oleic cottonseed oils produced by Hairpin RNA-mediated post-transcriptional gene-silencing. Plant Physiol. 129: 1732–1743.

Matthäus B. 1997. Antinutritive compounds in different oilseeds. Fett/Lipid 99: 170–174.

Meyer K., Cusumano J.C., Somerville C. and Chapple C.C.S. 1996. Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monooxygenases. PNAS 93: 6869–6874.

Milkowski C., Baumert A. and Strack D. 2000. Cloning and heterologous expression of a rape cDNA encoding UDP-glucose: sinapate glucosyltransferase. Planta 211: 883–886.

Milkowski C., Baumert A., Schmidt D., Nehlin L. and Strack D. 2004. Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. Plant J. 38: 80–92.

Murashige T. and Skoog F. 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant 15: 473–497.

Naczk M., Aramowicz A., Sullivan A. and Shahidi F. 1998. Current research developments on polyphenolics of rapeseeds/canola: a review. Food Chem. 62: 489–502.

Nair R.B., Joy R.W., Schnaider J., Shi X., Datla R.S.S., Keller W.A. and Selvaraj G. 1999. Metabolic Engineering of the Sinapine content of *Brassica napus* seeds. In: Rapeseed Congress (CD-Rom). Canberra.Australia.

Nair R.B., Joy R.W., Schnaider J., Kurylo E., Shi X., Datla R.S.S., Keller W.A. and Selvaraj G. 2000. Identification of a CYP84 family of cytochrome P450- dependent mono-oxygenase genes in *Brassica napus* and perturbation of their expression for engineering sinapine reduction in the seeds. Plant Physiol. 123: 1623–1634.

Ohlson R. 1978. Functional properties of rapeseed oil and protein product. In: Proceedings 5th International Rapeseed Congress. Malmö, Sweden, pp. 152–167.

Sambrook J., Fritsch E.F. and Maniatis T. 1989. Molecular Cloning A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Shahidi F. and Naczk M. 1992. An overview of the phenolics of canola and rapeseed: chemical, sensory and nutritional significance. J. Am. Oil Chem. Soc. 69: 917–924.

Smith N.A., Sigh S.P., Wang M.-B., Stoutjesdijk P.Green A., Waterhouse P.M. 2000. Total silencing by intron-spliced hairpin RNAs. Nature 407: 319–320.

Sozulski F. 1979. Organoleptic and nutritional effects of phenolic compounds on oilseed protein products: a review. JAOCS 56: 711–715.

Statistical Graphics Corp. 1997. Statgraphics Plus 3.0, Statistical graphic System by Statistical Graphic Corporation. Rockville.

Stoutjesdijk P., Surinder P.S., Liu Q., Hurlstone C.J., Waterhouse P.A. and Green A.G. 2002. hpRNA-mediated targeting of the *Arabidopsis* FAD2 gene gives highly efficient and stable silencing. Plant Physiol. 129: 1–9.

Strack D. 1983. Sinapine as a supply of choline for the biosynthesis of phosphatidylcholine in *Raphanus sativus* seedlings. Z. Naturforsch. 36c: 215–221.

Velasco L. and Möllers C. 1998. Nondestructive assessment of sinapic acid esters in *Brassica* species: II. Evaluation of germplasm and identification of phenotypes with reduced levels. Crop Sci. 38: 1650–1654.

Wang S., Oomah B.D., McGregor D.I. and Downey R.K. 1998. Genetic and seasonal variation in the sinapine content of seed from *Brassica* and *Sinapis* species. Can. J. Plant Sci. 78: 395–400.

Zum Felde T., Baumert A., Becker H.C. and Möllers C. 2003. Genetic variation, inheritance and development of NIRS-calibrations for sinapic acid esters in oilseed rape (*Brassica napus* L.). In: Proceedings 11th International Rapeseed Congress. Copenhagen.

Theor Appl Genet (2005) 111: 1553–1562
DOI 10.1007/s00122-005-0085-1

ORIGINAL PAPER

**Alexandra Hüsken · Alfred Baumert
Carsten Milkowski · Heiko C. Becker
Dieter Strack · Christian Möllers**

Resveratrol glucoside (Piceid) synthesis in seeds of transgenic oilseed rape (*Brassica napus L.*)

Received: 15 March 2005 / Accepted: 11 August 2005 / Published online: 14 September 2005
© Springer-Verlag 2005

Abstract Resveratrol is a phytoalexin produced in various plants like wine, peanut or pine in response to fungal infection or UV irradiation, but it is absent in members of the *Brassicaceae*. Moreover, resveratrol and its glucoside (piceid) are considered to have beneficial effects on human health, known to reduce heart disease, arteriosclerosis and cancer mortality. Therefore, the introduction of the gene encoding stilbene synthase for resveratrol production in rapeseed is a tempting approach to improve the quality of rapeseed products. The stilbene synthase gene isolated from grapevine (*Vitis vinifera L.*) was cloned under control of the seed-specific napin promotor and introduced into rapeseed (*Brassica napus L.*) by *Agrobacterium*-mediated co-transformation together with a ds-RNA-interference construct deduced from the sequence of the key enzyme for sinapate ester biosynthesis, UDP-glucose:sinapate glucosyltransferase (*BnSGT1*), assuming that the suppression of the sinapate ester biosynthesis may increase the resveratrol production in seeds through the increased availability of the precursor 4-coumarate. Resveratrol glucoside (piceid) was produced at levels up to 361 μg/g in the seeds of the primary transformants. This value exceeded by far piceid amounts reported from *B. napus* expressing *VST1* in the wild type sinapine background. There was no significant difference in other important agronomic traits, like oil, protein, fatty acid and glucosinolate content in comparison to the control plants. In the third seed generation, up to 616 μg/g piceid was found in the seeds of a homozygous T3-plant with a single transgene copy integrated. The sinapate ester content in this homozygous T3-plant was reduced from 7.43 to 2.40 mg/g. These results demonstrate how the creation of a novel metabolic sink could divert the synthesis towards the production of piceid rather than sinapate ester, thereby increasing the value of oilseed products.

Communicated by G. Wenzel

A. Hüsken · H. C. Becker · C. Möllers (✉)
Institute of Agronomy and Plant Breeding,
Georg-August-University, Von-Siebold-Str. 8,
37075 Gottingen, Germany
E-mail: cmoelle2@gwdg.de

A. Baumert · C. Milkowski · D. Strack
Department of Secondary Metabolism,
Leibniz Institute of Plant Biochemistry, Weinberg 3,
06120 Halle, Saale, Germany

Introduction

Resveratrol (*trans*-3,5,4′-trihydroxystilbene) is a phytoalexin, derived from the phenylpropanoid pathway. Several plant species like grape (Langcake and Pryce 1976), peanut and pine (Kindl 1985) synthesize resveratrol that is absent in members of the *Brassicaceae*. A correlation between fungal disease resistance and resveratrol production of grapevine cultivars has been reported (Langcake and McCarthy 1979; Stein and Blaich 1985). Resveratrol can also be induced by other stress factors like UV light (Schoeppner and Kindl 1978), wounding (Langcake 1981) and elicitor treatment (Melchior and Kindl 1991). The biosynthesis of resveratrol is catalysed by the enzyme stilbene synthase, which converts one molecule of 4-coumaroyl-CoA and three molecules of malonyl-CoA into resveratrol. As part of the phenylpropanoid pathway, these precursor molecules are present in all plant species as substrate for chalcone synthase, the key enzyme of the flavonoid pathway (Rupprich and Kindl 1978).

The gene encoding stilbene synthase (VST1) in grapevine (*Vitis vinifera*) has been expressed in various plant species (e.g. Hain et al. 1993; Stark-Lorenzen et al. 1997; Szankowski et al. 2003). These studies were mostly initiated to increase fungal disease resistance in crop species. A relationship between resveratrol production and enhanced fungal resistance has been shown (Hain et al. 1993; Leckband and Lörz 1998; Hipskind and Paiva 2000; Liang et al. 2000). Contrasting results were obtained in kiwi, where the production of resveratrol did not lead to an increased disease resistance (Kobayashi

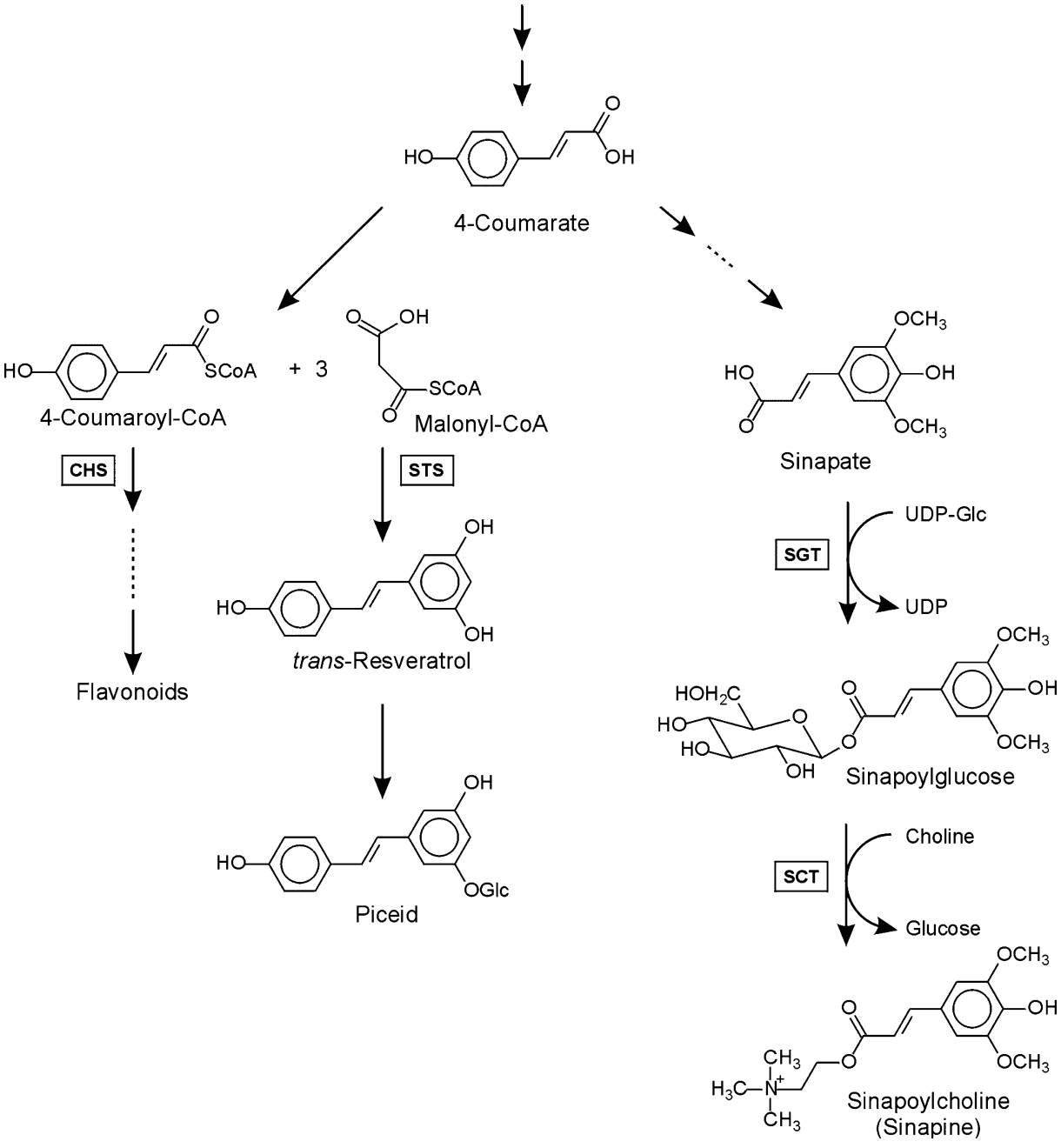


Fig. 1 Resveratrol and sinapate ester biosynthetic pathway in oilseed. The enzymes involved in sinapate ester metabolism are UDP-glucose (*UDP-Glc*): sinapate glucosyltransferase (*SGT*; EC 2.4.1.120), 1-*O*-sinapoyl-*β*-glucose: choline sinapoyl-transferase (*SCT*; EC 2.3.1.91). Resveratrol synthesis is catalysed by stilbene synthase (*STS*); the first step in flavonoid biosynthesis is catalysed by chalcone synthase (*CHS*)

et al. 2000). However, there are alternative reasons for producing resveratrol in transgenic plants, such as using this compound as nutraceutical or beneficial food component. Resveratrol and its glucoside (piceid) are known for their positive effects on human health. It has been shown that they reduce coronary heart disease mortality and arteriosclerosis (Manna et al. 2000), inhibit low-density lipoprotein oxidation (Manna et al. 2000) and eicosanoid synthesis (Pace-Asciak et al. 1995). An anti-cancer effect has also been reported (Jang et al. 1997).

In plant phenylpropanoid metabolism, 4-coumarate forms an important branching point, feeding biosyntheses of stilbenes, flavonoids and sinapate esters (Fig. 1). Seeds of *Brassica napus* accumulate mainly sinapoylcholine (sinapine) and sinapoylglucose accompanied by smaller amounts of other, not yet characterized, sinapate conjugates (Strack et al. 1983; Boucherau et al. 1991; Lorenzen et al. 1996). The total amount of sinapate esters reaches 1–2% of seed dry matter (Bell 1993). As energy-rich acyl donor, sinapoylglucose, which is

formed from UDP-glucose and sinapate by the enzyme UDP-glucose:sinapate glucosyltransferase (SGT; EC 2.4.1.120), plays a key role in sinapate ester metabolism (Fig. 1). In developing seeds, this glucose ester is transacylated by the enzyme sinapoylglucose:choline sinapoyltransferase (SCT, EC 2.3.1.91) to sinapine. During germination, sinapine is hydrolysed by a specific esterase (sinapine esterase, SCE; EC 3.1.1.49). The liberated sinapate is again conjugated by SGT to sinapoylglucose, which is converted to sinapoylmalate by the enzyme sinapoylglucose:malate sinapoyltransferase (SMT). Sinapoylmalate accumulates in epidermal and sub-epidermal tissues as an important UV protecting substance (Landry et al. 1995). Sinapate esters reduce the quality of the rapeseed meal, because they contribute to the bitter taste, astringency and dark colour of rapeseed products (Kozlowska et al. 1990; Naczk et al. 1998; Shahidi and Naczk 1992). In addition, sinapate esters may form complexes with proteins, thus lowering the nutritional value of the protein products (Kozlowska et al. 1990; Shahidi and Naczk 1992).

Since resveratrol and sinapate esters depend on the precursor 4-coumarate (Fig. 1), it was assumed that resveratrol production in the seeds of rape following transformation with the *VST1* gene could be enhanced by simultaneously suppressing the sinapate ester pathway. This hypothesis was supported by the fact, that transgenic rapeseed expressing a *SGT-ds-RNAi*-construct showed reduced levels of all sinapate containing compounds (Hüsken 2004). Plasmid constructs designed for seed-specific expression of *VST1* and silencing of *BnSGT1* by dsRNAi were used for *Agrobacterium*-mediated co-transformation. The present work was aimed at regenerating transgenic *Brassica napus* plants expressing the *VST1* gene and the *SGT-ds-RNAi* construct, and to identify transgenic lines with high levels of resveratrol and reduced levels of sinapate esters.

Materials and methods

Plasmid and bacterial strains

The vector pPSty5 carries the stilbene synthase gene (*Vst1*) from *Vitis vinifera L.* (Melchior and Kindl 1990) under the control of the napin promoter and *the NPTII*-gene conferring kanamycin resistance under control of the CaMV35S-promoter (Fig. 2). The full-length cDNA of the *SGT* gene from *Brassica napus* (*BnSGT1*) has been isolated (Milkowski et al. 2000) and a cDNA fragment covering the first 213 bp of the *BnSGT1* coding sequence was used to construct a suppression cassette under control of the seed-specific napin promoter which was finally inserted into binary vector pLH7000 (Hausmann 1999) resulting in the *dsRNAi* vector pLH-SGT-GUS (Hüsken 2004). *Agrobacterium tumefaciens* strain ATHV C58-C1 (Hellens et al. 2000) containing the binary vector pPSty5 (P. Jorasch, Universität Hamburg, unpublished data) and *Agrobacterium tumefaciens strain* AGL1 (Hellens et al. 2000) carrying the binary suppression plasmid pLH-SGT-GUS were used in the co-transformation experiments for the production of transgenic *Brassica napus* plants.

Plant material

Spring rapeseed (cv. Drakkar) with '00'-quality was used. Seeds were sterilized for 30 min in a 5% sodium hypochlorite solution and rinsed thrice in sterile distilled water. They were germinated on 1/2 MS-medium (Murashige and Skoog 1962) in a growth chamber at 20°C with a day length of 16 h.

Co-transformation, selection and regeneration

Agrobacterium-mediated transformation and plant regeneration were modified according to De Block et al. (1989). Hypocotyl segments were co-inoculated with ATHV C58C1 carrying the binary plasmid pPSty5 and AGL1 carrying the binary plasmid pLH-SGT-GUS. Transformed hypocotyl segments were regenerated in two steps on selective medium with 25 mg/l Kanamycin (pPSty5), 5 mg/l PPT (pLH-SGT-GUS) and 50 mg/l Kanamycin (Duchefa, Haarlem, The Netherlands), 10 mg/l PPT (Duchefa, Haarlem, The Netherlands). Three to six weeks after selection, calli with small shoots were regenerated. Shoots were separated and transferred to MS-medium (Murashige and Skoog 1962) for further regeneration and rooting. Surface-sterilized seeds of cv. Drakkar were germinated in vitro and plants obtained were propagated. These plants were transferred to the greenhouse together with the transformants (T1) and used as controls.

Fig. 2 Schematic drawing of the T-DNA region of the binary plasmid pPSty5. Sequences of functional importance are left border (*LB*) and right border (*RB*); CaMV 35S promoter (*p-35S*) and CaMV 35S terminator (*t-35S*). The seed-specific suppression cassette consists of a napin promoter fragment (*p-nap*) and the *VST1* gene from *Vitis vinifera*. Kanamycin resistance is conferred by the *nptII* gene from transposon Tn5 encoding neomycin-phosphotransferase. Recognition sites are shown for restriction endonucleases *Bam*HI (*B*) and *Pst*I (*P*)

Plant cultivation in the greenhouse and analysis of morphological traits

Putative transgenic plants (T1-plants) propagated in vitro were transferred to the greenhouse and cultivated at 16-h light with a photon flux density of 200–900 μmol m^{-2} s^{-1} and 8-h dark conditions. The main shoot of each plant was selfed for seed production (T2-seeds). Selected T2-plants were selfed to obtain T3-seeds and selected T3-plants were selfed to obtain T4-seeds. The transgenic T1-, T2- and T3-plants were compared to non-transformed control plants in a completely randomized design.

NPT-II-ELISA-test

Leaf tissue (50 mg) of putative transgenic plants was used. The NPT-II-ELISA-test was performed as described by the manufacturer (Agidia, Elkhart, USA). The absorbance of the reaction mixture was measured at room temperature at 450 nm. The standard curve was made following the manufacturer's instructions.

PAT-ELISA-test

Leaf tissue (100 mg) of putative transgenic plants was used. The PAT-ELISA-test was performed as described by the manufacturer (Steffens, Ebringen, Germany). The absorbance of the reaction mixture was measured at room temperature at 625 nm. The standard curve was made following the manufacturer's instructions.

DNA extraction, PCR and Southern blot analysis

To detect the introduced transgenes with PCR in the putative regenerated transgenic plants, DNA samples were extracted from leaf tissue using the DNAeasy-Mini-kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. A primer pair (*PAT*-fw 5′-ATG GGC CCA GAA CGA CGC CC-3′; *PAT*-rev 5′-GCG TGA TCT CAG ATC TCG GT-3′ and *NPTII*-fw 5′-ATC GGG AGC GGC GAT ACC GTA-3′; *NPTII*-rev 5′-GAG GCT ATT CGG CTA TGA CTG-3′) for amplification of the 498 bp *PAT*-fragment (pLH-7000-SGT-GUS resistance marker) and of the 700 bp *NPTII*-fragment (pPSty5 resistance marker) were designed and used for PCR. Southern blot analysis was performed to verify the integration of the transgenes and to determine the respective copy number. For Southern blot analysis, DNA samples were extracted from leaf tissue of primary transgenic plants (T1) according to the DNAeasy-Maxi-kit (Qiagen, Hilden, Germany). Five microgram genomic DNA was digested with *Eco*RI and separated by 0.8% agarose gel electrophoresis. Separated DNA fragments were transferred onto a positively charged nylon membrane (Amersham Biosciences, Freiburg, Germany). Probes (*PAT*- and *NPT-II*-resistance marker) were labelled by PCR (primer sequences see above) with digoxigenin using a DIG-HyPrime-kit (Roche, Mannheim, Germany), and a DIG Nucleic Acid Detection Kit (Roche, Mannheim, Germany) was used for filter hybridization and detection according to the manufacturer's instructions.

Kanamycin and phosphinothricin resistance of T2-seedlings

Primary transformants (T1-plants) were selfed to obtain T2 seeds. Seeds were sterilized for 30 min in a 5% sodium hypochlorite solution. Finally, the seeds were rinsed thrice in sterile distilled water. Fifty-five seeds per T1-plant were germinated on 1/2 MS-medium including 70 mg/l Kanamycin (Duchefa, Haarlem, The Netherlands) in a growth chamber at 20°C with a day length of 16 h. The development of antibiotic damage symptoms was scored 7 days post-treatment. After that, all T2-seedlings were transferred on 1/2 MS-medium including 10 mg/l PPT (Duchefa, Haarlem, Netherlands). The development of herbicide damage symptoms was scored 7 days post-treatment.

Resveratrol glucoside (piceid) and sinapate ester analysis

Grinded seed material (20 mg) was extracted with 1 ml Methanol/H_2O (4:1) in 2 ml-safe-lock tubes in the presence of zirconia beads (1 mm) using a bead beater (Bio Spec Products, Inc., Bartlesville, OK, USA). Reversed phase HPLC (Waters Separator 2795, Waters 2996 photodiode array detector and Waters 474 fluorescence detector) was carried out using a Nucleosil 5-μm C_{18} column (250×4 mm i.d.; Machery & Nagel, Düren, Germany). After centrifugation of the extracts 10 μl samples were injected. Separation was achieved using a 20-min linear gradient at a flow rate of 1 ml min^{-1} from 10 to 50% solvent B (acetonitrile) in solvent A (1.5% *o*-phosphoric acid in water). For detection of piceid, a Waters 474 fluorescence detector was used at 330 nm for excitation and 374 nm for emission and quantified by external standardization with piceid (Sigma, Taufkirchen, Germany). Piceid (*trans*-resveratrol-3-*O*-*β*-glucoside) from *Polygonum cuspidatum* was used as reference compound. Sinapate esters were detected using a photo diode array detector (PDA) at 330 nm and quantified by external standardization with 1-*O*-ß-D-sinapoylglucose, sinapine and sinapate, respectively. The total sinapate ester content (total SE) was calculated as sinapate equivalents.

Analysis of the seed quality traits

Seed samples were analysed for oil, protein, glucosinolate (GSL) and fatty acids by near-infrared reflectance

Table 1 Regeneration and transformation efficiency of *Brassica napus* T1-plants co-transformed with the pPSty5 + pLH-SGT-GUS plasmids

Experiment	Explants	25 mg Kan + 5 mg/l PPT	70 mg Kan + 10 mg/l PPT	ELISA[a] and PCR[a] positive	Transformation efficiency (%)
1	300	51	10	5	1.7
2	300	38	7	7	2.3
3	350	41	9	5	1.4
4	350	36	9	4	1.2
5	350	42	13	7	2.0
Sum/mean	1,650	208	48	28	1.8

[a]NPTII- and PAT-ELISA-Test and NPTII- and PAT-PCR

spectroscopy (NIRS) using the Raps2001.eqa. (Tillmann 2005).

Statistical analysis

All statistical parameters (mean, SD, χ^2) were calculated using the StatGraphics Plus for Windows 3.0 (STATISTICAL GRAPHICS CORP. 1997). For correlation analysis, Spearman rank correlation coefficients were used.

Results

Regeneration and transformation efficiency

A total of 48 NPT II- (70 mg/l kanamycin) and PPT-resistant (10 mg/l, Table 1) Drakkar plants were obtained from 1,650 explants. PCR analysis confirmed the presence of the *NPT II*- and *PAT*-marker genes in 28 of these plants, they showed a *NPT II*- and *PAT*-specific PCR-band of the expected size (Fig. 3) These plants also showed a NPT II- and PAT-ELISA signal (results not shown). The mean transformation efficiency was 1.8% with a range from 1.2 to 2.3% for the experiments (Table 1).

Piceid and sinapate ester content of T2-seeds

Twenty-eight NPT II- and PAT-ELISA-test positive and *NPT II*- and *PAT*-PCR-positive T1-plants were transferred to the greenhouse. The transgenic lines were fully fertile and normal in growth and morphology in comparison with the control plants (data not shown). There was no significant difference in other important agronomic traits, like oil, protein, fatty acid and glucosinolate content of the seeds (Table 2).

T2-seeds obtained from 28 T1-plants were analysed by HPLC for the accumulation of resveratrol and sinapate esters. The HPLC profiles from T2-seed extracts revealed a fluorescence signal with a shorter retention time than that of resveratrol (Fig. 4). This result suggested that the resveratrol has been modified by an endogenous glucosyltransferase activity to piceid. This was confirmed by HPLC-MS data (data not shown). The piceid content (Table 3) ranged from non-detectable amounts up to 361 µg/g seeds. T1-plants showed varying levels of total sinapate esters (Table 4). The average concentration of total sinapate esters of the control plants was 8.79 mg/g. The lowest control plant contained 7.28 mg/g total sinapate esters. The average concentration in the T1-plants was 6.14 mg/g and the lowest T1-plant contained 2.54 mg/g total sinapate esters. Among the transgenic T1-plants, a negative, not significant correlation between piceid and sinapate esters (Table 5) was found, which ranged from −0.34 to −0.36, and a high and significant correlation between sinapoylglucose and all other sinapate esters, which ranged from 0.88** to 0.97**. Furthermore, sinapine, the sum of other not characterized sinapate esters and the total sinapate esters content were highly correlated to each other. Table 6 shows the piceid content and sinapate ester composition of the three T1-plants with the highest piceid content. One of the T1-plants (1502.17) containing 361 µg/g piceid showed a sinapoylglucose content reduction to 0.48 mg/g, whereas the minimum of the controls

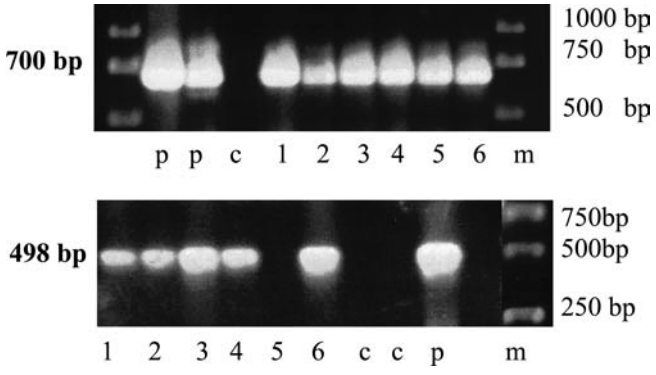


Fig. 3 PCR analysis of *Brassica napus* T1-plants co-transformed with pPSty5 and pLH-SGT-GUS plasmids. **a** Gel electrophoresis of PCR product from NPTII-gene with the expected length of 700 bp. **b** Gel electrophoresis of PCR product from PAT-gene with the expected length of 498 bp. *Lanes: 1–6* T1-plants; *c* untransformed control; *p* plasmid; *m* marker

Table 2 Comparison (mean ± SD) of seed quality traits between controls ($n = 30$) and transgenic T1-plants ($n = 28$) co-transformed with the pPSty5 + pLH-BnSGT-GUS plasmids

Trait	Control ($n = 30$)	1502 ($n = 28$)	Trait	Control ($n = 30$)	1502 ($n = 28$)
Oil (%)	44.9 ± 2.4	43.6 ± 2.9	18:1 (%)	68.5 ± 2.3	67.6 ± 2.9
Protein (%)	24.7 ± 1.9	25.7 ± 1.9	18:3 (%)	6.7 ± 0.8	6.9 ± 1.2
GSL (μmol/g)	22.4 ± 2.1	22.7 ± 3.2	22:1 (%)	0.0 ± 0.0	0.0 ± 0.0

GSL Glucosinolates, *18:1* oleic acid, *18:3* linolenic acid, *22:1* erucic acid

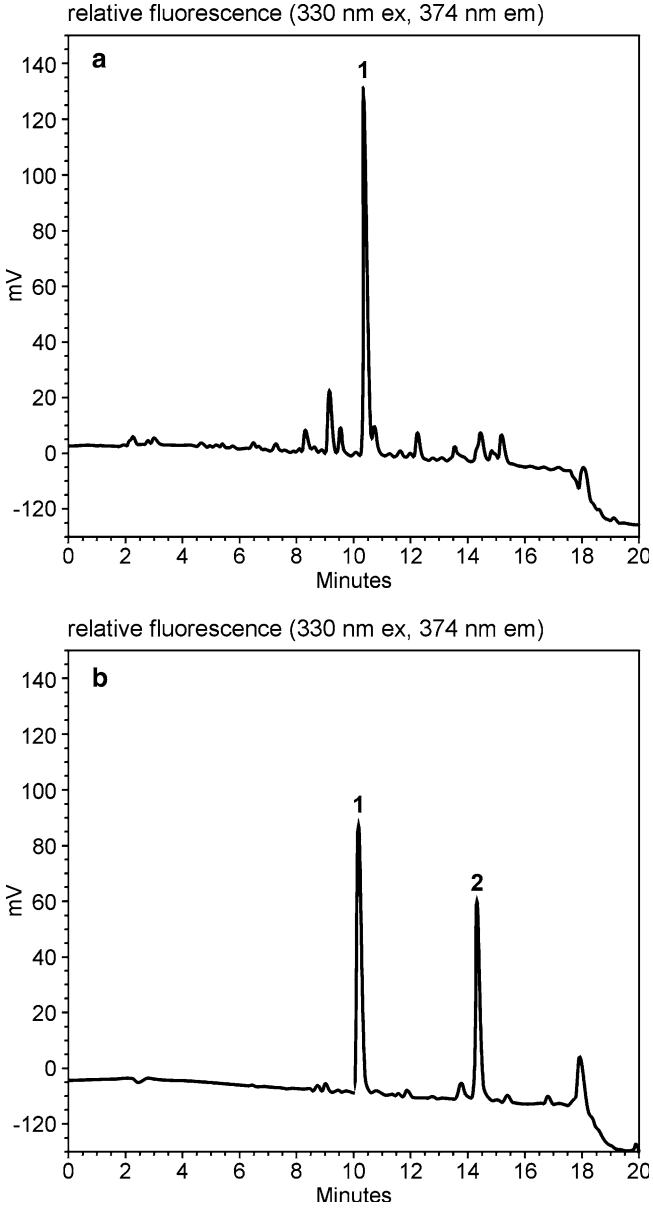


Fig. 4 Reversed-phase HPLC (fluorescence detection) of extracts (**a**) from transgenic seeds (T2). Elution pattern (**b**) of a mixture of piceid (*1*) and *trans*-resveratrol (*2*)

contained 1.21 mg/g. The sinapine content in seeds of this line showed a reduction to 5.00 mg/g, the total sinapate ester content was reduced to 4.00 mg/g.

Transgene copy number in high piceid lines

The three transgenic T1-plants with the highest piceid content as determined in the T2-seeds were characterized for their kanamycin- and PPT-resistance segregation in T2-seeds and analysed by Southern blot hybridisation for the integration of the transgenes in T1-plants. When testing the resistance to kanamycin and PPT of T2 -seeds ($n = 55$ per line) it was found, that 9–12 from 55 seedlings of each line were not resistant to kanamycin (Table 7), indicating a 3:1 inheritance pattern, suggesting that each of the three transformants contains only one T-DNA locus. After transfer to PPT-medium these seedlings showed similar segregation patterns (Table 7). These results correspond to an integrated single copy for each resistance marker gene (NPT II, PPT) as also determined by Southern blot (Table 7, blot not shown).

Piceid and sinapate ester content in the offspring of selected transgenic lines

From each of the three transgenic lines with the highest piceid content and single copy integration for the pPSty5 and pLH-SGT-GUS plasmid 20 T2-seeds were sown together with cv. Drakkar control plants ($n = 22$) in the greenhouse. All T2-plants were normal during the vegetative and reproductive phase of their growth in the greenhouse. There was no significant difference in other important agronomic traits, like oil, protein, fatty acid and glucosinolate content of harvested T3-seeds in comparison with the control plants (data not shown). The average piceid concentration in the T2-plants (T3-seeds) was 310 μg/g for line 1502.15, 203 μg/g for line 1502.16 and 182 μg/g for line 1502.17. The highest piceid content (424 μg/g) was detected in the T2-plant 1502.15.7 (Table 8). The average total sinapate ester content of the control plants was 8.40 mg/g. The average concentration of the T2-plants was 2.80 mg/g for line 1502.15, 3.60 mg/g for line 1502.16 and 4.30 mg/g for line 1502.17. In the T2-plant 1502.15.7 with the highest piceid content the total sinapate ester content was reduced to 1.30 mg/g (Table 8). In comparison to the lowest control plant (7.48 mg/g), the reduction was 83%. The sinapine content in this T2-plant (T3 seeds) of line 1502.15 was 1.90 mg/g. This means a relative reduc-

Table 3 Variation of the piceid content (μg/g) in T1-plants (T2-seeds) co-transformed with pPSty5 + pLH-SGT-GUS plasmids and of control plants

Plasmid	*n*	Mean	SD	Minimum	Maximum
pPSty5 + pLH-SGT-GUS	28	107.73	116.15	0.00	361.00
Control plants	30	0.00	0.00	0.00	0.00

n Number of plants, *SD* standard deviation

Table 4 Variation of the total sinapate ester content[a] (mg/g) in T1-plants (T2-seeds) co-transformed with pPSty5 + pLH-SGT-GUS plasmids and of control plants

Plasmid	*n*	Mean	SD	Minimum	Maximum
pPSty5 + pLH-SGT-GUS	28	6.14	2.70	2.54	10.46
Control plants	30	8.79	0.70	7.28	10.33

[a]Calculated as sinapate equivalents (mg/g)
n Number of plants, *SD* standard deviation

Table 5 Spearman rank correlations (r_s) between sinapate esters (SE) and piceid in T1-plants co-transformed with pPSty5 + pLH-SGT-GUS plasmids ($n = 28$)

	Sinapoylglucose	Sinapine	Other SE[a]	Total SE content[a]
Sinapine	0.93**			
Other SE[a]	0.92**	0.88**		
Total SE content[a]	0.97**	0.99**	0.94**	
Piceid	−0.34	−0.35	−0.34	−0.36

[a]Calculated as sinapate equivalents (mg/g)
**significant at $p < 0.01$

Table 6 Total amount of piceid (μg/g) and sinapate esters (SE) (mg/g) in the T2-seeds of *B. napus* T1-plants co-transformed with pPSty5 + pLH-SGT-GUS plasmids and of a control plant

Line	Piceid	Sinapoylglucose	Sinapine	Other SE[a]	Total SE content[a]
Control[b]	0.00	1.21	7.44	0.72	7.28
1502.15	351	0.48	4.46	0.19	3.68
1502.16	335	0.30	4.38	0.11	3.44
1502.17	361	0.48	5.00	0.12	4.00

[a]Calculated as sinapate equivalents (mg/g)
[b]Control plant with the lowest total SE content (from $n = 30$)

Table 7 Kanamycin and phosphinothricin (PPT) resistance of T2-seedlings and copy number as determined by Southern blot

Line	T2-seedlings ($n = 55$)		Segregation	T2-seedlings ($n = 55$)		Segregation	Copy number[b]	
	r	n.r.	(χ^2)	r[a]	n.r.[a]	(χ^2)	NPT II	PAT
1502.15	44	11	3:1 (0.7)	47	8	3:1 (3.2)	1	1
1502.16	46	9	3:1 (2.2)	47	8	3:1 (3.2)	1	1
1502.17	43	12	3:1 (0.3)	45	10	3:1 (1.3)	1	1

(χ^2) = 3.84 (5%, 1 DF) r/n.r. resistant/not resistant (70 mg/l Kanamycin)
[a]10 mg/l PPT
[b]Results of Southern blot

tion of 80% in comparison to the lowest control plant with 9.46 mg/g. The increase of the piceid content in the 60 T2-plants (T3-seeds) derived from 3 different T1-plants was inversely related to the total sinapate ester content with $r_s = -0.69$** (Fig. 5).

Twenty T2-plants derived from T1-plant 1502.17 could be grouped according to their piceid and total sinapate ester content (T3-seeds). By analysing the kanamycin and PPT-segregation of the T3-seedlings of these T2-plants it could be shown that the T2-plants that

Table 8 Piceid (μg/g), sinapine and total sinapate ester (SE) content (mg/g) in T3-seeds of selected T2-plants with the highest piceid content co-transformed with the pPSty5 + pLH-SGT-GUS plasmid and of a control plant

Line	Piceid	Sinapine	Percent	Total SE[a] content	Percent
Control[b]	0.00	9.46	100	7.48	100
1502.15.7	424	1.90	20	1.30	17
1502.16.8	308	2.21	23	1.51	20
1502.17.20	326	2.10	22	1.44	19

[a]Calculated as sinapate equivalents (mg/g)
[b]Control plant with the lowest total SE content (from $n = 22$)

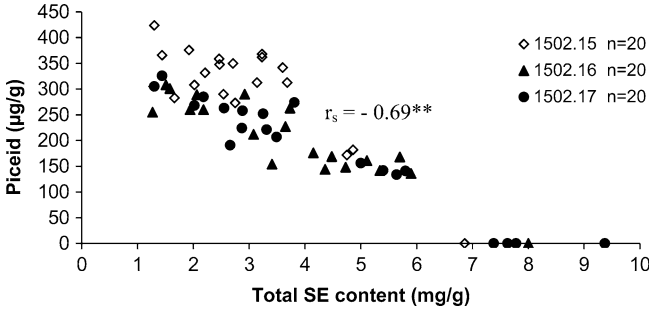


Fig. 5 Correlation between piceid and total sinapate ester (*SE*) content in T2-plants ($n = 60$) derived from three T1-plants with the highest piceid content

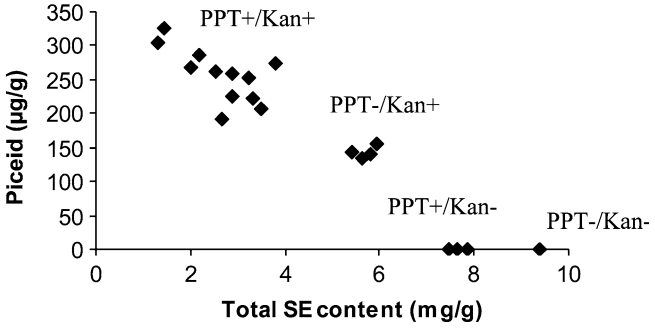


Fig. 6 Piceid (μg/g) and total sinapate ester (mg/g) content of T3-seeds obtained from T2-plants (line 1502.17, $n = 20$). *PPT* +/- Phosphinothricin resistant/not-resistant; *Kan* +/- kanamycin resistant/non-resistant

Table 9 Piceid (μg/g), total sinapate ester (SE) (mg/g) and sinapine content (mg/g) of T4-seeds of selected T3-plants with the highest piceid content, transformed with the pPSty5 + pLH-SGT-GUS plasmid and of control plants

Line	*n*	Mean	SD	Minimum	Maximum
Piceid					
Control	13	0	0	0	0
1502.15.7	20	388	81	221	616
Total SE[a]					
Control	13	8.65	0.60	7.43	9.93
1502.15.7	20	2.51	0.61	1.55	4.10
Sinapine					
Control	13	8.73	0.60	7.37	9.9
1502.15.7	20	3.64	0.88	2.26	5.94

[a]Calculated as sinapate equivalents (mg/g)
n Number of plants, *SD* standard deviation

were homozygous or hemizygous for both integrated genes (*VST1* + *SGT-dsRNAi*) had the highest piceid and the lowest total sinapate ester content (Fig. 6). The T2-plants that were homozygous or hemizygous for the *VST1* gene, but were lacking the *SGT-dsRNAi* gene had an intermediate piceid and total sinapate ester content. The T2-plants carrying only the *SGT-dsRNAi* gene had a lower total sinapate ester content than the plants which did not have either one of the two transgenes. These groups were showing the expected 9:3:3:1 segregation, respectively.

From T2-plant 1502.15.7 with the highest piceid content and single copy integration for the pPSty5 and pLH-SGT-GUS plasmids 20 T3-seeds were sown together with cv. Drakkar control plants ($n = 13$) in the greenhouse. All T3-plants were normal during the vegetative and reproductive phase of their growth. There was no significant difference in other important agronomic traits, like oil, protein, fatty acid and glucosinolate content of harvested T3-seeds in comparison with the control plants (data not shown). The average piceid concentration in the T3-plants (T4-seeds) was 388 μg/g (Table 9). The highest detected piceid concentration was 616 μg/g seeds. The average concentration of total sinapate esters of the control plants was 8.65 mg/g. The average concentration of the T3-plants was 2.51 mg/g (Table 9). In the T3-plant with the highest piceid content (616 μg/g) the total sinapate ester content was reduced to 2.40 mg/g. In comparison to the lowest control plant (7.43 mg/g), the reduction was 68%. In this T2-plant the sinapine content was reduced to 3.50 mg/g, which is a relative reduction of 52% in comparison to the lowest control plant with 7.37 mg/g. By analysing the PPT- and kanamycin-resistance of T4-seedlings ($n = 20$) from all T3-plants it was found that all 20 seedlings of each T3-plant were resistant to PPT and Kanamycin (data not shown), indicating the homozygosity of these T3-plants.

Discussion

The present study reveals that by creating the novel metabolic sink piceid the content of the undesirable sinapate esters is drastically reduced in transgenic seeds of oilseed rape as shown for segregating T2-plants (Fig. 6). This effect is obviously enhanced by the transgenic suppression of the endogenous BnSGT-gene (see Fig. 6). Since the content of sinapate or other SE (see 'other SE'

in Table 6) is also decreased in the transgenic lines, it appears that the common precursor 4-coumaroyl-CoA is used for enhanced production of piceid. Our results are supported by the observation of Orsini et al. (2003) who reported that the expression of the *VST1* gene alone (using plasmid pPSty5 as in our experiment) in cv. Drakkar has led to a maximum piceid content of 159 µg/g seeds among 23 transgenic lines analysed (T2-seeds). In the present experiments, a maximum piceid content of 321 µg/g in T2 seeds was found.

Among the transgenic rapeseed lines containing the *VST1* gene from *Vitis vinifera* and the *BnSGT1* suppression construct pLH-SGT-GUS one homozygous plant (T3-plant) with a single copy insertion was found to contain 616 µg piceid/g seeds. Resveratrol itself was not detected. When the *VST1* gene has been expressed in various plant species, an accumulation of resveratrol and/or piceid between 2 µg/g fresh weight leaves and 290 µg/g fresh weight leaves has been observed (Langcake and Pryce 1976; Fischer et al. 1997; Fettig and Hess 1998; Hipskind and Paiva 2000; Kobayashi et al. 2000).

Co-transformation and copy number

Komari et al. (1996) reported that tobacco and rice plants co-transformed with two separate T-DNAs either by single bacterial strain (containing two T-DNAs on separate plasmids) or by dual bacterial strain (each strain containing one T-DNA) infection contained one locus for each T-DNA. Linked loci were not favoured in either the dual method or the single strain method by using an octopine-type *Agrobacterium* for co-transformation. De Block and De Brouwer (1991) evaluated linkage of transgenes in *Brassica napus* transformants produced by the dual-infection method. They observed that linked loci were favoured. They used a derivative of a nopaline-type strain C58, and suggested that the use of nopaline-type strains in the dual method might favour linked co-transformation. Therefore, the use of C58C1 (nopaline-type) and AGL1 (succinamopine-type) in this experiment may be an important factor contributing to unlinked loci, i.e. progeny with a single copy integration for both T-DNAs segregating independently for both transgenes. The co-transformation efficiency in the present study was 1.8%, which is only half of that compared to single strain transformation experiments using the same SGT-plasmid in the same *Agrobacterium* strain and following the same transformation protocol as described in Hüsken (2004). Komari et al. (1996) reported that the frequency of co-transformation was relatively low using the dual method compared to the single-strain method.

There are conflicting reports about the relationship between transgene copy number and expression level. They have been shown to be negatively correlated, not correlated or positively correlated (Bauer et al. 1998; Hobbs et al. 1993; McCabe et al. 1999). Hipskind and Paiva (2000) reported that many of the higher piceid-accumulating transgenic alfalfa T1-plants contained only one copy of the *VST1* gene, as determined by Southern blot, while the lower accumulating lines contained mostly two or three copies. Reduced expression levels at higher copy numbers may be due to post-transcriptional-gene-silencing (Pickford and Cogoni 2003) or transcriptional-gene-silencing (Jakowitsch et al. 1999). This is in agreement with the present study, that the three selected T1-plants with the highest piceid content contained a single copy for the *VST1* gene.

Implications for seed quality improvement in rapeseed

It was shown, that it is possible to achieve a high amount of piceid (616 µg/g) and a drastical reduction of sinapate esters (2.40 mg/g) in seeds by co-transformation of rapeseed with *VST1* and silencing of *BnSGT1* using the seed-specific napin promotor. There was no indication that other important agronomic traits, like oil, protein, fatty acid and glucosinolate content of the seeds are affected by expressing the stilbene synthase gene and lowering the sinapate ester content. In tobacco and petunia, overexpression of stilbene synthase genes under control of a CaMV 35S promoter led to a substrate competition between stilbene and chalcon synthases causing male sterility (Fischer et al. 1997). In our experiment, there was no evidence for male sterility due to an overexpression of the stilbene synthase gene, which can be explained by the use of a seed specific napin promoter. Zum Felde (personal communication) studied the correlation of sinapate esters content with other seed quality traits in segregating doubled haploid winter rapeseed populations. He found no significant correlation between sinapate ester and oil, protein, fatty acid and glucosinolate content, respectively. Therefore, the high expression of piceid and the drastical suppression of sinapate esters in *B. napus* should not have any negative impact on other important seed quality traits.

Acknowledgements We thank Rosemarie Clemens and Nicole Ritgen-Homayounfar for technical assistance and José Orsini (Saatenunion Resistenzlabor, Germany) for his helpful hints regarding the transformation protocol. Many thanks also to Petra Jorasch (Universität Hamburg) for providing the binary plasmid pPSty5. This work is part of the research project 'NAPUS 2000—Healthy Food from Transgenic Rape Seeds', financially supported by the Bundesministerium für Bildung und Forschung.

References

Bauer M, Libantova J, Moravcikova J, Bekesiova I (1998) Transgenic tobacco plants constitutively expressing acidic chitinase from cucumber. Biologia 52:749–758

Bell JM (1993) Factors affecting the nutritional value of Canola Meal: a review. Can J Anim Sci 73:679–697

Bouchereau A, Hamelin J, Lamour I, Renard M, Larher F (1991) Distribution of sinapine and related compounds in seeds of Brassica and allied Genera. Phytochemistry 30:1873–1881

De Block M, De Brouwer D (1991) Two T- DNA's cotransformed into *Brassica napus* by a double *Agrobacterium tumefaciens* infection are mainly integrated at the same locus. Theor Appl Genet 82:257–263

De Block M, De Brouwer D, Tenning P (1989) Transformation of *Brassica napus* and *Brassica oleracea* using *Agrobacterium tumefaciens* and the expression of the *bar* and *neo* Genes in the transgenic Plants. Plant Physiol 91:694–701

Fettig S, Hess D (1998) Expression of a chimeric stilbene synthase gene in transgenic wheat lines. Transgenic Res 8:179–189

Fischer R, Budde I, Hain R (1997) Stilbene synthase gene expression causes changes in flower colour and male sterility in tobacco. Plant J 11:489–498

Hain R, Reif HJ, Krause E, Langebartels R, Kindl H, Vornam B, Wiese W, Schmelzer E, Schreier PH, Stöcker RH, Stenzel K (1993) Disease resistance results from foreign phytoalexin expression in a novel plant. Nature 361:153–156

Hausmann L, Töpfer R (1999) Entwicklung von Plasmid-Vektoren. In: BioEngineering für Rapssorten nach Maß. Vorträge für Pflanzenzüchtung 45:155–173

Hellens RP, Mullineaux P, Klee H (2000) A guide to *Agrobacterium* binary Ti vectors. Trends Plant Sci 5:446–451

Hipskind JD, Paiva NL (2000) Constitutive accumulation of a resveratrol-glucoside in transgenic alfalfa increases resistance to *Phoma medicaginis*. Mol Plant Microbe Interact 13:551–562

Hobbs SLA, Warkentin TD, DeLong CMO (1993) Transgene copy number can be positively or negatively associated with transgene expression. Plant Mol Biol 21:17–26

Hüsken A (2004) Untersuchungen zur Sinapinsäureestersuppression und Expression von Resveratrol in transgener Rapssaat (*Brassica napus L.*). Dissertation Universität Göttingen. http://www.sub.uni-goettingen.de

Jakowitsch J, Papp I, Moscone EA, van der Winden J, Matzke M, Matzke AJ (1999) Molecular and cytogenetic characterization of a transgene locus that induces silencing and methylation of homologous promoters in trans. Plant J 17:131–140

Jang M, Cai L, Udeani GO, Slowing KV, Thomas CF, Beecher CWW, Fong HHS, Farnswoth NR, Kinghorn AD, Metha RG, Moon RC, Pezzuto JM (1997) Cancer chemopreventive activity of piceid, a natural product derived from grapes. Science 275:218–220

Kindl H (1985) Biosynthesis of stilbenes. In: Higuchi T (ed) Biosynthesis and biodegradation of wood components. Academic Press, New York, pp 349–377

Kobayashi S, Ding CK, Nakamura Y, Nakajima I, Matsumoto R (2000) Kiwifruit (*Actinidia deliciosa*) transformed with a *Vitis* stilbene synthase gene produce piceid (resveratrol-glucoside). Plant Cell Rep 19:904–910

Komari T, Hiei Y, Saito Y, Murai N, Kumashiro T (1996) Vectors carrying two T-DNA for cotransformation of higher plants mediated by *Agrobacterium tumefaciens* and segregation of transformants free from selection markers. Plant J 10:165–174

Kozlowska H, Naczk M, Shahidi F, Zadernowski R (1990) Phenolic acids and tannins in rapeseed and canola. In: Shahidi, F (ed) Canola and Rapeseed. (Production, Chemistry, Nutrition and Processing Technology). Van Nostrand Reinhold, New York, pp 193–210

Landry LG, Chapple CCS, Last R (1995) *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. Plant Physiol 109:1159–1166

Langcake P (1981) Disease resistance of *Vitis* spp. and the production of stress metabolites piceid, ϵ-viniferin, α-viniferin and pterostilben. Physiol Plant Pathol 18:213–226

Langcake P, McCarthy WV (1979) The relationship of piceid production to infection of grapevine leaves by *Botrytis cinerea*. Vitis 18: 244–253

Langcake P, Pryce RJ (1976) The production of piceid by *Vitis vinifera* and other members of the *Vitaceae* as a response to infection or injury. Physiol Plant Pathol 9:77–86

Leckband G, Lörz H (1998) Transformation and expression of a stilbene synthase gene of *Vitis vinifera* L. in barley and wheat for increased fungal resistance. Theor Appl Genet 96:1001–1012

Liang H, Zheng J, Duan X, Sheng B, Jia S, Wang D, Ouyang J, Li J, Li L, Tian W, Jia X (2000) A transgenic wheat with a stilbene synthase gene resistant to powdery mildew obtained by biolistic method. Chin Sci Bull 45:634–638

Lorenzen M, Racicot V, Strack D, Chapple C (1996) Sinapate ester metabolism in wild type and a sinapoylglucose-accumulating mutant of *Arabidopsis*. Plant Physiol 112:1625–1630

Manna SK, Mukhopadhyay A and Aggarwal BB (2000) Piceid suppresses TNF-induced activation of nuclear transcription factors NF-*k*B, activator protein-1, and apoptosis: potential role of reactive oxygen intermediates and lipid peroxidation. J Immunol 164:6509–6519

McCabe MS, Mohapatra UB, Debnath SC, Power JB, Davey MR (1999) Integration, expression and inheritance of two linked T-DNA marker genes in transgenic lettuce. Mol Breed 5:329–344

Melchior F, Kindl H (1990) Grapevine stilbene synthase cDNA only slightly differing from chalcone synthase cDNA is expressed in *Escherichia coli* into catalytically active enzyme. FEBS Lett 268: 17–20

Melchior F, Kindl H (1991) Coordinate- and elicitor-dependent expression of stilbene synthase and phenylalanine ammonia lyase genes in *Vitis* cv. Optima. Arch Biochem Biophys 288:552–557

Milkowski C, Baumert A, Strack D (2000) Cloning and heterologous expression of a rape cDNA encoding UDP-glucose: sinapate glucosyltransferase. Planta 211:883–886

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15:473–497

Naczk M, Aramowicz A, Sullivan A, Shahidi F (1998) Current research developments on polyphenolics of rapeseed/canola: a review. Food Chem 62:489–502

Orsini J, Baumert A, Milkowski C, Weyen J, Leckband G (2003) *Agrobacterium*-mediated production of *Brassica napus* transgenic plants with a stilbene synthase gene. In: Proceedings 11th International Rapeseed Congress, Copenhagen, vol 1, BP3.30, p 155

Pace-Asciak CR, Hahn S, Diamanidis EP, Soleas G, Goldberg DM (1995) The red wine phenolics *trans*-piceid and quercetin block human platelet aggregation and eicosanoid synthesis: implications for protection against coronary heart disease. Clin Chim Acta 235:207–219

Pickford AS, Cogoni C (2003) RNA-mediated gene silencing. Cell Mol Life Sci 60:871–882

Rupprich N, Kindl H (1978) Stilbene synthase and stilbene carboxylate synthases. I. Enzymatic synthesis of 3,5,4,-trihydroxystilbene from p-coumaroyl-CoA and malonyl-CoA. Hoppe Seyler's Z Physiol Chem 359: 165–175

Schoeppner A, Kindl H (1978) Stilbene synthase (Pinosylvine synthase) and its function induction by ultraviolet light. FEBS Lett 108:125–133

Shahidi F, Naczk M (1992) An overview of the phenolics of canola and rapeseed: Chemical, sensory and nutritional significance. J Am Oil Chem Soc 69:917–924

Stark-Lorenzen P, Nelke P, Hänßler G, Mühlbach HP, Thomzik JE (1997) Tranfer of a grapevine stilbene synthase gene to rice (*Oryza sativa* L.). Plant Cell Rep 16:668–673

Statistical Graphics Corp. (1997) Statgraphics Plus 3.0, Statistical graphic System by Statistical Graphic Corporation, Rockville

Stein U, Blaich R (1985) Untersuchungen über Stilbenproduktion und Botrytisanfälligkeit bei *Vitis*-Arten. Vitis 24:75–87

Strack D, Knogge W, Dahlbender B (1983) Enzymatic synthesis of sinapine from 1-O-sinapoylglucose and choline by a cell-free system from developing seeds of red radish (*Raphanus sativus L.*). Zeitschrift für Naturforschung 38c:21–27

Szankowski I, Briviba K, Fleschhut J, Schönherr J, Jacobsen HJ, Kiesecker H (2003) Transformation of apple (*Malus domestica* Borkh.) with the stilbene synthase gene from grapevine (*Vitis vinifera* L.) and a PGIP gene from kiwi (*Actinidia deliciosa*). Plant Cell Rep 22:141–149

Tillmann P (2005) http://www.vdlufa.de/nirs. (Site visited 28 Feb 2005)

Planta (2007) 225:515–522
DOI 10.1007/s00425-006-0360-7

RAPID COMMUNICATION

Secondary product glycosyltransferases in seeds of *Brassica napus*

Juliane Mittasch · Dieter Strack · Carsten Milkowski

Received: 26 April 2006 / Accepted: 13 July 2006 / Published online: 15 August 2006
© Springer-Verlag 2006

Abstract This study describes a systematic screen for secondary product UDP-glycosyltransferases (UGTs; EC 2.4.1) involved in seed development of oilseed rape (*Brassica napus*) and was aimed at identifying genes related to *UGT84A9* encoding UDP-glucose:sinapate glucosyltransferase (EC 2.4.1.120), a proven target for molecular breeding approaches to reduce the content of anti-nutritive sinapate esters. By RT-PCR with primers recognizing the conserved signature motif of UGTs, 13 distinct ESTs could be generated from seed RNA. Sequence analysis allowed to assign the isolated ESTs to groups B, D, E, and L of the UGT family. In an alternative approach, two open reading frames related to *UGT84A9* were cloned from the *B. napus* genome and designated as *UGT84A10* and *UGT84A11*, respectively. Functional expression of *UGT84A10* revealed that the encoded enzyme catalyzes the formation of 1-*O*-acylglucosides (*β*-acetal esters) with several hydroxycinnamates whereas, in our hands, the recombinant UGT84A11 did not display this enzymatic activity. Semi-quantitative RT-PCR confirmed that the majority of potential UGTs specified by the isolated ESTs is differentially expressed. A pronounced transcriptional up-regulation during seed development was evident for *UGT84A9* and one EST (BnGT3) clustering in group E of UGTs. *UGT84A10* was highly induced in flowers and expressed to a moderate level in late seed maturation indicating a possible involvement in seed-specific sinapate ester biosynthesis.

