Differential Gene Expression and Subcellular Targeting of Arabidopsis Glutathione S-Transferase F8 Is Achieved through Alternative Transcription Start Sites*

Louise F. Thatcher,† Chris Carrie,‡ Carol R. Andersson,¶ Krishnapillai Sivasithamparam,§ James Whelan,* and Karam B. Singh†

From the †Commonwealth Scientific and Industrial Research Organisation, Plant Industry, Private Bag 5, Wembley, Western Australia 6913, Australia, ‡Soil Science and Plant Nutrition, School of Earth and Geographical Sciences and §Australian Research Council Centre of Excellence in Plant Energy Biology, M316, University of Western Australia, Crawley, Western Australia 6009, Australia

Glutathione S-transferases (GSTs) play major roles in the protection of plants from biotic and abiotic stresses through the detoxification of xenobiotics and toxic endogenous products. This report describes additional complexity in the regulation of the well characterized stress-responsive Arabidopsis thaliana GSTF8 promoter. This complexity results from the use of multiple transcription start sites (TSS) to give rise to alternate GSTF8 transcripts with the potential to produce two in-frame proteins differing only in their N-terminal sequence. In addition to the originally mapped TSS (Chen, W., Chao, G., and Singh, K. B. (1996) Plant J. 10, 955–966), a further nine TSS have been identified, with the majority clustered into a distinct group. The most 3′ TSS gives rise to the major message (GSTF8-S) and the shorter form of the protein, whereas those originating from upstream TSS (GSTF8-L) are more weakly expressed and encode for the larger form of the protein. Differential tissue-specific and stress-responsive expression patterns were observed (e.g. GSTF8-L is more highly expressed in leaves compared with roots, whereas GSTF8-S expression has the opposite pattern and is much more stress-responsive). Analysis of GSTF8-L and GSTF8-S proteins demonstrated that GSTF8-L is solely targeted to plastids, whereas GSTF8-S is cytoplasmic. In silico analysis revealed potential conservation of GSTF8-S across a wide range of plants; in contrast, conservation of GSTF8-L was confined to the Brassicaceae. These studies demonstrate that alternate TSS of the GSTF8 promoter are used to confer differential tissue-specific and stress-responsive expression patterns as well as to target the same protein to two different subcellular localizations.

Glutathione S-transferases (GSTs) are ubiquitous enzymes in animals and plants and protect tissues against oxidative damage or from toxic products produced during xenobiotic metabolism (2, 3). GSTs typically function by catalyzing the conjugation of GSH to a variety of electrophilic, hydrophobic substrates of endogenous or exogenous origin to render the substrate less toxic or to produce a more water-soluble conjugate (2, 4, 5). This detoxification process is part of a larger sequential three-phase detoxification process that also involves cytochrome P450s and ATP-binding cassette transporters (5–9). The phytotoxic compounds detoxified by GSTs may be direct products of microbial or animal attack, or they may be endogenous by-products of a defense or stress response (6). In addition to these compounds, GSTs can also detoxify exogenous chemicals, such as herbicides, in what is now a defining role of plant GSTs (5, 7, 8). The expression of GSTs can also be induced by a range of biotic and abiotic stresses, including herbicides, heavy metals, pathogen attack, and phytohormones (2, 4, 5, 7, 10).

In Arabidopsis thaliana, the GST superfamily consists of 53 genes, representing six classes (2, 11). A focus of our group’s research is a particular Arabidopsis Phi class GST termed GSTF8, which is used as a marker for early stress/defense responses. GSTF8 expression can be induced by a range of biotic and abiotic stresses, including H$_2$O$_2$, salicylic acid (SA), auxin, herbicides, and microbial infection (1, 10, 12–16). Although the endogenous function of GSTF8 is yet to be determined, it does exhibit transferase activity toward the model substrate 1-chloro-2,4-dinitrobenzene and glutathione-peroxidase activity toward cumene hydroperoxide and linoleic acid hydroperoxide, with the latter being the preferred substrate (10).

Until recently, the GSTF8 gene was thought to be transcribed from one major transcription start site (TSS). However, we noticed the potential for alternate TSS within the GSTF8 promoter through the observation of an expressed sequence tag (EST) with sequence 5′ to the mapped GSTF8 TSS. In animal models, the use of alternate TSS is not uncommon (17, 18); however, in plants this has been primarily observed for genes encoded on mitochondrial or chloroplast genomes (18–20). The use of multiple TSS can add flexibility to the ways in which a gene’s expression is regulated and can potentially affect translational efficiency; provide tissue, developmental, or signal specificity; or lead to the production of protein isoforms that differ in their N termini (17, 21, 22).

* This work was supported in part by the Grains Research and Development Corporation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Recipient of a Grains Research and Development Corporation Ph.D. scholarship and University of Western Australia specific scholarship.

‡ Present address: Food Standards Australia New Zealand, P.O. Box 7186, Canberra BC, Australian Capital Territory 2610, Australia.

§ To whom correspondence should be addressed. Tel.: 61-8-9333-6320; Fax: 61-8-9387-8991; E-mail: karam.singh@csiro.au.

¶ The abbreviations used are: GST, glutathione S-transferase; SA, salicylic acid; TSS, transcription start site(s); EST, expressed sequence tag; RLM, RNA ligase-mediated; RACE, rapid amplification of cDNA ends; ORF, open reading frame; AOX, alternative oxidase; RFP, red fluorescent protein.
Differential Expression and Targeting of GSTF8

In this report, we examine the potential for alternate GSTF8 TSS and their effects on GSTF8 expression or protein production. We demonstrate that there are multiple GSTF8 TSS upstream of the originally mapped TSS, with the majority of these upstream sites clustered into a distinct group. Furthermore, these TSS are differentially regulated; this occurs both at a quantitative level and in terms of tissue-specific and stress-responsive expression patterns. The most 3′ TSS gives rise to the predominant message and encodes for the shorter form of the protein, GSTF8-S, whereas those originating from upstream TSS encode for the larger form of the protein, GSTF8-L, that only differs from the short form by the presence of an N-terminal extension. Although in silico analysis of the GSTF8-L N-terminal region suggested that it may encode a plastidic or mitochondrial targeting signal (this paper) (10), we demonstrate that GSTF8-L is targeted to the plastid stroma, whereas GSTF8-S is cytoplasmic. A search for homologues of GSTF8 in other plant species revealed potential conservation of GSTF8-L only among the Brassicaceae, whereas GSTF8-S was conserved in a wide range of plants. We discuss the implications of our findings for redox regulation and detoxification processes in plants.

EXPERIMENTAL PROCEDURES

Plant Material—All experiments were conducted with the A. thaliana Columbia-0 (Col-0) ecotype. Luciferase reporter constructs were created by fusing GSTF8 promoter fragments with a promoterless luciferase+ (LUC+; Promega) coding sequence in the binary vector (pYRO), which contains BamHI and HindIII restriction sites for promoter cloning. pYRO is based on the binary vector pZIP21 with the HindIII site removed, and the LUC+ coding sequence cloned between the BamHI and KpnI sites (the Omega translational enhancer, Luc+ coding fragment, and EP3′ polyadenylation sequence were removed from pAtMDomega as a BamHI/KpnI fragment; pAtMDomega was kindly donated by Andrew Millar). Promoter fragments were amplified by PCR from genomic DNA using specific primer pairs with added restriction sites (see Table 1). The −814/+10 fragment was amplified using primers GSTF8 −814BamHI-F and GSTF8 +10HindIII-R and cloned into pYRO to create −800GSTF8-S::luc+. The −814/−83 GSTF8 promoter was inserted into pYRO using primers GSTF8 −814BamHI-F and GSTF8 −83HindIII-R to create −800GSTF8-L::luc+. All inserts were sequenced and transformed into Col-0, and homozygous T3 lines with one insert were screened for and used in subsequent experiments.

