Running title: Origin of Ethanolamine Moieties

Correspondence to: Andrew D. Hanson

University of Florida

Horticultural Sciences Department

P.O. Box 110690

Gainesville, FL 32611-0690, U.S.A.

Telephone: (352) 392 1928 ext 334

fax: (352) 392 5653

e-mail: adha@mail.ifas.ufl.edu

Subject Area: (4) Proteins, enzymes and metabolism

Title:

Evidence from Engineering That Decarboxylation of Free Serine Is the Major Source of Ethanolamine Moieties in Plants

Authors: **Denis Rontein [1,3], David Rhodes [2] and Andrew D. Hanson [1,4]**

[1] Horticultural Sciences Department, University of Florida, Gainesville, FL 32611-0690, U.S.A.

[2] Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN 47907, U.S.A.

[3] Present address: Librophyt, Centre de Cadarache, Bâtiment 185, DEVM, 13108 St. Paul-Lez-Durance, France

[4] Corresponding author; E-mail, adha@mail.ifas.ufl.edu; Fax, +1-352-392-5653.

Abbreviations: Etn, ethanolamine; PSD, phosphatidylserine decarboxylase; PtdEtn, phosphatidylethanolamine; PtdCho, phosphatidylcholine; PtdSer, phosphatidylserine; SDC, serine decarboxylase.

Abstract

Plants form ethanolamine (Etn) moieties by decarboxylating serine or phosphatidylserine (PtdSer), and use them to make phosphatidylethanolamine, phosphatidylcholine, choline, and glycine betaine. Serine decarboxylation is mediated by a serine decarboxylase (SDC) that is unique to plants and has a characteristic N-terminal extension. This extension was shown to have little influence on function of the enzyme in vitro. To explore the importance of SDC and its extension in vivo, native or truncated versions of the *Arabidopsis* enzyme were expressed in tobacco. Transgene expression increased SDC activity by up to ten-fold and free Etn level up to six-fold, but did not change levels of serine, choline, phosphocholine, or phosphatidyl bases. The truncated enzyme gave significantly higher Etn levels. These results show that SDC activity exerts substantial control over flux to Etn, and suggest that the enzyme's N-terminus may have a regulatory role. In complementary studies with *Arabidopsis*, we showed that a mutant with nine-fold elevated mitochondrial PtdSer decarboxylase activity had normal pools of serine, Etn, and Etn metabolites. Taken together, these data indicate that serine decarboxylation is the main source of Etn moieties in plants. The ability to enhance serine → Etn flux should advance engineering of choline and glycine betaine accumulation.

Keywords: Ethanolamine – Metabolic Engineering – PEST Motif – Phosphatidylserine Decarboxylase – Serine Decarboxylase – Tobacco.

Introduction

Ethanolamine (Etn) moieties are essential for the synthesis of phosphatidylethanolamine (PtdEtn) and phosphatidylcholine (PtdCho), as well as free choline and, in some species, glycine betaine (Mudd and Datko 1989, Rhodes and Hanson 1993) (Fig. 1). Beyond being crucial for membrane biogenesis, this area of plant metabolism has attracted much engineering interest because glycine betaine is a potent osmoprotectant that mitigates damage due to salinity, drought, and other stresses (Nomura et al. 1995, Sakamoto and Murata 2001, Rontein et al. 2002), and because choline is an essential nutrient for humans and other animals (Zeisel 2000, McNeil et al. 2001).

Fig. 1

Plants can synthesize Etn moieties by decarboxylating either free serine or phosphatidylserine (PtdSer) (Mudd and Datko 1989, Rontein et al. 2001, Rontein et al. 2003) (Fig. 1). Decarboxylation of serine is mediated by a soluble, pyridoxal phosphate-dependent serine decarboxylase (SDC) that is unique to plants and has been cloned from *Arabidopsis* and *Brassica napus* (Rontein et al. 2001). Decarboxylation of phosphatidylserine (PtdSer) is catalyzed by PtdSer decarboxylase (PSD), a membrane enzyme that uses a pyruvoyl cofactor, generated post-translationally from a serine residue (Voelker 1997). Mitochondrial PSDs have been cloned from *Arabidopsis* and tomato (*Lycopersicon esculentum*) and found to be like those of other eukaryotes (Rontein et al. 2003). Plants also have extramitochondrial PSDs (Rontein et al. 2003).

Although SDC and PSD genes are known and the recombinant enzymes have been characterized in vitro, it is not clear how much flux each enzyme carries in vivo. This is important from the standpoint of metabolic engineering because the supply of Etn units

becomes limiting in plants engineered to produce more choline and glycine betaine (McNeil et al. 2001). While there is no conclusive evidence on the relative fluxes via SDC and PSD, in-vivo radiotracer results are more readily reconciled with serine decarboxylation than with PtdSer decarboxylation (Mudd and Datko 1989, McNeil et al. 2000). One objective of the present study was accordingly to test this directly by overexpressing SDC and determining the pool sizes of Etn and its metabolites. We also took advantage of an *Arabidopsis* knockup mutant with enhanced mitochondrial PSD activity (Rontein et al. 2003) to investigate the effect of PSD overexpression on levels of Etn-related metabolites.

When SDCs were first cloned from *Arabidopsis* and *B. napus*, we noted a characteristic N-terminal extension that appeared not to be a targeting peptide (Rontein et al. 2001), and many EST sequences have since confirmed this region to be a general feature of plant SDCs. Because such extensions have regulatory roles in other plant enzymes (e.g., Villadsen and Nielsen 2001, Leegood and Walker 2003), another objective of this study was to probe the extension's function in vitro and in planta.