Keywords *Brassica* · Ester-forming glucosyltransferases · Hydroxycinnamates · Secondary product UDP-glucosyltransferases · Seed-specific gene expression

Abbreviations

EST	Expressed sequence tag
fw	Forward
HCA	Hydroxycinnamate
HCA-GT	Hydroxycinnamate glucosyltransferase
ORF	Open reading frame
rev	Reverse
UGT	UDP-glycosyltransferase

Introduction

Plants express an array of different secondary metabolite UDP-glycosyltransferases (UGTs). These enzymes catalyze the transfer of the sugar moiety from an UDP-activated precursor to low molecular-mass acceptors (for review see Vogt and Jones 2000; Gachon et al. 2005a). In the genome of Arabidopsis, Ross et al. (2001) recognized 107 UGT sequences defined by a conserved motif, the UGT signature, involved in binding the UDP moiety of the sugar nucleotide (Hughes and Hughes 1994). Later, this number of UGTs was updated by Paquette et al. (2003) who identified a total of 120 sequences. Analyses

Electronic Supplementary Material Supplementary material is available to authorised users in the online version of this article at http://dx.doi.org/10.1007/s00425-006-0360-7.

J. Mittasch · D. Strack · C. Milkowski (✉)
Department of Secondary Metabolism,
Leibniz Institute of Plant Biochemistry,
Weinberg 3, 06120 Halle (Saale), Germany
e-mail: carsten.milkowski@ipb-halle.de

revealed the presence of 14 distinct groups of Arabidopsis UGTs (Ross et al. 2001). An increasing number of these enzymes is being characterized by heterologous expression unraveling the catalytic activities in vitro.

In members of the Brassicaceae, the major phenylpropanoid sinapate is activated via the β-acetal ester 1-*O*-β-sinapoylglucose (Mock and Strack 1993). Accordingly, 1-*O*-β-sinapoylglucose has been shown to play a pivotal role as precursor for acyl transfer reactions leading to a vast variety of sinapate esters of which sinapoylcholine (sinapine) is the most abundant compound in seeds (Baumert et al. 2005). In oilseed rape (*Brassica napus*) the amount of sinapine makes up about 1% of seed dry weight (Kozlowska et al. 1990; Shahidi and Naczk 1992). Recent work with *B. napus* has shown that seed-specific post-transcriptional suppression of the gene *UGT84A9* encoding UDP-glucose:sinapate glucosyltransferase (Milkowski et al. 2000a) leads to a substantial decrease not only of sinapoylglucose but also of sinapine and related sinapate esters (Baumert et al. 2005; Hüsken et al. 2005). This defines *UGT84A9* as a major target gene for molecular breeding approaches to reduce antinutritive seed phenolics.

However, even the *UGT84A9*-suppressing plants showed residual enzyme activities leading to sinapate ester formation of about 20% of the unmodified control plants (Hüsken et al. 2005). Although this might be caused by non-complete suppression, the open question remains whether alternative glucosyltransferases could be involved in the formation of sinapoylglucose and related sugar esters during seed development.

In Arabidopsis four genes have been identified encoding UGTs, which catalyze the formation of hydroxycinnamate glucose esters (Milkowski et al. 2000b; Lim et al. 2001). Given the general redundancy of plant genomes illustrated by multigene families as well as the amphidiploid nature of the *B. napus* genome that is composed of two triplicated ancestor genomes (U 1935; Lagercrantz and Lydiate 1996), it seems likely that oilseed rape harbors several genes encoding ester-forming hydroxycinnamate glucosyltransferases (HCA-GTs). This led us to start an EST cloning approach to characterize UGTs expressed in developing seeds of *B. napus*. Additionally, by a candidate gene approach, two open reading frames (ORFs) related to *UGT84A9* were isolated from the genome and designated as *UGT84A10* and *UGT84A11*. Expression analysis was used to assess the involvement of different UGTs in seed development.

Material and methods

Plant material and growth conditions

Seeds of winter oilseed rape (*B. napus* L. var. *napus* cv. Express), forage kale (*Brassica oleracea* L. var. *medullosa* cv. Markola) and turnip (*Brassica rapa* L. var. *silvestris* cv. Rex) obtained from Norddeutsche Pflanzenzucht (Holtsee, Germany) were sown on soil and germinated in a growth chamber under constant light at 20°C for 4 days. Plants were grown in the greenhouse under a 16 h light regiment at 12–18°C. Leaves, buds, flowers, and seeds of *B. napus* harvested at different developmental stages were immediately frozen in liquid nitrogen and stored at −80°C.

Purification of nucleic acids

Purification of DNA and RNA was generally done by selective adsorption onto silica using appropriate commercial preparation kits (Qiagen, Hilden, Germany). Genomic DNA was isolated from young *B. napus* seedlings. Total RNA was extracted from different plant tissues and used to purify poly(A^+) RNA by selective binding to oligo-dT-Oligotex beads (Qiagen).

Cloning of UGT-ESTs from *B. napus* seeds

Total RNA from seven different seed stages (early globular to desiccated) covering the whole seed development of *B. napus* over 2 months (Milkowski et al. 2004) was pooled and reversely transcribed with the Omniscript RT Kit (Qiagen) and a customized poly-dT primer (CTA TCT CGG TCT $AA(T)_{15}$). For PCR amplification of UGT sequences, degenerate fw primers were derived from the peptide motif HCGWN (Mackenzie et al. 1997; CAY TGY GGN TGG AAT and CAY TGY GGN TGG AAC) and used in combination with customized poly-dT primer (rev) and seed cDNA as template in a reaction mixture containing Platinum PCR SuperMix High Fidelity (Invitrogen, Karlsruhe, Germany). To increase specificity, a touch down cycling modus was employed with annealing temperatures ranging from 55°C for the initial 5 cycles via 52 and 49°C (5 cycles each) to a final value of 45°C for another 25 cycles. Other cycling parameters were as follows. Denaturation for 30 s at 95°C and elongation for 60 s at 68°C. PCR products were cloned into plasmid pGEM-Teasy (Promega, Mannheim, Germany) and characterized by sequence analysis. Arabidopsis homologues were identified by performing BLASTN search using NCBI BLAST2.2.8 (Altschul et al. 1997) and the AGI genes data set at TAIR website (www. arabidopsis.org/Blast/).

Cloning of *UGT84A9*-related genes from *B. napus*

BLAST search of the database BrassicaDB from UK cropnet (Dicks et al. 2000) detected several genomic DNA fragments from *B. oleracea* with striking sequence similarity to *UGT84A9*. These sequences specified by GenBank Accession numbers BH516441, BH591016, BH560246, BH585199 and BH678779 (Ayele et al. 2005) were used to derive primers for PCR-based cloning strategies. As basic requirements, primers should amplify products from seed cDNA but not recognize *UGT84A9* encoding the predominant ester-forming HCA-GT in seeds. PCR test experiments revealed that primer pairs derived from database entries BH516441 and BH591016 met these requirements. Isolation of complete ORFs was done by a combined approach of cDNA cloning and genome walking. For cDNA cloning poly(A^+)RNA from *B. napus* leaves was subjected to reverse transcription with the Smart RACE cDNA Amplification Kit (BD Clontech, Heidelberg, Germany). Resulting RACE ready cDNA was used as template for 5′ and 3′ RACE-PCR with Advantage 2 Polymerase (BD Clontech). To clone genomic fragments, genome walker libraries were generated from *B. napus* DNA with Universal GenomeWalker Kit (BD Clontech) and screened by PCR using Genomic Advantage Polymerase Mix from the same supplier.

Cloning of *UGT84A10*

Sequence analysis revealed that the *B. oleracea* genomic fragment specified by GenBank Accession number BH591016 (Ayele et al. 2005) harbors the 3′ portion of a *UGT84A9*-related ORF. To amplify the full-length cDNA from *B. napus*, 5′ RACE was performed with the gene specific primer GGC GTT CCC ATT GTT TGT CTA CCG CAG TG (rev) derived from BH591016. The resulting fragment was cloned into pGEMTeasy (Promega) and sequenced (MWG-Biotech, Ebersberg, Germany). The full-length cDNA was amplified with 5′ RACE ready cDNA from leaves as template and the primer pair: ATG GAA TCG TCG CTG ACT C (fw) and TCA ACA GAT GTT GGC AAC CAA (rev). The PCR product was cloned into pGEMTeasy (Promega, Madison WI, USA) and sequenced. To amplify the chromosomal sequences, PCR was carried out with genomic DNA as template and specific primers derived from *UGT84A10* cDNA sequence covering the start codon and the stop codon, respectively (fw: ATG GAA TCG TCG CTG ACT CAT GTG ATG CT, rev: TCA ACA GAT GTT GGC AAC CAA CTT ATC CAC AA).

Cloning of *UGT84A11*

Sequence information of BH516441 (Ayele et al. 2005) was used to amplify the according full-length ORF by a combined strategy including cDNA cloning and genome walking. 3′ RACE was carried out with 3′ RACE ready cDNA from leaves as template and gene specific primer AGA CAA GCC AAG CAG ATC GTC GAA GGT GAG (fw). Nested PCR was done using a 50-fold dilution of cDNA synthesized by the first PCR and 3′ nested primer ACG ACG AGA TCG GGA CGG TTG CGT ATT TG. Since the 5′ RACE approach failed, a genome walking strategy was applied to amplify the 5′ end of the gene using the gene specific primer GGC AGG ACA TGA GTC TCC AGC TTC AGA T (rev). Resulting products were cloned into pGEMTeasy (Promega) and sequenced. The full-length cDNA was PCR-amplified using 5′RACE ready cDNA from leaves as template and the primer pair ATG GAG TTG GAA TCT TCT TCA CAT TCG AGT CAA GTT C (fw) and TCA GCC ACT GAT CAC GCC TAA CTT CTC CA (rev).

Expression of *UGT84A9*-like cDNAs in *E. coli*

Sequences were amplified from cloned cDNAs by PCR using ProofStart DNA Polymerase (Qiagen). Primers were designed to introduce upstream restriction sites for *Nco*I and downstream sites for *Sal*I into the PCR products. Amplification of *UGT84A10* was done with primers: GTA CCA TGG AAT CGT CGC TGA CTC (fw) and GTA GTC GAC ACA GAT GTT GGC AAC CAA (rev). *UGT84A11* was amplified with primers TGA CCA TGG AGT TGG AAT CTT CTT CAC ATT CGA GTC CAG TTC (fw) and GTA GTC GAC GCC ACT GAT CAC GCC TAA CTT CTC CA (rev). PCR products were subject to restriction by *Nco*I and *Sal*I and inserted into similarly cleaved plasmid vector pET28a(+); Novagen (Madison, WI, USA). Ligation mixtures were used to transform *Escherichia coli* strain BL21 Codon Plus (DE3)-PR; Stratagene (LaJolla CA, USA). Cultivation of *E. coli* transformants, induction of cDNA expression and HCA-GT assays were done as described previously (Milkowski et al. 2000a, b).

Semi-quantitative transcript profiling

Reverse transcription reactions were carried out with 4 µg of total RNA using the Omniscript RT Kit (Qiagen) following the protocol provided by the manufacturer. 50-fold dilutions of cDNA were used for each PCR amplification done with Taq PCR Mastermix Kit

(Qiagen) and specific primers. To minimize the likelihood of non-specific annealing, primer sequences were chosen from divergent regions within the whole set of isolated *B. napus* UGT-EST sequences identified by multiple alignment using the CLUSTALW algorithm. Positive control PCR was carried out with a primer pair recognizing ubiquitin. To test the RNA template mixture for possible DNA contamination, PCR was included using a primer pair designed to amplify a genomic *UGT84A9* fragment. Primer sequences are given in Table S1.

Sequence analysis and distance tree construction

Sequence analysis of DNA and proteins was done using the software packages Clone Manager (Durham, NC, USA) and DNASTAR LASERGENE (Madison, WI, USA). For tree construction, partial amino acid sequences of the ORFs starting with the UGT signature motif HCGWNS and including 70 residues were initially aligned using the program ClustalW. Distance analysis and tree construction were done with the program Treecon (Van de Peer and de Wachter 1994). Bootstrap analysis consisted of 100 replicates.

Accession numbers and UGT nomenclature

The isolated sequences described in this paper have been deposited in EMBL Nucleotide Sequence Database and have been assigned the following accession numbers. *UGT84A10*, AM231594; *UGT84A11*, AM231595; BnGT1, AM231596; BnGT2, AM231597; BnGT3, AM231598; BnGT4, AM231599; BnGT5, AM231600; BnGT6, AM231601; BnGT7, AM231602; BnGT8, AM231603; BnGT9, AM231604; BnGT10, AM231605; BnGT11, AM231606; BnGT12, AM231607. UGT names were given by the UGT nomenclature committee (http://www.som.flinders.edu.au/FUSA/ClinPharm/UGT/nomenc.htm). *UGT84A9* encodes UDP-glucose:sinapate glucosyltransferase from *B. napus* that was previously described as *SGT1* (GenBank Accession AF287143; Milkowski et al. 2000a).

Results and discussion

Identification of UGT-ESTs from seeds of *B. napus*

To detect the widest possible range of UGT-ESTs, primers recognizing the signature motif of all UGTs (Mackenzie et al. 1997) were used in RT-PCR with a mixed RNA pool representing seed development from the early globular stage to desiccated ripe seeds. This approach led to the cloning and sequence analysis of 132 ESTs. Of these, about one fourth (32 cDNAs) shared sequence similarity with plant secondary product UGTs. It is noteworthy that the implementation of a touch down protocol for PCR led to a marked increase in the proportion of putative UGT fragments from 17 to about 37%.

The isolated 32 putative UGT-cDNAs represent 13 distinct classes of ESTs. For each of these, sequence analysis revealed ORFs starting with the UGT signature motif. One of the EST classes was found to specify the previously described *UGT84A9* (Milkowski et al. 2000a). The remaining clones denominated as BnGT1-BnGT12 carry ORFs for peptides, which align to UGTs from Arabidopsis, with the exception of BnGT5 and BnGT12. These ESTs were found to accumulate nonsense mutations revealing that they could represent pseudogenes. Consequently, BnGT5 and BnGT12 were excluded from further analysis. The EST fragments BnGT4 and BnGT7 harbor ORFs that appear remarkably short with regard to functional UGTs, which contain more than 100 amino acid residues spanning the distance from the HCGWN motif to the C terminus. This is most probably due to internal priming events during RT-PCR. To establish the relationship of the isolated ESTs to proven UGTs, a distance-based un-rooted tree was constructed (Fig. 1). The isolated sequences could be assigned to established groups of the UGT family, with exception of BnGT4. Groups D, E, and L are represented each by three of the isolated ESTs whereas only one (BnGT2) was found in group B. Regarding sinapine synthesis, it is noticeable that—with exception of *UGT84A9* itself—we could not identify seed-ESTs related to ester-forming HCA-GTs catalyzing the biosynthesis of sinapoylglucose. Group L from which ester-forming UGTs have been described, contains besides *UGT84A9* the seed-ESTs BnGT9 which is related to an anthranilate UGT (Quiel and Benders 2003) and BnGT7. For the BnGT7 homologue, UGT74D1, no activity has been described, so far. However, related enzymes UGT74B1 and UGT74C1 were reported to function in glucosylation of glucosinolate precursors (Grubb et al. 2004; Gachon et al. 2005b). From *B. napus*, a related gene encoding thiohydroximate *S*-glucosyltransferase (S-GT) has been described (Marillia et al. 2001). However, BnGT7 shares only 52 and 64% sequence identity with UGT74B1 and UGT74C1, respectively, and only 52% with S-GT. In contrast, BnGT7 shares 89% amino acid identity with the non-characterized UGT74D1. This reveals that BnGT7 is not an allele of S-GT and that there is only distant relationship to UGT74B1 and UGT74C1.

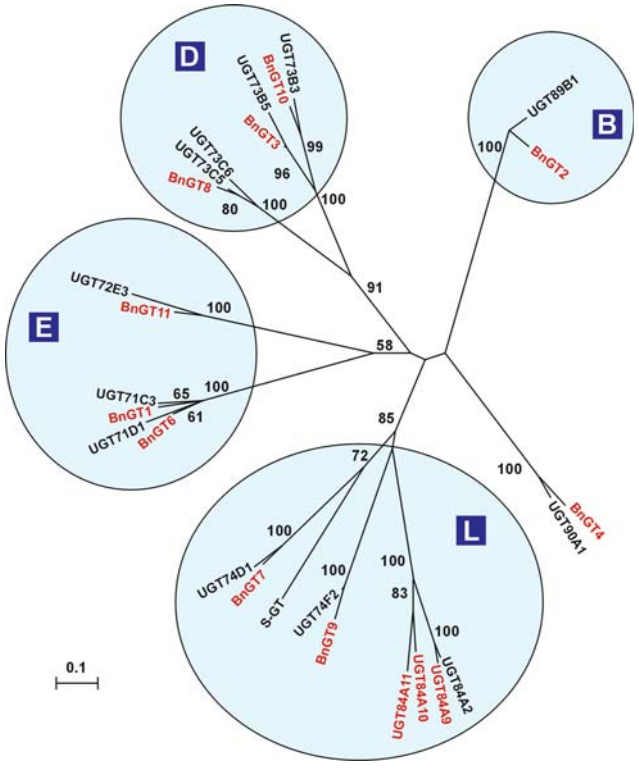


Fig. 1 Composite tree of UGT-ESTs from *Brassica napus* and Arabidopsis homologues. A multiple alignment of peptide sequences covering 70 amino acid residues starting with the UGT signature motif was generated by CLUSTALW. The tree was derived by neighbor-joining distance analysis using Treecon (van de Peer and de Wachter 1994). *Bootstrap* values over 50% are indicated. The scale represents 0.1 fixed mutations per site. Groups B, D, E, and L established for the Arabidopsis UGT superfamily (Li et al. 2001) are indicated. Seed-ESTs and full-length cDNAs from *Brassica napus* isolated in this work are marked in *red*. Arabidopsis sequences are referred according to the UGT nomenclature. UGTs with proven activity: UGT89B1, hydroxybenzoate GT (Lim et al. 2002); UGT73C6, flavonol 3-*O*-glycoside GT (Jones et al. 2003); UGT73C5, brassinosteroid GT (Poppenberger et al. 2005); UGT72E3, hydroxycinnamate, hydroxycinnamyl alcohol and aldehyde GT (Lim et al. 2005); UGT74F2, anthranilate GT (Quiel and Benders 2003); UGT84A2, sinapate GT (Milkowski et al. 2000b; Lim et al. 2001). S-GT refers to thiohydroximate *S*-glucosyltransferase from *Brassica napus* (Marillia et al. 2001)

Interestingly, BnGT8 clusters in group D with two functionally proven enzymes, UGT73C5 and UGT73C6, which have been characterized as brassinolide UGT (Poppenberger et al. 2005) and flavonol 3-*O*-glycoside 7-*O*-glucosyltransferase (Jones et al. 2003), respectively. Although the available sequence data suggest the brassinolide UGT as closest homologue, the presence of kaempferol 3-*O*-sophoroside 7-*O*-*β*-glucopyranoside and related 7-*O*-glucosides in *B. napus* seeds (Baumert et al. 2005) requires an enzymatic activity similar to UGT73C6 for which a full-length BnGT8 is a possible candidate to be tested. In the same UGT group, the ESTs BnGT3 and BnGT10 were found as closest homologues to UGTs involved in establishing resistance to the pathogen *Pseudomonas* (UGT73B5 and UGT73B3; Langlois-Meurinne et al. 2005). In group E, BnGT11 shows close relationship to UGT72E3 catalyzing the *O*-glucosylation of hydroxycinnamates, hydroxycinnamyl alcohols and aldehydes (Lim et al. 2005). Regarding BnGT4 even a thorough analysis could not revise the isolated position in the tree. For the closest homologue of BnGT4, UGT90A1, the acceptor specificity has not been described, so far. Of the second-best aligning proteins, UGT90A4 and UGT90A2, neither has been functionally characterized.

Cloning and characterization of *UGT84A9*-like genes from *B. napus*

By RACE-PCR and genome walking techniques full-length cDNAs were isolated from *B. napus* according to both BH516441and BH591016 and designated *UGT84A10* and *UGT84A11*, respectively. *UGT84A10* carries an ORF for a protein of 472 amino acids with a calculated molecular mass of 53,056 Da and an isoelectric point (pI) of 5.53. *UGT84A11* specifies an ORF for a protein of 476 amino acids. The molecular mass and the pI were calculated to be 52,619 Da and 5.66, respectively.

Both *UGT84A10* and *UGT84A11* share sequence identities in the range of 60–80% with *UGT84A9* and Arabidopsis HCA-GTs at both nucleotide and derived amino acid level (Fig. 1, 2). Among each other, *UGT84A10* and *UGT84A11* share sequence identity of 60%. To prove functionality, cDNAs of *UGT84A10* and *UGT84A11* were expressed in *E. coli*. Crude protein extracts of host cells were assayed for UGT activities with cinnamate and hydroxycinnamates as potential substrates (Fig. 3). UGT84A10 was found to catalyze the formation of HCA glucose esters with preference for ferulate and sinapate as glucose acceptors followed by 4-coumarate. The enzyme showed only minor activities toward cinnamate and caffeate. Thus, according to the acceptor specificity, UGT84A10 differs from UGT84A9 and its Arabidopsis ortholog UGT84A2 which both display a pronounced preference for sinapate (Milkowski et al. 2000a, b). Due to a less efficient IPTG induction, the catalytic activity of recombinant UGT84A10 toward sinapate was remarkably low compared to the values reported for UGT84A9 and UGT84A2. With protein extracts from *E. coli* expressing *UGT84A11*, no product formation was found when assayed with UDP-glucose and cinnamate or hydroxycinnamates although the presence of

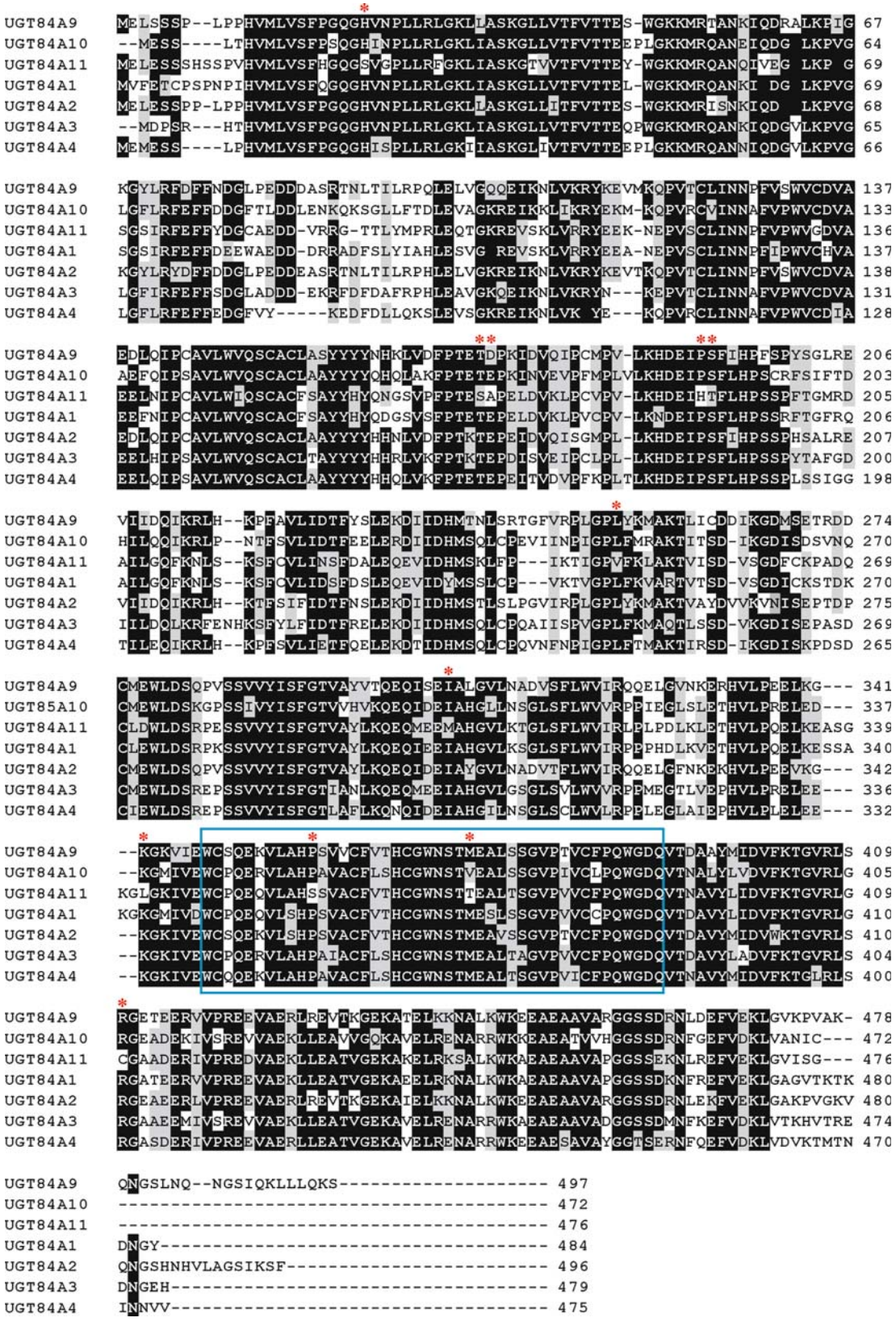


Fig. 2 Alignment of the deduced amino acid sequence of UGT84A9 with those of related potential HCA-GTs from *Brassica napus* and Arabidopsis. Fully conserved residues are shaded in *black*. *Grey shading* indicates a similarity significance value of at least 70%. Substitutions in UGT84A11 within otherwise conserved sequence blocks are indicated by *red asterisks*. The PSPG box motif is marked by a frame

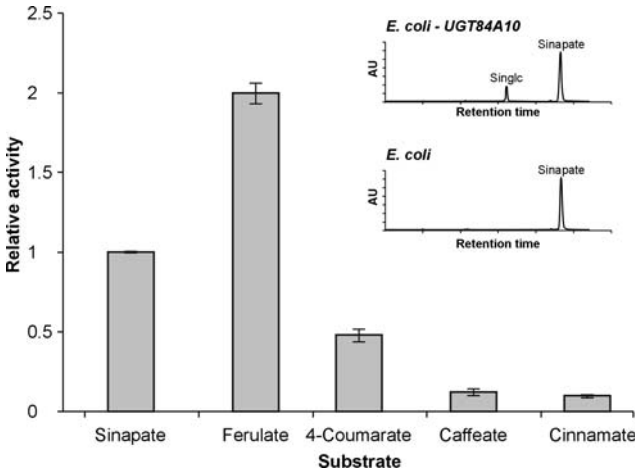


Fig. 3 Acceptor specificity of the recombinant UGT84A10 enzyme. Mean values of the amount of product formation were determined from three independent measurements (mean ± SD). Relative activities were calculated according to the formation of sinapoylglucose [1.9 ± 0.5 pkat (mg protein)$^{-1}$], set to 1.0. HPLC chromatograms of sinapate GT assays conducted with extracts of *Escherichia coli* harboring the empty vector pET28 (*bottom*) and pET28-*UGT84A10* (*top*) showing the production of sinapoylglucose (*Singlc*) are displayed in the inserted figure. In the reaction mixtures, no other product was detectable, excluding the possible formation of *O*-glucosides

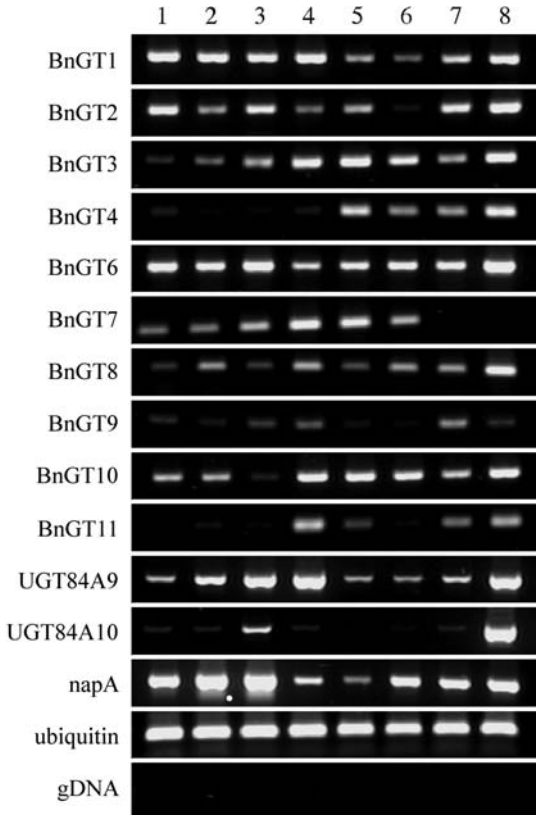


Fig. 4 RT-PCR analysis of *Brassica napus* UGT expression. RNA was extracted from the tissues indicated and used for RT-PCR with specific primers to evaluate UGT gene expression. Results for the *napA* gene (GenBank Accession J02798; Josefsson et al. 1987) encoding a major seed specific storage protein are included. Expression of ubiquitin is given as positive control. PCR with primers recognizing genomic sequences of *UGT84A9* was used to demonstrate the absence of genomic DNA in the samples (*gDNA*). Tissues analyzed: *lane* 1 seeds at early globular to torpedo stage; *lane* 2 seeds at cotyledonary stage; *lane* 3 seeds at well-developed mature embryo stage; *lane* 4 seedlings, 2 days after sowing; *lane* 5 seedlings, 7 days after sowing; *lane* 6 rosette leaves; *lane* 7 flower buds; *lane* 8 flowers

soluble recombinant protein was evident by Western-blot assay (data not shown). Sequence analysis revealed that the ORF derived from *UGT84A11*-cDNA carries amino acid substitutions in sequence blocks that are highly conserved among proven HCA-GTs. These might interfere with functionality (Fig. 2). However, to prove catalytic activity of UGT84A11, a broader spectrum of potential substrates should be tested.

Tissue-specific expression of secondary product GTs

To assess expression of the generated ESTs and the newly identified *UGT84A10*, RT-PCR analyses were carried out using RNA extractions from different tissue types (Fig. 4). RT-PCR reveals that the majority of potential UGTs is differentially expressed. Only one EST (BnGT6) was abundantly detected in all tissues analyzed. With regard to seed maturation, nearly constant high transcript levels were evident for BnGT1 and BnGT2 whereas transcriptional up-regulation was pronounced for BnGT3 and *UGT84A9*. However, for both ESTs, the highest abundance was reached in young seedlings. The expression of BnGT10 and *UGT84A10* was restricted to specific stages of seed development with preference for early (BnGT10) or late (*UGT84A10*) stages of seed maturation. A characteristic feature of the *UGT84A10* gene is the distinct expression in flowers. However, since we used whole open flowers for analysis, the measured expression might reveal transcriptional activity of *UGT84A10* in the very early stage of seed development. Unlike the *napA* gene that encodes a major seed storage protein (Josefsson et al. 1987) and shows the paradigmatic expression pattern of a gene induced in seed maturation, all of the respective UGTs reach their highest expression level in tissues different from seeds.

Expression data raise the question whether UGT84A10 could contribute to sinapate ester accumulation in seeds of *B. napus*. To address this question, expression analysis of *UGT84A10* in a *UGT84A9*-suppressing genetic background (Hüsken et al. 2005) is currently performed in our lab.

Acknowledgments We thank Claudia Horn for technical assistance, Regina Weiß for EST sequencing and Christine Kaufmann for preparing the figures. This work is part of the research project "NAPUS 2000—Healthy Food from Transgenic Rape Seeds", financially supported by the Bundesministerium für Bildung und Forschung.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Ayele M, Haas BJ, Kumar N, Wu H, Xiao Y, Van Aken S, Utterback TR, Wortman JR, White OR, Town CD (2005) Whole genome shotgun sequencing of *Brassica oleracea* and its application to gene discovery and annotation in *Arabidopsis*. Genome Res 15:487–495

Baumert A, Milkowski C, Schmidt J, Nimtz M, Wray V, Strack D (2005) Formation of a complex pattern of sinapate esters in *Brassica napus* seeds, catalysed by enzymes of a serine carboxypeptidase-like acyltransferase family. Phytochemistry 66:1334–1345

Dicks J, Anderson M, Cardle L, Cartinhour S, Couchman M, Davenport G, Dickson J, Gale M, Marshall M, May S, McWilliam H, O'Malia A, Ougham H, Trick M, Walsh S, Waugh R (2000) UK CropNet: a collection of databases and bioinformatics resources for crop plant genomics. Nucleic Acids Res 28:104–107

Gachon CMM, Langlois-Meurinne M, Saindrenan P (2005a) Plant secondary metabolism glycosyltransferases: the emerging functional analysis. Trends Plant Sci 10:542–549

Gachon CMM, Langlois-Meurinne M, Henry Y, Saindrenan P (2005b) Transcriptional co-regulation of secondary metabolism enzymes in Arabidopsis: functional and evolutionary implications. Plant Mol Biol 58:229–245

Grubb CD, Zipp BJ, Ludwig-Müller J, Masuno MN, Molinski TF, Abel S (2004) Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis. Plant J 40:893–908

Hughes J, Hughes MA (1994) Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (*Manihot esculenta* Crantz) cotyledons. DNA Seq 5:41–49

Hüsken A, Baumert A, Strack D, Becker HC, Möllers C, Milkowski C (2005) Reduction of sinapate ester content in transgenic oilseed rape (*Brassica napus*) by dsRNAi-based suppression of *BnSGT1* gene expression. Mol Breed 16:127–138

Jones P, Messner B, Nakajima J-I, Schäffner AR, Saito K (2003) UGT73C6 and UGT78D1, glycosyltransferases involved in flavonol glycoside biosynthesis in *Arabidopsis thaliana*. J Biol Chem 278:43910–43918

Josefsson L-G, Lenman M, Ericson ML, Rask L (1987) Structure of a gene encoding the 1.7 S storage protein, napin, from *Brassica napus*. J Biol Chem 262:12196–12201

Kozlowska H, Naczk M, Shahidi F, Zadernowski R (1990) Phenolic acids and tannins in rapeseed and canola. In: Shahidi F (eds) Canola and rapeseed. Production, chemistry, nutrition and processing technology. Van Nostrand Reinhold, New York, pp 193–210

Lagercrantz U, Lydiate DJ (1996) Comparative genome mapping in Brassica. Genetics 144:1903–1910

Langlois-Meurinne M, Gachon CMM, Saindrenan P (2005) Pathogen-responsive expression of glycosyltransferase genes *UGT73B3* and *UGT73B5* is necessary for resistance to *Pseudomonas syringae* pv. *tomato* in Arabidopsis. Plant Physiol 139:1890–1901

Li Y, Baldauf S, Lim E-K, Bowles DJ (2001) Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. J Biol Chem 276:4338–4343

Lim E-K, Doucet CJ, Li Y, Elias L, Worrall D, Spencer SP, Ross J, Bowles DJ (2002) The activity of *Arabidopsis* glycosyltransferases towards salicylic acid, 4-hydroxybenzoic acid, and other benzoates. J Biol Chem 277:586–592

Lim E-K, Jackson RG, Bowles DJ (2005) Identification and characterisation of Arabidopsis glycosyltransferases capable of glucosylating coniferyl aldehyde and sinapyl aldehyde. FEBS Lett 579:2802–2806

Lim E-K, Li Y, Parr A, Jackson R, Ashford DA, Bowles DJ (2001) Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in Arabidopsis. J Biol Chem 276:4344–4349

Mackenzie P, Owens I, Burchell B, Bock K, Bairoch A, Belanger A, Fournel-Gigleux S, Green M, Hum D, Iyanagi T et al (1997) The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. Pharmacogenetics 7:255–269

Marillia E-F, MacPherson JM, Tsanga EWT, van Audenhove K, Keller WA, GrootWassink JWD (2001) Molecular cloning of a *Brassica napus* thiohydroximate *S*-glucosyltransferase gene and its expression in *Escherichia coli*. Physiol Plant 113:176–184

Milkowski C, Baumert A, Schmidt D, Nehlin L, Strack D (2004) Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. Plant J 38:80–92

Milkowski C, Baumert A, Strack D (2000a) Cloning and heterologous expression of a rape cDNA encoding UDP-glucose:sinapate glucosyltransferase. Planta 211:883–886

Milkowski C, Baumert A, Strack D (2000b) Identification of four Arabidopsis genes encoding hydroxycinnamate glucosyltransferases. FEBS Lett 486:183–184

Mock H-P, Strack D (1993) Energetics of the uridine 5′-diphosphoglucose: hydroxycinnamic acid acyl-glucosyltransferase reaction. Phytochemistry 32:575–579

Paquette S, Møller BL, Bak S (2003) On the origin of family 1 plant glycosyltransferases. Phytochemistry 62:399–413

Poppenberger B, Fujioka S, Soeno K, George GL, Vaistij FE, Hiranuma S, Seto H, Takatsuto S, Adam G, Yoshida S, Bowles D (2005) The UGT73C5 of Arabidopsis thaliana glucosylates brassinosteroids. Proc Natl Acad Sci USA 102:15253–15258

Quiel JA, Benders J (2003) Glucose conjugation of anthranilate by the Arabidopsis UGT74F2 glucosyltransferase is required for tryptophan mutant blue fluorescence. J Biol Chem 278:6275–6281

Ross J, Li Y, Lim E-K, Bowles DJ (2001) Higher plant glycosyltransferases. Genome Biol 2:reviews3004.1–3004.6

Shahidi F, Naczk M (1992) An overview of the phenolics of canola and rapeseed: chemical, sensory and nutritional significance. J Am Oil Chem Soc 69:917–924

U N (1935) Genomic analysis of *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. Jpn J Bot 7:389–452

Van de Peer Y, de Wachter R (1994) TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput Appl Biosci 10:569–70

Vogt T, Jones P (2000) Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. Trends Plant Sci 5:380–386

Planta (2008) 227:375–385
DOI 10.1007/s00425-007-0624-x

ORIGINAL ARTICLE

The genes *BnSCT1* and *BnSCT2* from *Brassica napus* encoding the final enzyme of sinapine biosynthesis: molecular characterization and suppression

Diana Weier · Juliane Mittasch · Dieter Strack · Carsten Milkowski

Received: 30 May 2007 / Accepted: 30 August 2007 / Published online: 20 September 2007
© Springer-Verlag 2007

Abstract This study describes the molecular characterization of the genes *BnSCT1* and *BnSCT2* from oilseed rape (*Brassica napus*) encoding the enzyme 1-*O*-sinapoyl-*β*-glucose:choline sinapoyltransferase (SCT; EC 2.3.1.91). SCT catalyzes the 1-*O*-*β*-acetal ester-dependent biosynthesis of sinapoylcholine (sinapine), the most abundant phenolic compound in seeds of *B. napus*. GUS fusion experiments indicated that seed specificity of *BnSCT1* expression is caused by an inducible promoter confining transcription to embryo tissues and the aleurone layer. A dsRNAi construct designed to silence seed-specifically the *BnSCT1* gene was effective in reducing the sinapine content of *Arabidopsis* seeds thus defining SCT genes as targets for molecular breeding of low sinapine cultivars of *B. napus*. Sequence analyses revealed that in the allotetraploid genome of *B. napus* the gene *BnSCT1* represents the C genome homologue from the *B. oleracea* progenitor whereas *BnSCT2* was derived from the *Brassica* A genome of *B. rapa*. The *BnSCT1* and *BnSCT2* loci showed colinearity with the homologous *Arabidopsis SNG2* gene locus although the genomic microstructure revealed the deletion of a cluster of three genes and several coding regions in the *B. napus* genome.

D. Weier · J. Mittasch · D. Strack · C. Milkowski (✉)
Department of Secondary Metabolism,
Leibniz Institute of Plant Biochemistry,
Weinberg 3, 06120 Halle (Saale), Germany
e-mail: Carsten.Milkowski@ipb-halle.de

Present Address:
D. Weier
Department of Molecular Genetics,
Leibniz Institute of Plant Genetics and Crop Plant Research,
Corrensstr. 3, 06466 Gatersleben, Germany

Keywords *Brassica* · *Arabidopsis* · Phenylpropanoid metabolism · Sinapoylglucose:choline sinapoyltransferase · Gene structure · Promoter · RNAi suppression

Abbreviations

SCT	1-*O*-sinapoyl-*β*-glucose:choline sinapoyltransferase
BAC	Bacterial artificial chromosome
SE	Sinapate esters
SC	Sinapoylcholine
SG	1-*O*-sinapoyl-*β*-glucose

Introduction

Like many other Brassicaceae plants, the agronomically valuable crop oilseed rape (*Brassica napus*) accumulates high amounts of sinapate esters (SE) in organ-specific complex patterns. In *B. napus* as in the model plant *Arabidopsis* the most prominent SE compounds are sinapoylmalate, one of the UV shielding substances (Sheahan 1996) that accumulates in epidermal and sub-epidermal leaf tissues, and sinapoylcholine (sinapine) that is specifically produced in seeds (Shahidi and Naczk 1992; Mock et al. 1992; Hause et al. 2002). The sinapine concentration in *B. napus* reaches values in the range of 1–2% (w/w) of seed mass thus representing by far the major seed phenolic compound (Bell 1993). In *B. napus* seeds, sinapine is accompanied by minor amounts of several other sinapate esters, mostly sugar and flavonoid glycoside esters (Baumert et al. 2005), which together make up the agronomically important seed trait of total seed SE content. SE compounds are considered anti-nutritional mainly by the fact that they confer astringency and low digestibility to the seed meal that is rich in proteins of a well-balanced composition for animal feeding and potential human food supplements (Ohlson 1978; Sozulski

1979; Shahidi and Naczk 1992). Therefore, a substantial reduction of seed SEs, especially of the sinapine content, is regarded a main requirement for an efficient use of rape seed meal. Unfortunately, the *B. napus* cultivars analyzed so far did not display stable low SE genotypes making conventional breeding approaches unfeasible. This led us to investigate genes involved in sinapine biosynthesis of *B. napus* with the aim to define targets for molecular breeding approaches to reduce the seed SE content.

The biosynthesis of SE compounds is fueled by the network of phenylpropanoid metabolism (Ruegger and Chapple 2001; Nair et al. 2004) providing sinapate that is conjugated via 1-*O*-sinapoyl-*β*-glucose with various acyl acceptors (Tkotz and Strack 1980; Strack et al. 1983; Baumert et al. 2005). The use of 1-*O*-sinapoylglucose as acyl donor instead of the more common coenzyme A thioester marks an alternative pathway of transacylation described e.g. in the indoleacetate metabolism (Kesy and Bandurski 1990) and in other parts of plant secondary metabolism (Mock and Strack 1993; Li et al. 1999). Accordingly, acyltransferases involved in glucose ester-dependent transacylations are distinguishable from CoA-dependent enzymes by their homology with hydrolytic serine carboxypeptidases as has initially been shown for isobutyroyl transferases from wild tomato (Li and Steffens 2000). Therefore, these enzymes were assigned to the family of serine carboxypeptidase-like (SCPL) proteins.

Regarding the transacylation steps from 1-*O*-sinapoylglucose towards sinapine and sinapoylmalate, the analysis of *Arabidopsis* mutants displaying the sinapoylglucose accumulator (*sng*) phenotype has led to the identification of the respective genes—*SNG1* encoding sinapoylglucose:malate sinapoyltransferase (SMT; EC2.3.1.92) and *SNG2* encoding 1-*O*-sinapoyl-*β*-glucose:choline sinapoyltransferase (SCT; EC2.3.1.91). Sequence analysis revealed both SMT and SCT as members of the SCPL protein family (Lehfeldt et al. 2000; Shirley et al. 2001). In a parallel approach, a cDNA was isolated from *B. napus* seeds sharing about 85% sequence identity with *SNG2* from *Arabidopsis* and designated as *BnSCT1* (Accession number AY383718) expressing SCT enzyme activity (Milkowski et al. 2004).

The aim of the present work was to identify and characterize the homoeologous SCT genes, *BnSCT1* and *BnSCT2*, from the allotetraploid genome of *B. napus*. *BnSCT1* was chosen to demonstrate functionality of the encoded SCT enzyme and to prove seed-specificity of the promoter. dsRNAi-mediated suppression of *BnSCT1* led to decreased amounts of accumulating sinapine. The microstructure of both genomic *BnSCT1* and *BnSCT2* loci has been elucidated and compared to the *SNG2* locus of *Arabidopsis*.

Materials and methods

Plant material

Wild-type plants of *Arabidopsis thaliana* L. Heynh. ecotype Columbia as well as the SCT-deficient *Arabidopsis sng2* null mutant SALK_002255 were obtained from the Nottingham *Arabidopsis* Stock Centre (NASC; http://www.arabidopsis.info/). *Arabidopsis* plants were cultivated in a soil:vermiculite mixture (4:1) under a photoperiod of 16 h light/8 h dark at 23°C and a photon flux density of 250 µmol m^{-2} s^{-1} in a plant growth chamber (Percival Scientific, Perry, IA, USA).

Seeds of winter oilseed rape (*B. napus* L. var. *napus* cv. Express), forage kale (*B. oleracea* L. var. *medullosa* cv. Markola) and turnip (*B. rapa* L. var. *silvestris* cv. Rex) obtained from Norddeutsche Pflanzenzucht (Holtsee, Germany) were sown on soil and germinated under constant light at 20°C for 5 days. Plants were grown in the greenhouse at 12–18°C under a 16 h light regiment.

Tobacco plants (*Nicotiana tabacum* L. cv. Samsun) obtained from Vereinigte Saatzuchten eG (http://www.vs-ebstorf.de) were grown on soil under a photoperiod of 16 h light/8 h dark at 23°C in the greenhouse. Photon flux density for all plants cultivated in the greenhouse was between 200 and 900 µmol m^{-2} s^{-1}.

Construct for complementation of *Arabidopsis sng2* mutant

The coding part of *BnSCT* cDNA including 8 bp upstream the translation start was amplified by PCR with primers attaching restriction sites for *Sma*I and *Bam*HI to the 5′ and 3′ ends of the product (primer pair: 5′Sma-Kompl 5′-TATCCCGGGCGGAGAAAATGAGAA-3′; 3′Bam-Kompl 5′-TATGGATCCTCAGAGAGATTCACCATCAA-3′). The PCR product was cloned in a modified pBluescript plasmid (Stratagene, La Jolla, CA, USA) between the seed specific *napin* promoter (Kridl et al. 1991) and the *nos* terminator (Chen et al. 2003). The expression cassette was then transferred as *Spe*I-*Hin*dIII fragment to the binary vector pLH7000 (Hausmann and Töpfer 1999) cleaved by the same restriction enzymes.

Construct for transient expression of *BnSCT* in tobacco leaves

The coding part of *BnSCT* cDNA including 8 bp upstream the translation start was transcriptionally fused to the promoter of Rubisco small subunit (*rbcS1*) from *Asteraceous chrysanthemum* (Outchkourov et al. 2003) by cloning into the *Not*I site of plasmid pImpact1.1 (Plant Research International, Wageningen, The Netherlands). The whole *BnSCT1* expression cassette was then introduced as *Asc*I–*Pac*I

fragment into the binary vector pBINPLUS (Plant Research International; van Engelen et al. 1995).

Construct for dsRNAi suppression of *BnSCT*

A subfragment of *BnSCT* cDNA covering the first 332 bp of the reading frame was amplified by PCR using the following primer pairs:

5′Nco-RNAi(as) 5′-GTACCATGGATGAGAAATCTTTACTTTCTAGTC-3′ and
3′Sma-RNAi(as) 5′-GTACCCGGGAGCTCTAAAGGAGGC-3′;
5′Nhe-RNAi(s) 5′-GTAGCTAGCATGAGAAATCTTTACTTTCTAGTC-3′ and
3′Bam-RNAi(s) 5′-GTAGGATCCAGCTCTAAAGGAGGCAGTGTC-3′.

The PCR products are flanked by restriction sites for *Nco*I and *Sma*I or *Nhe*I and *Bam*HI allowing to assemble a suppression cassette consisting of the 332 bp *BnSCT1* fragment in antisense (as *Sma*I–*Nco*I fragment), a subfragment of the bacterial *gus*A gene (Chuang and Meyerowitz 2000) as spacer element and the *BnSCT1* fragment in sense orientation (as *Nhe*I–*Bam*HI fragment). This dsRNAi cassette was inserted between the seed specific *napin* promoter and the *nos* terminator. Assembling of these elements was done in pBluescript from which the whole suppression construct was cut as *Spe*I–*Hin*dIII fragment and ligated with the similarly cleaved binary vector pLH7000 (Hausmann and Töpfer 1999) to give the *BnSCT* suppression plasmid.

Promoter GUS fusions

A DNA fragment covering 729 bp upstream from the translation initiation site of *BnSCT1* was amplified using the primer pair:

5′Xho-pSCT 5′-CTCGAGTAGATCATGTCGAAGGTGTTG-3′
3′Bam-pSCTBn 5′-GGATCCTTTCTCCGCTTCTTGG-3′.

The PCR product was fused as *Xho*I–*Bam*HI fragment to the promoterless *uidA* gene from *E. coli* encoding *β*-glucuronidase. The *BnSCT1*::*uidA* cassette was transferred to the binary plasmid pGreen (Hellens et al. 2000).

Transformation of *Arabidopsis*

Constructs for transformation of *Arabidopsis* were introduced into *Agrobacterium tumefaciens* EHA105 (Hood et al. 1993) by electroporation. Plant transformation was performed by the floral dip method (Clough and Bent 1998). Transgenic plants (T1) harboring the *bar* gene that confers resistance against phosphinothricine (D Halluin et al. 1992) were identified by spraying soil-grown seedlings with a 5,000-fold dilution of the herbicide BASTA (Hoechst Schering AgrEvo GmbH, Sinsheim, Germany). For selection of transgenic plants expressing the *nptII* gene, T1 seeds were surface-sterilized in 70% ethanol for 2 min followed by a mixture of 0.15% (v/v) Tween 20 and 12% (w/v) household bleach (sodium hypochlorid) for 10 min. Seeds were rinsed thoroughly with sterile water and after swelling over night at 4°C plated on modified MS medium (Duchefa, Haarlem, The Netherlands; Murashige and Skoog 1962) supplemented with 500 μg ml^{-1} carbenicillin and 25 μg ml^{-1} kanamycin. After scoring the development of antibiotic damage symptoms for 14 days post treatment, kanamycin resistant plants were transferred to soil.

Transient transformation of *N. tabacum*

Agrobacterium tumefaciens GV2260 (McBride and Summerfelt 1990) harboring the binary plasmid construct was used to transiently transform tobacco (*N. tabacum* L. cv. Samsun) by infiltration of 10-week-old leaves as described previously (Kapila et al. 1997). After 5 days of incubation under greenhouse conditions, infected leaf areas were cut out for further analysis.

BAC library screen

A genomic BAC library from *B. napus* cv. Express was screened by hybridization of a ^{33}P-labeled *BnSCT1*-cDNA as probe to high-density colony macroarrays according to the protocol given by the provider [German Resource Center for Genome Research (RZPD), Berlin, Germany]. BAC DNA was prepared with the BAC-MAX™ DNA Purification Kit (EPICENTRE, Madison, WI, USA).

BAC clones were screened for full length *BnSCT* by PCR with primers recognizing the cDNA ends of both *BnSCT1* and *BnSCT2* genes:

5′SCTBn 5′-ATGAGAAATCTTTACTTTCTAGTC-3′
3′SCTBn 5′-TCAGAGAGATTCACCATCAATCC-3′

For BAC classification, PCR was done with a forward primer recognizing both *BnSCT* genes (5′SCTuni) and reverse primers binding selectively in the 3′UTR of either *BnSCT1* (3′UTR1) or *BnSCT2* (3′UTR2). The selective 3′UTR primers were derived from 5′ incomplete *BnSCT*-cDNAs isolated from a seed-cDNA library of *B. napus* (Milkowski et al. 2004).

5′SCTuni 5′-GTTCAAAAGATGGATTGATGGTGAATCTC-3′

3′UTR1 5′-TATAAAAGGGCATTGTTATTATTAATTTTTTTAC-3′
3′UTR2 5′-AATAAAAGGTTATTATTATTTTTAATTTTTGAC-3′

Sequence analysis and tree construction

Sequence analysis of DNA and proteins was done using the software package Clone Manager (Scientific & Educational Software, Cary, NC, USA). For distance tree construction, genomic DNA sequences from translation start to stop codons were aligned by the program CLUSTALW. Distance analysis and tree construction were done with the program Treecon (van de Peer and Wachter 1994). Bootstrap analysis consisted of 100 replicates.