Green fluorescent protein (GFP) vectors for protein localization studies were constructed from the binary vector pBINFG-P-AOX (gift from Orinda Chew, University of Western Australia). pBINFG-P-AOX is based on pGEMGFP-AXX, which contains m-gfp5 (23). Sequences corresponding to either full-length, the first 100 amino acids of, or the first 50 amino acids of GSTF8-L or GSTF8-S were amplified with 5′ BamHI and 3′ EcoRI sites by PCR using cDNA from 14-day-old Col-0 leaf tissue with the respective primers as noted in Table 1. Note that the full-length constructs do not include the GSTF8 stop codon. The soybean alternative oxidase (AOX) signal peptide sequence was removed from pBINFG-P-AOX by BamHI/EcoRI restriction digestion, and replaced with the GSTF8 PCR products cloned upstream and in-frame of the mgfp5 gene. To use the empty pBINFG vector as a 355 GFP control, an ATG linker was cloned between the 355 promoter and the mgfp5 gene. The oligonucleotides 5′-GATCATGG-3′ and 5′-AATCCATG-3′ with BamHI and EcoRI sites, respectively, were annealed together, phosphorylated, and ligated into BamHI/EcoRI-digested pBINFG-P-AOX with the AOX signal sequence removed. Constructs were sequenced and transformed into Col-0. T2 lines were selected and used for subsequent experiments.

For mutagenesis of the upstream GSTF8-L ATG codon in the GFP vector, the ATG sequence was replaced with TTG. Analysis of full-length 35S GSTF8-L::TTG was performed with 35S GSTF8-L and 35S GSTF8-S by transformation of these constructs into Arabidopsis suspension cells as described previously (24), using small subunit of Rubisco RFP as a control for plastid localization.

Plant Growth Conditions—Seeds were surface-sterilized, stratified, and plated onto 10-cm square agar plates containing 1× Murashige and Skoog salts as described previously (13). Plates for the luciferase assay were supplemented with 50 μM luciferin (Biosynth AG) added after autoclaving the medium. The plates were incubated vertically under a 16-h light/8-h dark cycle at 22°C.

Plant Treatments—Salicylic acid was from Aldrich, and the herbicide Dicamba was obtained from the Western Australian Herbicide Resistance Initiative (Perth, Australia). Treatment plates were performed by diluting concentrated stocks in water, and 10 ml of the resulting dilution (1 mM SA, 7 mM Dicamba) was poured onto the Murashige and Skoog plates (100 cm2) containing the growing plants. After 40 min of incubation at room temperature, the excess liquid was discarded. For seedlings to be imaged using an in vivo bioluminescence assay, plates were transferred to the camera to be photographed such that after acquiring the bioluminescence image, 1 h had elapsed.

Bioluminescence and Luciferase Assays—Bioluminescence was captured from seedlings grown on Murashige and Skoog plates supplemented with luciferin by imaging in an EG & G Berthold molecular light imager as previously described (13, 15). Results are presented with the bioluminescence (pseudocolored blue) image superimposed onto the fluorescence (white) image or graphically as light units/seedling, as determined using Winlight32 software (version 2.7; Berthold Technologies). The biochemical luciferase assay was performed as described by Chen and Singh (12).

Fluorescence Microscopy—Seven-day-old T2 GSTF8-GFP seedlings grown upright on Murashige and Skoog plates with selection were suspended in H2O and visualized using confocal analysis performed with a TCS SP2 AOBS Multiphoton confocal microscope (Leica, Bensheim, Germany) with the following filter settings: for GFP, excitation 501 nm and emission 550 nm; for chlorophyll autofluorescence (TRITC), excitation 615 nm and emission 685 nm. A HCX PL APO CS 40.0 × 1.25 OIL PH objective was used with digital zoom. Images were edited with Leica confocal software, Adobe Photoshop 6.0, and/or Confo-
Differential Expression and Targeting of GSTF8

| Primer sequence | Sequence |
|-----------------|----------|
| 5‘-RLM-RACE     |          |
| GSTF8-T-specific outer | GGGGAGATCTATGATATATATAGAAGGG |
| GSTF8-L-specific outer | GAAAGATCTATGATATATATAGAAGGG |
| GSTF8-T-specific inner | CCTAGAGCTTGGATCTTCTCAGAAAGAG |
| GSTF8-L-specific inner | GAGCAAGCTTTATGCTCCATTTTCGAGATC |
| Luciferase reporter gene constructs |          |
| GSTF8-814BamHI-F | GCCGGATCCATGACATGTATATAGC |
| GSTF8+10HindIII-R | GCCGAACCTTGGATCTTCTCAGAAAGAG |
| GSTF8-83HindIII-R | GCTAGACCTTGGATCTTCTCAGAAAGAG |
| GFP fusion constructs |          |
| GSTF8-L-50aa-F    | GAAAGGAATGACTACATGAGAACATTTCAAGCT |
| GSTF8-L-50aa-R    | CGTGGATCTTGGATCTTCTCAGAAAGAG |
| GSTF8-L-100aa-F   | GAAAGGAATGACTACATGAGAACATTTCAAGCT |
| GSTF8-L-100aa-R   | CGTGGATCTTGGATCTTCTCAGAAAGAG |
| GSTF8-L-full-length-F | GAAAGGAATGACTACATGAGAACATTTCAAGCT |
| GSTF8-L-full-length-R | CGTGGATCTTGGATCTTCTCAGAAAGAG |
| GSTF8-S-full-length-F | GAAAGGAATGACTACATGAGAACATTTCAAGCT |
| GSTF8-S-full-length-R | CGTGGATCTTGGATCTTCTCAGAAAGAG |
| Real time RT-PCR   |          |
| GSTF8-T-F         | CCCCCTCAGATAGAAC |
| GSTF8-T-R         | GAAAGGAATGACTACATGAGAACATTTCAAGCT |
| GSTF8-L-F         | GACTCTGCTTGGATCTTCAGAA |
| GSTF8-L-R         | ATCTGCTTGGATCTTCTCAGAA |
| EF1α-F           | ATCTGCTTGGATCTTCTCAGAA |
| EF1α-R           | ATCTGCTTGGATCTTCTCAGAA |

For analysis of GFP in suspension cells, fluorescence patterns were visualized 24 h after transformation using an Olympus BX61 fluorescence microscope and imaged using Cell imaging software.

**In Vitro Chloroplast Imports**—In vitro chloroplast import studies were carried out as previously described (23).

*Mapping Alternate GSTF8 Transcription Start Sites*—5‘-RNA ligase-mediated rapid amplification of cDNA ends (5‘-RLM-RACE) was carried out with the Ambion FirstChoice RLM-RACE kit (Ambion) according to the manufacturer’s instructions with the following modifications. Total RNA was extracted using a Gentra Purescript RNA isolation kit (Gentra Systems) and checked for DNA contamination as previously described (16), followed by poly(A) mRNA isolation using an Ambion MicroPoly(A) Purist kit. cDNA synthesis was either carried out according to the RLM-RACE kit or modified for use with avian myeloblastosis virus reverse transcriptase (Promega) or Moloney murine leukemia virus reverse transcriptase (Promega) with an intermediate partial heat denaturation step during two rounds of reverse transcription. The modified protocol consisted of 2 μl of ligated RNA, 1 mm random decamers (RLM-RACE kit), or GSTF8-L-specific oligonucleotide (5‘-GAATGAGGAAAGTATGTG-3′), 0.8 mm dNTPs (Promega), and 10 units of RNasin (Promega) in a total volume of 20 μl with nuclease-free H2O. cDNA synthesis was performed at 42 °C using 10 units of avian myeloblastosis virus reverse transcriptase (Promega) or Moloney murine leukemia virus reverse transcriptase (Promega) with 1 × appropriate reverse transcription buffer according to Huttemann (25). Nested PCRs were carried out according to the RLM-RACE kit using Platinum TaqDNA polymerase (Invitrogen) (annealing temperature 60 °C; 1.5 mm final MgCl2 concentration). The outer PCR consisted of 5‘-RACE outer primer with GSTF8-T- or GSTF8-L-specific outer primers (Table 1). The inner PCR consisted of 1 μl of the outer PCR as template with 5‘-RACE inner primer and GSTF8-T- or GSTF8-L-specific inner primers (Table 1). To facilitate cloning of PCR products, the RACE inner primer has a 5‘ BamHI site, whereas the GSTF8-specific inner primers have 5‘ HindIII sites. PCR products were cloned into pBSK11+ (Stratagene), then more than one clone for each insert sequenced. The specificity of the 5‘-RLM-RACE products was confirmed by conducting the controls recommended by the kit manufacturer.