Results

The N-terminal extension of SDC

Alignment of *Arabidopsis* and *B. napus* SDC sequences with those of closely related group II amino acid decarboxylases revealed N-terminal extensions of ~85 residues, of which about the first 60 are much less conserved than the others (Rontein et al. 2001). Fig. 2A confirms and extends this observation by adding SDC orthologs from tomato, *Medicago truncatula*, and soybean, all of which have extensions with roughly the same size and pattern of sequence conservation as in *Arabidopsis* and *B. napus*. When the extensions of

Fig. 2

the five sequences in Fig. 2A were analyzed using ExPASy proteomics tools, none was found to have the features of N-terminal sorting signals (plastid, mitochondrial, or secretory pathway), or to contain predicted membrane-spanning domains. Nor did their variable regions show conserved secondary structure, potential phosphorylation sites, or any other known motif. However, all five extensions contained a strong PEST sequence (PEST score +6.32 to + 17.48) in the conserved region where they join the rest of the protein (Fig. 2A). PEST sequences are considered to target proteins for rapid destruction or – more rarely – for proteolytic processing (Rechsteiner and Rogers 1996, Chen et al. 2002).

To test whether the N-terminal extension is essential for catalytic function, a truncated version of *Arabidopsis* SDC (beginning at Met 58) was expressed in *Escherichia coli* and compared to the full length recombinant enzyme. SDC activity in *E. coli* extracts was similar for both proteins [0.33 and 0.54 μmol (mg protein)$^{-1}$ min^{-1} for full length and truncated enzymes, respectively, at saturating serine concentration] and the apparent K_m value for the truncated enzyme measured in desalted extracts (29 ± 4 mM) was quite comparable to that for the intact enzyme (12 ± 2 mM; Rontein et al. 2001). This evidence for the catalytic competence of the shorter protein led us to compare its effects with those of the full length enzyme when overexpressed in tobacco.

Overexpression of full length and truncated Arabidopsis SDCs in tobacco

cDNAs encoding the complete *Arabidopsis* SDC or the truncated version just described were cloned into the plant expression vector pBI121, which contains the CaMV 35S promoter (Fig. 2B). To improve protein expression, the *Arabidopsis* 5' leader was replaced by the Ω translational enhancer (Dowson-Day et al. 1993), and the sequence context of the

start codon was designed to conform to the tobacco consensus (Nuccio et al. 2000). These constructs were introduced into tobacco, and kanamycin-resistant transformants were selected and screened for SDC activity. The populations of transformants containing the entire or truncated *Arabidopsis* SDC had mean SDC activities that were similar and significantly higher than that for the control population harboring empty vector (Table 1). Both versions of SDC were thus expressed at similar levels in the plant host, as in *E. coli*. There were no visible differences in growth rate, color, or development between the SDC transgenic populations and the empty vector control population, indicating that neither SDC construct was deleterious.

Table 1

From the populations expressing full length or truncated cDNAs, we selected three lines with the highest SDC activities for further study, together with representative empty vector controls. RNA gel blot analysis confirmed a high level of transgene expression in the plants carrying *Arabidopsis* SDC constructs; the hybridizing band was – as predicted – of slightly lower mass in plants carrying the truncated cDNA (Fig. 3). The SDC activities in the SDC transgenics were about four- to ten-fold those in empty-vector controls (Fig. 3); the control values were comparable to those reported previously for tobacco (Rontein et al. 2001).

Fig. 3

Effects of SDC overexpression on pools of ethanolamine and its metabolites

The three transgenic lines with the highest SDC activities (F3, carrying the full length cDNA and T2 and T3, carrying the truncated cDNA) were analyzed for free Etn, its precursor serine, and its metabolites PtdEtn, choline, phosphocholine, and PtdCho (Fig. 4). Empty vector line V3 was analyzed as a benchmark. SDC overexpression had no impact

Fig. 4

on serine levels (Fig. 4A), but significantly increased free Etn levels by 2.6- to 6.6-fold (Fig. 4B). There were no consistent and significant effects on levels of PtdEtn (Fig. 4C) or choline and it metabolites (Fig. 4D). A noteworthy feature of the Etn data (Fig. 4B) is that the level in transgenic line T2 was more than double that in line F3, although the SDC activities in these lines were similar (Fig. 3). We will return to this point in the Discussion.

Pool sizes of ethanolamine metabolites in an Arabidopsis SDC knockup mutant

Our previous work (Rontein et al. 2003) identified a T-DNA insertional mutant of *Arabidopsis* with a nine-fold elevation of mitochondrial PSD activity, and showed its leaf PtdEtn and PtdCho levels to be normal. As this mutant afforded a convenient opportunity to test the effect of PSD overexpression on Etn and its soluble metabolites, we measured the levels of these compounds in mutant and wild type leaves (Table 2). The mutant showed no differences from the wild type for any of the metabolites examined (serine, Etn, choline, and phosphocholine).

Table 2

Digital Northern analysis of SDC and PSD expression

Because EST abundance generally reflects the metabolic flux that enzymes carry (Mekhedov et al. 2000), we determined the abundance in NCBI and TIGR databases of ESTs from eudicot SDC and PSD genes. This digital Northern analysis showed that SDC ESTs are 35-fold more abundant than mitochondrial PSD ESTs (Fig. 5). Eudicot SDC ESTs are also seven-fold more abundant than those encoding a second type of PSD, shown to be extramitochondrial (Rontein et al. 2003). In *Arabidopsis* leaves, extramitochondrial PSD activity is higher than mitochondrial activity (Rontein et al. 2003), which agrees with the relative abundances of their respective ESTs in eudicots (Fig. 5). A cautionary point to

Fig. 5

note about the relative importance of SDC and PSDs is that it may be different in monocots, since these show no massive preponderance of SDC ESTs (Fig. 5).

Discussion

The results presented here indicate that SDC is a major source of Etn moieties in tobacco, since its overexpression caused a large expansion of the free Etn pool. That this expansion was not accompanied by an increase in the level of PtdEtn is in agreement with the lack of effect of exogenously supplied Etn on PtdEtn levels in tobacco (McNeil et al. 2001) and, more generally, with the view that phospholipid synthesis is governed more by the CDP-base formation step than by base or phosphobase synthesis (Kinney et al. 1987). Similarly, that SDC overexpression did not increase the level of choline or its metabolites agrees with the results of supplying Etn exogenously to tobacco (Nuccio et al. 1998), and with evidence that the synthesis of choline moieties is regulated primarily by the activity of the enzyme that methylates the Etn moiety, phosphoethanolamine *N*-methyltransferase (McNeil et al. 2001).

