The Plant Journal (2001) **28**(1), 83–94

The *sng2* mutant of *Arabidopsis* is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose:choline sinapoyltransferase

Amber M. Shirley, Colleen M. McMichael and Clint Chapple*
Department of Biochemistry, Purdue University, West Lafayette, IN 47907, USA

Received 30 April 2001; revised 27 June 2001; accepted 27 June 2001
*For correspondence (fax: +1 765 496 7213; email: chapple@purdue.edu)

Summary

Serine carboxypeptidase-like (SCPL) proteins have traditionally been assigned roles in the hydrolytic processing of proteins; however, several SCPL proteins have recently been identified as catalysts in transacylation reactions of plant secondary metabolism. The novel functions of these enzymes suggest a catalytic diversity for plant SCPL proteins that extends beyond simple hydrolysis reactions. Characterization of the *Arabidopsis sng2* (*sinapoylglucose accumulator 2*) mutant has identified another SCPL protein involved in plant secondary metabolism. The *sng2* mutant was isolated by screening seed extracts for altered levels of sinapate esters, a group of phenylpropanoid compounds found in *Arabidopsis* and some other members of the Brassicaceae. Homozygous *sng2* seeds accumulate sinapoylglucose instead of sinapoylcholine, and have increased levels of choline and decreased activity of the enzyme sinapoylglucose:choline sinapoyltransferase (SCT). Cloning of the *SNG2* gene by a combination of map-based and candidate gene approaches demonstrates that SCT is another member of the growing class of SCPL acyltransferases involved in plant secondary metabolism.

Keywords: sinapoylcholine, sinapine, secondary metabolism, *Arabidopsis*, serine carboxypeptidase-like protein.

Introduction

Serine carboxypeptidases are often thought to play a role in the processing and degradation of proteins and peptides because they are capable of catalysing the cleavage of C-terminal peptide bonds. As a result, when proteins or cDNAs that exhibit homology to serine carboxypeptidases have been isolated from tissues actively engaged in protein turnover, it has often been assumed that these serine carboxypeptidase-like (SCPL) proteins are playing a role in proteolytic processes (Bamforth *et al.*, 1979; Bradley, 1992; Walker-Simmons and Ryan, 1980; Washio and Ishikawa, 1994). Similarly, the annotation of many SCPL genes throughout the *Arabidopsis* genome might be taken to imply that their encoded proteins are involved in protein degradation. The sequence similarity between these gene products and genuine serine carboxypeptidases has been shown to be potentially misleading by the recent identification of two SCPL acyltransferases involved in plant secondary metabolism (Lehfeldt *et al.*, 2000; Li and Steffens, 2000) and an SCPL hydroxynitrile lyase involved in cyanogenic glycoside degradation (Wajant *et al.*, 1994). These examples indicate that SCPL enzymes have a broader catalytic capability than is implied by their similarity to serine carboxypeptidases.

With the completion of the *Arabidopsis* Genome Initiative (AGI), 46 SCPL genes have been annotated in the *Arabidopsis* genome. The large number of genes in this family and the developing body of literature implicating SCPL proteins in non-traditional reactions suggests that at least some of these 46 *Arabidopsis* genes may encode proteins that function in reactions other than protein processing. One of these SCPL genes was recently shown to be required for the synthesis of sinapoylmalate, a UV-protective phenylpropanoid accumulated by *Arabidopsis* and some other members of the Brassicaceae. In contrast to the function suggested by its homology, the encoded enzyme, sinapoylglucose:malate sinapoyltransferase (SMT), catalyses a transesterification reaction (Lehfeldt *et al.*, 2000). A defect in

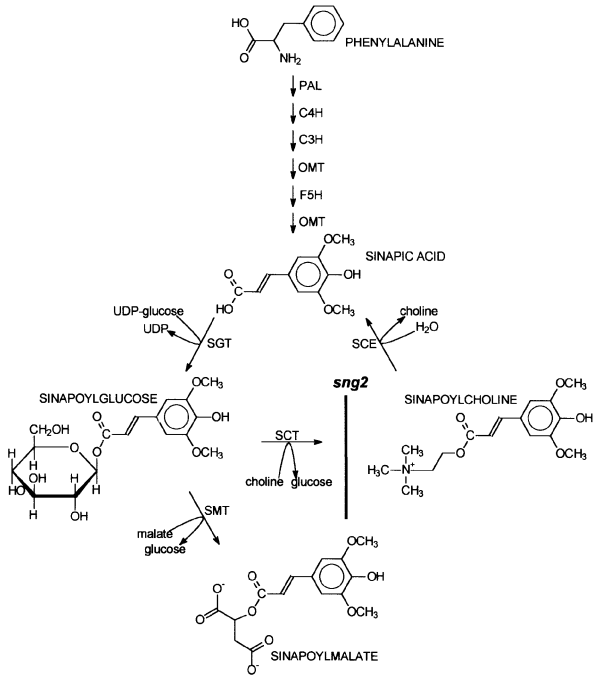

Figure 1. The pathway of sinapate ester biosynthesis.
The enzymes required for the conversion of phenylalanine to sinapic acid are phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), *p*-coumarate 3-hydroxylase (C3H), caffeic acid/5-hydroxyferulic acid *O*-methyltransferase (OMT) and ferulate 5-hydroxylase (F5H). The enzymes unique to sinapate ester biosynthesis are UDP-glucose:sinapic acid glucosyltransferase (SGT), sinapoylglucose:malate sinapoyltransferase (SMT), sinapoylglucose:choline sinapoyltransferase (SCT) and sinapoylcholinesterase (SCE). The biochemical block in the *sng2* mutant is indicated with a vertical line across the step catalysed by SCT.

the *SMT* gene causes the *sng1* mutant to accumulate sinapoylglucose in its leaves in place of sinapoylmalate (Lehfeldt *et al.*, 2000; Lorenzen *et al.*, 1996).

In addition to the sinapoylmalate found in leaves, *Arabidopsis* also accumulates a seed-specific sinapate ester, sinapoylcholine, also known as sinapine. We predicted that the final enzyme of the seed-specific pathway of sinapate ester metabolism, sinapoylglucose:choline sinapoyltransferase (SCT), might also be an SCPL protein because both SMT and SCT use sinapoylglucose as a substrate and catalyse similar transacylation reactions (Figure 1). In this paper, we describe the isolation and characterization of the *Arabidopsis sng2* mutant, which accumulates sinapoylglucose in its seeds in place of sinapoylcholine. By cloning the *SNG2* gene, we demonstrate that it encodes SCT and that SCT is an SCPL protein.

Results

Isolation of the sng2 *mutant*

To identify mutations affecting the accumulation of seed-specific sinapate esters, we screened methanolic seed extracts from 3000 M_3 families by TLC. One mutant was identified that showed greatly reduced levels of sinapoylcholine but accumulated high levels of a compound that co-chromatographed with sinapoylglucose (Figure 2a). In crosses to Columbia wild-type, the mutant phenotype segregated as a single, recessive Mendelian gene (Table 1). Quantitative HPLC analysis of Landsberg *erecta* wild-type and mutant seed extracts confirmed that mutant seeds accumulated reduced levels of sinapoylcholine and instead accumulated a compound that co-chromatographed with sinapoylglucose, in an amount almost equivalent to the amount of sinapoylcholine normally accumulated in the wild-type (Figure 2b; Table 2).

To identify the compound accumulated in the mutant, seed extracts were analysed by LC/MS. The peak that co-chromatographs with sinapoylglucose in the mutant gave the prominent parental (M-H)-ion at m/z 385 expected for sinapoylglucose as well as a fragment ion at m/z 205, consistent with the loss of glucose from the parent molecule. The TLC, HPLC and LC/MS results were consistent with the hypothesis that the mutant accumulates 1-*O*-sinapoyl-β-D-glucose. Based upon this phenotype, the mutant was designated *sng2* for *sinapoylglucose accumulator 2*.