**Real Time Reverse Transcription-PCR Analysis**—Leaf samples for developmental analysis were harvested from Col-0 seedlings at the indicated times. Two biological replicates were taken for each sample, consisting of tissue pooled from five individual plants. For a comparison between leaf and root tissue, seedlings were grown upright on Murashige and Skoog plates, with three biological replicates taken for each sample consisting of tissue pooled from 10 individual seedlings. Hypocotyls were included in the root samples. Extraction of total RNA, cDNA synthesis, and quantitative reverse transcription-PCR were performed as previously described (16), with the exception that samples were normalized to the EF1α (elongation factor 1α) gene and that the thermocycling reactions were carried out on MyiQ with MyiQ System Software 1.1.410 (Bio-Rad). Primer efficiencies were calculated using LinRegPCR software 7.5 (26). The gene-specific primer pairs used for real time PCR are noted in Table 1.

**Data Base Searches and Sequence Analysis**—Putative AtGSTF8 homologues were obtained through Homologene and Unigene at NCBI (available on the World Wide Web), and by BLASTing the AtGSTF8-L protein sequence against the Phytoome v2 database (available on the World Wide Web) or TIGR gene indices (available on the World Wide Web). Tentative consensus and ESTs from *Brassica* with homology to AtGSTF8 were obtained
Differential Expression and Targeting of GSTF8

a

\[ \begin{align*}
&\text{cDNA} \\
&\text{EST} \\
&\text{TAIR clone 11128453} \\
&\text{TAIR clone R12910}
\end{align*} \]

\[ \begin{align*}
\text{GSTF8-L ATG} & \quad \text{GSTF8-S ATG} \\
\text{-100} & \quad \text{-100}
\end{align*} \]

b

\[ \begin{align*}
\text{GSTF8-L} \\
0 & \quad \text{MGAIQRLLFLPSIKSIIHSHLLSNFKNKRSNSSSSSSIIMASIKVGVPSTMATATLYEKDQFELIPVDMRAGAHKQAHLALNPF} \\
100 & \quad \text{CQIPAELDQDLTLFEZRAITQYLAEEYSEKGEKLISQDCCKVKATTNWLQVEQQFDTPNASKLAFERVFKGMPHMTDPAVQELEGKLQKLVLYEAR} \\
200 & \quad \text{LAKSFLAGDSFTLADLHHLPAIHYLLGTDKVLDFSFPKVEWIKSIARPWAWKVIDLQKQ}
\end{align*} \]

\[ \begin{align*}
\text{GSTF8-S} \\
0 & \quad \text{-----------------------------------MASIKVGVPSTMATATLYEKDQFELIPVDMRAGAHKQAHLALNPF} \\
100 & \quad \text{CQIPAELDQDLTLFEZRAITQYLAEEYSEKGEKLISQDCCKVKATTNWLQVEQQFDTPNASKLAFERVFKGMPHMTDPAVQELEGKLQKLVLYEAR} \\
200 & \quad \text{LAKSFLAGDSFTLADLHHLPAIHYLLGTDKVLDFSFPKVEWIKSIARPWAWKVIDLQKQ}
\end{align*} \]

c

\[ \begin{align*}
\text{TATA} & \quad \text{TATA} \\
-400 & \quad \text{-341} \\
\text{TATA} & \quad \text{TATA} \\
-300 & \quad \text{-172} \quad \text{-165} \quad \text{-146} \quad \text{-143} \quad \text{-130} \quad \text{-128} \quad \text{-127} \\
\text{TATA} & \quad \text{TATA} \\
-200 & \quad \text{7} \quad \text{6} \quad \text{5} \quad \text{4} \quad \text{3} \\
\text{TATA} & \quad \text{TATA} \\
-100 & \quad \text{-93} \quad \text{2} \\
\text{TATA} & \quad \text{TATA} \\
\text{ATCCAATTCACATTGCAGGATACAGAAGCGACATTCCATAATATGATCGTGGTGT AGGTGTTGATCTTCT}
\end{align*} \]

\[ \begin{align*}
\text{GSTF8-L ATG} \\
+1 & \quad \text{GSTF8-S ATG} \\
\text{ATCCAATTCACATTGCAGGATACAGAAGCGACATTCCATAATATGATCGTGGTGT AGGTGTTGATCTTCT}
\end{align*} \]

d

\[ \begin{align*}
\text{141bp} \\
\text{10} \quad \text{9} \quad \text{8} \quad \text{3/4/5} \quad \text{2} \quad \text{1} \\
\text{100bp}
\end{align*} \]
by BLASTing the AtGSTF8-L coding or protein sequence against the Phytome data base, the *Brassica* data base (available on the World Wide Web), or GenBank™ *Brassica* sequences. Amino acid and nucleotide alignments were performed with ClustalW and Align (available on the World Wide Web) and visualized with BoxShade 3.21 (available on the World Wide Web).

**RESULTS**

*Identification of Alternate GSTF8 Transcription Start Sites—*

The GSTF8 gene was originally thought to encode a protein of 215 amino acids corresponding to the TAIR (The Arabidopsis Information Resource) clone R12910. However, analysis of ESTs and cDNAs indicated the presence of a cDNA and three ESTs with sequence 5′ to the mapped GSTF8 +1 TSS (1). This TAIR cDNA clone 11128453 is predicted to encode a 263-amino acid GSTF8 protein. We designated the larger GSTF8 protein as GSTF8-L and the smaller GSTF8 protein as GSTF8-S (Fig. 1, a and b). The ATG codon of GSTF8-L is in-frame with GSTF8-S, thereby forming a protein identical to GSTF8-S except for an N-terminal extension. The mapped GSTF8-S +1 TSS seems to be the predominant TSS, with 36 ESTs in the TAIR data base with 5′ ends positioned between the GSTF8-S +1 TSS and the GSTF8-S ATG start codon. In contrast, there are only three ESTs with sequence 5′ to the GSTF8-S +1 TSS, and only two of these are upstream of the GSTF8-L ATG start codon (Fig. 1a).

The 5′ end of the most upstream GSTF8 EST is only 13 bp upstream of the GSTF8-L ATG (Fig. 1a), and, based on the average length of eukaryotic 5′-untranslated regions (27) is unlikely to represent the true 5′ end of the GSTF8 5′-untranslated region. To map TSS upstream of the GSTF8-L ATG start codon, 5′-RLM-RACE was used with tissue sourced from a variety of conditions to enhance the identification of alternate GSTF8 TSS. Initial analysis, using primers specific to the total GSTF8 transcript population, only detected the already mapped GSTF8-S +1 TSS. This appeared to be accounted for by the greater abundance of the shorter GSTF8 transcript as well as high G + C content upstream of the GSTF8-S +1 TSS, which may have led to the formation of stable RNA secondary structures that may have inhibited cDNA synthesis (25, 28–30).

