The Arabidopsis phi class glutathione transferase *At*GSTF2: binding and regulation by biologically active heterocyclic ligands

David P. Dixon*, Jonathan D. Sellars† and Robert Edwards*

*Centre for Bioactive Chemistry, School of Biological and Biomedical Sciences, and †Chemistry Department, Durham University, South Road, Durham, DH1 3LE, UK

Corresponding author:
Prof. Robert Edwards
Tel: +44 (0)1904 462415
Fax: +44 (0)1904 462111
Email: robert.edwards@fera.gsi.gov.uk
Address: Food and Environment Research Agency, Sand Hutton, York, YO41 1LZ, UK

Synopsis

The plant-specific phi class of glutathione transferases (GSTFs) are often highly stress-inducible and expressed in a tissue-specific manner, suggestive of important protective functions. To date, these functions remain unknown, although roles in the binding and transport of reactive metabolites have been proposed. Using a sensitive and selective binding screen, we have probed the *Arabidopsis thaliana* GSTFs for natural product ligands from bacteria and plants. Uniquely, when overexpressed in bacteria, GSTF2 and GSTF3 bound a series of heterocyclic compounds including lumichrome, harmane, norharmane and indole-3-aldehyde. When screened against total metabolite extracts from *A. thaliana*, GSTF2 was also found to selectively bind the indole-derived phytoalexin camalexin as well as the flavonol quercetin-3-*O*-rhamnoside. In each case, isothermal titration calorimetry revealed high affinity binding (typically K_d < 1 μM), which was enhanced in the presence of glutathione and by the other heterocyclic ligands. With GSTF2, these secondary ligand associations resulted in an allosteric enhancement in glutathione conjugating activity. Together with the known stress responsiveness of GSTF2 and its association with membrane vesicles, these results are suggestive of roles in regulating the binding and transport of defence-related compounds *in planta*.

Short title: Arabidopsis GST ligands

Key words: defence; flavonols; global metabolite profiling; indoles; ligand fishing; ligandin; phytoalexin

Abbreviations used: CDNB, 1-chloro-2,4-dinitrobenzene; DHAR, dehydroascorbate reductase; DTT, dithiothreitol; GST, glutathione transferase; HBS, Hepes-buffered saline; IAA, indole-3-acetic acid; ITC, isothermal titration calorimetry; NPA, 1-*N*-naphthylphthalamic acid.

Introduction

In plants, glutathione transferases (GSTs; EC 2.5.1.18) are a diverse group of proteins, with the respective super-families described in *Arabidopsis thaliana* [1], rice [2, 3], poplar [4], maize and soybean [5]. Plant GSTs exist in seven distinct subfamilies, termed the phi (F), tau (U), theta (T), zeta (Z), lambda (L), dehydroascorbate reductase (DHAR) and TCHQD classes [1, 6]. Whereas the DHARs and GSTZs have roles in ascorbic acid and tyrosine metabolism respectively [6], the endogenous functions of the other classes, notably the large and plant-specific groups of phi and tau GSTs, are poorly understood [7]. This is in contrast to the known roles of GSTFs and GSTUs in catalysing the conjugation of xenobiotics with the tripeptide glutathione (GSH), with the resulting detoxification determining the selectivity of several major herbicides [8].

Using Arabidopsis as a model system, the functional genomics of the GSTFs and GSTUs are currently the subject of considerable interest. Proteomic and transcriptomic studies in Arabidopsis plants and cultures have shown that these proteins accumulate in response to xenobiotics [9, 10], plant hormones [11], infection [12, 13], as well as illumination and environmental stress [14]. In the few instances where progress has been made in defining function, the Arabidopsis *At*GSTs have been found to be involved in signalling and transport, rather than in conjugating natural products. For example, *At*GSTF12 is involved in anthocyanin and proanthocyanidin accumulation [15], while *At*GSTU20 modulates responses to light reception [16]. Based on these observations, the latest evidence suggests that plant GSTs have evolved regulatory non-catalytic functions, which are most likely an extension of their ability to selectively bind biologically active ligands. For example, tau class GSTs isolated from Arabidopsis and maize (*Zea mays*) have recently been shown to selectively bind fatty acid and porphyrin natural products respectively. When *At*GSTUs were incubated with plant extracts, family members bound oxidized fatty acid derivatives in a protein-ligand specific manner [17]. Binding partners included unstable oxylipins with known roles as stress signalling agents [18, 19]. Similarly, the expression of maize *Zm*GSTUs in the chloroplasts of Arabidopsis plants resulted in the accumulation of porphyrins due to these proteins binding unstable porphyrinogen precursors [17, 20]. In both cases, similar types of metabolites were found to hyper-accumulate in bacteria when these same enzymes were over-expressed in *Escherichia coli*. Thus, when *Zm*GSTU enzymes were over-expressed in *E. coli,* the bacteria accumulated porphyrin intermediates in a de-regulated manner, due to their

binding of the heme precursor harderoporphyrinogen [20]. When the Arabidopsis tau *At*GSTUs were over-expressed in *E. coli,* a range of unusual oxidised fatty acids were determined, which showed similarities in terms of chain length and oxygenation to the metabolites identified in binding studies with plant extracts [17]. These studies have demonstrated that a combination of determining metabolite perturbation following GST expression in bacteria and the characterisation of bound ligands on incubation with plant extracts are powerful and complementary approaches in identifying classes of chemicals which are selectively recognized by these proteins. Importantly, the identification of binding partners by 'ligand fishing' can determine new functions for GSTs. For example, using such approaches we have recently identified a sub-set of flavonols as specific binding partners of Arabidopsis and wheat lambda GSTs that are associated with a novel glutathionyl-flavonol reductase activity [21].