Expressing intact and truncated *Arabidopsis* SDC in *E. coli* and in tobacco established that at least the major part of its N-terminal extension is dispensable for catalytic function both in vitro and in planta. Since similar extensions are present in all available SDC sequences, this region is presumably maintained by selection, even if much of its amino acid sequence is not conserved. This raises the question of what the extension does. Our data point to several possibilities for further investigation. Thus, it may be significant that when intact and truncated SDC are overexpressed at comparable levels, the expansion of the free Etn pool is significantly greater in the case of the truncated enzyme (compare the data for lines

F3 and T2 in Fig. 3 and Fig. 4B). This could connote reduced sensitivity of the truncated enzyme to either (a) feedback inhibition by Etn or its derivatives (e.g., phospho- or CDP-Etn), or (b) an inactivating modification triggered by an accumulation of these metabolites. The presence of a strong PEST motif where the extension joins the rest of the protein suggests that the extension might under some circumstances be cleaved off in vivo, in a process analogous to the proteolytic modification of starch phosphorylase in the region of its PEST sequence (Chen et al. 2002). In this case, our truncated SDC would mimic a form that exists naturally. Alternatively, the PEST sequence could in certain conditions simply target SDC for rapid degradation (Rechsteiner and Rogers 1996), and it is conceivable that removing the extension distal to the PEST motif (as in our truncated construct) interferes with the degradation process and so stabilizes the enzyme. Whatever the case, it is clear that from an engineering standpoint the truncated protein is preferable to the intact one, for it results in enzyme activities that are at least as high, and Etn levels that are higher.

In contrast to SDC in tobacco, overexpression of mitochondrial PSD in *Arabidopsis* did not increase the levels of Etn and its soluble metabolites or, as shown previously (Rontein et al. 2003), the levels of PtdEtn or PtdCho. As the SDC and PSD results were obtained with different species they are clearly not directly comparable. However, *Arabidopsis* and tobacco have roughly comparable SDC activities (Fig. 3C, and Rontein et al. 2001) and similar demands for Etn moieties since neither accumulates glycine betaine (Rhodes and Hanson 1994). In addition, EST abundance data support the view that SDC is a more important source of Etn than PSDs in eudicots in general. Together, these lines of evidence make it fair to assume that the relative contributions of SDC and mitochondrial PSD to the

production of Etn moieties in the two species are also similar, and hence that SDC contributes much more than mitochondrial PSD in both of them.

Finally, it should be noted that the demonstration that SDC overexpression enhances Etn supply is the third milestone in the process of engineering enhanced metabolic flux all the way from serine to choline and glycine betaine (Fig. 1). The first milestone was to install glycine betaine synthesis in plants that naturally lack it by introducing choline-oxidizing enzymes (Hayashi et al. 1997; Nuccio et al. 1998; Huang et al. 2000). The second was to boost choline synthesis from Etn moieties by overexpressing phosphoethanolamine N-methyltransferase (McNeil et al. 2001). A future milestone will be to combine all three transgenic modifications in one plant.

Materials and Methods

Expression of truncated Arabidopsis SDC in E. coli

The SDC coding sequence, starting at Met 58 and ending at the stop codon, was PCR-amplified from an *Arabidopsis* SDC plasmid template (Rontein et al. 2001) using *pfu* polymerase (Stratagene) and the primers 5'-CATGCCATGGTTCTCGGTAGGAAT-3' (forward) and 5'-TGGTGCTCGAGTCACTTGTGAGCTGGACA-3' (reverse), digested with *Nco*I and *Xho*I, and ligated between the *Nco*I and *Xho*I sites of pET28b (Novagen). The construct was electroporated into *E. coli* strain DH10B and then, after sequence verification, into strain BL21 (DE3). For protein production, 50-ml cultures were grown at 37°C to an A_{600} of 0.6 in LB medium containing 100 μg ml^{-1} kanamycin; isopropyl β-D-1-thiogalactopyranoside was then added (1 mM final concentration) and incubation was con-

tinued at 37°C for 3 h. Proteins were extracted and desalted as previously described (Rontein et al. 2001).

Expression of full length and truncated Arabidopsis SDC in tobacco

Full length and truncated coding regions were amplified as above using the forward primers 5'-GGGTATTTTTACAACAATTACCAACAACAACAAACAACAAACAACA-TTACAATTACTATTTACAATTACAAAAATGGTTGGATCTTTGGAATCT-3' and 5'--GGGTATTTTTACAACAATTACCAACAACAACAAACAACAAACAACATTACAA-TTACTATTTACAATTACAAAAATGGTTCTCGGTAGGAATATA-3', respectively. To enhance translation, both these primers contain the tobacco mosaic virus Ω sequence (Dowson-Day et al. 1993) and a sequence context around the start codon that matches the consensus for tobacco and dicotyledons in general (Koziel et al. 1996; Nuccio et al. 2000). The reverse primer for both constructs was 5-TGACGAGCTCTCACTTGTGAGCTGGA-CAG-3'. The amplicons were digested with *Sac*I, and ligated between the *Sma*I and *Sac*I sites of pBI121 (Clontech), which places them behind a CaMV 35S promoter and in front of a NOS terminator. After sequencing, the SDC constructs (and pBI121 from which the β-glucuronidase gene had been excised) were introduced into *Agrobacterium tumefaciens* strain ABI. Tobacco *cv.* Wisconsin 38 was then transformed as described (Horsch et al. 1985). Transformants were cultured in Magenta boxes and clonally propagated as detailed previously (Nuccio et al. 1998). Plants two to four weeks of age were taken for analyses.