Using a modified protocol, which included an intermediate partial heat denaturation step during reverse transcription, nine TSS upstream of the GSTF8-S +1 TSS were mapped, and these appeared to form clusters on the GSTF8 promoter and downstream of potential TATA boxes (identified using Plant-CARE) (31) (Fig. 1c). Interestingly, there appeared to be bias toward detection of upstream TSS (8 of 10 TSS) in leaf tissue. A schematic representation of the mapped GSTF8 TSS relative to the two potential GSTF8 ATG start codons is depicted in Fig. 1d. The upstream GSTF8 transcripts contain several ATG codons upstream of the GSTF8-L ATG that may form small open reading frames (ORFs), although these would only be produced if transcription initiates at the most 5′ TSS (TSS10). A comparison of these small ORFs with proteins in GenBank™ (all nonredundant entries) or TAIR (AGI protein data set) using blastp, however, showed little or no similarities with other proteins (data not shown). It should be noted that no stop codons lie in-frame between the GSTF8-L ATG and GSTF8-S ATG codons.
Diffential Expression and Targeting of GSTF8

**Fig. 2**

As calculated using LinRegPCR software 7.5 (26). As shown in GSTF8-S we are unable to determine the relative amount of only the
tain the same sequence downstream of the

**GSTF8-L**

Initial observations from 5
ine the expression of the
time PCR was used with a primer pair specific to the upstream
transcript population. Since all of the

**GSTF8**

transcripts, quantitative real
time PCR was used with a primer pair specific to the upstream
**GSTF8** transcript and a primer pair to detect the entire **GSTF8**
transcript population. Since all of the **GSTF8** transcripts con-
tain the same sequence downstream of the **GSTF8** +1 TSS, we are unable to determine the relative amount of only the
**GSTF8** transcript. Both primer pairs have similar efficiencies,
as calculated using LinRegPCR software 7.5 (26). As shown in
**Fig. 2a**, in leaf tissue, the relative abundance of the longer
**GSTF8** transcript (**GSTF8-L**) was significantly less than that of

**DNA**

**Transcription start sites**

**-814**

**GSTF8 promoter**

**GSTF8** gene

**-814**

**GSTF8 promoter**

**luciferase**

**-800GSTF8-S**

**-800GSTF8-L**

**a**

**b**

4 days 7 days 10 days

**-800GSTF8-S**

**-800GSTF8-L**

**c**

4 days 7 days 10 days

**-800GSTF8-L**

**d**

Relative luciferase activity (U/ml, 20 sec/mg protein)

**-800GSTF8-S**

**-800GSTF8-L**

**FIGURE 3.** The **GSTF8-S** and **GSTF8-L** promoters show differential tissue-specific expression patterns. a, schematic representation of the −800GSTF8-S:Luc + (−814/+10) and −800GSTF8-L:Luc + (−814/−83) constructs with respect to the identified **GSTF8** TSS. The −800GSTF8-S construct is a transcriptional fusion of the −800 **GSTF8** promoter to the luciferase reporter and follows transcriptional activity from any of the **GSTF8** TSS but only translation from the **GSTF8-L** ATG. The −800GSTF8-L construct follows transcriptional activity from TSS upstream (2−10 in the figure) of the **GSTF8-S** +1 TSS (1 in the figure). b and c, **in vivo** luciferase imaging of 4–10-day-old −800GSTF8-S (b) and −800GSTF8-L (c) seedlings. Three independent lines (comprising 10–20 seedlings) for each construct were analyzed at each age, with seedlings from a representative line shown. Promoter activity, as determined by pseudocolored bioluminescence (blue and green signal) is overlaid on a photomage of the seedlings. d, relative −800GSTF8-S and −800GSTF8-L promoter activity in 4-day-old whole

Differential Tissue-specific Expression of GSTF8 Transcripts—
Initial observations from 5′-RACE analysis suggested that the
**GSTF8-L** transcript was less abundant than **GSTF8-S**. To exam-
ine the expression of the **GSTF8** transcripts, quantitative real
time PCR was used with a primer pair specific to the upstream
**GSTF8** transcripts and a primer pair to detect the entire **GSTF8**
transcript population. Since all of the **GSTF8** transcripts con-
tain the same sequence downstream of the **GSTF8** +1 TSS, we are unable to determine the relative amount of only the
**GSTF8-S** transcript. Both primer pairs have similar efficiencies,
as calculated using LinRegPCR software 7.5 (26). As shown in
**Fig. 2a**, in leaf tissue, the relative abundance of the longer
**GSTF8** transcript (**GSTF8-L**) was significantly less than that of

with respect to the **GSTF8-S** +1 TSS such that the 3′ end of the
construct is 10 bp downstream of the TSS (TSS2) closest to the
**GSTF8-L** ATG start codon.

Using a previously described **in vivo** luciferase imaging system (13, 15), basal −800GSTF8-S and −800GSTF8-L promoter activity in seedlings from 4 to 10 days of age was monitored (Fig. 3, b and c). −800GSTF8-S activity remained high in the main root throughout early development with expression in the cotyledons increasing at 10 days (Fig. 3b). At 4 days of age, −800GSTF8-L activity was highest within the leaves and the root tip (Fig. 3c). At 7 days of age, activity was still high within the leaves and root tip, but it had also increased around the hypocotyl, and this pattern was maintained at 10 days. Similar

The total amount of **GSTF8** (GSTF8-T). Along with EST data, this sug-
gests that **GSTF8-S** must represent the majority of total **GSTF8**
transcripts. Since **GSTF8-T** has previously been shown to be highly expressed in root tissue (1), a com-
parison between the abundance of the **GSTF8** transcripts in leaves ver-
sus roots was carried out (Fig. 2, b and c). As expected, the amount of
**GSTF8-T** was higher in roots, and this gradually increased with age
(Fig. 2b). However, the opposite expression pattern was detected for
**GSTF8-L** with expression up to 5 times higher in leaves compared
with roots, and this difference was more pronounced in younger plants
(Fig. 2c).

**GSTF8-L** and **GSTF8-S** Promoters Show Differential Spatial Expression Patterns—Since **GSTF8-L** was more
highly expressed in leaves compared with roots, and vice versa for
**GSTF8-S**, the regulation of the **GSTF8-L** and **GSTF8-S** promoters was examined, particularly in terms of
their spatial expression patterns. To conduct these experiments, **GSTF8**
promoter constructs linked to an enhanced version of luciferase
(Luc+) were created (Fig. 3a). The −800GSTF8-S:Luc + (−800GSTF8-S) construct contains **GSTF8** pro-
moter sequence from −814 to +10 with respect to the **GSTF8-S** +1 TSS. Luciferase expression from
−800GSTF8-S will not result if translation starts at the **GSTF8-L**
ATG codon, since the luciferase ORF will be out of frame. The
−800GSTF8-L:Luc + (−800GSTF8-L) construct contains **GSTF8** promoter sequence from −814 to −83 (also

28920 JOURNAL OF BIOLOGICAL CHEMISTRY

VOLUME 282 • NUMBER 39 • SEPTEMBER 28, 2007
results for both promoters were seen in plants at 14 and 21 days of age (data not shown). The overall strength of the −800GSTF8-S promoter was also significantly stronger than the −800GSTF8-L promoter (Fig. 3d). Therefore, these promoter studies that allow one to unequivocally follow transcription from the alternate TSS reflect the earlier transcript levels and expression patterns that were observed.

