Enzyme activities and subcellular localization of members of the Arabidopsis glutathione transferase superfamily

David P. Dixon, Timothy Hawkins, Patrick J. Hussey and Robert Edwards*
School of Biological and Biomedical Sciences, Durham University, Durham DH1 3LE, UK

Received 14 October 2008; Revised 1 December 2008; Accepted 22 December 2008

Abstract

Enzyme screens with Strep-tagged recombinant proteins and expression studies with the respective green fluorescent protein (GFP) fusions have been employed to examine the functional activities and subcellular localization of members of the Arabidopsis glutathione transferase (GST) superfamily. Fifty-one of 54 GST family members were transcribed and 41 found to express as functional glutathione-dependent enzymes in Escherichia coli. Functional redundancy was observed and in particular three theta (T) class GSTs showed conserved activities as hydroperoxide-reducing glutathione peroxidases (GPOXs). When expressed in tobacco as GFP fusions, all three GSTTs localized to the peroxisome, where their GPOX activity could prevent membrane damage arising from fatty acid oxidation. Through alternative splicing, two of these GSTTs form fusions with Myb transcription factor-like domains. Examination of one of these variants showed discrete localization within the nucleus, possibly serving a role in reducing nucleic acid hydroperoxides or in signalling. Based on this unexpected differential sub-cellular localization, 15 other GST family members were expressed as GFP fusions in tobacco. Most accumulated in the cytosol, but GSTU12 localized to the nucleus, a family member resembling a bacterial tetrachlorohydroquinone dehalogenase selectively associated with the plasma membrane, and a lambda GSTL2 was partially directed to the peroxisome after removal of a putative chloroplast transit peptide. Based on the results obtained with the GSTTs, it was concluded that these proteins can exert identical protective functions in differing subcellular compartments.

Key words: Alternative splicing, confocal microscopy, glutathione peroxidase, green fluorescent protein, lipid hydroperoxides, Nicotiana benthamiana, peroxisome, Strep tag.

Introduction

Glutathione transferases (GSTs; EC 2.5.1.18) in plants are a superfamily of proteins which can be clustered into six groupings based on similarities in sequence and gene organization, namely the phi (F), tau (U), theta (T), zeta (Z), lambda (L), and dehydroascorbate reductase (DHAR) classes (Dixon et al., 2002b). Recent informatic studies in Arabidopsis thaliana (L.) Heynh have also revealed a further single representative of an additional type of GST which most closely resembles the bacterial tetrachlorohydroquinone dehalogenase (TCHQD) (Edwards and Dixon, 2005). Recently progress has been made in defining the function of some of the smaller groups of plant GSTs. Thus GSTZs appear to function as glutathione (GSH)-dependent isomerases in tyrosine catabolism (Dixon and Edwards, 2006), while the DHARs are active in ascorbic acid recycling (Dixon et al., 2002a). By contrast, the functions of the majority of plant GSTs remain unknown. Using Arabidopsis as an example, individual GSTs have been shown to be involved in flavonoid metabolism (Kitamura et al., 2004) and signalling (Chen et al., 2007), with many studies also implicating them in a wide range of responses to stress, infection, and plant hormones (Moons, 2005). An underlying problem in studying the functions of these proteins by classical molecular genetic methods is that multiple members of this large gene family can have identical roles (Alfenito et al., 1998), with this redundancy concealing function on gene disruption (Bouché and Bouchez, 2001).
As an alternative to genetic studies, it has been of interest to use a combination of biochemical and histochemical methods to investigate the functions of individual GSTs from *Arabidopsis*. As a first step, the members of this superfamily of genes encoding functional GSH-dependent transferases and peroxidases needed to be identified. At the next level of enquiry, it was then helpful to know where individual GSTs are expressed in the plant cell, as that could focus the future search for likely physiologically relevant substrates or binding partners. The importance of understanding the sub-cellular localization of GSTs when defining the function of these proteins has recently been demonstrated. Thus, when maize tau class GSTs were expressed in bacteria they were found to interact with porphyrinogen intermediates of haem biosynthesis, catalysing their GSH conjugation (Dixon et al., 2008). Similarly, the transplastomic expression of these proteins in tobacco caused similar changes in porphyrin metabolism. However, despite the apparent selectivity of these GSTs for binding and conjugation of porphyrin intermediates, this does not appear to be their function in planta due to their exclusive expression in the cytosol where they are physically separated from mainstream tetrapyrrole biosynthesis in the chloroplast (Tanaka and Tanaka, 2007).