RNA gel blot analysis

Total RNA was extracted from tobacco leaves using RNeasy Plant Mini Kits (Qiagen), and separated by agarose gel electrophoresis using the following buffers: buffer A – 200 mM

MOPS, 50 mM Na-acetate, 10 mM EDTA, pH 7.0; buffer B – 0.16% v/v saturated bromophenol blue, 4 mM EDTA, pH 8.0, 0.89 mM formaldehyde, 20% v/v glycerol, 31% v/v formamide, 40% v/v buffer A. RNA dissolved in 20% v/v of buffer B was loaded onto a 1.2% agarose gel containing 10% v/v buffer A, 0.22 M formaldehyde, and 10 μg ml^{-1} ethidium bromide. The gel was run in 10% v/v buffer A plus 0.25 M formaldehyde. RNA was transferred by capillary elution to a nitrocellulose membrane (Protran BA 85, Schleicher & Schuell). Hybridization was at 65°C in 6× SSC, 2× Denhardt's reagent, and 0.1% SDS. The blot was washed once in 1× SSC, 0.1% SDS at room temperature, followed by three washes in 0.2× SSC, 0.1% SDS at 65°C. The SDC probe, corresponding to the last 930 bp of the SDC ORF, was amplified using *Pfu* polymerase (Stratagene) and the primers 5'-TGGTACTGAAGGCAACCT-3' (forward) and 5'-TGTCCAGCTCACAA-GTGA-3' (reverse). The amplicon was gel-purified, and labeled with ^{32}P using a Random Primed DNA Labeling Kit (Roche).

SDC extraction and assay

Tobacco leaf tissue (100 mg) was ground in liquid N_2 and extracted in five volumes of 200 mM K-phosphate, pH 7.2, 15 mM Na-ascorbate, 2.5 mM dithiothreitol, 0.2 mM pyridoxal 5'-phosphate, and 0.5% (w/v) polyvinylpolyprrolidone. After centrifuging (16,000×*g*, 10 min), extracts were desalted on 1-ml spin columns of Sephadex G-25 resin equilibrated with 50 mM K-phosphate, pH 7.2, 0.1 mM pyridoxal 5'-phosphate, and 2.5 mM dithiothreitol. Protein concentrations were determined with the Bradford reagent (BioRad). SDC activity was assayed by measuring formation of [^{14}C]Etn from L-[U-^{14}C]serine (160 μCi μmol^{-1}) (Rontein et al. 2001). Except for kinetic measurements, the serine concentration in

assays was 50 μM. SDC activity data were subjected to one-way analysis of variance after log transformation.

Growth of the Arabidopsis mitochondrial PSD mutant

A T-DNA insertional mutant of *Arabidopsis* (*Arabidopsis thaliana* L. (Heyn), ecotype Columbia) was obtained from the Torrey Mesa Research Institute (line SAIL_508_C12) and characterized as a mitochondrial PSD knockup mutant (Rontein et al. 2003). Homozygous mutant and wild type plants were grown in Super Fine Germination Mix (Fafard, Agawam, MA, U.S.A.) at 22°C in 12-h days (80-150 μE m^{-2} s^{-1}) and irrigated with water. Rosette leaves were harvested at day 21, pooling leaves from five plants into 0.5-g batches for metabolite analysis. Leaves were frozen in liquid N_2 at once and stored at –80°C until used.

Isolation and determination of metabolites

Frozen leaf tissue (0.5 g) was extracted using a methanol-chloroform-water procedure, and then processed essentially as described (Nuccio et al. 1998). Briefly, the methanol-chloroform phase (containing PtdEtn and PtdCho) was evaporated, then hydrolyzed in 4 N HCl at 110°C for 18 h. The hydrolysate was lyophilized, dissolved in water and applied to a 1-ml BioRad AG-1 (OH^-) column in series with a 1-ml BioRex-70 (H^+) column. Etn and Cho were eluted from the latter with 1 N HCl. Serine (and other amino acids) was isolated from the aqueous phase by passing a sample through a 1-ml BioRad AG-50 (H^+) column, and eluting with 6 M NH_4OH. To isolate Etn, choline, and their phospho derivatives, a sample of the aqueous phase was applied to a 1-ml AG-1 (OH^-) column in series with a 1-ml BioRex 70 (H^+) column. Etn and Cho were eluted from BioRex 70 with 1 N HCl. Phosphobases were eluted from AG-1 with 2.5 N HCl, and hydrolyzed in 6 N HCl at 110°C for

36 h. After lyophilizing, the hydrolysate was applied to a second AG-1 (OH^-)/BioRex-70 (H^+) column series, and the free bases were eluted from BioRex-70 with 1 N HCl. [It should be noted that AG-1 (OH^-) binds reducing sugars; these undergo Maillard reactions with amino groups during the hydrolysis step, reducing the recovery of the Etn moiety of phosphoethanolamine. In the present series of experiments, phosphoethanolamine recoveries were too low and variable to allow reliable quantification.] Choline was assayed by the enzymatic method of Nie et al. (1993), as adapted by Nuccio et al. (1998). Eluates containing amino acids or Etn were dried and derivatized to N-(O,S)-heptafluorobutyryl isobutyl esters by reaction with isobutanol:acetyl chloride (5:1, v/v, 120°C, 20 min), followed by heptafluorobutyric anhydride (120°C, 10 min) (Rhodes et al. 1986). After evaporating excess heptafluorobutyric anhydride, samples were redissolved in 100 μL of ethyl acetate:acetic anhydride (1:1, v/v) for analysis by GC and by electron ionization GC-MS as described (Rhodes et al. 1986), except that a GCQ mass spectrometer (Thermo Finnigan, San Jose, CA) was used for GC-MS. Etn gave the following derivative: CF_3-CF_2-CF_2-(C=O)-NH-CH_2-CH_2-O-(C=O)-CF_2-CF_2-CF_3 (molecular weight M = 453, $M+H^+$ = 454) using the same derivatization scheme as for amino acids (above). Because this derivative is highly volatile, a starting oven temperature of 85°C was used to help separate it from the solvent; the oven temperature program used was 85°C for 4 min, then rising to 250°C at 8°C min^{-1}. The GC was calibrated with an external amino acid standard mixture (AA-S-18; Sigma) spiked with 3,4-dehydro-D,L-proline (2.5 nmol μl^{-1} of each amino acid derivative in ethyl acetate:acetic anhydride, 1:1, v/v), and Etn.HCl (7.5 nmol μl^{-1}). Representative plant samples were spiked with known amounts of Etn and 3,4-dehydro-D,L-proline to

determine recoveries, which were then applied to the data. Etn data were subjected to one-way analysis of variance after log transformation.