Using the combination of metabolic perturbation following bacterial expression and ligand fishing using plant extracts we now report on the identification of a new set of binding partners of the Arabidopsis phi class protein *At*GSTF2. This GST has been identified in a range of studies as being regulated by pathogen attack, plant hormones, heavy metals and xenobiotics [10, 12, 22]. In contrast, to date its functional roles have remained unknown, despite its strong association with plant stress responses.

Experimental

Materials

Camalexin was prepared as described [23], with its properties agreeing with those previously determined [24]. 1-Acetyl-β-carboline was synthesised from tryptophan and methylglyoxal [25].

GST cloning, expression, purification and ligand analysis

Constructs for the bacterial expression of N-terminally *Strep*-tagged *At*GSTs were synthesised and expressed in *E. coli* as described previously [1], with any disruption to metabolism in the cultures determined by HPLC-MS as described [17]. Site-directed mutagenesis was performed using overlap extension PCR and appropriate mutagenic oligonucleotide primers. *Strep*-tagged proteins were purified using standard procedures [1]. For expression of untagged GSTF2, the coding sequence was excised from pET-STRP3 with *Nde*I and *Sal*I, and ligated into pET-24a (Novagen) *Nde*I and *Xho*I sites, generating pET24a-F2. Untagged GSTF2 was expressed as for tagged proteins and purified from soluble extracts using Orange A agarose (Millipore), with elution using 5 mM GSH in Hepes-buffered saline (HBS; 20 mM Hepes–NaOH, 150 mM NaCl, 1 mM EDTA, pH 7.5). To remove GSH and any bound ligands, purified *At*GSTF2 was buffer exchanged into HBS acidified to pH 5.5 with acetic acid then repeatedly concentrated by ultrafiltration prior to dilution with acidified HBS. Finally, *At*GSTF2 was desalted into the required buffer by gel filtration through a HiTrap desalting column (GE Healthcare). As an alternative method, *At*GSTF2 was extensively dialysed against HBS prior to use.

Enzyme and ligand binding assays

GST activity was determined with 1-chloro-2,4-dinitrobenzene (CDNB) as substrate [26]. GST(CDNB) assays in the presence of ligands were performed using 1 mM GSH in Hepes-buffered saline (HBS; 20 mM Hepes, 150 mM NaCl, 1 mM EDTA pH 7.6) for consistency with the binding studies. Isothermal titration calorimetry (ITC) experiments were performed at 25 °C in HBS buffer and analysed using a VP-ITC instrument with Origin 7.0 software (GE Healthcare). In addition to titrating ligand into protein, control titrations of ligand into buffer and buffer into protein were performed. Ligand concentrations were determined from absorbance values using extinction coefficients which were empirically determined for the ligands in HBS as ε_{348}(norharmane) = 3.98 mM^{-1}, ε_{348}(harmane) = 4.44 mM^{-1}, ε_{352}(lumichrome) = 9.13 mM^{-1} and ε_{317}(camalexin)= 1.23 mM^{-1}. For other ligands, literature extinction coefficients were used, while the concentration of GSH was determined gravimetrically. The content of *Strep-At*GSTF2 protein was based on a calculated ε_{280} of 20.34 mM^{-1}.

Gel filtration was performed using a Superdex 200 10/30 GL column (GE Healthcare), with HBS plus 2.5 mM DTT, and 1 mM GSH as running buffer at a flow rate of 0.5 ml/min. The effect of ligands on the elution of *At*GSTF2 was tested by adding 10 μM lumichrome to the running buffer. Protein elution was monitored by determining the absorbance at 280 nm, after calibrating the column with aldolase, bovine serum albumin, ovalbumin, chymotrypsinogen and ribonuclease A.

Results

Ligand binding to At*GSTFs in* E. coli

As an extension to our previously reported studies with the *At*GSTU*s* [17], all the members of the *At*GSTFs that expressed as soluble *Strep*-tagged recombinant proteins (*At*GSTs F2-F10 and F14) were transformed into *E. coli* as described previously [1]. After growth under inducing conditions to stationary phase, the bacteria were harvested and solvent-extracted prior to analysis by HPLC, coupled with UV and MS detection. Unlike the GSTUs, the expression of GSTs F4, F5, F6, F7, F8, F9, F10 and F14 did not give any major perturbations in metabolic profiles as compared with the control bacteria. In contrast, examination of the metabolome profiles from *E. coli* expressing *At*GSTF2 [AGI code: At4g02520] and *At*GSTF3 [AGI code: At2g02930] revealed an identical group of five novel UV-absorbing peaks (**1** to **5**; Fig. 1A), which were not observed in the extracts from the control bacteria. In each case, these novel metabolites were tentatively identified by tandem and accurate mass electrospray MS as *N*-containing heterocyclic natural products (Fig. 2); namely norharmane (**1**), harmane (**2**), lumichrome (**3**), indole-3-aldehyde (**4**) and 1-acetyl-β-carboline (**5**). Identities were subsequently confirmed by comparing their spectroscopic properties with those of authentic standards. To distinguish between metabolite accumulation due to a selective binding to the GST and perturbations caused by a non-specific ectopic protein expression, *Strep*-tagged *At*GSTF2 and *At*GSTF3 were affinity-purified from *E. coli* and then extracted with methanol and analysed for co-purifying ligands. Compounds **1**

to **4** were recovered bound to both *At*GSTF2 and *At*GSTF3 (Fig. 1B,C), showing that their accumulation was due to their selective associations with the respective proteins. In view of the identical ligand binding profiles obtained with the two Arabidopsis GSTFs, further detailed characterization was focused on *At*GSTF2, which on the basis of its known regulation and stress inducibility was the best characterized of the two proteins [12].