Sinapate ester profiles of the wild-type and the sng2 *mutant*

The accumulation of sinapate esters in *Arabidopsis* seedlings and maturing embryos is developmentally regulated, both qualitatively and quantitatively (Lorenzen *et al.*, 1996; Ruegger *et al.*, 1999). The sinapate ester profiles of mutant and wild-type tissues were examined in order to determine whether the *sng2* mutation altered the accumulation of sinapate esters in tissues other than mature seeds. For these experiments, seedlings were grown on modified MS medium that contained nitrate as its sole nitrogen source because ammonia has been previously shown to inhibit sinapoylmalate accumulation and decrease SMT activity in radish seedlings (Dahlbender and Strack, 1984; Strack *et al.*, 1986). In Landsberg *erecta* wild-type seedlings, sinapoylcholine levels decreased during the first 3 days after imbibition and there was a transient rise in sinapoylglucose levels (Figure 3a). The decrease in sinapoylglucose levels on days 4–6 was accompanied by an increase in sinapoylmalate content. These results were consistent with previous analyses of the Columbia wild-type ecotype (Lorenzen *et al.*, 1996). In the *sng2* mutant, the total level of sinapate esters in seeds is nearly equivalent to that of wild-type (Table 2; Figure 3b,d). As mutant seedlings developed, sinapoylglucose levels remained elevated until after day 3 when, as in wild-type, sinapoylmalate levels increased (Figure 3b). These data indicate that the *sng2* mutation, and the absence of sinapoylcholine in the *sng2*

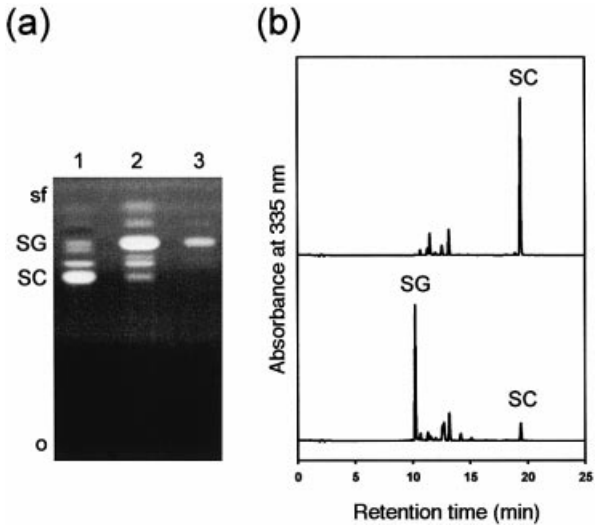

Figure 2. Analysis of sinapate ester content in wild-type and *sng2* seeds. (a) Seed extracts were prepared from Landsberg *erecta* wild-type (lane 1) and *sng2* (lane 2). A leaf extract from the sinapoylglucose-accumulating *sng1-1* mutant (lane 3) was included as a sinapoylglucose standard. Extracts were analysed on silica gel plates with the mobile phase *n*-butanol:acetic acid:water (5:2:3 v/v/v). Sinapoylcholine (SC) and sinapoylglucose (SG) were visualized under 312 nm UV light. o, origin; sf, solvent front. (b) Seed extracts from Landsberg *erecta* wild-type (top) and *sng2* (bottom) were analysed by HPLC with UV detection at 335 nm.

mutant, do not have pleiotropic effects on the developmental regulation of sinapate ester biosynthesis and turnover in seedlings.

In order to examine the impact of the *sng2* mutation on sinapate ester accumulation in maturing embryos, the sinapate ester content of siliques developing on individual inflorescences was quantified; siliques at the bottom of the rachis contain the most mature embryos, whereas siliques at the top of the inflorescence represent the most immature embryos. In the wild-type, sinapoylglucose levels increase during early maturation. As siliques mature, a decrease in sinapoylglucose levels is accompanied by an increase in sinapoylcholine content (Figure 3c). In contrast, in the *sng2* mutant, sinapoylcholine levels remain very low throughout embryonic development and sinapoylglucose content increases to near the same level as the sinapoylcholine found in the wild-type (Figure 3d). These results indicate that the *sng2* mutant fails to accumulate sinapoylcholine to a significant degree at any developmental time point examined, and that, in the absence of sinapoylcholine, total levels of sinapate esters in the mutant remain essentially unchanged.

Biochemical profiling of the wild-type and sng2 *mutant*

In an attempt to identify the underlying reason why the *sng2* mutant fails to accumulate sinapoylcholine, the free choline levels of the mutant and the wild-type were quantified by plasma desorption mass spectrometry. As choline is a substrate for SCT, the lack of sinapoylcholine accumulation in the mutant could be explained by a defect in choline biosynthesis; however, the free choline content of the mutant was twice that of wild-type, indicating that a lack of choline is not responsible for the *sng2* phenotype (Table 2). Interestingly, the excess of free choline in the mutant was equal to the amount of sinapoylcholine normally present in the wild-type. The presence of both sinapoylglucose and choline in *sng2* seeds suggests that *sng2* seeds should be able to synthesize sinapoylcholine unless they lack a functional SCT enzyme; therefore, the ability of wild-type and mutant enzyme extracts to synthesize sinapoylcholine was quantified. The results of *in vitro* enzyme assays from crude seed extracts revealed that the specific activity of SCT in the mutant was only 1% that of the wild-type (Table 2). The data from these biochemical analyses suggest that the *SNG2* gene is required for SCT expression or activity.

Table 1 Segregation of the *sng2* allele in an F_2 population derived from a *sng2* × Columbia cross

Genotype	Observed	Expected[a]
SNG2/SNG2	52	57
SNG2/sng2	105	114
sng2/sng2	71	57

The *sng2* phenotype was scored in F_3 progeny by dissecting individual embryos from the siliques developing on F_2 plants and evaluating their sinapate ester content by TLC.
[a]Values expected for inheritance of a recessive allele of a Mendelian gene ($\chi^2 = 0.7$, $P > 0.4$).

SNG2 *encodes a serine carboxypeptidase-like protein*

The cloning of the *SNG1* gene demonstrated that it encodes SMT, and revealed a role for an SCPL protein in sinapate ester metabolism (Lehfeldt *et al.*, 2000). Because the reactions catalysed by SMT and SCT are similar transacylation reactions that use sinapoylglucose as an activated sinapate donor, we predicted that SCT might also be an SCPL acyltransferase. This hypothesis is supported by experiments in which *Arabidopsis* seed extracts were treated with PMSF, an inhibitor of enzymes with active site serine residues, including serine carboxypeptidases. PMSF treatment inhibited the activity of the *Arabidopsis* SCT by 25%, similar to the results previously obtained with the *Brassica napus* SCT enzyme (Vogt *et al.*, 1993).