Bioinformatics

BLAST searches of EST databases were performed at the websites of the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov) and The Institute for Genomic Research (www.tigr.org/tdb/tgi). The significance of EST abundance data was analyzed at the Information Génomique et Structurale website (igs-server.cnrs-mrs.fr/~audic/significance.html) (Audic and Claverie 1997). Structural features of the N-terminal region of SDC sequences were analyzed using ExPASy proteomics tools (us.expasy.org/tools).

Acknowledgments

We thank Barb Wood (Department of Horticulture and Landscape Architecture, Purdue University) for technical assistance in Etn and amino acid analysis. This work was supported in part by the Florida Agricultural Experiment Station, by an endowment from the C.V. Griffin, Sr. Foundation and by USDA NRI CGP grant no. 2001-35100-10620 (to A.D.H), and approved for publication as Journal Series no. R-09627.

References

Audic, S. and Claverie, J.M. (1997) The significance of digital gene expression profiles. *Genome Res.* 7: 986-995.

Chen, H.M., Chang, S.C., Wu, C.C., Cuo, T.S., Wu, J.S. and Juang, R.H. (2002) Regulation of the catalytic behaviour of L-form starch phosphorylase from sweet potato roots by proteolysis. *Physiol Plant.* 114: 506-515.

Delhaize, E., Hebb, D.M., Richards, K.D., Lin, J.M., Ryan, P.R. and Gardner, R.C. (1999) Cloning and expression of a wheat (*Triticum aestivum* L.) phosphatidylserine synthase cDNA. Overexpression in plants alters the composition of phospholipids. *J. Biol. Chem.* 274: 7082-7088.

Dowson-Day, M.J., Ashurst, J.L., Mathias, S.F., Watts, J.W., Wilson, T.M. and Dixon, R.A. (1993) Plant viral leaders influence expression of a reporter gene in tobacco. *Plant Mol. Biol.* 23: 97-109.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G. and Fraley, R.T. (1985) A simple and general method for transferring genes into plants. *Science* 227: 1229-1231.

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

Kinney, A.J., Clarkson, D.T. and Loughman, B.C. (1987) The regulation of phosphatidylcholine biosynthesis in rye (*Secale cereale*) roots. Stimulation of the nucleotide pathway by low temperature. *Biochem. J.* 242: 755-759.

Koziel, M.G., Carozzi, N.B.and Desai, N. (1996) Optimizing expression of transgenes with an emphasis on post-transcriptional events. *Plant Mol. Biol.* 32: 393-405.

Leegood, R.C. and Walker, R.P. (2003) Regulation and roles of phospho*enol*pyruvate carboxykinase in plants. *Arch. Biochem. Biophys.* 414: 204-210.

Hayashi, H., Alia, Mustardy, L., Deshnium, P., Ida, M. and Murata, N. (1997) Transformation of *Arabidopsis thaliana* with the *codA* gene for choline oxidase; accumulation of glycinebetaine and enhanced tolerance to salt and cold stress. *Plant J.* 12: 133-142.

McNeil, S.D., Nuccio, M.L., Ziemak, M.J. and Hanson, A.D. (2001) Enhanced synthesis of choline and glycine betaine in transgenic tobacco plants that overexpress phosphoethanolamine *N*-methyltransferase. *Proc. Natl. Acad. Sci. USA* 98: 10001-10005.

McNeil, S.D., Rhodes, D., Russell, B.L., Nuccio, M.L., Shachar-Hill, Y. and Hanson, A.D. (2000) Metabolic modeling identifies key constraints on an engineered glycine betaine synthesis pathway in tobacco. *Plant Physiol.* 124: 153-162.

Mekhedov, S., de Ilarduya, O.M. and Ohlrogge, J. (2000) Toward a functional catalog of the plant genome. A survey of genes for lipid biosynthesis. *Plant Physiol.* 122: 389-402.

Mudd, S.H. and Datko, A.H. (1989) Synthesis of ethanolamine and its regulation in *Lemna paucicostata*. *Plant Physiol.* 91: 587-597.

Nie, Y., He, J.L. and Hsia, S.L. (1993) A micro enzymic method for determination of choline-containing phospholipids in serum and high density lipoproteins. *Lipids* 28: 949-951.

Nomura, M., Ishitani, M., Takabe, T., Rai, A.K. and Takabe, T. (1995) *Synechococcus* sp. PCC7942 transformed with *Escherichia coli bet* genes produces glycine betaine from choline and acquires resistance to salt stress. *Plant Physiol.* 107: 703-708.

Nuccio, M.L., McNeil, S.D., Ziemak, M.J., Hanson, A.D., Jain, R.K. and Selvaraj, G. (2000) Choline import into chloroplasts limits glycine betaine synthesis in tobacco: analysis of plants engineered with a chloroplastic or a cytosolic pathway. *Metab. Eng.* 2: 300-311.

Nuccio, M.L., Russell, B.L., Nolte, K.D., Rathinasabapathi, B., Gage, D.A. and Hanson, A.D. (1998) The endogenous choline supply limits glycine betaine synthesis in transgenic tobacco expressing choline monooxygenase. *Plant J.* 16: 487-496.

Rechsteiner, M. and Rogers, S.W. (1996) PEST sequences and regulation by proteolysis. *Trends Biochem. Sci.* 21: 267-271.

Rhodes, D., Handa, S. and Bressan, R.A. (1986) Metabolic changes associated with adaptation of plant cells to water stress. *Plant Physiol.* 82: 890-903.

Rhodes, D. and Hanson, A.D. (1993) Quaternary ammonium and tertiary sulfonium compounds in higher plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 44: 357-384.