GSTF8-L and GSTF8-S Promoters Show Differential Spatial Expression Patterns in Response to Biotic and Abiotic Stress—Chen and Singh (12) have previously shown in transgenic seedlings that the −800GSTF8-S promoter is inducible in roots following SA treatment. Since the −800GSTF8-S and −800GSTF8-L promoters show differential tissue-specific expression patterns, we were interested to determine the effect of SA on −800GSTF8-L activity (Fig. 4, a and b). Interestingly, although the −800GSTF8-S promoter was highly responsive to SA with the highest expression in the roots (Fig. 4a), the −800GSTF8-L promoter exhibited little response, with only a 2-fold induction occurring in leaves (Fig. 4b). To determine if the −800GSTF8-L promoter was responsive to abiotic stress, −800GSTF8-L and −800GSTF8-S seedlings were treated with the herbicide Dicamba (Fig. 4, c and d). As for the SA treatment, the −800GSTF8-S promoter was highly responsive to Dicamba, and again this was more pronounced in roots (Fig. 4c). In contrast, −800GSTF8-L promoter activity only increased by 1.7-fold in the leaves following Dicamba treatment. Interestingly, a repression in −800GSTF8-L promoter activity occurred in the roots (Fig. 4, d and e), in striking contrast to the expression seen with the −800GSTF8-S promoter (Fig. 4, c and e) following Dicamba treatment.

GSTF8-L Is Predicted to Contain an Organelle Targeting Signal—The majority of nuclear encoded mitochondrial or plastid targeted proteins contain N-terminal extensions that encode for signal peptides (32). Since GSTF8-L contains an N-terminal extension relative to GSTF8-S, we were interested to determine if it also encoded a signal peptide. Analysis of full-length GSTF8-S and GSTF8-L proteins with six protein localization and organelle targeting prediction programs suggested that GSTF8-L contains a potential chloroplast or mitochondrial targeting signal and that targeting to the chloroplast may direct it to the chloroplast stroma (Table 2). Although GSTF8-S was predicted by some programs to be targeted to mitochondria or peroxisomes, most programs predicted that it contains no organelle targeting signal and is most likely cytoplasmic.
Several independent studies of Arabidopsis chloroplast proteomes have identified peptides common to GSTF8 (Table 3). However, the identified peptides are common to both GSTF8-L and GSTF8-S, with none of the peptides containing amino acids specific for GSTF8-L, consistent with removal of the targeting signal of GSTF8-L following import into the chloroplast.

*GSTF8-L and GSTF8-S Direct Differential GFP Targeting to the Chloroplast and Cytoplasm, Respectively*—To determine if and where GSTF8-L and GSTF8-S are targeted within the cell, we constructed GSTF8-GFP fusions. Full-length GSTF8-S or GSTF8-L proteins fused to the N terminus of GFP (35S GSTF8-S-full-length GFP and 35S GSTF8-L-full-length GFP) were constructed under the control of a constitutive promoter (35S CaMV) and used to stably transform Arabidopsis plants. Nontransformed seedlings and seedlings containing nontargeted 35S GFP were used as controls. As a control for mitochondrial targeting, Arabidopsis plants were also transformed with the mitochondrial targeting peptide; S, signal peptide; Other, any other location; ER, endoplasmic reticulum.

| Prediction program | Reference | Prediction details | GSTF8 sequence and its predicted cellular location (and/or probability of subcellular targeting) | GSTF8-L |
|--------------------|-----------|--------------------|------------------------------------------------------------------------------------------------|--------|
| ChloroP            | 33        | cTP present, score | Yes, 0.58                                                                                     |        |
| MitoProt           | 34        | Probability of export to mitochondria | 0.84, 0.81                                                                                   |        |
| Predotar           | 35        | Chloroplast score  | 0.15, 0.33                                                                                      |        |
|                    |           | ER score           | 0.00, 0.19                                                                                       |        |
|                    |           | Elsewhere score    | 0.03, 0.11                                                                                      |        |
|                    |           | Predicted subcellular location | None                                                                                     |        |
| IPSORT             | 36        | Highest predicted potential protein localization site and score | No signal, mitochondrial, or chloroplast transit peptides |        |
| PSORT              | 37        | 35S GSTF8-L-full-length GFP | Microbody (peroxisome) 0.56 | Possibly mitochondrial | Chloroplast transit peptide | Chloroplast stroma 0.53 |
| TargetP            | 38        | cTP score          | 0.03, 0.91                                                                                      |        |
|                    |           | mTP score          | 0.10, 0.11                                                                                      |        |
|                    |           | SP score           | 0.53, 0.01                                                                                      |        |
|                    |           | Other score        | 0.56, 0.03                                                                                      |        |
|                    |           | Predicted subcellular location | Any other location                                                                                   | Chloroplast |

| Reference | Sample location | Identified peptides |
|-----------|----------------|---------------------|
| 39, 40    | Total chloroplast envelope | GMFGMTTDDAFAVQELEGK |
| 41; also see Ref. 42 | Chloroplast thylakoids | AITQYLAEYSEK |
| 43        | Chloroplast stroma | AITQYLAEYSEGERGK |
| K. J. van Wijk, personal communication; see also Ref. 42 | Chloroplast stroma | ATTNWLVQVEGQQUFDPNASK |
|           | -               | GMFGMTTDDAFAVQELEGK |
|           | -               | QRAHLALNPFGQIQALEGDI/TLFESR |

TABLE 3
Proteomic identification of GSTF8 from chloroplasts

Highlighted are details of peptides unique to GSTF8 that have been identified from chloroplast proteomes. These peptides are common to both GSTF8-L and GSTF8-S. —, unavailable peptide details.

TABLE 2
Predicted targeting of GSTF8-S and GSTF8-L subcellular targeting

Shown is predicted targeting of GSTF8-S and GSTF8-L using the prediction programs ChloroP, MitoProt, Predotar, IPSORT, PSORT, and TargetP. The full-length amino acid sequences of GSTF8-L and GSTF8-S were analyzed using these programs. Only the highest prediction is shown for IPSORT. cTP, chloroplast transit peptide; mTP, mitochondrial targeting peptide; S, signal peptide; Other, any other location; ER, endoplasmic reticulum.

| Prediction program | Reference | Prediction details | GSTF8 sequence and its predicted cellular location (and/or probability of subcellular targeting) | GSTF8-L |
|--------------------|-----------|--------------------|------------------------------------------------------------------------------------------------|--------|
| ChloroP            | 33        | cTP present, score | Yes, 0.58                                                                                     |        |
| MitoProt           | 34        | Probability of export to mitochondria | 0.84, 0.81                                                                                   |        |
| Predotar           | 35        | Chloroplast score  | 0.15, 0.33                                                                                      |        |
|                    |           | ER score           | 0.00, 0.19                                                                                      |        |
|                    |           | Elsewhere score    | 0.03, 0.11                                                                                      |        |
|                    |           | Predicted subcellular location | None                                                                                     |        |
| IPSORT             | 36        | Highest predicted potential protein localization site and score | No signal, mitochondrial, or chloroplast transit peptides |        |
| PSORT              | 37        | 35S GSTF8-L-full-length GFP | Microbody (peroxisome) 0.56 | Possibly mitochondrial | Chloroplast transit peptide | Chloroplast stroma 0.53 |
| TargetP            | 38        | cTP score          | 0.03, 0.91                                                                                      |        |
|                    |           | mTP score          | 0.10, 0.11                                                                                      |        |
|                    |           | SP score           | 0.53, 0.01                                                                                      |        |
|                    |           | Other score        | 0.56, 0.03                                                                                      |        |
|                    |           | Predicted subcellular location | Any other location                                                                                   | Chloroplast |

Differential Expression and Targeting of GSTF8

Several independent studies of Arabidopsis chloroplast proteomes have identified peptides common to GSTF8 (Table 3). However, the identified peptides are common to both GSTF8-L and GSTF8-S, with none of the peptides containing amino acids specific for GSTF8-L, consistent with removal of the targeting signal of GSTF8-L following import into the chloroplast.