With the expectation that the *SNG2* locus was likely to encode SCT, we attempted to isolate the *SNG2* gene using a combination of map-based cloning and candidate gene approaches. In an initial survey of an *sng2* mapping

Table 2 Biochemical characterization of wild-type and *sng2* seeds

Line	Sinapoylglucose[a]	Sinapoylcholine[a]	Choline[a]	SCT activity[b]
SNG2/SNG2	ND	33.22 ± 1.00	34.06 ± 2.39	30.61 ± 0.38
sng2/sng2	19.30 ± 3.97	5.16 ± 0.08	60.18 ± 8.56	0.30 ± 0.01

[a]nmol mg^{-1} ± SE
[b]pkat mg^{-1} ± SE
ND, not detectable

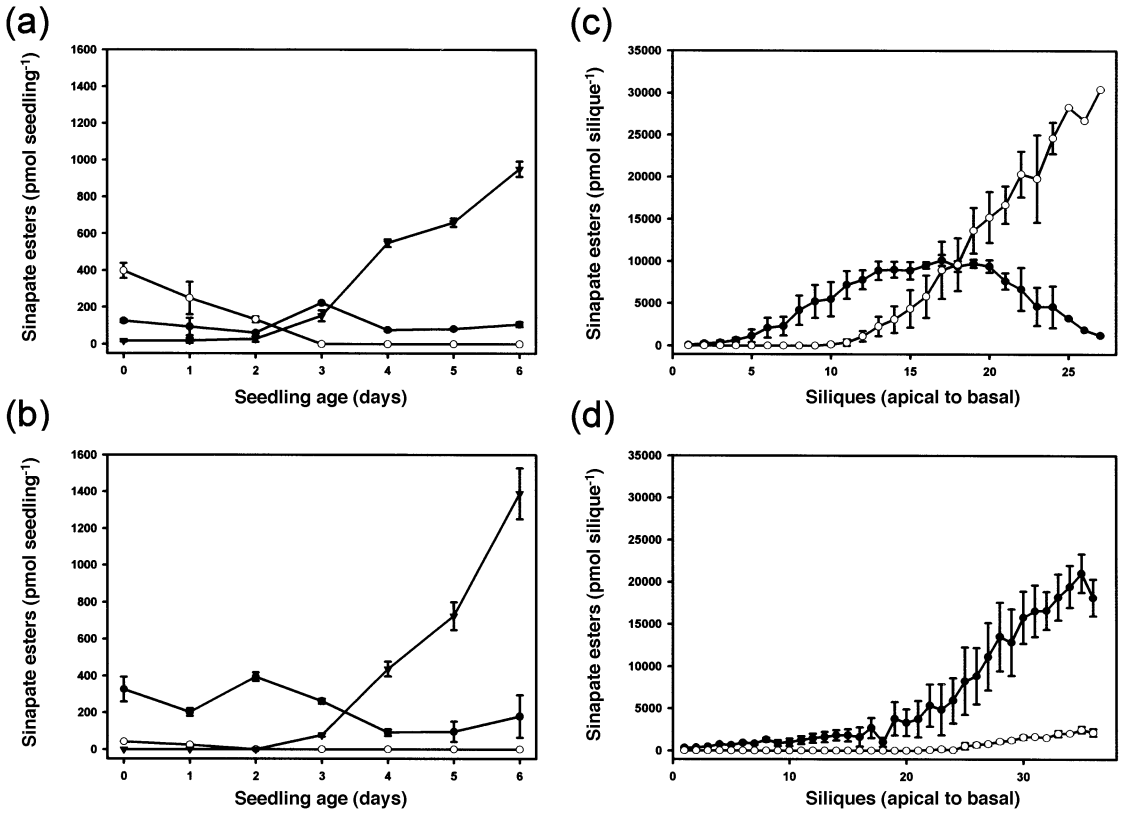

Figure 3. Developmental changes in sinapate ester content in wild-type and *sng2* seedlings and embryos.
Sinapate esters from Landsberg *erecta* wild-type (a) and *sng2* (b) seedlings grown on modified Murashige and Skoog agar plates (Lorenzen *et al.*, 1996) and Landsberg *erecta* wild-type (c) and *sng2* (d) siliques developing on individual inflorescences were analysed by HPLC with UV detection at 335 nm and quantified using the extinction coefficient of sinapic acid. Siliques at the bottom of the inflorescence represent the oldest developmental time point, whereas siliques at the top of the inflorescence contain the most immature embryos. Each point represents the average of three seedlings or siliques ± SE. (▼) sinapoylmalate; (●) sinapoylglucose; (○) sinapoylcholine.

population, the *SNG2* gene was mapped near the RFLP marker m217 (one recombinant chromosome out of 38 total chromosomes). In an analysis of a larger mapping population, the *SNG2* locus demonstrated linkage to the SSLP marker nga151 (18 recombinant chromosomes out of 388 total chromosomes) and tight linkage to the SSLP marker nga249 (no recombinant chromosomes out of 384 total chromosomes) (Figure 4).

Before the sequencing in this region of chromosome V was complete, a candidate SCPL gene (*T22D6.200*) was identified on two overlapping bacterial artificial chromosomes (BACs), F24L14 and F27P17 (Figure 4), by probing a BAC library of Col-0 genomic DNA (Mozo *et al.*, 1998) with cDNA from eight SCPL ESTs with no known genomic position. Sequencing of this gene from the *sng2* mutant did not reveal an EMS-induced mutation, and transformation of *sng2* plants with a binary vector carrying the wild-type allele did not result in complementation of the mutant phenotype (data not shown). After the sequence for this area of chromosome V was released, BAC F17I14, which carried a gene annotated as an SCPL, was identified near BACs

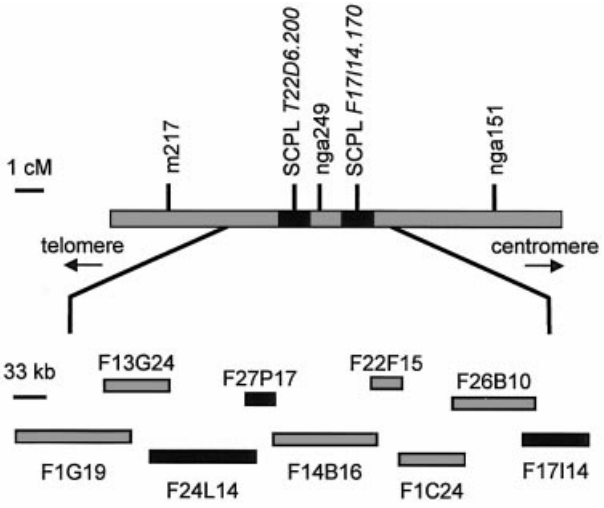

Figure 4. Mapping of the *SNG2* locus.
The *SNG2* locus was initially mapped near the RFLP marker m217 (one recombinant chromosome out of 38 total chromosomes). Further mapping was carried out with the SSLP markers nga249 (no recombinant chromosomes out of 384 total chromosomes) and nga151 (18 recombinant chromosomes out of 388 total chromosomes). The region flanking nga249 is covered by the BAC contig illustrated at the bottom of the figure. Two SCPL genes, *T22D6.200* and *F17I14.170*, near the nga249 marker were identified, and the latter has been shown to correspond to the *SNG2* locus.

F24L14 and F27P17 (Figure 4). BAC F17I14 had not hybridized in the BAC library screening described above.