Rontein, D., Basset, G. and Hanson, A.D. (2002) Metabolic engineering of osmoprotectant accumulation in plants. *Metab. Eng.* 4: 49-56.

Rontein, D., Nishida, I., Tashiro, G., Yoshioka, K., Wu, W.I., Voelker, D.R., Basset, G. and Hanson, A.D. (2001) Plants synthesize ethanolamine by direct decarboxylation of serine using a pyridoxal phosphate enzyme. *J. Biol. Chem.* 276: 35523-35529.

Rontein, D., Wu, W.I., Voelker, D.R. and Hanson, A.D. (2003) Mitochondrial phosphatidylserine decarboxylase from higher plants. Functional complementation in yeast, localization in plants, and overexpression in Arabidopsis. *Plant Physiol.* 132: 1678-1687.

Sakamoto, A. and Murata, N. (2001) The use of bacterial choline oxidase, a glycinebetaine-synthesizing enzyme, to create stress-resistant transgenic plants. *Plant Physiol.* 125: 180-188.

Villadsen, D. and Nielsen, T.H. (2001) N-terminal truncation affects the kinetics and structure of fructose-6-phosphate 2-kinase/fructose-2,6-bisphosphatase from *Arabidopsis thaliana*. *Biochem. J.* 359: 591-597.

Vincent, P., Maneta-Peyret, L., Sturbois-Balcerzak, B., Duvert, M., Cassagne, C. and Moreau, P. (1999) One of the origins of plasma membrane phosphatidylserine in plant cells is a local synthesis by a serine exchange activity. *FEBS Lett.* 464: 80-84.

Voelker, D.R. (1997) Phosphatidylserine decarboxylase. *Biochim. Biophys. Acta* 1348: 236-244.

Zeisel, S.H. (2000) Choline: an essential nutrient for humans. *Nutrition* 16: 669-671.

Table 1 SDC activities in tobacco plantlets transformed with native or truncated *Arabidopsis* SDC, or with empty vector

Construct	Population size	SDC Activity [pmol (mg protein)$^{-1}$ h^{-1}]	
		Population mean	Standard error
Empty vector	24	10.8	0.9
Full length SDC	42	26.7	3.7
Truncated SDC	44	35.3	6.1

SDC activity was measured in desalted extracts of duplicate or triplicate leaf samples from the indicated number of independent transformants for each construct. The SDC activity means of the SDC populations are significantly different ($P = 0.05$) from the empty vector control, but not from each other.

Table 2 Quantification of serine, ethanolamine and soluble ethanolamine metabolites in leaves of wild type *Arabidopsis* and a knockup PSD mutant

Metabolite	Metabolite content [nmol (g fresh weight)$^{-1}$]	
	Wild type	Knockup mutant
Serine	2070 ± 50	3000 ± 400
Etn	255 ± 102	193 ± 46
Choline	332 ± 26	306 ± 28
Phosphocholine	49 ± 3	64 ± 8

Values are means ± SE for three independent samples, each comprising 0.5 g of rosette leaves from 21-day old plants.

Figure Legends

Fig. 1 The synthesis and metabolism of ethanolamine moieties in plants. Dotted arrows show alternative routes to PtdSer via base exchange (BE) or PtdSer synthase (PSS), which both occur in plants (Vincent et al. 1999, Delhaize et al. 1999). The CDP-base intermediates between phospho- and phosphatidyl bases are omitted for simplicity. P, phospho; Ptd, phosphatidyl; SDC, serine decarboxylase; PSD, PtdSer decarboxylase; Cho, choline; Etn, ethanolamine; Me, methyl; GlyBet, glycine betaine.

Fig. 2 Structure of SDCs and plant expression constructs. (A) Alignment of the N-terminal regions of *Arabidopsis* and *B. napus* SDCs, and orthologous sequences from three other dicotyledons. Identical residues are shaded in black, similar residues in gray. Dashes are gaps introduced to maximize alignment. The angled arrow marks the start of the N-terminal extension relative to histidine decarboxylases from *Morganella morganii* and other bacteria (the closest relatives of SDC). The bar indicates a strong PEST motif (PEST scores +6.32 to +17.48). The open triangle shows the first residue of the truncated *Arabidopsis* SDC. At, *Arabidopsis* SDC (GenBank AAK77493); Bn, *Brassica napus* SDC (GenBank BAA78331); Le, *Lycopersicon esculentum* (TIGR contig TC124323); Mt, *Medicago truncatula* (TIGR contig TC85674); Gm, *Glycine max* (TIGR contig TC160683). All three TIGR contigs comprise multiple ESTs authenticating the sequence of the region of interest. (B) Constructs used to express the full-length or truncated *Arabidopsis* SDC in tobacco. The arrow represents the transcription start site. Ω is the Ω translational enhancer sequence from tobacco mosaic virus. Numbers refer to amino acids.

Fig. 3 Expression of full length (F) or truncated (T) *Arabidopsis* SDC cDNAs in tobacco leaves. Data for empty vector controls (V) are included for comparison. (A) RNA gel blot showing SDC mRNA levels. Lanes contained 5 μg of total RNA; the blot was probed with the final 930 bp of the *Arabidopsis* SDC ORF. (B) RNA loading control for the RNA gel blot, stained with ethidium bromide. (C) SDC activities in desalted leaf extracts, measured at 50 μM serine. Values are means and SE of measurements on three individual plants from each transgenic line. Means marked by the same letter are not significantly different at $P = 0.05$ according to Duncan's multiple range test.

Fig. 4 Levels of ethanolamine and related metabolites in leaves of tobacco lines transformed with full length (F3) or truncated (T2, T3) *Arabidopsis* SDC cDNAs, or with empty vector (V3). The lines are the same as those with corresponding designations in Fig. 3. Data are means and SE of determinations on three individual plants from each transgenic line, and are expressed on a fresh weight (FW) basis. A, Free serine. B, Free Etn. C, PtdEtn. D, Free choline, phosphocholine, and PtdCho. For free Etn, means marked by the same letter are not significantly different at $P = 0.05$ according to Duncan's multiple range test. For other metabolites, differences between lines were not significant.