**GSTF8-L and GSTF8-S Direct Differential GFP Targeting to the Chloroplast and Cytoplasm, Respectively**—To determine if and where GSTF8-L and GSTF8-S are targeted within the cell, we constructed GSTF8-GFP fusions. Full-length GSTF8-S or GSTF8-L proteins fused to the N terminus of GFP (35S GSTF8-S-full-length GFP and 35S GSTF8-L-full-length GFP) were constructed under the control of a constitutive promoter (35S CaMV) and used to stably transform Arabidopsis plants. Nontransformed seedlings and seedlings containing nontargeted 35S GFP were used as controls. As a control for mitochondrial targeting, Arabidopsis plants were also transformed with the mitochondrial targeting peptide; S, signal peptide; Other, any other location; ER, endoplasmic reticulum.

Seven-day-old transgenic seedlings were examined for GFP fluorescence using confocal microscopy. Since imaging of GFP in leaf mesophyll cells can be difficult due to the large density of chloroplasts in these cells and the resulting high background autofluorescence, we imaged GFP in hypocotyl epidermal cells, where the chloroplasts are smaller and the chlorophyll signal is weaker (44). No signal was detected in the GFP channel of nontargeted seedlings (Fig. 5, e and c). The control 35S GFP construct, as expected, gave fluorescence throughout the nucleolus and the cytoplasm (Fig. 5e). Note that due to the highly vacuolate nature of the cells, GFP in the cytoplasm is seen as a narrow zone around the vacuoles. The overlaying of GFP fluorescence (green) and chloroplast autofluorescence (red) images confirms that no GFP is detected in the chloroplasts (Fig. 5, e–g).

Examination of 35S GSTF8-S-full-length GFP revealed that it has the same GFP pattern as the 35S GFP construct and is localized throughout the nucleolus and the cytoplasm (Fig. 5, i–k). Examination of 35S GSTF8-L-full-length GFP, however, revealed that it is localized to the chlorophyll-containing plastids (Fig. 5, m–o). Where the GFP and chlorophyll signals overlap, the merged image is yellow. Since some of the subcellular targeting programs predicted GSTF8-L to be targeted to the mitochondria (Table 2), a comparison of 35S GSTF8-L-full-length GFP localization was made with the mitochondrial localized 35S AOX-GFP (Fig. 5, q–s). GFP targeted under the direction of the AOX presequence produced a pattern of numerous small spots, as seen previously in soybean and tobacco by Chew et al. (23). A comparison of 35S GSTF8-S-full-length GFP and 35S AOX-GFP signals clearly indicates that the longer GSTF8 protein is not targeted to the mitochondria. Leaf epidermal and mesophyll cells were also examined for GFP localization and produced similar results to Fig. 5, but the results were not as clear due to the strength of the chlorophyll signal (data not shown).
To determine if the 35S GSTF8-L-full-length GFP could be targeted to nongreen plastids, imaging was conducted on seedling roots just below the crown (Fig. 5, d, h, l, p, and t). Fig. 5l shows that full-length GSTF8-S-GFP is also cytoplasmic in roots, whereas Fig. 5p shows that full-length GSTF8-L-GFP is targeted to the non-chlorophyll-containing plastids. Note that GFP signal in the root plastids is not circular due to the irregular shape and size of nongreen plastids, which are typically smaller than chloroplasts (45).

In Vitro Imports Confirm Targeting of GSTF8-L to Plastids—To confirm that GSTF8-L can be taken up into plastids and to determine if the signal peptide is removed during this process, in vitro import assays were carried out with the two radio-labeled precursor proteins with isolated chloroplasts (Fig. 6). In vitro translations of the radiolabeled GSTF8-L and GSTF8-S proteins produced products of the expected sizes, 29.2 and 24.07 kDa, respectively (Fig. 6a). Although a product of similar size to GSTF8-S was identified from translation of GSTF8-L, this may represent leaky scanning of the upstream ATG in vitro, since in vivo we did not observe a cytosolic location for 35S GSTF8-L-full-length GFP (Fig. 5).

In vitro import assays of the GSTF8 proteins revealed that GSTF8-L was imported into a protease-protected location, with the 29-kDa precursor cleaved to a protein with an apparent molecular mass of 24 kDa (Fig. 6b). This finding verifies previous proteomic data that suggested removal of the GSTF8-L targeting signal (Table 3). The GSTF8-S protein was not taken up by chloroplasts, since no protease-protected band was apparent (Fig. 6b), consistent with the GFP cytosolic localization data.

Analysis of GSTF8-L Reveals That the Plastid Targeting Signal Is Unique to the GSTF8-L N Terminal and Also Directs Targeting to Stromules—To determine if mutation of the GSTF8-L ATG results in a loss of the targeting signal, the upstream ATG was converted to TTG such that translation should proceed from the second methionine (the GSTF8-S ATG). In vitro translation of radiolabeled GSTF8-L-TTG only produced a protein of similar molecular mass to GSTF8-S (Fig. 6a) and when fused to GFP produced a cytosolic localization pattern, unlike that of the wild-type GSTF8-L but similar to that seen with GSTF8-S (Fig. 7a). Thus, both in vitro and in vivo assays indicate that GSTF8-L

---

**FIGURE 5.** GSTF8-L is exclusively targeted to plastids, whereas GSTF8-S is cytoplasmic. Confocal single scans through hypocotyl or root epidermal tissue. For hypocotyl tissue, the green channel in the first column (a, e, i, m, and q) shows GFP expression, the red channel in the middle column (b, f, j, n, and r) shows chlorophyll autofluorescence, and the last column (c, g, k, o, and s) shows a merge of the green and red channels. For root tissue, only the combined channels are shown (d, h, l, p, and t). Shown is a representative image from the analysis of two independent T2 lines for each construct. Scale bar, 20 μm. a–d, nontransformed Arabidopsis shows no GFP in the green channel. e–h, 35S GFP expression is localized within the cytoplasm and nucleoplasm. i–l, 35S GSTF8-S-full-length GFP is localized within the cytoplasm with the same expression pattern as 35S GFP. m–p, 35S GSTF8-L-full-length GFP is localized to plastids within the hypocotyl and roots. q–t, 35S AOX-GFP is localized to mitochondria.
is targeted to chloroplasts, with the targeting signal unique to GSTF8-L.

To determine which region unique to GSTF8-L contained the plastid targeting signal, hypocotyls and roots of transgenic plants containing GFP fused to the first 100 or first 50 amino acids of GSTF8-L (35S GSTF8-L-100-GFP and 35S GSTF8-L-50-GFP) were examined. As shown in Fig. 7b, the first 50 amino acids of the GSTF8 N terminus were sufficient to give plastid-specific expression of GFP, confirming that the plastid targeting signal is located within the N-terminal extension. Similar results were also seen for GFP fused to the first 100 amino acids of GSTF8-L (data not shown).

Fig. 7b also shows targeting of GSTF8-L-50-GFP to stromules (stroma-filled tubules) in hypocotyl epidermal cells. Stromules are thin protrusions from the plastid surface into the surrounding cytoplasm (45–47). Since chloroplast thylakoids are not present in stromules (46), no chlorophyll autofluorescence is seen in stromules in the red channel of Fig. 7b. Targeting to stromules was also seen with the 100-amino acid and full-length GSTF8-L-GFP constructs (data not shown), thus demonstrating that GSTF8-L is targeted to the plastid stroma.