Several approaches were then taken to determine whether the identified candidate gene corresponded to *SNG2*. First, RNA gel blot hybridization analysis was used to quantify *F17I14.170* transcript from various tissues, including young leaves, mature leaves, senescent leaves, 10-day-old seedlings, stem tissue, siliques, flowers and roots. This experiment indicated that *F17I14.170* is expressed highly only in siliques (Figure 5), consistent with the pattern expected for a gene encoding SCT. Expression of *F17I14.170* in the *sng2* mutant was also examined by quantifying the abundance of the transcript by RNA gel blot hybridization of total silique RNA from both Landsberg *erecta* wild-type and the *sng2* mutant. Using cyclophilin as an internal standard for loading, expression in the mutant was approximately 15% of wild-type (data not shown). As EMS-induced mutations often lead to mRNA destabilization, the failure of *F17I14.170* transcript to significantly accumulate in the mutant also suggests that *F17I14.170* represents *SNG2*. Finally, the putative *sng2* genomic allele was amplified by PCR and sequenced. A G → A nonsense mutation corresponding to position 1197 of the cDNA was identified that alters Trp399 of the inferred protein to a stop codon, truncating the last 66 amino acids of the protein (Figure 6). It should be noted that sequencing of the RT–PCR product obtained for the putative *SNG2* allele revealed that several portions of

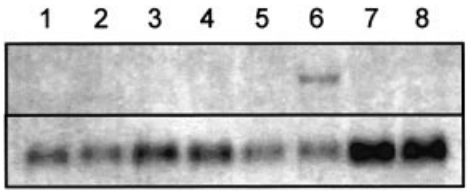

Figure 5. Expression analysis of the *SNG2* gene.
RNA was prepared from various tissues and probed with the putative SNG2 cDNA (expressed sequence tag 309H12T7; Genbank accession number AA394342) (top). Expression of cyclophilin (bottom) was used as a loading control. Lane 1, young leaves; lane 2, mature leaves; lane 3, senescent leaves; lane 4, 10-day-old seedlings; lane 5, stems; lane 6, siliques; lane 7, flowers; lane 8, roots.

genomic sequence that had been annotated as introns by the AGI are actually exons (Figure 6). Inclusion of these nucleotide sequences indicated that F17I14.170 shares all three conserved serine, aspartic acid and histidine residues (S178, D388 and H442 in the F17I14.170 sequence) found in carboxypeptidase Y. The amino acids missing from the inferred mutant protein include His442, which aligns with His398 in carboxypeptidase Y, one of the three amino acids that comprise the active site catalytic triad in carboxypeptidase Y. This mutation also introduces an *Sau*3AI polymorphism in the putative *sng2* allele, which was used to generate a cleavable amplified polymorphic sequence marker. DNA amplified from the wild-type was cleaved by *Sau*3AI, whereas it was not cleaved when mutant DNA was used as the template for PCR (data not shown). This finding unambiguously demonstrates that the mutation identified by sequencing was not introduced by the PCR process.

The final evidence that *F17I14.170* corresponds to the *SNG2* locus was provided by complementation of the *sng2* mutant phenotype. A genomic clone of *F17I14.170* was isolated from BAC F17I14, and was used to generate a binary vector construct with 1.1 kb of promoter sequence, 2.6 kb of genomic sequence, and 0.3 kb of 3′ sequence. The resulting pBI101-*SNG2* vector was transformed into *sng2* plants by the floral dip method (Clough and Bent, 1998). When sinapate ester levels of T_2 embryos from kanamycin-resistant T_1 plants were analysed by TLC and HPLC, sinapoylcholine biosynthesis was not restored in transformed lines carrying the pBI101 empty vector; however, 20 independent lines transformed with the pBI101-*SNG2* construct contained wild-type levels of sinapoylcholine (Figure 7).

The SNG2 *locus encodes SCT*

The evidence described above indicates that *SNG2* encodes an SCPL protein, which is likely to be SCT. Several years ago, sequences from *Brassica* SCT tryptic peptides were published (Vogt *et al.*, 1993), but to date have not been used to clone the *SCT* gene. The *F17I14.170*

cDNA sequence contained amino acid sequences corresponding to all four of the tryptic peptide fragments, with two amino acid insertions in the *B. napus* sequence and two amino acid substitutions (Figure 6), strongly suggesting that the *SNG2* gene encodes SCT. Nevertheless, to conclusively demonstrate that the *SNG2* locus encodes SCT, the function of the SNG2 protein was determined following expression of the *SNG2* cDNA in *Escherichia coli*. As SMT carries an N-terminal pro-peptide (Lehfeldt *et al.*, 2000), the algorithm described by Nielsen *et al.* (1997) and available at the SignalP website (http://www.cbs.dtu.dk/services/SignalP/) was used to predict which amino acids might comprise a similar pro-peptide in SCT (Figure 6). The portion of the *SNG2* cDNA corresponding to the inferred mature protein was then expressed under the control of the T7 promoter in pET28A. SNG2 protein did not significantly accumulate in BL21 (DE3) host cells (data not shown). Visible accumulation of protein required the use of BL21-CodonPlus™ (DE3)-RIL host cells that compensate for rare arginine, isoleucine and leucine codons. Using this host, a novel protein with a molecular mass of approximately 45 kDa, in keeping with the mass of 48 kDa expected for the modified SNG2 polypeptide, was

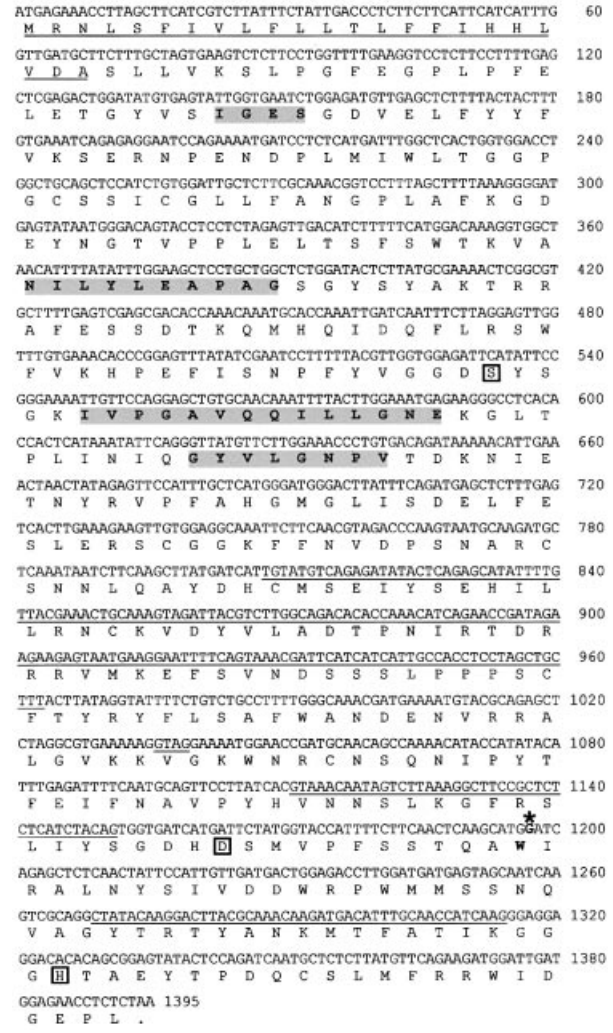

Figure 6. The *SNG2* cDNA and its deduced amino acid sequence.
The nucleotide sequences not found in the cDNA annotation at GenBank for *F17I14.170* are underlined. The nucleotide altered in the *sng2* mutant is denoted by an asterisk. Mutation of this site from guanine to adenine in the *sng2* mutant alters the trytophan residue in bold to a stop codon. The predicted N-terminal amino acid signal sequence is double underlined. Shaded amino acid sequences correspond to tryptic peptide fragments obtained from SCT purified from *Brassica napus* (Vogt *et al.*, 1993). These residues correspond to the following *B. napus* fragments: IDGES, NILYLEDPAG, IVPGAAQQILLGNRK and GYVLGNPAV. The boxed amino acids align with the conserved catalytic triad of carboxypeptidase Y. The GenBank accession number for the *SNG2* cDNA is AY033947.