Identification of Arabidopsis ligands of AtGSTF2

To complement and extend the bacterial screening, plants were screened for compounds that bound strongly to *At*GSTF2. Two global metabolite profiling approaches were adopted, namely an *in vivo* method where the *Strep*-tagged enzymes were plant-expressed and bound ligands identified after affinity purification of the proteins and an *in vitro* method where the bacterially-expressed, purified enzymes were incubated with plant extracts [21]. In both cases, a *Strep*-tagged *At*GSTU19 protein was used as a 'ligand-fishing' control. For *in vivo* ligand fishing, *Strep*-tagged *At*GSTF2 was stably expressed in Arabidopsis, but poor expression levels hampered protein and ligand recovery. As a result, this method was not pursued further. Instead the *in vitro* method was employed, with extracts from Arabidopsis passed over affinity-immobilised *At*GSTF2, which had been purified from bacteria grown in M9 medium (see below). As a further control, an additional phi-class enzyme (*At*GSTF6) was included. This approach readily identified plant-derived ligands of *At*GSTF2 that were not ligands of *At*GSTU19 (Fig. 3). Surprisingly, *At*GSTF6 pulled down a near-identical profile of ligands to *At*GSTF2. In particular, the abundances of two compounds **6** and **7** (Fig. 2), were enhanced at least 100-fold on binding to *At*GSTF2 and *At*GSTF6 compared with their retention by the *At*GSTU19 control. Compound **6** had a distinctive UV/Vis absorbance spectrum and was tentatively identified as the well known Arabidopsis flavonoid quercetin-3-*O*-rhamnoside (quercitrin). Compound **7** had a much weaker UV signature that nevertheless gave abundant ions by MS, and was tentatively identified as the indole-derived antifungal phytoalexin camalexin. The identities of both compounds were subsequently confirmed by comparison with authentic standards. *At*GSTU19 previously bound metabolites from tobacco [17] and also specifically bound Arabidopsis metabolites that the *At*GSTFs did not, though the characterisation of these is beyond the scope of this report.

At*GSTF2-ligand binding studies*

ITC was used to examine the affinity and stoichiometry of binding of the ligands identified in *E. coli* and Arabidopsis to *At*GSTF2. In preparing the proteins for ITC studies, it became apparent that when the *Strep*-tagged *At*GSTF2 was purified from *E. coli* under standard conditions, the protein could not be used in titration assays, due to the high levels of pre-bound ligands. To minimize the likelihood of such contamination, *At*GSTF2 was expressed in bacteria grown in M9 minimal broth instead of the normal 'rich' medium used, to reduce the availability of preformed indole compounds. In an alternative approach, when *At*GSTF2 was purified from bacteria grown in rich medium, the protein was either extensively dialysed prior to use, or acidified to pH 5.5 with acetic acid to precipitate the protein and release associated ligands, prior to rapid desalting

and renaturation. Binding studies with lumichrome, norharmane and harmane showed that all three ligands bound with high affinity to *At*GSTF2, with the association being strongly promoted in the presence of 1 mM GSH (Table 1). However, increasing the concentration of GSH to 5mM did not further enhance ligand binding. In contrast, indole-3-aldehyde bound with more moderate affinity, while other *N*-containing heterocyclic molecules including riboflavin, kinetin, *trans*-zeatin, indole-3-acetic acid, harmine and tetrahydroharmane carboxylic acid showed no obvious binding by ITC. To rule out any effect of the *Strep*-tag on ligand association, untagged native *At*GSTF2 was purified from recombinant bacteria using Orange A affinity chromatography [27]. When assayed by ITC with lumichrome, binding to the native protein was essentially identical to that observed with the *Strep*-tagged form, confirming the affinity tag did not affect the results obtained. It was then of interest to determine the stoichiometry of ligand binding to *At*GSTF2, which normally associates to form homodimers [28]. In the case of norharmane, harmane and indole-3-aldehyde each ligand bound in an approximately 1:1 ratio with the *At*GSTF2 polypeptides, consistent with one molecule of ligand binding per GST monomer. In contrast, lumichrome bound to *At*GSTF2 with a lower stoichiometry (0.36:1). To examine the possibility that this binding-ratio arose from *At*GSTF2 forming trimeric or other unexpected oligomers, the protein was analysed by gel filtration chromatography in the presence and absence of lumichrome. In both cases, *At*GSTF2 eluted as a single peak immediately following an ovalbumin standard (43 kDa), consistent with the protein remaining as a dimer on ligand binding. It was concluded that a single lumichrome molecule was binding per protein dimer, although the determined stoichiometry was sufficiently low that inaccuracies in determining active protein and ligand concentrations could not satisfactorily account for the deviation from the expected 0.5:1 ratio for associations with a dimer.

Further binding studies with *At*GSTF2 focused on the optimal ligands harmane and lumichrome, and the effect of their binding on GST activity towards the model substrate CDNB. In both cases stoichiometric binding by each ligand enhanced the rate of CDNB conjugating activity by 60%. Kinetic analysis showed the addition of harmane (52.9 μM) did not cause a major decrease in substrate binding affinity, with the apparent K_m toward GSH being unchanged (15.5 μM), while the apparent K_m toward CDNB increased slightly from 1.6 mM to 2.4 mM. Instead, the observed activation was due to an increase in turnover (k_{cat}) in each case. There was no obvious explanation for this non-competitive activation, but possibilities include enhanced catalysis, a reduction in the binding affinity for the dinitrobenzylglutathione product of CDNB conjugation, or displacement of an unobserved non-competitive inhibitor from the enzyme active site. The observed increase in apparent K_m(CDNB) pointed to decreased product binding being the most likely explanation. To confirm that the association of these ligands with the protein did not involve covalent modification, electrospray mass spectrometry of intact, ligand-treated *At*GSTF2 was performed. These experiments demonstrated that the protein had not undergone any irreversible modifications.