Putative Homologues of GSTF8-L Exist in Brassica—To evaluate whether homologues of GSTF8-L and/or GSTF8-S exist in other plant species, we searched for known or predicted proteins from available database sources. As shown in Fig. 8a, AtGSTF8-S homologues were identified in a range of dicot and monocot plants. In contrast, the only proteins that contained N-terminal extensions came from Brassica. Fig. 8, a and b, shows good conservation between AtGSTF8-L and the other Brassica proteins, and subcellular prediction programs predicted that, like GSTF8-L, these related proteins also contain chloroplast and/or mitochondrial targeting signals. Further analysis of...
Brassica napus tentative consensus and ESTs revealed the potential for long and short forms of the protein, which aligned with GSTF8-L and GSTF8-S from A. thaliana (Fig. 8c). This finding was also apparent in Brassica rapa and Brassica oleracea (data not shown), suggesting that the production of alternate GSTF8 transcripts occurs throughout the Brassicaceae. These results suggest that although GSTF8-S is conserved among a wide range of plants, GSTF8-L is only conserved among the Brassicaceae.

**DISCUSSION**

This report demonstrates that the A. thaliana GSTF8 gene is regulated through multiple TSS, which are differentially regulated in tissues and in response to stress and lead to the production of proteins that are targeted to different subcellular compartments. GSTF8 was originally thought to be transcribed from one major TSS (1); however, we have been able to identify an additional nine GSTF8 TSS, with the majority clustered into a group around −130/−140 bp. *In silico* analysis of GSTF8 using neural network promoter prediction (available on the World Wide Web) (48) identified several upstream TSS that correlated well with the regions detected experimentally by 5′-RACE. Therefore, the GSTF8 promoter has regions of transcriptional activity rather than one static site. Multiple TSS may arise from loose specification of TSS, a need for differential gene expression, and/or a need for regulating translation efficiency.
ciency (18–20, 49, 50). Our analysis of transcripts from the alternate GSTF8 TSS suggested the production of two forms of the GSTF8 protein, both identical except for the larger version containing an N-terminal extension. We detected no alternate splicing events during our 5′-RACE analysis or GFP fusion protein construction. Furthermore, analysis of EST and cDNA entries in the TAIR data base shows no alternatively spliced GSTF8 transcripts, and in silico splice site prediction programs failed to identify potential splice sites within the 5′-untranslated regions of the upstream GSTF8 transcripts (data not shown). Although several AUG codons upstream of the GSTF8-L AUG are present and may translate to form small ORFs, these small ORFs do not seem to code for functional proteins. Although it remains to be investigated, these small ORFs may play other roles (22), such as inhibiting downstream GSTF8 translation.

Further analysis of the GSTF8-L N-terminal using subcellular prediction programs suggested that it may encode a chloroplast or mitochondrial targeting signal and that GSTF8-S is most likely cytosolic. The most common mechanism for targeting proteins with the same or similar functions to different subcellular compartments is to encode each by a different gene to produce distinct isoforms of a protein, each with a specific targeting signal (51, 52). However, multiple subcellular targeting of a protein can also be achieved from a single gene, with most reported cases in plants being in regard to co-targeting to the plastid or mitochondrial targeting signal and that GSTF8-S is most likely cytosolic. The most common mechanism for targeting proteins with the same or similar functions to different subcellular compartments is to encode each by a different gene to produce distinct isoforms of a protein, each with a specific targeting signal (51, 52). However, multiple subcellular targeting of a protein can also be achieved from a single gene, with most reported cases in plants being in regard to co-targeting to the mitochondria and chloroplasts, primarily through the use of an ambiguous targeting sequence that is recognized by both organelles (32, 52, 53). Dual targeting of a protein from a single gene can also arise from the use of alternate translation initiation codons (52, 54, 55). Using GSTF8-GFP fusions and in vitro plastid imports, we demonstrated that alternate targeting of GSTF8 to different subcellular compartments is achieved from a single gene through the use of alternate TSS. GSTF8-L is solely targeted to plastids, including chloroplasts in leaves and hypocotyls, and nongreen plastids in roots. Although we did not determine which type of plastid GSTF8-L is targeted within roots, it is most likely amyloplasts or proplastids in the meristematic tissues. In addition to this, the GSTF8-L N-terminal extension was sufficient to target GFP to plastids, and mutation of the upstream GSTF8-L ATG caused translation to initiate at the GSTF8-S ATG, resulting in a loss of the plastid targeting signal. These results are consistent with our in silico analysis of the GSTF8-L protein and published plastid proteome data. GSTF8-L was also predicted to be localized to the chloroplast stroma, which is also the most likely observation from the confocal analysis of the GSTF8-L-GFP fusions and the presence of GFP signal within the stromules. In contrast to GSTF8-L, GSTF8-S was localized to the cytoplasm in leaves, hypocotyls, and roots.

The use of multiple TSS to encode cytosolic and/or plastidic isoforms as seen for GSTF8 has also been demonstrated for the Arabidopsis GSH1 (γ-glutamylcysteine synthetase) and GSH2 (glutathione synthetase 2) genes (56), which catalyze the sequential biosynthesis of GSH (57, 58). Multiple TSS are used to generate several GSH2 transcripts, but only the most 5′ transcript encodes for a GSH2 protein containing a plastid targeting signal, with the other GSH2 transcripts lacking parts of this signal sequence and the resulting protein localized to the cytosol (56). For both GSH2 and GSTF8, the protein containing an N-terminal extension is encoded by a rarer transcript(s) and is targeted to plastids, whereas the isoform lacking the extension is cytosolic and encoded by a more abundant transcript(s). The use of alternate GSTF8 TSS also results in the production of differentially regulated transcripts. GSTF8-L expression was highest in leaves compared with roots, whereas the opposite pattern was observed for GSTF8-S. Similar tissue-specific expression patterns occur for GSH1 through the use of two TSS. However, differential targeting does not result, with both GSH1 transcripts encoding for the same plastid targeted protein (56).

Using transgenic seedlings containing −800GSTF8-S and −800GSTF8-L promoter constructs, we were able to follow temporal and tissue-specific promoter activity that regulates GSTF8-S and GSTF8-L expression. Analysis of −800GSTF8-L activity confirmed our previous transcript analysis data but also indicated that expression within the root was mostly confined to the root tip, probably in meristematic tissue, and within the hypocotyl or crown. Analysis of −800GSTF8-S indicated the highest activity throughout the root and root tip. The promoters also exhibited quite distinct responses to biotic and abiotic stress, with the −800GSTF8-S promoter being quite responsive to both biotic and abiotic stresses, whereas the −800GSTF8-L promoter had only a minor response to stress that was confined to leaves. Moreover, in the case of roots, −800GSTF8-L promoter activity was repressed in response to stress. These interesting results demonstrate that although the GSTF8-L and GSTF8-S promoters share the same cis-acting promoter sequences, their effects on the alternate TSS are quite different, leading to distinct tissue-specific and stress-responsive expression patterns.

The use of multiple TSS to direct stress-responsive GST expression has also been described for the Tau class GSTs, Bronze2 (Bz2) from maize, and GmGST26-A (also known as Gmhsop26-A and GH2/4) from soybean (59–62). Under control or heat shock conditions, Bz2 is transcribed from up to three TSS (61). However, following treatment with the heavy metal cadmium, an additional upstream TSS is used and also results in the increased presence of an unspliced Bz2 transcript. A similar scenario occurs for GmGST26-A expression, which can be transcribed from three alternate TSS and whose intron removal is also inhibited by cadmium (59, 60). For both Bz2 and GmGST26-A, the unspliced messages appear to encode for truncated versions of these proteins (59, 61, 63). For Bz2, however, the truncated protein has not been detected, suggesting that the unspliced transcripts are most likely noncoding (64). Our in silico analysis of transcripts originating from the alternate Bz2 and GmGST26-A TSS indicates that neither would result in the production of proteins with N-terminal extensions, suggesting a cytosolic localization for these two GSTs (data not shown).