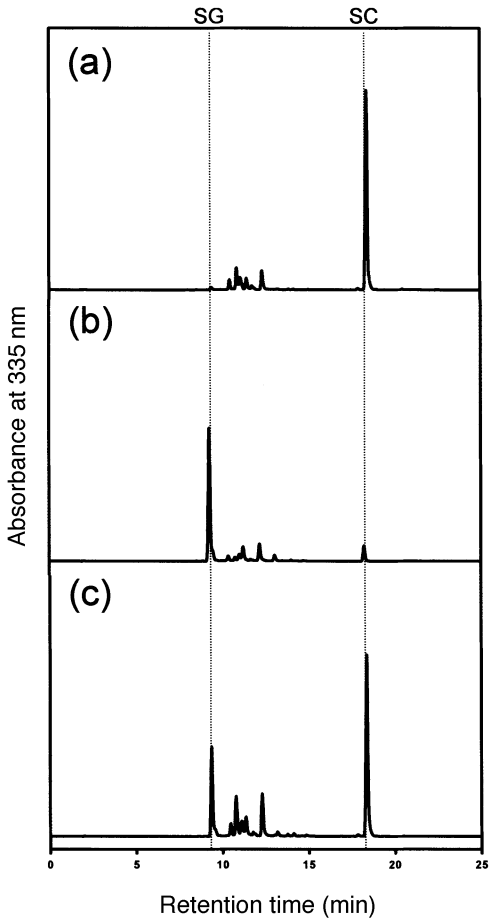

Figure 7. Complementation of the *sng2* mutant.
Seed extracts were prepared from Landsberg *erecta* wild-type (a), *sng2* transformants carrying the pBI101-empty vector (b), and *sng2* transformants carrying the pBI101-*SNG2* transgene (c) and were analysed by HPLC with UV detection at 335 nm. SG, sinapoylglucose; SC, sinapoylcholine. To test for complementation, T_2 families were identified that segregated 3:1 for pBI101-*SNG2* or the pBI101 transgene. Seeds from these populations were then analysed for their sinapate ester content. The presence of sinapoylglucose in the pBI101-*SNG2* extracts is explained by the percentage of *sng2* seeds within the population that have lost the transgene through segregation.

visible in the insoluble fraction of cells containing the pET28A-*SNG2* vector induced at 37°C. The band was absent in all fractions from induced cells containing the pET28A empty vector. SNG2 protein was not present in the soluble fraction, suggesting that most of the protein was present in inclusion bodies (Figure 8). Unfortunately, when the soluble fraction of these cell extracts was analysed for SCT activity by HPLC, no enzymatic synthesis of sinapoylcholine was observed, probably because little if any protein had folded into its native state under these induction conditions. As induction at lower temperature and the use of lower concentrations of isopropyl-β-D-thiogalactopyranoside also failed to lead to the production of soluble active enzyme, inclusion bodies were isolated, and the recombinant protein was denatured, refolded, and partially purified by anion exchange chromatography. When the renatured protein was used for SCT assays, HPLC analysis revealed that a compound that co-chromatographs with sinapoylcholine and has a spectrum consistent with known standards of sinapoylcholine was formed in the presence of sinapoylglucose and choline when incubated at 30°C (Table 3). In assays lacking enzyme, trace amounts of sinapoylcholine were identified by HPLC and matrix-assisted laser desorption ionization (MALDI)/MS, presumably reflecting spontaneous transacylation. Although the amount of sinapoylcholine produced in SCT assays was low, probably due to inefficient protein refolding, these data provide conclusive proof that the *SNG2* gene encodes SCT.

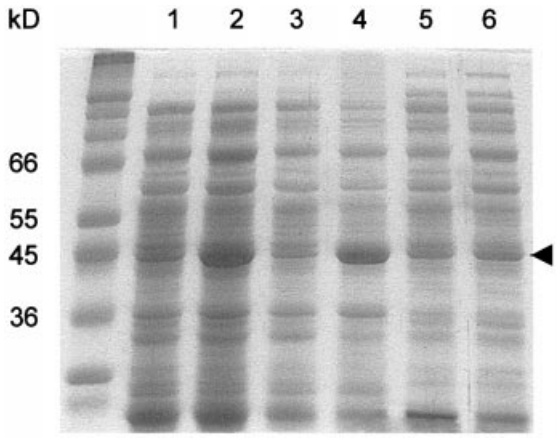

Figure 8. Analysis of *SNG2* expressed in *E. coli.*
SDS–PAGE analysis of *E. coli* harbouring pET28a (lanes 1, 3 and 5) and the *SNG2* expression vector pET28a-*SNG2* (lanes 2, 4 and 6) for total protein (lanes 1 and 2), insoluble fractions (lanes 3 and 4) and soluble fractions (lanes 5 and 6).

Discussion

The isolation of the *SNG2* gene demonstrates that SCT is another example from an emerging class of acyltransferases that are homologous to serine carboxypeptidases. Only two other examples of SCPL proteins that catalyse transacylation reactions *in vivo* are known. One of these is SMT, the enzyme responsible for the final step in sinapoylmalate biosynthesis (Lehfeldt *et al.*, 2000). The other is an isobutyryl acyltransferase involved in the synthesis of 2,3,4-tri-*O*-isobutyrylglucose in wild tomato trichomes (Li and Steffens, 2000). Both of these enzymes are similar to SCT in that they use 1-*O*-acylglucosides as activated acyl donors in the transacylation reactions that they catalyse. Before the genes encoding these enzymes were isolated, it had not been anticipated that these acyltransferases would be similar to serine carboxypeptidases as carboxypeptidases catalyse distinctly different reactions in which they act as hydrolases that use proteins and peptides as substrates. Interestingly, carboxypeptidase Y has been shown to catalyse transacylation reactions with amino acids or amino acid derivatives as nucleophiles *in vitro*; however, it can do so only under alkaline conditions that are not physiologically relevant (Breddam *et al.*, 1980; Widmer *et al.*, 1980). Thus, although SCT, SMT and the *Lycopersicon pennellii* isobutyryl acyltransferase share significant amino acid identity with

Table 3 Analysis of SCT activity in *E. coli* expressing the *SNG2* gene

	Substrates added		
Source of protein	Sinapoylglucose	Choline	Sinapoylcholine formed[a]
pET28a-*SNG2*	+	+	560 ± 45
pET28a-empty vector	+	+	73 ± 5
Boiled pET28a-*SNG2*	+	+	74 ± 7
pET28a-*SNG2*	+	–	0 ± 0
pET28a-*SNG2*	–	+	0 ± 0
pET28a-*SNG2*	–	–	0 ± 0
Buffer alone	+	+	55 ± 6

[a]pmol ± SE (n = 3)

carboxypeptidase Y, including at least three inferred active site residues, the structures of the plant proteins must differ from classical carboxypeptidases in such a way as to facilitate the binding of their very different substrates and to favour the transacylation that these enzymes catalyse.

A common feature of serine carboxypeptidases as well as other members of the α/β hydrolase superfamily of proteins is the presence of conserved serine, histidine and aspartic acid active site residues that make up the so-called catalytic triad (Bech and Breddam, 1989; Hayashi *et al.*, 1973, 1975; Liao and Remington, 1990; Liao *et al.*, 1992; Ollis *et al.*, 1992). Presumably, SCT also utilizes the three conserved residues for catalysis, perhaps by nucleophilic attack of S178 on the ester bond of sinapoylglucose, thereby generating a sinapoylated enzyme intermediate. This mechanism would be consistent with previous studies on the reaction kinetics of SCT isolated from *B. napus* which indicated that SCT has a double displacement mechanism of catalysis (Vogt *et al.*, 1993), typical of serine carboxypeptidases. Despite these similarities, the reaction catalysed by SCT is distinct from that catalysed by carboxypeptidase Y in that carboxypeptidase Y uses water as a nucleophile to cleave the acyl intermediate formed in the first phase of the reaction that it catalyses. In contrast, SCT would have to exclude water from its active site in order to use the hydroxyl of choline as a nucleophile. Without the exclusion of water, SCT would be expected to function as a sinapoylglucose esterase instead of an acyltransferase. Thus, if the catalytic mechanism of SCT is similar to that of carboxypeptidase Y, SCT must employ specific residues or perhaps protein conformations that prevent water from rapidly hydrolysing the sinapoylated enzyme intermediate. Alternatively, S178 may act as a base during catalysis, deprotonating the hydroxyl of choline, thereby facilitating its direct nucleophilic attack on the ester linkage of sinapoylglucose, and obviating the need for exclusion of water from the enzyme's active site. Such a mechanism has recently been postulated for MhpC, a C–C hydrolase from *E. coli* which, like SCT, SMT and carboxypeptidase Y, employs a catalytic triad (Fleming *et al.*, 2000).