Relatively little attention has been paid to the cellular localization of plant GSTs, as the majority lack obvious sub-cellular targeting sequences and are therefore predicted to accumulate in the cytosol. In *Arabidopsis*, exceptions to this are seen with members of the DHAR (e.g. DHAR3) and lambda (GSTL2) classes where unambiguous N-terminal targeting peptides direct these proteins to the plastid/mitochondrion (Dixon et al., 2002a). The phi class GSTF8 also contains a plastid-targeting peptide, though differential mRNA splicing leads to the majority of the derived transcripts lacking this signal, giving cytosolic protein (Thatcher et al., 2007). Despite their predicted cytosolic localization, several proteomic studies in *Arabidopsis* and other plants have consistently shown that various organelar preparations contain high titres of specific GSTs. For example, comprehensive proteomic analyses in *Arabidopsis* have identified GSTs F2, F8, F9, F10, U19, U20, L2, and DHAR3 in the chloroplast (Zybalov et al., 2008), GSTF5 and GSTF6 in the mitochondrion (Heazlewood et al., 2004), and GSTs F2, F6, F7, F9, and F10 in the vacuole (Carter et al., 2004). Earlier studies in other plants have also reported specific GSTs accumulating in the nucleus (Takahashi et al., 1995) and apoplast (Flury et al., 1996). The presence of these GSTs in organelles may in part be due to cross-contamination from the cytosol, where these hydrophobic proteins can be present at high concentrations (Sappl et al., 2004). Alternatively, it is also possible that these GSTs associate with different sub-cellular fractions as a consequence of their association with other proteins or substrates/ligands and as such an understanding of their sub-cellular localization could be very important in unravelling their endogenous function.

Reported here are the cloning and expression of the GST superfamily from *Arabidopsis* and the functional screening of the recombinant proteins of currently unknown function for GSH-dependent enzyme activities. The subcellular localization of *Arabidopsis* GST family members which lack obvious plastid/mitochondrial targeting sequences has been examined by generating N-terminal green fluorescent protein (GFP) fusions and determining their localization using confocal microscopy after transiently expressing the fusions in *Nicotiana benthamiana*. A total of 18 family members were selected based on (i) a careful examination of the respective sequences of these proteins to look for C-terminal targeting domains and (ii) proteomics literature suggesting that ‘cytosolic’ *Arabidopsis* GSTs, or orthologues in other plants, were associated with specific organelles. In each case, a library of GSTs cloned into a custom Strep-tagged bacterial pET-derived expression plasmid was used, which allowed for the straightforward sub-cloning of sequences into the plant transformation binary vector.

**Materials and methods**

**Vectors**

The expression vectors pET-STRP3 and BIN-STRP3 were synthesized as described (Dixon et al., 2008; Supplementary data available at *JXB* online). For plant expression of N-terminal GFP fusions, the vector BIN-STRP3 was modified such that a GFP tag was introduced between the Strep-tag and the cloned protein. This construct was prepared by amplifying GFP from a plasmid containing smGFP (EMBL accession U70495) (Davis and Vierstra, 1998) using the oligonucleotides GCGCGCGATCGTGAAGAAGGAGAAGAAC and GCGCGGCCCTGCAAGCCCTTAAATTTGTATAGTTCATCCATGC. The PCR product was digested using *Pvu*I and *Sfi*I and ligated into BIN-STRP3 digested with *Pac*I and *Sfi*I to give BIN-STRP3–GFP. Vectors (for sequences, see Supplementary data available at *JXB* online) are available upon request.

**Cloning of GSTs**

Most GST sequences were PCR-amplified from cDNA prepared from a combination of *Arabidopsis* plants and root cultures (Dixon et al., 2002a), using KOD polymerase (Novagen, Nottingham, UK) and the primers detailed in Table S1 in Supplementary data available at *JXB* online. To amplify GSTF12, mature *Arabidopsis* plants were placed under high light conditions (25 °C, 200 μE m⁻² s⁻¹) for 48 h prior to RNA extraction and cDNA synthesis. Amplified GST sequences were digested with restriction sites engineered into the primers and cloned into similarly digested pET-STRP3. Sequences were sub-cloned from pET-vectors to BIN-vectors typically using *Pac*I and *Bst*I restriction sites (Fig. 1A).

**GST expression in Escherichia coli**

The pET-STRP3 constructs were transformed into *E. coli* strain Tuner(DE3) (Novagen) containing the pRARE plasmid from strain Rosetta (Novagen). Cultures were grown in LB broth containing appropriate antibiotics and...
0.1 mM IPTG for 16–24 h at 30 °C with shaking at 200 r.p.m. in the dark. Cells were harvested by centrifugation and re-suspended in 5% of the original culture volume in cold (4 °C) HEPES-buffered saline (HBS = 20 mM HEPES–NaOH, 150 mM NaCl, 1 mM EDTA, pH 7.5). After adding DTT (1 mM) and avidin (8 μg ml⁻¹) to remove endogenous biotin and biotinylated proteins, the cells were sonicated and centrifuged, and the lysate loaded (1 ml min⁻¹) onto a 1 ml Strep-Tactin macroprep column (Stratech Scientific Ltd, Soham, UK) pre-equilibrated at 22 °C with HBS. After washing with HBS, recombinant protein was eluted with HBS containing 2.5 mM desthiobiotin and the column regenerated with HBS (10 ml) containing 1 mM 2-(4-hydroxy-benzeneazo)-benzoic acid. Recombinant protein eluting in a single 1 ml fraction was flash-frozen in 10% v/v glycerol.