BnSCT enzyme assay

Enzyme reaction mixtures contained 2 mM sinapoylglucose, 10 mM choline chloride and 100–150 µg plant protein in a total volume of 100 µl sodium phosphate buffer (100 mM, pH 7.0). Mixtures were incubated for 60 min at 30°C, and the reactions were terminated by adding 10 µl trifluoroacetic acid (TFA). To detect product formation, mixtures were centrifuged and aliquots of the supernatants were analyzed by HPLC as described (Milkowski et al. 2004). Sinapoylglucose for use in enzyme assays was purified from leaves of the *Arabidopsis* mutant *sng1* (Lorenzen et al. 1996). Plant protein extracts were prepared from tobacco leaves by disrupting infected leaf sectors of 100 mg fresh weight with a glass homogenizer (VWR International, Vienna, Austria) in 1 ml of 100 mM sodium phosphate buffer (pH 7.0). After centrifugation (5 min at 3,000*g* and 4°C), the supernatants were desalted using PD-10 Sephadex G-25 columns (Amersham Bioscience, Munich, Germany). Diluted extracts were concentrated by ultrafiltration through the Amicon Ultra 4 filter device with a molecular weight cut off (MWCO) of 10 kDa (Millipore, Schwalbach, Germany). Protein concentrations were determined by the Bradford method (Bradford 1976) with bovine serum albumin as standard. Assays were done with protein extracts from 5 infected leaf areas.

Secondary metabolite analysis

For extraction of metabolites, leaf and seed tissues were suspended in 80% (v/v) aqueous methanol and disrupted by vigorous shaking in a Bead Beater (Biospec Products, Bartlesville, OK, USA) for 2 min at maximum speed in the presence of zirconia beads (1 mm in diameter). Homogenates were centrifuged and aliquots of the supernatants analyzed by HPLC on a 5-µm Nucleosil C18 column (250 mm × 4 mm i.d.; Macherey-Nagel, Düren, Germany). A 20-min linear gradient was applied at a flow rate of 1 ml min^{-1} from 10 to 50% solvent B (acetonitrile) in solvent A (1.5% *o*-phosphoric acid in water). The compounds were photometrically detected at 330 nm and quantified by external standardization with authentic samples (Baumert et al. 2005) from our collection of phenylpropanoid compounds (not documented).

Analysis of nucleic acids

Molecular cloning experiments and restriction analyses were performed according to the standard protocols (Sambrook et al. 1989). Routine sequencing of DNA was done by a commercial supplier (MWG-BIOTECH AG, Ebersberg, Germany). Sequences of DNA and proteins were analyzed by use of the software package Clone Manager (Scientific & Educational Software). Leader peptides for subcellular targeting and cleavage sites were predicted using the algorithm described by Nielsen et al. (1997), available at http://www.cbs.dtu.dk/services/SignalP-2.0/.

GUS assay

For histochemical detection of GUS activity (Jefferson et al. 1987), seeds were perforated with a needle and transferred to a GUS staining solution (100 mM sodium phosphate buffer, pH 7.0, 0.5 mM 5-bromo-4-chloro-3-indolyl-*β*-D-glucuronide (X-Gluc; Duchefa), 1 mM EDTA, 0.05% (v/v) Triton X-100, 0.1 mM $K_3[Fe(CN)_6]$, 0.1 mM $K_4[Fe(CN)_6]$). Vacuum was applied for 1 h before incubation was started at 37°C in the dark for 48 h. Chlorophyll was removed by incubation in 70% (v/v) aqueous ethanol at room temperature overnight. Observation of whole seeds, dissected embryos and seed coats were performed with a stereo microscope (Zeiss Stemi 2000-C; Carl Zeiss, Jena, Germany).

For cellular analyses, seeds were fixed in 0.3% (v/v) formaldehyde, 1 mM EDTA in 50 mM sodium phosphate buffer, pH 7.0. After incubation for 30 min at room temperature, samples were washed in sodium phosphate buffer and incubated in GUS staining solution at 37°C in the dark for 48 h. Embedding and sectioning was done according to Hause et al. (1996). For microscopy of seed cross sections Zeiss Axioskop II (Carl Zeiss) was used.

Accession numbers

The isolated sequences described in this paper have been deposited in EMBL Nucleotide Sequence Database (http://www.ebi.ac.uk/embl/index.html) and have been assigned the following accession numbers. *BoSCT*, AM706347;

BrSCT, AM706348; *BnSCT1*, AM706349; *BnSCT2*, AM706350.

Results

BnSCT1 complements the *sng2* mutant phenotype of *Arabidopsis*

To provide the genetic proof that the previously isolated *BnSCT1* (AY383718; Milkowski et al. 2004) represents the *SNG2* orthologue from *B. napus*, we fused the *BnSCT1*-cDNA to the seed-specific napin promoter and introduced the resulting expression cassette into the binary vector pLH7000 to generate plasmid pLHBnSCT1. This construct was then used to transform the SCT-deficient *Arabidopsis sng2* null mutant SALK_002255 that seed-specifically accumulates sinapoylglucose instead of sinapine. Control lines were generated by transforming SALK_002255 with the empty vector pLH7000. Phosphinothricine-resistant seedlings (T1) were selected for further growth and tested for transgene insertion by PCR amplification of the resistance-mediating *bar* gene from genomic DNA (data not shown). Transgenic plants were cultivated in the greenhouse, and methanolic seed extracts (T2) were profiled for soluble sinapate esters.

For SALK_002255/pLHBnSCT1 transgenic lines, HPLC profiles revealed a partial reconstitution of the sinapate ester pattern found in *Arabidopsis* wild-type seeds (Fig. 1). T2 seeds of these lines contained significant amounts of sinapine (Table 1). The recovery of sinapine accumulation in the mutant transformants was accompanied by a decrease in seed sinapoylglucose content. These data demonstrate that *BnSCT1*- cDNA complements the *sng2* mutation in *Arabidopsis*.

Expression of *BnSCT1* causes SCT activity in tobacco

Sequence analyses predicted that SCT is subject to post-translational modifications like glycosylation and endoproteolytic cleavage requiring an eukaryotic expression system to produce the mature functional enzyme. Previous experiments with the *SNG2* gene indicated that functional expression in yeast requires distinct modifications like mutant host strains defective in vacuole formation (Shirley and Chapple 2003). This led us to choose tobacco (*N. tabacum*) as a heterologous plant host for functional expression of *BnSCT1*. Transient transformation of *BnSCT1*-cDNA under control of a strong Rubisco promoter conferred SCT activity of about 20.6 pkat mg^{-1} protein to a leave extract of *N. tabacum* (Fig. 2). These data provide the unequivocal biochemical proof that the *BnSCT1* cDNA encodes the functional SCT from *B. napus*.

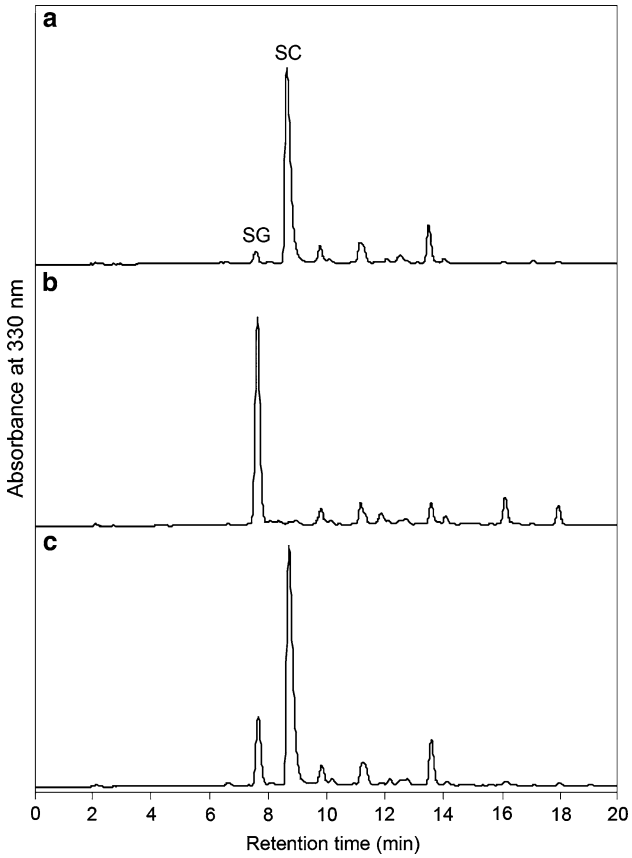


Fig. 1 Complementation of an *Arabidopsis sng2* mutant by *BnSCT1*-cDNA. HPLC diagrams of methanolic seed extracts from representative plants of *Arabidopsis* Col-O wild type (**a**) and the *sng2* mutant line SALK_002255 (**b**) both transformed with the empty vector pLH7000 and SALK_002255 expressing *BnSCT1* from pLHBnSCT1 (**c**). *SG* 1-*O*-sinapoyl-*β*-glucose; *SC* sinapoylcholine (sinapine)

BnSCT1 expression is controlled by a seed specific promoter

Transgenic *Arabidopsis* lines were generated carrying *BnSCT1* promoter::*GUS* gene fusion cassettes for histochemical analysis of *BnSCT1* promoter activity. Six independent lines transformed with the *GUS* gene under control of a 729 bp *BnSCT1* promoter fragment showed *GUS* expression exclusively in seeds without marked variation in strength (not shown). GUS activity was evident in both the embryo and the surrounding tissues (Fig. 3). Control plants transgenic for the promoterless *GUS* gene did not reveal any detectable reporter gene expression. Microscopic analysis of embryo sections showed that *BnSCT1* promoter activity revealed by *GUS* expression was uniformly distributed in cotyledons and the hypocotyl. In the surrounding tissues consisting of the single layer of peripheral endosperm cells (aleurone layer) and seed coat, the *BnSCT1* promoter activity was

Table 1 Contents of sinapoylglucose and sinapine (mg g^{-1} seed) in selected T2 seed samples ($n = 20$) of *Arabidopsis* Col-0 wild type transformed with the empty vector pLH7000 and the *sng2* mutant line SALK_002255 expressing *BnSCT1* from pLHBnSCT1

Arabidopsis line	Sinapoylglucose				Sinapoylcholine			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Col-0/pLH7000	0.71	0.66	0.3	1.8	10.77	0.88	8.5	11.7
SALK002255/pLH7000	11.33	1.59	8.7	14.4	ND			
SALK002255/pLHBnSCT1	4.31	2.74	2.6	10.3	9.42	2.74	3.5	13.0

ND not detected

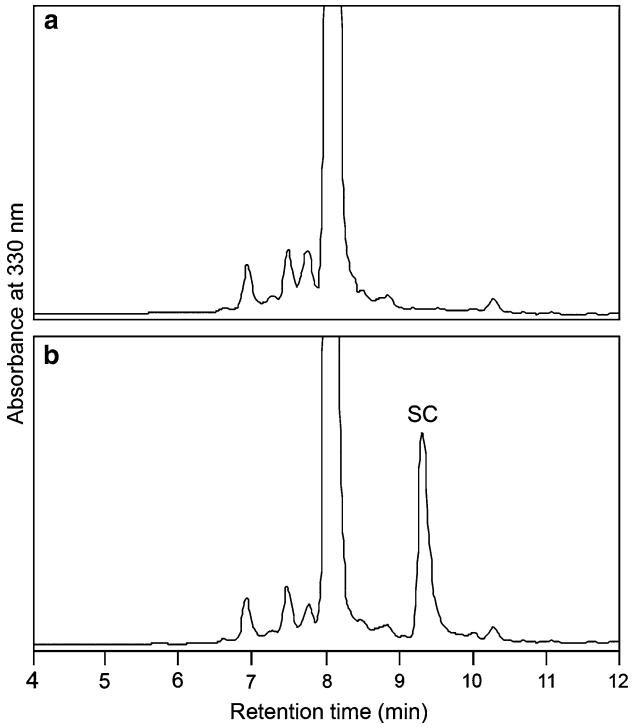


Fig. 2 Heterologous expression of *BnSCT1* in *Nicotiana tabacum*, cv. Samsun. HPLC diagrams of enzyme assays showed the formation of sinapine (*SC*) in the transgenic host (**b**) whereas sinapine synthesis was not observed when extracts from host tissues harboring the empty vector were analyzed (**a**). For SCT activity, a mean value of 20.6 ± 2.1 pkat mg^{-1} protein was calculated from assaying protein extracts of five infected leaf areas

restricted to the aleurone cell layer. These results confirm that *BnSCT1* gene expression is driven by an inducible promoter that is specific for the tissues derived from double fertilization.

BnSCT1 imparts dsRNAi suppression of *SNG2* in *Arabidopsis*

We used a partial cDNA sequence of 332 bp from *BnSCT1* starting with the translation initiation codon to assemble a dsRNAi suppression cassette fused to the seed-specific napin promoter. The resulting construct was used to transform *Arabidopsis*. Seeds produced by the first transgenic generation (T2 seeds) were tested for sinapate ester accumulation by HPLC analysis. The frequency distribution for the sinapine content showed a clear shift to lower values in the transgenic offspring harboring the *BnSCT1* suppression cassette (*BnSCTi* plants) compared to the control plants transformed with the empty vector (Fig. 4a). Results revealed that 88% of the transgenic *BnSCTi* plants displayed reduced sinapine levels in comparison to the control plants. The strongest suppression was indicated by a reduction in sinapine content of 48% compared to the control plant showing the lowest level (Table 2). The reduced sinapine content was counterbalanced by a shift to higher contents of sinapoylglucose in the *BnSCTi* plants (Fig. 4b).

Characterization of the genes *BnSCT1* and *BnSCT2*

BnSCT gene fragments were amplified from genomic DNA of allotetraploid *B. napus* and the diploid ancestors *B. oleracea* and *B. rapa* by PCR with primers (5′SCTBn; 3′SCTBn) recognizing the 5′ and 3′ terminal sequences of the previously identified *BnSCT* reading frame (AY383718). Analysis of the cloned PCR products revealed for *B. napus* two sequence variants designated as *BnSCT1* (Accession number AM706349; 2.875 bp) and *BnSCT2* (AM706350; 2.864 bp) whereas from both *B. rapa* and *B. oleracea* genomes only one respective sequence type could be isolated—*BrSCT* (AM706348; 2.864 bp) and *BoSCT* (AM706347; 2861 bp), respectively. The four identified *Brassica* SCT sequences show similar exon-intron patterns each consisting of 14 exons and 13 introns. A comparison analysis revealed that the *Brassica SCT* gene structure resembled that of the *Arabidopsis* homologue *SNG2* (Shirley et al. 2001) except for the second intron, which covers more than twice as much nucleotides in the *Brassica* genes than in *Arabidopsis* (Fig. 5).

Sequence information indicated that the genomic *BnSCT1* fragment carried the coding sequence previously identified as *BnSCT* (AY383718; Milkowski et al. 2004). *BnSCT1* represents the *Brassica* C genome homologue

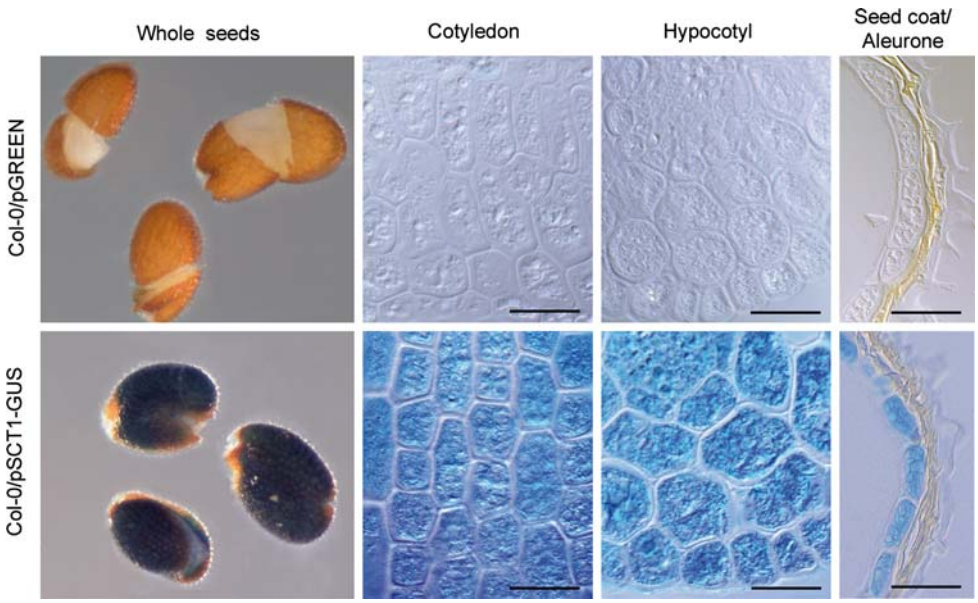


Fig. 3 Histochemical staining for GUS activity in seeds of transgenic *Arabidopsis* lines. Whole seeds were perforated to facilitate infiltration of GUS solution. Blue staining indicates GUS gene expression in embryonic tissues and aleurone layer of plants transformed with the *BnSCT1* promoter-GUS gene fusion (Col-0/pSCT1-GUS) whereas in control plants carrying the empty vector (Col-0/pGREEN) no stain was detectable. *Bars* represent 20 µm

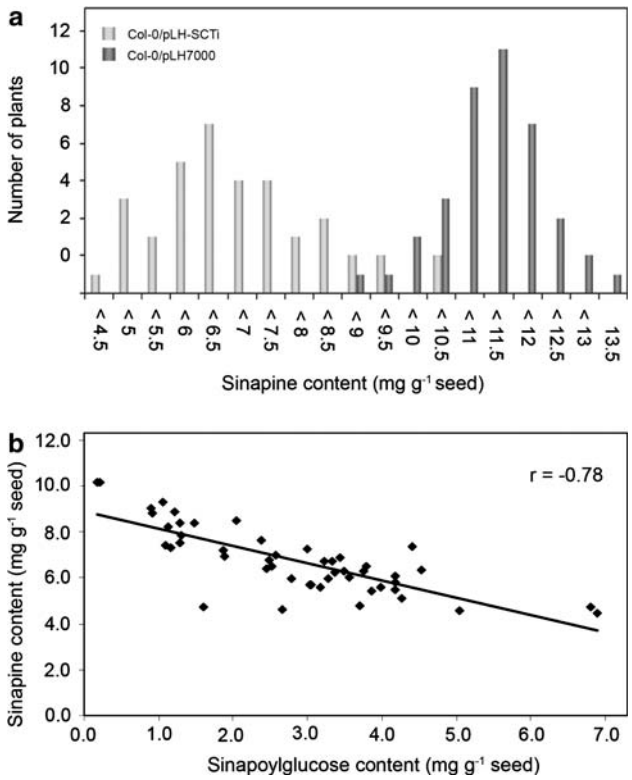


Fig. 4 Suppression of sinapine biosynthesis in *Arabidopsis* by *BnSCT1*-dsRNAi Frequency distribution for the sinapine content in T2 seeds of transgenic *Arabidopsis* plants carrying the *BnSCT1*-dsRNAi construct pLH-SCTi or the empty vector pLH7000 (**a**). The reduction of the sinapine content was accompanied by an increase in the amount of sinapoylglucose as was corroborated by the inverse correlation between sinapine and sinapoylglucose accumulation in the transgenic population (**b**)

of SCT provided by *B. oleracea* whereas *BnSCT2* is derived from *B. rapa* carrying the *Brassica* A genome (Fig. 6).

Table 2 Contents of sinapoylglucose and sinapine (mg g^{-1} seed) in selected T2 seed samples ($n = 50$) of BnSCTi (Col-0/pLH-SCTi) and control (Col-0/pLH7000) plants

Arabidopsis line	Sinapoylglucose		Sinapoylcholine	
	Min	Max	Min	Max
Col-0/pLH7000	0.51	2.89	8.52	14.93
Col-0/pLH-SCTi	0.17	6.89	4.43	10.17

Genomic structure of *BnSCT* loci

Screening of a genomic BAC library from *B. napus* with *BnSCT1* cDNA as a probe resulted in the isolation of 47 positive BAC clones. By PCR with primers recognizing the cDNA ends of both *BnSCT1* and *BnSCT2*, fragments were amplified with eight of these BACs as template indicating a full length *BnSCT* gene on each of these BAC clones. Gene discrimination was achieved by PCR with specific primers binding selectively to the 3′UTR of either *BnSCT1* or *BnSCT2*. Results revealed that six of the pre-selected *BnSCT* BACs carried the gene *BnSCT1* whereas two harbored *BnSCT2* (data not shown). Subcloning led to the identification of a 6.969 bp *Eco*RI fragment designated as BnSCT-BAC5 that represents the group of *BnSCT1* BACs and a 4.526 bp *Nco*I fragment termed BnSCT-BAC4 for the *BnSCT2* BACs (Fig. 7). Sequence analysis led to the identification of partial open reading frames adjacent to *BnSCT1* and *BnSCT2* with sequence identities to *Arabidopsis* genes in the range of 80–90% indicating orthology (Schmidt 2002). On BnSCT-BAC5, upstream from *BnSCT1*, a partial sequence of the gene *sdh3* encoding a succinate dehydrogenase subunit of the mitochondrial respiratory complex II (Figueroa et al. 2002) was detected. The sequence identity

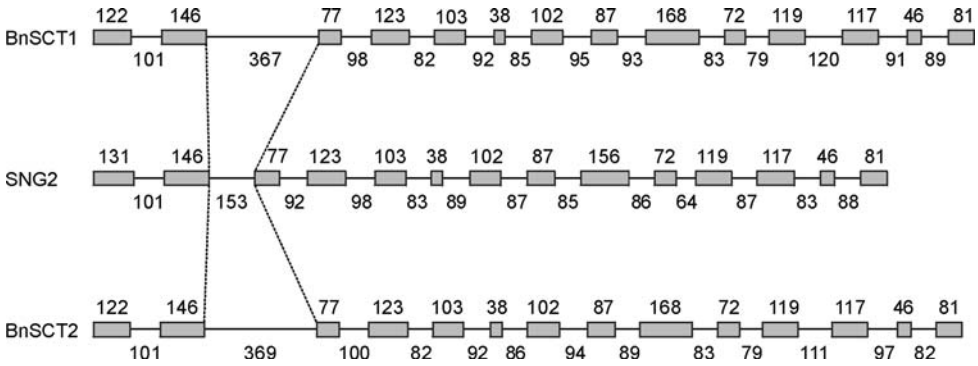


Fig. 5 Structure of *BnSCT1* and *BnSCT2* genes and the *Arabidopsis* homologue *SNG2*. Exons are depicted by *boxes*. The sizes of exons and introns are given as numbers of nucleotides. Genes are drawn to scale. Intron 2 differing remarkably in size between *Brassica* and *Arabidopsis* is highlighted

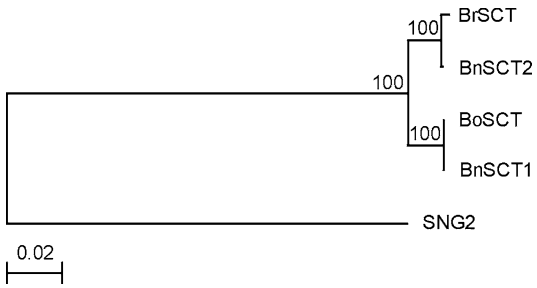


Fig. 6 Distance tree of *SCT* genes from *Arabidopsis* and Brassicaceae plants. A multiple alignment of nucleotide sequences covering the genes from translation initiation to the stop codon was generated by CLUSTALW. The tree was derived by neighbor-joining distance analysis using Treecon (van de Peer and de Wachter 1994). Bootstrap values are indicated. The *scale* represents 0.02 fixed mutations per site. *BrSCT*, *B. rapa SCT* (Accession number AM706348); *BnSCT2*, *B. napus SCT2* (AM706350, nucleotide positions 1007–3872); *BoSCT*, *B. oleracea SCT* (AM706347); *BnSCT1*, *B. napus SCT1* (AM706349, nucleotide positions 2135–5010)

starts in intron1 of *sdh3* and includes exon 2, intron 2 and part of exon 3. Upstream from this *sdh3* subsequence, the 3′ sequence portion of *HSC70-5* encoding a heat shock protein (Sung et al. 2001) was identified. Downstream from *BnSCT1* we detected a partial sequence of the *PPA* gene coding for an inorganic pyrophosphatase (Schulze et al. 2004). This incomplete *PPA* gene copy was found to consist of exons 2–6 lacking the first exon. The upstream region of this partial *PPA* is characterized by the presence of the 3′ portion of *B. napus mdh1*gene (Accession number AJ242713), the homologue of *Arabidopsis mbNAD-MDH* (At5g09660) encoding microbody NAD-dependent malate dehydrogenase. Sequence analysis of the BnSCT-BAC4 fragment revealed differences in the genomic microstructure compared to BnSCT-BAC5. In detail, the *PPA* partial sequence is located closer to the *BnSCT2* gene than to *BnSCT1*, and the alignment with the *Arabidopsis PPA* breaks within exon 5, i.e. BnSCT-BAC4 carries a significantly smaller part of *PPA*.

Comparison of BnSCT-BAC5 and BnSCT-BAC4 with the homologous genomic region from *Arabidopsis* chromosome 5 reveals that even for these partial genes colinearity has been conserved (Fig. 7). However, at the *Arabidopsis SNG2* locus, three additional genes (At5g09610-30) have been detected interspersed between *mtHSC70* and *SNG2*. These genes of which no function has been described yet,

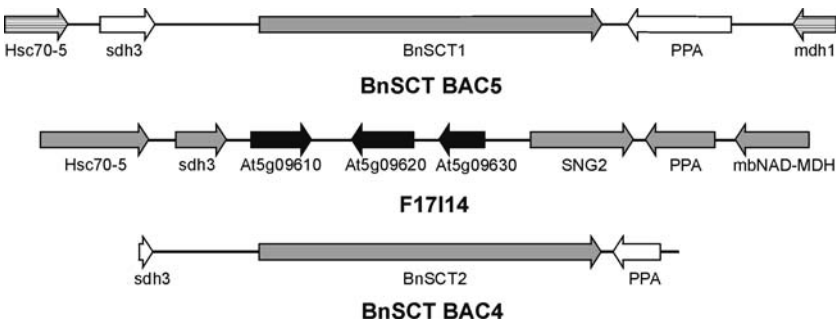


Fig. 7 Microstructures of the *BnSCT1*and *BnSCT2* genomic regions from *B. napus* and the homologous *SNG2* locus from *Arabidopsis*. The 6.969 bp BAC5 genomic fragment from *B. napus* (Accession number AM706349) shown in the *upper panel* carries the 3′ portion of *mtHSC70* (nucleotide positions 1-531), a subsequence from *sdh3* (758-1264), the *BnSCT1* gene (2135-5010), a partial *PPA* sequence (5215-6098), and the 3′ portion of *MDH* (6870-6969). The 4.526 bp BAC4 genomic fragment from *B. napus* (AM706350) shown in the *bottom panel* carries a subsequence of *sdh3* (1-117), the *BnSCT2* gene (1007-3872) and a *PPA* partial sequence (3973-4368). The homologous genomic fragment from *Arabidopsis* is drawn according to BAC clone F17I14 (AL353994), nucleotide positions 73000-53170. Genes are shown as pentagons pointing in the direction of transcription. *White boxes* represent partial genes. 3′ gene portions produced by BAC restriction are marked by *striped boxes*. For *Arabidopsis* genes *shaded in black*, homologues could not be detected in the genomic *B. napus* regions. DNA fragments are not drawn to scale

are absent from both homologous *B. napus* loci. The microstructural diversity found between the *Arabidopsis SNG2* genomic locus and the homologous *B. napus* regions reflects rearrangements during genome evolution that most probably have led to the loss of coding sequences including entire genes upstream from *BnSCT1* and *BnSCT2*.

Discussion

In *B. napus*, sinapine remains a principal antinutritive factor and thereby compromises the use of the valuable seed protein in animal feeding and prevents it from being regarded as human food supplement (Sozulski 1979). Hence, a substantial decrease of sinapine content would be of far reaching economic impact. Recently, a strong reduction of the seed sinapine level was achieved by seed-specific suppression of *UGT84A9* encoding UDP-glucose:sinapate glucosyltransferase (Hüsken et al. 2005). However, it becomes increasingly evident that further improvement of crop plants will depend not only on the manipulation of single metabolic steps but on the engineering of whole metabolic networks. This requires a detailed knowledge of the full complement of genes involved, including expression regulation, enzyme localization and the direction of metabolite fluxes. Since *B. napus*—like most of the crop plants—has been characterized as a polyploid organism, genome organization provides an additional layer of complexity. In this context, it is of crucial importance to identify the various genomic loci of a gene and to prove the extent to which each of them contributes to the trait of interest. Here we describe the genes *BnSCT1* and *BnSCT2* from *B. napus* encoding the enzyme 1-*O*-sinapoyl-β-glucose:choline sinapoyltransferase that catalyzes the final step of sinapine biosynthesis.

B. napus is an allotetraploid species ($2n = 38$, genome constitution AACC), which has evolved through hybridization and polyploidization between the two diploid species *B. rapa* ($2n = 20$, AA) and *B. oleracea* ($2n = 18$, CC) (U 1935; Schenck and Röbbelen 1982). A previous Southern Blot analysis of these *Brassica* genomes with *BnSCT1* cDNA as probe revealed two *BnSCT* loci corresponding to single loci in *B. rapa* and *B. oleracea* (Milkowski et al. 2004). PCR amplification of genomic *SCT* sequences from *B. napus*, *B. rapa* and *B. oleracea* as well as screening of a genomic BAC library from *B. napus* confirmed this view. We could isolate two *SCT* genes from *B. napus* named *BnSCT1* and *BnSCT2* and one corresponding genomic *SCT* sequence from each *B. rapa* and *B. oleracea*. Sequence analysis of a set of *BnSCT* cDNAs from seeds of *B. napus* detected both *BnSCT1* and *BnSCT2* sharing about 99% of identical nucleotides in the exon regions. As functionality of *BnSCT1* became evident from heterologous expression and enzyme assay, the sequence and expression data reveal that *BnSCT2* encodes a second enzymatically active SCT, and both genes contribute to sinapine biosynthesis during seed maturation. With regard to transgenic strategies for establishing the low sinapine trait, the high level of sequence identity should afford effective dsRNAi suppression of both *BnSCT1* and *BnSCT2* by a single dsRNAi cassette. For conventional breeding based on TILLING approaches, the number of active gene loci is one critical point. Southern Blot analyses and sequence inspection corroborate the presence of only two *BnSCT* gene loci in the genome of *B. napus*.

Sequence analysis of BAC subfragments revealed an insight into the microstructure of *BnSCT1* and *BnSCT2* genomic loci. The comparison with the homologous *SNG2* locus of *Arabidopsis* (Shirley et al. 2001) confirmed the general picture of a considerable microsyntenic colinearity of the *Arabidopsis* and *Brassica* genome segments (Sadowski et al. 1996; Cavell et al. 1998) that, however, includes minor deletions, insertions and translocations. Besides some subtle distinctions in spacing regions between genes, the most striking difference between the *Arabidopsis SNG2* locus and the *BnSCT* loci is the absence of three complete genes, the homologues of AT5g09610–30, from the *Brassica* loci and the loss of considerable sequence portions rendering the genes contiguous to *BnSCT*, *PPA* and *SDH3*, nonfunctional. The absence of *Arabidopsis* homologues from *Brassica* genomic loci has been reported in previous studies. One example is the genomic segment carrying *ABI1-Rps2-Ck1* on *Arabidopsis* chromosome 4 (Quiros et al. 2001). None of the isolated homologous *B. napus* loci carried all three genes. For *Brassica* genomes, mapping experiments revealed a triplicated structure (Lagercrantz and Lydiate 1996). More detailed studies indicated that in any one of the triplicated *Brassica* regions, one or several homologues of the *Arabidopsis* genes may be missing (Schmidt 2002).

In previous studies we characterized BnSCT as a seed-specific enzyme that is strongly induced at the transcript level when the seed enters the maturation phase (Milkowski et al. 2004). Seed specificity of expression has also been reported for the *Arabidopsis* homologue of *BnSCT1*, the *SNG2* gene (Fraser et al. 2005). GUS fusion experiments presented in this work indicate that the 729 bp sequence upstream from the translation start is sufficient to mediate seed specificity of *BnSCT1* expression. GUS expression driven by the *BnSCT1* promoter was found to be evenly distributed in the embryo and the aleurone layer. Similar spatial and temporal expression patterns have been described for genes involved in the accumulation of seed storage reserves as proteins and lipids. Seed specificity of these genes is dependent on distinct promoter *cis* elements like the RY/G complex and the B-box (Bäumlein et al. 1986;

Ellerström et al. 1996), which interact with transcription factors of the B3 family like FUSCA3 (FUS3; Luerßen et al. 1998) and ABSCISIC ACID INSENSITIVE3 (ABI3; Giraudat et al. 1992). Together with the B3 transcription factor LEAFY COTYLEDON2 (LEC2; Meinke et al. 1994; Stone et al. 2001) and the HAP3-like factor LEAFY COTYLEDON1 (LEC1; Meinke 1992; Meinke et al. 1994) these proteins regulate seed development and therefore might be involved in *BnSCT1* expression as well. This hypothesis is corroborated by the finding that seed-specific expression of the *Arabidopsis* homologue *SNG2* was found to be severely reduced in the *Arabidopsis fus3* mutant genetic background (data not shown). However, sequence analysis of the 729 bp *BnSCT1* upstream segment neither revealed a complete RY element in an appropriate position nor other *cis* elements described for seed specificity. Future experiments including promoter deletion and transient expression assays in host cells providing distinct transcription factors (Reidt et al. 2000) will help to elucidate the molecular mechanisms mediating seed specificity of the *BnSCT1* promoter.

In conclusion, as sinapine constitutes by far the highest concentration of *B. napus* seed SE compounds, suppression of its biosynthesis appears to be an effective means to reduce the agronomically undesirable seed SE content. Initial dsRNAi experiments in *Arabidopsis* indicate that silencing of *SCT* genes could be used to specifically reduce the seed sinapine content. However, since in the transgenic lines the decrease of sinapine was accompanied by a corresponding increase in the metabolic precursor sinapoylglucose, the method turned out to be ineffective for a complete suppression of SE accumulation. Thus, it becomes obvious that the most substantial decrease of SE content will require a combination of different approaches. Nevertheless, *B. napus* lines suppressing the *BnSCT* genes will be an important resource for future breeding programs that should involve genetic crossing with lines suppressing the biosynthetic pathway toward sinapate esters at different metabolic steps.

Acknowledgments The authors thank the Norddeutsche Pflanzenzucht Hans-Georg Lembke KG (NPZ), Hohenlieth, for supply of seed samples. This study was part of the research project "NAPUS 2000—Healthy Food from Transgenic Rape Seeds" and was financially supported by the Bundesministerium für Bildung und Forschung (BMBF). We thank Bettina Hause for helpful discussions during the course of this work and Claudia Horn and Sylvia Vetter for excellent technical assistance.

References

Baumert A, Milkowski C, Schmidt J, Nimtz M, Wray V, Strack D (2005) Formation of a complex pattern of sinapate esters in *Brassica napus* seeds, catalysed by enzymes of a serine carboxypeptidase-like acyltransferase family. Phytochemistry 66:1334–1345

Bäumlein H, Wobus U, Pustell H, Kafatos FC (1986) The legumin gene family: structure of a B-type gene of *Vicia faba* and a possible legumin gene specific regulatory element. Nucleic Acids Res 14:2707–2720

Bell JM (1993) Factors affecting the nutritional value of canola meal: a review. Can J Anim Sci 73:679–697

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254

Cavell AC, Lydiate DJ, Parkin IA, Dean C, Trick M (1998) Collinearity between a 30-centimorgan segment of *Arabidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome. Genome 41:62–69

Chen PY, Wang CK, Soong SC, To KY (2003) Complete sequence of the binary vector pBI121 and its application in cloning T-DNA insertion from transgenic plants. Mol Breed 11:287–293

Chuang C-F, Meyerowitz EMO (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. Proc Natl Acad Sci USA 97:4985–4990

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J16:735–743

D Halluin K, De Block M, Denecke J, Janssens J, Leemans J, Reynaerts A, Botterman J (1992) The *bar* gene as selectable and screenable marker in plant engineering. Methods Enzymol 216:415–426

Ellerström M, Stålberg K, Ezcurra I, Rask L (1996) Functional dissection of a napin gene promoter: identification of promoter elements required for embryo and endosperm specific transcription. Plant Mol Biol 32:1019–1027

Figueroa P, Léon G, Elorza A, Holuigue L, Araya A, Jordana X (2002) The four subunits of mitochondrial respiratory complex II are encoded by multiple nuclear genes and targeted to mitochondria in *Arabidopsis thaliana*. Plant Mol Biol 50:725–734

Fraser CM, Rider LW, Chapple C (2005) An expression and bioinformatics analysis of the *Arabidopsis* serine carboxypeptidase-like gene family. Plant Physiol 138:1136–1148

Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM (1992) Isolation of the *Arabidopsis ABI3* gene by positional cloning. Plant Cell 4:1251–1261

Hause B, Demus U, Teichmann C, Partier B, Wasternack C (1996) Developmental and tissue-specific expression of JIP-23, a jasmonate-inducible protein of barley. Plant Cell Physiol 37:641–649

Hause B, Meyer K, Viitanen PV, Chapple C, Strack D (2002) Immunolocalization of 1-*O*-sinapoylglucose:malate sinapoyltransferase in *Arabidopsis thaliana*. Planta 215:26–32

Hausmann L, Töpfer R (1999) Entwicklung von Plasmid-Vektoren. BioEngineering für Rapssorten nach Maß. Vorträge für Pflanzenzüchtung 45:155–173

Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM (2000) pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. Plant Mol Biol 42:819–832

Hood EE, Gelvin SB, Melchers LS, Hoekema A (1993) New agrobacterium helper plasmids for gene transfer to plants. Transgenic Res 2:208–218

Hüsken A, Baumert A, Strack D, Becker HC, Möllers C, Milkowski C (2005) Reduction of sinapate ester content in transgenic oilseed rape (*Brassica napus*) by dsRNAi-based suppression of *BnSGT1* gene expression. Mol Breed 16:127–138

Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions:β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6:3901–3907

Kapila J, De Rycke R, Van Montagu M, Angenon G (1997) An agrobacterium-mediated transient gene expression system for intact leaves. Plant Sci 122:101–108

Kesy LM, Bandurski RS (1990) Partial purification and characterization of insol-3-ylacetylglucose:*myo*-inositol indol-3-ylacetyltransferase

(indoleacetic acid-inositol synthase). Plant Physiol 94:1598–1604

Kridl JC, McCarter DW, Rose RE, Scherer DE, Knutzon DS, Radke SE, Knauf VC (1991) Isolation and characterization of an expressed napin gene from *Brassica rapa*. Seed Sci Res 1:209–219

Lagercrantz U, Lydiate DJ (1996) Comparative genome mapping in *Brassica*. Genetics 144:1903–1910

Lehfeldt C, Shirley AM, Meyer K, Ruegger MO, Cusumano JC, Viitanen PV, Strack D, Chapple C (2000) Cloning of the *SNG1* gene of *Arabidopsis* reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. Plant Cell 12:1295–1306

Li AX, Steffens JC (2000) An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. Proc Natl Acad Sci USA 97:6902–6907

Li AX, Eannetta N, Ghangas GS, Steffens JC (1999) Glucose polyester biosynthesis. Purification and characterization of a glucose acyltransferase. Plant Physiol 121:453–460

Lorenzen M, Racicot V, Strack D, Chapple C (1996) Sinapic acid ester metabolism in wild type and a sinapoylglucose-accumulating mutant of *Arabidopsis*. Plant Physiol 112:1625–1630

Luerßen H, Kirik V, Hermann P, Miséra S (1998) The *FUSCA3* gene of *Arabidopsis thaliana* encodes a product with partial homology to VP1/ABI3-like regulatory proteins. Plant J 15:755–764

McBride KE, Summerfelt KR (1990) Improved binary vectors for agrobacterium-mediated plant transformation. Plant Mol Biol 14:269–276

Meinke DW (1992) A homeotic mutant of *Arabidopsis thaliana* with leafy cotyledons. Science 258:1647–1650

Meinke DW, Franzmann LH, Nickle TC, Yeung EC (1994) *Leafy Cotyledon* mutants of *Arabidopsis*. Plant Cell 6:1049–1064

Milkowski C, Baumert A, Schmidt D, Nehlin L, Strack D (2004) Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. Plant J 38:80–92

Mock H-P, Strack D (1993) Energetics of the uridine 5′-diphosphoglucose: hydroxycinnamic acid acyl-glucosyltransferase reaction. Phytochemistry 32:575–579

Mock H-P, Vogt T, Strack D (1992) Sinapoylglucose:malate sinapoyltransferase activity in *Arabidopsis thaliana* and *Brassica napus*. Z Naturforsch 47C:680–682

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15:473–497

Nair RB, Bastress KL, Ruegger MO, Denault JW, Chapple C (2004) The *Arabidopsis thaliana REDUCED EPIDERMAL FLUORESCENCE1* gene encodes an aldehyde dehydrogenase involved in ferulic acid and sinapic acid biosynthesis. Plant Cell 16:544–554

Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1–6

Ohlson R (1978) Functional properties of rapeseed oil and protein product. In: Proceedings 5th International Rapeseed Congress, Malmö, Sweden, pp 152–167

Outchkourov NS, Peters J, de Jong J, Rademakers W, Jongsma MA (2003) The promoter–terminator of chrysanthemum rbcS1directs very high expression levels in plants. Planta 216:1003–1012

Quiros CF, Grellet F, Sadowski J, Suzuki T, Li G, Wroblewski T (2001) *Arabidopsis* and *Brassica* comparative genomics: sequence, structure and gene content in the *ABI-Rps2-Ck1* chromosomal segment and related regions. Genetics 157:1321–1330

Reidt W, Wohlfarth T, Ellerström M, Czihal A, Tewes A, Ezcurra I, Rask L, Bäumlein H (2000) Gene regulation during late embryogenesis: the RY motif of maturation-specific gene promoters is a direct target of the FUS3 gene product. Plant J 21:401–408

Ruegger M, Chapple C (2001) Mutations that reduce sinapoylmalate accumulation in *Arabidopsis thaliana* define loci with diverse roles in phenylpropanoid metabolism. Genetics 159:1741–1749

Sadowski J, Gaubier P, Delseny M, Quiros CF (1996) Genetic and physical mapping in *Brassica* diploid species of a gene cluster defined in *Arabidopsis thaliana*. Mol Gen Genet 251:298–306

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

Schenck HR, Röbbelen G (1982) Somatic hybrids by fusion of protoplasts from *Brassica oleracea* and *B.* campestris. Z Pflanzenzücht 89:278–288

Schmidt R (2002) Plant genome evolution: lessons from comparative genomics at the DNA level. Plant Mol Biol 48:21–37

Schulze S, Mant A, Kossmann J, Lloyd JR (2004) Identification of an *Arabidopsis* inorganic pyrophosphatase capable of being imported into chloroplasts. FEBS Lett 565:101–105

Shahidi F, Naczk M (1992) An overview of the phenolics of canola and rapeseed: chemical, sensory and nutritional significance. J Am Oil Chem Soc 69:917–924

Sheahan JJ (1996) Sinapate esters provide greater UV-B attenuation than flavonoids in *Arabidopsis thaliana* (Brassicaceae). Am J Bot 83:679–686

Shirley AM, Chapple C (2003) Biochemical characterization of sinapoylglucose:choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism. J Biol Chem 278:19870–19877

Shirley AM, McMichael CM, Chapple C (2001) The *sng2* mutant of *Arabidopsis* is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose:choline sinapoyltransferase. Plant J 28:83–94

Sozulski F (1979) Organoleptic and nutritional effects of phenolic compounds on oilseed protein products: a review. JAOCS 56:711–715

Stone SL, Kwong LW, Yee KM, Pelletier J, Lepiniec L, Fischer RL, Goldberg RB, Harada JJ (2001) *LEAFY COTYLEDON2* encodes a B3 domain transcription factor that induces embryo development. Proc Natl Acad Sci USA 98:11806–11811

Strack D, Knogge W, Dahlbender B (1983) Enzymatic synthesis of sinapine from 1-*O*-sinapoylglucose and choline by a cell-free system from developing seeds of red radish (*Raphanus sativus* L. var. sativus). Z Naturforsch 38c:21–27

Sung DY, Vierling E, Guy CL (2001) Comprehensive expression profile analysis of the *Arabidopsis* Hsp70 gene family. Plant Physiol 126:789–800

Tkotz N, Strack D (1980) Enzymatic synthesis of sinapoyl-L-malate from 1-sinapoylglucose and L-malate by a protein preparation from *Raphanus sativus* cotyledons. Z Naturforsch 35c:835–837

U N (1935) Genomic analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilisation. Jpn J Bot 7:389–452

van de Peer Y, de Wachter R (1994) TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput Appl Biosci 10:569–70

van Engelen FA, Moltho JW, Conner AJ, Nap JP, Pereira A, Stiekema WJ (1995) pBinplus: an improved plant transformation vector based on pBin19. Transgenic Res 4:288–290

The Plant Journal (2008) **53**, 802–813 doi: 10.1111/j.1365-313X.2007.03374.x

Role of a GDSL lipase-like protein as sinapine esterase in Brassicaceae

Kathleen Clauß[1], Alfred Baumert[1], Manfred Nimtz[2], Carsten Milkowski[1] and Dieter Strack[1,*]
[1]*Leibniz Institute of Plant Biochemistry (IPB), Weinberg 3, D-06120 Halle (Saale), Germany, and*
[2]*Helmholtz Centre for Infection Research (HCI), Inhoffenstraße 7, D-38124 Braunschweig, Germany*

Received 8 October 2007; accepted 6 November 2007.
*For correspondence (fax +49 345 5582 1509; e-mail dieter.strack@ipb-halle.de).

Summary

The seeds of most members of the Brassicaceae accumulate high amounts of sinapine (sinapoylcholine) that is rapidly hydrolyzed during early stages of seed germination. One of three isoforms of sinapine esterase activity (BnSCE3) has been isolated from *Brassica napus* seedlings and subjected to trypsin digestion and spectrometric sequencing. The peptide sequences were used to isolate *BnSCE3* cDNA, which was shown to contain an open reading frame of 1170 bp encoding a protein of 389 amino acids, including a leader peptide of 25 amino acids. Sequence comparison identified the protein as the recently cloned BnLIP2, i.e. a GDSL lipase-like protein, which displays high sequence identity to a large number of corresponding plant proteins, including four related Arabidopsis lipases. The enzymes belong to the SGNH protein family, which use a catalytic triad of Ser-Asp-His, with serine as the nucleophile of the GDSL motif. The corresponding *B. napus* and Arabidopsis genes were heterologously expressed in *Nicotiana benthamiana* leaves and proved to confer sinapine esterase activity. In addition to sinapine esterase activity, the native *B. napus* protein (BnSCE3/BnLIP2) showed broad substrate specificity towards various other choline esters, including phosphatidylcholine. This exceptionally broad substrate specificity, which is common to a large number of other GDSL lipases in plants, hampers their functional analysis. However, the data presented here indicate a role for the GDSL lipase-like BnSCE3/BnLIP2 as a sinapine esterase in members of the Brassicaceae, catalyzing hydrolysis of sinapine during seed germination, leading, via 1-*O*-sinapoyl-β-glucose, to sinapoyl-L-malate in the seedlings.

Keywords: sinapate ester metabolism, sinapine esterase, *Brassica napus,* Arabidopsis, GDSL lipase, molecular evolution.

Introduction

Germinating seeds of members of the Brassicaceae family express high activity of an esterase that catalyzes hydrolysis of the major seed phenylpropanoid component sinapine (sinapoylcholine) to liberate sinapate and choline (Nurmann and Strack, 1979; Strack *et al.*, 1980; Tzagoloff, 1962). Based on the substrate specificity the enzyme has been classified as a 'sinapine esterase' (SCE; EC 3.1.1.49). Sinapate released by *in vivo* SCE activity is re-esterified in the young seedlings, via 1-*O*-sinapoyl-β-glucose, with L-malate to form the UV-shielding compound sinapoyl-L-malate, which accumulates in epidermal tissues (Landry *et al.*, 1995; Sheahan, 1996). The free choline feeds into phosphatidylcholine biosynthesis (Strack, 1981).

Formation of sinapoyl-L-malate is catalyzed by the enzymes UDP-glucose:sinapate glucosyltransferase (SGT; EC 2.4.1.120) and 1-*O*-sinapoyl-β-glucose:L-malate sinapoyltransferase (SMT; EC 2.3.1.92). The same mechanism has been described for sinapine formation in developing seeds involving the action of SGT and 1-*O*-sinapoyl-β-glucose:choline sinapoyltransferase (SCT; EC 2.3.1.91). Recently, a detailed study on molecular regulation during sinapate ester metabolism was performed in *Brassica napus* (Milkowski *et al.*, 2004).

Initial studies on molecular evolution of the enzymes involved in sinapate ester metabolism in the Brassicaceae family indicated that SMT (Lehfeldt *et al.*, 2000) and SCT (Milkowski *et al.*, 2004; Shirley *et al.*, 2001) belong to a new class of serine carboxypeptidase-like (SCPL) proteins (Milkowski and Strack, 2004) originally defined by isobutyroyl transferases from *Lycopersicon penellii* (Li and Steffens,

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

2000). It has been speculated that SCE might also share the characteristics of SCPL proteins (Fraser *et al.*, 2005), but differing from the acyltransferases in retaining its hydrolytic activity. This assumption could not be verified in the present study.

Here we describe separation of three protein fractions exhibiting SCE activity from germinating seeds of *B. napus*, i.e. BnSCE1, BnSCE2 and BnSCE3. The protein with the highest activity (BnSCE3) was isolated and used for peptide sequencing followed by cDNA cloning. The deduced amino acid sequence of the full-length clone identified BnSCE3 as the recently cloned BnLIP2 (GenBank accession no. AY870270; Ling *et al.*, 2006), a member of the large protein family of GDSL lipases (Akoh *et al.*, 2004). However, its lipase activity and role in plant metabolism have not yet been experimentally proven. In our hands, BnSCE3/BnLIP2 showed, in addition to activity with sinapine and other hydroxycinnamate choline esters, a broad substrate specificity towards various other phenolic choline esters, and, surprisingly, also towards phosphatidylcholine. Our data suggest that the GDSL-like protein previously described as BnLIP2 has either gained a specific role as sinapine esterase during evolution of plant secondary metabolism in Brassicaceae, or accepts sinapine and other phenolic choline esters due to its inherent broad substrate specificity.

Results and discussion

Purification of BnSCE

Protein extracts from 2-day-old *Brassica napus* seedlings express high 'sinapine esterase' activity (Milkowski *et al.*, 2004). To purify the 'sinapine esterase', 18 g of ammonium sulphate-precipitated protein from these seedlings was subjected to a combination of chromatographic separation steps, including adsorption, ion-exchange and size-exclusion techniques (Table 1). Figure 1 shows a typical elution profile from Q-Sepharose, by which three protein fractions with esterase activity towards sinapine were separated, i.e. BnSCE1, BnSCE2 and BnSCE3. The protein fraction with the highest enzymatic activity, BnSCE3, was purified to near homogeneity with a 2300-fold enrichment and a yield of 5% of the extracted enzyme activity. Final chromatography on Mono-Q of the BnSCE3 fraction yielded 400 μg protein, from which aliquots were resolved into two bands on SDS–PAGE (Figure 2). The molecular mass of the major band was estimated to be about 44 kDa. Gel filtration of this protein on a Superose 12 HR 10/30 column gave a mass value of 47 kDa, indicating that the enzyme has a monomeric structure. The very faint second band in lane 8 has been tentatively ascribed to a glycosylated isoform of the major one.

Table 1 Purification scheme of BnSCE from *B. napus*

Purification step	Total protein (mg)	Total activity (nkat)	Specificity (nkat/mg)	Enrichment (fold)	Yield (%)
BnSCE1/BnSCE2/BnSCE3					
Ammonium sulphate	18 000	1080	0.06	1	100
Phenylsepharose	8400	773	0.092	1.5	72
Q-Sepharose					
BnSCE1	2610	117	0.045	0.75	11
BnSCE2	823	123	0.15	2.5	11
BnSCE3	656	366	0.56	9.3	34
BnSCE3					
Sephadex G-75	91	295	3.24	54	27
Q-Sepharose	30	182	6.1	102	17
Phenylsepharose	3.0	103	34	566	9.5
Superdex G-75	1.8	86	48	800	8.0
Mono-Q	0.40	55	138	2300	5.1

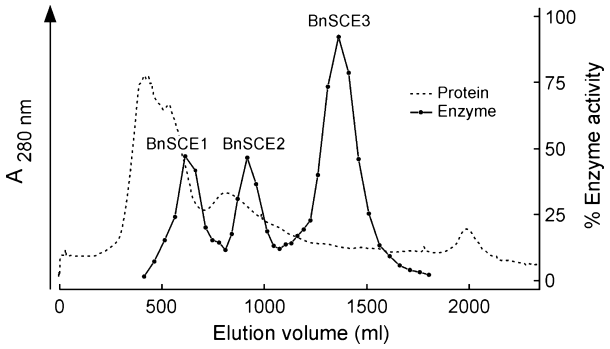


Figure 1. Ion-exchange chromatography of BnSCE proteins on Q-Sepharose resulting in separation of three sinapine esterase activities: BnSCE1, BnSCE2 and BnSCE3.

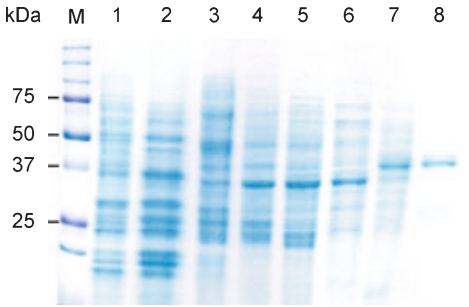


Figure 2. SDS-PAGE analysis of the BnSCE purification steps from two-day-old *B. napus* seedlings. *M*, Precision Plus Protein™ Standards Dual Color, BIO-RAD; *lane 1*, crude protein extract; *lane 2*, protein solution after Phenylsepharose FF 50/20; *lane 3*, Q-Sepharose 50/20; *lane 4*, Sephadex G-75 50/60; *lane 5*, Q-Sepharose 16/10; *lane 6*, Phenylsuperose 5/5; *lane 7*, Superdex G-75 16/60; *lane 8*, Mono-Q.

cDNA cloning and sequence analysis

To obtain sequence information for the purified BnSCE3, the major protein band in lane 8 of the SDS–PAGE

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd,

(Figure 2) was excised from the Coomassie-stained gel and subjected to trypsin digestion and mass spectrometric sequencing. Sequences of five peptides were determined, and were used to derive degenerate primers for RT-PCR with RNA from *B. napus* seedlings. This approach led to the amplification of a 471 bp cDNA carrying the 3′ portion of an open reading frame. The 5′ cDNA sequence was amplified by 5′ RACE using poly(A^+) RNA extracted from 2-day-old seedlings and a set of gene-specific primers (R1, R2, R3, R4; see Experimental procedures for sequences). PCR using one of these gene-specific primers (R2) led to the amplification of a 1037 bp cDNA fragment that had a 57 bp sequence overlap with the 471 bp 3′ cDNA fragment and a putative in-frame translation initiation codon near the 5′ end. The full-length *BnSCE3* cDNA was isolated from seedling RNA by RT-PCR using specific primers derived from the 5′ RACE fragment and the 471 bp partial sequence. The resulting cDNA was shown to contain an open reading frame that consists of 1170 bp and encodes a protein of 389 amino acid residues. The derived amino acid sequence carried all sub-sequences determined by peptide sequencing of the purified BnSCE3 protein (Figure 3). Sequence-based prediction of the molecular mass of the mature enzyme revealed a value of 40.46 kDa, and the isoelectric point (pI) was 4.69. The calculated mass differs by about 4–7 kDa from the estimated values from SDS–PAGE and gel filtration. This indicates that the native enzyme is a glycoprotein, which is in accordance with the potential *N*-glycosylation sites identified in the amino acid sequence (see below).