Even with the availability of a completely sequenced genome, without the insights provided by the *sng1* and *sng2* mutant phenotypes, the functions of the genes encoding SMT and SCT would not have been readily apparent. An understanding of the mechanism that this novel group of proteins uses to carry out transacylation will be fundamental to identifying additional SCPL proteins that function as acyltransferases. In the *Arabidopsis* genome, there are presently 46 genes annotated as encoding SCPL proteins. Preliminary phylogenetic analysis of their sequences does not group SCT, SMT and the *Lycopersicon pennellii* isobutyryl acyltransferase into a distinct clade of proteins that might include other putative acyltransferases. Thus, the features of SCPL acyltransferases that distinguish them from traditional serine carboxypeptidases may consist of only a few key residues, or may be elements of secondary or tertiary structure that are not revealed by analysis of their primary sequence.

In addition to providing insight into the role of SCPL proteins in plant metabolism, the *sng2* mutant also provides an opportunity to study the function of sinapoylcholine accumulation in seeds of the Brassicaceae. Radiotracer feeding experiments with ^{14}C-choline and ^{14}C-ethanolamine indicated that choline released from the hydrolysis of sinapoylcholine in germinating seeds is incorporated into phosphatidylcholine (Strack, 1981). Despite the clear importance of membrane lipid synthesis in developing seedlings, greatly reduced levels of sinapoylcholine in *sng2* seeds appear to have no effect on germination or seedling growth, and the accumulation of free choline in the mutant indicates that sinapoylcholine is not an obligatory storage form for choline in *Arabidopsis*. These findings are consistent with previous studies on the *fah1* mutant which completely lacks sinapate esters at all developmental stages, and also accumulates free choline in its seeds instead of sinapoylcholine (Chapple *et al.*, 1992). Interestingly, some residual sinapoylcholine is found in *sng2* seeds, despite the fact that the *sng2-1* allele is probably null because it lacks the conserved, catalytic histidine residue. This finding may indicate that there are other acyltransferases that are partially redundant with SNG2. On the other hand, in *in vitro* reactions, we consistently observed the formation of low levels of sinapoylcholine from sinapoylglucose and choline in the absence of enzyme. These data suggest that spontaneous non-enzymatic synthesis may account for the low levels of sinapoylcholine found in *sng2* seeds.

The absence of obvious phenotypes in mutants lacking sinapoylcholine is also a finding of potential agronomic importance. Oilseed rape or canola (*Brassica* sp.) accumulates sinapoylcholine in seeds, and when post-crushing canola meal is used in poultry feed, by-products of sinapoylcholine degradation impart a fishy taint to eggs (Hobson-Frohock *et al.*, 1977). The examination of breeding lines of *B. napus* and *B. campestris* for genetic variation in seed sinapoylcholine accumulation has not identified significant variation for the trait (Vogt *et al.*, 1993). The isolation of the sinapoylcholine-deficient *sng2* mutant strongly suggests that it should be possible to manipulate sinapoylcholine levels in *Brassica* crops, and the cloning of the *SNG2* gene provides the tools necessary for a genetic engin-

eering approach to this problem using antisense or co-suppression technology.

Experimental procedures

Plant material

Arabidopsis thaliana L. Heynh. ecotypes Landsberg *erecta* or Columbia were cultivated at a light intensity of 100 μE m^{-2} sec^{-1} at 23°C under a photoperiod of 16 h light/8 h dark in Redi-Earth potting mix (Scotts-Sierra Horticultural Products, Marysville, OH, USA). For seedling material to be used in the analysis of sinapate esters, seeds were surface-sterilized for 10 min in a 2:1 v/v mixture of 0.1% Triton-X 100 and household bleach. Seeds were rinsed thoroughly with sterile water and plated onto modified MS medium (Murashige and Skoog, 1962) (ammonia-free medium to which an additional 20.6 mM potassium nitrate was added in place of ammonium nitrate) containing 0.7% agar.

Genetic methods

For mutant screening, seed (5–10 mg) from M_3 families of EMS-mutagenized Landsberg *erecta* populations were extracted in 200 μl of 50% methanol. An 5 μl aliquot of the extract was analysed by silica gel TLC using a mobile phase of *n*-butanol:acetic acid:water (5:2:3, v/v/v). After development of the chromatograms, sinapate esters were visualized under UV light. Putative mutants were re-tested by the same procedure in the next generation. Homozygous *sng2* progeny were back-crossed to wild-type four times to remove unlinked background mutations.

Sinapate ester analysis

For analysis of seed sinapate ester content, Landsberg *erecta* and *sng2* seeds (10 mg) were extracted in 1.5 ml 50% methanol containing 1.5% v/v acetic acid. For analysis of sinapate ester content throughout silique development, single siliques from 5-week-old plants were extracted beginning with the bottom silique that had just started to turn brown and ending with the first expanding silique. Each silique was extracted in 50 μl 50% methanol containing 1.5% v/v acetic acid. For analysis of germinating seedlings, single seedlings were extracted in 50 μl 50% methanol containing 1.5% v/v acetic acid. In each case, a 20 μl aliquot was analysed by HPLC on a Puresil™ C18 column (Waters) (1200 nm pore size, 5 μm particle size) using a 23 min gradient from 1.5% acetic acid, 0.05% SDS to 30% acetonitrile in 1.5% acetic acid, 0.05% SDS at a flow rate of 1 ml min^{-1} using UV detection at 335 nm. Sinapate esters were quantified using the extinction coefficient of sinapic acid.

Liquid chromatography/mass spectrometry (LC/MS)

The LC/MS analysis was carried out using a Finnigan MAT LCQ mass spectrometer (Thermoquest Corp. San Jose, CA, USA). A 20 μl aliquot of the sample was injected onto a C18 column (Waters) (1200 nm pore size, 5 μm particle size) (flow rate 0.5 ml min^{-1}) using a 1.5% acetic acid/acetonitrile gradient (starting with 95% of 1.5% acetic acid and ramping to 95% of acetonitrile over 32 min). The sample was ionized using negative ion electrospray ionization (ESI) with the LCQ being scanned to 2000 amu. The source voltage was set at 4 kV and the capillary voltage at 20 V. The background source pressure was 1.5×10^{-5} torr as read by an ion gauge. The drying gas used in this study was nitrogen.

Matrix-assisted laser desorption ionization/mass spectrometry (MALDI/MS)

The MALDI/MS results were obtained using a Voyager mass spectrometer (PerSeptive Biosystems, Framingham, MA, USA). This instrument utilizes a nitrogen laser (337 nm UV laser) for ionization with a time-of-flight mass analyser. The sample and matrix were mixed in a ratio of 1 μl to 1 μl on the sample plate. This mixture was allowed to air dry prior to analysis. The matrix used for this sample was α-cyano-4-hydroxy cinnamic acid.

Choline determination

Choline was extracted from Landsberg *erecta* wild-type and *sng2* seeds and analysed by plasma desorption/MS as described previously (Yang *et al.*, 1995).