**Enzyme assays**

GST activity was determined with 1-chloro-2,4-dinitrobenzene (CDNB) and benzylisothiocyanate (BITC) as substrates (Dixon et al., 1998). Glutathione peroxidase (GPOX) activity was determined with cumene hydroperoxide and fatty acid hydroperoxides prepared as described previously (Edwards and Dixon, 2005). Assays with fatty acid hydroperoxides used these substrates at a final concentration of 0.2 mM in the presence of 0.1% (v/v) Triton X-100.

**GFP–GST expression in Nicotiana benthamiana Domin**

GST coding sequences sub-cloned into binary vectors were transformed into Agrobacterium tumefaciens GV3101:pMP90 (Koncz and Schell, 1986). Cultures were mixed 1:1 with similarly prepared cultures containing the construct 35S:p19, to provide co-expression of the tomato bushy stunt virus p19 protein which suppressed gene silencing (Voinnet et al., 2003). Cultures were infiltrated into the lower surface of N. benthamiana leaves using a syringe (Wroblewski et al., 2005). Transformed leaf tissue was harvested 2–7 d later and used immediately for imaging, or was stored at –80 °C until required for protein

---

**Fig. 1.** Strep-tag vector design and BIN-STRP3–GFP-mediated expression in plants. (A) Overview of vector design, showing promoters, restriction enzyme sites useful for cloning, and left and right borders (LB, RB) of T-DNA region for the bacterial (pET-STRP3) and plant (BIN-STRP3, BIN-STRP3–GFP) expression vectors. (B) Coomassie Brilliant Blue-stained SDS–PAGE gel showing affinity purification of GFP–GST fusion proteins from N. benthamiana leaves infiltrated with BIN-STRP3–GFP constructs. Purifications are shown for the GST fusions as labelled and for a negative control. Lanes for each purification: C, crude protein extract (equivalent to 1 mg FW tissue); U, unbound fraction (equivalent to 1 mg FW tissue); P, purified fraction (equivalent to 100 mg FW tissue). Molecular weight markers (M) are shown with masses (in kDa) as shown. The identities of purified proteins were confirmed by peptide mass fingerprinting. The vertical bar marks the removal of superfluous gel lanes.
purification. For co-localization studies, similar *A. tumefaciens* cultures containing the constructs px-rk or pm-rk for labelling of peroxisomes and plasma membrane, respectively (Nelson et al., 2007), obtained from NASC (Scholl et al., 2000), were co-infiltrated. Both constructs expressed the red fluorescent protein mCherry, with either a C-terminal SKL motif added for peroxisomal targeting (px-rk) or as a fusion with the plasma membrane aquaporin AtPIP2a (pm-rk) for plasma membrane targeting. After 2–3 d, the transformed leaves were analysed by laser scanning confocal microscopy using a Zeiss LSM 510 Meta instrument with a plan Neofluar ×40/NA 1.3 oil immersion lens. Where required, tissue samples were counterstained for 5–20 min using 1 mg ml⁻¹ propidium iodide and then washed with water before visualization. GFP fusion fluorescence was imaged using excitation with a 488 nm laser, coupled with a 505–530 nm band pass filter, propidium iodide was imaged using a 543 nm laser with a 560-nm-long pass filter, and mCherry fusions were imaged using a 543 nm laser with a 650-nm-long pass filter. For each fluorophore combination, controls lacking each fluorophore in turn confirmed negligible channel crosstalk and auto-fluorescence (except from chloroplasts).

**GFP–GST purification**

GSTs were purified from frozen plant tissue after extraction using 4 v/v 100 mM TRIS-Cl pH 7.5 containing 150 mM NaCl, 1 mM EDTA, 10 mM sodium ascorbate, 50 μg ml⁻¹ avidin, 10 μg ml⁻¹ bovine pancreatic DNase I, 10 μg ml⁻¹ bovine pancreatic RNase A, and 5% w/v polyvinylpolypyrrolidone. After filtration through miracloth (Calbiochem, Nottingham, UK) and clarification by centrifugation (15 000 g, 20 min, 4°C), the Strep-GSTs were affinity purified as described for the bacterially expressed enzymes. Purified proteins were concentrated to ~100 μl by ultrafiltration through a 10 kDa cutoff membrane (2 ml Vivaspin; Sartorius Stedim UK Ltd, Epsom, UK).

**Results**

As a first step in studying the whole family of *Arabidopsis* GSTs, the coding sequence for each GST was amplified, with PCR products indicating that the respective gene was indeed transcribed. These clones were then expressed in *E. coli* and purified recombinant proteins assayed to survey for typical GST activities. The cloning and characterization of the GSTZs (Dixon and Edwards, 2006), DHARs, and GSTLs from *Arabidopsis* (Dixon et al., 2002a) has been reported previously. Their further analysis was restricted to their amplification and cloning into the new expression vectors to study their subsequent sub-cellular localization in *planta*. GSTZ1, GSTL1, GSTL2, GSTL3, and DHARs 1–3 were successfully amplified, all without any N-terminal targeting peptides. In the case of DHAR4 and GSTZ2, no amplification products were obtained and, since there is no evidence of transcripts for these genes, it was concluded that they were probably pseudogenes.

cDNAs encoding all 28 GSTUs and 12 of the GSTFs were amplified from *Arabidopsis*. Of the products obtained, GSTF8 was cloned as the splice variant without a plastidic targeting peptide, since this has been shown to be the major form in *vivo* (Thatcher et al., 2007). GSTF13 could not be amplified and no attempt was made to clone GSTF1 (Bartling et al., 1993), which is not encoded within the sequenced *Arabidopsis* genome (Wagner et al., 2002). The polypeptide sequences of the cloned GSTs matched those predicted from the TAIR 8 genome annotation (www.arabidopsis.org, April 2008 release), with two exceptions. GSTU17 had a frame-shift, but was identical to EMBL cDNA sequence AF288191, while GSTU21 had a 6 bp downstream shift in the start of the second exon. Thus, *gstf13* was the only possible pseudogene of those tested.