The binding of other ligands to *At*GSTF2 was then examined. Firstly, the effect of lumichrome and GSH on the binding of harmane to *At*GSTF2 was

investigated using ITC (Table 1). The presence of lumichrome and GSH both increased harmane binding affinity, but by different mechanisms. Lumichrome reduced the stoichiometry and associated heat release of harmane binding to *At*GSTF2 by one third, while GSH showed no such effect. Complementary binding studies also demonstrated that both lumichrome and harmane increased binding affinity for GSH from K_d = 22 μM to K_d = ~ 5 μM (Fig. 4, Table 1). From these studies it was clear that the binding of ligands and GSH to *At*GSTF2 was cooperative. The binding of the phytoalexin camalexin to *At*GSTF2 showed a different response, with an apparent unusual stoichiometry of 1.45:1. Intriguingly, camalexin prevented any thermally visible binding of harmane to *At*GSTF2. Similarly the presence of harmane prevented any ITC-detectable association with camalexin, confirming that the two ligands shared a (presumably single) binding site. Quercitrin showed stoichiometric binding to *At*GSTF2, albeit with a somewhat weaker affinity (K_d = 6 μM) than determined with the heterocyclic ligands. Binding by quercitrin was abrogated by pre-binding of harmane, again suggesting that both ligands were binding at the same site.

Mutagenesis

To better understand the results obtained from binding studies, the active sites of *At*GSTF2 and *At*GSTF3 were examined, making use of the available crystal structure for *At*GSTF2 (PDB accession 1BX9). Of particular interest were residues around the active site that were conserved in these two GSTs but altered in related enzymes that did not show the selective binding to harmane or lumichrome and could therefore be involved in the ligand binding observed. Phe-123 was a promising candidate since this residue is substituted with an isoleucine in *At*GSTF4 and *At*GSTF5, being enzymes that do not bind these heterocycles. On generating the mutant *At*GSTF2-123I, the purified enzyme showed considerably higher CDNB-conjugating activity than the parent *At*GSTF2. In addition, unlike *At*GSTF2, the conjugating activity of the mutant was inhibited by high levels of harmane (Fig. 5). As compared to *At*GSTF2, *At*GSTF2-123I bound harmane and lumichrome with only slightly reduced affinity and with similar heat release (Table 1), showing that this mutation had not disrupted ligand binding. However, overall a lower binding stoichiometry was observed, suggesting that only about half the protein was active, which was indicative of the inefficient conformational folding of the mutated protein.

Discussion

The functional characterisation of GST superfamily members in Arabidopsis and other plants remains challenging, at least in part due to the potential functional redundancy resulting from the expansion of GST family gene numbers. Genetic approaches have had limited success at deducing GST function. For example, *At*GSTF12 disruption caused an obvious pigment phenotype in Arabidopsis seeds that could be linked to transport of anthocyanins and proanthocyanidins [15], although even here the precise function of the GST remains unknown. Altering the expression of other GSTs to perturb function has generally been unsuccessful. For example, the RNAi

knockdown of multiple GSTFs followed by detailed analysis failed to detect any major disruption in metabolism or physiology [29]. Instead, information on plant GST function has come through unexpected associations with other proteins of interest, such as the discovery of *At*GSTU20 interacting with phytochrome signalling components [16]. In an attempt to apply an alternative systematic strategy to define GST function we have developed a ligand fishing approach that uses global metabolite profiling of extracts purified by virtue of their affinity to a protein of interest, to identify tightly-binding ligands. We have already shown that this approach works to identify not only ligands of lambda-class GSTs, but also to subsequently identify a natural function for these enzymes [21]. Here, we have applied this approach to Arabidopsis phi-class GSTs, in particular *At*GSTF2, raising some interesting possibilities concerning the function of this protein *in planta*.

An initial metabolite profiling screen of bacterially expressed *At*GSTFs identified *At*GSTF2 (and its close homologue *At*GSTF3) as causing an unusual accumulation of heterocyclic compounds in the host bacteria, due to their tight binding as ligands to the ectopically expressed proteins. None of the identified ligands were bacterial in origin and instead originated from the rich culture medium and showed similarities to plant natural products. The four indole derivatives identified were probably formed through tryptophan reacting with other medium components during the autoclaving of the medium [25, 30]. Lumichrome was most likely present as a photodegradation product of the riboflavin present in yeast extract [31]. In each case, harmane, norharmane and lumichrome all bound to *At*GSTF2 far more tightly than had been described for previously identified interactions of this protein with auxins and flavonoids [32]. Although lumichrome was not identified as an *At*GSTF2 ligand in the Arabidopsis extracts, it is a well known plant metabolite with an established activity as a rhizobial signalling agent [33]., It is therefore conceivable that *At*GSTF2 could regulate the biological activity and associated photoactivation of lumichrome *in planta* by being involved in its transport and storage. Intriguingly, *At*GSTF2 is known to be up-regulated in response to pathogens, with riboflavin implicated in the induction of systemic disease resistance in plants [34], most likely following its transformation to lumichrome. The identified indole derivatives harmane, norharmane and 1-acetyl-β-carboline are unusual natural products that are not known to occur in Arabidopsis. However, Arabidopsis does contain a variety of indole-derived defence-related metabolites that these ligands could be mimicking, and the availability of pure standards made these useful for *in vitro* binding studies.