Enzyme assays

Enzyme extraction and assay conditions were based upon those used for purification and assay of SCT from *B. napus* (Vogt *et al.*, 1993). For assay from crude seed extracts, Landsberg *erecta* or *sng2* seeds were frozen in liquid nitrogen and ground to a fine powder. The powder was stirred for 20 min in 5 vol 100 mM potassium phosphate buffer (pH 6.8) containing 20 mM sodium chloride and 4% w/v insoluble polyvinylpolypyrrolidone. The samples were filtered through Miracloth (Calbiochem, La Jolla, CA, USA) and centrifuged for 20 min at 13 000 ***g***. The supernatant was made to 0.1% w/v protamine sulphate, stirred for 20 min, and centrifuged for 20 min at 13 000 *g*. The supernatant was again filtered through Miracloth, and the protein was precipitated by adding ammonium sulphate to 85% saturation, followed by centrifugation for 20 min at 13 000 *g*. The pellet was resuspended in 100 mM potassium phosphate buffer (pH 7.0) containing 50 mM sodium chloride, desalted using PD-10 Sephadex® G-25M Columns (Supelco, Bellefonte, PA, USA) into 100 mM potassium phosphate buffer (pH 7.0) and used for determination of SCT activity. Each assay contained 40 μl 2.5 mM sinapoylglucose, 8 μl 100 mM choline chloride, and 32 μl protein extract. The assays were incubated for 60 min at 30°C, stopped by the addition of 420 μl cold 50% methanol, and analysed by HPLC. Sinapoylglucose for use in enzyme assays was purified from the *sng1* mutant of *Arabidopsis* (Lorenzen *et al.*, 1996). Protein content was determined by the bicinchonic acid assay procedure (Pierce) with BSA as a standard. For PMSF inhibition assays, 1 μl 0.1 M PMSF in ethanol was added to protein extracts and the reaction was pre-incubated for 30 min at 30°C before addition of substrates.

For assay of SCT activity from heterologously expressed, denatured, and refolded *SNG2* in *E. coli*, each assay contained 75 μl *E. coli* extract, 5 μl 100 mM choline chloride and 20 μl 2.5 mM sinapoylglucose. The assays were incubated at 30°C for 4 h, stopped by the addition of 75 μl 100% methanol, and analysed by HPLC.

Analysis of nucleic acids

For DNA gel blot analysis, BAC DNA, isolated using a Plasmid Midi Kit for Purification of BAC DNA (Qiagen, Santa Clarita, CA, USA), and DNA extracted from plant material (Doyle and Doyle,

1990) were digested with restriction endonucleases, electrophoretically separated, transferred to Hybond N+ membrane (Amersham) and hybridized according to standard protocols (Sambrook *et al.*, 1989) with DNA probes made using the DECAprime II system (Ambion, Austin, TX, USA). RNA was extracted from tissues (Goldsbrough and Cullis, 1981), electrophoretically separated, transferred to Hybond N+ membrane, and hybridized with radiolabelled probes prepared as described above from expressed sequence tag 309H12T7 (Genbank accession number AA394342). Hybridization to target RNA was quantified using a Typhoon 8600 Variable Mode Imager (Molecular Dynamics, Sunnyvale, CA, USA). Sequencing was performed using a thermosequenase flourescent-labelled primer cycle sequencing kit (Amersham, Piscataway, NJ, USA) with the IRD-800 M13forward and IRD-700 M13 reverse primers (LICOR, Lincoln, NB, USA). The reaction products were analysed with the LongReadIR DNA 4200 automated DNA sequencer (LICOR).

Molecular cloning of the SNG2 *gene*

A mapping population of 228 F_2 plants was generated by self-pollination of F_1 plants derived from a *sng2*/*sng2* × Columbia wild-type cross. The genotype at the *SNG2* locus for all individuals in the population was determined by TLC analysis of embryos from the F_3 generation derived from self-pollination of the F_2 generation. To roughly map the *SNG2* locus, 19 of the *sng2*/*sng2* lines were used for analysis with the *Arabidopsis* RFLP marker set (Fabri and Schäffner, 1994) obtained from the *Arabidopsis* Biological Resource Center (ABRC, Ohio State University). Subsequent mapping was performed using the simple sequence length polymorphism markers nga249 and nga151 (Bell and Ecker, 1994).

RT–PCR of the sng2 *allele*

The cDNA for the Landsberg *erecta* allele of *F17I14.170* was amplified from total silique RNA using the Access RT–PCR System (Promega, Madison, WI, USA). Primers designed from the AGI annotated cDNA for F17I14.170 were used for the reaction. The forward primer (5′-CACTAAGAAGAAGCAGAG-3′) was designed 21 bp upstream of the predicted start codon, and the reverse primer (5′-CACAGAATAGCATGTGGC-3′) was designed 45 bp downstream of the predicted stop codon. The RT–PCR product was subcloned into the pGEM-T Easy vector using the pGEM-T Easy Vector System (Promega).

Complementation of the sng2 *mutant*

Standard techniques were used for DNA manipulations (Sambrook *et al.*, 1989). For construction of pBI101-*SNG2*, an 8.2 kb *Sac*II fragment of BAC F17I14 containing *F17I14.170* was subcloned into the *Sac*II site of pBS KS$^-$. The pBI101-*SNG2* construct was generated by cloning the 4.0 kb *Avr*II/*Sal*I fragment of the subcloned BAC fragment into the *Xba*I/*Sal*I sites of pBI101 (Clontech, Palo Alto, CA, USA). Plant transformation was performed by floral dip (Clough and Bent, 1998), and transformations with the pBI101-empty vector were included as controls. Transformed seedlings (T_1) were identified by selection on MS medium containing 50 mg l^{-1} kanamycin and 200 mg l^{-1} timentin and were transferred to soil. The sinapoylcholine content of T_2 embryos and seeds was determined by TLC and HPLC as previously described.

Constructs for expression of SNG2 *in* E. coli

Two oligonucleotides were designed to amplify a truncated fragment of the *SNG2* cDNA encoding a protein lacking the predicted SNG2 signal peptide. The N-terminal oligonucleotide (5′-CCATGGCTTTGCTAGTGAAGTCTC-3′) incorporated a start codon and the restriction site *Nco*I (CCATGG) and altered the putative N-terminal serine of the mature protein to an alanine. The C-terminal oligonucleotide (5′-GTCGACTTAGAGAGGTTCT-CCATC-3′) incorporated a *Sal*I restriction site after the stop codon. The *SNG2* gene was amplified by PCR, subcloned into *Eco*RV-digested pBS KS$^-$, and sequenced. The modified coding sequence was excised by *Nco*I/*Sal*I digestion and cloned into the *Nco*I/*Sal*I-digested pET28A vector (Novagen, Madison, WI, USA) to yield pET28A-*SNG2*. For analysis of *SNG2* expression and activity, the *E. coli* host BL21-CodonPlus^{TM} DE3)-RIL (Stratagene, La Jolla, CA, USA) was transformed with the empty pET28A vector and pET28A-*SNG2*.