Sequence comparison by tblastn (http://www.ncbi.nlm.nih.gov/BLAST/) identified the BnSCE3 protein as the previously submitted GDSL lipase-like BnLIP2 (accession no. AY871275), based on a sequence identity of 97%. The cDNA was recently cloned from germinating seeds of *B. napus*, named BnLIP2 and mistakenly published as BnLIP1 (Ling *et al.*, 2006), a homologue of the real *BnLIP1* (Ling *et al.*, 2006b). The database entry describes BnLIP2 as a putative lipase, but this is based exclusively on sequence identities without any experimental evidence.

The protein domain architecture derived from the deduced amino acid sequence of *B. napus* BnSCE3/BnLIP2 was examined by computer-assisted protein family analysis using the database Pfam (http://www.sanger.ac.uk/Software/Pfam/). The results showed that the sequence contains a conserved GDSL lipase family domain spanning residues 35–361. A GDSL/I motif was detected between amino acids 39 and 42 (Figure 3).

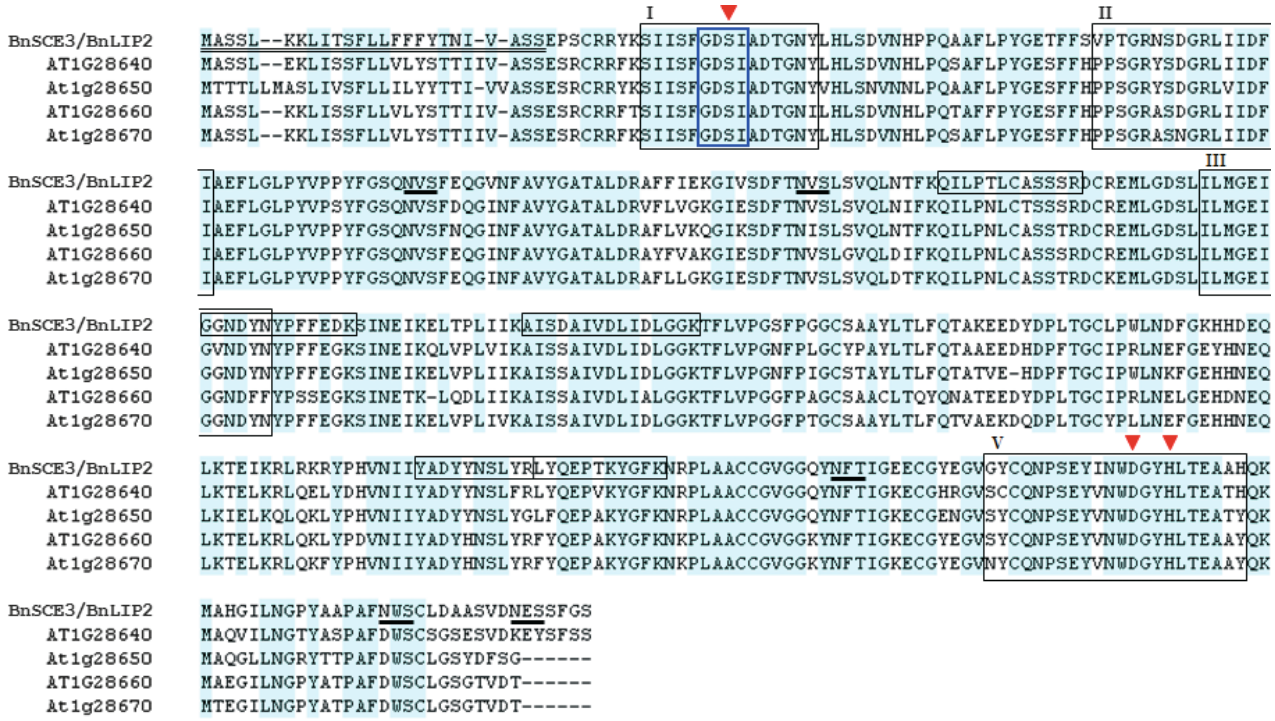


Figure 3. Multiple alignment of deduced amino acid sequence of BnLIP2/BnSCE3 with four putative Arabidopsis GDSL lipases (At1g28640, At1g28650, At1g28650, At1g28670). The N-terminal leader peptide (first 25 amino acids) of BnSCE3/BnLIP2 is double underlined. Fully conserved residues are shaded in blue. Four conserved blocks in the SGNH-hydrolases family are boxed by black lines (I, II, III, V; Akoh *et al.*, 2004; Ling *et al.*, 2006). Amino acid residues forming the catalytic triad (Ser, Asp, His) in the consensus sequences of blocks I and V are marked by red triangles. One of the Gly residues in block II and Asn in block V potentially may also act as catalytic residues. The GDSL motif (blue framed box) and four other sub-sequences (black framed boxes) were identified by mass spectrometric sequencing after trypsin digestion. Putative *N*-glycosylation sides are underlined.

© 2008 The Authors

Inspection of the deduced BnSCE3/BnLIP2 amino acid sequence using the sorting server TargetP (Emanuelsson *et al.*, 2000) predicts that the N-terminus of the protein contains a signal sequence of 25 residues, which routes the protein to the secretory pathway with a probability of 0.983 (>0.800 = strongest prediction; Figure 3). This is in accordance with a probable localization of the assumed natural substrate sinapine to vacuoles of germinating seeds. In the amino acid sequence, five potential *N*-glycosylation sites, NX(S/T), were identified (Figure 3).

GDSL lipases constitute a subclass of lipolytic enzymes characterized by a distinct GDSL sequence motif that is different from the GXSXG motif found in many lipases (Akoh *et al.*, 2004). Within the GDSL subclass, BnSCE3/BnLIP2 belongs to the subgroup of SGNH hydrolases that are characterized by the presence of the four strictly conserved residues Ser, Gly, Asn and His in four conserved peptide blocks (Ling *et al.*, 2006). The active centre of these enzymes is formed by a catalytic triad consisting of the seryl residue from the GDSL motif, which is part of conserved peptide block I, and aspartyl and histidyl residues that are both located in conserved peptide block V. This implies that, in BnSCE3/BnLIP2, the catalytic triad is most probably composed of Ser41, Asp345 and His348, i.e. Ser16-Asp320-His323 of the mature protein (Figure 3).

Role of GDSL lipases in plant metabolism

The sequence identity of BnSCE3 with BnLIP2 was the first indication that the 'sinapine esterase' is related to lipases instead of being a member of the serine carboxypeptidase-like (SCPL) protein family as previously speculated based on expression analysis of the Arabidopsis SCPL gene family (Fraser *et al.*, 2005). Given the broad substrate specificity of GDSL lipases, which may be due to the remarkable flexibility of the active site of the protein (Akoh *et al.*, 2004), we conclude that members of this enzyme family may have been recruited during evolution for functions in plant secondary metabolism such as sinapine hydrolysis.

Protein sequences related to BnSCE3/BnLIP2 were found in *Arabidopsis thaliana*, *Oryza sativa*, *Triticum aestivum* and *Zea mays*, with identities of up to 82%. A multiple alignment of the deduced amino acid sequences of BnSCE3/BnLIP2 and four putative Arabidopsis GDSL lipases is shown in Figure 3. These putative Arabidopsis GDSL lipases display the highest sequence identities to BnSCE3/BnLIP2, with scores of 672 for At1g28670, 655 for At1g28640, 642 for At1g28650 and 634 for At1g28660, and E-values of zero. All four Arabidopsis homologues are arranged in tandem repeats on chromosome 1.

The GDSL lipase subclass includes bacterial and a large number of plant proteins that have little sequence identity to true lipases (Akoh *et al.*, 2004). Their function in seed germination and plant development remains unknown, and functional analysis is hampered by their exceptional broad substrate specificity (Ling *et al.*, 2006).

Transcript accumulation and heterologous expression of BnSCE3/BnLIP2 *and its Arabidopsis homologues*

Upon measurement of SCE activity during *B. napus* seed germination, we found that the maximum value was reached in seedlings 2 days after sowing (Milkowski *et al.*, 2004), correlating well with a strong decrease in the content of sinapine as one of the assumed BnSCE substrates. This developmental regulation of enzyme activity is roughly reflected by the *BnSCE3/BnLIP2* transcript abundance, which showed a sharp induction 2 days after sowing. Figure 4 shows results from RT-PCR analyses of *BnSCE3/BnLIP2* expression during *B. napus* seedling development, together with a modified diagram of SCE activity and sinapine degradation based on Milkowski *et al.* (2004). A high level of transcript was maintained until 6 days after sowing, followed by a decrease up to 12 days of seedling development. The decline in enzyme activity was found to be more rapid during these developmental stages, indicating that additional post-transcriptional regulation of *BnSCE3/BnLIP2* activity may occur. Moreover, the enzyme could be involved in other as yet unknown metabolic pathways in seedlings (Ling *et al.*, 2006), especially in the light of its expression

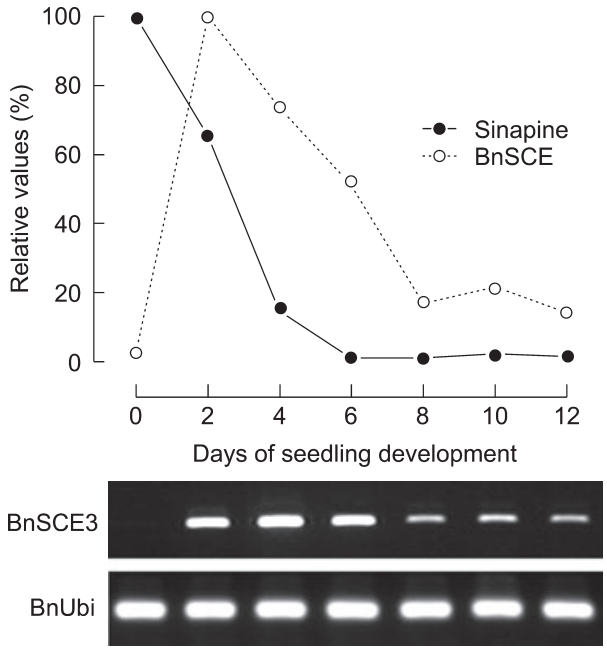


Figure 4. RT-PCR analysis of *BnLIP2/BnSCE3* expression along with a modified diagram after Milkowski *et al.* (2004) showing enzyme activities (100 = 511 pkat mg^{-1} protein) and amount of sinapine (100 = ca. 135 nmol $individuum^{-1}$) in *B. napus* seeds and developing seedlings. Expression of the *ubiquitin* gene is given as control. Staining was achieved with ethidium bromide.

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd,

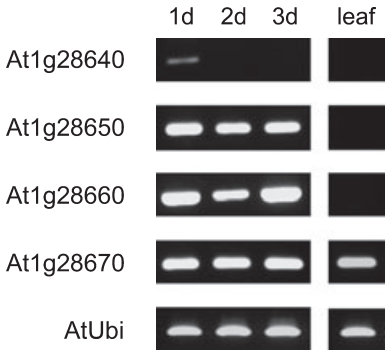


Figure 5. RT-PCR analysis of expression of four putative Arabidopsis GDSL lipases (*At1g28640*, *At1g28650*, *At1g28660*, *At1g28670*) in early seedling stages and leaf samples. Expression of the *ubiquitin* gene is given as control. Tissues analyzed were developing seedlings 1–3 days after sowing (1d, 2d, 3d) and adult leaves (leaf). Staining was achieved with ethidium bromide.

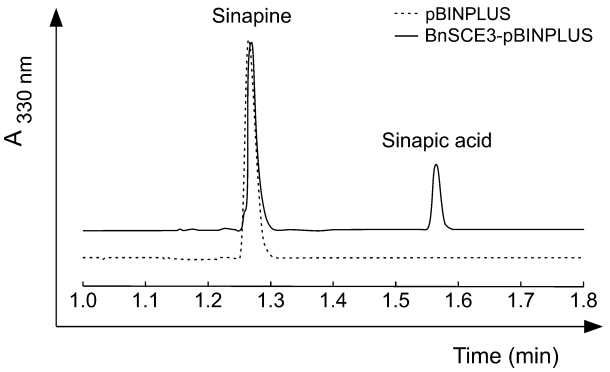


Figure 6. Traces of HPLC analyses of sinapine esterase assays run with protein extracts of *Nicotiana benthamiana* leaves harboring the empty vector pBINPLUS and the expression cassette BnSCE3-pBINPLUS.

in all mature *B. napus* tissues, including root, stem, leaf, bud, flower and pod (Ling *et al.*, 2006). In germinating seeds, BnSCE3/BnLIP2 may be involved in oil body mobilization. It has been shown, for example, that phospholipase A activity is associated with oil bodies in cotyledons of *Helianthus annuus* (Gupta and Bhatla, 2007) and *Cucumis sativus* (May *et al.*, 1998).

For comparison, the transcript accumulation of the four Arabidopsis *BnSCE3/BnLIP2* homologues encoding putative GDSL lipases was analyzed (Figure 5). During early seedling development, characterized by rapid sinapine degradation (Lorenzen *et al.*, 1996), we found high transcript levels for *At1g28650*, *At1g28660* and *At1g28670* in seedlings up to 3 days old. For the putative lipase gene *At1g28640*, weak transcript abundance was restricted to the very early stage of germination, i.e. 1 day after sowing, indicating that this gene may not be involved in sinapine hydrolysis. Only *At1g28670* was transcribed in seedlings as well as in adult leaves, in which sinapine is not detectable as a possible substrate. Thus, *At1g28670* and *BnSCE3/BnLIP2* may encode GDSL lipase-like proteins that are involved in lipid metabolism (Ling *et al.*, 2006) or other functions of the mature plant as discussed by Akoh *et al.* (2004).

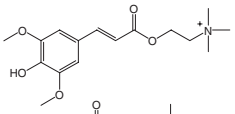
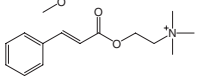
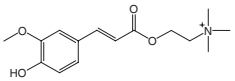
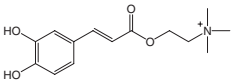
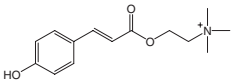
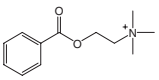
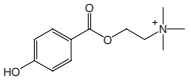
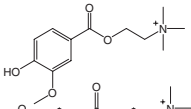
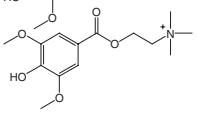
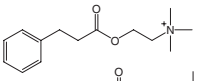
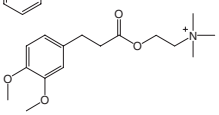
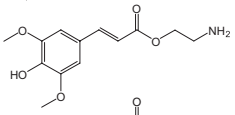
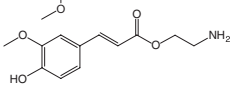
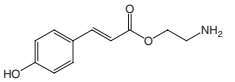
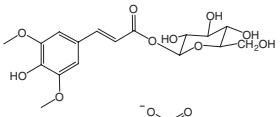
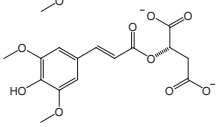
To provide experimental proof that the cloned cDNA from *B. napus* and the four putative Arabidopsis homologues encode enzymes that exhibit the expected sinapine esterase activity, cDNA fragments encoding the mature proteins were fused to the strong *rbcs* promoter and cloned into the binary vector pBINPLUS. Transiently transformed leaves of *Nicotiana benthamiana* (Kapila *et al.*, 1997) were subjected to protein extraction, and the presence of sinapine esterase activity was proven by incubating the protein with sinapine as substrate. Enzyme activities were found in extracts from *N. benthamiana* leaves expressing BnSCE3/BnLIP2 as demonstrated in Figure 6. The Arabidopsis proteins At1g28650, At1g28660 and At1g28670, expressed in these leaves, also showed esterase activity, with specific activities (nkat mg^{-1} protein) of 2.1, 27.1 and 9.0, respectively. No SCE activity was detected for At1g28640.

Substrate specificity of native BnSCE3/BnLIP2 and inhibition of enzymatic activity

To analyze the substrate specificity of BnSCE3/BnLIP2, the enzyme was assayed with 16 potential phenolic substrates (Table 2) and also with some lipids, e.g. phosphatidylcholine (Figure 7). The phenolic substrates were chosen to allow systematic comparison of the influence of ring substituents and side-chain structures, as well as the nature of the ester-bound moieties. The results summarized in Table 2 show that, among the hydroxycinnamate esters, the enzyme prefers sinapine as substrate, followed by cinnamoylcholine for BnSCE3 and feruloylcholine for BnSCE1 and BnSCE2. Shortening of the side chain of the phenylpropanoid substrates drastically reduced the hydrolytic activity. Compounds with a C_1 instead of a C_3 side chain, e.g. syringoylcholine, were found to be almost inactive as substrates. In contrast, elimination of the C_3 side-chain double bond, as in 3(3,4-dimethoxyphenyl)-propionylcholine and 3-phenylpropionylcholine, revealed much higher relative specific activity compared with that towards sinapine. Thus, 3-phenylpropionylcholine was hydrolyzed with 50-fold higher activity compared to the related structure cinnamoylcholine. Replacing the choline moiety by ethylamine led to a strong decrease in enzyme activity. These results indicate that specificity is directed towards the phenyl-C_3 part of the phenolic substrate and the conjugating moiety, predominantly choline. To resolve some of the contradictions in the substrate specificities listed in Table 2, the kinetic constants of BnSCE3/BnLIP2 for 3-phenylpropionylcholine in comparison with sinapine and cinna-

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd,

Table 2 Substrate specificity of BnSCE3/BnLIP2 towards various phenolic choline esters and related compounds

Substrate	Structure	Relative activity (%)		
		BnSCE1	BnSCE2	BnSCE3
(Hydroxy)cinnamoylcholines				
Sinapine (sinapoylcholine)		100 ± 28.4	100 ± 13.3	100 ± 2.3
Cinnamoylcholine		15 ± 0.6	12 ± 0.3	51 ± 2.0
Feruloylcholine		59 ± 15	40 ± 5.4	37 ± 0.9
Caffeoylcholine		33 ± 3.0	22 ± 1.8	25 ± 0.7
4-Coumaroylcholine		34 ± 3.0	22 ± 1.1	20 ± 2.1
(Hydroxy)benzoylcholines				
Benzoylcholine		2.8 ± 0.04	2.0 ± 0.2	18 ± 0.5
4-Hydroxybenzoylcholine		0.4 ± 0.2	0.4 ± 0.01	4.0 ± 0.4
Vanilloylcholine		ND	ND	1.0 ± 0.2
Syringoylcholine		ND	ND	ND
Phenylpropionylcholines				
3-Phenylpropionylcholine		300 ± 6.0	337 ± 46	2398 ± 515
3(3′,4-Dimethoxyphenyl)-propionylcholine		78 ± 12	189 ± 2.0	1422 ± 255
Hydroxycinnamoylethanolamines				
Sinapoylethanolamine		7 ± 0.01	14 ± 0.3	15 ± 0.4
Feruloylethanolamine		13 ± 0.4	10 ± 0.1	12 ± 0.6
4-Coumaroylethanolamine		3 ± 0.2	2.6 ± 0.3	7 ± 0.8
Other sinapoyl conjugates				
Sinapoylglucose		ND	ND	ND
Sinapoylmalate		ND	ND	ND

Data are the mean relative activity of three assays ± SD (n = 3); 100% = 0.3 nkat mg^{-1} with BnSCE1, 0.8 nkat mg^{-1} with BnSCE2, 139 nkat mg^{-1} with BnSCE3; ND, no product detected.

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd,

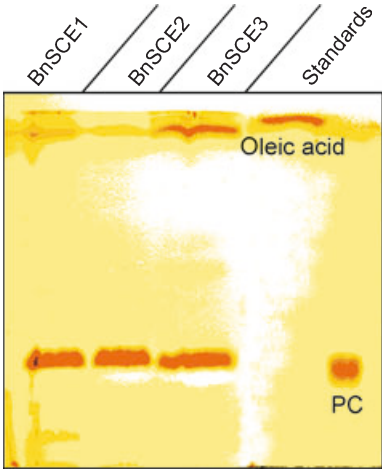


Figure 7. Thin layer chromatography (TLC) in hexan/diethyl ether/acetic acid (85:15:1) of phosphatidylcholine (PC) assays with protein extracts of BnSCE1 (184 μg), BnSCE2 (95 μg) and the purified BnSCE3 (1 μg). The applied standards were phosphadidylcholine (1,2-dioleoyl-PC) and oleic acid.

moylcholine were determined (Table 3). The data show that the high conversion rate of the enzyme with the phenylpropionyl conjugate corresponds to a V_{max} value that is 40 times higher than those obtained with the two other substrates. It should be noticed, however, that there is no difference in the catalytic efficiencies with sinapine and 3-phenylpropionylcholine, as the affinity of the protein (K_m) towards sinapine is 40 times higher than that for 3-phenylpropionylcholine.

Interestingly, two other non-choline sinapoyl conjugates, i.e. 1-*O*-sinapoyl-β-glucose and sinapoyl-L-malate, major phenolic components of the sinapate ester pathway in the seedlings (Milkowski *et al.*, 2004), were not accepted as substrates by BnSCE3/BnLIP2. This excludes a function of the enzyme as non-specific general esterase, and more specifically also as a cholinesterase (EC 3.1.1.8). The latter was also excluded for *Raphanus sativus* SCE activity (Nurmann and Strack, 1979), which did, however, hydrolyze sinapoylglucose. Our previous assumption that this activity may be caused by a contaminating hydrolase activity has thus been corroborated.

The results listed in Table 2 indicate that the presence of choline moieties is a basic structural feature recognized by BnSCE3/BnLIP2. Given the homology of this enzyme with lipases, we tested phosphatidylcholine as potential substrate in comparison with some neutral lipids, e.g. a digalactosyldiacylglycerol and a triacylglycerol. Thin-layer chromatography of the reaction mixtures showed that all three BnSCE proteins were able to hydrolyze phosphatidylcholine (1,2-dioleoyl-PC), as confirmed by the appearance of oleic acid as a reaction product (Figure 7). The neutral lipids tested were not accepted (data not shown). This result indicates a phospholipase A-like catalysis activity (Wang, 2001) for BnSCE3/BnLIP2.

Due to the sequence identity with GDSL lipases, the enzyme most probably uses a catalytic triad of Ser16-Asp320-His323 of the mature protein, with the seryl side chain as the nucleophile. To prove the involvement of a catalytic seryl residue in sinapine hydrolysis, the enzyme was subjected to treatment with phenylmethylsulfonylfluoride (PMSF), a potent inhibitor that phosphorylates the seryl residues of enzymes; this treatment has previously been applied to lipolytic GDSL proteins (e.g. Teissére *et al.*, 1995). Treatment of BnSCE3 with 1, 10 and 50 mM PMSF led to strong decreases in sinapine esterase activity, with 50% inhibition at about 8 mM PMSF (Figure 8). Thus, Ser16 of the mature protein of the GDSL motif may be part of the catalytic triad in BnSCE3/BnLIP2.

As the enzyme was active towards phosphatidylcholine (Figure 7), we addressed the question of whether this activity may be due to the broad substrate specificity of a phospholipase (Akoh *et al.*, 2004; Huang, 1993; Wang, 2001), or an extended function of a GDSL lipase as a 'sinapine esterase' in Brassicaceae. Therefore, we tested possible sinapine hydrolysis using commercially available phospholipase A2 from cobra venom (Sigma-Aldrich, München, Germany) and protein preparations from GDSL lipase-rich cotyledons of *H. annuus* and *C. sativus*, according to the methods described by Gupta and Bhatla (2007) and May *et al.* (1998), respectively. Phospholipase A2 did not accept sinapine as a possible substrate (results not shown), and, although the protein preparations showed hydrolytic activity towards sinapine, this was less than 1% of that determined for *B. napus* (Milkowski *et al.*, 2004). The specific function of BnSCE3/BnLIP2 as 'sinapine esterase' is further supported by results from an earlier study indicating a correlation between the occurrence of sinapine and esterase activity in members of the Brassicaceae (Strack *et al.*, 1980).

Substrate	K_m (μM)	V_{max} (nkat mg^{-1})	V_{max}/K_m (nkat mg^{-1} μM^{-1})	Efficiency (%)
Sinapine (sinapoylcholine)	2.6 ± 0.9	66 ± 3.3	25.4	100
Cinnamoylcholine	11 ± 0.9	59 ± 1.4	5.4	21
3-Phenylpropionylcholine	101 ± 14	2525 ± 124	25.0	98

Data were obtained and evaluated using an enzyme kinetics tool (SigmaPlot; Systat Software, Inc.).

Table 3 Apparent K_m and V_{max} values for the native BnSCE3/BnLIP2 towards selected choline esters tested for enzyme specificity (Table 2)

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd,

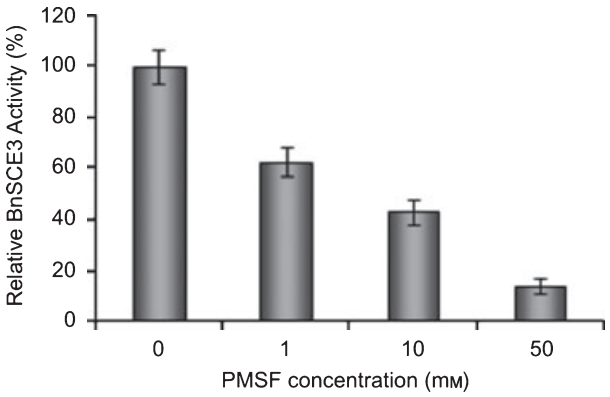


Figure 8. Inhibition of BnSCE3/BnLIP2 enzyme activity by the serine inhibitor PMSF; mean relative activity of three assays ± SD ($n = 3$); 100% = 232 nkat mg^{-1} protein.

In more applied studies focusing on the potential utilization of *B. napus* not only as oilseed crop but also as valuable protein supply in human nutrition, we have expressed *BnSCE3/BnLIP2* in developing seeds of transgenic lines. This may suppress the sinapine accumulation that compromises the use of the seed meal as human food supplement (Ismail *et al.*, 1981). In these studies, Arabidopsis plants were transformed with *Agrobacterium tumefaciens* harboring a *BnSCE3/BnLIP2* expression plasmid under the control of the seed-specific napin promoter. Our results show that sinapine accumulation in some of the transgenic lines was drastically reduced, as shown in Figure 9 for a line that showed a residual amount of sinapine that was less than 5% of that of control plants. In HPLC analyses of the transgenic seed extracts, sinapine could scarcely be detected, and there is no accumulation of another sinapate ester compensating for the strong suppression of sinapine accumulation. This result strongly supports the evidence presented in this paper that the GDSL lipase-like BnSCE3/BnLIP2 in *B. napus* is involved in sinapine hydrolysis.

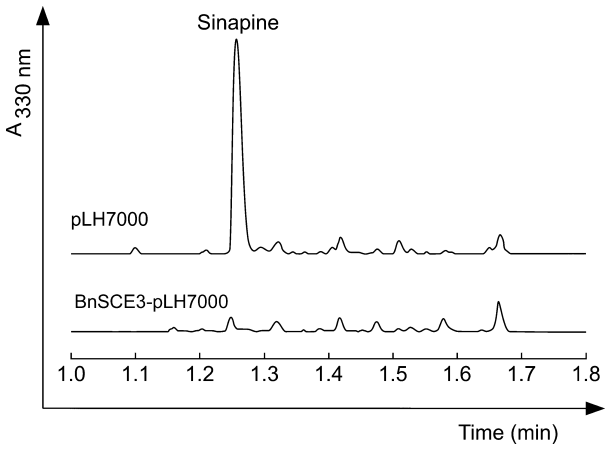


Figure 9. Traces of HPLC analyses of methanolic seed extracts from transformed Arabidopsis plants harboring the empty vector pLH7000 or the expression cassette *BnSCE3*-pLH7000. The seeds expressing the *BnSCE3* gene reveal a strong suppression of sinapine formation.

This result indicates that members of the GDSL lipase family have gained new functions in plant phenylpropanoid metabolism. Together with a UDP-glucose:sinapate glucosyltransferase (SGT) and two sinapoylglucose-dependent sinapoyltransferases (SMT, SCT), characterized as SCPL-like acyltransferases (Milkowski and Strack, 2004; Milkowski *et al.*, 2004; Stehle *et al.*, 2006) that catalyze the formation of sinapoyl-L-malate (Lehfeldt *et al.*, 2000) in the seedling cotyledons and sinapine (Shirley *et al.*, 2001) in the seeds, these enzymes form the sinapate ester pathway in members of the Brassicaceae.

Further studies are required to determine the nature of the interactions of BnSCE3/BnLIP2 with choline esters, and the conformational changes due to substrate binding (Akoh *et al.*, 2004) according to the induced fit mechanism (Koshland, 1958). Elucidating the structures of the protein–ligand complexes, e.g. with hydroxycinnamoyl- and phenylpropionylcholines as well as phosphatidylcholine, will lead to an understanding of the factors that contribute to the plasticity and extraordinarily broad substrate specificity of GDSL lipases.

Experimental procedures

Plant material and growth conditions

Seeds of winter oilseed rape (*Brassica napus* L. var. *napus* cv. Express) were obtained from the Norddeutsche Pflanzenzucht (http://www.norddeutschepflanzenzucht.de). For enzyme purification, seeds were germinated for 2 days under greenhouse conditions at 12–18°C on moist filter paper in closed plastic boxes. Seedlings were grown for up to 12 days in the greenhouse under a 16 h light regimen. After harvest at various developmental stages, germinating seeds and seedlings were immediately frozen and stored at –80°C. Arabidopsis plants (*Arabidopsis thaliana* L. Heynh., wild-type Columbia) were grown in growth chambers on soil under a 16 h light regimen at 24°C. Transient heterologous expression of *BnSCE3* was performed in host plants of *Nicotiana benthamiana* L. grown under greenhouse conditions as described for *B. napus*.

Purification of BnSCE

All protein manipulations were performed at 4°C. Protein concentrations were determined according to the method described by Bradford (1976), using bovine serum albumin as standard. Purification steps were performed using an ÄKTAexplorer (Amersham Biosciences, http://www5.amershambiosciences.com/). The chromatographic columns were purchased from GE Healthcare Life Sciences (http://www.gehealthcare.com). The buffers used were buffer A, 0.1 M Tris–HCl, pH 7.5; buffer B, 0.02 M Tris–HCl, pH 7.5; buffer C, 0.02 M Tris–HCl, pH 7.5, 1 M ammonium sulphate; buffer D, 0.02 M Tris–HCl, pH 7.5, 1 M NaCl. Buffers were supplemented with 10% glycerol, 0.5 mM EDTA and 2 mM 2-mercaptoethanol. Purification of BnSCE3 involved eight chromatographic steps.

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd,

Step 1. Three kilograms of 2-day-old seedlings were ground to a fine powder in a mortar with liquid nitrogen. Buffer A was added, and the homogenate treated with an Ultra-Turrax homogenizer (Janke & Kunkel GmbH & Co. KG-IKA-Labortechnik (http://www.ika.net) and centrifuged at 27 485 ***g*** for 10 min. To the supernatant, 1% protamine sulfate (PS) solution was added to a final concentration of 0.05%, stirred for 15 min and centrifuged at 27 485 ***g*** for 10 min. The collected protein was precipitated at 80% ammonium sulphate saturation and collected by centrifugation at 27 485 ***g*** for 15 min.

Step 2. The resulting pellet was dissolved in buffer C, and the enzyme solution was applied in five runs to a phenyl Sepharose FF 50/20 column equilibrated with buffer C. Unbound protein was washed off with the same buffer. Bound protein was eluted with a 300 ml gradient from 0 to 70% buffer B in C, followed by 400 ml at 70% buffer B in C with a flow rate of 5 ml min^{-1}. Fractions (10 ml) were collected, and those containing enzyme activities were pooled, precipitated with ammonium sulphate at 80% saturation, and re-dissolved in buffer B.

Step 3. After desalting using a Sephadex G-25 column the protein solution was applied in 2 runs to a Q-Sepharose FF 50/20 column equilibrated with buffer B. Unbound protein was washed off using the same buffer. Bound protein was eluted by applying a 200-ml gradient from buffer 0 to 15% buffer D in B, followed by 1600 ml to 35% buffer D in B and 200 ml of 100% buffer D at a flow rate of 8 ml min^{-1}. Fractions (10 ml) were collected. Enzyme activity was detected in three fractions: BnSCE1 eluted between 500 and 700 ml, BnSCE2 between 800 and 950 ml, and BnSCE3 between 1200 and 1320 ml. These fractions were pooled separately, and the protein was precipitated with ammonium sulphate at 80% saturation and re-dissolved in buffer B.

Step 4. Purification was continued using fraction BnSCE3, which was applied to a Sephadex G-75 50/60 column equilibrated with buffer B. Bound protein was eluted with the same buffer at a flow rate of 5 ml min^{-1} and 10 ml fractions were collected. The esterase activity eluted between 350 and 450 ml. Active fractions were pooled and concentrated by ultrafiltration through Centriprep-10 (Millipore; http://www.millipore.com).

Step 5. The protein was applied to a HiLoad 16/10 Q-Sepharose HP anion-exchange column and eluted with a 100-ml gradient from 0 to 20% buffer D in B, followed by a 250-ml gradient from 20 to 40% buffer D in B at a flow rate of 3 ml min^{-1} (5 ml fractions). The active fractions that eluted between 175 and 230 ml were pooled, concentrated by ultrafiltration through Centripep YM-10 Centrifugal Filter Units (Millipore) and re-dissolved in buffer B.

Step 6. Buffer B was replaced by buffer C and the enzyme solution was then applied to a phenylsepharose HR 5/5 column that had been equilibrated with buffer C. Elution was achieved with a 60 ml linear gradient from buffer C to 100% buffer B at a flow rate of 0.5 ml min^{-1}. Fractions containing enzyme activities were pooled and concentrated by ultrafiltration through Centripep YM-10 Centrifugal Filter Units (Millipore).

Step 7. The concentrated protein was loaded at a flow rate of 1 ml min^{-1} onto a HiLoad 16/60 Superdex G-75 column equilibrated with buffer B. The collected fractions were assayed for BnSCE3 activity. Active fractions were concentrated by ultrafiltration through Centricon 10 Centrifugal Filter Units (Millipore).

Step 8. The enzyme was applied to a Mono-Q HR 5/5 column that had been equilibrated with buffer B, and eluted with a 60-ml gradient from 0 to 30% buffer D in B at a flow rate of 1 ml min^{-1}. Active fractions were concentrated, the buffer was replaced by 0.01 M aqueous ammonium sulphate, and the protein was lyophilized in 100 μg portions.

Assay for sinapine esterase activity

For standard assay of sinapine esterase activity, enzyme solutions were incubated for 10 min at 30°C in 100 mM Tricine buffer, pH 8.5, containing 250 μM sinapine thiocyanate isolated from *B. napus* seeds, and various chemically synthesized phenolic choline esters (see below) in a total volume of 200 μl. Reactions were terminated by the addition of 10 μl trifluoroacetate. After centrifugation, at 20 817 ***g*** for 2 min, enzyme activities were determined by liquid chromatography using an ACQUITY UPLC (Waters; http://www.waters.com). Aliquots of the reaction mixtures (2–5 μl) were injected onto an ACQUITY UPLC C18 (1.7 μm) column (50 mm × 2.1 mm internal diameter). The liberated phenolic moieties were analyzed by a 2-min elution from phosphoric acid (1.5% aqueous solution) with an acetonitrile linear gradient from 0 to 30% buffer D in B at a flow rate of 1 ml min^{-1}. The column eluates were monitored through max-plot UV spectroscopy by an ACQUITY uPLC Photodiode Array (PDA) Detector (Waters). Identification and quantification were achieved using sinapate and the phenolics used in choline ester syntheses as standards.

Determination of molecular mass

The molecular mass of native BnSCE3 was determined by FPLC on a Superose 12 HR 10/30 gel filtration column calibrated using a set of reference proteins including bovine serum albumin (67 kDa), chymotrypsinogen (25 kDa), ovalbumin (43 kDa) and ribonuclease A (13.7 kDa) from a gel filtration calibration kit (LMW, Amersham) (for chromatographic parameters, see step 7 of the protein purification protocol). The molecular mass of denatured BnSCE3 was determined by 10% w/v linear gradient SDS–PAGE (Laemmli, 1970) in comparison with a Bio-Rad molecular weight protein mixture (Precision Plus Protein Standard Dual Color; Bio-Rad Laboratories, http://www.bio-rad.com/). Proteins were visualized using Coomassie brilliant blue.

Mass spectrometric sequencing

The protein band on SDS–PAGE corresponding to BnSCE3 was excised, washed with water, reduced, carboxyamidomethylated and digested with trypsin, and the resulting peptides were extracted and purified according to standard protocols (e.g. Shevchenko *et al.*, 1996). A QTOF II mass spectrometer (Micromass; http://www.micromass.com) equipped with a nanospray ion source and gold-coated capillaries (Protana; http://www.protana.com) was used for electrospray MS/MS of peptides. For collision-induced dissociation (CID) experiments, multiple charged parent ions were selectively transmitted from the quadrupole mass analyzer into the collision cell (collision energy 25–35 eV for optimal fragmentation). The resulting daughter ions were then separated using the orthogonal TOF mass analyzer. These MS/MS spectra were enhanced (Max. Ent. 3; Micromass) and used for manual *de novo*

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd,

peptide sequencing using the PepSeq program (Micromass). The isomeric amino acids leucine and isoleucine could not be distinguished by the low-energy MS/MS techniques used.

RNA isolation

Extraction of total RNA from *B. napus* and Arabidopsis was performed by selective adsorption onto a silica matrix using a kit for plant RNA purification (Qiagen, http://www.qiagen.com/). Total RNA from 2-day-old *B. napus* seedlings was used to enrich poly(A^+) RNA by selective binding to oligo(dT) Oligotex beads (Qiagen). Purification of total RNA from germinating seeds, seedlings and leaves of Arabidopsis was done as described previously (Vicient and Delseny, 1999).

Isolation of BnSCE3 *cDNA, sequence analyses and alignment*

Three peptide sequences derived from MS sequencing of purified BnSCE3 were used (bold underlined) to derive degenerated primers (**NYADYYN**SLYR: 5′-AAYTAYGCYGAYTAYTACAAC-3′, 5′-AAYTAYGCRGAYTAYTACAAC-3′, 5′-AAYTAYGCYGAYTAYTACAAT-3′, 5′-AAYTAYGCRGAYTAYTACAAT-3′, 5′-AAYTAYGCYGAYTAYTATAAC-3′, 5′-AAYTAYGCRGAYTAYTATAAC-3′, 5′-AAYTAYGCYGAYTAYTATAAT-3′, 5′-AAYTAYGCRGAYTAYTATAAT-3′; G**GNDYNYP**FFEDK: 5′-GGYAAYGAYTAYAAYTACC-3′, 5′-GGRAAYGAYTAYAAYTACC-3′, 5′-GGYAAYGAYTAYAAYTATCC-3′, 5′-GGRAAYGAYTAYAAYTATCC-3′; L**YQEPTKYG**FK: 5′-TAYCARGARCCRACRAARTACGG-3′, 5′-TAYCARGARCCYACYAARTACGG-3′, 5′-TAYCARGARCCRACYAARTACGG-3′, 5′-TAYCARGARCCYACRAARTACGG-3′). The resulting oligonucleotides were used in reverse transcription reactions as forward primers, combined with a customized poly(dT) reverse primer and RNA extracted from *B. napus* seedlings 1 or 2 days after germination. RT-PCR was performed using the Omniscript reverse transcription kit (Qiagen) according to the manufacturer's protocol. Amplified cDNA fragments were subcloned into plasmid pGEM-T Easy (Promega, http://www.promega.com/) and subjected to sequence analysis. A 5′ incomplete 471 bp cDNA fragment with sequence similarity to putative lipases was used to derive gene-specific reverse primers (R2, 5′-GCCATCTTCTGGTGTGCGGCTTC-3′; R3, 5′-CCGTTGAGTATGCCATGAGCCATCTTCTG-3′; R4, 5′-GCAGCGTCAAGGCAGGACCAGTTG-3′) to amplify a full-length cDNA by 5′ RACE PCR from RNA from *B. napus* seedlings at 2 days of germination using the BD Smart Race cDNA amplification kit (BD Clontech, http://www.clontech.com/). Nested PCR was performed using 5′ nested reverse primer R1 (5′-CTCCAACTCCACAACAAGCAGCCAAAG-3′) and a 50-fold dilution of cDNA obtained from the first PCR run as template. The resulting products were cloned into pGEM-T Easy (Promega) and sequenced. The full-length cDNA was PCR-amplified using 5′ RACE-ready cDNA from *B. napus* seedlings at 2 days after germination as template and the primer pair BnSCEall_F (5′-ATGGCTTCTTCACTAAAGAAGCTTATC-3′) and BnSCEall_R (5′-TCAACTACCGAAAGAAGATTCGTTATC-3′). Sequence analysis and alignments of DNA and proteins were performed using the software package Clone Manager (Sci-Ed software; http://www.scied.com).

Semi-quantitative transcript analysis of BnSCE and Arabidopsis homologues

Reverse transcription reactions were performed with 1 μg of total RNA using the Omniscript reverse transcription kit (Qiagen) according to the manufacturer's protocol. Fifty-fold dilutions of cDNA were used for each PCR amplification performed using Platinum PCR SuperMix High Fidelity (Invitrogen, http://www.invitrogen.com/) and primer pair LIP2RT_F 5′-CAGTTGGATGTAGCGCCGCGTATC-3′ and LIP2RT_R 5′-CTTCTGGTGTGCGGCTTCGGTTATATG-3′ for amplification of *BnSCE3/BnLIP2*, and primer pairs NM631spez_F 5′-TGTTAACCACCTTCCTCAATCC-3′, NM631spez_R 5′-GTGAGATACGCGGGATAGCAT-3′; NM632spez_F 5′- AACGTGAGCTTCAATCAAGG-3′, NM632spez_R 5′-GATATGCCGTGGAACATCCTATTG-3′; NM633spez_F 5′-GATGTCGGCGCTTTACATCTATC-3′, NM633spez_R 5′-GATGATCAAATCCTGTAGTTTGGTTTC-3′ and NM634spez_F 5′-CCTAACTTATGCGCCTCGTC-3′, NM634spez_R 5′-CGTTGAGCAATGGGTAACAAC-3′ for amplification of the putative Arabidopsis lipases. Control experiments were performed using primer pairs recognizing *ubiquitin* expression: UbiRT_F 5′-ACTTCACTTGGTGCTCAGGC-3′ and UbiRT_R 5′-CCTTGACGTTGTCAATGGTG-3′ for Arabidopsis and UbiRT_F 5′-AAGCCAAGATCCAGGACAAAG-3′ and UbiRT_R 5′-CGAGCCAAAGCCATCAAAGAC-3′ for *B. napus*. The RT-PCR runs consisted of 25–40 cycles, depending on the linear range of PCR amplification for the individual genes. Each cycle included incubations at 94°C for 15 sec, 55°C for 30 sec and 72°C for 42 sec. An additional cycle at 72°C for 10 min was run after the last cycle to allow trimming of incomplete polymerizations. The gels were stained with ethidium bromide.

Heterologous expression of BnSCE3 *and Arabidopsis homologues in* Nicotiana benthamiana *leaves*

For heterologous expression, cDNAs were fused to the strong promoter of the Rubisco gene encoding the small subunit (*rbcs*) from *Asteraceous chrysanthemum* (Outchkourov *et al.*, 2003) by subcloning into plasmid pImpact1.1 (Plant Research International; http://www.pri.wur.nl/uk). Expression cassettes consisting of the *rbcs* promoter, coding sequence and *rbcs* transcription terminator were then transferred to the binary vector pBINPLUS (Plant Research International), and the resulting recombinant vectors were transformed into *Agrobacterium tumefaciens* strain GV2260 (McBride and Summerfelt, 1990).

The BnSCE coding sequence was PCR-amplified from cloned cDNA using primers BNSCErub_F (5′-ACTGGGATCCGAAAATGGCTTCTTCACTAAAGAAGCTTATC-3′) and BNSCErub_R (5′-GCAGCGGCCGCTCAACTACCGAAAGAAGATTCGTTATC-3′) designed to introduce an upstream *Bam*HI restriction site and a downstream *Not*I site (underlined) into the PCR product. The BnSCE coding sequence was inserted as a *Bam*HI–*Not*I fragment into plasmid pImpact1.1, from which the resulting expression cassette was isolated by *Asc*I–*Pac*I restriction and inserted into similarly cleaved pBINPLUS.

Coding sequences of the Arabidopsis genes *At1g28640*, *At1g28650*, *At1g28660* and *At1g28670* were PCR-amplified from seedling cDNAs using gene-specific primers (given below) designed to introduce an upstream recognition site for endonuclease *Xba*I and a downstream *Not*I site (*At1g28640*: 640rub_F (5′-ACTGTCTAGAGAAAATGGCTTCTTCACTGGAGAAGCTTATTTC-3′) and 640rub_R (5′-GCAGCGGCCGCTTAACTGCTGAAAGAATACTCTTTATCCACTGATTC-3′), *At1g28650*: 650rub_F (5′-ACTGTCTAGAGAAAATGACGACGACTCTCCTCATGG-3′) and 650rub_R (5′-GCAGCGGCCGCTTATCCACTGAAATCATAAGAGCCAAGG-3′), *At1g28660*: 660rub_F (5 ′-ACTGTCTAGAGAAAATGGCTTCTTCACTGAAGAAGCTTATC-3′) and 660rub_R (5′-GCAGCGGCCGCTTATGTATCCACTGTACCAGAGCCAAG-3′) *At1g28670*: 670rub_F (5′-ACTGTCTAGAGAAAATGGCTTCTTCACTGAAGAAGCTTATCTC-3′) and 670rub_R (5′-GCAGCGGCCGCTTATGTATCCACTGTACCAGAGCCAAGG-3′).

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd,

Coding sequences were inserted as *Xba*I–*Not*I fragments into pImpact1.1, and expression cassettes were transferred to the binary vector pBINPLUS as described above. PCR was performed using Platinum PCR SuperMix High Fidelity (Invitrogen). For *Nicotiana benthamiana* infiltration, recombinant agrobacteria were grown overnight in 5 ml YEB medium supplemented with 25 mg ml^{-1} rifampicin, 50 mg ml^{-1} kanamycin and 50 mg ml^{-1} carbenicillin at 28°C with gentle shaking. After addition of 20 μM acetosyringone, 10 mM glucose and 10 mM MES, pH 5.6, cultures were grown for another 24 h. Bacteria were then pelleted at 2755 ***g*** for 8 min and resuspended in a mixture of 1 ml YEB, 2 ml 2x infiltration medium (100 g l^{-1} saccharose, 3.6 g l^{-1} glucose and 8.6 g l^{-1} MS medium, pH 5.6) and 2 ml water to an OD_{600} of 1. After addition of 0.2 μM acetosyringon, the resuspended agrobacteria were used to infiltrate the lower part of *Nicotiana benthamiana* leaves using a 1 ml syringe (Roth; http://www.carl-roth.de). As a negative control, agrobacteria harboring the empty vector pBINPLUS were used. Transformed plants were cultivated for 5 days under greenhouse conditions. For protein extraction, 400 mg of inoculated leaf sectors were cut, and homogenized in 2.5 ml extraction buffer (0.1 M Tricine, pH 8.5). Cell debris was pelleted by centrifugation at 20 817 ***g*** for 5 min, and the cleared supernatant was desalted using a PD-10 column (Sephadex G25 matrix; GE Healthcare Life Sciences). The soluble proteins were concentrated using a Vivaspin concentrator (Viva-Science AG; http://www.vivascience.de).

Substrate specificity of the native BnSCE

The incubation mixture was the same as that used in the BnSCE assay. A total of 16 phenolic substrates (structures in Table 2) and some lipids, e.g. phosphatidycholine (Figure 8), were tested as possible substrates. 1-*O*-sinapoyl-β-glucose and sinapoyl-L-malate were from our collection of sinapate esters isolated from *B. napus* seedlings. The substrate concentration in the assay mixtures was 0.4 mM. To ensure reaction linearity, the amount of protein varied with regard to the enzyme activity for the various substrates. Each value for enzyme activity represents the mean of three determinations. Activities were related to that determined using sinapine. The kinetic parameters of the enzyme with sinapine, cinnamoylcholine and 3-phenylpropionylcholine as substrates were determined using Michaelis–Menten and Lineweaver–Burk plots (Segel, 1975), using the program SigmaPlot with the corresponding enzyme kinetics tool (Systat Software, Inc.; http://www.systat.de).

Synthesis of phenolic choline esters

The phenolic choline and ethanolamine esters were synthesized according to a slightly modified bromocholine bromide method (Clausen *et al.*, 1983). The phenolic acids (1 mmol) and bromocholine bromide (1.5 mmol) were suspended in 50 ml water. After addition of 2.5 mmol thriethylamine, the mixture was refluxed for 8 h. The solution was then concentrated in a rotary evaporator (BÜCHI Labortechnik; http://www.buechigmbh.de). The synthesized esters were purified from the reaction mixture by preparative HPLC (Waters) using a DELTAPAK C_{18} column (15 μm, 30 cm × 5 cm internal diameter) with 1% aqueous HOAc (solvent A) and MeCN (solvent B), using various gradients from A to (A + B) at a flow rate of 20 ml min^{-1}. The ester eluates were monitored using a diode array detector, evaporated to dryness and crystallized from acetone as acetate salts, yielding between 40% and 70%. These compounds, as well as sinapate liberated in the enzyme assays, were identified by mass spectrometry (data not shown).

Assay for lipolytic activity

The enzymatic assays for lipolytic activity of the three BnSCEs (BnSCE1, BnSCE2, BnSCE3) were carried out using phosphatidylcholine (1,2-dioleoyl-glycero-3-phosphocholine, 20 μl), a digalactosyldiacylglycerol (90 μl) and a triacylglycerol (20 μl) from the lipid collection of P. Dörmann (Max-Planck Institute, Golm, Germany). The organic phase of the lipid solutions was evaporated and the residue dissolved in diethyl ether (2 ml). Aliquots (10 μl) of the enzyme fractions (BnSCE1, 84 μg; BnSCE2, 95 μg; BnSCE3, 1 μg) were added and complemented with 490 μl lipase buffer containing 40 mM Tris–HCl (pH 7.2), 2.5 mM $CaCl_2$ and 50 mM H_3BO_3. As a control, water was used instead of the enzymes. The reactions were stirred for 20 min using a vortex. The organic phases were evaporated and the samples reduced to 500 μl. Lipid extraction was performed using 3 ml chloroform/methanol (2:1 v/v). After centrifugation at 1793 ***g*** for 4 min, the organic phases were evaporated and the residues resuspended in 40 μl chloroform/methanol (2:1 v/v), and separated by thin-layer chromatography (TLC) on Si250-PA silica gel (Baker; http://www.mallbaker.com): phosphatidylcholine and digalactosyldiacylglycerol with hexane/diethyl ether/acetic acid 85:15:1 (v/v/v); triacylglycerol with chloroform/methanol/ammonium hydroxide 65:35:5 (v/v/v). The compounds were visualized by exposure to iodine vapour.

Inhibition of BnSCE3/BnLIP2 activity

The serine protease inhibitor PMSF (Sigma-Aldrich) was added to the enzyme assay, and the mixture was incubated for 5 min at 30°C. Three inhibitor concentrations were tested (1, 10 and 50 mM). Assays were carried out with 0.5 μg of the purified enzyme. The reactions were stopped by adding 10 μl of trifluoroacetic acid (TFA), and the enzyme activity was determined using the HPLC-based standard sinapine esterase assay.

Expression of BnSCE3/BnLIP2 *in Arabidopsis seeds*

For expression of *BnSCE3/BnLIP2* in Arabidopsis seeds, the cDNA was fused to the seed-specific promoter of the *napin* gene from *B. napus* and cloned into plasmid pLH7000 (Hausmann and Töpfer, 1999). The expression cassette, consisting of the *nap* promoter, coding sequence and *nap* transcription terminator, was transferred into *Agrobacterium tumefaciens* strain EHA105 (Hood *et al.*, 1993).

The *BnSCE3/BnLIP2* coding sequence was PCR-amplified from cloned cDNA using primers BNSCEnap_F (5′-ACTGCCCGGGGAAAATGGCTTCTTCACTAAAGAAGCTC-3′) and BNSSCEnap_R (5′-GCAGGATCCTCAACTACCGAAAGAAGATTCGTTATC-3′), designed to introduce an upstream *Sma*I restriction site and a downstream *Bam*HI site (underlined) into the PCR product. PCR was performed using Platinum PCR SuperMix High Fidelity (Invitrogen). The coding sequence was inserted as an *Sma*I–*Bam*HI fragment into vector pLH7000.

For transformation of Arabidopsis buds, recombinant *A. tumefaciens* strain EHA105 was grown overnight in 4 ml LB medium supplemented with 25 mg ml^{-1} rifampicin and 100 mg ml^{-1} spectinomycin at 28°C with gentle shaking. The inoculated culture (400 ml LB medium) was grown to an OD_{600} of 0.8. The bacteria were then pelleted at 8000 ***g*** for 8 min and resuspended in a mixture of 1 vol 5% sucrose and 200 μl l^{-1} Silwet L-77 (Lehle Seeds; http://www.lehleseeds.com). The Arabidopsis buds were incubated with the bacterial solution for 2 min. As a control, bacteria harboring the empty vector pLH7000 were used. Transformed plants were cultivated for 2 days in a greenhouse.

© 2008 The Authors

Acknowledgements

We thank Jürgen Schmidt (Leibniz Institute of Plant Biochemistry, Halle (Saale), Germany) for analyses of the synthesized phenolic esters, and Peter Dörmann (Max-Planck Institute for Molecular Plant Physiology, Potsdam-Golm, Germany) for help in determination of lipolytic enzyme activity. Our thanks are also due to Thomas Hartmann (Braunschweig University) for critical reading of the manuscript. The seeds of *Brassica napus* were kindly provided by the Norddeutsche Pflanzenzucht (Holtsee, Germany). This study was supported by the Deutsche Forschungsgemeinschaft (Bonn, Germany).

References

Akoh, C.C., Lee, G.-C., Liaw, Y.-C., Huang, T.-H. and Shaw, J.-F. (2004) GDSL family of serine esterases/lipases. *Progr. Lipid Res.* **43**, 534–552.

Bradford, M.M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein–dye binding. *Anal. Biochem.* **72**, 248–254.

Clausen, S., Olsen, O. and Sørensen, H. (1983) Separation of aromatic choline esters by high-performance liquid chromatography. *J. Chromatogr.* **260**, 193–199.

Emanuelsson, O., Nielson, H., Brunak, S. and von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **300**, 1005–1016.

Fraser, C.M., Rider, L.W. and Chapple, C. (2005) An expression and bioinformatics analysis of the Arabidopsis serine carboxypeptidase-like gene family. *Plant Physiol.* **138**, 1136–1148.

Gupta, A. and Bhatla, S.C. (2007) Preferential phospholipase A_2 activity on the oil bodies in cotyledons during seed germination in *Helinathus annuus* L. cv. *Morden. Plant Sci.* **172**, 535–543.

Hausmann, L. and Töpfer, R. (1999) Entwicklung von Plasmid-Vektoren. BioEngineering für Rapssorten nach Maß. *Vortr. Pflanzenzüchtg.* **45**, 155–173.