The binding studies with lumichrome and *At*GSTF2 were mechanistically revealing, as they demonstrated that the association with one ligand caused the protein to both increase its enzymic conjugating activity and enhance its binding to secondary ligands, notably harmane. Both harmane and lumichrome enhanced *At*GSTF2 enzyme activity, without affecting the affinity of the enzyme for its substrates. Further studies using the mutant *At*GSTF2-123I showed that for both wild-type and mutant enzyme, low concentrations (< 5 μM) of harmane enhanced enzyme activity while higher levels further increased activity for the wild-type enzyme but slowly reduced activity for the mutant (Fig. 5). Based on the kinetic data, we postulate that binding of harmane/lumichrome to *At*GSTF2

promotes the release of the DNB-GSH reaction product, thereby enhancing enzyme turnover. It is unclear whether the non-substrate ligands bind adjacent to the active site and directly influence catalysis, or bind further away and act as allosteric effectors. The competitive nature of ligand binding suggests these compounds associate at the same location. In addition, ligand and GSH binding are cooperative, suggesting interaction, and with the association enhancing catalysis without obviously altering substrate affinity. The two-phase effects seen on catalysis with increasing harmane concentration suggest the presence of a high-affinity harmane binding site that activates the enzyme and a much lower affinity binding site (or sites) that activates the wild-type enzyme but inhibits the 123I mutant. The nature of these binding sites awaits further investigation.

The association of *At*GSTF2 with multiple ligands had previously been observed in structural biology studies, which showed each monomer to bind two molecules of the inhibitor *S*-hexylglutathione [28]. Biochemical results consistent with *At*GSTF2 having separate binding sites for the ligands indole-3-acetic acid (IAA) and *N*-1-naphthylphthalamic acid (NPA) have also been reported [32]. The current study provides direct evidence of how ligand-binding interactions with *At*GSTF2 could modulate signalling events, with one group of compounds (eg: lumichrome) influencing the binding to other biologically active natural products (eg: indole derivatives). Alternatively, such binding could directly modulate the conjugating activity of *At*GSTF2 toward molecules involved in signalling, thereby altering their activity. Consistent with *At*GSTF2 having a signalling role, previous studies have shown the respective gene to be responsive to diverse stimuli, including plant hormones (auxins, ethylene) and pathogen attack [12, 32]. A regulatory role for *At*GSTF2 in development and stress tolerance has also been suggested by the observation that Arabidopsis plants expressing sense and antisense constructs of the orthologous gene from *Brassica juncea* showed subtle changes in flowering time, along with differences in shoot regeneration and stress resistance [35].

The binding of *At*GSTF2 to indoles, such as camalexin in Arabidopsis extracts, and the β-carboline alkaloids harmane, norharmane and 1-acetyl-β-carboline, and indole-3-aldehyde in *E. coli* cultures was particularly interesting. *At*GSTF2 was originally identified as a membrane-associated protein which could be labelled with azido-activated IAA in photoaffinity experiments [36]. Our own ITC studies showed that neither IAA, kinetin, nor *trans*-zeatin were tightly bound by *At*GSTF2, supporting earlier suggestions that this protein is unlikely to have a physiological role in binding these plant hormones [32]. However, *At*GSTF2 could bind other plant indole-derived metabolites sufficiently well to exert a physiological effect. For example the *At*GSTF2 ligand indole-3-aldehyde is a known plant metabolite [37], with its application causing an inhibition of lateral bud growth [38]. More strikingly, tight (K_d = 1.2 μM) binding of *At*GSTF2 to camalexin is very suggestive of a physiological role, perhaps in the intracellular transport of this reactive and defence-inducible secondary metabolite. A strong link between camalexin and glutathione already exists, with good evidence that the sulfur in camalexin is introduced through glutathionylation of a cytochrome P_{450}-activated precursor [39, 40]. This conjugation may occur spontaneously, but it is likely that the reaction is

more carefully regulated using a GST and as such, *At*GSTF6 has been implicated in this role [39]. There is a strong link between *At*GSTF2 and *At*GSTF6 (see below) and it is possible that *At*GSTF2 also catalyses glutathione conjugation of the camalexin precursor. A similar conjugation occurs during (indole) glucosinolate synthesis [41], providing another potential role for *At*GSTF2. GSTs may therefore be involved both in the synthesis of these indole-derived defence compounds, and in the subsequent intracellular transport of these compounds, as illustrated using camalexin as an example in Fig. 6.

*At*GSTF2 was also shown, along with *At*GSTF6, to selectively bind the flavonol quercitrin in Arabidopsis extracts, even though it was present in very low concentrations. *At*GSTF2 has also been shown to interact with flavonoids in earlier studies [32], with the binding of kaempferol inhibiting associations with NPA. This binding was shown to be competitive, with 100 µM kaempferol inhibiting binding to either NPA, or IAA by two thirds [32]. In the current studies, the binding of *At*GSTF2 to quercitrin was weaker than that observed with the *N*-containing ligands, suggesting that binding to camalexin and related molecules would be more likely than associations with flavonols. However, previous studies in Arabidopsis have shown that *At*GSTF2 had an altered sub-cellular localisation in flavonoid-deficient mutants as compared with wild type plants, suggesting flavonol-binding was physiologically relevant [32]. It is feasible that *At*GSTF2 functions to transport flavonoids under normal conditions, but in response to pathogen attack also transports the now abundant indole-derived phytoalexins, with its expression level increased to cope with the increased demand. The turnover of molecules for an intracellular transport process will be orders of magnitude slower than a typical enzyme catalytic turnover. Therefore intracellular transport proteins would be expected to be present at much higher abundance than enzymes in the same pathway. When coupled with the high levels of synthesis of camalexin during a defence response, very high amounts of transporter would be needed, and this is indeed observed for *At*GSTF2 and also for other GSTs, which are easily affinity-purified from relatively small amounts of Arabidopsis tissue [17].