Fig. 5 Digital gene expression profiles for SDC and PSDs for dicotyledons and monocotyledons. The GenBank dBEST database was searched for ESTs homologous to SDC, mitochondrial PSD (PSD-m), and extramitochondrial PSDs (PSD-e) from *Arabidopsis*. EST frequencies are expressed as % of the total number of EST entries from dicotyledons (1.57×10^6) or monocotyledons (1.41×10^6). Values not marked by the same letter differ significantly at $P \leq 0.01$ (Audic and Claverie 1997).

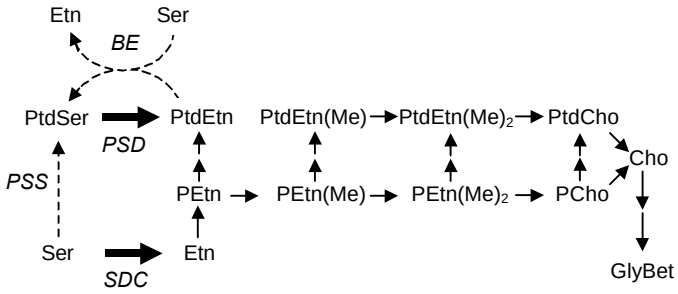
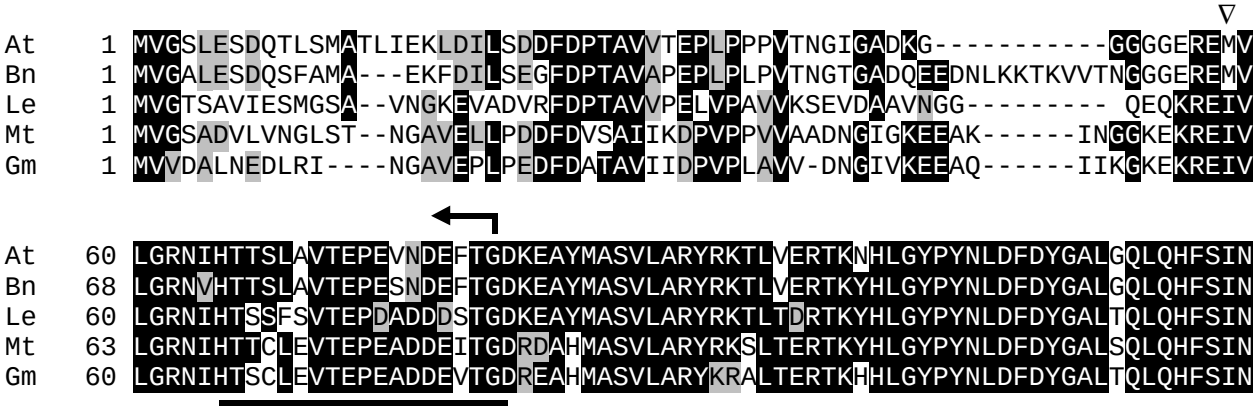

Etn
Ser
BE
PtdSer
PSD
PtdEtn
PtdEtn(Me)
PtdEtn(Me)2
PtdCho
Cho
PSS
PEtn
PEtn(Me)
PEtn(Me)2
PCho
Ser
SDC
Etn
GlyBet

Figure 1

A


∇
At 1 MVGSLESDQTLSMATLIEKLDILSDDFDPTAVVTEPLPPPVTNGIGADKG----------GGGGEREMV
Bn 1 MVGALESDQSFAMA---EKFDILSEGFDPTAVAPEPLPLPVTNGTGADQEEDNLKKTKVVTNGGGEREMV
Le 1 MVGTSAVIESMGSA--VNGKEVADVRFDPTAVVPELVPAVVKSEVDAAVNGG--------- QEQKREIV
Mt 1 MVGSADVLVNGLST--NGAVELLPDDFDVSAIIKDPVPPVVAADNGIGKEEAK------INGGKEKREIV
Gm 1 MVVDALNEDLRI----NGAVEPLPEDFDATAVIIDPVPLAVV-DNGIVKEEAQ------IIKGKEKREIV
At 60 LGRNIHTTSLAVTEPEVNDEFTGDKEAYMASVLARYRKTLVERTKNHLGYPYNLDFDYGALGQLQHFSIN
Bn 68 LGRNVHTTSLAVTEPESNDEFTGDKEAYMASVLARYRKTLVERTKYHLGYPYNLDFDYGALGQLQHFSIN
Le 60 LGRNIHTSSFSVTEPDADDDSTGDKEAYMASVLARYRKTLTDRTKYHLGYPYNLDFDYGALTQLQHFSIN
Mt 63 LGRNIHTTCLEVTEPEADDEITGDRDAHMASVLARYRKSLTERTKYHLGYPYNLDFDYGALSQLQHFSIN
Gm 60 LGRNIHTSCLEVTEPEADDEVTGDREAHMASVLARYKRALTERTKHHLGYPYNLDFDYGALTQLQHFSIN