Hood, E.E., Gelvin, S.B., Melchers, L.S. and Hoekema, A. (1993) New Agrobacterium helper plasmids for gene transfer to plants. *Transgenic Res.* **2**, 208–218.

Huang, A.H.C. (1993) Lipases. In *Lipid Metabolism in Plants* (Moore, T.S., ed.). Boca Raton, FL: CRC Press, pp. 473–503.

Ismail, F., Vaisey-Genser, M. and Fyfe, B. (1981) Bitterness and astringency of sinapine and its components. *J. Food Sci.* **46**, 1241–1244.

Kapila, J., De Rycke, R., Van Montagu, M. and Angenon, G. (1997) An *Agrobacterium*-mediated transient gene expression system for intact leaves. *Plant Sci.* **122**, 101–108.

Koshland, D.E. (1958) Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl Acad. Sci. USA*, **44**, 98–104.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Landry, L.G., Chapple, C.C.S. and Last, R. (1995) Arabidopsis mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. *Plant Physiol.* **109**, 1159–1166.

Lehfeldt, C., Shirley, A.M., Meyer, K., Ruegger, M.O., Cusumano, J.C., Viitanen, P.V., Strack, D. and Chapple, C. (2000) Cloning of the *SNG1* gene of Arabidopsis reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. *Plant Cell*, **12**, 1295–1306.

Li, A.X. and Steffens, J.C. (2000) An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. *Proc. Natl Acad. Sci. USA*, **97**, 6902–6907.

Ling, H., Zuo, K., Zhao, J., Qin, J., Qiu, C., Sun, X. and Tang, K. (2006) Isolation and characterization of a homologous to lipase gene from *Brassica napus*. *Russ. J. Plant Physiol.* **53**, 366–372.

Lorenzen, M., Racicot, V., Strack, D. and Chapple, C. (1996) Sinapic acid ester metabolism in wild type and a sinapoylglucose-accumulating mutant of Arabidopsis. *Plant Physiol.* **112**, 1625–1630.

May, C., Preisig-Müller, R., Höhne, M., Gnau, P. and Kindl, H. (1998) A phospholipase A_2 is transiently synthesized during seed germination and localized to lipid bodies. *Biochim. Biophys. Acta*, **1391**, 267–276.

McBride, K.E. and Summerfelt, K.R. (1990) Improved binary vectors for Agrobacterium-mediated plant transformation. *Plant Mol. Biol.* **14**, 269–276.

Milkowski, C. and Strack, D. (2004) Serine carboxypeptidase-like acyltransferases. *Phytochemistry* **65**, 517–524.

Milkowski, C., Baumert, A., Schmidt, D., Nehlin, L. and Strack, D. (2004) Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. *Plant J.* **38**, 80–92.

Nurmann, G. and Strack, D. (1979) Sinapine esterase. Part I. Characterization of sinapine esterase from cotyledons of *Raphanus sativus*. *Z. Naturforsch.* **34c**, 715–720.

Outchkourov, N.S., Peters, J., de Jong, J., Rademakers, W. and Jongsma, M.A. (2003) The promoter-terminator of chrysanthemum *rbcsS1* directs very high expression levels in plants. *Planta*, **216**, 1003–1012.

Segel, I.H. (1975) *Enzyme Kinetics. Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*. Toronto, Canada: John Wiley & Sons, Inc.

Sheahan, J.J. (1996) Sinapate esters provide greater UV-B attenuation than flavonoids in *Arabidopsis thaliana* (Brassicaceae). *Am. J. Bot.* **83**, 679–686.

Shevchenko, A., Wilm, M., Vorm, O. and Mann, M. (1996) Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels. *Anal. Chem.* **68**, 850–858.

Shirley, A.M., McMichael, C.M. and Chapple, C. (2001) The *sng2* mutant of Arabidopsis is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose:choline sinapoyltransferase. *Plant J.* **28**, 83–94.

Stehle, F., Brandt, W., Milkowski, C. and Strack, D. (2006) Structure determinants and substrate recognition of serine carboxypeptidase-like acyltransferases from plant secondary metabolism. *FEBS Lett.* **580**, 6366–6374.

Strack, D. (1981) Sinapine as a supply of choline for the biosynthesis of phophatidylcholine in *Raphanus sativus* seedlings. *Z. Naturforsch.* **36c**, 215–221.

Strack, D., Nurmann, G. and Sachs, G. (1980) Sinapine esterase. Part II. Specificity and change of sinapine esterase activity during germination of *Raphanus sativus*. *Z. Naturforsch.* **35c**, 963–966.

Teissére, M., Borel, M., Caillol, B., Nari, J., Gardies, A.M. and Noat, G. (1995) Purification and characterization of a fatty acyl-ester hydrolase from post-germinated sunflower seeds. *Biochim. Biophys. Acta*, **1255**, 105–112.

Tzagoloff, A. (1962) Metabolism of sinapine in mustard plants. II. Purification and some properties of sinapine esterase. *Plant Physiol.* **38**, 207–213.

Vicient, C.M. and Delseny, M. (1999) Isolation of total RNA from *Arabidospsis thaliana* seeds. *Anal. Biochem.* **268**, 412–413.

Wang, X. (2001) Plant phospholipases. *Annu. Rev. Plant Physiol. Biol.* **52**, 211–231.

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd,

Planta (2008) 228:663–674
DOI 10.1007/s00425-008-0768-3

ORIGINAL ARTICLE

The role of UDP-glucose:hydroxycinnamate glucosyltransferases in phenylpropanoid metabolism and the response to UV-B radiation in *Arabidopsis thaliana*

Dirk Meißner · Andreas Albert · Christoph Böttcher · Dieter Strack · Carsten Milkowski

Received: 27 May 2008 / Accepted: 3 June 2008 / Published online: 18 June 2008
© Springer-Verlag 2008

Abstract *Arabidopsis* harbors four UDP-glycosyltransferases that convert hydroxycinnamates (HCAs) to 1-*O*-*β*-glucose esters, UGT84A1 (encoded by At4g15480), UGT84A2 (At3g21560), UGT84A3 (At4g15490), and UGT84A4 (At4g15500). To elucidate the role of the individual UGT84A enzymes *in planta* we analyzed gene expression, UGT activities and accumulation of phenylpropanoids in *Arabidopsis* wild type plants, *ugt* mutants and overexpressing lines. Individual *ugt84A* null alleles did not significantly reduce the gross metabolic flux to the accumulating compounds sinapoylcholine (sinapine) in seeds and sinapoylmalate in leaves. For the *ugt84A2* mutant, LC/MS analysis revealed minor qualitative and quantitative changes of several HCA choline esters and of disinapoylspermidine in seeds. Overexpression of individual *UGT84A* genes caused increased enzyme activities but failed to produce significant changes in the pattern of accumulating HCA esters. For UGT84A3, our data tentatively suggest an impact on cell wall-associated 4-coumarate. Exposure of plants to enhanced UV-B radiation induced the UGT84A-encoding genes and led to a transient increase in sinapoylglucose and sinapoylmalate concentrations.

Keywords *Arabidopsis* · Hydroxycinnamate:UDPG glucosyltransferases · Sinapate ester · *ugt84A* mutants · *UGT84A* overexpression · UV-B stress

Abbreviations

SGT	UDP-glucose:sinapate glucosyltransferase
HCA	Hydroxycinnamates
OE	Overexpression
KO	Knock out
UGT	UDP-glycosyltransferase

Introduction

In plant secondary metabolism, glycosylation is a frequent chemical modification that affects metabolic activity and bioavailability of natural products and contributes to the extraordinary diversity of these compounds (Vogt and Jones 2000). Common reactions include the conjugation of the sugar moiety with hydroxyl or carboxyl groups of the aglycon to form *O*-glycosides or -esters. In Brassicaceae plants it has been shown that 1-*O*-sinapoylglucose serves as sinapoyl donor in acyltransfer reactions leading to a wide

Electronic supplementary material The online version of this article (doi:10.1007/s00425-008-0768-3) contains supplementary material, which is available to authorized users.

D. Meißner · D. Strack · C. Milkowski (✉)
Department of Secondary Metabolism,
Leibniz Institute of Plant Biochemistry,
Weinberg 3, 06120 Halle (Saale), Germany
e-mail: carsten.milkowski@uni-jena.de;
Carsten.Milkowski@ipb-halle.de

A. Albert
Department of Environmental Engineering,
Institute of Soil Ecology, Helmholtz Zentrum München,
German Research Center for Environmental Health,
85764 Neuherberg, Germany

C. Böttcher
Department of Stress and Developmental Biology,
Leibniz Institute of Plant Biochemistry,
Weinberg 3, 06120 Halle (Saale), Germany

Present Address:
C. Milkowski
Institute of General Botany and Plant Physiology,
Friedrich Schiller-University Jena, Am Planetarium 1,
07743 Jena, Germany

array of sinapate esters (Tkotz and Strack 1980; Lorenzen et al. 1996; Shirley et al. 2001; Baumert et al. 2005).

In developing seeds, sinapoylglucose feeds the synthesis of sinapoylcholine (sinapine), the major seed phenylpropanoid that accumulates to levels of 1–2% of seed dry weight, and the production of related esters (Baumert et al. 2005). During germination, sinapine is hydrolyzed by a sinapine esterase activity (Strack et al. 1980; Clauß et al. 2008). The liberated sinapate is again activated as sinapoylglucose and subsequently conjugated to L-malate. The resulting sinapoylmalate accumulates in vacuoles of sub-epidermal tissues and provides shielding against the deleterious effects of UV-B radiation (Landry et al. 1995; Sheahan 1996; Lehfeldt et al. 2000; Hause et al. 2002). In addition, HCA moieties are constituents of the cell wall, mostly esterified with polysaccharides and thereby assumed to counterbalance cell expansion (Tan et al. 1991, 1992; Iiyama et al. 1994; Piber and Koehler 2005; Philippe et al. 2007). Given the metabolic network of sinapate ester biosynthesis, it became evident that UGTs involved in sinapoylglucose formation might be promising targets to engineer the sinapate ester content in Brassicaceae. Previously, we could isolate a cDNA from oilseed rape (*Brassica napus*) that encodes a functional UDP-glucose:sinapate glucosyltransferase (BnSGT1; Milkowski et al. 2000a). According to the established UGT nomenclature this enzyme was classified as UGT84A9. Seed-specific silencing of the UGT84A9-encoding gene has proved successful in establishing lines of *B. napus* with substantially decreased seed content of sinapate esters (Baumert et al. 2005; Hüsken et al. 2005). The low sinapine trait is regarded as a major requirement for an efficient use of the valuable *B. napus* seed protein (Shahidi and Naczk 1992). From *Arabidopsis* four UGT84A9 homologs designated as UGT84A1, UGT84A2, UGT84A3 and UGT84A4 (referred to as UGT84A1-A4 enzymes) were identified (Milkowski et al. 2000b). UGT84A2 encoded by At3g21560 catalyzed in vitro preferentially the formation of sinapoylglucose whereas the remaining glucosyltransferases (UGT84A1, A3, and A4 encoded by At4g15480, At4g15490, and At4g15500, respectively) did not show pronounced substrate specificity toward any of the hydroxycinnamates (Milkowski et al. 2000b; Lim et al. 2001).

Compared to a wealth of reported in vitro data, only few studies are available on the biological role of UGTs. In maize, the conjugation of indole-3-acetic acid via 1-*O*-indolylacetylglucose to inositol plays a role in growth hormone homoeostasis (Kesy and Bandurski 1990; Kowalczyk and Bandurski 1991; Szerszen et al. 1994). UGT74B1 (Grubb et al. 2004) and UGT73C5 (Poppenberger et al. 2005) play a role in auxin and brassinosteroid homoeostasis. Jones et al. (2003) elucidated the substrate specificities of two flavonol glycosyltransferases *in planta*. UGT74F2 and UGT74F1 were identified as anthranilate-UGTs (Quiel and Benders 2003). By combining transcriptome and metabolome analyses, UGT75C1 was characterized as anthocyanin 5-*O*-glycosyltransferase in *Arabidopsis* (Tohge et al. 2005). Recently, the *bright trichomes 1* (*brt1*) mutant phenotype could be attributed to a mutation in *At3g21560* (UGT84A2) leading to the accumulation of a fluorescent putative polyketide in trichomes (Sinlapadech et al. 2007). However, residual SGT activity and sustained sinapate ester accumulation in the *brt1* null mutant suggest functional redundancy among the *Arabidopsis* UGT84A1-A4 enzymes.

To alleviate the damaging effects of enhanced UV-B radiation, plants make use of UV-screening pigments (Lois 1994; Reuber et al. 1996; Bharti and Khurana 1997; Bieza and Lois 2001). For *Arabidopsis*, the role of flavonoids and sinapate esters in UV-B stress adaptation was deduced from the analysis of mutants impaired in the biosynthesis of these compounds like *transparent testa tt4* and *tt5* (Li et al. 1993) or *fah1* (Chapple et al. 1992; Landry et al. 1995; Sheahan 1996). These mutant-based analyses, however, were hampered by the fact that the genetic background specifically interfered with the effects of mutations. Moreover, the individual mutants displayed different patterns of redistribution of other phenolics affecting the UV-B screening capacity (Booij-James et al. 2000). This made it difficult to dissect the relative effect of flavonoids and sinapate esters on UV stress adaptation.

This work was aimed at elucidating the role of UGT84A1-A4 *in planta*. The study describes a systematic analysis integrating transcript abundances with UGT enzyme activities and phenylpropanoid accumulation in *Arabidopsis* wild type plants, individual *ugt84A* null mutants and *UGT84A* overexpressing lines. Furthermore, we analyzed the influence of enhanced UV-B radiation on *UGT84A1-A4* expression and the accumulation of the UV-B-screening compound sinapoylmalate.

Materials and methods

Plant material and growth conditions

Arabidopsis thaliana L. Heynh. wild type plants of ecotypes Columbia (Col-0) and Nossen (Nos-0) were used as well as *Arabidopsis* mutants in the Col-0 background harboring T-DNA insertions in genes *At4g15480* (*ugt84A1*), *At4g15490* (*ugt84A3*) and *At4g15500* (*ugt84A4*), and a transposon-tagged mutant line in the Nos-0 background with the insertion in gene *At3g21560* (*ugt84A2*). Seeds of Col-0 and the mutant lines *ugt84A1* (Gabi Kat 765F10), *ugt84A3* (Salk_045492) and *ugt84A4* (Gabi Kat 826F10) were obtained from the Nottingham Arabidopsis Stock Centre (NASC; http://www.arabidopsis.info/). Seeds of Nos-0 and the *ugt84A2* mutant (11-5836-1) were purchased from RIKEN BioResource Center (http://www.brc.

riken.jp). Plants were sown on a soil:vermiculite mixture (4:1, w/w) and cultivated in a growth chamber under a photoperiod of 16 h light/8 h dark at 24°C and a photon flux density of about 130 µmol m^{-2} s^{-1}. For selection of transgenic plants, surface-sterilized seeds were germinated on synthetic media solidified by 1% Agar [Murashige and Skoog (MS) media, 0.05% MES, pH 5.7, supplemented with 1% sucrose; Murashige and Skoog 1962] containing 50 µg ml^{-1} of kanamycin. Seedlings were grown for 7 days and then transferred to soil:vermiculite. For analyses of transcript levels, enzyme activities and metabolite content, seedlings were harvested 1 week after sowing. Mature leaves were taken from rosettes of 4-week-old plants.

UV-B exposure

Arabidopsis plants were grown in the greenhouse for 3 weeks and then transferred to the sun simulator facilities of the German Research Center for Environmental Health (Neuherberg, Germany). The natural photobiological environment was provided using a combination of four lamp types (metal halide lamps, quartz halogen lamps, blue fluorescent tubes and UV-B fluorescent tubes) to obtain a natural balance of simulated global radiation throughout the ultraviolet to infrared spectrum. The lamp types are arranged in several groups to realize the natural variations of the solar irradiance by switching appropriate group of lamps on and off. The short-wave cut-off was shaped by selected borosilicate and soda-lime glass filters (Döhring et al. 1996). The plants were exposed to radiation for 16 h per day in the sun simulators at climate conditions of 24°C temperature and 70% relative humidity. The light intensity in the sun simulators was stepwise increased during a 4-day-period of acclimatization up to a mean value of PAR (photosynthetically active radiation, 400–700 nm) of 500 µmol m^{-2} s^{-1} and UV-A radiation (280–315 nm) of 10.5 W m^{-2}. After this acclimatization in the absence of UV-B radiation (280–315 nm), the plants were treated by adding UV-B radiation of 1.25 W m^{-2} corresponding to a biologically effective UV-B radiation of 400 mW m^{-2} (Caldwell 1971) for 14 h per day. The spectroradiometric measurements were performed with the double monochromator system (Bentham, Reading, UK). Variations of the integrated values were less than 15%. Control plants were grown under the same conditions in the absence of UV-B radiation. Plants were harvested after 10 h and 6 days of UV-B treatment.

RNA isolation

RNA from seedlings and leaves was isolated by selective adsorption onto silica using the RNeasy® Plant Mini Kit (Qiagen, Hilden, Germany). RNA purification from seeds was done as described earlier (Vicient and Delseny 1999). Both protocols used DNaseI treatment according to the cleanup protocol of the RNeasy® Plant Mini Kit to remove residual DNA contamination.

RT-PCR

A total of 1 µg of total RNA was subjected to first-strand cDNA synthesis using the Omniscript RT Kit (Qiagen) following the protocol provided by the manufacturer. 2 µl of cDNA were used as template for PCR amplification done with Go-Taq-Polymerase (Promega, Madison, WI, USA). PCR conditions were as follows: 95°C for 2 min; 30 cycles of 95°C for 45 s; 56°C for 30 s, 72°C for 1 min; and a final elongation step of 10 min at 72°C. The following gene-specific primers were used:

At4g15480 (*UGT84A1*), 5′-GGTAGTGCCAAGGGAGGAAGTG-3′ (fw) and
5′-CCATTATCTTTAGTCTTCGTTACTC-3′ (rev);
At3g21560 (*UGT 84A2*), 5′-GTTAGTGCCGAGGGAGGAAG-3′ (fw) and
5′-CAGCCAAGACATGATTATGAC-3′ (rev);
At4g15490 (*UGT 84A3*), 5′-CAAGTGACTGATGCGGTGTACTTG-3′ (fw) and
5′-TCACATGTTTCGTAACCAACTTGTC-3′ (rev);
At4g15500 (*UGT 84A4*), 5′-CGTGGAGCTTCCGATGAGAG-3′ (fw) and
5′-CACGACATTATTAATGTTTGTCATTGTC-3′ (rev);
actin, 5′-TGTATGTCGCCATCCAAGCTG-3′ (fw) and
5′-CCGCTCTGCTGTTGTGGTG-3′ (rev).

A total of 10 µl of PCR mixture was loaded onto a 1.2% agarose gel and the amplification products were separated by gel electrophoresis. The 200 bp PCR products were visualized by ethidium bromide staining.

Quantitative real time RT-PCR (qPCR)

Quantitative PCR analyses with cDNA as template were performed using the qPCR™ Mastermix Plus for SYBR® Green I (Eurogentec) in a Mx3005P™ Cycler (Stratagene, La Jolla, CA, USA) according to the protocols given by the supplier. Transcript levels were normalized to 18S rRNA transcript levels. The following primers were used.

UGT 84A1, 5′-GATGACCGGAGAGCTGATTTCT-3′ (fw) and
5′-AACGCTCTCTAGGTGAGCAATG-3′ (rev);
UGT 84A2, 5′-TGAAGACGACGAAGCTAGCAGA-3′ (fw) and
5′-AGATGTGGTCGGAGGATGGTTA-3′ (rev);
UGT 84A3, 5′-ATCTCGTTAAGAGATATAACAAGGAGCC-3′ (fw) and

5′-CGTTGTTTATGAGACACGTCACC-3′ (rev);
UGT 84A4, 5′-TCGGGCTAGGTTTTCTCCG-3′ (fw) and
5′-TGTAGACAAATCCATCCTCGAAGA-3′rev;
18S rRNA, 5′-ATCAGCTCGCGTTGACTACGTC-3′ (fw) and
5′-CCGGATCATTCAATCGGTAGG-3′ (rev).

Average results of three biological replicates, each measured by three technical replicates, are given.

Constructs for overexpression of the *UGT84A* genes in *Arabidopsis*

Molecular cloning experiments were performed according to the standard protocols (Sambrook et al. 1989). Total DNA from leaf tissue was isolated with the DNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. The coding regions of the *Arabidopsis* genes At4g15480 (UGT84A1), At3g21560 (UGT 84A2), At4g15490 (UGT 84A3) and At4g15500 (UGT 84A4) were amplified by PCR with Platinum PCR SuperMix High Fidelity (Invitrogen, Karlsruhe, Germany) from genomic DNA. The following gene-specific primers were used.

UGT 84A1, 5′-ACA AACATACAAATAATTCT-3′ (fw) and
5′-CAAACATGTGCCGAAGCC-3′ (rev);
UGT 84A2, 5′-ATGGAGCTAGAATCTTCTCC-3′ (fw) and
5′-TTAAAAGCTTTTGATTGATCC-3′ (rev);
UGT 84A3, 5′-GATGGAAAGCACTGGAAAGG-3′ (fw) and
5′-GATAAGAAGACAATACG-3′ (rev);
UGT 84A4, 5′-ATGGAGATGGAATCGTCGTTACC TC-3′ (fw) and
5′-TTACACGACATTATTAATGTTTGTC-3′ (rev).

Polymerase chain reaction products were cloned into plasmid pGEM-Teasy (Promega) and characterized by sequence analysis. The four UGT84A reading frames were then transcriptionally fused to the promoter of Rubisco small subunit (rbcS1) from *Asteraceous chrysanthemum* (Outchkourov et al. 2003) by cloning into the *Not*I site of plasmid pImpact1.1 (Plant Research International, Wageningen, The Netherlands). The expression cassettes were then introduced as *Asc*I–*Pac*I fragments into the binary vector pBINPLUS (Plant Research International; van Engelen et al. 1995).

Transformation of *Arabidopsis*

Binary plasmid constructs for transformation of *Arabidopsis* were introduced into *Agrobacterium tumefaciens* strain EHA105 (Hood et al. 1993) by electroporation. *Arabidopsis* plants were transformed by the floral dip method (Clough and Bent 1998). For selection of transgenic plants T1 seeds were surface-sterilized in 70% ethanol for 2 min followed by a mixture of 0.15% (v/v) Tween 20 and 12% (w/v) household bleach (sodium hypochloride) for 10 min. Seeds were rinsed thoroughly with sterile water and after swelling over night at 4°C plated on modified MS medium (Duchefa, Haarlem, The Netherlands) supplemented with 500 μg ml^{-1} carbenicillin and 25 μg ml^{-1} kanamycin. After scoring the development of antibiotic damage symptoms for 14 days post treatment, kanamycin resistant plants were transferred to soil:vermiculite (4:1, w/w).

UGT activity assays

Arabidopsis protein extracts were prepared by disrupting leaf material of 100 mg fresh weight with a glass homogenizer (VWR International, Vienna, Austria) in 1 ml of 100 mM sodium phosphate buffer (pH 6.0). After centrifugation (2 min, 14,000*g*, 4°C), the supernatants were desalted using PD-10 Sephadex G-25 columns (Amersham Bioscience, Munich, Germany). Extracts were concentrated by ultrafiltration through the Amicon Ultra 4 filter device with a molecular weight cut off (MWCO) of 10 kDa (Millipore, Schwalbach, Germany). Protein concentrations were determined by the Bradford method (Bradford 1976) with bovine serum albumin as standard. The assay mixtures contained 4 mM UDP-glucose, 2 mM HCA glucose acceptor (4-coumarate, caffeate, ferulate or sinapate) and 80 μg plant protein in a total volume of 150 μl of 100 mM MES buffer (pH 6.0), 1 mM 2-mercaptoethanol, 0.5 mM EDTA and 10% (v/v) glycerol. Enzyme reaction mixtures were incubated at 30°C for 1 h. The reactions were terminated by adding 10 μl trifluoroacetic acid (TFA). To determine product formation, the mixtures were centrifuged and aliquots of the supernatants were subjected to HPLC analysis as described (Milkowski et al. 2004).

Secondary metabolite analysis

Extracts from plant tissues in 80% (w/v) methanol were prepared by mechanical disruption in a Bead Beater (Biospec Products, Bartlesville, OK, USA) and analyzed by HPLC as described previously (Weier et al. 2007). Hydroxycinnamates were photometrically detected at 330 nm and quantified by external standardization with authentic samples (Baumert et al. 2005) from our collection of phenylpropanoid compounds (not documented). Flavonoids were photometrically characterized as kaempferol or quercetin derivatives by comparison with the absorption spectra of the authentic aglycones. In each case, six biological replicates were analyzed. To measure cell wall-esterified

phenylpropanoids, methanol-extracted tissues were subject to alkaline hydrolysis according to Tan et al. (2004). Hydrolysis products were analyzed by HPLC.

LC–MS analysis

Sample preparation

A total of 10 mg of seeds were homogenized in 500 µl methanol/water, 4:1 (v/v) with 0.3 g zirconia beads in a MiniBeadBeater. After centrifugation, 400 µl of the supernatant was evaporated to dryness in a vacuum centrifuge. The pellets were resuspended in 400 µl methanol/water, 1:9 (v/v) and filtered through a 0.45 µm PTFE syringe filter.

Data acquisition

A total of 1 µl of seed extract was separated using a Ultimate capillary LC system (Dionex) equipped with a modified C_{18} column (GROMSIL ODS 4 HE, 150 mm, 0.3 mm, 3 µm particle size, 120 Å pore size, guard column 10 mm, 0.3 mm; Alltech Grom GmbH, Rottenburg-Hailfingen, Germany) applying the following binary gradient at a flow rate of 5 µl min^{-1}: 0–5 min, isocratic 95% A (water/formic acid, 99.9:0.1, v/v), 5% B (acetonitrile/formic acid, 99.9:0.1, v/v); 5–45 min, linear from 5 to 55% B. Eluted compounds were detected by an API QSTAR Pulsar Hybrid quadrupole time-of-flight mass spectrometer (Applied Biosystems) equipped with an IONSPRAY electrospray ionization source in positive ion mode. Instrument settings were as follows: ion spray voltage 5.5 kV, DP1 50 V, DP2 15 V, FP 220 V, nebulizer gas (N_2) 25 arbitrary units, curtain gas (N_2) 20 arbitrary units, collision gas (N_2) 4 arbitrary units, pulser frequency 9.986 kHz, accumulation time 2 s. Ions were detected in "enhance all" mode from *m/z* 75 to 1,000. For each genotype three biologically independent samples were analyzed in two technical replications.

Data analysis

Raw data files were centroided and converted to mzData format. Afterwards, peak picking, non-linear retention time correction and alignment were performed using the XCMS package (http://metlin.scripps.edu/download/). The resulting data matrix [$(m/z, t_r)'$ sample] was imported into Microsoft Excel. For further analysis only consistent signals were considered, which were at least present in a single genotype in all of the biological and technical replicates. For pairwise comparison of two genotypes fold changes and *P* values (*t* test, two-sided, unequal variance) were calculated for each signal. A signal was regarded differential, when fold change was higher than two at a significance level of $P < 0.01$. Putative identification of compounds related to differential signals was performed by calculation of elemental composition and interpretation of tandem-MS data.

Results

Activity of ester-forming HCA-UGTs in *Arabidopsis*

Protein extracts of *Arabidopsis* seeds, seedlings and mature leaves were incubated with UDP-glucose and several HCAs. HPLC analysis of the formed HCA glucose esters revealed highest UGT activities in young seedlings and seeds (Fig. 1). Mature leaves contained particularly low ester-forming HCA-UGT activity. Moreover, the analyses showed a clear shift of UGT specificities when comparing seeds and seedlings to mature leaves. In seeds and seedlings, the enzymatic activity to form sinapoylglucose was predominating followed by that to synthesize 4-coumaroylglucose and feruloylglucose. In contrast, protein extracts of mature leaves displayed the highest activity toward 4-coumarate, and the activity toward ferulate had dropped to a minimum value.

The real time qPCR assay confirmed that *UGT84A2* was the predominantly expressed *UGT84A* gene in *Arabidopsis* with highest transcript abundance in young seedlings as reported previously (Sinlapadech et al. 2007). We detected transcripts of all four *UGT84A* genes in seeds, seedlings and mature leaves indicating a concomitant expression (Electronic supplementary material 1).

Impact of overexpression and null alleles on UGT84A1-A4 enzyme activity and metabolite composition *in planta*

By overexpression (OE) of individual UGT84A genes we were able to modulate the HCA-UGT activity of *Arabidop-*

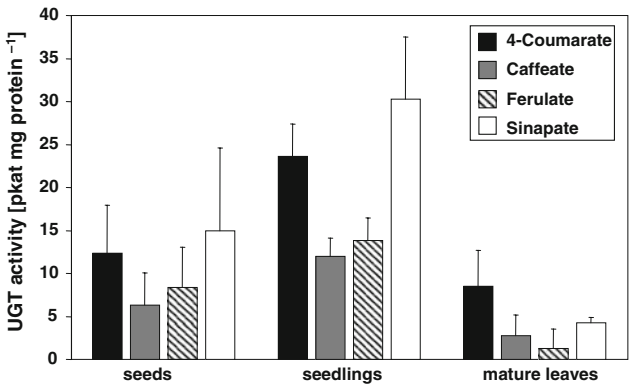


Fig. 1 UGT activity toward hydroxycinnamates in *Arabidopsis*. Protein extracts were prepared from the plant organs indicated. UGT activities were calculated by quantification of 1-*O*-glucose esters formed with the substrates 4-coumarate, caffeate, ferulate and sinapate. The data represent the mean ± SE of six samples

sis plants (Electronic supplementary material 2; Fig. 2). Compared to the *Arabidopsis* wild type, crude protein extracts from leaves of UGT84A1OE plants displayed increased activity toward 4-coumarate and ferulate, the preferred in vitro substrates of UGT84A1 (Milkowski et al. 2000a, b; Lim et al. 2001). Accordingly, UGT84A2OE caused increasing activity toward sinapate whereas leaves of UGT84A3OE displayed a significantly higher capacity to synthesize feruloylglucose. UGT84A4OE plants did not display changes in UGT specificity of leave extracts. This was most likely due to a less pronounced overexpression in this transgenic line (Electronic supplementary material 2).

Compared to the wild type, the overexpressing lines UGT84A1OE, UGT84A2OE, and UGT84A3OE exhibited increased levels of sinapoylglucose in seeds and seedlings (Fig. 3a). The seed sinapine content was slightly increased in UGT84A1OE, UGT84A2OE, and UGT84A3OE. Mature leaves of UGT84A2OE showed a slight increase in the sinapoylmalate content (Fig. 3b). Remarkably, the HPLC analyses failed to detect any HCA glucose esters different from sinapoylglucose in the soluble metabolite pool of the UGT84AOE plants. Analysis of HCAs liberated from cell walls revealed for UGT84A3OE a significant increase in cell wall-associated 4-coumarate, ferulate and sinapate (Fig. 4). Overexpression of UGT84A1 and UGT84A2 resulted both in a significant increase of cell wall-bound sinapate. Given the impact of UGT84A3 overexpression on the amount of cell wall-associated HCAs, we analyzed the cell wall fraction of the insertion mutant UGT84A3KO. HPLC-based quantification of hydrolytically released HCAs showed a significant decrease of the 4-coumarate content in cell walls of UGT84A3KO whereas the levels of ferulate and sinapate were not changed compared to the wild type (Fig. 5). This corroborates the importance of UGT84A3 for the cell wall content of HCAs, especially 4-coumarate.

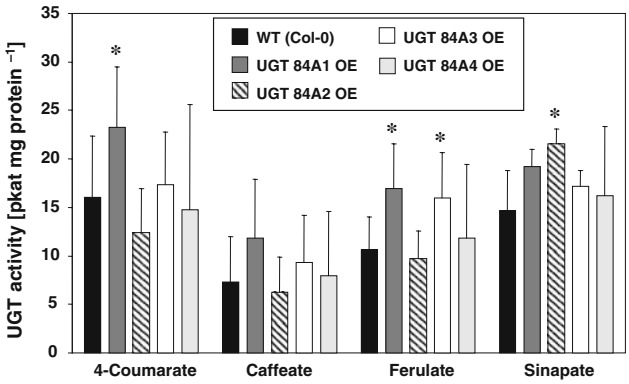


Fig. 2 UGT activity of *Arabidopsis* plants overexpressing the genes At4g15480 (UGT84A1OE), At3g21560 (UGT84A2OE), At4g15490 (UGT84A3OE) and At4g15500 (UGT84A4OE). Protein extracts were prepared from leaves. UGT activities were calculated by quantification of 1-*O*-glucose esters formed with the substrates 4-coumarate, caffeate, ferulate and sinapate. The data represent the mean ± SE of six samples. Significant changes compared to the wild type (*t* test, $P < 0.05$) are indicated by *asterisks*

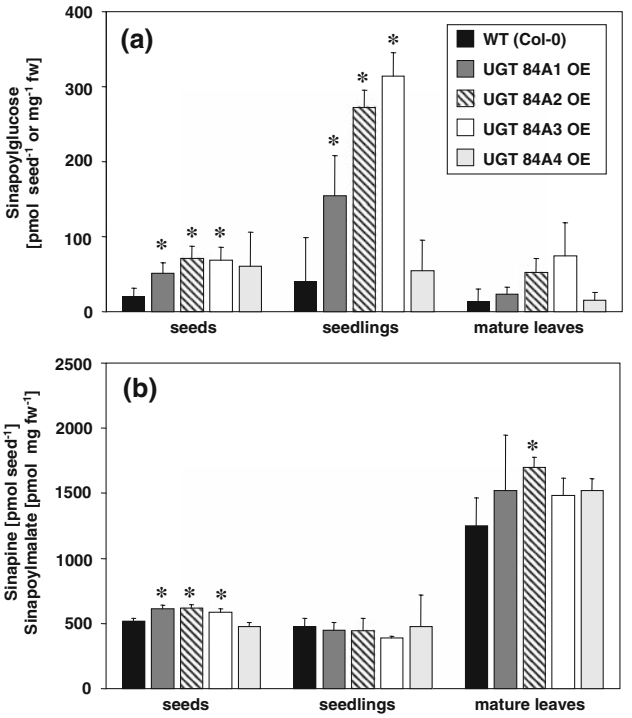


Fig. 3 Impact of *UGT84A* overexpression on the accumulation of sinapoylglucose (**a**) and sinapine or sinapoylmalate (**b**) in *Arabidopsis*. Methanolic extracts from the plant organs indicated were analyzed by HPLC. Sinapoylglucose was quantified for seeds, seedlings and mature leaves. Sinapine content was calculated for seeds whereas for seedlings and mature leaves sinapoylmalate was measured. Data represent the mean ± SE from six samples. Significant changes compared to the wild type (*t* test, $P < 0.05$) are indicated by *asterisks*

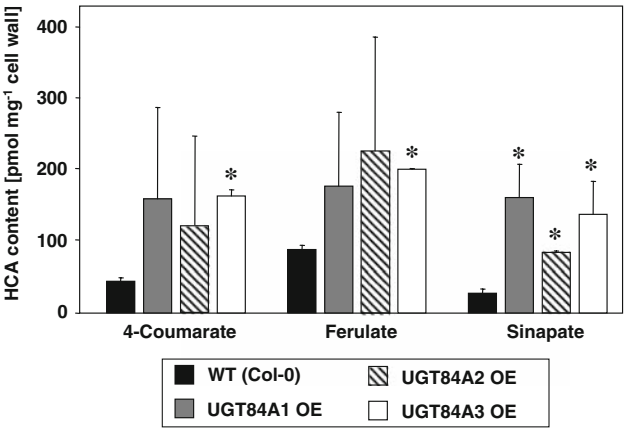


Fig. 4 Impact of UGT84A overexpression on the content of cell wall-bound hydroxycinnamates. Methanolic extracts from mature leaves were subject to alkaline hydrolysis and the liberated compounds were analyzed by HPLC. Data represent the mean ± SE of six samples. Significant changes compared to the wild type (*t* test, $P < 0.05$) are indicated by *asterisks*

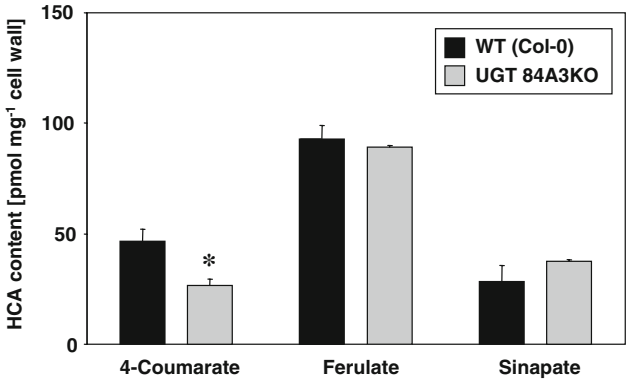


Fig. 5 Impact of *ugt84A3* null mutation on cell wall-bound hydroxycinnamates from mature leaves. Data represent the mean ± SE of six samples. Significant changes compared to the wild type (*t* test, $P < 0.05$) are indicated by *asterisk*

Among the *ugt84A* mutant lines (KO lines; S2), only the null allele of *UGT84A2* provoked a decrease of ester-forming HCA-UGT activity *in planta*. This was in the range of the previously reported allelic *brt1* mutant (Sinlapadech et al. 2007). Accordingly, HPLC analysis revealed a slightly reduced sinapoylmalate content in seedlings and adult leaves whereas the seeds seemed to accumulate sinapine as the wild type (data not shown).

To elucidate metabolic changes in a highly sensitive manner we performed untargeted LC-ESI-QTOF-MS-based metabolite profiling of UGT84A2KO and *Arabidopsis* Col-0 wild type seeds (von Roepenack-Lahaye et al. 2004). By this approach we were able to detect and align about 1,200 distinct mass signals from methanolic extracts of both mutant and wild type seeds. Based on accurate masses obtained from QTOF instruments, the annotation of differential mass signals and the putative identification of corresponding metabolites were pursued by calculation of elemental composition and interpretation of collision-induced dissociation mass spectra (MS data see Electronic supplementary material 3). Pair-wise comparison of metabolite profiles applying fold-change and *P* value cut-offs of 2 and 0.01, respectively, revealed that in UGT84A2KO seeds 40 mass signals were less abundant than in the wild type. On the other hand, 48 mass signals displayed stronger intensity in UGT84A2KO seeds. We were able to annotate 11 of the mass signals less abundant in UGT84A2KO seeds. These were found to be associated with feruloylcholine and hydroxyferuloylcholine (Table 1). Likewise, the annotation of 19 mass signals increased in UGT84A2KO seeds led to the putative identification of five choline esters derived from sinapate or syringate as well as *N*,*N*′-di(sinapoyl)spermidine (Table 2; Fig. 6).

Table 1 Identified metabolites showing decreased levels in *Arabidopsis* UGT84A2KO seeds

No.	Compound	t_r (min)	Quantifier ion *m/z*	Fold change	*P*
1	FC	19.7	280.15 $(M)^+$	5.0	3.5×10^{-8}
2	OH-FC	21.5	296.15 $(M)^+$	7.9	2.8×10^{-8}

FC Feruloylcholine

Role of UGT84A1-A4 enzymes in the response to UV-B radiation

We studied gene expression, enzyme activity and metabolite composition in *Arabidopsis* leaves upon UV-B exposure. As shown by semi-quantitative RT-PCR, exposure of plants to increased UV-B radiation for 10 h triggers the transcriptional induction of all UGT84A1-A4 encoding genes in *Arabidopsis* leaves (Fig. 7). Highest transcript levels were measured for UGT84A2 and UGT84A3. UGT84A1 for which the transcript level in leaves under control conditions was near the detection limit showed a clear cDNA signal after 10 h of UV-B exposure. UGT84A4 was only slightly induced by 10 h UV-B treatment. For UGT84A2 and UGT84A3 the increased transcript levels were found to be maintained after 6 days of UV-B exposure. In contrast, the transcript levels of UGT84A1 and UGT84A4 had decreased below the detection limit of RT-PCR at this time point. The transcript abundance of the *SNG1* gene encoding sinapoylglucose:malate sinapoyltransferase (SMT)—the final enzyme in the biosynthesis of sinapoylmalate—was not induced after 10 h of UV-B exposure and only slightly increased by long-term UV-B stress. Chalcone synthase (CHS), the first committed enzyme of flavonoid biosynthesis, displayed a clear transcriptional induction at 10 h of increased UV-B treatment. After 6 days of UV-B stress, CHS transcript abundance had further increased. In accordance with the induced transcription, *Arabidopsis* leaves exposed for 10 h to UV-B radiation showed a significant increase in ester forming UGT activity towards sinapate (Fig. 8). After 6 days of UV-B exposure, however, this enzymatic SGT activity was only slightly above the untreated control although the transcript levels for *UGT84A2* and *UGT84A3* were still high (Fig. 7). To investigate the impact of UV stress-mediated gene regulation on phenylpropanoid accumulation we measured the leaf content of sinapate esters and flavonoid conjugates. For quantitative analysis the latter were specified as kaempferol and quercetin derivatives according to the spectral features of the aglycones. As the *Arabidopsis* leaf flavonoids are mainly composed of kaempferol and quercetin derivatives (Veit and Pauli 1999) this approach proved sufficient to

Table 2 Identified metabolites accumulating to higher levels in *Arabidopsis* UGT84A2KO seeds

No.	Compound	t_r (min)	Quantifier ion *m/z*	Fold change	*P*
3	SC 4-*O*-Hexoside	17.5	472.22 $(M)^+$	16.9	1.1×10^{-7}
4	SyC	18.7	284.15 $(M)^+$	12.4	1.7×10^{-5}
5	SyC 4-*O*-hexoside	10.2	446.20 $(M)^+$	9.5	6.4×10^{-4}
6	SC(4-O-*β*)G	25.4	506.24 $(M)^+$	2.4	4.2×10^{-7}
7	SC(4-*O*-*β*)G 4'-*O*-hexoside	21.6	668.29 $(M)^+$	7.2	3.1×10^{-5}
8	*N*,*N*′-Di(sinapoyl)spermidine	27.2	558.28 $(M + H)^+$	2.4	2.2×10^{-5}

SC Sinapoylcholine, *FC* Feruloylcholine, *SyC* Syringoylcholine, *G* Guaiacyl

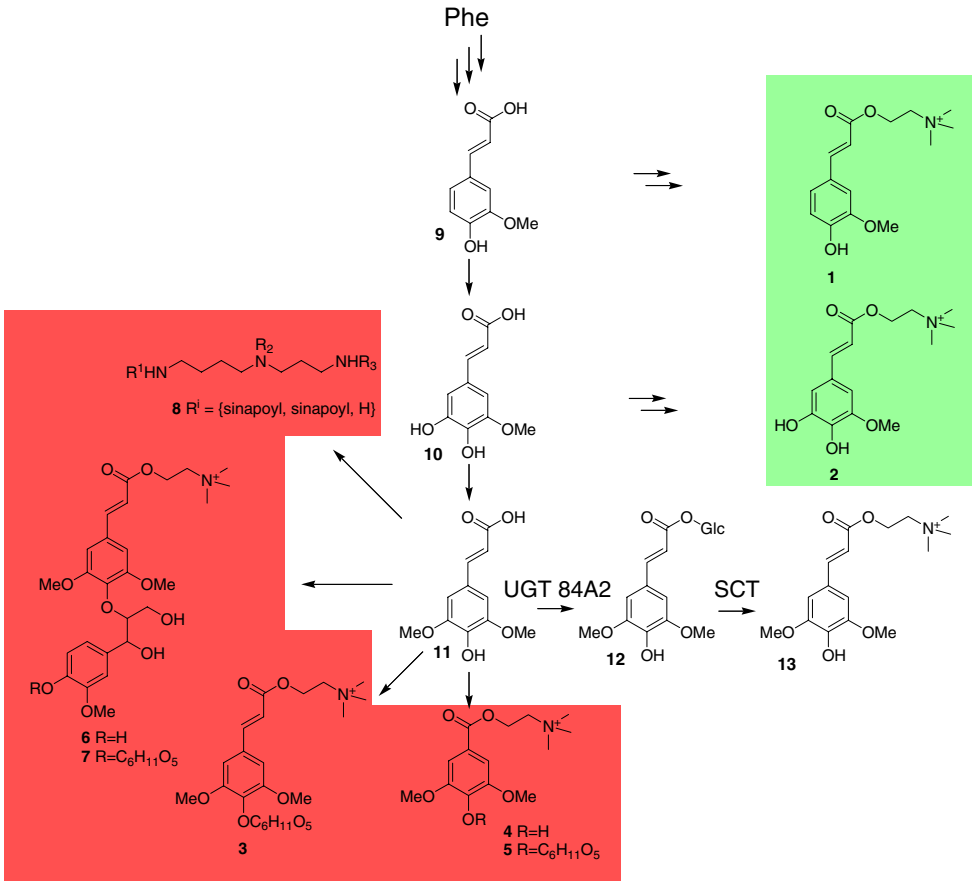


Fig. 6 Impact of the *ugt84A2* null mutation on HCA metabolism in seeds. Compounds with increased concentrations compared to the wild type are highlighted in *red*. Metabolites for which the concentration was decreased in *ugt84A2* seeds are underlayed with *green*. *Phe* L-Phenylalanine, *SCT* 1-*O*-sinapoyl-*β*-D-glucose:choline sinapoyltransferase (EC 2.3.1.91), *9* ferulate, *10* 5-hydroxyferulate, *11* sinapate, *12* 1-*O*-sinapoylglucose, *13* sinapoylcholine. Substances *1*–*8*, see Tables 1 and 2

evaluate the relative impact of flavonoids without identifying the full range of individual compounds. After 10 h of UV-B exposure the analysis revealed an increase in sinapoylglucose and sinapoylmalate as well as in the flavonoid derivatives compared to the untreated control plants (Fig. 9a). In contrast, after 6 days of UV-B treatment, the level of sinapate esters had dropped to that of control plants whereas the flavonoid content further increased. Interestingly, in the *ugt84A2* null mutant we found a UV-B stress-dependent accumulation kinetics of sinapate esters similar to the wild type (Fig. 9b). This indicates a functional redundancy of UGT84A enzymes to maintain the adaptation to short term UV-B exposure. Among the flavonoids, the strongest increase was found in the level of quercetin derivatives.

Discussion

The impact of UGT84A1-A4 enzymes on plant HCA metabolism: functional redundancy versus specific role

The sensitive real time qPCR revealed that UGT84A1-A4 encoding genes were expressed concomitantly in seeds, seedlings and adult leaves. Together with the broad substrate specificity, this provides the biochemical basis for a

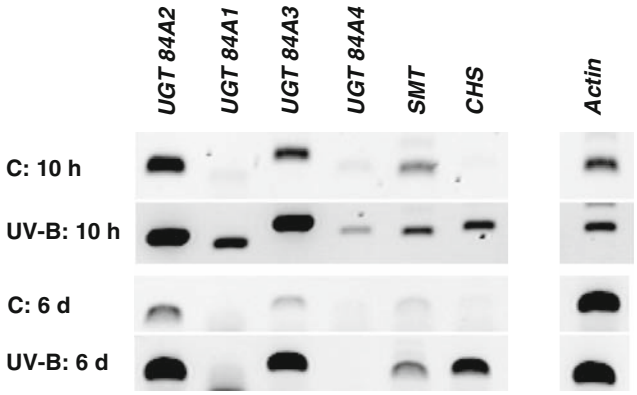


Fig. 7 RT-PCR analysis of gene expression in leaves of *Arabidopsis* Col-0 wild type exposed to UV-B radiation. RNA from rosette leaves of plants exposed to increased UV-B radiation for 10 h and 6 days, respectively, was used in RT-PCR (30 cycles) with gene-specific primers. Expression of actin is given as control. *SMT* Sinapoylglucose:malate sinapoyltransferase, *CHS* chalcone synthase, *C* control plants without UV-B treatment, *UV-B* plants exposed to UV-B radiation

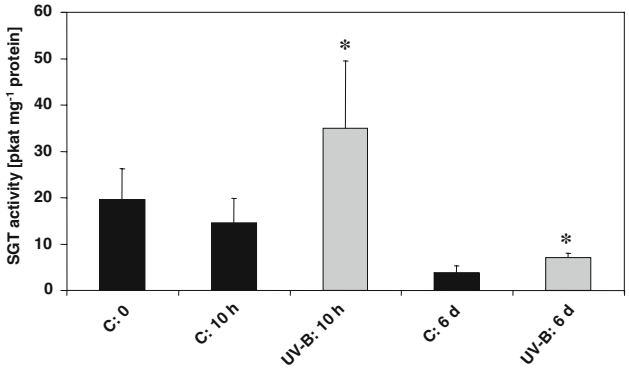


Fig. 8 UGT activity toward sinapate in *Arabidopsis* leaves exposed to UV-B radiation. Enzyme activities of leaf protein extracts were calculated by quantification of 1-*O*-sinapoylglucose formation. The data represent the mean ± SE of six samples. Significant changes (*t* test, $P < 0.05$) are indicated by *asterisks*. *C* Control plants without UV-B treatment, *UV-B* plants exposed to UV-B radiation

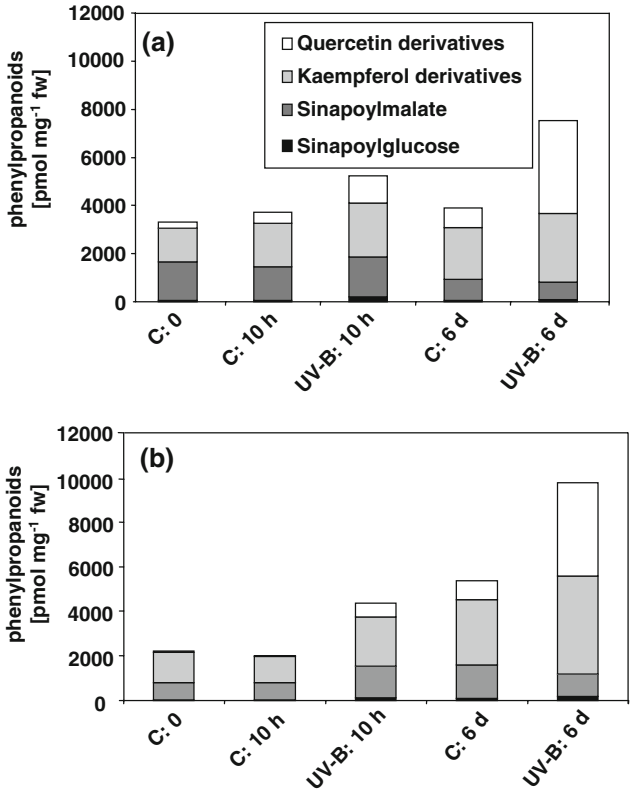


Fig. 9 Impact of UV-B radiation on the phenylpropanoid content in leaves. Methanolic extracts were prepared from leaves of *Arabidopsis* Col-0 (**a**) and the *ugt84A2* null mutant (**b**) exposed under UV-B radiation as indicated and analyzed by HPLC. Data represent the mean from six samples. Standard errors were in the range of 10% of the mean values. *C* Control plants without UV-B treatment, *UV-B* plants exposed to UV-B radiation

possible functional redundancy of UGT84A1-A4 enzymes in *Arabidopsis*. To evaluate the degree of redundancy we analyzed the content of HCA esters in individual UGT84AKO and OE lines. The phenylpropanoid metabolism of *Arabidopsis* and *B. napus* is characterized by a remarkable metabolic flux toward sinapine in seeds and sinapoylmalate in vegetative tissues. Our analysis showed that neither overexpression nor null mutations of individual genes were effective in diverting or disrupting this metabolic path in *Arabidopsis*. Even in the *ugt84A2* null mutant defective in the major *Arabidopsis*-SGT we found only a slight decrease in the soluble sinapoylmalate pool, and the sinapine accumulation in seeds was not affected. This is in accordance with the results reported for the *brt1* mutant impaired in *UGT84A2* (Sinlapadech et al. 2007) and corroborates the assumption of a remarkable functional redundancy among the UGT84A1-A4 enzymes. Interestingly, the pronounced functional redundancy that governs HCA glucose ester formation in *Arabidopsis* could not be found in the closely related crop plant *B. napus*. In *B. napus*, the seed-specific silencing of *UGT84A9*, the *UGT84A2* homolog, has proved sufficient to reduce sinapine accumulation by about 60% and to decrease the total seed sinapate ester content by up to 80% (Hüsken et al. 2005). Accordingly, a RT-PCR based screening of *B. napus* seed ESTs for UGT-cDNAs failed to detect *UGT84A* homologs different from *UGT84A9* (Mittasch et al. 2007).

Overexpression of the *UGT84A* genes led to increased levels of sinapoylglucose in seeds and seedlings. The increase in seedlings was more pronounced than in seeds, most likely due to the relatively low SMT activity at the young seedling developmental stage. This accumulation, however, indicates that the supply of sinapoylglucose is not limiting for transacylation toward sinapine in seeds and sinapoylmalate in seedlings, respectively. Unexpectedly, we could not detect HCA glucose esters different from

sinapoylglucose in the UGT84OE lines. This reveals that the supply of HCAs different from sinapate might be insufficient. On the other hand, we cannot exclude that potentially formed HCA glucose esters could serve as transient metabolic intermediates, which would be sequestered to the cell wall and would therefore not be detectable in the soluble pool of phenylpropanoids.

Since the UGT84A1-A4-encoding genes were concomitantly transcribed in *Arabidopsis* we cannot give predictions on the extent to which the several UGT84A enzymes contribute to this redundancy in plant metabolism. However, given the weak expression in seeds, seedlings and mature leaves, UGT84A4 seems to be of less functional importance in these plant organs. Sinlapadech et al. (2007) reported on high transcript abundance of *UGT84A4* in roots that could be confirmed by our real time RT qPCR approach (not shown). This might point to a specific function of this enzyme in root metabolism and should be subject of future investigations. To further address the question of redundant UGT84A1-A4 functions *in planta*, double mutants would be an appropriate tool. Unfortunately, this approach is restricted to the combinations of the *ugt84A2* mutant allele with either *ugt84A1*, *ugt84A3*, or *ugt84A4*. Combining of mutant alleles *ugt84A1*, *ugt84A3* and *ugt84A4* among each other by genetic crossing would be hampered by the tandem arrangement of these genes in the *Arabidopsis* genome. On the other hand, to measure the relative impact of the several UGT84A1-A4 enzymes on the formation of sinapoylglucose in leaves and seeds it might be helpful to introduce the individual mutant alleles into the genetic background of the *Arabidopsis* mutants *sng1* and *sng2*. These mutants accumulate sinapoylglucose in leaves (*sng1*) or in seeds (*sng2*) caused by mutations in the genes encoding sinapoylglucose:malate sinapoyltransferase (SMT) or sinapoylglucose:choline sinapoyltransferase (SCT) (Lehfeldt et al. 2000; Shirley et al. 2001).

UGT84A3 was the only UGT84A enzyme in *Arabidopsis* for which our data tentatively suggested a specific role in addition to its impact on the plant soluble sinapate ester pool. This was illustrated by the *ugt84A3* null allele that caused a significantly reduced content of cell wall associated 4-coumaroyl conjugates that could not be counterbalanced by functional UGT84A1, UGT84A2 or UGT84A4 enzymes in the mutant. Moreover, the UGT84A3OE line displayed an increase in cell wall associated 4-coumarate. So far, our data reveal UGT84A3 as a potential target for molecular approaches aimed at changing the hydroxycinnamate composition of cell walls.