For the theta GSTs, three genes lie adjacent to one another on chromosome 5 (*GSTT1 = At5g51210; GSTT2 = At5g41240; GSTT3 = At5g41220*). GSTT2 and GSTT3 are unusual in being annotated as GSTs with a C-terminal Myb-like DNA-binding protein fusion (TAIR release 8). An examination of the C-terminal sequences of the three previously identified *Arabidopsis* GSTTs (Wagner et al., 2002) identified C-terminal SKM or SKI motifs, which would be anticipated to target these proteins to the peroxisome (Reumann, 2004). However, with the GSTT2 and GSTT3 genes the C-terminal Myb-like extensions masked the putative peroxisome targeting signal. PCR-based cloning successfully amplified the short forms of all three genes, with GSTT3 amplified by mis-priming with GSTT2 primers, which introduced an additional methionine at the N-terminus but did not alter the C-terminus. The resulting amplification products encoded the 28 kDa proteins GSTT1, GSTT2, and GSTT3, respectively (for sequences, see Supplementary data available at *JXB* online). A long splice variant of GSTT3 could also be amplified, encoding the 68 kDa GST–Myb fusion GSTT3L. Although not isolated, splice variants of GSTT2 encoding the GST–Myb fusion GSTT2L have also been reported in the DNA databases.

In total, *Arabidopsis* GSTs were found to be transcribed, with the respective coding sequences recovered by PCR amplification. A custom vector system was developed to provide inducible high level expression of GSTs in recombinant *E. coli*, using an N-terminal Strep-tag II (Skerra and Schmidt, 1999). This PET-STRP3 vector was engineered to minimize restriction enzyme incompatibility by utilizing flanking 6 bp cutters for routine use, with additional flanking 8 bp rare cutters introduced for problem sequences (Fig. 1A). Using this vector, all cloned GSTs were expressed in *E. coli*. Focusing on the phi, tau, and theta class enzymes, a total of 41 GSTs could be purified as soluble proteins using Strep-Tactin affinity chromatography (Figs S1 and S2 in Supplementary data available at *JXB* online). Only GSTF11 and GSTF12 were undetectable in the soluble fraction. In all cases, MS analysis of purified GSTs showed they had undergone cleavage of the N-terminal methionine, with the revealed alanine residue of the Strep-tag oligopeptide.
bacterial expression vector, allowing Novel binary vectors were constructed to complement the expression vectors (ASWSHPQFEK) undergoing partial N-acetylation (Table S2), an unusual post-translational modification in E. coli (Charbaut et al., 2002). Several GSTs were found to co-purify protein contaminants (Figs S1, S2A in Supplementary data available at JXB online), which were identified by peptide mass fingerprinting. Notably, GSTU12, GSTT1, and GSTT2 were found to co-purify ribosomal proteins, and GSTF5 co-purified the chaperone GroEL. A number of preparations contained a 70 kDa polypeptide identified as the E1 subunit of pyruvate dehydrogenase, which was therefore presumed to be a non-specific contaminant. For the theta class GSTs, ethidium bromide counterstaining of SDS–PAGE gels showed that GSTT1, GSTT2, and GSTT3L contained nucleic acids, while GSTT3 showed only minor contamination (Fig. S2B).

To test for functionality, recombinant phi, tau, and theta class GSTs were assayed for activity toward typical GST substrates. GSH-conjugating activity was assessed using the model xenobiotic substrate CDNB as well as BITC, the latter being a potential natural GST substrate in cruciferous plants derived from the degradation of glucosinolates (Gil and MacLeod, 1980). GSH-dependent peroxidase (GPOX) activity was routinely determined with cumene hydroperoxide (Edwards and Dixon, 2005). To examine the linkage between enzyme activity and relatedness, the results of these assays are represented with the GSTs clustered on the basis of their polypeptide sequence similarities (Fig. 2). Several of the 41 GSTFs and GSTUs could not be assayed. GSTF11, GSTF12, and GSTF13 could not be cloned, or only expressed as insoluble proteins, while GSTF4, GSTF10, and GSTU15 were only obtained in very low yields. Of the remaining 35 recombinant proteins, 32 showed GSH-dependent enzyme activity, with only GSTF5, GSTF14, and GSTU11 being inactive toward the three substrates tested. Closely related GSTs showed a similar spectrum of enzyme activities, although absolute values for specific activities between enzymes varied. The conjugation of CDNB was the most commonly observed enzyme activity, though it varied over three orders of magnitude within the superfamilly (Fig. 2). BITC was a more discriminating GST substrate, being acted on by most tau class enzymes but rarely by the GSTFs. The majority of the GSTs also had some GPOX activity, with GSTU25 having particularly high activity. When the three GSTTs were assayed, these enzymes were shown to be highly active as GPOXs. To explore these activities in greater detail, two fatty acid hydroperoxides were prepared and found to be excellent GPOX substrates. This was in contrast to GSTU25 which, although showing high GPOX activity towards cumene hydroperoxide, showed little activity toward the fatty acid hydroperoxides that are likely substrates of GPOXs in planta (Bartling et al., 1993).