The ligand binding results for *At*GSTF6 are intriguing. *At*GSTF2 and *At*GSTF6 are closely linked and their functions may well overlap. The two genes are closely related [1], and have closely matching expression patterns [40]. It is clear from the Arabidopsis ligand fishing results that the two enzymes bind similarly to indole and flavonol metabolites. It is perhaps surprising then that *At*GSTF6 did not bind similar ligands to *At*GSTF2 when expressed in *E. coli*, but it is reasonable to assume that differences between the proteins altered binding to non-target bacterial ligands without substantially altering binding to physiological ligands. It is therefore likely that the bacterial screens have missed other ligand-binding GSTs, and that these will only be detected using plant-based ligand fishing approaches.

GSTs are typically thought of as detoxifying conjugating enzymes. However, neither *At*GSTF2 nor *At*GSTF6 has been shown to form conjugates *in planta*, with the recombinant enzymes showing little activity towards model substrates such as CDNB [1]. None of the identified ligands appeared sufficiently electrophilic to allow GSH conjugation and while *At*GSTF2 did not

catalyse the conjugation of any of the ligands identified, their binding to the protein was shown to be promoted in the presence of GSH. This provides further evidence that such GSTs retain a functional dependence on GSH even when performing non-catalytic binding functions. While the studies described here cannot unequivocally identify a function for *At*GSTF2 and other GSTFs *in planta*, they do give a mechanistic context for how this protein could perform a regulatory role in responding to different classes of ligands through alterations in structure and function. When considered along with the strong and rapid induction of *At*GSTF2 in response to pathogens [12] and its association with plasma membrane vesicles [36] this does suggest that this protein performs a regulatory transport function involving the export of small bioactive natural products during plant stress.

Funding
This work was supported by the Biotechnology and Biological Sciences Research Council [grant number BBC51227X1]. The authors thank the BBSRC for supporting a research development fellowship awarded to RE.

References

1 Dixon, D. P., Hawkins, T., Hussey, P. J. and Edwards, R. (2009) Enzyme activities and subcellular localization of members of the *Arabidopsis* glutathione transferase superfamily. J. Exp. Bot. **60**, 1207-1218

2 Soranzo, N., Gorla, M. S., Mizzi, L., De Toma, G. and Frova, C. (2004) Organisation and structural evolution of the rice glutathione *S*-transferase gene family. Mol. Genet. Genomics. **271**, 511-521

3 Jain, M., Ghanashyam, C. and Bhattacharjee, A. (2010) Comprehensive expression analysis suggests overlapping and specific roles of rice glutathione *S*-transferase genes during development and stress responses. BMC Genomics. **11**, doi:10.1186/1471-2164-1111-1173

4 Lan, T., Yang, Z.-L., Yang, X., Liu, Y.-J., Wang, X.-R. and Zeng, Q.-Y. (2009) Extensive functional diversification of the *Populus* glutathione *S*-transferase supergene family. Plant Cell. **21**, 3749-3766

5 McGonigle, B., Keeler, S. J., Lau, S. M. C., Koeppe, M. K. and O'Keefe, D. P. (2000) A genomics approach to the comprehensive analysis of the glutathione *S*-transferase gene family in soybean and maize. Plant Physiol. **124**, 1105-1120

6 Dixon, D. P., Lapthorn, A. and Edwards, R. (2002) Plant glutathione transferases. Genome Biol. **3**, reviews3004.3001-reviews3004.3010

7 Frova, C. (2006) Glutathione transferases in the genomics era: New insights and perspectives. Biomol. Eng. **23**, 149-169

8 Edwards, R. and Dixon, D. P. (2005) Plant glutathione transferases. In Methods in Enzymology: Glutathione Transferases and Gamma-Glutamyl Transpeptidases. pp. 169-186

9 DeRidder, B. P., Dixon, D. P., Beussman, D. J., Edwards, R. and Goldsbrough, P. B. (2002) Induction of glutathione *S*-transferases in *Arabidopsis* by herbicide safeners. Plant Physiol. **130**, 1497-1505

10 Smith, A. P., DeRidder, B. P., Guo, W. J., Seeley, E. H., Regnier, F. E. and Goldsbrough, P. B. (2004) Proteomic analysis of *Arabidopsis* glutathione *S*-transferases from benoxacor- and copper-treated seedlings. J. Biol. Chem. **279**, 26098-26104

11 Moons, A. (2005) Regulatory and functional interactions of plant growth regulators and plant glutathione *S*-transferases (GSTs). In Plant Hormones. pp. 155-202

12 Lieberherr, D., Wagner, U., Dubuis, P.-H., Métraux, J.-P. and Mauch, F. (2003) The rapid induction of glutathione *S*-transferases *AtGSTF2* and *AtGSTF6* by avirulent *Pseudomonas syringae* is the result of combined salicylic acid and ethylene signaling. Plant Cell Physiol. **44**, 750-757

13 Sappl, P. G., Oñate-Sánchez, L., Singh, K. B. and Millar, A. H. (2004) Proteomic analysis of glutathione *S*-transferases of *Arabidopsis thaliana* reveals differential salicylic acid-induced expression of the plant-specific phi and tau classes. Plant Mol. Biol. **54**, 205-219