Metabolic changes in *ugt84A2* mutant seeds

Since sinapate esters provide anti-nutritive characteristics to the seed protein fraction of crop plants like *B. napus*, there is a persistent interest in manipulating the seed phenylpropanoid metabolic network to reduce sinapate ester biosynthesis and accumulation (Shahidi and Naczk 1992; Hüsken et al. 2005). This led us to thoroughly analyze the sinapate ester content of *Arabidopsis ugt84A2* mutant seeds. By sensitive LC/MS analysis we could detect changes in the concentration of compounds whose biosyntheses could be attributed to sinapate or ferulate. The measurements revealed for *ugt84A2* mutant seeds increased levels of certain aromatic choline esters and disinapoylspermidine and, on the other hand, reduced levels of feruloylcholine and hydroxyferuloylcholine. The changes in the seed metabolite profile may reflect two metabolic strategies taken by the plants to cope with a potential transient accumulation of the reactive sinapate triggered by its reduced conversion to sinapoylglucose in the *ugt84A2* mutant. A first strategy would include the reinforced channeling of sinapate into more inert choline esters like sinapoyl-4-*O*-hexoside and related compounds, syringylcholine or disinapoylspermidine (Fig. 6; Table 2). The seed soluble metabolic pool obviously contains sufficient free choline and spermidine to counterbalance an increased sinapate supply by conjugation. Secondly, the reduced levels of feruloylcholine and hydroxyferuloylcholine (Fig. 6; Table 1) may reflect a negative metabolic feedback regulation caused by accumulating sinapate. This would lead to a reduced metabolite flux through the phenylpropanoid pathway towards sinapate, likewise decreasing the steady state levels of ferulate and hydroxyferulate, the precursors of accumulating feruloylcholine and hydroxyferuloylcholine.

From the eight compounds identified (Fig. 6; Electronic supplementary material 3), three had been reported previously from other plants. Feruloylcholine was isolated from seeds of *Cleome punges* (Pagani and Romussi 1970), 4-hexosyloxy-3,5-dimethoxy-cinnamoylcholine (SC 4-*O*-hexoside) was found in seeds of *Alliaria officinalis* (Larsen et al. 1983), and *N,N*-di(sinapoyl)spermidine was characterized from pollen of *Hippeastrum* × *hortorum* (Youhnovski et al. 1998).

UGT84A1-A4 enzymes mediate transient response to UV-B exposure

The transcript abundances in leaves exposed to UV-B stress for 10 h indicated that the induced biosynthesis of sinapoylmalate was mainly triggered by up-regulation of the genes encoding UGT84A2, UGT84A3 and UGT84A1, respectively, whereas expression of the final enzyme in sinapoylmalate formation, SMT encoded by *SNG1*, seems not to be involved. While—in accordance with gene expression—the SGT activity of leaves was found to be remarkably increased after 10 h of UV-B exposure, only trace activity was found in leaves exposed for 6 days

although the transcript abundance of both UGT84A2 and UGT84A3 were still kept at induced levels. This may point to a post-transcriptional mechanism involved in down-regulation of sinapate ester biosynthesis during long-term UV-B exposure. Regarding metabolite accumulation, the continuously increasing amount of flavonoids was in contrast to that of sinapate esters which exhibited a decline under long term UV-B treatment. This indicates that UGT84A-mediated sinapate ester biosynthesis is involved in the short-term UV-B response whereas the increased formation of flavonoids, particularly of quercetin derivatives, contributes to a sustained adaptation strategy in *Arabidopsis*.

Conclusions

The UGT84A1-A4 enzymes display a remarkable functional overlap in plant HCA metabolism. This enables the plant to maintain the gross metabolite flux toward the major sinapate esters, sinapine and sinapoylmalate, when individual UGT84A1-A4 genes are mutated. Mutation of the UGT84A2-encoding gene was accompanied by minor changes in the seed metabolic network including the increased channeling of sinapate into several choline esters and spermidine derivatives and by the decrease of ferulate derivatives. Our results tentatively suggest a specific role for UGT84A3 in the accumulation of cell wall-bound 4-coumarate. During UV-B exposure, UGT84A1-A4 enzymes mediate the transient increase in the UV-screening compound sinapoylmalate. Our results indicate that sinapoylmalate plays a role mainly in the transient UV-B stress response whereas flavonoid derivatives are continuously involved and confer long-term protection.

Acknowledgments The authors thank Dr. Werner Heller (German Research Center for Environmental Health, Neuherberg, Germany) for help with the UV-B exposure experiments and for inspiring discussions during the course of this work. The technical assistance of Franziska Götsch is greatly acknowledged. This study was supported by the Deutsche Forschungsgemeinschaft (Bonn, Germany).

References

Bieza K, Lois R (2001) An *Arabidopsis* mutant tolerant to lethal ultraviolet-B levels shows constitutively elevated accumulation of flavonoids and other phenolics. Plant Physiol 126:1105–1115

Bharti AK, Khurana JP (1997) Mutants of *Arabidopsis* as tools to understand the regulation of phenylpropanoid pathway and UVB protection mechanisms. Photochem Photobiol 65:765–776

Baumert A, Milkowski C, Schmidt J, Nimtz M, Wray V, Strack D (2005) Formation of a complex pattern of sinapate esters in Brassica napus seeds, catalysed by enzymes of a serine carboxypeptidase-like acyltransferase family. Phytochemistry 66:1334–1345

Booij-James IS, Dube SK, Jansen MAK, Edelman M, Mattoo AK (2000) Ultraviolet-B radiation impacts light-mediated turnover of the photosystem II reaction center heterodimer in *Arabidopsis* mutants altered in phenolic metabolism. Plant Physiol 124:1275–1284

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254

Caldwell MM (1971) Solar ultraviolet radiation and the growth and development of higher plants. In: Giese AC (ed) Photophysiology. Academic Press, New York, p 6 131–177

Chapple CC, Vogt T, Ellis BE, Somerville CR (1992) An *Arabidopsis* mutant defective in the general phenylpropanoid pathway. Plant Cell 4:1413–1424

Clauß K, Baumert A, Nimtz M, Milkowski C, Strack D (2008) Role of a GDSL lipase-like protein as sinapine esterase in Brassicaceae. Plant J 53:802–813

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J 16:735–743

Döhring T, Köfferlein M, Thiel S, Seidlitz HK (1996) Spectral shaping of artificial UV-B irradiation for vegetation stress research. J Plant Physiol 148:115–119

Grubb CD, Zipp BJ, Ludwig-Müller J, Masuno MN, Molinski TF, Abel S (2004) *Arabidopsis* glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis. Plant J 40:893–908

Iiyama K, Lam T, Stone BA (1994) Covalent cross-links in the cell wall. Plant Physiol 104:315–320

Jones P, Messner B, Nakajima J-I, Schäffner AR, Saito K (2003) UGT73C6 and UGT78D1, glycosyltransferases involved in flavonol glycoside biosynthesis in *Arabidopsis thaliana*. J Biol Chem 278:43910–43918

Hause B, Meyer K, Viitanen PV, Chapple C, Strack D (2002) Immunolocalization of 1-*O*-sinapoylglucose:malate sinapoyltransferase in *Arabidopsis thaliana*. Planta 215:26–32

Hood EE, Gelvin SB, Melchers LS, Hoekema A (1993) New *Agrobacterium* helper plasmids for gene transfer to plants. Transgenic Res 2:208–218

Hüsken A, Baumert A, Strack D, Becker HC, Möllers C, Milkowski C (2005) Reduction of sinapate ester content in transgenic oilseed rape (*Brassica napus*) by dsRNAi-based suppression of *BnSGT1* gene expression. Mol Breed 16:127–138

Kesy JM, Bandurski RS (1990) Partial purification and characterization of indol-3-ylacetylglucose:myo-inositol indol-3-ylacetyltransferase (indoleacetic acid-inositol synthase). Plant Physiol 94:1598–1604

Kowalczyk S, Bandurski RS (1991) Enzymic synthesis of 1-*O*-(indol-3-ylacetyl)-beta-D-glucose. Purification of the enzyme from *Zea mays*, and preparation of antibodies to the enzyme. Biochem J 279:509–514

Landry LG, Chapple CCS, Last RL (1995) *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. Plant Physiol 109:1159–1166

Larsen LM, Olsen O, Plöger A, Sorensen H (1983) Sinapine-*O*-β-D-glucopyranoside in seeds of *Alliaria officinalis*. Phytochemistry 22:219–222

Lehfeldt C, Shirley AM, Meyer K, Ruegger MO, Cusumano JC, Viitanen PV, Strack D, Chapple C (2000) Cloning of the *SNG1* gene of *Arabidopsis* reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. Plant Cell 12:1295–1306

Li J, Ou-Lee T-M, Raba R, Amundson RG, Last RL (1993) *Arabidopsis* flavonoid mutants are hypersensitive to UV-B irradiation. Plant Cell 5:171–179

Lim E-K, Li Y, Parr A, Jackson R, Ashford DA, Bowles DJ (2001) Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in *Arabidopsis*. J Biol Chem 276:4344–4349

Lois R (1994) Accumulation of UV-absorbing flavonoids induced by UV-B radiation in *Arabidopsis thaliana* L. Planta 194:498–503

Lorenzen M, Racicot V, Strack D, Chapple C (1996) Sinapic acid ester metabolism in wild type and a sinapoylglucose-accumulating mutant of *Arabidopsis*. Plant Physiol 112:1625–1630

Milkowski C, Baumert A, Schmidt D, Nehlin L, Strack D (2004) Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. Plant J 38:80–92

Milkowski C, Baumert A, Strack D (2000a) Cloning and heterologous expression of a rape cDNA encoding UDP-glucose:sinapate glucosyltransferase. Planta 211:883–886

Milkowski C, Baumert A, Strack D (2000b) Identification of four *Arabidopsis* genes encoding hydroxycinnamate glucosyltransferases. FEBS Lett 486:183–184

Mittasch J, Strack D, Milkowski C (2007) Secondary product glycosyltransferases in seeds of *Brassica napus*. Planta 225:515–522

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15:473–497

Outchkourov NS, Peters J, de Jong J, Rademakers W, Jongsma MA (2003) The promoter–terminator of chrysanthemum rbcS1 directs very high expression levels in plants. Planta 216:1003–1012

Pagani F, Romussi G (1970) Composition of *Cleome punges*. III. Principal chemical components of the plant and the synthesis of related compounds of potential pharmacological interest. Farmaco (Sci) 25:727–748

Piber M, Koehler P (2005) Identification of dehydro-ferulic acid-tyrosine in rye and wheat: evidence for a covalent cross-link between arabinoxylans and proteins. J Agric Food Chem 53:5276–5284

Poppenberger B, Fujioka S, Soeno K, George GL, Vaistij FE, Hiranuma S, Seto H, Takatsuto S, Adam G, Yoshida S, Bowles D (2005) The UGT73C5 of *Arabidopsis thaliana* glucosylates brassinosteroids. Proc Natl Acad Sci USA 102:15253–15258

Quiel JA, Benders J (2003) Glucose conjugation of anthranilate by the *Arabidopsis* UGT74F2 glucosyltransferase is required for tryptophan mutant blue fluorescence. J Biol Chem 278:6275–6281

Reuber S, Bornman JF, Weissenböck G (1996) Phenylpropanoid compounds in primary leaf tissues of rye (*Secale cereale*): light responses of their metabolism and the possible role in UVB protection. Physiol Plant 97:160–168

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual, vol 2nd. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Shahidi F, Naczk M (1992) An overview of the phenolics of canola and rapeseed: chemical, sensory and nutritional significance. J Am Oil Chem Soc 69:917–924

Sheahan JJ (1996) Sinapate esters provide greater UV-B attenuation than flavonoids in *Arabidopsis thaliana* (Brassicaceae). Am J Bot 83:679–686

Sinlapadech T, Stout J, Ruegger MO, Deak M, Chapple C (2007) The hyper-fluorescent trichome phenotype of the *brt1* mutant of *Arabidopsis* is the result of a defect in a sinapic acid: UDPG glucosyltransferase. Plant J 49:655–668

Shirley AM, McMichael CM, Chapple C (2001) The *sng2* mutant of *Arabidopsis* is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose:choline sinapoyltransferase. Plant J 28:83–94

Strack D, Nurmann G, Sachs G (1980) Sinapine esterase. Part II. Specificity and change of sinapine esterase activity during germination of Raphanus sativus. Z Naturforsch 35c:963–966

Szerszen JB, Szczyglowski K, Bandurski RS (1994) *iaglu*, a gene from *Zea mays* involved in conjugation of growth hormone indole-3-acetic acid. Science 265:1699–1701

Tan KS, Hoson T, Masuda Y, Kamisaka S (1991) Correlation between cell wall extensibility and the content of diferulic and ferulic acids in cell walls of *Oryza sativa* coleoptiles grown under water and in air. Physiol Plant 83:397–403

Tan KS, Takayuki H, Yoshio M, Seiichiro K (1992) Involvement of cell-wall-bound diferulic acid in light-induced decrease in growth rate and cell wall extensibility of *Oryza* coleoptiles. Plant Cell Physiol 33:103–108

Tohge T, Nishiyama Y, Hirai MY, Yano M, Nakajima J, Awazuhara M, Inoue E, Takahashi H, Goodenowe DB, Kitayama M, Noji M, Yamazaki M, Saito K (2005) Functional genomics by integrated analysis of metabolome and transcriptome of *Arabidopsis* plants overexpressing an MYB transcription factor. Plant J 42:218–235

Tkotz N, Strack D (1980) Enzymatic synthesis of sinapoyl-L-malate from 1-sinapoylglucose and L-malate by a protein preparation from *Raphanus sativus* cotyledons. Z Naturforsch 35c:835–837

van Engelen FA, Moltho JW, Conner AJ, Nap JP, Pereira A, Stiekema WJ (1995) pBinplus: an improved plant transformation vector based on pBin19. Transgenic Res 4:288–290

Veit M, Pauli GF (1999) Major flavonoids from *Arabidopsis thaliana* leaves. J Nat Prod 62:1301–1302

Vicient CM, Delseny M (1999) Isolation of total RNA from *Arabidopsis thaliana* seeds. Anal Biochem 268:412–413

Vogt T, Jones P (2000) Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. Trends Plant Sci 5:380–386

von Roepenack-Lahaye E, Degenkolb T, Zerjeski M, Franz M, Roth U, Wessjohann L, Schmidt J, Scheel D, Clemens S (2004) Profiling of *Arabidopsis* secondary metabolites by capillary liquid chromatography coupled to electrospray ionization quadrupole time-of-flight mass spectrometry. Plant Physiol 134:548–559

Weier D, Mittasch J, Strack D, Milkowski C (2007) The genes *BnSCT1* and *BnSCT2* from *Brassica napus* encoding the final enzyme of sinapine biosynthesis: molecular characterization and suppression. Planta 227:375–385

Youhnovski N, Bigler L, Werner C, Hesse M (1998) On-line coupling of high-performance liquid chromatography to atmospheric pressure chemical ionization mass spectrometry (HPLC/APCI-MS and MS/MS). The pollen analysis of *Hippeastrum* × *hortorum* (Amaryllidaceae). Helv Chim Acta 81:1654–1671

FEBS Letters 580 (2006) 6366–6374

Structure determinants and substrate recognition of serine carboxypeptidase-like acyltransferases from plant secondary metabolism

Felix Stehle[a], Wolfgang Brandt[b], Carsten Milkowski[a], Dieter Strack[a,*]

[a] *Department of Secondary Metabolism, Leibniz Institute of Plant Biochemistry (IPB), Weinberg 3, D-06120 Halle (Saale), Germany*
[b] *Department of Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry (IPB), Weinberg 3, D-06120 Halle (Saale), Germany*

Received 16 October 2006; accepted 23 October 2006

Available online 10 November 2006

Edited by Ulf-Ingo Flügge

Abstract **Structures of the serine carboxypeptidase-like enzymes 1-*O*-sinapoyl-β-glucose:L-malate sinapoyltransferase (SMT) and 1-*O*-sinapoyl-β-glucose:choline sinapoyltransferase (SCT) were modeled to gain insight into determinants of specificity and substrate recognition. The structures reveal the α/β-hydrolase fold as scaffold for the catalytic triad Ser-His-Asp. The recombinant mutants of SMT Ser173Ala and His411Ala were inactive, whereas Asp358Ala displayed residual activity of 20%. 1-*O*-sinapoyl-β-glucose recognition is mediated by a network of hydrogen bonds. The glucose moiety is recognized by a hydrogen bond network including Trp71, Asn73, Glu87 and Asp172. The conserved Asp172 at the sequence position preceding the catalytic serine meets sterical requirements for the glucose moiety. The mutant Asn73Ala with a residual activity of 13% underscores the importance of the intact hydrogen bond network. Arg322 is of key importance by hydrogen bonding of 1-*O*-sinapoyl-β-glucose and L-malate. By conformational change, Arg322 transfers L-malate to a position favoring its activation by His411. Accordingly, the mutant Arg322Glu showed 1% residual activity. Glu215 and Arg219 establish hydrogen bonds with the sinapoyl moiety. The backbone amide hydrogens of Gly75 and Tyr174 were shown to form the oxyanion hole, stabilizing the transition state. SCT reveals also the catalytic triad and a hydrogen bond network for 1-*O*-sinapoyl-β-glucose recognition, but Glu274, Glu447, Thr445 and Cys281 are crucial for positioning of choline.**
© 2006 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Arabidopsis; Homology modeling; Serine carboxypeptidases; Acyltransferases; Phenylpropanoid metabolism; Molecular evolution; *Brassica napus*

1. Introduction

In various plant species, acyltransferases have been described that accept β-acetal esters (1-*O*-β-glucose esters) as acyl donors instead of coenzyme A thioesters. This alternative mechanism of acyl transfer has proven to be relevant for biosyntheses of plant secondary compounds [1-4 and the literature cited therein]. Analysis of an acyltransferase from wild tomato catalyzing the glucose ester-dependent transfer of isobutyryl moieties revealed sequence similarity to serine carboxypeptidases (SCPs) [1]. Subsequently, this unexpected classification as serine carboxypeptidase-like (SCPL) protein could be confirmed for sinapoyltransferases from phenylpropanoid metabolism of Brassicaceae plants [2–4]. This relationship indicates that glucose ester-dependent acyltransferases and serine carboxypeptidases share common ancestry and have developed by divergent evolution.

SCPs are hydrolases that catalyze the stepwise removal of C-terminal amino acids from peptides. Given their involvement in fundamental physiological processes such as programmed cell death, hormone signaling and seed development [5], it seems most likely that SCPL acyltransferases can be considered as derived enzymes which have been recruited from hydrolytic ancestors to take over acyl transfer functions in secondary metabolic pathways. This hypothesis is corroborated by the finding that plant SCPL proteins form a large and diverse group of enzymes for which different clustering of functionally proven carboxypeptidases and acyltransferases became evident [6]. It is remarkable, that SCPL acyltransferases isolated so far are not able to catalyze peptide or ester hydrolysis. This raises the fundamental question on molecular mechanisms that drive the functional shift from hydrolase to acyltransferase activity [6,7].

Sequence analyses reveal that SCPL acyltransferases harbor sequence motifs proven to be crucial for hydrolase activities of SCPs. Most notably, the amino acids that build the catalytic triad in SCPs are completely conserved in type Ser-Asp-His and relative sequence position. Moreover, inhibitors affecting the catalytic seryl residue have been shown to be effective in decreasing acyltransferase activity [1–3]. This indicates that ester-dependent acyltransferases employ the same chemistry as carboxypeptidases including nucleophilic attack onto the carbonyl carbon of the substrate that leads via a tetrahedral transition state to an acylenzyme intermediate. However, primary structure gives neither a clue on substrate specificity of SCPL acyltransferases nor on substrate recognition and reaction mechanism, which favor transacylation over hydrolysis. Thus, unraveling mechanistic aspects that drive the functional shift from hydrolase to acyltransferase would require studies on protein structure. Since large scale heterologous expression of SCPL acyltransferases has not been established so far, we

*Corresponding author. Fax: +49 345 5582 1509.
E-mail addresses: wolfgang.brandt@ipb-halle.de (W. Brandt), dieter.strack@ipb-halle.de (D. Strack).

Abbreviations: SMT, 1-*O*-sinapoyl-β-glucose:L-malate sinapoyltransferase; SCT, 1-*O*-sinapoyl-β-glucose:choline sinapoyltransferase; SCP, serine carboxypeptidase

0014-5793/$32.00 © 2006 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2006.10.046

challenged this question by homology modeling and site-directed mutagenesis.

It is quite clear that a homology modeling cannot replace the quality of a solved X-ray structure, but analyses of the models based on PROCHECK and PROSA indicate a native-like fold that of course does not exclude that some regions are slightly incorrectly modeled. This will, however, not affect the main conclusions based on the models. In addition, this seems to be a promising approach because serine carboxypeptidases used as templates for modeling studies are among the best characterized enzymes. Besides the catalytic triad, the oxyanion hole stabilizing the transition state and a hydrogen bond network for substrate recognition are functionally defined regions conserved in all serine carboxypeptidases. Sequence analyses indicate that these functional elements are also conserved in the acyltransferases, subjected to structure modeling in this work.

This study provides a first insight into structure-function relationships of SCPL acyltransferases gained by structure modeling of 1-*O*-sinapoyl-β-glucose:L-malate sinapoyltransferase (SMT; EC 2.3.1.92) and 1-*O*-sinapoyl-β-glucose:choline sinapoyltransferase (SCT; EC 2.3.1.91) as well as site-directed mutageneses of SMT. The enzymes accept the glucose ester as acyl donor and transfer the sinapoyl moiety to either L-malate (SMT), resulting in the UV-shielding compound sinapoyl-L-malate that accumulates in epidermal tissues, or to choline (SCT) giving rise to sinapoylcholine (sinapine), a characteristic seed constituent of members of the Brassicaceae (Fig. 1).

Our results indicate that SCPL acyltransferases use the SCP catalytic triad for enzymatic activity. The models reveal amino acid residues that form the oxyanion hole in SCPL acyltransferases and present evidence for protein regions most likely involved in the recognition of 1-*O*-sinapoyl-β-glucose and the acyl acceptors L-malate and choline.

2. Material and methods

2.1. Heterologous expression of AtSMT cDNA

The AtSMT cDNA was inserted as *Bam*HI–*Xba*I fragment into a multicopy derivative of the yeast plasmid pYES2 (Invitrogen). A detailed description of expression constructs and optimized culture conditions will be published elsewhere. Site-directed mutagenesis was performed with the QuickChange© XL Site-Directed Mutagenesis Kit (Stratagene) according to the protocol given by the supplier. All constructs, wildtype AtSMT and the mutant variants (Ser173Ala, His411Ala, Asp358Ala, Lys268Glu, Arg322Asp, Arg322Glu, His272Asp, Asp172Glu/Ser175Ala) were proven by sequence analysis and then used to transform competent *Saccharomyces cerevisiae* INVSc1 cells (MATa *his3Δ1 leu2 trp1-289 ura3-52*/MATα *his3Δ1 leu2 trp1-289 ura3-52*; Invitrogen), prepared with the *S. cerevisiae* EasyCom™ Kit (Invitrogen).

For expression analysis, recombinant *S. cerevisiae* cells were grown and induced according to the pYES2 manual (Invitrogen). Cells from 100-ml cultures were harvested by centrifugation and the pellets resuspended in 4 ml of ice cold lysis buffer (100 mM potassium phosphate, pH 6.0, 10 mM L-malate, 150 mM ammonium sulphate, 0.1% (v/v) Triton X-100, 1 mM EDTA, 1 mM DTT). Cell lysis was achieved by vortexing the suspension 10 times for 30 s in the presence of acid-washed glass beads (0.25–0.5 mm in diameter). Cell debris were pelleted by centrifugation (30 000 × *g*, 4 °C) and the supernatant was used for Western blot analyses and enzyme assays. Protein concentration was determined according to Bradford [8] with BSA as standard. For Western blot analysis of recombinant SMT, the protocol described by Hause et al. [9] was used.

2.2. Enzyme assays

One hundred microlitres of SMT assay mixture (100 mM L-malate, 1 mM 1-*O*-sinapoyl-β-glucose, 350 μg protein in lysis buffer) was incubated at 30 °C for 1 h. The reaction was terminated by adding TFA to a final concentration of 1% (v/v) and the product formation was analyzed by HPLC [2]. To study the effects of potential inhibitors, phenylmethylsulfonyl fluoride (PMSF) and diisopropylfluoro phosphate (DFP) were added to crude desalted protein extracts from *Arabidopsis thaliana* and *Brassica napus* leaves to a final concentration of 10 mM (PMSF) and 1 mM (DFP), respectively. After preincubation for 30 min at 30 °C, the mixtures were subjected to SMT assays.

2.3. Sequence analysis

Protein sequences were compared to sequence databases with Blastp [10] and 3D-PSSM [11,12]. For multiple sequence alignments, the program CLUSTALW with the BLOSUM62 matrix [13,14] was used.

2.4. Molecular modeling

For homology modeling of AtSMT, BnSCT and AtSCT structures the program Modeller8v1 [15] was used. The models were refined using the Charmm force field [16] with Born approximation for effective

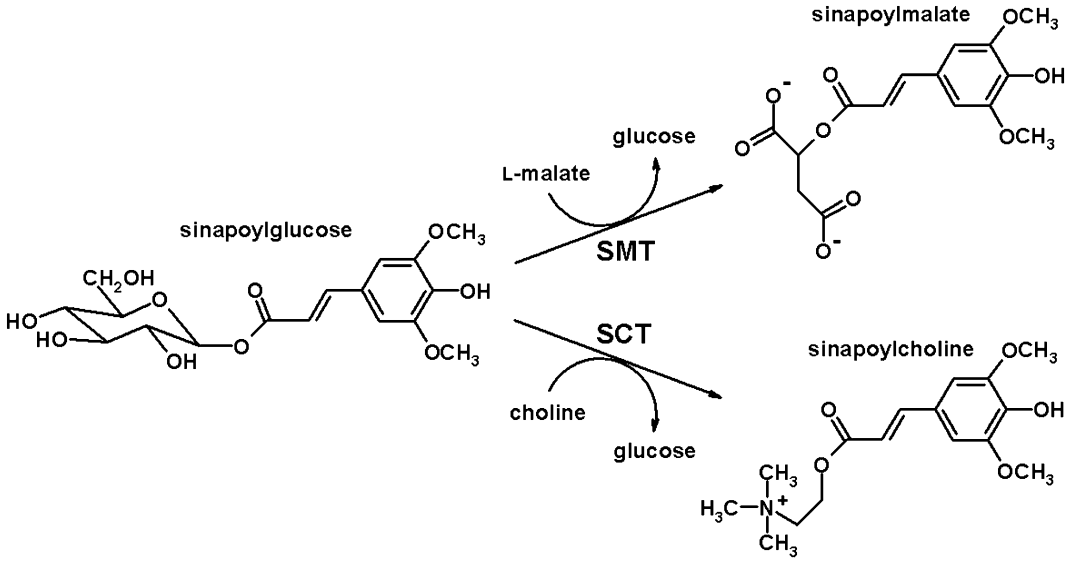


Fig. 1. Acyl transfer reactions catalyzed by SMT and SCT in members of the Brassicaceae.

electrostatics in solution [17] included in MOE© (Molecular Operating Environment). Stereochemical quality was proven by the program PROCHECK [18]. The Ramachandran Plot was used to test the backbone dihedral angle distribution of all amino acid residues. The quality of the fold was inspected with PROSA [19]. Docking studies to investigate protein–ligand interactions were performed using the automatic docking function of the program GOLD (GGOLD©[20]: Genetic Optimized Ligand Docking, Cambridge Crystallographic Data Centre, 1998, Cambridge, UK). For each ligand, 3–6 slightly different docking arrangements were produced from which the one with the best fitting score was taken for further refinement by energy minimization. For this purpose, the complete model with docked substrates was refined using the TRIPOS force field [21] and Gasteiger charges [22] implemented in the SYBYL molecular modeling package (SYBYL©[23], SYBYL© Tripos Associates Inc.) at Silicon Graphics workstations.

3. Results and discussion

3.1. Template proteins for homology modeling

Database searches for AtSMT homologues with known X-ray structures indicated the human 'protective protein' (HPP; pdb-code 1ivy) [24] with an amino acid sequence identity of 28% over a stretch of 452 residues (*e*-value of 1*e*-105) as template for structure modeling as well as the yeast serine carboxypeptidases CPY (pdb-code 1ysc) [25] and Kex1p (pdb-code 1ac5) [26] displaying identities with the query sequence of 25% (*e*-value of 2*e*-98) and 19% (*e*-value of 2*e*-77), respectively. The alignment of SCPL acyltransferases with the chosen template molecules emphasizes the overall conservation of sequence elements with proven functionality for SCPs like the catalytic triad and the oxyanion hole (Fig. 2). A highly conserved sequence block covers amino acid residues that form part of the hydrogen bond network which mediates recognition of the peptide carboxylate in SCPs [27]. To describe sequence elements unambiguously, the following nomenclature has been adopted: real amino acid sequence number in AtSMT (number in BnSCT, number in AtSCT) and BnSCT (number in AtSCT).

3.2. Model structures of AtSMT, BnSCT and AtSCT

Structure models for AtSMT and BnSCT (Figs. 3A and B) have been developed using the program MODELLER (Protein Data Bank entries 2DRF, 2DTP and 2DRG). The Ramachandran plot revealed for the AtSMT model 81.5% of backbone dihedral angles in most favored regions, whereas 16.4% were positioned in additionally allowed regions and 2.1% in generously allowed regions. For the model of BnSCT, 81.3% of backbone dihedral angles were found to occupy most favored regions; 17.8% and 0.9% were classified into the less preferred regions. All other stereo-chemical parameters were inside the quality range expected for a structure with a 2.0-Å resolution (data not shown). Probable correctness of the model folds is indicated by Prosa II plots [19] and conservation of the disulfide bridges. Previous studies indicate that mature AtSMT is a monomer [9], whereas the AtSCT is subjected to post-translational endoproteolysis generating two subunits that form a heterodimeric structure [28]. Accordingly, AtSMT is modeled as monomer and BnSCT/AtSCT as heterodimer formed via endoproteolytic excision of residues 282 (285) to 311 (310) as proposed for AtSCT [28] and described for a group of SCPs, e.g. the CPDW-II [29]. The computer-predicted central folds of AtSMT and BnSCT/AtSCT are nearly identical with those observed in SCPs [25,30,31] designated as "α/β-hydrolase fold" [32], a structure element that provides an ideal scaffold for the arrangement of several catalytic triads and has therefore been found in several enzyme groups [6].

The model structures of AtSMT and BnSCT suggest the formation of three disulfide bridges (S1, S2 and S3; Fig. 3) in both acyltransferases that should also be applied for AtSCT. The number of disulfide bridges in the template proteins varies between three for Kex1p to four in HPP and five in CPY. However, in all template proteins the disulfide bridge designated as S1 is highly conserved. In the model structures of acyltransferases S1 is formed between Cys78 (79, 82) and Cys323 (354, 352). The disulfide bond S3 can be found in all templates, but its position in the alignment is not strictly conserved. In the acyltransferase structure models S3 is formed between Cys279 (281, 284) and Cys289 (321, 320). For the heterodimeric BnSCT the model reveals that S1 and S3 are used to cross-link the subunits as was shown for CPDW-II [29]. The conservation of the position of the amino acid residues of S2 is similar to S3 but in the three-dimensional structures the positions are quiet different in all template proteins.

3.3. Catalytic triad of SMT

Sequence analysis indicated for all SCPL acyltransferases analyzed so far the presence of highly conserved amino acid residues that build the catalytic centre of SCPs. Accordingly, this suggests that AtSMT employs a catalytic triad consisting of Ser173-His411-Asp358 as charge relay system to perform acyl transfer activity (Fig. 4). Initial evidence for a seryl residue being part of the catalytic centre was gained by inhibitor studies with DFP or PMSF known to modify seryl side chains covalently [1,2]. To verify this assumption for AtSMT, protein extracts from *A. thaliana* and *B. napus* leaves were treated with both inhibitors and subjected to SMT assay. Treatment with 10 mM PMSF led to a decrease in SMT activity of 30%, whereas the presence of 1 mM DFP caused a drop of 95% compared to SMT activity in untreated extracts. This strongly suggests that Ser173 forms part of the catalytic centre in AtSMT.

To ascertain the relative impact of the individual members of the potential catalytic triad on AtSMT activity, recombinant mutant variants Ser173Ala, His411Ala and Asp358Ala were produced by *Saccharomyces cerevisiae* cells and tested for activity. The cells were harvested after cultivation in galactose-supplemented media for 36 and 48 h, respectively. The activity ratios of wild type enzyme and mutant variants at both harvest times (data not shown) indicate that the SMT cDNAs were expressed at similar levels. This was confirmed by Western blot analyses (Fig. 5A).

Fig. 5B summarizes the results from site-directed mutageneses of the catalytic triad and amino acids assumed to be responsible for substrate recognition (the following paragraph) of AtSMT. For the mutant variants Ser173Ala and His411Ala no activity was observed, whereas the mutant protein Asp358Ala kept SMT activity reduced to 20% of the wild type level. This residual activity is most probably due to the remaining ability of His411 to abstract a proton from the seryl hydroxyl group albeit restricted by the loss of stabilizing Asp358. A similar effect has been described for mutant variants of SCPs lacking the active site aspartate [33].

These results clearly indicate that Ser173, His411 and Asp358 form the reacting part in the active centre of AtSMT,

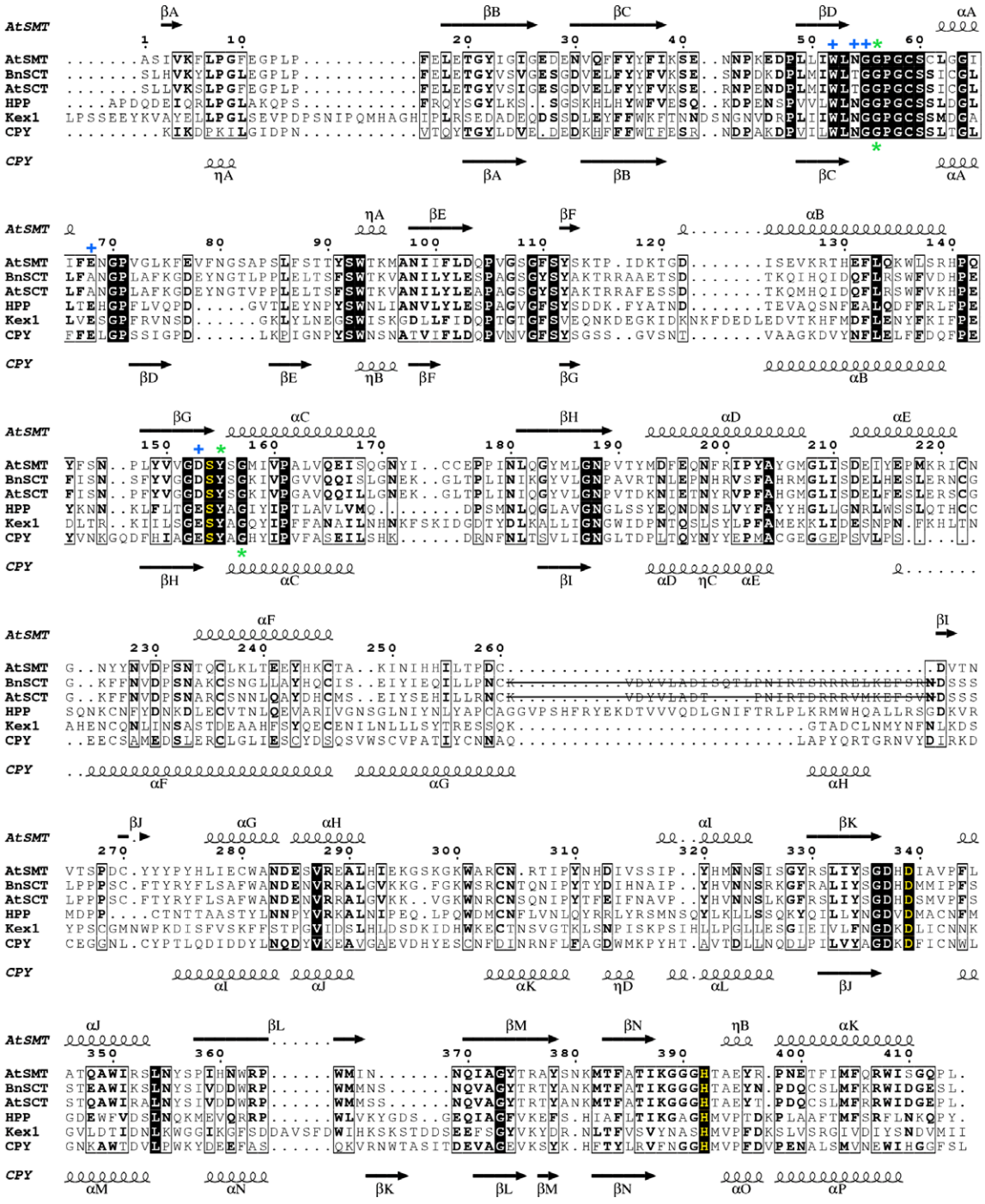


Fig. 2. Alignment of amino acid sequences of SCPL acyltransferases with those of serine carboxypeptidases used as templates for structure modeling. N-terminal leader peptides are removed. Fully conserved residues are shaded in black. Amino acid residues forming the catalytic triad are depicted in yellow. Residues are marked that are involved in substrate recognition (hydrogen bond network, blue crosses) and stabilization of the transition state (oxyanion hole, green asterisks in SMT/SCT above and in CPY below the sequences). Secondary structure elements are shown for AtSMT and CPY. Endopeptides cleaved out during enzyme maturation are crossed out. AtSMT (*Arabidopsis thaliana* SMT; GenBank accession number AF275313) was aligned with BnSCT (*Brassica napus* SCT; GenBank accession number AY383718), AtSCT (*Arabidopsis thaliana* SCT; GenBank accession number AY033947), HPP (human protective protein; NM000308.1), Kex1p (NC001139.7) and CPY (*Saccharomyces* carboxypeptidase Y; NC001145.2) by CLUSTALW using the BLOSUM62 matrix.

most probably by assembling a catalytic triad to be used as charge relay system. Given the complete conservation of these residues, this suggests that glucose ester-dependent acyltransferases have adopted the catalytic centre of SCPs to carry out acyltransferase reactions.

3.4. Substrate specificity of SMT

With regard to substrate recognition, the structure model of AtSMT provides evidence that the residues Glu215 and Arg219 (Fig. 4) play an important role in positioning the sinapoyl moiety. Glu215 accepts a hydrogen bond from the

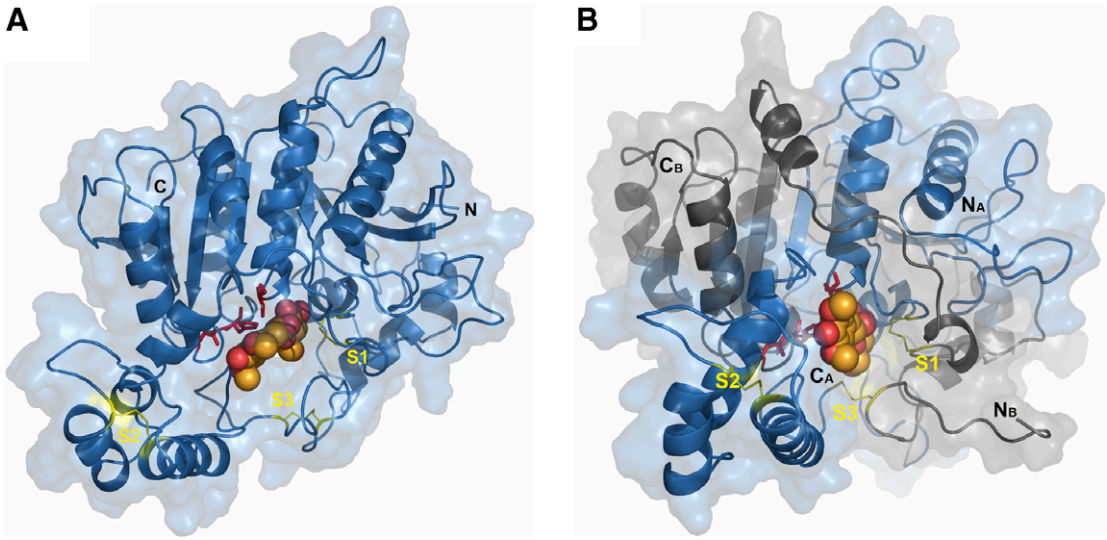


Fig. 3. (A) Proposed α/β hydrolase fold of the ternary structure of AtSMT and (B) quaternary structure of BnSCT, each with the docked ligand 1-*O*-sinapoyl-β-glucose. Disulfide bridges are labeled from S1 to S3. The catalytic triad is colored in red. The internal BnSCT sequences 282–311 was removed as indicated in the alignment (Fig. 2).

hydroxyl group at C4 of the aromatic ring, whereas Arg219 donates hydrogen bonds to the same substituent and to a neighboring methoxyl group.

Based on these results, some other hydroxycinnamate glucose esters proven to be accepted as acyldonors in SMT assays [34–36] were tested to fit into the model structure. Using the program GOLD, the potential acyl donors 1-*O*-feruloyl-, -caffeoyl- and -4-coumaroyl-β-glucose were docked to the active site. The calculated interaction energies of the individual ligands reveal a strong preference of AtSMT for the sinapoyl moiety (data not shown). Among the alternative substrates, the strengths of calculated enzyme–ligand interactions were found to be in the same range as the specificity values reported from enzyme assays [28,34–36]. Given the model structure, the alternative hydroxycinnamoyl donors appear to be less favored to establish hydrogen bonds, especially with Arg219 and Glu215.

The sequence alignment (Fig. 2) reveals that the amino acid residues building the hydrogen bond network for carboxylate recognition in position S1′ of CPY (Trp49, Asn51, Gly52, Glu65, Glu145) are conserved in AtSMT, except Glu145, which has been replaced with aspartate. This raises the question whether this conserved sequence motif could be employed for substrate recognition by AtSMT. The structure model of AtSMT indicates the formation of a hydrogen bond network resembling that of CPY. Thus, the side chain of Trp71 donates a hydrogen bond to the amide oxygen of Asn73 that on the other hand forms hydrogen bonds to the carboxylate of Glu87 and to the hydroxyl group at C3 of the glucose moiety from the substrate, 1-*O*-sinapoyl-β-glucose (Figs. 4A and B). Moreover, the hydroxyl groups at C2 and C3 of the glucose moiety donate hydrogen bonds to the carboxylate of Asp172. Unlike its counterpart in CPY, Gly74 does not contribute to the hydrogen bond network in AtSMT. The exchange of the glutamyl residue that is highly conserved among SCPs at the sequence position preceding the active site seryl residue (172) by the smaller aspartate in AtSMT affords the necessary space to accommodate the glucose moiety of the acyl donor. Remarkably, this substitution seems to be highly conserved for both proven SCPL acyltransferases and sequences that sub-cluster with these in the family of SCPL proteins [6]. Thus, the characteristic serine peptidase motif surrounding the catalytic seryl residue is considered as Gly-Glu-Ser-Tyr-Ala [37], whereas the homologous region has changed to Gly-Asp-Ser-Tyr-Ser in sinapoyltransferases. Accordingly, the double mutant variant of AtSMT Asp172Glu/Ser175Ala displayed only a residual acyl transfer activity of 15% compared to the wild type enzyme. These results suggest that the substitution of glutamate by aspartate in the catalytic centre might be the hallmark of glucose ester-dependent acyltransferases.

The model structure of AtSMT suggests that the hydrogen bond network is also used to recognize the second substrate, L-malate, once the substrate ester bond is cleaved and glucose dissociated from the active centre giving rise to the acylenzyme intermediate. Hydrogen bonds are formed between the protonated carboxyl group of L-malate and the side chains of Asn73 and Asp172 (Figs. 4C and D). This bonding pattern resembles substrate recognition that has been established between SCPs and the peptide carboxylate [27]. The mutant variant Asn73Ala where no hydrogen bond network can be formed shows only 13% residual activity (Fig. 5). This strongly indicates that the hydrogen bond network in AtSMT is still functional. However, the modeled AtSMT structure defines Arg322 as the amino acid residue with key importance for substrate recognition (Fig. 4). The side chain of Arg322 donates two hydrogen bonds to both the glucose moiety of 1-*O*-sinapoyl-β-glucose and the carboxylate of L-malate. After the glucose has left the active site, Arg322 undergoes a conformational change (Figs. 4B and C), that positions L-malate in a way that, together with the hydrogen bonds to Asn73 and Asp172, enables the catalytic His411 to abstract a proton from the hydroxyl group of the second substrate. Moreover, the thereby activated L -malate is then brought into a perfect position to attack the acylenzyme complex to form the product sinapoyl-L-malate. The outstanding importance of Arg322 is

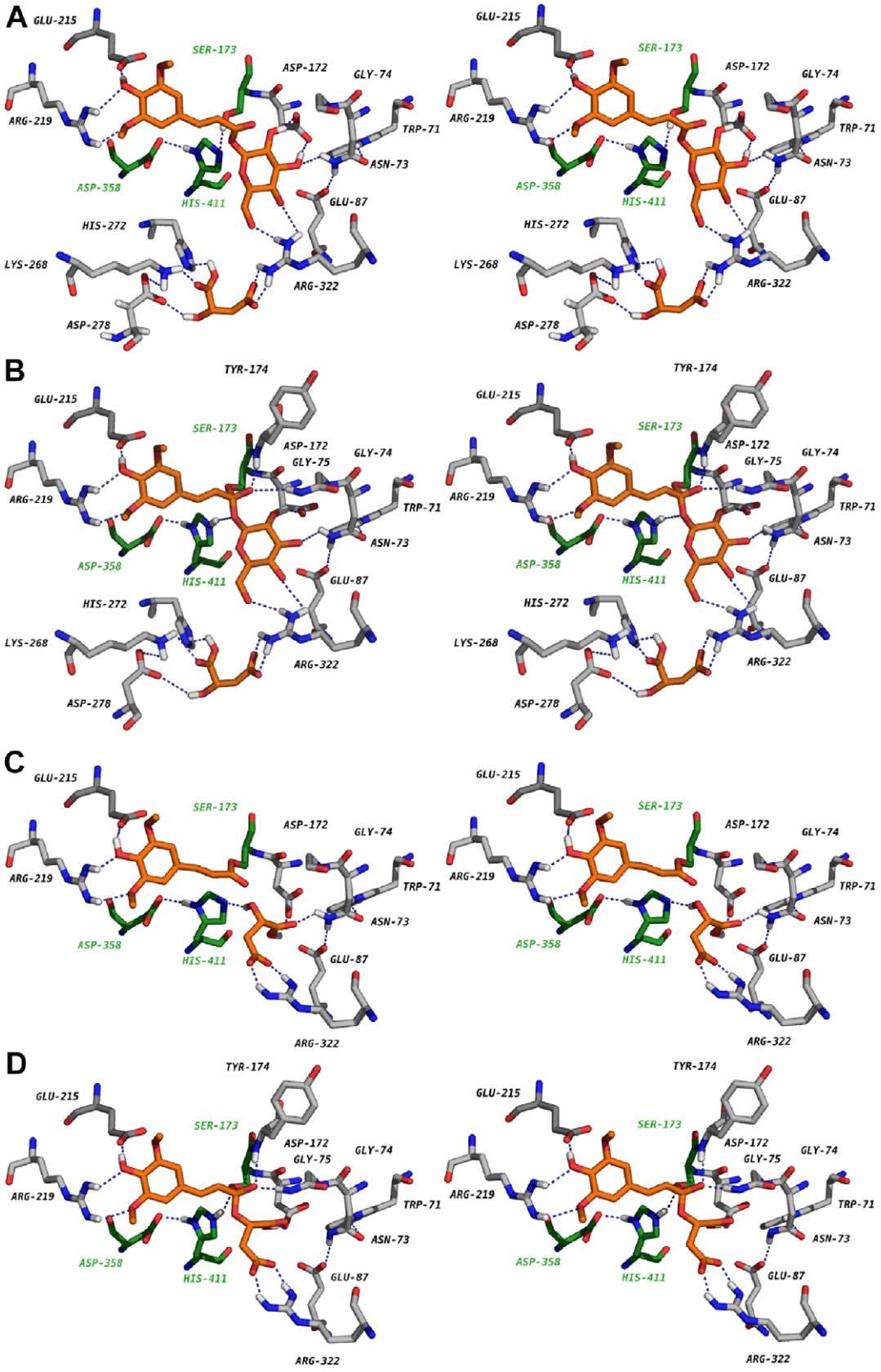


Fig. 4. Stereo pictures of AtSMT structure models with real amino acid sequence numbers. Residues forming the catalytic triad (Ser173, His411, Asp358) are labeled in green. The substrates 1-*O*-sinapoyl-β-glucose and L-malate are colored in orange. Dotted lines indicate hydrogen bonds. (A) Active site with the docking interaction between the substrates 1-*O*-sinapoyl-β-glucose and L-malate; (B) model of a potential transition state, stabilized by the oxyanion hole where hydrogen bonds are formed between the negatively charged tetrahydral state oxygen and the backbone nitrogens of amino acid residues Gly75 and Tyr174; (C) acylenzyme complex where the second substrate L-malate is activated by abstraction of a proton by the catalytic His411; (D) model of the potential second transition state, stabilized by the oxyanion hole as described above.

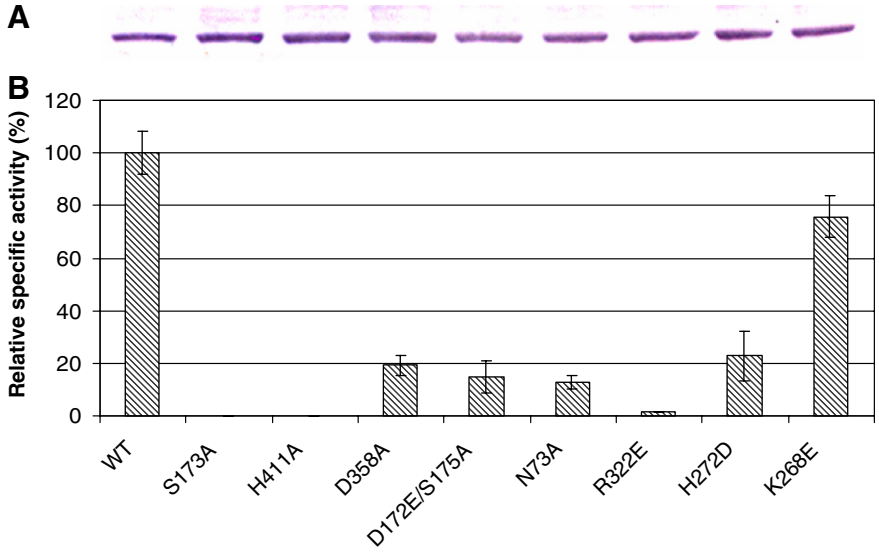


Fig. 5. (A) Western blots of recombinant SMT wild type (WT) and mutant variants with rabbit polyclonal monospecific anti-SMT antibodies according to Hause et al. [9]. Lanes were loaded with 20 μg crude protein extracts from SMT-transformed yeast cells. Blots correspond to the relative specific activities of the recombinant enzymes in panel (B). Mean values (±S.D.) were calculated from three independent enzyme assays; 100% = 14 pkat/mg protein.

corroborated by the fact that the mutant variant Arg322Glu of AtSMT is nearly inactive with a residual activity of 1%.

Inspecting the model structure, it became obvious that three additional amino acid residues may be involved in the initial recognition of the second substrate L-malate: His272, Lys268 and Asp278 (Figs. 4A and B). Activity assays of mutant variants His272Asp and Lys268Glu indicate that both residues play a role in substrate recognition (Fig. 5). The substitution His272Asp led to a decrease of SMT activity by 78% with regard to the wild type enzyme. The finding that the His272 homologous residue in CPY (Tyr256) forms part of the S2′ binding pocket supports the importance of this sequence element. The enzyme activity of AtSMT mutant variant Lys268Glu was found to be reduced by 25%.

Crystal structures of SCPs reveal the oxyanion hole as an important structural feature required for stabilization of the negatively charged tetrahedral transition state. In CPY the oxyanion hole (compare Figs. 4B and D) is formed by the backbone amide hydrogens of Gly53 and Gly149 [25]. Accordingly, the AtSMT structure model suggests that the oxyanion

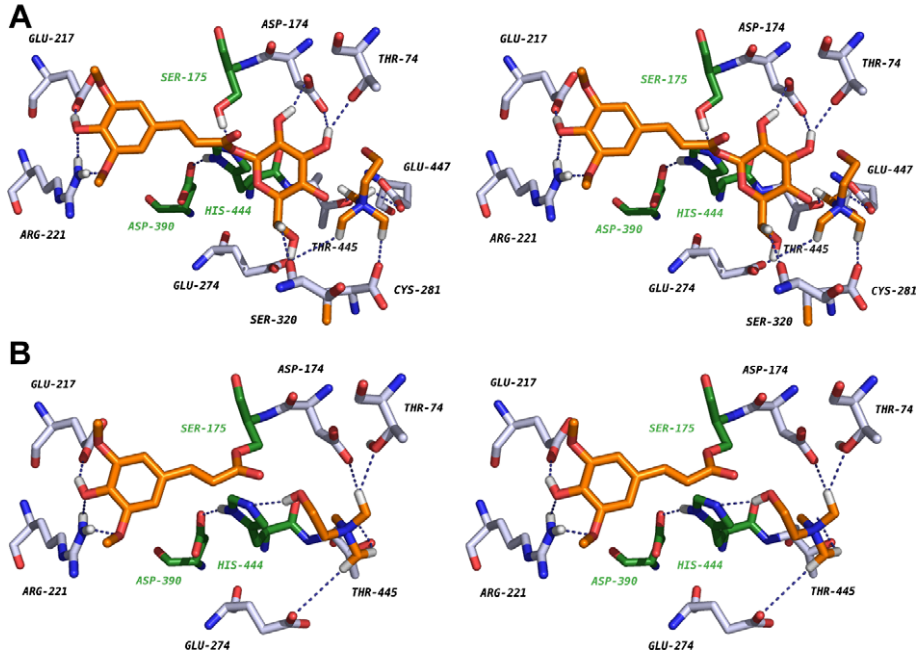


Fig. 6. Stereo pictures of BnSCT structure models with real amino acid sequence numbers, structures are colored as in Fig. 4. (A) Proposed active site of BnSCT with the docked substrates 1-*O*-sinapoyl-β-glucose and choline and (B) the proposed BnSCT acylenzyme complex.

hole is built by the backbone amide hydrogens of Gly75 and Tyr174. The alignment defines Gly75 (real sequence numbers) of AtSMT as homologous residue to Gly53 in CPY (Fig. 2). The second constituent of the oxyanion hole in AtSMT (Tyr174) has obviously been shifted by two positions towards the active site seryl residue compared to CPY.

An important question to be addressed is why the acylenzymes are not hydrolyzed analogous to the peptide hydrolysis of SCPs. As speculated earlier [6], there may be either an exclusion of water from the catalytic centre by hydrophobic shielding or a higher nucleophilic potential (if activated by the catalytic histidine and aspartate) of the second substrates, L-malate or choline, than water. There is no reliable explanation yet. However, one possible explanation for the prevention of hydrolysis could be the deprotonized state of Asn73 in the AtSMT model. This would favor the binding of an oxonium ion that is unable to attack the acylenzyme complex.