**Generation and testing of custom plant expression vectors**

Novel binary vectors were constructed to complement the bacterial expression vector, allowing Agrobacterium-mediated constitutive expression of Strep-tagged proteins in plants (Fig. 1A). BIN-STRP3 allowed plant-based expression of the same tagged protein as given by pET-STRP3 in bacteria, and BIN-STRP3-GFP provided an additional N-terminal GFP tag for visualization studies while retaining the Strep-tag for easy gel-based detection and/or purification if required. For example, GSTs could be recovered after transient expression in N. benthamiana and shown to be intact (Fig. 1B) and fluorescent, giving confidence that in vivo fluorescence data was not compromised by proteolytic cleavage of the GFP-GST fusion. Both vectors facilitated sub-cloning from the corresponding pET-STRP3 expression vector by incorporating compatible rare-cutting restriction sites.

**GSTTs as Arabidopsis GST family members with C-terminal targeting motifs**

Having established a tractable expression system which allowed for both the functional expression of GSTs in E. coli and the definition of their sub-cellular localization in planta it was then of interest to determine whether the information derived from these two experimental systems could provide insight into the roles of GST family members. Examining the functional activities of GSTs and the potential for differential sub-cellular localization, it was apparent that the GSTTs represented a class of proteins which had conserved enzyme activities and the potential for targeting to different subcellular organelles. Thus, GSTs T1, T2, and T3 contained C-terminal domains corresponding to likely peroxisomal targeting motifs (C-terminal SKI, SKM, and SKM motifs, respectively), whereas GSTT2L and GSTT3L contained polypeptide extensions resembling Myb transcription factors which were more likely to direct the fusion proteins to the nucleus. To investigate their sub-cellular localization, GFP fusions of GSTT1, GSTT2, GSTT3 (short form), and GSTT3L (transcription factor fusion) were transiently expressed in Nicotiana benthamiana, and their localization then monitored by laser scanning confocal microscopy. GFP–GSTT1 (Fig. 3A), GFP–GSTT2 (Fig. 3B), and GFP–GSTT3 (Fig. 3C) were each confined to small organelles which, on the basis of their mobility, number, and size, together with the putative C-terminal targeting sequence of these GSTs, were assumed to be peroxisomes. This was confirmed (Fig. S3 in Supplementary data available at JXB online) by co-localization with the peroxisomal marker px-rk (Nelson et al., 2007). By contrast, GFP–GSTT3L localized solely to the nucleus (Fig. 3D, Fig. S4), giving a punctate expression pattern and showing exclusion from the nucleolus.

**Localization of other plant GSTs**

In addition to the theta class enzymes, a range of phi (F2, F6, F8, F9, and F12) and tau (U2, U7, U9, U11, U12, U19, and U28) GSTs were also tested for sub-cellular localization using GFP tagging. In each case, these GSTs lacked characterized C-terminal targeting motifs, and the N-terminal transit peptide of GSTF8 was removed prior to
Fig. 2. Phylogenetic tree and enzyme activities of bacterially expressed recombinant Arabidopsis phi (F) and tau (U) class GSTs assayed with 1-chloro-2,4-dinitrobenzene (CDNB), benzyl isothiocyanate (BITC), and cumene hydroperoxide (cumene-OOH). Also shown are theta (T) class GSTs (with GSTU25 for comparison), assayed for GPOX activity with the additional substrates \((13S,9Z,11E)-13\)-hydroperoxy-9,11-octadecadienoic acid (18:2-OOH) and \((13S,9Z,11E,15Z)-13\)-hydroperoxy-9,11,15-octadecatrienoic acid (18:3-OOH).