14 Glombitza, S., Dubuis, P., Thulke, O., Welzl, G., Bovet, L., Götz, M., Affenzeller, M., Geist, B., Hehn, A. and Asnaghi, C. (2004) Crosstalk and differential response to abiotic and biotic stressors reflected at the

transcriptional level of effector genes from secondary metabolism. Plant Mol. Biol. **54**, 817-835
15 Kitamura, S., Shikazono, N. and Tanaka, A. (2004) TRANSPARENT TESTA 19 is involved in the accumulation of both anthocyanins and proanthocyanidins in Arabidopsis. Plant J. **37**, 104-114
16 Chen, I.-C., Huang, I.-C., Liu, M.-J., Wang, Z.-G., Chung, S.-S. and Hsieh, H.-L. (2007) Glutathione *S*-transferase interacting with far-red insensitive 219 is involved in phytochrome A-mediated signaling in Arabidopsis. Plant Physiol. **143**, 1189-1202
17 Dixon, D. P. and Edwards, R. (2009) Selective binding of glutathione conjugates of fatty acid derivatives by plant glutathione transferases. J. Biol. Chem. **284**, 21249-21256
18 Dueckershoff, K., Mueller, S., Mueller, M. J. and Reinders, J. (2008) Impact of cyclopentenone-oxylipins on the proteome of *Arabidopsis thaliana*. Biochim. Biophys. Acta. **1784**, 1975-1985
19 Mueller, S., Hilbert, B., Dueckershoff, K., Roitsch, T., Krischke, M., Mueller, M. J. and Berger, S. (2008) General detoxification and stress responses are mediated by oxidized lipids through TGA transcription factors in *Arabidopsis*. Plant Cell. **20**, 768-785
20 Dixon, D. P., Lapthorn, A., Madesis, P., Mudd, E. A., Day, A. and Edwards, R. (2008) Binding and glutathione conjugation of porphyrinogens by plant glutathione transferases. J. Biol. Chem. **283**, 20268-20276
21 Dixon, D. P. and Edwards, R. (2010) Roles for stress-inducible lambda glutathione transferases in flavonoid metabolism in plants as identified by ligand fishing. J. Biol. Chem. **285**, 36322-36329
22 DeRidder, B. P. and Goldsbrough, P. B. (2006) Organ-specific expression of glutathione *S*-transferases and the efficacy of herbicide safeners in Arabidopsis. Plant Physiol. **140**, 167-175
23 Ayer, W. A., Craw, P. A., Ma, Y. and Miao, S. (1992) Synthesis of camalexin and related phytoalexins. Tetrahedron. **48**, 2919-2924
24 Fürstner, A. and Ernst, A. (1995) Syntheses of camalexin, indolopyridocoline and flavopereirine. Tetrahedron. **51**, 773-786
25 Nemet, I. and Varga-Defterdarović , L. (2007) Methylglyoxal-derived β-carbolines formed from tryptophan and its derivates in the Maillard reaction. Amino Acids. **32**, 291-293
26 Dixon, D. P., Cole, D. J. and Edwards, R. (1998) Purification, regulation and cloning of a glutathione transferase (GST) from maize resembling the auxin-inducible type-III GSTs. Plant Mol. Biol. **36**, 75-87
27 Dixon, D., Cole, D. J. and Edwards, R. (1997) Characterisation of multiple glutathione transferases containing the GST I subunit with activities toward herbicide substrates in maize (*Zea mays*). Pestic. Sci. **50**, 72-82
28 Reinemer, P., Prade, L., Hof, P., Neuefeind, T., Huber, R., Zettl, R., Palme, K., Schell, J., Koelln, I., Bartunik, H. D. and Bieseler, B. (1996) Three-dimensional structure of glutathione *S*-transferase from *Arabidopsis thaliana* at 2.2-Å resolution - structural characterization of herbicide-conjugating plant glutathione *S*-transferases and a novel active-site architecture. J. Mol. Biol. **255**, 289-309

29 Sappl, P. G., Carroll, A. J., Clifton, R., Lister, R., Whelan, J., Millar, A. H. and Singh, K. B. (2009) The Arabidopsis glutathione transferase gene family displays complex stress regulation and co-silencing multiple genes results in altered metabolic sensitivity to oxidative stress. Plant J. **58**, 53-68
30 Herraiz, T. (2000) Analysis of the bioactive alkaloids tetrahydro-ß-carboline and ß-carboline in food. J. Chromatogr. **881**, 483-499
31 Treadwell, G. E. and Metzler, D. E. (1972) Photoconversion of riboflavin to lumichrome in plant tissues. Plant Physiol. **49**, 991
32 Smith, A. P., Nourizadeh, S. D., Peer, W. A., Xu, J., Bandyopadhyay, A., Murphy, A. S. and Goldsbrough, P. B. (2003) *Arabidopsis AtGSTF2* is regulated by ethylene and auxin, and encodes a glutathione *S*-transferase that interacts with flavonoids. Plant J. **36** 433-442
33 Matiru, V. N. and Dakora, F. D. (2005) The rhizosphere signal molecule lumichrome alters seedling development in both legumes and cereals. New Phytologist. **166**, 439-444
34 Dong, H. and Beer, S. (2000) Riboflavin induces disease resistance in plants by activating a novel signal transduction pathway. Phytopathology. **90**, 801-811
35 Gong, H. B., Jiao, Y. X., Hu, W. W. and Pua, E. C. (2005) Expression of glutathione *S*-transferase and its role in plant growth and development *in vivo* and shoot morphogenesis *in vitro*. Plant Mol. Biol. **57**, 53-66
36 Zettl, R., Schell, J. and Palme, K. (1994) Photoaffinity labeling of *Arabidopsis thaliana* plasma membrane vesicles by 5-azido-[7-^{3}H]indole-3-acetic acid: Identification of a glutathione *S*-transferase. Proc. Natl. Acad. Sci. USA. **91**, 689-693
37 Nakabayashi, R., Kusano, M., Kobayashi, M., Tohge, T., Yonekura-Sakakibara, K., Kogure, N., Yamazaki, M., Kitajima, M., Saito, K. and Takayama, H. (2009) Metabolomics-oriented isolation and structure elucidation of 37 compounds including two anthocyanins from *Arabidopsis thaliana*. Phytochem. **70**, 1017-1029
38 Nakajima, E., Nakano, H., Yamada, K., Shigemori, H. and Hasegawa, K. (2002) Isolation and identification of lateral bud growth inhibitor, indole-3-aldehyde, involved in apical dominance of pea seedlings Phytochem. **61**, 863-865
39 Su, T. B., Xu, J. A., Li, Y. A., Lei, L., Zhao, L., Yang, H. L., Feng, J. D., Liu, G. Q. and Ren, D. T. (2011) Glutathione-indole-3-acetonitrile Is required for camalexin biosynthesis in *Arabidopsis thaliana*. Plant Cell. **23**, 364-380
40 Dixon, D. P., Skipsey, M. and Edwards, R. (2010) Roles for glutathione transferases in plant secondary metabolism. Phytochem. **71**, 338-350
41 Sonderby, I. E., Geu-Flores, F. and Halkier, B. A. (2010) Biosynthesis of glucosinolates - gene discovery and beyond. Trends Plant Sci. **15**, 283-290