E.coli *growth conditions and preparation of* E. coli *extracts*

For heterologous expression of *SNG2*, an overnight culture of bacteria grown at 37°C was diluted 200-fold into fresh Luria Bertani (LB) medium and grown at 37°C to an OD_{600} of 0.6. Cells were subsequently induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and grown for 6 h at 30°C. Cells were harvested and lysed in 20 ml of 20 mM Tris–HCl, pH 8.0, and 500 mM NaCl using a French press. The cell lysate was cleared by centrifugation at 14 000 *g* at 4°C for 30 min. Supernatant (soluble protein fraction) and pellet (insoluble protein fraction) were analysed by SDS–PAGE. For heterologous expression of *SNG2* for enzyme assays, an overnight culture of bacteria grown at 37°C was diluted 200-fold into 500 ml of fresh LB medium and grown at 37°C until the OD_{600} of the culture was between 0.4 and 0.6. Cells were subsequently induced with 1 mM IPTG and grown for 3 h at 37°C. The cells were pelleted at 5000 *g* for 10 min, resuspended in 25 ml lysis buffer (25 mM Tris-acetate (pH 7.5), 1 mM EDTA, 0.1% Triton X-100, 0.1 mg ml^{-1} lysozyme, 0.01 mg ml^{-1} RNAase A, 0.05 mg ml^{-1} DNAase I and 2 mM magnesium chloride) and incubated for 10 min on ice. The insoluble fraction was pelleted at 15 000 *g* for 10 min and washed three times in 10 ml wash buffer 1 (50 mM TrisHCl (pH 7.7), 0.3 M sodium chloride, 1 mM EDTA, 0.1% Triton X-100). The pellet was washed with 5 ml wash buffer 2 (wash buffer 1 + 5 mM DTT) and finally resuspended in 1.5 ml wash buffer 2 containing 5% glycerol. The protein content of the sample was quantified on an SDS–PAGE gel using serial dilutions of the inclusion body suspension, after which the sample was stored at –80°C. For renaturation of SNG2 protein, isolated *E. coli* inclusion bodies (approximately 1.2 mg total protein) were resuspended in 0.64 ml 100 mM Tris–HCl (pH 8) containing 8 M urea, 1 mM EDTA and 20 mM DTT. The sample was incubated at room temperature for 2 h with occasional vortex mixing. The solubilized inclusion bodies were then diluted fivefold with 100 mM Tris–HCl (pH 8) containing 8 M urea and 1 mM EDTA to a final protein concentration of approximately 0.4 mg ml^{-1}. Following denaturation, protein folding was initiated using the so-called rapid dilution technique (Rudolph and Lilie, 1996). Eight aliquots (0.4 ml each) of denatured inclusion body protein were slowly added at 15 min intervals to 100 ml of vigorously stirred 100 mM Tris–HCl (pH 8) containing 0.2 mM EDTA, 15% v/v glycerol, 0.01% v/v Tween-20, 3 mM reduced glutathione and 0.6 mM oxidized glutathione at room temperature. The mixture was then incubated for 16 h at

room temperature without stirring. Active recombinant SCT was partially purified by anion exchange chromatography by application of the entire refolding mixture to a column containing 2 ml settled bed volume of QAE Sephadex A-50 (Sigma, St Louis, MO, USA) that had been equilibrated with 50 mM Tris–HCl (pH 7.5). The resin was then washed with 3 ml 50 mM Tris–HCl (pH 7.5) and the eluent discarded. The SCT protein was eluted from the column with two 3 ml washes of 50 mM Tris–HCl (pH 7.5) containing 100 mM NaCl.

Acknowledgements

We thank Dr David Rhodes for his assistance in the choline measurement experiments and Joanne Cusumano for her preparation of the RNA gel blot. This work was supported by grants from the Division of Energy Biosciences, United States Department of Energy, and the National Science Foundation to C.C., and graduate fellowships from the United States Department of Agriculture and Purdue University to A.M.S. This is journal paper number 16609 from the Purdue University Agricultural Experiment Station.

References

Bamforth, C.W., Martin, H.L. and Wainwright, T. (1979) A role for carboxypeptidase in the solubilization of barley β-glucan. *J. Inst. Brew.* **85**, 334–338.

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.

Bell, C.J. and Ecker, J.R. (1994) Assignment of 30 microsatellite loci to the linkage map of Arabidopsis. *Genomics*, **19**, 137–144.

Bradley, D. (1992) Isolation and characterization of a gene encoding a carboxypeptidase Y-like protein from *Arabidopsis thaliana*. *Plant Physiol.* **98**, 1526–1529.

Breddam, K., Widmer, F. and Johansen, J.T. (1980) Influence of the substrate structure on carboxypeptidase Y catalyzed peptide bond formation. *Carlsberg Res. Commun.* **45**, 361–367.

Chapple, C.C.S., Vogt, T., Ellis, B.E. and Somerville, C.R. (1992) An Arabidopsis mutant defective in the general phenylpropanoid pathway. *Plant Cell*, **4**, 1413–1424.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743.

Dahlbender, B. and Strack, D. (1984) Nitrogen nutrition and the accumulation of free and sinapoyl-bound malic acid in *Raphanus sativus* cotyledons. *Planta*, **161**, 142–147.

Doyle, J.J. and Doyle, J.L. (1990) A rapid total DNA preparation procedure for fresh plant tissue. *Focus*, **12**, 13–15.

Fabri, C.O. and Schäffner, A.R. (1994) An *Arabidopsis thaliana* RFLP mapping set to localize mutations to chromosomal regions. *Plant J.* **5**, 149–156.

Fleming, S.M., Robertson, T.A., Langley, G.J. and Bugg, T.D.H. (2000) Catalytic mechanism of a C–C hydrolase enzyme: evidence for a *gem*-diol intermediate, not an acyl enzyme. *Biochemistry*, **39**, 1522–1531.

Goldsbrough, P.B. and Cullis, C.A. (1981) Characterization of the genes for ribosomal RNA in flax. *Nucl. Acids Res.* **9**, 1301–1309.

Hayashi, R., Moore, S. and Stein, W.H. (1973) Serine at 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-L-phenylalanine. *J. Biol. Chem.* **250**, 5221–5226.

Hobson-Frohock, A., Fenwick, G.R., Heaney, R.K., Land, D.G. and Curtis, R.F. (1977) Rapeseed meal and egg taint: association with sinapine. *Br. Poult. Sci.* **18**, 539–541.

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 acyltransferse 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**, 6209–6907.

Liao, D.-I. and Remington, S.J. (1990) Structure of wheat serine carboxypeptidase II at 3.5-Å resolution. A new class of serine proteinase. *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.

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.

Mozo, T., Fischer, S., Shizuya, H. and Altmann, T. (1998) Construction and characterization of the IGF Arabidopsis BAC library. *Mol. Gen. Genet.* **258**, 562–570.

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

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.

Ollis, D.L., Cheah, E., Cygler, M., *et al.* (1992) The α/β hydrolase fold. *Protein Eng.* **5**, 197–211.

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

Ruegger, M., Meyer, K., Cusumano, J.C. and Chapple, C. (1999) Regulation of ferulate-5-hydroxylase expression in Arabidopsis in the context of sinapate ester biosynthesis. *Plant Physiol.* **119**, 101–110.

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.

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., Reinicke, J. and Takeuchi, S. (1986) Evidence of a relationship between malate metabolism and activity of 1-sinapoylglucose: L-malate sinapoyltransferase in radish (*Raphanus sativus* L.) cotyledons. *Planta*, **167**, 212–217.

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

Wajant, H., Mundry, K. and Pfizenmaier, K. (1994) Molecular cloning of hydroxynitrile lyase from *Sorghum bicolor* (L.). Homologies to serine carboxypeptidases. *Plant Mol. Biol.* **26**, 735–746.

Walker-Simmons, M. and Ryan, C.A. (1980) Isolation and

properties of carboxypeptidase from leaves of wounded tomato plants. *Phytochemistry.* **19**, 43–47.

Washio, K. and Ishikawa, K. (1994) Organ-specific and hormone-dependent expression of genes for specific serine carboxypeptidases during development and following germination of rice grains. *Plant Physiol.* **105**, 1275–1280.

Widmer, F., Breddam, K. and Johansen, J.T. (1980) Carboxypeptidase Y catalyzed peptide synthesis using amino acid alkyl esters as amine components. *Carlsberg Res. Commun.* **45**, 453–463.

Yang, W., Nadolska-Orczyk, A., Wood, K.V., *et al.* (1995) Near-isogenic lines of maize differing for glycinebetaine. *Plant Physiol.* **107**, 621–630.

GenBank accession number AY033947 (*SNG2* cDNA).