Each GST name is suffixed with a key describing the associated analysis: C, successfully cloned; A, purified recombinant protein assayed; P, recombinant protein purified but not assayed; I, recombinant protein totally insoluble; Y, recombinant protein not available in...
testing. GFP–GST fusions of all the GSTFs and GSTs U2, U7, U9, U11, U19, and U28 all localized to the cytosol. In most cases GFP fusions were also found at low levels in the nucleus, presumably due to passive diffusion, as illustrated by GFP–GSTF2 and GFP–GSTU19 (Fig. 4A, B), with both of these GSTs being described as abundantly expressed proteins in proteomic studies (Sappl et al., 2004; Smith et al., 2004). GFP fusions of GSTs U7, U9, U17, and U28 were practically absent from the nuclei (Fig. 4C). By contrast, GFP–GSTU12 localized entirely to the nucleus (Fig. 4D), confirmed by counterstaining with propidium iodide. As compared with related GSTs, examination of the GSTU12 polypeptide sequence revealed an N-terminal extension of 25 amino acid residues containing a putative nuclear localization signal (KKRKK), although it is beyond the scope of this work to test whether this signal is necessary and sufficient for nuclear import. Sub-cellular localization studies were then extended to other GSTs containing unusual sequence motifs. The lambda protein GSTL2 (without its N-terminal signal peptide) was selected since it has a potential peroxisomal targeting sequence (ARL) at its C-terminus. GFP–GSTL2 was found to localize to both the cytosol and peroxisomes (Fig. 4E), confirming the presence of a peroxisomal targeting signal. The unusual GST superfamily member TCHQD was also studied. GFP–TCHQD localized to the plasma membrane (Fig. 4F), co-localizing (Fig. S3) with the diagnostic marker pm-rk (Nelson et al., 2007). Table 1 summarizes the distribution of non-cytosolic Arabidopsis GSTs.

**Discussion**

Differential targeting of the Arabidopsis GST superfamily

The Arabidopsis genome contains 54 identified members of the GST superfamily. Of these 51 are transcribed and at least 41 encode functional GSH-dependent enzymes when sufficient yield/purity for assay; S, subcellular localization studied; a dash = analysis not performed. ND, No activity detected; NA, not assayed. † GST concentration overestimated due to co-purifying chromophores, so activity is underestimated. ‡ Variation in the replicates from the mean values quoted was <10% in all cases.

---

**Fig. 3.** Laser scanning confocal micrographs showing intracellular localization of N-terminally GFP-tagged GSTs transiently expressed in N. benthamiana (green channel), with counterstaining with propidium iodide (magenta channel). GSTs shown are GSTT1 (A), GSTT2 (B), GSTT3 (C), and GSTT3L (D). Scale bars represent 100 μM.
assayed as GSTs or GPOXs (GSTFs, GSTUs, GSTTs; Fig. 2) thioltransferases or reductases (DHARs and GSTLs) (Dixon et al., 2002) or isomerases (GSTZs) (Dixon and Edwards, 2006). Among the GSTFs, GSTUs, and GSTTs it was clear from the enzyme screen (Fig. 2) that family members of related sequence have a broadly similar spectrum of activities as GSTs and GPOXs when assayed with model substrates. While it is conceivable that each enzyme has highly selective activities directed toward a well-defined group of natural substrates, the present results would suggest that it is more likely that there is a substantial overlap of activities and functional redundancy within the superfamily. Expression of particular GST isoenzymes in different plant organs and tissues and in response to different stimuli (e.g. constitutive Vs stress-inducible) have been proposed as reasons to explain this apparent redundancy (Edwards et al., 2000; Frova, 2003). The results presented here suggest that there is an additional reason for apparent functional redundancy within the family; namely, multiple GSTs with similar activities may be required to

Fig. 4. Laser scanning confocal micrographs showing intracellular localization of N-terminally GFP-tagged GSTs transiently expressed in *N. benthamiana* (green channel), with some panels counterstained with propidium iodide (magenta channel). GSTs shown are GSTF2 (A), GSTU19 (B), GSTU28 (C; nuclei and cell walls counterstained), GSTU12 (D; cell walls counterstained); GSTL2 (E), and TCHQD (F). N, nucleus. Scale bars represent 100 μM.
perform similar functions in different subcellular compartments. In this respect, the GSTTs serve as an excellent example.

The three GSTTs in Arabidopsis have clearly arisen through gene duplication (Cannon et al., 2004), with the enzymes showing conservation in both sequence and enzyme activity. Thus, all three GSTTs were highly active as GPOXs when assayed with organic hydroperoxides. The GFP-fusion protein expression studies in planta showed that, whereas GSTT1 and the short splice variants of GSTT2 and GSTT3 would express their GPOX activity in the peroxisomes, GSTT3L, with its C-terminal Myb-like extension, was directed to the nucleus where it accumulated. This distribution closely resembles that of nuclear speckles, thought to be storage sites for pre-mRNA splicing-related proteins and also other proteins (Shaw and Brown, 2004). Although theta class GSTs have been described in mammals (Jowsey et al., 2001), insects (Ding et al., 2003), and fungi (Bryant et al., 2006), these enzymes have not attracted the same level of attention as other GST classes. Mammalian GSTTs characteristically exhibit dehalogenase activity, are highly active GPOXs, and show low affinity binding of GSH and little transferase activity toward CDNB (Landi, 2000). Similarly, plant GSTTs are also highly active as GPOXs, while showing limited GST activities (Dixon et al., 1999). However, whereas GSTTs in animals are stress responsive proteins (Landi, 2000), the respective genes in Arabidopsis were unresponsive to chemical and abiotic stress treatments (Dixon et al., 2002a; Wagner et al., 2002). In Arabidopsis, the GSTTs were expressly targeted to either the peroxisomes or the nucleus. The expression of GPOX-active GSTTs in peroxisomes presumably reflects the oxidative environment within this compartment which leads to the generation of phytotoxic fatty acid hydroperoxides, which are substrates of these enzymes (Fig. 2). Thus, the peroxisomal localization of GSTT1, GSTT2, and GSTT3 would most likely be associated with a protective function in maintaining the integrity of the peroxisomal membranes. Theta class GSTs in other plants also have C-terminal motifs that are likely to result in peroxisomal targeting, including GSTT proteins from rice (GenBank AAK98534), soybean (GenBank AAG34813), and Medicago truncatula (GenBank ABE92134). However, GSTTs in animals and fungi have no such obvious targeting sequences and have not been reported to be associated with peroxisomes. This is in contrast to other classes of GSTs, such as the kappa class proteins in mammals (Morel et al., 2004), and omega-class GSTs in Saccharomyces cerevisiae (Barreto et al., 2006), which are known to be targeted to the peroxisomes. This suggests that, despite their sequence and apparent functional conservation across the phyla, GSTTs in plants have been recruited to perform unique roles in peroxisomal metabolism.