B

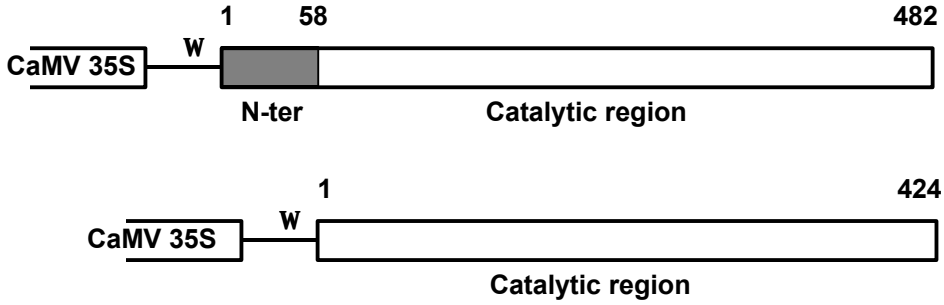

1
58
482
W
CaMV 35S
N-ter
Catalytic region
1
424
W
CaMV 35S
Catalytic region

Figure 2

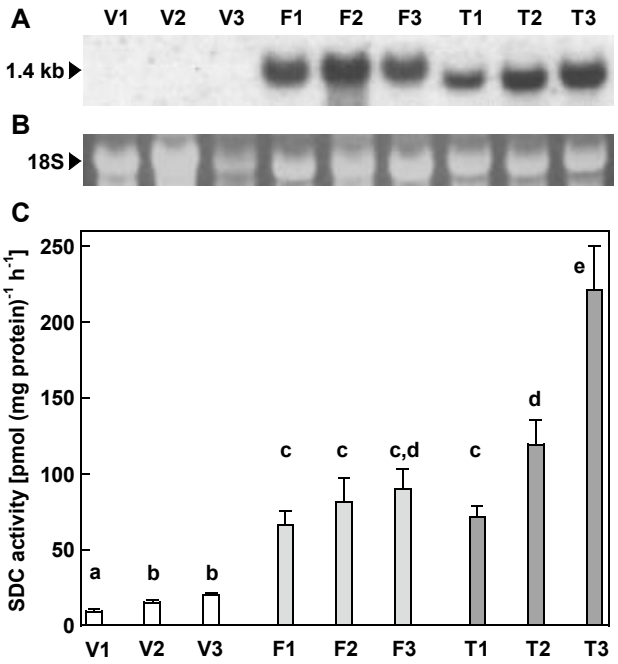

A
V1 V2 V3 F1 F2 F3 T1 T2 T3
1.4 kb
B
18S
C
SDC activity [pmol (mg protein)-1 h-1]
250
200
150
100
50
0
a b b c c c,d c d e
V1 V2 V3 F1 F2 F3 T1 T2 T3

Figure 3

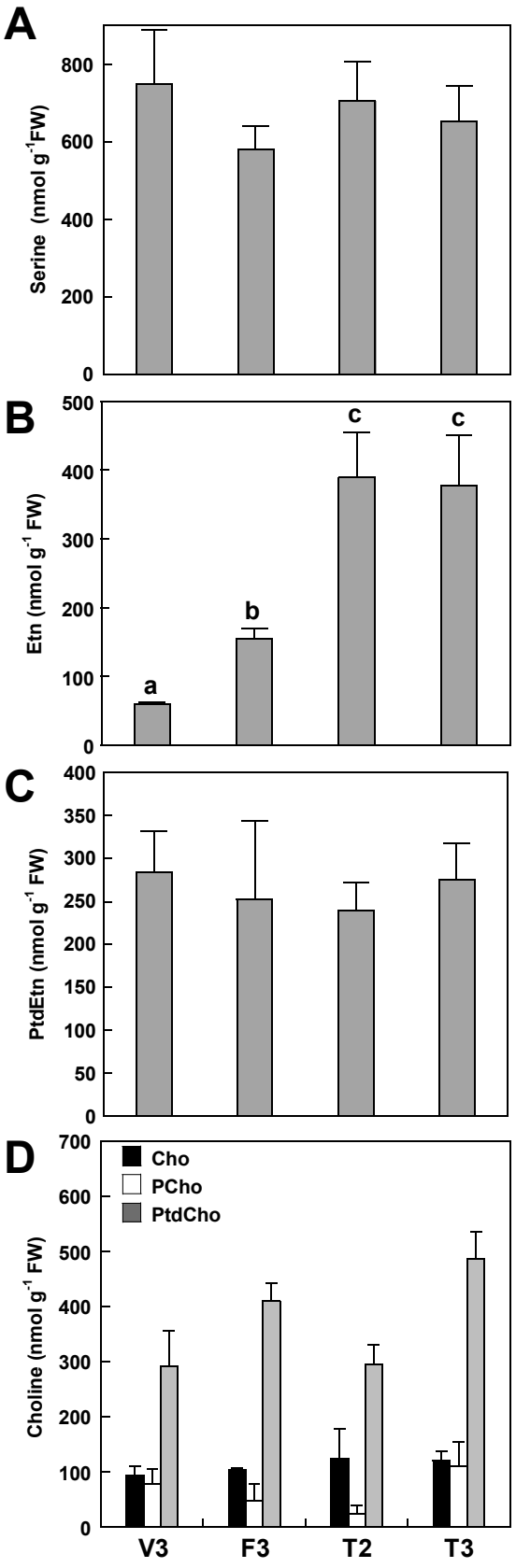

A
Serine (nmol g-1FW)
0
200
400
600
800
B
Etn (nmol g-1 FW)
0
100
200
300
400
500
a
b
c
c
C
PtdEtn (nmol g-1 FW)
0
50
100
150
200
250
300
350
400
D
Choline (nmol g-1 FW)
0
100
200
300
400
500
600
700
Cho
PCho
PtdCho
V3
F3
T2
T3

Figure 4

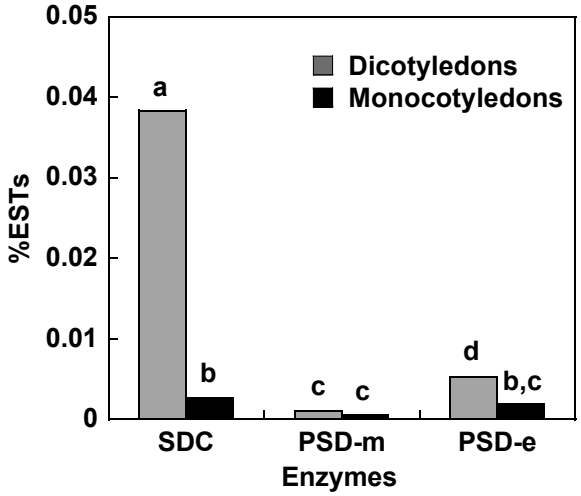

0.05
0.04
0.03
0.02
0.01
0
%ESTs
Dicotyledons
Monocotyledons
a
b
c
c
d
b,c
SDC
PSD-m
PSD-e
Enzymes

Figure 5