Current knowledge of plant GSTs indicates that nearly all are soluble cytoplasmic enzymes (2, 5, 7, 10, 59, 62, 65), with the studies presented here representing the first case of a plant GST being targeted both to a subcellular organelle and to the cytoplasm. Although GSTF8 has been identified in plastid pro-
omes by MS (39–42) and even in the nuclear proteome (66), in another study of proteins purified from Arabidopsis chloroplasts, no GSTF8 was detected (67). GSTF8 was also not detected in the mitochondrial proteome (68). In this report, we verify the plastid proteomic data but show that only GSTF8-L is targeted to plastids with the signal peptide subsequently cleaved. Analysis of other Arabidopsis GSTs suggests that GSTL2 and DHAR3 also contain potential chloroplast or mitochondrial transit peptides through the presence of N-terminal extensions (2, 69). Although GSTL2 (70) and DHAR3 (39, 67) have been detected in the plastidic proteome by MS, a direct analysis of DHAR3 subcellular localization by in vitro import studies failed to detect any import of DHAR3 into either mitochondria or chloroplasts (71). These results demonstrate some of the problems associated with contamination and purification of subcellular organelles (11, 72).

The roles of the alternate GSTF8 proteins are unknown, and attempts to determine their function have been hindered by the observation that GSTF8 knock-out plants with no detectable GSTF8 protein expression (13) do not show any obvious phenotypes. This suggests compensation of GSTF8 activity by other genes (e.g., by one or more of the other 13 Phi GSTs). Danpures (51) suggests that multicompartmentalized proteins frequently have different, although similar, functions in different compartments, but this may not be the case for GSTF8. From our search of available plant sequence information, we only identified putative AtGSTF8 proteins with N-terminal extensions from Brassica species, whereas GSTF8-S was conserved across a wide range of species. Moreover, EST and tentative consensus sequence analysis from B. napus, B. rapa, and B. oleracea suggested the production of two protein forms, as for AtGSTF8, with one lacking the N-terminal extension. Our observation that GSTF8-S is highly conserved across a wide range of plants suggests a general role that is carried across the monocot/dicot divide. GSTF8-S is cytosolic and highly expressed in roots, where it may act to detoxify products of oxidative stress, such as toxic lipid hydroperoxide derivates, resulting from pathogen attack or xenobiotic chemicals in the soil. In contrast, the conservation observed for GSTF8-L suggests that it has a function unique to Brassicaceae, and this occurs in the chloroplasts. Interestingly, GSTF8 was recently identified in the stromal proteome of plants subject to long term cold exposure, and although not yet shown, it was suggested that GSTF8 may act to reduce lactoylglutathione toxicity (43). Future work on the alternate GSTF8 proteins will involve elucidating their functions within the Brassicaceae family and how their expression is regulated.

Acknowledgments—We thank Dr. Stephen Powles for the herbicides used in this paper, Dr. Orinda Chew for providing the AOXGFP vectors, Dr. Manny Delhaize for providing the ShMTP1-GFP seeds as GFP controls, Drs. Jonathan Anderson and Harvey Millar for helpful comments on the manuscript, and members of the Singh laboratory for useful discussions. The confocal microscopy was carried out at the Centre for Microscopy and Microanalysis/Biomedical Image and Analysis Facility, University of Western Australia.

6 P. G. Sappl, A. H. Millar, and K. B. Singh, unpublished results.
Differential Expression and Targeting of GSTF8

40. Peltier, J., Ytterberg, J., Sun, O., and van Wijk, K. (2004) J. Biol. Chem. 279, 49367–49383
41. Giacomelli, L., Rudella, A., and van Wijk, K. J. (2006) Plant Physiol. 141, 685–701
42. Friso, G., Ytterberg, A. J., Giacomelli, L., Peltier, J. B., Rudella, A., Sun, Q., and van Wijk, K. J. (2004) Plant Cell 16, 478–499
43. Goulas, E., Schubert, M., Kieselbach, T., Kleczkowski, L. A., Gardestrom, P., Schroder, W., and Hurry, V. (2006) Plant J. 47, 720–734
44. Logan, D. C., and Leaver, C. J. (2000) J. Exp. Bot. 51, 865–871
45. Kohler, R. H., and Hanson, M. R. (2000) J. Cell Sci. 113, 81–89
46. Natesan, S. K., Sullivan, J. A., and Gray, J. C. (2005) J. Exp. Bot. 56, 787–797
47. Waters, M. T., Rupert, G. F., and Pyke, K. A. (2004) Plant J. 39, 655–667
48. Reese, M. G. (2001) Comput. Chem. 26, 51–56
49. Lupold, D. S., Caiole, A. G., and Stern, D. B. (1999) J. Biol. Chem. 274, 3897–3903
50. Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki, Y., Nakamura, Y., Suyama, A., and Sugano, S. (2001) EMBO Rep. 2, 388–393
51. Danpure, C. J. (1995) Trends Cell Biol. 5, 230–238
52. Silva-Filho, M. C. (2003) Curr. Opin. Plant Biol. 6, 589–595
53. Peeters, N., and Small, I. (2001) Biochim. Biophys. Acta 1541, 54–63
54. Small, I., Wintz, H., Akashi, K., and Mireau, K. (1998) Plant Mol. Biol. 38, 265–277
55. Souciet, G., Menand, B., Ovesna, J., Cosset, A., Dietrich, A., and Wintz, H. (1999) Eur. J. Biochem. 266, 848–854
56. Wachter, A., Wolf, S., Steininger, H., Bogs, J., and Rausch, T. (2005) Plant J. 41, 15–30
57. Foyer, C. H., Theodoulou, F. L., and Delrot, S. (2001) Trends Plant Sci. 6, 486–492
58. Meyer, A. J., and Hell, R. (2005) Photosynth. Res. 86, 435–457
59. Czarnecka, E., Nagao, R. T., Key, J. L., and Gurley, W. B. (1988) Mol. Cell. Biol. 8, 1113–1122
60. Hagen, G., Uhrhammer, N., and Guilfoyle, T. J. (1988) J. Biol. Chem. 263, 6442–6446
61. Marrs, K. A., and Walbot, V. (1997) Plant Physiol. 113, 93–102
62. Ulmasov, T., Ohmiya, A., Hagen, G., and Guilfoyle, T. (1995) Plant Physiol. 108, 919–927
63. Czarnecka, E., Edelman, L., Schoff, F., and Key, J. L. (1984) Plant Mol. Biol. 3, 45–58
64. Paino, C. F., and Walbot, V. (2003) Plant Mol. Biol. 53, 75–86
65. Flury, T., Wagner, E., and Kreuz, K. (1996) Plant Physiol. 112, 1185–1190
66. Bae, M. S., Cho, E. J., Choi, E., and Park, O. K. (2003) Plant J. 36, 652–663
67. Kleffmann, T., Russenberger, D., von Zychlinski, A., Christopher, W., Sjo-lander, K., Gruss, W., and Baginsky, S. (2004) Curr. Biol. 14, 354–362
68. Millar, A. H., Sweetlove, L. J., Giegé, P., and Leaver, C. J. (2001) Plant Physiol. 127, 1711–1727
69. Dixon, D., Davis, B., and Edwards, R. (2002) J. Biol. Chem. 277, 30859–30869
70. Peltier, J. B., Cai, Y., Sun, Q., Zabrousakov, V., Giacomelli, L., Rudella, A., Ytterberg, A. J., Rutschow, H., and van Wijk, K. J. (2006) Mol. Cell. Proteomics 5, 114–133
71. Chew, O., Whelan, J., and Millar, A. H. (2003) J. Biol. Chem. 278, 46869–46877
72. Heazlewood, J. L., Tonti-Filippini, J., Verboom, R. E., Millar, A. H. (2005) Plant Physiol. 139, 598–609