The accumulation of GSTTs in the nucleus is less easy to rationalize. The localization of GSTT3L to this organelle is clearly determined by the presence of the unusual C-terminal extension which resembles a Myb DNA-binding protein. GSTTs have previously been reported in the nucleus in animal cells. Thus, in hepatocytes, GSTTs are expressed evenly in both the cytosol and the nucleus (Liteplo et al., 1998; Sherratt et al., 1998), with one report of preferential nuclear expression in specialized murine tissues (Quondamatteo et al., 1998). This suggests that GSTTs may have originally evolved functions which are required both in the cytosol and in the nucleus. In plants, the latent ability of GSTTs to bind to nucleic acids may offer a clue as to the functional significance of the selective localization of GSTT isoforms. When affinity purified from E. coli, both GSTT1 and GSTT2 preparations were heavily contaminated with nucleic acid and ribosomal polypeptide components (Fig. S2 in Supplementary data available at JXB online). It is likely that the RNA-binding properties of these proteins are due to their basic nature (calculated pI of pH 9.50 and pH 9.32 for GSTT1 and GSTT2, respectively) giving rise to non-specific interactions with nucleic acids. The short splice variant of GSTT3 did not bind nucleic acids when expressed in bacteria, presumably due to its less basic nature (pI 8.89), whereas GSTT3L with its C-terminal extension showed obvious nucleic acid contamination. These results clearly demonstrate that GSTT3L binds nucleic acids as a consequence of its transcription factor-like domain (although this is unlikely to be as clear-cut for GSTT2L since GSTT2 retains RNA binding). This may suggest that for GSTT3L to exhibit a selective function in the nucleus there has been selective pressure to decrease the non-specific binding of nucleic acid to the GSTT domain, as achieved through the replacement of basic residues. Instead, the ability to selectively bind to nucleic acids has been introduced through the fusion with the Myb-like domain. While the function of GSTT3L is yet to be determined, it is speculated that these proteins either play a redox-sensitive role in controlling transcription, or are involved in specialized nucleic acid repair. With respect to the latter hypothesis, it is known that mammalian GSTTs reduce DNA hydroperoxides formed in vitro through X-ray radiation damage (Tan et al., 1988). It is therefore possible that

### Table 1. Summary of GSTs found to accumulate outside the cytosol, with proposed targeting signal

<table>
<thead>
<tr>
<th>GST (gene)</th>
<th>Location</th>
<th>Likely signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSTT1 (At5g41210)</td>
<td>Peroxisome</td>
<td>C-term SKI</td>
</tr>
<tr>
<td>GSTT2 (At5g41240)</td>
<td>Peroxisome</td>
<td>C-term SKM</td>
</tr>
<tr>
<td>GSTT3 (At5g41220)</td>
<td>Peroxisome</td>
<td>C-term SKM</td>
</tr>
<tr>
<td>GSTT3L (At5g41220)</td>
<td>Nucleus (speckles?)</td>
<td>C-term Myb-like domain</td>
</tr>
<tr>
<td>GSTU12 (At5g41210)</td>
<td>Nucleus</td>
<td>Nuclear localization signal</td>
</tr>
<tr>
<td>GSTT2 (At3g55040)</td>
<td>Chloroplast*</td>
<td>N-term targeting peptide*</td>
</tr>
<tr>
<td>GSTT1 (At5g41210)</td>
<td>Peroxisome</td>
<td>C-term AR</td>
</tr>
<tr>
<td>DHAR3 (At5g16710)</td>
<td>Chloroplast*</td>
<td>N-term targeting peptide*</td>
</tr>
<tr>
<td>TCHQD (At1g77290)</td>
<td>Plasma membrane</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

* Previously identified targeting peptides, giving an assumed chloroplastic location (Dix e et al., 2002a).
targeted by the C-terminal Myb-like domain, the function of GSTT3L is to repair actively transcribed DNA segments which are subject to elevated levels of oxidative damage. Examination of other plant GSTT sequences failed to identify similar fusions in other species, so the *Arabidopsis* GSTT fusions may have a very specialized role that most plants either do not require or have an alternative system in place.