3.5. Structural features of BnSCT

Structure models of both BnSCT and AtSCT suggest that these acyltransferases mainly employ amino acid residues for catalytic function that are homologous to those implicated in AtSMT activity. Thus, recognition of the acyl donor 1-*O*-sinapoyl-β-glucose is facilitated through Glu217 in BnSCT (220 in AtSCT) and Arg221 (224) by establishing hydrogen bonds to the sinapoyl moiety (Fig. 6). The glucose moiety is recognized by hydrogen-bonding Thr74 (77) and Asp174 (177). Both these residues are homologous to those involved in glucose recognition of AtSMT (Asn73, Asp172). Additionally, Ser320 (319) was found to stabilize the glucose moiety in SCT enzymes. Furthermore, the residues which form the oxyanion hole Tyr176 (179) and Gly89 (92) correspond to those of AtSMT. However, a striking difference between SCT and SMT structure models involves binding of the second substrate choline. The choline moiety is first recognized by the negatively charged side chains of Glu274 (277), Glu447 (445), Cys281 (284) and by Thr445 (443) (Fig. 6A), whereas Glu274 (277) is the homologous residue to His272 (AtSMT) and Tyr256 (CPY) forming parts of the S2′ binding pocket. After cleavage of the ester bond of the acyl donor leading to dissociation of glucose from the active centre, choline moves towards this position forced by Asp174 (177) and is activated by the catalytic histidine similar to L-malate in AtSMT (Fig. 6B).

4. Conclusion

This study provides structure models of the plant SCPL acyltransferases SMT and SCT based on homology with serine carboxypeptidases that are among the best characterized proteins. The models confirm the main functional elements conserved within the SCPL protein family, i.e. hydrolase fold, catalytic triad, oxyanion hole and substrate recognition site. This indicates overall reliability of the models even though the placement of single amino acid side chains might intrinsically be problematic with modeling approaches. Site-directed mutagenesis supports the model-based hypothesis about structure–function relationships of SMT, provided that the substitutions introduced do not disturb the general structure of the protein.

Acknowledgement: Research on SCPL acyltransferases is supported by the DFG priority program 1152, "Evolution of Metabolic Diversity".

References

[1] Li, A.X. and Steffens, J.C. (2000) An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. Proc. Natl. Acad. Sci. USA 97, 6902–6907.
[2] Lehfeldt, C., Shirley, A.M., Meyer, K., Ruegger, M.O., Cusumano, J.C., Viitanen, P.V., Strack, D. and Chapple, C. (2000) Cloning of the SNG1 gene of Arabidopsis reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. Plant Cell 12, 1295–1306.
[3] Shirley, A.M., McMichael, C.M. and Chapple, C. (2001) The sng1 mutant of Arabidopsis is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose:choline sinapoyltransferase. Plant J. 28, 83–94.
[4] Milkowski, C., Baumert, A., Schmidt, D., Nehlin, L. and Strack, D. (2004) Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. Plant J. 38, 80–92.
[5] Schaller, A. (2004) A cut above the rest: the regulatory function of plant proteases. Planta 220, 183–197.
[6] Milkowski, C. and Strack, D. (2004) Serine carboxypeptidase-like acyltransferases. Phytochemistry 65, 517–524.
[7] Steffens, J.C. (2000) Acyltransferases in protease's clothing. Plant Cell 12, 1253–1255.
[8] Bradford, H. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254.
[9] Hause, B., Meyer, K., Viitanen, P.V., Chapple, C. and Strack, D. (2002) Immunolocalization of 1-*O*-sinapoylglucose:malate sinapoyltransferase in *Arabidopsis thaliana*. Planta 215, 26–32.
[10] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
[11] Fischer, D., Barret, C., Bryson, K., Elofsson, A., Godzik, A., Jones, D., Karplus, K.J., Kelley, L.A., MacCallum, R.M., Pawlowski, K., Rost, B., Rychlewski, L. and Sternberg, M. (1999) CAFASP-1: critical assessment of fully automated structure prediction methods. Proteins: Struct. Funct. Genet. S3, 209–217.
[12] Kelley, L.A., MacCallum, R.M. and Sternberg, M.J.E. (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM. J. Mol. Biol. 299, 499–520.
[13] Henikoff, S. and Henikoff, J.G. (1992) Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89, 10915–10919.
[14] Henikoff, S. and Henikoff, J.G. (1993) Performance evaluation of amino acid substitution matrices. Proteins 17, 49–61.
[15] Sali, A. and Blundell, T.L. (1993) Comparative protein modeling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779–815.
[16] MacKerell Jr., A.D., Bashford, D., Bellott, M., Dunbrack Jr., R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher III, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Miorkiewicz-Kuczera, J., Yin, D. and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B 102, 3586–3616.
[17] Onufriev, A., Bashford, D. and Case, D.A. (2000) Modification of the generalized born model suitable for macromolecules. J. Phys. Chem. B 104, 3712–3720.
[18] Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. 26, 283–291.
[19] Sippl, M.J. (1990) Calculation of conformational ensemble from potentials of mean force. An approach to the knowledge-based prediction of local structures in globular proteins. J. Mol. Biol. 213, 859–883.

[20] GOLD, Cambridge Crystallographic Data Centre (CCDC). Available from: http://www.ccdc.cam.ac.uk/.
[21] Clark, M., Cramer Jr., R.D. and van Opdenbosch, N.J. (1989) Validation of the general purpose tripos 5.2 Force Field. J. Comput. Chem. 10, 982–1012.
[22] Gasteiger, J. and Marsili, M. (1980) Iterative partial equalization of orbital electronegativity – a rapid access to atomic charges. Tetrahedron 36, 3219–3228.
[23] SYBYL© Tripos Associates Inc., S. L., MO, USA. Available from: http://www.tripos.com.
[24] Rudenko, G., Bonten, E.J., d'Azzo, A. and Hol, W.G.J. (1995) Three-dimensional structure of the human 'protective protein': structure of the precursor form suggests a complex activation mechanism. Structure 3, 1249–1259.
[25] Endrizzi, J.A., Breddam, K. and Remington, S.J. (1994) 2.8-Å Structure of yeast serine carboxypeptidase. Biochemistry 33, 11106–11120.
[26] Shilton, B.H., Li, Y., Tessier, D., Thomas, D.Y. and Cygler, M. (1996) Crystallization of a soluble form of the Kex1p serine carboxypeptidase from *Saccharomyces cerevisiae*. Protein Sci. 5, 395–397.
[27] Mortensen, U.H., Remington, S.J. and Breddam, K. (1994) Site-directed mutagenesis on (serine) carboxypeptidase Y. A hydrogen bond network stabilizes the transition state by interaction with the C-terminal carboxylate group of the substrate. Biochemistry 33, 508–517.
[28] Shirley, A.M. and Chapple, C. (2003) Biochemical characterization of sinapoylglucose:choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism. J. Biol. Chem. 278, 19870–19877.
[29] Breddam, K., Sørensen, S.B. and Svendsen, I. (1987) Primary structure and enzymatic properties of carboxypeptidase II from wheat bran. Carlsberg Res. Commun. 52, 297–311.
[30] Liao, D.J., Breddam, K., Sweet, R.M., Bullock, T. and Remington, S.J. (1992) Refinded atomic model of wheat serine carboxypeptidase II at 2.2-Å resolution. Biochemistry 31, 9796–9812.
[31] Bullock, T.L., Branchaud, B. and Remington, S.J. (1994) Structure of the complex of L-benzylsuccinate with wheat serine carboxypeptidase II at 2.0 Å resolution. Biochemistry 33, 11127–11134.
[32] Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G. and Goldman, A. (1992) The α/β hydrolase fold. Protein Eng. 5, 197–211.
[33] Craik, C.S., Roczniak, S., Largman, C. and Rutter, W.J. (1987) The catalytic role of the active site aspartic acid in serine proteases. Science 237, 909–913.
[34] Gräwe, W. and Strack, D. (1986) Partial purification and some properties of 1-sinapoylglucose:choline sinapoyltransferase (sinapine synthase) from seeds of *Raphanus sativus* L. and *Sinapis alba* L.. Z. Naturforsch. 41c, 28–33.
[35] Gräwe, W., Bachhuber, P., Mock, H.P. and Strack, D. (1992) Purification and characterization of sinapoylglucose:malate sinapoyltransferase from *Raphanus sativus*. Planta 187, 236–241.
[36] Vogt, T., Aebersold, R. and Ellis, B. (1993) Purification and characterization of sinapine synthase from seeds of *Brassica napus*. Arch. Biochem. Biophys. 300, 622–628.
[37] Derewenda, U., Brzozowski, A.M., Lawson, D.M. and Derewenda, Z.S. (1992) Catalysis at the interface: the anatomy of the conformational change in a triglyceride lipase. Biochemistry 31, 1532–1541.

Phytochemistry 69 (2008) 1826–1831

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Activities of Arabidopsis sinapoylglucose:malate sinapoyltransferase shed light on functional diversification of serine carboxypeptidase-like acyltransferases

Felix Stehle [a], Wolfgang Brandt [b], Jürgen Schmidt [b], Carsten Milkowski [a,1], Dieter Strack [a,*]

[a] Department of Secondary Metabolism, Leibniz Institute of Plant Biochemistry (IPB), Weinberg 3, D-06120 Halle (Saale), Germany
[b] Department of Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry (IPB), Weinberg 3, D-06120 Halle (Saale), Germany

ARTICLE INFO

Article history:
Received 14 January 2008
Received in revised form 25 February 2008
Available online 17 May 2008

Keywords:
Serine carboxypeptidase-like acyltransferase
Arabidopsis
Brassicaceae
Gene cluster
Molecular evolution
Homology modeling

ABSTRACT

Analysis of the catalytic properties of the serine carboxypeptidase-like (SCPL) 1-*O*-sinapoyl-β-glucose:L-malate sinapoyltransferase (SMT) from Arabidopsis showed that the enzyme exhibits besides its primary sinapoylation of L-malate, minor hydrolytic and disproportionation activities, producing free sinapic acid and 1,2-di-*O*-sinapoyl-β-glucose, respectively. The ability of the enzyme to liberate sinapic acid from the donor molecule 1-*O*-sinapoyl-β-glucose indicates the existence of a short-lived acylenzyme intermediate in the proposed random sequential bi–bi mechanism of catalysis. SMT-catalyzed formation of disinapoylglucose has been corroborated by docking studies with an established homology structure model that illustrates the possible binding of two 1-*O*-sinapoyl-β-glucose molecules in the active site and the intermolecular reaction of the two glucose esters. The *SMT* gene is embedded in a tandem cluster of five SCPL sinapoyltransferase genes, which encode enzymes with high amino acid sequence identities and partially overlapping substrate specificities. We assume that in recent duplications of genes encoding SCPL proteins, neofunctionalization of the duplicates to accept 1-*O*-sinapoyl-β-glucose as acyl donor was gained first, followed by subfunctionalization leading to different acyl acceptor specificities.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Decisive steps in the evolution of genetic systems are gene duplication and subsequent neo- and subfunctionalization of the enzymes encoded by gene duplicates (Ohno, 1970; Roth et al., 2007). Thus, new enzymes emerge during speciation, often forming large gene families encoding enzymes with either new stringent functions or extended substrate specificities. Genetic dynamics is especially important in the evolution of plant secondary metabolism (Hartmann, 2007). The extraordinary diversity that characterizes plant chemistry is brought about by gene families coding for enzymes that modify various structure skeletons of phenylpropanoids, isoprenoids, alkaloids or polyketides, amounting to more than 200,000 compounds so far. Besides hydroxylation, methylation, glycosylation or prenylation, acylation particularly plays an important role to gain versatile properties of secondary metabolites. The latter includes reactions with acetate, malonate, malate or hydroxybenzoates and hydroxycinnamates.

In various plant species, transfer of acyl groups is driven by energy-rich β-acetal esters (1-*O*-β-glucose esters) instead of the more common coenzyme A thioesters (Steffens, 2000; Strack and Mock, 1993). An isobutyroyltransferase from wild tomato (Li and Steffens, 2000) and sinapoyltransferases of members of the Brassicaceae, i.e. 1-*O*-sinapoyl-β-glucose:L-malate sinapoyltransferase (SMT) and 1-*O*-sinapoyl-β-glucose:choline sinapoyltransferase (SCT) (Lehfeldt et al., 2000; Shirley et al., 2001; Fraser et al., 2007; Weier et al., 2008), use the β-acetal esters as acyl donors. These acyltransferases have sequence similarities to serine carboxypeptidases (SCPs), classifying them as serine carboxypeptidase-like (SCPL) proteins. The Arabidopsis genome encodes a large family of annotated SCPL proteins with a distinguished number of proteins that cluster with proven SCPL acyltransferases (Milkowski and Strack, 2004; Fraser et al., 2005).

The high overall sequence identity between the SCPL acyltransferases and their presumed origin from SCPs poses questions on molecular mechanisms that drive the functional shift from hydrolase to acyltransferase activity. First studies with artificial peptides established no concomitant peptidase activities for glucose ester-dependent acyltransferases (Li and Steffens, 2000; Shirley and Chapple, 2003).

The *SMT* gene in Arabidopsis is embedded in a tandemly arranged cluster of five SCPL acyltransferase genes, and it has been discussed that encoded enzymes of this cluster may exhibit overlapping substrate specificities (Fraser et al., 2007). Analysis of accumulating sinapate esters in transgenic Arabidopsis mutants indicated that SMT may also catalyze the formation of 1,2-di-*O*-sinapoyl-β-glucose (**3**), which is the defined activity of

* Corresponding author. Fax: +49 345 5582 1509.
E-mail address: dieter.strack@ipb-halle.de (D. Strack).
[1] Present address: Institute of General Botany, University Jena, Am Planetarium 1, D-07743 Jena, Germany.

0031-9422/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2008.03.021

1-*O*-sinapoyl-β-glucose:1-*O*-sinapoyl-β-glucose sinapoyltransferase (SST) from *Raphanus sativus* (Dahlbender and Strack, 1986) and Arabidopsis (Fraser et al., 2007).

We report herein on the variable activities of the recombinant Arabidopsis SMT that produces not only sinapoyl-L-malate (**2**) but also free sinapic acid (**4**) and 1,2-di-*O*-sinapoyl-β-glucose (**3**). To substantiate the latter, we performed docking studies applying our previously established SMT structure model (Stehle et al., 2006). The resulting docking arrangements clearly demonstrate the binding of two 1-*O*-sinapoyl-β-glucose molecules (**1**) in the active site of the enzyme and illustrate the intermolecular reaction of the two glucose esters to form the disinapoylated product. In addition, based on hydrolytic activity of the SMT, we postulate the existence of a short-lived acylenzyme intermediate in a random sequential bi–bi mechanism of enzyme catalysis. Acylenzyme formation corresponds to that of SCPs; however, the SMT kinetics is in contrast to their double displacement (ping–pong) mechanism. We conclude that the Arabidopsis sinapoyltransferases, classified as SCPL proteins, originated from more recent gene duplications and are probably still at a stage of evolutionary diversification.

2. Results and discussion

2.1. SMT mediates SST activity and hydrolysis of sinapoylglucose

In SMT assays lacking the preferred acceptor L-malate, we observed by chromatographic and diode array detection analyses (HPLC–DAD) two minor products. One of these products, less polar than sinapoyl-L-malate (**2**), with absorption characteristics of sinapate esters (RT, 25.5 min; λ_{max}, 330 nm), was shown to be identical with 1,2-di-*O*-sinapoyl-β-glucose (**3**) by co-chromatography with standard compounds and LC–ESI-MS/MS analysis. The data are in agreement with those of the same compound isolated from *R. sativus* (Strack et al., 1984), *Brassica napus* (Baumert et al., 2005) and Arabidopsis (Fraser et al., 2007). The second minor reaction product was identified as free sinapic acid (**4**) (RT, 15.4 min; λ_{max}, 324 nm), indicating hydrolytic activity of SMT towards 1-*O*-sinapoyl-β-glucose (**1**).

The negative ion ESI mass spectrum of sinapic acid (**4**) showed a $[M-H]^-$ ion at *m/z* 223, prominent key ions at *m/z* 208 ($[M-H-Me]^-$), 193 ($[M-H-CH_2O]^-$), 179 ($[M-H-CO_2]^-$) and a base peak ion at *m/z* 164 ($[M-H-CO_2-Me]^-$). While 1-*O*-sinapoyl-β-glucose (**1**) ($[M-H]^-$ at *m/z* 385) displayed a significant loss of 162 mass units (hexose moiety) to form the ion at *m/z* 223 ([sinapate–H]$^-$) and the base peak ion at *m/z* 205, the ESI-CID mass spectrum of 1,2-di-*O*-sinapoyl-β-glucose (**3**) ($[M-H]^-$ at *m/z* 591) is characterized by key ions at *m/z* 367 ($[M-H-sinapate]^-$) and 223.

Fig. 1 illustrates the three activities of the recombinant SMT. Formation of 1,2-di-*O*-sinapoyl-β-glucose (**3**) is the prime activity of SST, first identified from *R. sativus* (Dahlbender and Strack, 1986) and recently also proposed for Arabidopsis (Fraser et al., 2007). The activity of SST was shown to be highest near pH 8.0, with residual activities of 25% at pH 6.0 (Dahlbender and Strack, 1986). When we measured SMT activity at pH 8.0, formation of sinapoyl-L-malate (**2**) decreased to about 1%, whereas that of

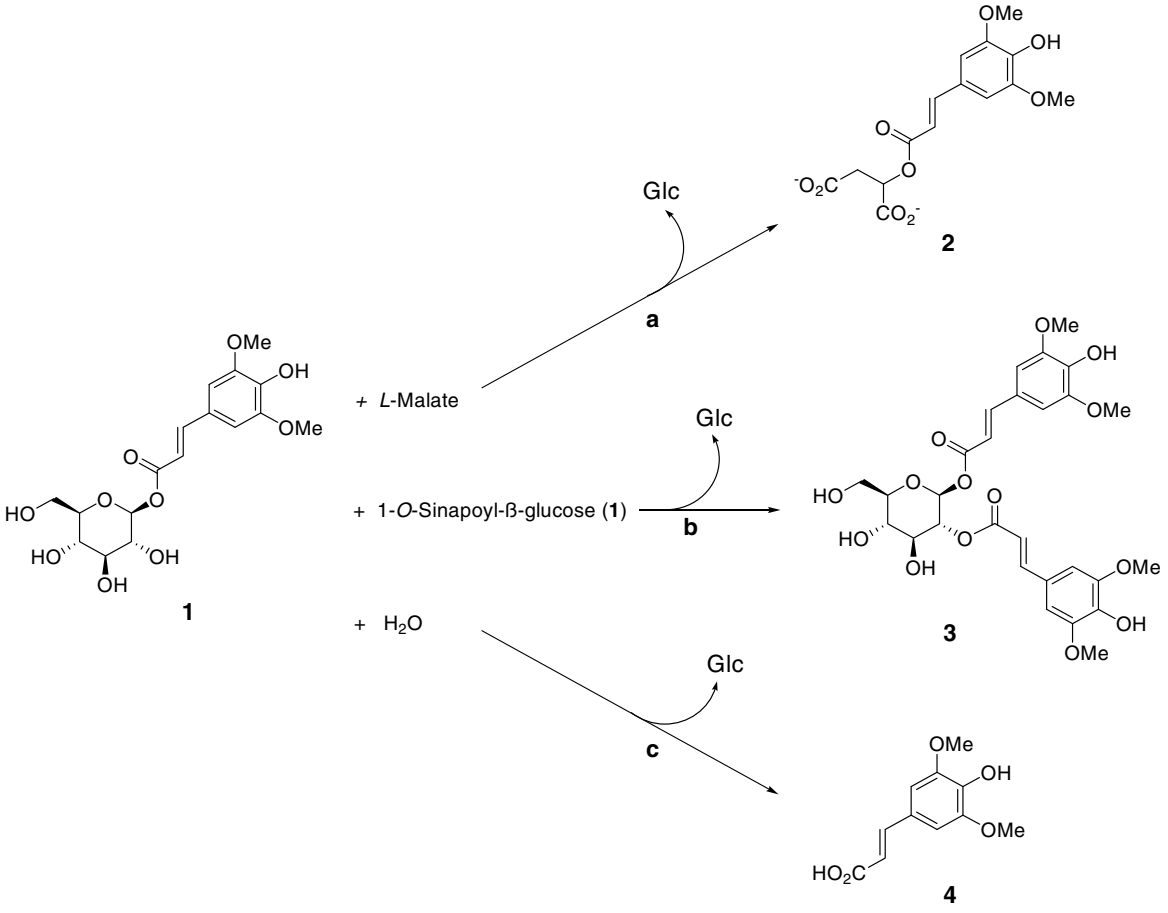


Fig. 1. Scheme of the SMT-catalyzed reactions. (a) Primary activity of SMT in the formation of sinapoyl-L-malate (**2**) with 1-*O*-sinapoyl-β-glucose (**1**) and L-malate under standard conditions (pH 6.0); (b) formation of 1,2-di-*O*-sinapoyl-β-glucose (**3**) in the absence or weak binding of L-malate at pH 8.0; (c) minor hydrolytic SMT activity under standard conditions to liberate sinapic acid (**4**), strongly increasing in the absence of L-malate and at pH 8.0 (see Table 1).

1,2-di-*O*-sinapoyl-β-glucose (**3**) increased, along with free sinapic acid (**4**). Table 1 summarizes the different SMT activities at pHs 6.0 and 8.0 with and without L-malate, using the recombinant protein produced in *Saccharomyces cerevisiae* cultures.

2.2. The SMT structure model supports SST activity

The unexpected SMT-catalyzed formation of 1,2-di-*O*-sinapoyl-β-glucose (**3**) prompted us to examine this reaction using our previously developed Arabidopsis SMT structure model (Stehle et al., 2006). Although we recently established a high yield expression system for heterologous production of active SMT (Stehle et al., 2008), we were not able yet to produce this protein in amounts that would facilitate crystallographic approaches for protein structure elucidation. However, further studies on expression optimization possibly pave the way to experimentally access structure–function relationships of the Arabidopsis SMT.

Table 1
SMT activities (pkat mg^{-1} protein; mean ± s.d.) at pHs 6.0 and 8.0 with and without L-malate, catalyzing the formation of sinapoyl-L-malate (**2**), sinapic acid (**4**) and 1,2-di-*O*-sinapoyl-β-glucose (**3**)

Substrates		Products		
Sinapoylglucose (**1**)	Malate	Sinapoylmalate (**2**)	Sinapic acid (**4**)	Disinapoylglucose (**3**)
pH 6.0				
+	+	23,069 ± 129	189 ± 37	–
+	–	–	2268 ± 18	27 ± 4
pH 8.0				
+	+	185 ± 14	1934 ± 5	173 ± 4
+	–	–	1903 ± 40	186 ± 1

Presence and absence of compounds is marked by "+" and "–", respectively.
Minor amounts of non-enzymatically liberated sinapic acid (**4**) in control assays containing heat-denatured enzyme was taken into consideration.

We performed docking studies with our homology structure model to simulate the possible binding of a second molecule of 1-*O*-sinapoyl-β-glucose (**1**). These analyses suggest that there are no marked alterations in the protein structure compared to the structure model with bound L-malate (Fig. 2A). The docking arrangement with the highest fitness-value of the second 1-*O*-sinapoyl-β-glucose (**1**), located close to the first molecule, is shown in Fig. 2B. The binding of the acyl donor is not affected and that of the acceptor molecule is achieved by the formation of only two hydrogen bonds, instead of five bonds with L-malate to Lys268, His272, Asp278 and Arg322. The first bond with the acceptor 1-*O*-sinapoyl-β-glucose (**1**) is formed between Gly75 and the C3 hydroxyl group of the glucose moiety and the second between the ε-nitrogen of Arg322 and the C3 methoxyl group of the sinapoyl moiety. This weak binding of the acyl acceptor compared to the strong fixation of L-malate may explain why measurable *in vitro* disproportionation reactions of the SMT are detectable only in the absence of L-malate.

When the acylenzyme is formed, the second 1-*O*-sinapoyl-β-glucose molecule (**1**) slightly changes the docking arrangement in the active site due to more available space (Fig. 2B). Then the acceptor molecule is stabilized by π–π-interactions between the hydrophobic benzyl ring and the imidazol ring of His272. Further stabilization of the sinapoyl moiety is achieved by hydrogen bonds between the hydroxyl and methoxyl groups and the amino acid residues Thr412 and Glu414, respectively. Arg322 forms hydrogen bonds to the carbonyl oxygen, while the glucose moiety is still fixed at Gly75. This arrangement enables the catalytic His411 to abstract a proton from the C2 hydroxyl group. As a result, the activated acceptor 1-*O*-sinapoyl-β-glucose (**1**) is in an optimal po-

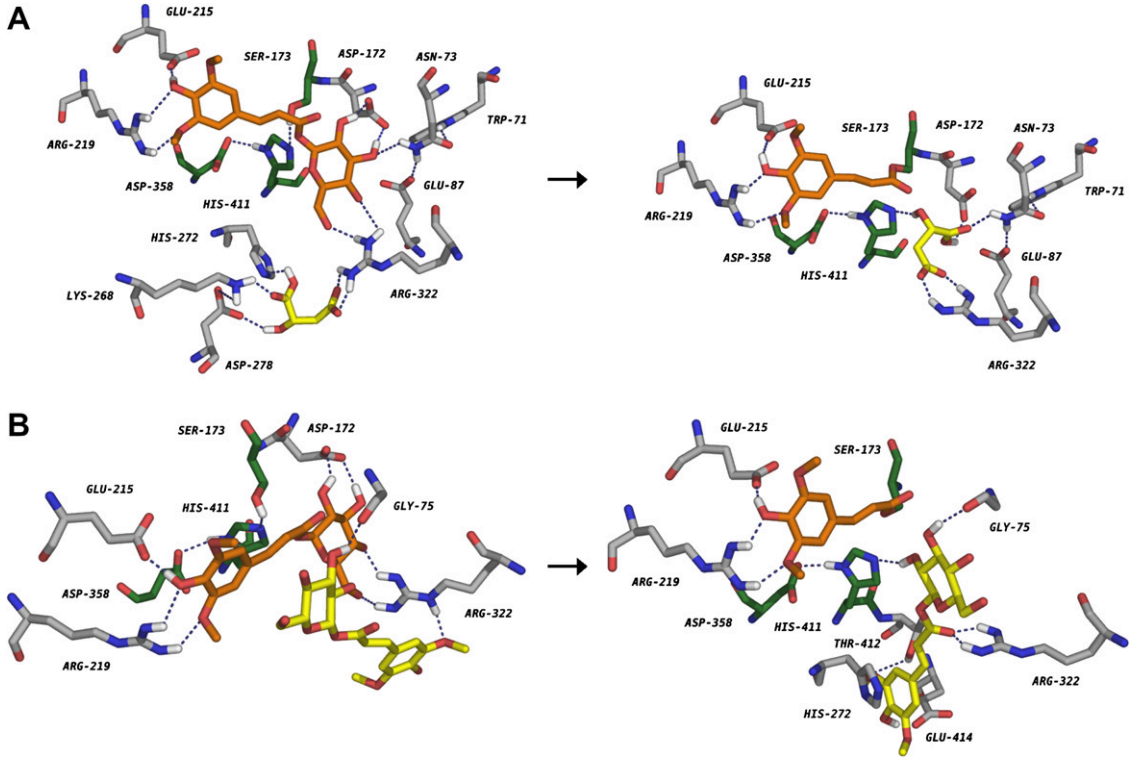


Fig. 2. Homology-based structure model of the active site of the SMT with the docked acyl donor (orange) and acyl acceptors (yellow). The residues of the catalytic triad Ser173, His411 and Asp358 are highlighted in green. Dotted lines indicate hydrogen bonds. (A) SMT with the two docked substrates 1-*O*-sinapoyl-β-glucose (**1**) and L-malate (left) and the corresponding acylenzyme complex (right) (after Stehle et al., 2006); (B) SMT with the two docked 1-*O*-sinapoyl-β-glucose molecules (**1**) (left) and the corresponding acylenzyme complex (right). For better visualization of hydrogen bonds, a different view is shown for the disproportionation reaction.

sition to attack the acylenzyme complex to form 1,2-di-*O*-sinapoyl-β-glucose (**3**).

2.3. SMT employs a short-lived acylenzyme complex

It was previously shown that the SMT reaction follows a random sequential bi–bi mechanism (Stehle et al., 2008). Both substrates have to be present in the active site before the reaction is initiated. However, in the SMT-mediated SST reaction the displacement of the weakly bound second 1-*O*-sinapoyl-β-glucose molecule (**1**) by water is facilitated. This leads to hydrolysis resulting in an increased liberation of sinapic acid (**4**) (Table 1) that points to a pivotal question concerning our model of the SMT kinetic mechanism. So far, based on modeling studies and kinetic analysis of sinapoyl-L-malate (**2**) formation (Stehle et al., 2006, 2008), we were not able to distinguish clearly between two conceivable modes for the SMT kinetic mechanism – including an acylenzyme complex (sinapoylseryl-SMT), characteristic for the serine carboxypeptidases (Douglas et al., 1976), or the catalytic serine acting as a base, described for a C–C hydrolase (Fleming et al., 2000). In standard enzyme assays (pH 6.0), containing L-malate as second substrate, only minute amounts of liberated sinapic acid (**4**) are detectable. Thus, water is probably not recognized as acyl acceptor and unable to initiate the reaction. Hence it follows that if the SMT activity would depend on a general base mechanism, hydrolysis should not proceed, given the sequential reaction mechanism of the enzyme, demanding the presence of both substrates in the active site that is not accomplished when water displaces the acyl acceptor. However, the observation of the hydrolytic reaction of SMT most likely indicates the existence of a short-lived acylenzyme intermediate that can be attacked by a water molecule.

We achieved probably weak binding of L-malate at pH 8.0, the pH optimum determined for SST activity from *R. sativus* (Dahlbender and Strack, 1986), and observed the formation of 1,2-di-*O*-sinapoyl-β-glucose (**3**) and sinapic acid (**4**) in appreciable amounts, along with sinapoyl-L-malate (**2**). At this pH value L-malate is mostly deprotonated and therefore His272 in the structure with the two docked substrates and Asp172 in the acylenzyme complex are no longer able to stabilize the acceptor by hydrogen bonds (Fig. 2A). This assumption was supported by an enzymatic activity of 185 pkat mg^{-1} protein for sinapoyl-L-malate (**2**) formation, indicating a strong decrease to a residual activity value below 1% compared to the activity at pH 6.0, whereas the hydrolytic activity increased about 12-fold (Table 1). Additionally, the SMT catalyzed the disproportionation reaction of two 1-*O*-sinapoyl-β-glucose molecules (**1**) resulting in an activity of 170 pkat mg^{-1} protein.

In enzyme assays lacking L-malate, hydrolase activity was remained at the same level at both pH values, whereas the formation of 1,2-di-*O*-sinapoyl-β-glucose (**3**) at pH 8.0 increased about sevenfold, comparable to the activity of sinapoyl-L-malate (**2**) formation at pH 8.0. These results show that SMT specificity towards L-malate is favoured by more acidic pH values (optimum at pH 6.0), which is

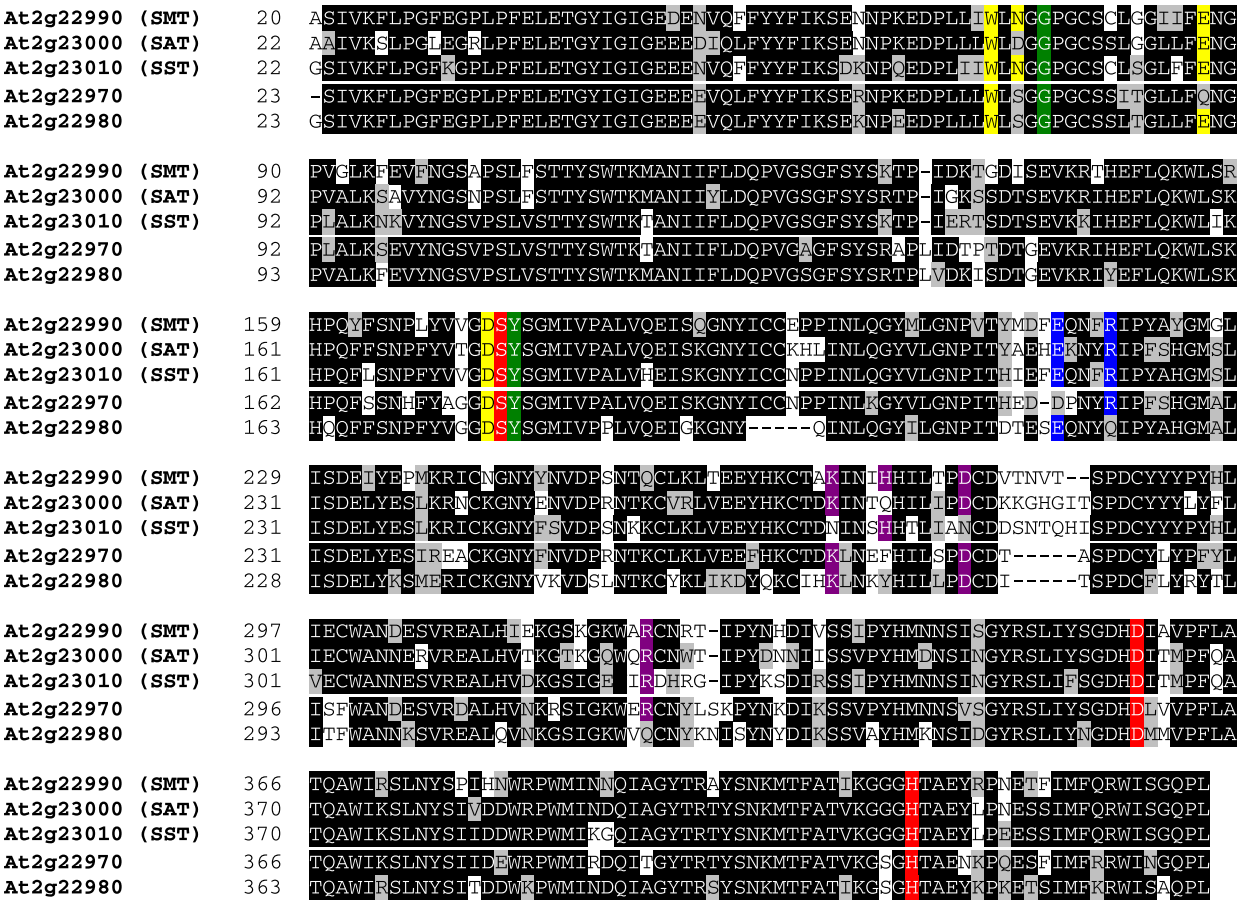


Fig. 3. Alignment of amino acid sequences of the Arabidopsis SCPL acyltransferases, encoded by the SCPL gene cluster on chromosome 2. The cDNA-encoding amino acid sequence of SMT was aligned with the sequences of the identified SAT and SST (Fraser et al., 2007) and the two putative acyltransferases encoded by At2g22970 and At2g22980 by CLUSTAL W using the BLOSUM 62 matrix (Henikoff and Henikoff, 1992, 1993). The N-terminal leader sequences have been removed. Fully conserved residues are shaded in black and those in grey indicate conservation of at least 50%. Residues shaded in different colours are part of the reaction mechanism and substrate recognition that was found out for the SMT (Stehle et al., 2006); the catalytic triad in red, the oxyanion hole in green, recognition of the sinapoyl moiety of the donor molecule in blue, the hydrogen bond network in yellow and the primary L-malate recognition in the SMT in magenta.

consistent with SMT localization to the vacuoles of Arabidopsis leaves (Hause et al., 2002). Nevertheless, SMT does not exhibit strict substrate specificity *in planta*, which was shown in analysis of an *sng1-6* transgenic line, harbouring the *SMT* gene (Fraser et al., 2007), that still accumulated 1,2-di-*O*-sinapoyl-β-glucose (**3**).

2.4. SCPL sinapoyltransferases are still on evolutionary diversification

The *SMT* gene is one of five SCPL acyltransferase genes on Arabidopsis chromosome 2. Of these genes, three have been considered as potentially mediating SST activity *in planta* (Fraser et al., 2007). While a major SST activity has been postulated for the protein encoded by At2g23010, minor SST activities have also been ascribed to the At2g22980 gene product and to SMT (At2g22990). For the latter, *in vitro* activity has been proven in a recent biochemical approach (Stehle et al., 2008).

In a homology-based SMT structure model, amino acid residues of the active site could be identified. By site-directed mutagensis these were proven to play an important role in enzyme catalysis (Stehle et al., 2006). Particularly, the residues for recognition of 1-*O*-sinapoyl-β-glucose (**1**) as well as those forming the hydrogen bond network and the oxyanion hole are conserved in the SMT-related enzymes SST and SAT (1-*O*-sinapoyl-β-glucose:anthocyanin acyltransferase) (Fig. 3). Most interestingly, these enzymes display at a conserved position the functionally important residue Asn73 (SMT), involved in binding and approximation of both 1-*O*-sinapoyl-β-glucose (**1**) and L-malate. Among the functionally important amino acid residues, only the SMT-Asn73 homolog, part of the hydrogen bond network, has changed to Asp in the SAT enzyme.

In contrast, the residues that potentially play a role for L-malate recognition of SMT (Stehle et al., 2006) are neither completely conserved in the two related acyltransferases, SAT and SST, nor in the two putative acyltransferases, encoded by At2g22970 and At2g22980 (Fraser et al., 2007). At2g22980 was assumed to encode an enzyme with minor SST activity, but the function of At2g22970 protein is unknown. The amino acid sequence derived from the At2g22970 ORF carried the conserved residues Lys268 (SMT), Asp278 (SMT) and Arg322 (SMT) involved in L-malate recognition by SMT. Moreover, the At2g22970 sequence contains the conserved residues Trp71 (SMT) and Asp172 (SMT) that form together with Asn73 (SMT) the hydrogen bond network in SMT. In At2g22970, however, the Asn73 (SMT) homolog has changed to a Ser residue.

In a current study on molecular evolution of the SCPL acyltransferase genes in Arabidopsis, we cloned the At2g22970 cDNA and were able to prove the formation of sinapoyl-L-malate (**2**) from 1-*O*-sinapoyl-β-glucose (**1**) and L-malate in activity assays containing the recombinant At2g22970 protein, when we specifically replaced some amino acids with those implicated in the SMT-like reaction (not documented). The mutant variants were generated based on the alignment of this protein with the SMT amino acid sequence (Fig. 3) and a model structure using SMT as template protein (Stehle et al., 2006). These results clearly classify this protein as another SCPL sinapoyltransferase, but a targeted metabolomics approach on Arabidopsis mutants and detailed biochemical work is needed to understand the nature of this enzyme.

The multiple alignments in Fig. 3 suggests that the most diverse sequence positions are located near the residues involved in acyl acceptor recognition, indicating that this region may be a target for evolution-driven structural alterations without changing the chemical identity of the functional residues. Interestingly, these amino acid residues are also highly conserved in the related SCT – except of those for acyl acceptor recognition (Stehle et al., 2006). This is remarkable, since the SCT clusters in the SCPL acyltransferase family (Milkowski and Strack, 2004; Fraser et al., 2005) but a post-translational endoproteolysis during enzyme maturation distinguished this protein from the other group members. Nevertheless, the present results establish the molecular basis required for gaining further insights into acyl donor and acceptor specificities for SCPL acyltransferases.

3. Concluding remarks

So far the ability to catalyze hydrolysis is only proven for the SMT, but it is most likely, that other SCPL acyltransferases retained functionally redundant hydrolytic activities. Interestingly, enzymes with bifunctional activities, transacylation and hydrolysis, have also been described from other systems. Kowalczyk et al. (2003) showed that indole-3-acetyl transferase from *Zea mays* catalyzes both synthesis and hydrolysis of indole-3-acetylinositol. In *S. cerevisiae* two acyl-coenzyme A:ethanol *O*-acyltransferases, Eht1 and Eeb1, with an α/β-hydrolase fold and a catalytic triad consisting of the amino acid residues Ser–Asp–His exhibit acyltransferase as well as esterase activities (Saerens et al., 2006). Therefore, the change from a hydrolytic α/β-hydrolase enzyme to an acyltransferase seems to be a common phenomenon.

The different enzymatic activities of SMT suggest that evolution of this enzyme has not yet reached strict substrate specificity and loss of hydrolytic properties. We assume that *SMT* and the other four SCPL acyltransferase genes on Arabidopsis chromosome 2 originated from more recent gene duplications and might be still at a stage of evolutionary diversification. Because of the high sequence similarities of all five proteins (Fig. 3), encoded by tandemly arranged genes, and the fact that the enzymes all use 1-*O*-sinapoyl-β-glucose (**1**) and exhibit partially overlapping acyl acceptor specificities, it is likely that neofunctionalization of the duplicates towards acyl donor specificity was gained first, followed by subfunctionalization resulting in specificities for different acyl acceptors.

4. Experimental

4.1. Enzyme assays

The SMT protein used in this study has been produced in *S. cerevisiae* cultures, transformed with the Arabidopsis SMT cDNA that has been optimized for high expression levels (Stehle et al., 2008). The enzyme assays with the recombinant protein contained 1 mM 1-*O*-sinapoyl-β-glucose (**1**) in the presence or absence of 10 mM L-malate in a total volume of 100 μl 100 mM MES buffer (pH 6.0) or 100 mM Tris–HCl (pH 8.0), both containing 5% (v/v) DMSO. After incubation at 30 °C at varying time intervals, the reaction was terminated by adding TFA to a final concentration of 10% (v/v). Product formation was analyzed by HPLC. After centrifugation, assay aliquots were injected onto a Nucleosil C18 column (5 μm; 250 × 4 mm i.d.; Macherey-Nagel, Düren, Germany). Separation was achieved using a 40-min linear gradient at a flow rate of 1 ml min^{-1} from 10% to 40% MeCN in 1.5% aq. H_3PO_4. Compounds were photometrically detected (maxplot between 210 and 500 nm) by a Waters 2996 photodiode array detector. Co-chromatography was achieved by adding standard compounds (Baumert et al., 2005) to the terminated enzyme assay before HPLC analysis. The quantitative data, calculated as equivalents of 1-*O*-sinapoyl-β-glucose (**1**), represent mean values (±s.d.) from three independent measurements.

4.2. LC–ESI-MS/MS

The negative ion ESI mass spectra of sinapic acid (**4**), 1-*O*-sinapoyl-β-glucose (**1**) and 1,2-di-*O*-sinapoyl-β-glucose (**3**) were obtained from a Finnigan MAT TSQ Quantum Ultra AM system equipped with a hot ESI source (HESI), electrospray voltage 3.0 kV, sheath gas: nitrogen; vaporizer temperature: 50 °C;

capillary temperature: 250 °C. The MS system is coupled with a Surveyor Plus micro-HPLC (Thermo Electron), equipped with a Ultrasep ES RP18E column (5 μm, 1 × 100 mm, SepServ). HPLC separation was achieved by using a 15-min linear gradient at a flow rate of 50 μl min^{-1} from 10% to 95% MeCN in 0.2 % aq. HOAc, with the latter held at 95% MeCN for another 15 min. The collision-induced dissociation (CID) mass spectra of sinapic acid (**4**) and 1-*O*-sinapoyl-β-glucose (**1**) were recorded during the HPLC run with a collision energy of 15 eV for the $[M-H]^-$-ions at *m*/*z* 223 (**4**) and 385 (**1**) by using a skimmer voltage of 25 eV, respectively. For 1,2-di-*O*-sinapoyl-β-glucose (**3**) a collision energy of 20 eV was used (collision gas: argon; collision pressure: 1.5 mTorr).

Sinapic acid (**4**): RT_{HPLC} = 11.7 min, ESI-CID mass spectrum (*m*/*z*, rel. int. (%)): 223 ($[M-H]^-$, 41), 208 (97), 193 (39), 179 (12), 164 (100), 149 (77), 121 (15).

1-O-Sinapoyl-β-glucose (**1**): RT_{HPLC} = 4.6 min, ESI-CID mass spectrum (*m*/*z*, rel. int. (%)): 385 ($[M-H]^-$, 18), 265 (3), 247 (10), 223 (24), 205 (100), 190 (22).

1,2-Di-O-sinapoyl-β-glucose (**3**): RT_{HPLC} = 15.2, ESI-CID mass spectrum (*m*/*z*, rel. int. (%)): 591 ($[M-H]^-$, 14), 367 ([M−H–sinapate]$^-$, 84), 223 ([sinapate–H]$^-$, 100), 205 (6).

4.3. Sequence analysis

The multiple sequence alignment was generated with the program CLUSTAL W (Thompson et al., 1994) using the BLOSUM 62 matrix (Henikoff and Henikoff, 1992, 1993).

4.4. Docking studies

Based on the recently published SMT homology model (pdb = 2drf; Stehle et al., 2006), docking studies were preformed to predict putative binding of a second 1-*O*-sinapoyl-β-glucose molecule (**1**). For this purpose, L-malate at the active site has been replaced by 1-*O*-sinapoyl-β-glucose (**1**). Subsequently, two different docking investigations using the automatic docking function of the program GOLD (GGOLD© Genetic Optimized Ligand Docking, Cambridge Crystallographic Data Center, 1998, Cambridge, UK) were carried out. The first one was done with 1-*O*-sinapoyl-β-glucose (**1**) as acyl donor bound to the enzyme as in the original model. The second docking run was performed with the acylenzyme intermediate, where the donor sinapoyl moiety is covalently linked to the catalytic active Ser173. In each case, 30 different docking arrangements were produced.

Acknowledgement

Research on SCPL acyltransferases is supported by the DFG priority program 1152, "Evolution of Metabolic Diversity".

References

Baumert, A., Milkowski, C., Schmidt, J., Nimtz, M., Wray, V., Strack, D., 2005. Formation of a complex pattern of sinapate esters in *Brassica napus* seeds, catalyzed by enzymes of a serine carboxypeptidase-like acyltransferase family? Phytochemistry 66, 1334–1345.

Dahlbender, B., Strack, D., 1986. Purification and properties of 1-(hydroxycinnamoyl)-glucose:1-(hydroxycinnamoyl)-glucose hydroxycinnamoyltransferase from radish seedlings. Phytochemistry 25, 1043–1046.

Douglas, K.T., Nakagawa, Y., Kaiser, E.T., 1976. Mechanistic studies of carboxypeptidase Y. Kinetic detection of an acyl-enzyme intermediate in trimethylacetate esterase action. J. Am. Chem. Soc. 98, 8231–8236.

Fleming, S.M., Robertson, T.A., Langley, G.J., Bugg, T.D., 2000. Catalytic mechanism of a C-C hydrolase enzyme: evidence for a gem-diol intermediate, not an acyl enzyme. Biochemistry 39, 1522–1531.

Fraser, C.M., Rider, L.W., Chapple, C., 2005. An expression and bioinformatics analysis of the Arabidopsis serine carboxypeptidase-like gene family. Plant Physiol. 138, 1136–1148.

Fraser, C.M., Thompson, M.G., Shirley, A.M., Ralph, J., Schoenherr, J.A., Sinlapadech, T., Hall, M.C., Chapple, C., 2007. Related Arabidopsis serine carboxypeptidase-like sinapoylglucose acyltransferases display distinct but overlapping substrate specificities. Plant Physiol. 144, 1986–1999.

GOLD, Cambridge Crystallographic Data Centre (CCDC), 1998. Available from: <http://www.ccdc.cam.ac.uk/>.

Hartmann, T., 2007. From waste products to ecochemicals: fifty years research of plant secondary metabolism. Phytochemistry 68, 2831–2846.

Hause, B., Meyer, K., Viitanen, P.V., Chapple, C., Strack, D., 2002. Immunolocalization of 1-*O*-sinapoylglucose:malate sinapoyltransferase in *Arabidopsis thaliana*. Planta 215, 26–32.

Henikoff, S., Henikoff, J.G., 1992. Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89, 10915–10919.

Henikoff, S., Henikoff, J.G., 1993. Performance evaluation of amino acid substitution matrices. Proteins 17, 49–61.

Kowalczyk, S., Jakubowska, A., Zielínska, E., Bandurski, R.S., 2003. Bifunctional indole-3-acetyl transferase catalyses synthesis and hydrolysis of indole-3-acetyl-*myo*-inositol in immature endosperm of *Zea mays*. Physiol. Plant. 119, 165–174.

Lehfeldt, C., Shirley, A.M., Meyer, K., Ruegger, M.O., Cusumano, J.C., Viitanen, P.V., Strack, D., Chapple, C., 2000. Cloning of the *SNG1* gene of Arabidopsis reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. Plant Cell 12, 1295–1306.

Li, A.X., Steffens, J.C., 2000. An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. Proc. Natl. Acad. Sci. USA 97, 6902–6907.

Milkowski, C., Strack, D., 2004. Serine carboxypeptidase-like acyltransferases. Phytochemistry 65, 517–524.

Ohno, S., 1970. Evolution by Gene Duplication. Springer-Verlag, New York.

Roth, C., Rastogi, S., Arvestad, L., Dittmar, K., Light, S., Ekman, D., Liberles, D.A., 2007. Evolution after gene duplication: models, mechanisms, sequences, systems, and organisms. J. Exp. Zool. B: Mol. Dev. Evol. 308, 58–73.

Saerens, S.M., Verstrepen, K.J., Van Laere, S.D., Voet, A.R., Van Dijck, P., Delvaux, F.R., Thevelein, J.M., 2006. The *Saccharomyces cerevisiae EHT1* and *EEB1* genes encode novel enzymes with medium-chain fatty acid ethyl ester synthesis and hydrolysis capacity. J. Biol. Chem. 281, 4446–4456.

Shirley, A.M., Chapple, C., 2003. Biochemical characterization of sinapoylglucose: choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism. J. Biol. Chem. 278, 19870–19877.

Shirley, A.M., McMichael, C.M., Chapple, C., 2001. The *sng2* mutant of Arabidopsis is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose:choline sinapoyltransferase. Plant J. 28, 83–94.

Steffens, J.C., 2000. Acyltransferases in protease's clothing. Plant Cell 12, 1253–1256.

Stehle, F., Brandt, W., Milkowski, C., Strack, D., 2006. Structure determinants and substrate recognition of serine carboxypeptidase-like acyltransferases from plant secondary metabolism. FEBS Lett. 580, 6366–6374.

Stehle, F., Stubbs, M.T., Strack, D., Milkowski, C., 2008. Heterologous expression of a serine carboxypeptidase-like acyltransferase and characterization of the kinetic mechanism. FEBS J. 275, 775–787.

Strack, D., Dahlbender, B., Grotjahn, L., Wray, V., 1984. 1,2-Disinapoylglucose accumulated in cotyledons of dark-grown *Raphanus sativus* seedlings. Phytochemistry 23, 657–659.

Strack, D., Mock, H.-P., 1993. Hydroxycinnamic acids and lignins. In: Dey, P.M., Harborne, J.B. (Eds.), Methods in Plant Biochemistry, vol. 9. Academic Press, New York, NY, pp. 45–97.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acid Res. 22, 4673–4680.

Weier, D., Mittasch, J., Strack, D., Milkowski, C., 2008. The genes *BnSCT1* and *BnSCT2* from *Brassica napus* encoding the final enzyme of sinapine biosynthesis – molecular characterization and suppression. Planta 227, 375–385.

the FEBS Journal

Heterologous expression of a serine carboxypeptidase-like acyltransferase and characterization of the kinetic mechanism

Felix Stehle[1], Milton T. Stubbs[2], Dieter Strack[1] and Carsten Milkowski[1]

1 Department of Secondary Metabolism, Leibniz Institute of Plant Biochemistry (IPB), Halle (Saale), Germany
2 Institute of Biochemistry and Biotechnology, Martin-Luther-University Halle-Wittenberg, Germany

Keywords
acyltransferase; enzymatic kinetic mechanism; heterologous expression; molecular evolution; serine carboxypeptidase-like proteins

Correspondence
D. Strack, Department of Secondary Metabolism, Leibniz Institute of Plant Biochemistry (IPB), Weinberg 3, 06120 Halle (Saale), Germany
Fax: +49 345 5582 1509
Tel: +49 345 5582 1500
E-mail: dieter.strack@ipb-halle.de

(Received 13 November 2007, revised 13 December 2007, accepted 14 December 2007)

doi:10.1111/j.1742-4658.2007.06244.x

In plant secondary metabolism, β-acetal ester-dependent acyltransferases, such as the 1-*O*-sinapoyl-β-glucose:L-malate sinapoyltransferase (SMT; EC 2.3.1.92), are homologous to serine carboxypeptidases. Mutant analyses and modeling of Arabidopsis SMT (AtSMT) have predicted amino acid residues involved in substrate recognition and catalysis, confirming the main functional elements conserved within the serine carboxypeptidase protein family. However, the functional shift from hydrolytic to acyltransferase activity and structure–function relationship of AtSMT remain obscure. To address these questions, a heterologous expression system for *AtSMT* has been developed that relies on *Saccharomyces cerevisiae* and an episomal *leu2-d* vector. Codon usage adaptation of *AtSMT* cDNA raised the produced SMT activity by a factor of approximately three. N-terminal fusion to the leader peptide from yeast proteinase A and transfer of this expression cassette to a high copy vector led to further increase in *SMT* expression by factors of 12 and 42, respectively. Finally, upscaling the biomass production by fermenter cultivation lead to another 90-fold increase, resulting in an overall 3900-fold activity compared to the *AtSMT* cDNA of plant origin. Detailed kinetic analyses of the recombinant protein indicated a random sequential bi-bi mechanism for the SMT-catalyzed transacylation, in contrast to a double displacement (ping-pong) mechanism, characteristic of serine carboxypeptidases.

Plant secondary metabolism generates large amounts of low molecular weight products whose exceptional diversity results from combinatorial modification of common molecular skeletons, including hydroxylation and methylation as well as glycosylation and acylation. Accordingly, plants have evolved large gene families of modifying enzymes with distinct or broad substrate specificities. With regard to acylations, most acyltransfer reactions described so far to be involved in plant secondary metabolism are catalyzed by enzymes that accept coenzyme A thioesters [1]. As an alternative, β-acetal esters (1-*O*-acyl-β-glucoses) function as activated acyl donors. In maize, the transfer of the indolylacetyl moiety from 1-*O*-indolylacetyl-β-glucose to inositol plays a role in hormone homoeostasis [2–4] and, in Arabidopsis, the UV-protecting phenylpropanoid ester sinapoyl-L-malate is produced by transfer of the sinapoyl moiety of 1-*O*-sinapoyl-β-glucose to

Abbreviations
AtSMT, Arabidopsis SMT; CAI, codon usage adaptation index; CPY, carboxypeptidase Y; DPAP B, aminopeptidase B; ER, endoplasmic reticulum; OCH1, initiation-specific α-1,6-mannosyltransferase; PEP4, proteinase A; PHA L, phytohemagglutinin L; SCPL, serine carboxypeptidase-like; SMT, 1-*O*-sinapoyl-β-glucose:L-malate sinapoyltransferase; SRP, signal recognition particle; SST, 1-*O*-sinapoyl-β-glucose:1-*O*-sinapoyl-β-glucose sinapoyltransferase; SUC2, yeast invertase 2.