**Localization of other GSTs**

Most of the GSTs tested accumulated in the cytosol, as expected based on their lack of any obvious targeting sequences. This does suggest that members of this cytosolic group of GSTs which are reported to associate with mitochondrial, chloroplastic, and vacuolar proteomes do so as a result of contamination of these organelles with these relatively abundant and hydrophobic proteins (Carter et al., 2004; Heazlewood et al., 2004; Zybaïlov et al., 2008). Similar conclusions have been made regarding GSTA1-1 in rats, which though known to be expressed in the cytosol accumulate to high levels in the outer nuclear envelope (Stella et al., 2007). While this may explain the presence of relatively abundant cytosolic GSTs as proteome components of organelles in *Arabidopsis*, it is also possible that family members which are present in lower abundance in the cell are carried into organelles through interactions with explicitly targeted proteins. Intriguingly, some, but not all of the predominantly cytosolic GFP–GST fusions were also present at low levels in the nucleus, consistent with their import by passive diffusion. It is possible that the GFP–GSTUs which solely localized to the cytosol did so as a consequence of the size of the fusion protein precluding nuclear import, with the smaller native GSTs more likely to enter the pores. Certainly the GFP–GST fusions appeared to be processed as fully functional polypeptides which, in view of their stability and yield, were most likely expressed as homodimers (Fig. 1). In any event, the ability of several of the GSTS to move between the nucleus and cytosol should be taken into account when considering their potential ligands and roles in inter-compartmental transport and signalling.

The present studies have shown that several GSTs previously assumed to be cytosolic are targeted elsewhere. GSTU12 was clearly nuclear-localized and, to date, the only other known tau class GST known to accumulate in this compartment is ParA from tobacco (Takahashi et al., 1995). Unlike GSTU12, ParA has no obvious localization signal and the functional significance of this nuclear targeting is unknown. The roles for GSTU12 being directed to the nucleus must be distinct from that inferred for GSTT3L. Unlike GSTT3, GSTU12 has no activity as a GPOX and therefore would be unlikely to be involved in reducing either DNA- or lipid-derived hydroperoxides. This GST may instead function to regulate the activity of transcriptional regulators within the nucleus, with a precedent being the interaction of human GSTP1-1 with c-Jun N-terminal kinase (Wang et al., 2001).

The other intriguing subcellular localization of the *Arabidopsis* GST family members was seen in the selective accumulation of TCHQD at the plasma membrane. Previously, only GSTF2 has been reported to associate with the plasma membrane in *Arabidopsis* (Smith et al., 2003), though in the present study this localization was not observed (Fig. 4A). This difference has not been investigated further but could be due to the lack of appropriate hydrophobic GSTF2 ligands in tobacco leaves that promote membrane association in *Arabidopsis* roots. The functional significance of membrane-localized TCHQD is unclear, not least because the respective recombinant enzyme had no activity as a GST or GPOX. This GST family member showed similarity (25% identity) to the TCHQD identified in *Sphingobium chlorophenolicum*, an enzyme which catalyses the reductive dehalogenation of TCHQ and trichlorohydroquinone, which are key steps in the degradation of the pesticide pentachlorophenol (Warner et al., 2008). The endogenous functions of this enzyme in bacteria are unknown, but its evolutionary origins and enzymic mechanism are related to the GSTZ isomerases (Anandarajah et al., 2000). In bacteria, TCHQDs are soluble proteins, and it is most likely that the membrane-associated orthologue in *Arabidopsis* has been recruited for quite different functions. Significantly, the knowledge of the subcellular localization of this GST and the development of vectors allowing for the expression of Strep-tagged fusion proteins in planta now provide a set of unique tools to probe for the binding partners and ligands of this protein and other family members.

**Supplementary data**

sequences.txt FASTA-format nucleotide and polypeptide sequences of GST coding sequences cloned into pET-STRP3.

data.txt FASTA-format nucleotide sequences of vectors pET-STRP3, BIN-STRP3, and BIN-STRP3–GFP.

table S1. Primers used for PCR amplification of GSTs.

table S2. Theoretical and measured masses for purified Strep-tagged GSTs expressed in *E. coli* and plants.

Figure S1. Coomassie Brilliant Blue-stained SDS–PAGE gels of Strep-tagged *Arabidopsis* phi and tau class GSTs purified from over-expressing *E. coli*.

Figure S2. Coomassie Brilliant Blue-stained SDS–PAGE gels of Strep-tagged *Arabidopsis* theta class GSTs purified from over-expressing *E. coli*, showing RNA binding.

Figure S3. Laser scanning confocal micrographs showing colocalization of GFP–GST fusions with organelar markers.

Figure S4. Laser scanning confocal micrographs showing localization of GFP–GSTT3L to the nucleus, counterstained with propidium iodide.

**Acknowledgements**

This work was supported by the Biotechnology and Biological Sciences Research Council through grant...
BBC51227X1 and a research development fellowship awarded to RE. The authors acknowledge Dr Mike Deeks for his assistance with confocal microscopy.

References


Liteplo RG, Long GW, Meek ME. 1998. Relevance of carcinogenicity bioassays in mice in assessing potential health risks associated with exposure to methylene chloride. Human & Experimental Toxicology 17, 84.