L-malate [5,6]. There are various other acyltransferases accepting β-acetal esters that have been described [7]. Investigations of these enzymes at the molecular level are so far restricted to isobutyroyl transferases from wild tomato [8] and two sinapoyl transferases from Brassicaceae, namely 1-*O*-sinapoyl-β-glucose:choline sinapoyltransferase from Arabidopsis (AtSCT; EC 2.3.1.91) [9,10] and *Brassica napus* (BnSCT) [11–13], as well as 1-*O*-sinapoyl-β-glucose:L-malate sinapoyltransferase from Arabidopsis (AtSMT; EC 2.3.1.92) [6,14]. Most interestingly, these enzymes have been characterized by sequence analyses as serine carboxypeptidase-like (SCPL) proteins, indicating the evolutionary recruitment of β-acetal ester-dependent acyltransferases from hydrolytic enzymes of primary metabolism [6,8,15]. Although the analyzed SCPL acyltransferases have maintained the nature and configuration of the Ser-His-Asp catalytic triad from hydrolases, designed to perform nucleophilic cleavage of peptide or ester bonds, these enzymes have lost hydrolytic activity towards peptide substrates [8]. Site-directed mutagenesis studies revealed that the catalytic triad, especially its nucleophilic seryl residue, is crucial for acyl transfer [14].

We have chosen the enzyme AtSMT [6] to elucidate molecular changes that convert a hydrolytic enzyme into an acyltransferase and to unravel the reaction mechanism adopted for the β-acetal ester-dependent acyl transfer (Fig. 1). Previously described functional expression assays with isobutyroyl transferase from wild tomato and AtSCT favor *Saccharomyces cerevisiae* as heterologous host for SCPL acyltransferases [8]. Similar approaches with AtSMT in our laboratory, however, resulted in a weak expression level, barely sufficient to prove and characterize enzyme activity. The previously reported functional expression of *AtSMT* in *Escherichia coli* [6] could not be confirmed in our hands. Since we were unable to refold SMT inclusion bodies produced in *E. coli*, prokaryotic expression systems does not appear to be suitable for the production of active AtSMT protein. This is in accordance with the results from structure modeling of AtSMT [14] that indicated three disulfide bridges in the protein, thus excluding correct AtSMT maturation in any prokaryotic cytose expression system. Moreover, the presence of a N-terminal leader peptide for translocation into the endoplasmic reticulum (ER), as well as the localization of the mature AtSMT enzyme to vacuoles [16], reveals post-translational modifications as being an integral part of functional *SMT* expression. Since extensive kinetic studies and crystallographic approaches essentially depend on a more efficient expression system, we optimized heterologous production of AtSMT by systematic adaptation of critical parameters-like plasmid copy number, leader peptide and codon usage. In the present study, we describe the impact of these modifications on the yield of functional AtSMT protein. In conclusion, we report on an efficient heterologous expression system for *AtSMT* in *S. cerevisiae*. The produced AtSMT was used for kinetic studies that indicate a random sequential bi-bi mechanism for the acyl transfer.

Results

Expression of *AtSMT* in different eukaryotic hosts

To identify the best-performing heterologous host for expression of *AtSMT*, insect cells and Baker's yeast were tested. For all expression constructs, the unmodified *AtSMT* cDNA was used, including the original leader peptide sequence. In *Nicotiana tabacum*, transient transformation of *AtSMT*-cDNA under control of a strong Rubisco promoter failed to produce SMT activity in transgenic leaf sectors (data not shown).

Spodoptera frugiperda Sf9 insect cells, however, infected with a baculovirus-based *AtSMT* expression vector, were shown to produce functional SMT protein. The transgenic cells excreted the recombinant

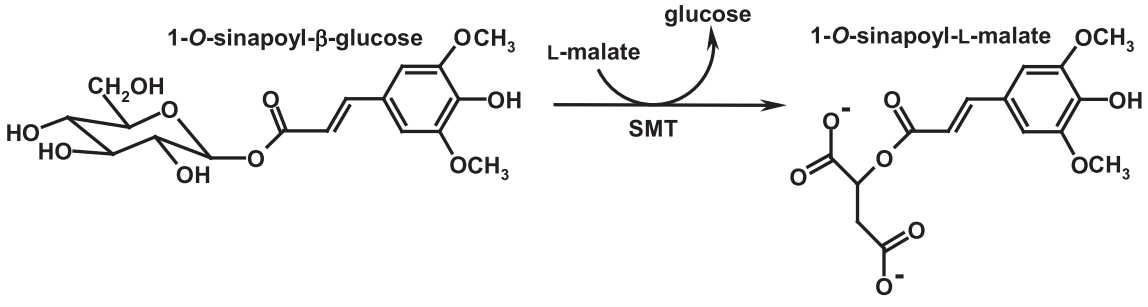


Fig. 1. Scheme of the acyltransfer reaction catalyzed by SMT.

enzyme resulting in an overall SMT activity of 220 pkat·L^{-1} culture in the growth media. Only a minor activity of approximately 6 pkat·L^{-1} culture was found as intracellular SMT activity.

Saccharomyces cerevisiae INVSc1 cells carrying the *AtSMT* cDNA fused to the *GAL1* promoter did not develop detectable SMT activities after induction by galactose (Fig. 2). This led us to optimize the sequence motif near the ATG translation initiation codon of *AtSMT* according to the consensus sequence proposed by Kozak [17]. The resulting sequence (ATAATGG) differed from the original *AtSMT* cDNA with regard to the second codon (GGT, Gly versus AGT, Ser) and conferred minimum amounts of SMT activity of approximately 20 pkat·L^{-1} culture (Fig. 2).

Although the *AtSMT* expression level in yeast was below that of *Sf9* insect cells, we decided to optimize the former system because of the well-established methods to change important expression parameters, such as cultivation conditions or gene dosage, and to upscale biomass production by fermenter cultivation for *S. cerevisiae*.

Optimization of *AtSMT* expression in *S. cerevisiae*

Sequence optimization

Efficient heterologous protein production requires that the gene to be expressed is adapted to the needs of the host organism, particularly to its codon preference calculated as codon usage adaptation index (CAI) [18]. For *S. cerevisiae*, the *AtSMT* cDNA sequence revealed a CAI of 75%. Therefore, an optimized yeast *SMT* sequence (*ySMT*) was designed with a CAI of 97% for *S. cerevisiae* (GENEOPTIMIZER software; GENEART, Regensburg, Germany; see supplementary Fig. S1). Moreover, this sequence lacks all other elements that potentially interfere with gene expression in yeast such as potential polyadenylation signals, cryptic splice donor sites and prokaryotic inhibitory sequence motifs (not documented). The *ySMT* cDNA was fused to the similarly optimized AtSMT leader sequence (*ySMT*-*ySMT*) and inserted into expression plasmid pYES2. *Saccharomyces cerevisiae* cells harboring the resulting plasmid expressed functional SMT of approximately 65 pkat·L^{-1} culture (Fig. 2). This indicates a

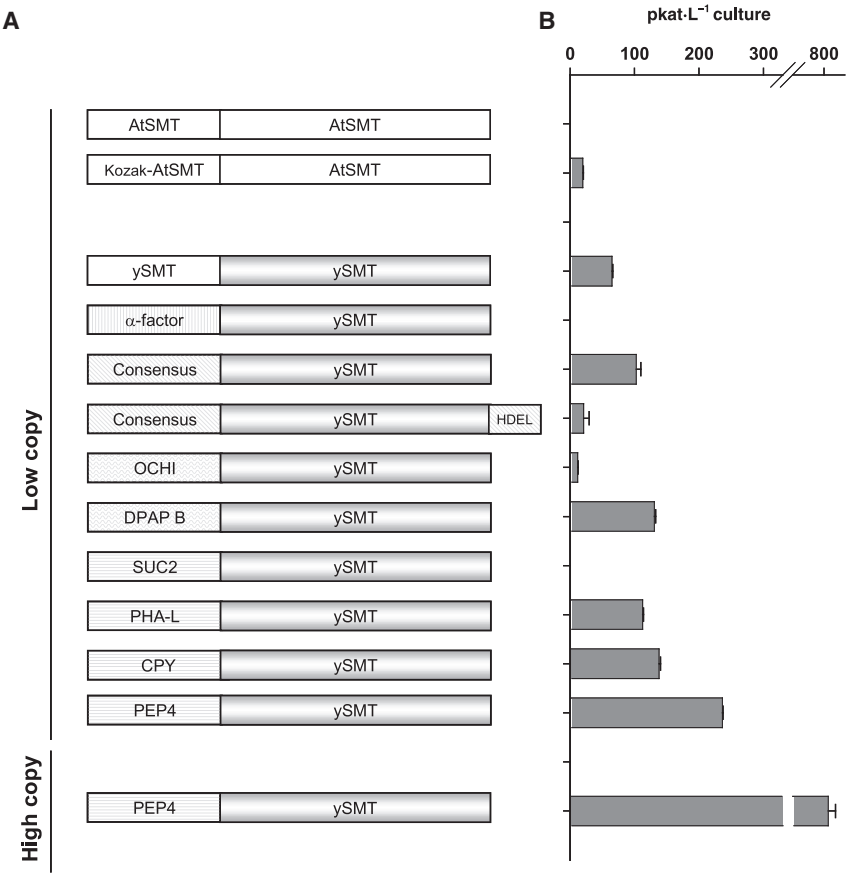


Fig. 2. Optimization of *SMT* expression in *S. cerevisiae* INVSc1. Primary structure schemes of expressed *SMT* sequence variants (A) and resulting expression strength (B) expressed as SMT activity·L^{-1} culture. The data represent the mean ± SD from three independent measurements. Kozak, Kozak-consensus sequence; α-factor, mating-factor (amino acids 1–89); Consensus, artificial consensus-signal peptide (amino acids 1–19); HDEL, ER-retention signal.

three-fold increase in SMT production with regard to the *AtSMT* sequence.

Signal peptide

In Arabidopsis, *AtSMT* is translated into a precursor protein and delivered to the ER by a 19-amino acid N-terminal signal peptide that is removed upon translocation. After folding and glycosylation, the enzyme is transported to the vacuole, most likely via the Golgi apparatus [6,16]. Since an imperfect recognition of the Arabidopsis signal peptide might account for low expression levels, we tested several leader peptides (Fig. 2) whose efficiency for heterologous protein production in yeast had been described. Signal sequences were fused to *ySMT* and inserted into plasmid pYES2 for transformation of *S. cerevisiae* INVSc1 cells.

To facilitate secretion of SMT protein into the medium, the pre–pro sequence of yeast mating pheromone α-factor [19] was tested. Expression studies, however, failed to detect SMT activity in the culture medium of transformed yeast cells.

For delivering the SMT protein to the ER, a 19-amino acid consensus signal peptide (*Consensus-ySMT*) [20] was used. This fusion led to an intracellular SMT activity in the range of 100 pkat·L^{-1} culture, indicating a 1.5-fold increase compared to the reference construct (*ySMT-ySMT*). To foster the localization of SMT into the ER, this construct was provided with a 3′-sequence extension encoding the ER retention signal HDEL [21,22]. The resulting C-terminal extension of these four amino acids led to a decrease of SMT activity by 80%.

In an approach to retain the mature SMT in specific sub-cellular compartments, the *ySMT* sequence was fused to transmembrane domains. For delivery to the Golgi apparatus and integration into the vesicle membrane, a fusion with the leader of the initiation-specific α-1,6-mannosyltransferase (OCH1; amino acids 1–30) [23] was applied. Vacuolar localization was accomplished by a partial sequence of dipeptidyl aminopeptidase B (DPAP B; amino acids 26–40) [24]. The expression levels detected were 12 pkat·L^{-1} culture with *OCH1-ySMT* and 130 pkat·L^{-1} culture with *DPAP B-ySMT*.

To deliver the mature SMT to the lumen of the yeast vacuole, we constructed N-terminal fusions with a set of signal peptides including those of yeast enzyme invertase 2 (SUC2; amino acids 1–19) [25], proteinase A (PEP4; amino acids 1–21) [26] and carboxypeptidase Y (CPY; amino acids 1–19) [27]. As a plant source, the pre–pro sequence of phytohemagglutinin L (PHA L; amino acids 1–63) [28], a seed lectin from bean (*Phaseolus vulgaris*), was used and shown to mediate SMT activity of 110 pkat·L^{-1} culture. Expression quantification revealed the highest SMT activity for the *PEP4* fusion construct (240 pkat·L^{-1} culture). This indicated an increase in production of functional SMT to approximately 400%. Medium yields were achieved with the *CPY-ySMT* fusion resulting in SMT activity of 140 pkat·L^{-1} culture, whereas the *SUC2-ySMT* construct turned out to be inactive.

With the aim of facilitating the subsequent purification of the produced SMT protein, the best-performing fusion construct (*PEP4-ySMT*) was provided with a 6xHis tag at the C-terminus. This modification, however, was shown to abolish SMT activity (not documented).

Gene dosage

To increase the copy number of the episomal 2μ expression plasmid pYES2, the *leu2-d* gene [29] was amplified from plasmid p72UG [30] and inserted into pYES2. The resulting plasmid pDIONYSOS (see supplementary Fig. S2) was shown to complement the *leu2* mutant *S. cerevisiae* INVSc1, indicating a high copy number (see supplementary Fig. S3). To demonstrate whether this increase in expression plasmid copy number would yield enhanced SMT activity via the gene dosage effect, the best performing fusion construct, *PEP4-ySMT*, was cloned into pDIONYSOS, and the resulting expression construct was used to transform *S. cerevisiae* INVSc1. The SMT activity assayed in the crude protein extract from these cells indicated a four-fold higher SMT yield compared to the pYES2-based expression of *PEP4-ySMT* (Fig. 2).

Determination of the kinetic mechanism

Increase in biomass production was obtained by fermenter cultivation of *S. cerevisiae* INVSc1 (pDIONYSOS:*PEP4-ySMT*). Cells were induced at an attenuance of 35 at $D_{600\ nm}$ and kept under inducing galactose concentrations until an attenuance of 45 at $D_{600\ nm}$ was reached. To purify the SMT activity, the protein crude extract was applied to a combination of heat treatment and chromatographic separation steps, including hydrophobic interaction, ion exchange and size exclusion techniques (Table 1). The protein fraction with the highest SMT activity was purified with a 1600-fold enrichment and a yield of 9% of the extracted enzyme activity (Fig. 3).

The *in vitro* kinetics of SMT was examined by assaying the conversion of 1-*O*-sinapoyl-β-glucose (sinapoylglucose = singlc) to 1-*O*-sinapoyl-L-malate

Table 1. Purification scheme of the recombinant SMT.

Purification step	Total protein (mg)	Total activity (nkat)	Specific activity (nkat·mg^{-1})	Enrichment (fold)	Yield (%)
Crude extract	5700	523	0.1	1	100
Heat treatment	2565	470	0.2	2	90
Butyl FF	360	163	0.5	5	31
Sephadex 200	123	89	0.7	7	17
Heat treatment	37	80	2.2	22	15
Q-Sepharose	0.3	47	157	1570	9

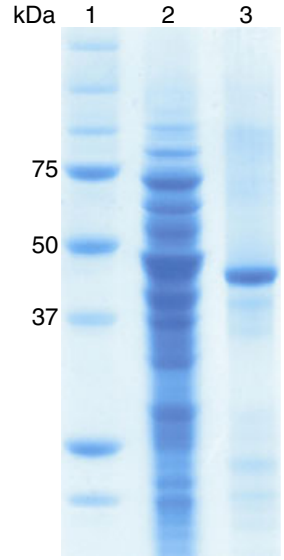


Fig. 3. Protein purification. Proteins were separated on a NuPAGE 12% Bis-Tris Gel (Invitrogen) under denaturing conditions and stained with Coomassie brilliant blue R-250. Lane 1, molecular mass markers; lane 2, *S. cerevisiae* crude cell extract; lane 3, AtSMT protein purified from *S. cerevisiae* by a combination of heat treatment and chromatographic separation steps, including hydrophobic interaction, ion exchange and size exclusion techniques.

(sinapoylmalate) in the presence of L-malate (mal). The enzymatically produced sinapoylmalate was analyzed by HPLC. Compared to previous reports [31], the change of the buffer system towards 0.1 M MES (pH 6.0) proved crucial for maintaining Michaelis–Menten kinetics over broad substrate concentration ranges (Fig. 4A,B). To prevent precipitation of the substrate sinapoylglucose or the product sinapoylmalate, the final dimethylsulfoxide concentration was adjusted to 5% (v/v) in the reaction mixtures. Dimethylsulfoxide does not interfere with SMT activity when present in concentrations of up to 8% (v/v) in the assay mixture (data not shown). To calculate the initial reaction velocities as a function of substrate concentration, the formation of sinapoylmalate was quantified at five different concentrations for both sinapoylglucose and L-malate, whereas the respective second substrate was kept constant at five different concentration levels (Fig. 4). To keep steady state conditions, reactions were stopped after 2, 4 and 6 min, respectively. Furthermore, no product inhibition could be observed when the substrates were saturated and only weak inhibition was detected when the substrates were present in the $K_{A(singlc)}$ or $K_{B(mal)}$ range (not shown).

In the double-reciprocal plots according to Lineweaver and Burk (Fig. 4, insets), the graphs were not parallel but tended to intersect. Since these graphs do not intersect at the ordinate, the maximal velocity is not constant at different substrate concentrations. Thus, the present data provide strong evidence for a random sequential bi-bi mechanism, excluding a possible order bi-bi reaction [32]. Furthermore, forcing a common intercept point using an enzyme kinetic tool (SIGMAPLOT; Systat Software, San Jose, CA, USA), the graphs fit very well with those of the measured data (not shown). The dissociation constants of the individual substrates [$K_{A(singlc)}$ and $K_{B(mal)}$] determined by Florini–Vestling plots (see supplementary Fig. S4) were found to be 115 ± 7 μM for sinapoylglucose and 890 ± 30 μM for L-malate and the ternary complex dissociation constants [$\alpha K_{A(singlc)}$ and $\alpha K_{B(mal)}$] were determined to be 3700 μM for sinapoylglucose and 12 500 μM for L-malate (see supplementary Fig. S5). The maximal catalytic activity (V_{max}) and the catalytic efficiency (k_{cat}) were found to be 370 nkat·mg^{-1} and 1.7 s^{-1}, respectively. These values (Table 2) are comparable to the kinetic parameters reported for the *Raphanus sativus* SMT [31]. In contrast to the latter, however, our data on the recombinant SMT from Arabidopsis do not support substrate inhibition by L-malate up to concentrations exceeding the $K_{B(mal)}$ value by the factor of 100 (data not shown).

Substrate specificity for L-malate

Some molecules structurally related to L-malate were tested as potential acyl acceptors or inhibitors in the SMT reaction. Activity assays reaction mixtures contained 1 mM sinapoylglucose and 10 mM or 50 mM of the related structures. Inhibition assays were performed with 10 mM of the potential inhibitors in the standard reaction mixture (1 mM sinapoylglucose and 10 mM L-malate; Table 3).

To assess the role of the L-malate carboxyl groups, (*S*)-2-hydroxyburate and (*R*)-3-hydroxybutyrate were tested as possible acyl acceptors. With regard to L-malate, a methyl group in each of these derivatives

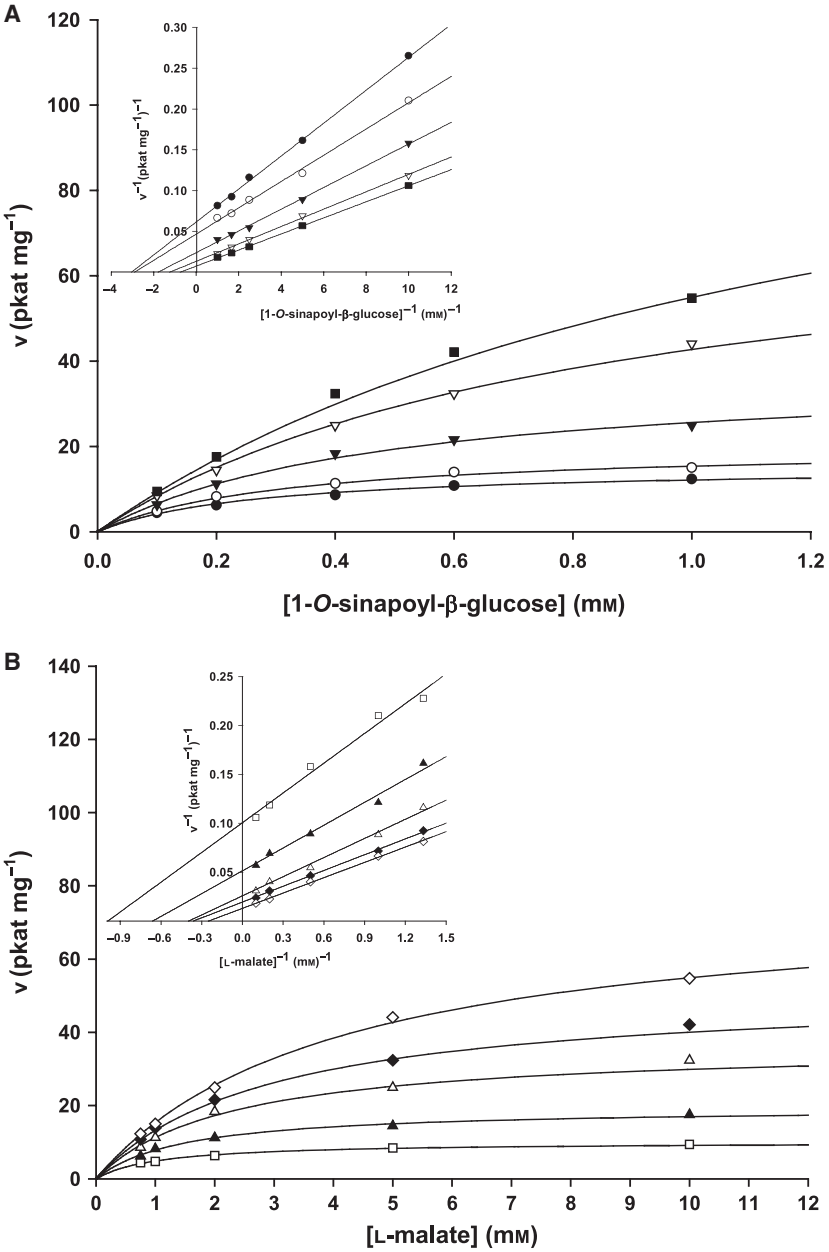


Fig. 4. v/s-Plots of SMT reaction with insets of plots displaying corresponding Lineweaver–Burk plots. Dependence of enzyme activity on sinapoylglucose concentrations in the presence of L-malate at 0.75 mM (●); 1.0 mM (○), 2.0 mM (▼), 5 mM (▽) and 10 mM (■) in (A) and on L-malate concentrations in the presence of sinapoylglucose at 0.1 mM (□), 0.2 mM (▲), 0.4 mM (△), 0.6 mM (◆) and 1 mM (◇) in (B).

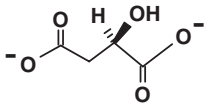
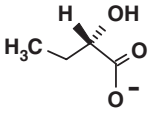
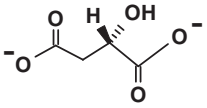
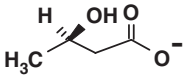
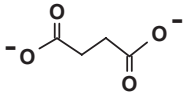
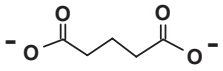
substitutes one of the two carboxyl groups, whereas the conformation of the reactive hydroxyl group is kept (cf. Table 3). SMT activity assays with these compounds failed to produce reaction products, even with incubation times of up to 60 min (not documented). This indicates that neither (*S*)-2-hydroxybutyrate nor (*R*)-3-hydroxybutyrate are suitable acyl acceptors for the SMT. However, inhibition studies revealed both of these compounds as weak, most likely competitive inhibitors decreasing the SMT activity by approximately 12% (Table 3). A slightly more effective inhibitor was glutarate with the carbon-chain elongated by one CH_2 group compared to L-malate but without a reactive hydroxyl group. Succinate, a derivative differing from L-malate only by the absence of the reactive

Table 2. Kinetic parameters of the recombinant AtSMT with sinapoylglucose and L-malate as substrates.

Substrate	K (μM)	αK (μM)	V_{max}/K (nkat·mg^{-1}·μM^{-1})	k_{cat}/K (μM^{-1}·s^{-1})
Sinapoylglucose	115 ± 7	3700[a]	3200	15
L-Malate	890 ± 30	12500[a]	420	2

[a] Standard derivation < ± 1%.

Table 3. Competitive inhibition with 10 mM of compounds structurally related to L-malate. Activities are expressed as % values (mean ± SD) compared to control assays without inhibitor (100 = 54.7 pkat·mg^{-1}).

Substrate	Inhibitor	Activity (%)
	D-(+)-Malate 	92.8 ± 1.7
	(*S*)-2-Hydroxybutyrate 	87.8 ± 0.1
L-(–)-Malate 	(*R*)-3-Hydroxybutyrate 	87.3 ± 0.7
	Succinate 	79.0 ± 0.2
	Glutarate 	84.6 ± 0.5

hydroxyl group, was the best inhibitor among the compounds tested, accounting for a decrease of SMT activity by 21%. The lowest inhibition of SMT activity was measured with the D-malate isomer.

In assays lacking L-malate, we found surprisingly a product less polar than sinapoylmalate. This compound could be identified as 1,2-di-*O*-sinapoyl-β-glucose by co-chromatography with standard compounds isolated from *B. napus* seeds [33]. The structure of this product was identified by LC-ESI-MS/MS (not documented). The MS data are in accordance with those obtained with 1,2-di-*O*-sinapoyl-β-glucose isolated from *R. sativus* [34]. Formation of this compound is catalyzed by an enzyme classified as 1-*O*-sinapoyl-β-glucose:1-*O*-sinapoyl-β-glucose sinapoyltransferase (SST) [35].

Discussion

Optimization of heterologous *AtSMT* expression

The heterologous production of functional AtSMT requires an eukaryotic expression system that facilitates post-translational processing such as the formation of disulfide bridges. Likewise, it should be accessible to upscaling procedures in order to yield protein amounts in the range required for comprehensive kinetic measurements and crystallization. For functional expression of the related sinapoyltransferase SCT, Shirley and Chapple [10] adopted the *S. cerevisiae vpl1* mutant [36], known to excrete large amounts of the homologous yeast carboxypeptidase (CPY) to the medium when expressed from a multicopy vector [30]. However, to avoid the laborious enrichment and purification procedures for protein isolation from culture medium, we decided to develop an expression system for intracellular protein production in *S. cerevisiae*. Our results revealed the codon usage of the Arabidopsis gene as well as the nature of the signal peptide and the sequence motif around the translation start as critical parameters for efficient expression of *AtSMT* in yeast. Although codon usage optimization can be calculated by CAI values, the best-performing signal peptide had to be determined empirically. We found that the signal peptide of yeast vacuolar proteinase A (PEP4) facilitated SMT expression most efficiently followed by DPAP B. Both these sequences are characterized by high hydrophobicities resembling that of the original AtSMT signal peptide. Since high hydrophobicity is correlated with the signal recognition particle (SRP)-dependent translocation [37], this suggests that SRP-dependent targeting supports SMT expression in *S. cerevisiae*. On the other hand, the SMT fusion with the SRP-dependent SUC2 signal peptide failed to express the functional enzyme, whereas the SRP-independent CPY signal sequence mediated *SMT* expression levels in the range of DPAP B. This indicates that other sequence determinants, whose

characteristics remain elusive, affect the efficiency of protein secretion and may even outperform the impact of SRP-dependence. Interestingly, C-terminal extension of the SMT sequence with both the ER retention signal and the 6xHis tag led to severe reduction of SMT activity, thus revealing the requirement of a native C-terminus.

Kinetic studies

The sinapoylglucose-dependent sinapoyltransferases SMT and SCT are homologous to SCPs. Peptide hydrolysis catalyzed by the latter follows a double displacement ping-pong mechanism. The kinetic examination of SCT from *B. napus* [11] and Arabidopsis [10] suggested that these enzymes have kept the SCP double displacement mechanism for acyl transfer. These results raise questions with regard to a proposed random bi-bi mechanism for the related SMT from *R. sativus* [31]. However, if indeed the SCT reaction follows the double displacement mechanism, it requires the formation of a sinapoylated enzyme (i.e. the acylenzyme complex) that is subsequently cleaved by the incoming acyl acceptor choline. To prevent hydrolysis of the acylenzyme, the exclusion of water is required. From the data so far available, the molecular mechanisms for water exclusion cannot be explained and will thus remain elusive until elucidation of the structure of SCT by a crystallographic approach.

The kinetic data obtained in the present study for the SMT reaction are consistent with a random sequential bi-bi mechanism (Fig. 5), partly confirming the results obtained with SMT from *R. sativus* [31]. Although the ratios of $K_{A(singlc)}/\alpha K_{A(singlc)}$ and $K_{B(mal)}/\alpha K_{B\ (mal)}$ are not equal (as is stipulated by the scheme of random binding in Fig. 5), this discrepancy can be ascribed to the partial deprotonated state of L-malate. Since there is no indication for a ping-pong mechanism, the intersections in insets Fig. 4 could not be the result of product inhibition.

Under the assay conditions applied, the interaction of SMT with L-malate may be hampered by the fact that both L-malate carboxyl groups are largely deprotonated. Thus, at pH 6.0, the C4-carboxyl group of L-malate (pK_a 3.46) should be almost completely deprotonated, whereas the C1-carboxyl group (pK_a 5.1) should be deprotonated to more than 50%. Our modeling studies as well as site-directed mutagenesis and substrate specificity analysis revealed the interaction of AtSMT with the protonated C1 carboxyl group as being essential for substrate recognition [14]. Hence, the presence of deprotonated L-malate species up to 50% should reduce the binding frequency of protonated L-malate accordingly giving rise to the apparent preference of AtSMT for sinapoylglucose in the assays. The data for substrate activation by sinapoylglucose and for substrate inhibition by L-malate from the *R. sativus* enzyme [31] could not be verified for the AtSMT.

The random sequential bi-bi mechanism of AtSMT catalysis requires both substrates, sinapoylglucose and L-malate, bound in an enzyme–donor–acceptor complex before transacylation starts. The structure homology model recently developed for AtSMT [14] supports this assumption. The formation of a very short-lived acylenzyme that is not reflected by the kinetic measurements would be accompanied by a conformational change that brings the bound acyl acceptor L-malate in a position favoring the nucleophilic attack onto the acylenzyme, as previously proposed by homology modeling [14], thus excluding water as a possible second

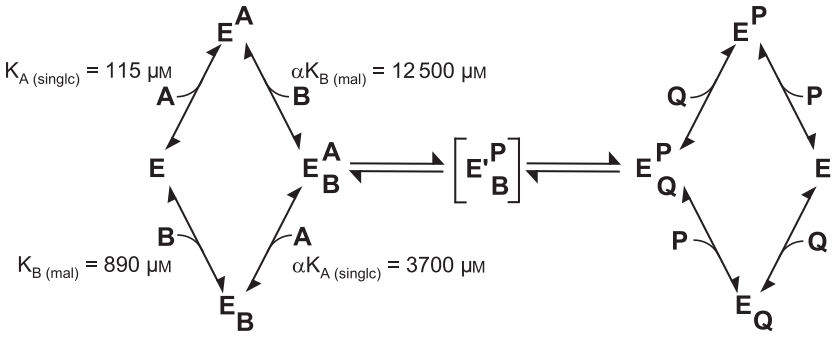


Fig. 5. Kinetic model of the SMT reaction mechanism including the putative acyl-enzyme complex. E, enzyme; A, acyl-group donor (sinapoylglucose); B, acyl-group acceptor as nucleophil (L-malate); P, released product (β-glucose); Q, released product (sinapoylmalate) of transacylation; EAB, enzyme–donor–acceptor complex; E′, putative acyl–enzyme complex; E′PB, putative acyl–enzyme–acceptor complex; $K_{A(singlc)}$, dissociation constant for sinapoylglucose and $K_{B(mal)}$ for L-malate; $\alpha K_{A(singlc)}$, ternary complex dissociation constant for sinapoylglucose and $\alpha K_{B(mal)}$ for L-malate.

substrate. However, we cannot completely exclude a different activation mode [38] involving a direct interaction with the acyl acceptor L-malate leading to proton abstraction by the active site seryl alkoxide acting as a base. The thereby activated L-malate would then attack directly the ester carbonyl of sinapoylglucose, in accordance with the postulated random sequential bi-bi mechanism.

Investigation of the substrate specificity of AtSMT towards L-malate revealed structural features required for the interaction of the acyl acceptor with the enzyme. The lack of enzymatic activity with compounds structurally related to L-malate, (*S*)-2-hydroxybutate and (*R*)-3-hydroxybutyrate, as well as the weak inhibition mediated by both compounds, indicates inadequate competitive binding to the enzyme. Hence, both carboxyl groups of L-malate appear to be crucial determinants for the interaction with the enzyme. This is corroborated by the SMT structure model that indicates recognition and binding of both carboxyl groups by hydrogen bonds [14]. Substitution of the amino acid residues Arg322 and Asn73 of SMT predicted to be mainly involved in L-malate recognition and binding resulted in strong reduction of enzyme activity. The inhibition of SMT catalysis by D-malate reveals the positioning of the reactive hydroxyl group as another structure determinant required for interaction with SMT.

Based on metabolite analysis of a transgenic SST Arabidopsis insertion mutant, it was hypothesized that SMT is able to catalyze the disproportionation of two sinapoylglucose molecules in the formation of 1,2-*O*-disinapoyl-β-glucose [39]. In the present study, we provide the biochemical proof of this enzymatic activity. Further investigations, including docking studies with the AtSMT structure model [14], will help to elucidate the molecular mechanism of this disproportionation reaction.

Conclusions

In the present study, we describe the development of a yeast expression system for heterologous production of functional SMT from Arabidopsis. A substantial increase in the yield of produced active SMT required the concerted optimization of codon usage, the N-terminal signal peptide and gene dosage. Upscaling of the produced biomass by fermenter cultivation led to the heterologous production of SMT amounts that will facilitate future crystallographic approaches for protein structure elucidation. Hence, the expression optimization described herein paves the way to experimentally access definite structure–function relationships of AtSMT whose investigation is a prerequisite for understanding the adaptation of hydrolases to catalyze acyltransfer reactions.

The kinetic characterization of AtSMT reaction revealed a random sequential bi-bi mechanism. The presence of both sinapoylglucose and L-malate in the active site may favor acyl transfer over hydrolysis by facilitating proximity. However, based on these kinetic data, at the molecular level, it is not possible to distinguish between the existence of a short-lived acyl-enzyme and a direct attack of the activated acyl acceptor L-malate.

Experimental procedures

Plant material and yeast cells

Tobacco plants (*Nicotiana tabacum* L. cv. Samsun) obtained from Vereinigte Saatzuchten eG (http://www.vs-ebstorf.de) were grown on soil under an 16 : 8 h light/dark photoperiod at 23 °C in the greenhouse. Photon flux density for all plants cultivated in the greenhouse was in the range 200–900 $\mu mol \cdot m^{-2} \cdot s^{-1}$. The *S. cerevisiae* strain INVSc1 (MATa *his3Δ1 leu2 trp1-289 ura3-52/*MATα *his3Δ1 leu2 trp1-289 ura3-52*) was obtained from Invitrogen (Carlsbad, CA, USA) and cultivated at 30 °C in synthetic or complete growth media (Sigma-Aldrich, St Louis, MO, USA) supplemented as required for *AtSMT* expression.

Oligonucleotides

Primers used to amplify SMT variants for the different expression constructs are provided in the supplementary (Table S1).

Expression of *AtSMT* in *N. tabacum*

The coding part of *AtSMT* cDNA, including 10 bp upstream the translation start, was transcriptionally fused to the promoter of Rubisco small subunit (*rbcS1*) from *Chrysanthemum morifolium* [40] by cloning into the *Not*I site of plasmid pImpact1.1 (Plant Research International, Wageningen, the Netherlands). The whole *AtSMT* expression cassette was then introduced as *Asc*I-*Pac*I fragment into the binary vector pBINPLUS (Plant Research International) [41]. The resulting *AtSMT* expression plasmid was transformed into *Agrobacterium tumefaciens* GV2260 [42] and used to transiently transform tobacco (*N. tabacum* L. cv. Samsun) by infiltration of 10-week-old leaves as described previously [43]. After 5 days of incubation, infected leaf areas were cut out for further analysis. For protein extraction, 1 g of fresh weight of leaf material was disrupted in 2 volumes of ice-cold extraction buffer (100 mM sodium phosphate, pH 6.0) by mortar and pestle.

After centrifugation at 10 000 ***g*** for 30 min at 4 °C the crude supernatant was used for SMT activity analysis.

Expression of *AtSMT* in *S. frugiperda Sf9* cells

Expression of *AtSMT* in insect cells was performed using the BD BaculoGold™ Baculovirus Expression Vector System (BD Biosciences, San Jose, CA, USA) according to the manufacturer's instructions. The *AtSMT* cDNA including 10 bp of the 5′-UTR was cloned as *Xba*I-*Not*I fragment into the baculovirus transfer vector pVL1393. The resulting plasmid was used for co-transfection of *S. frugiperda Sf9* cells together with BaculoGold baculovirus DNA. The recombinant baculovirus was amplified and used to infect freshly seeded insect cells, which were then incubated at 27 °C for 3 days. For protein extraction, cells of a 50 mL *Sf9* recombinant suspension culture with a cell density of 2×10^6 were harvested by centrifugation (5 min at 450 ***g*** and room temperature), transferred to fresh TC-100 medium (Invitrogen) and infected with 5 mL of the virus stock. After approximately one-third of the cells were lyzed (72 h of incubation), they were harvested and pelleted. The cells were resuspended in 1.5 mL of buffer (100 mM sodium phosphate buffer, pH 6.0) and disrupted with a glass homogenizer (VWR, Darmstadt, Germany). After centrifugation for 20 min at 10 000 ***g*** and 4 °C the supernatant was subjected to SMT activity analysis.

Expression of AtSMT in *S. cerevisiae*

For transformation, competent cells of *S. cerevisiae* INVSc1 were prepared using the *S. cerevisiae* EasyCom™ Kit (Invitrogen) and transformed according to the protocol given by the supplier. *Saccharomyces cerevisiae* cells harboring *AtSMT* expression plasmids were grown in synthetic drop out medium without uracil or leucine to an attenuance of 1 at $D_{600\ nm}$. Induction of *AtSMT* expression was initiated by adding galactose to a final concentration of 4% (w/v). Cells were cultivated in the presence of the inductor for additional 36 h and then harvested and disrupted as described previously [14]. For cells excreting AtSMT, the growth medium was buffered with NaOH and citric acid (pH 5.8) as described previously [30]. For protein enrichment, the culture supernatant was cleared by centrifugation and concentrated with Amicon Ultra-15 filters with a MWCO of 30 000 kDa (Millipore, Billerica, MA, USA). The 100-fold concentrated supernatant was dialyzed twice against 100 mM sodium phosphate buffer (pH 6.0) and then used for activity measurements.

Constructs for expression of SMT in *S. cerevisiae*

AtSMT cDNA variants designed for expression in *S. cerevisiae* were amplified by PCR with primers attaching restriction sites for *Hin*dIII and *Xba*I to the 5′- and 3′-ends of the product. By cloning as *Hin*dIII-*Xba*I fragments into the expression vectors pYES2 (Invitrogen) or pDIONYSOS, the PCR products were transcriptionally fused to the galactose-inducible yeast *GAL1* promoter. Nucleotide sequences encoding N-terminal signal peptides were included in forward PCR primers, except for the long pre–pro sequences of mating pheromone α-factor and *PHA-L*. Both pre–pro sequences were synthesized by GENEART and linked to the cDNA encoding the mature SMT by PCR. Modifications of the 5′-UTR were introduced via PCR by modified forward primers. Design and synthesis of the *AtSMT* sequence adapted to the codon usage of *S. cerevisiae* was performed by GENEART.

Construction of the multicopy-plasmid pDIONYSOS

The *leu2-d* marker gene was amplified from plasmid p72UG [30] by PCR with primers incorporating flanking *Bsp*HI restriction sites and cloned into the *Bsp*HI-digested 2μ plasmid pYES2 (Invitrogen).

Yeast fermentation

For recombinant protein production, *S. cerevisiae* INVSc1 cells harboring the pDIONYSOS-based *SMT* expression plasmid were cultivated in a 10 L Biostat ED fermentor (B. Braun Biotech International GmbH, Melsungen, Germany) at 30 °C and pH 5.0 in a glucose-limited growth medium [44]. During cultivation, the dissolved oxygen tension was measured and used to adjust automatically the stirring or airflow rate to keep the dissolved oxygen tension value above 50%. After 1 h of cultivation, glucose feeding was started. To avoid the Crabtree effect [45,46], the concentration of sugars was kept below 0.1 $mg \cdot L^{-1}$. After the culture had reached an attenuance of 35 at $OD_{600\ nm}$, the glucose supply was stopped and induction of SMT expression was started by feeding galactose. Cells were harvested from cultures with an attenuance of 45 at $OD_{600\ nm}$ by centrifugation for 30 min at 8000 ***g*** and 4 °C. The cell pellet was shock-frozen in liquid nitrogen and stored at −80 °C.

Purification of SMT

Yeast cells collected from fermentation were resuspended in 70 mL of phosphate buffer (100 mM sodium phosphate (pH 6.0), 0.1% (v/v) Triton X-100, 1 mM EDTA and 1 mM dithiothreitol) and disrupted in a bead beater (Biospec Products, Bartelsville, OK, USA). To pellet the cell debris, the lysate was centrifuged at 10 000 ***g*** and 4 °C for 20 min. The supernatant was incubated with 0.05% (w/v) protamine sulfate under continuous stirring for 20 min at

room temperature followed by centrifugation for 20 min at 10 000 ***g*** and 4 °C and another incubation at 55 °C for 10 min. After centrifugation at 10 000 ***g*** and 4 °C for 20 min, the supernatant was brought to 1.3 M ammonium sulfate and applied to a Butyl Sepharose FF column (40 mL bead volume; GE Healthcare Bio-Sciences, Uppsala, Sweden). Linear gradient elution was applied using buffer A (20 mM sodium phosphate, 1.3 M ammonium sulfate, pH 6.0) and buffer B (20 mM sodium phosphate, pH 6.0). Fractions displaying SMT activities were pooled and the protein was precipitated by adding ammonium sulfate to 85% saturation under continuous stirring for 30 min on ice. The protein precipitate was pelleted by centrifugation for 20 min at 10 000 ***g*** and 4 °C. After resuspension in 20 mM sodium phosphate buffer, the protein was applied to a pre-equilibrated Superdex 200 26/60 size exclusion-column (GE Healthcare Bio-Sciences). Protein was eluted with 20 mM sodium phosphate buffer. The pooled fractions exhibiting SMT activities were incubated at 55 °C for 10 min and centrifuged (10 000 ***g*** for 20 min and 4 °C). The supernatant was loaded onto a Q-Sepharose (16/10) anion-exchange column (GE Healthcare Bio-Sciences). The protein was eluted by a linear gradient using buffer A (20 mM sodium phosphate buffer, pH 6.0) and buffer B (20 mM sodium phosphate pH 6.0, 0.5 M NaCl). The active fractions were pooled and dialyzed twice against 1 L of 100 mM MES buffer (pH 6.0) and then used for enzyme kinetic studies.

SMT activity assay

Enzyme reaction mixtures contained the substrates sinapoylglucose (0.1–1.0 mM) and L-malate (0.75–10 mM) in a total volume of 200 μL 0.1 M Mes buffer (pH 6.0) containing 5% (v/v) dimethylsulfoxide. Adding SMT protein started the reaction. Reaction mixtures were incubated at 30 °C. Samples of 50 μL were taken every 2 min and mixed with an equal volume of 100% (v/v) methanol to stop the reaction. Product formation was analyzed by HPLC as previously described [6]. SMT activity was calculated from the slope of a plot through the origin of the sinapoyl-L-malate product peak areas versus the reaction time at three selected time points (2, 4 and 6 min) for each donor and acceptor concentration. Data were evaluated using SIGMAPLOT and applying the corresponding enzyme kinetics tool (Systat Software, San Jose, CA, USA). Analysis of the SMT expression levels were performed as described previously [14]. SMT kinetic mechanism and the kinetic parameters were determined by plots according to Michaelis and Menten as well as Lineweaver and Burk [32]. Dissociation constants were estimated from Florini–Vestling plots (reciprocal intersections; see supplementary Fig. S4) [47]. Ternary complex dissociation constants were determined by calculating the reciprocal intersections of the reciprocal apparent maximal catalytic activity (V_{maxapp}^{-1}) versus reciprocal substrate concentration plots (see supplementary Fig. S5) [32].

Acknowledgements

We thank the Carlsberg research Center for the generous gift of the CPY p72UG plasmid, Andreas Gesell (University of Victoria, Canada) and Doreen Floß (Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany) for excellent assistance with the SMT expression in insect cells and tobacco leaves, respectively, as well as Narendar K. Khatri (University of Oulu, Finland) and Kathrin Schröder-Tittmann (Martin-Luther-University Halle-Wittenberg, Halle, Germany) for helpful advice in bioreactor SMT cultivation. We are especially grateful to Stephan König (Martin-Luther-University, Halle, Germany) for critical discussions on enzyme kinetics. This work was supported by the DFG priority program 1152 (Evolution of Metabolic Diversity).

References

1 Strack D & Mock H-P (1993) Hydroxycinnamic acids and lignins. In *Methods in Plant Biochemistry* Vol 9 (Dey PM & Harbonne JB, eds), pp. 45–97. Academic Press, New York, NY.

2 Kesy JM & Bandurski RS (1990) Partial purification and characterization of indol-3-ylacetylglucose:myo-inositol indol-3-ylacetyltransferase (indoleacetic acid-inositol synthase). *Plant Physiol* **94**, 1598–1604.

3 Kowalczyk S & Bandurski RS (1991) Enzymic synthesis of 1-*O*-(indol-3-ylacetyl)-beta-D-glucose. Purification of the enzyme from *Zea mays*, and preparation of antibodies to the enzyme. *Biochem J* **279**, 509–514.

4 Szerszen JB, Szczyglowski K & Bandurski RS (1994) iaglu, a gene from *Zea mays* involved in conjugation of growth hormone indole-3-acetic acid. *Science* **265**, 1699–1701.

5 Strack D (1982) Development of 1-*O*-sinapoyl-β-D-glucose: L-malate sinapoyltransferase activity in cotyledons of red radish (*Raphanus sativus* L. var. sativus). *Planta* **155**, 31–36.

6 Lehfeldt C, Shirley AM, Meyer K, Ruegger MO, Cusumano JC, Viitanen PV, Strack D & Chapple C (2000) Cloning of the *SNG1* gene of Arabidopsis reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. *Plant Cell* **12**, 1295–1306.

7 Milkowski C & Strack D (2004) Serine carboxypeptidase-like acyltransferases. *Phytochemistry* **65**, 517–524.

8 Li AX & Steffens JC (2000) An acyltransferase catalyzing the formation of diacylglucose is a serine

carboxypeptidase-like protein. *Proc Natl Acad Sci USA* **97**, 6902–6907.
9 Shirley AM, McMichael CM & Chapple C (2001) The *sng2* mutant of Arabidopsis is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose:choline sinapoyltransferase. *Plant J* **28**, 83–94.
10 Shirley AM & Chapple C (2003) Biochemical characterization of sinapoylglucose:choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism. *J Biol Chem* **278**, 19870–19877.
11 Vogt T, Aebershold R & Ellis B (1993) Purification and characterization of sinapine synthase from seeds of *Brassica napus*. *Arch Biochem Biophys* **300**, 622–628.
12 Milkowski C, Baumert A, Schmidt D, Nehlin L & Strack D (2004) Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. *Plant J* **38**, 80–92.
13 Weier D, Mittasch J, Strack D & Milkowski C (2008) The genes *BnSCT1* and *BnSCT2* from *Brassica napus* encoding the final enzyme of sinapine biosynthesis – molecular characterization and suppression. *Planta* **227**, 375–385.
14 Stehle F, Brandt W, Milkowski C & Strack D (2006) Structure determinants and substrate recognition of serine carboxypeptidase-like acyltransferases from plant secondary metabolism. *FEBS Lett* **580**, 6366–6374.
15 Milkowski C & Strack D (2004) Serine carboxypeptidase-like acyltransferases. *Phytochemistry* **65**, 517–524.
16 Hause B, Meyer K, Viitanen PV, Chapple C & Strack D (2002) Immunolocalization of 1-*O*-sinapoylglucose:malate sinapoyltransferase in *Arabidopsis thaliana*. *Planta* **215**, 26–32.
17 Kozak M (1983) Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiol Rev* **47**, 1–45.
18 Sharp PM & Li WH (1987) The Codon Adaptation Index – a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res* **15**, 1281–1295.
19 Bitter GA, Chen KK, Banks AR & Lai PH (1984) Secretion of foreign proteins from *Saccharomyces cerevisiae* directed by alpha-factor gene fusions. *Proc Natl Acad Sci USA* **81**, 5330–5334.
20 Clements JM, Catlin GH, Price MJ & Edwards RM (1991) Secretion of human epidermal growth factor from *Saccharomyces cerevisiae* using synthetic leader sequences. *Gene* **106**, 267–271.
21 Munro S & Pelham HR (1987) A C-terminal signal prevents secretion of luminal ER proteins. *Cell* **48**, 899–907.
22 Monnat J, Neuhaus EM, Pop MS, Ferrari DM, Kramer B & Soldati T (2000) Identification of a novel saturable endoplasmic reticulum localization mechanism mediated by the C-terminus of a Dictyostelium protein disulfide isomerase. *Mol Biol Cell* **11**, 3469–3484.
23 Graham TR & Krasnov VA (1995) Sorting of yeast alpha-1,3-mannosyltransferase is mediated by a lumenal domain interaction, and a transmembrane domain signal that can confer clathrin-dependent Golgi localization to a secreted protein. *Mol Biol Cell* **6**, 809–824.
24 Roberts CJ, Nothwehr SF & Stevens TH (1992) Membrane protein sorting in the yeast secretory pathway: evidence that the vacuole may be the default compartment. *J Cell Biol* **119**, 69–83.
25 Chang CN, Matteucci M, Perry LJ, Wulf JJ, Chen CY & Hitzeman RA (1986) *Saccharomyces cerevisiae* secretes and correctly processes human interferon hybrid proteins containing yeast invertase signal peptides. *Mol Cell Biol* **6**, 1812–1819.
26 Ammerer G, Hunter CP, Rothman JH, Saari GC, Valls LA & Stevens TH (1986) *PEP4* gene of *Saccharomyces cerevisiae* encodes proteinase A, a vacuolar enzyme required for processing of vacuolar precursors. *Mol Cell Biol* **6**, 2490–2499.
27 Jung G, Ueno H & Hayashi R (1999) Carboxypeptidase Y: structural basis for protein sorting and catalytic triad. *J Biochem (Tokyo)* **126**, 1–6.
28 Tague BW, Dickinson CD & Chrispeels MJ (1990) A short domain of the plant vacuolar protein phytohemagglutinin targets invertase to the yeast vacuole. *Plant Cell* **2**, 533–546.
29 Erhart E & Hollenberg CP (1983) The presence of a defective *LEU2* gene on 2μ DNA recombinant plasmids of *Saccharomyces cerevisiae* is responsible for curing and high copy number. *J Bacteriol* **156**, 625–635.
30 Nielsen TL, Holmberg S & Petersen JG (1990) Regulated overproduction and secretion of yeast carboxypeptidase Y. *Appl Microbiol Biotechnol* **33**, 307–312.
31 Gräwe W, Bachhuber P, Mock H-P & Strack D (1992) Purification and characterization of sinapoylglucose:malate sinapoyltransferase from *Raphanus sativus* L. *Planta* **187**, 236–241.
32 Segel IH (1975) *Enzyme Kinetics. Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*. John Wiley & Sons, Inc., Wiley-Interscience, Toronto.
33 Baumert A, Milkowski C, Schmidt J, Nimtz M, Wray V & Strack D (2005) Formation of a complex pattern of sinapate esters in *Brassica napus* seeds, catalyzed by enzymes of a serine carboxypeptidase-like acyltransferase family? *Phytochemistry* **66**, 1334–1345.
34 Dahlbender B & Strack D (1986) Purification and properties of 1-(hydroxycinnamoyl)-glucose:1-(hydroxycinnamoyl)-glucose hydroxycinnamoyltransferase from radish seedlings. *Phytochemistry* **25**, 1043–1046.

35 Strack D, Dahlbender B, Grotjahn L & Wray V (1984) 1,2-disinapolylglucose accumulated in cotyledons of dark-grown *Raphanus sativus* seedlings. *Phytochemistry* **23**, 657–659.
36 Rothman JH & Stevens TH (1986) Protein sorting in yeast: mutants defective in vacuole biogenesis mislocalize vacuolar proteins into the late secretory pathway. *Cell* **47**, 1041–1051.
37 Ng DT, Brown JD & Walter P (1996) Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J Cell Biol* **134**, 269–278.
38 Fleming SM, Robertson TA, Langley GJ & Bugg TD (2000) Catalytic mechanism of a C-C hydrolase enzyme: evidence for a gem-diol intermediate, not an acyl enzyme. *Biochemistry* **39**, 1522–1531.
39 Fraser CM, Thompson MG, Shirley AM, Ralph J, Schoenherr JA, Sinlapadech T, Hall MC & Chapple C (2007) Related Arabidopsis serine carboxypeptidase-like sinapoylglucose acyltransferases display distinct but overlapping substrate specificities. *Plant Physiol* **144**, 1986–1999.
40 Outchkourov NS, Peters J, de Jong J, Rademakers W & Jongsma MA (2003) The promoter-terminator of chrysanthemum rbs1 directs very high expression levels in plants. *Planta* **216**, 1003–1012.
41 van Engelen FA, Molthoff JW, Conner AJ, Nap JP, Pereira A & Stiekema WJ (1995) pBINPLUS: an improved plant transformation vector based on pBIN19. *Transgenic Res* **4**, 288–290.
42 McBride KE & Summerfelt KR (1990) Improved binary vectors for Agrobacterium-mediated plant transformation. *Plant Mol Biol* **14**, 269–276.
43 Kapila J, De Rycke R, Van Montagu M & Angenon G (1997) An Agrobacterium-mediated transient gene expression system for intact leaves. *Plant Science* **122**, 101–108.
44 Verduyn C, Postma E, Scheffers WA & Van Dijken JP (1992) Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* **8**, 501–517.
45 De Deken RH (1966) The Crabtree effect: a regulatory system in yeast. *J Gen Microbiol* **44**, 149–156.
46 Crabtree HG (1929) Observations on the carbohydrate metabolism of tumours. *Biochem J* **23**, 149–156.
47 Florini JR & Vestling CS (1957) Graphical determination of the dissociation constants for two-substrate enzyme systems. *Biochim Biophys Acta* **25**, 575–578.

Supplementary material

The following supplementary material is available online:

Fig. S1. Alignment of the Arabidosis SMT-cDNA (AtSMT) and the codon usage optimized cDNA (*ySMT*) for expression in *S. cerevisiae*.

Fig. S2. Map of plasmid pDIONYSOS used to express SMT in yeast.

Fig. S3. Indication of high gene dosage of *leu2-d* by increased plasmid copy number.

Fig. S4. Estimation of the dissociation constants.

Fig. S5. Estimation of the ternary complex dissociation constants.

Table S1. Oligonucleotides used for expression constructs and vector optimization.

This material is available as part of the online article from http://www.blackwell-synergy.com

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.