Interaction (+ cofactors)	Parameter N	K_a (μM^{-1})	ΔH (kcal/mol)
GSTF2 + harmane	0.97 ± 0.01	2.09 ± 0.14	-15.0 ± 0.2
GSTF2 + harmane (+ GSH)	0.98 ± 0.01	3.68 ± 0.19	-16.9 ± 0.1
GSTF2 + harmane (+ lumichrome)	0.62 ± 0.01	3.61 ± 0.47	-11.4 ± 0.2
GSTF2 + harmane (+ GSH + lumichrome)	0.65 ± 0.01	3.15 ± 0.24	-10.5 ± 0.1
GSTF2 + lumichrome (+ GSH)	0.36 ± 0.01	6.09 ± 0.25	-31.8 ± 0.2
GSTF2 + norharmane (+ GSH)	0.88 ± 0.02	0.95 ± 0.09	-16.1 ± 0.6
GSTF2 + indole-3-aldehyde (+ GSH)	0.90 ± 0.09	0.09 ± 0.01	-13.7 ± 1.8
GSTF2 + quercitrin (+ GSH)	0.86 ± 0.06	0.16 ± 0.02	-21.1 ± 2.1
GSTF2 + camalexin (+ GSH)	1.45 ± 0.02	0.84 ± 0.08	-9.3 ± 0.2
GSTF2 + GSH	1[†]	0.044 ± 0.001	-5.2 ± 0.1
GSTF2 + GSH (+53 μM harmane)	1[†]	0.22 ± 0.03	-5.5 ± 0.2
GSTF2 + GSH (+ 64 μM lumichrome)	1[†]	0.17± 0.02	-5.7 ± 0.2
GSTF2-123I + harmane (+ GSH)	0.48 ± 0.06	2.84 ± 0.30	-15.2 ± 0.2
GSTF2-123I + lumichrome (+ GSH)	0.14 ± 0.01	3.57 ± 0.26	-35.9 ± 0.8

Table 1. ITC-derived thermodynamic parameters for titration of the compounds listed into *At*GSTF2 and its point mutant *At*GSTF2-123I.
In each case, parameters were obtained from fits to a one set of sites binding model, with errors representing standard error in the fitted model. N = ligand:GST binding stoichiometry. Certain experiments contained GSH (1 mM) and/or lumichrome (23.5 μM). [†] Stoichiometry unreliably determined at low affinity, so fixed at 1:1.

Figure Legends

Figure 1. Reversed-phase HPLC analysis of *At*GST ligands from *E. coli*. A) Extracts from bacteria expressing *At*GSTF12 (control; grey lower line) or *At*GSTF2 (black upper line), showing accumulation of metabolites **1** to **5**. B) Ligands from purified *At*GSTF2, and C) *At*GSTF3 were similarly analysed. Altered retention times between panel A and panels B and C were due to changing from a 2.1 x 50 mm column to a 2.1 x 100 mm column.

Figure 2. Structures of the identified *At*GSTF2 ligands. Ions in bold are the parent mass ions (MH^+), with associated collision-induced dissociation fragments.

Figure 3. Reversed-phase HPLC analysis of *At*GSTF2 and *At*GSTF6 ligands isolated from Arabidopsis extracts. Bacterially-expressed *Strep*-tagged enzymes were affinity immobilised and incubated with metabolite extracts from Arabidopsis. Ligands were recovered from affinity-eluted *At*GSTF2 and *At*GSTF6 (black lines) and compared with the compounds isolated from the control protein *At*GSTU19 (grey line). Identified Arabidopsis-derived ligands are numbered (see Figure 2). Ion counts for the *m/z* 201 ion corresponding to camalexin (peak 7) are shown inset.

Figure 4. The heat released on titrating GSH into a solution of 13.41 µM *At*GSTF2 in the presence of no additive (♦), 52.8 µM harmane (▲) or 64.3 µM lumichrome (■) as determined by ITC. The associated lines indicate best fit curves for a one set of sites binding model.

Figure 5. Effect of harmane on CDNB-conjugating activity of *At*GSTF2 and its point mutant *At*GSTF2-123I.

Figure 6. Potential roles for GSTs in the synthesis and delivery of camalexin in Arabidopsis. As well as a putative catalytic role introducing sulfur into the molecule, GSTFs may also be involved in transporting both the glutathione conjugate intermediate and the end product to their required locations.

Figure 1.

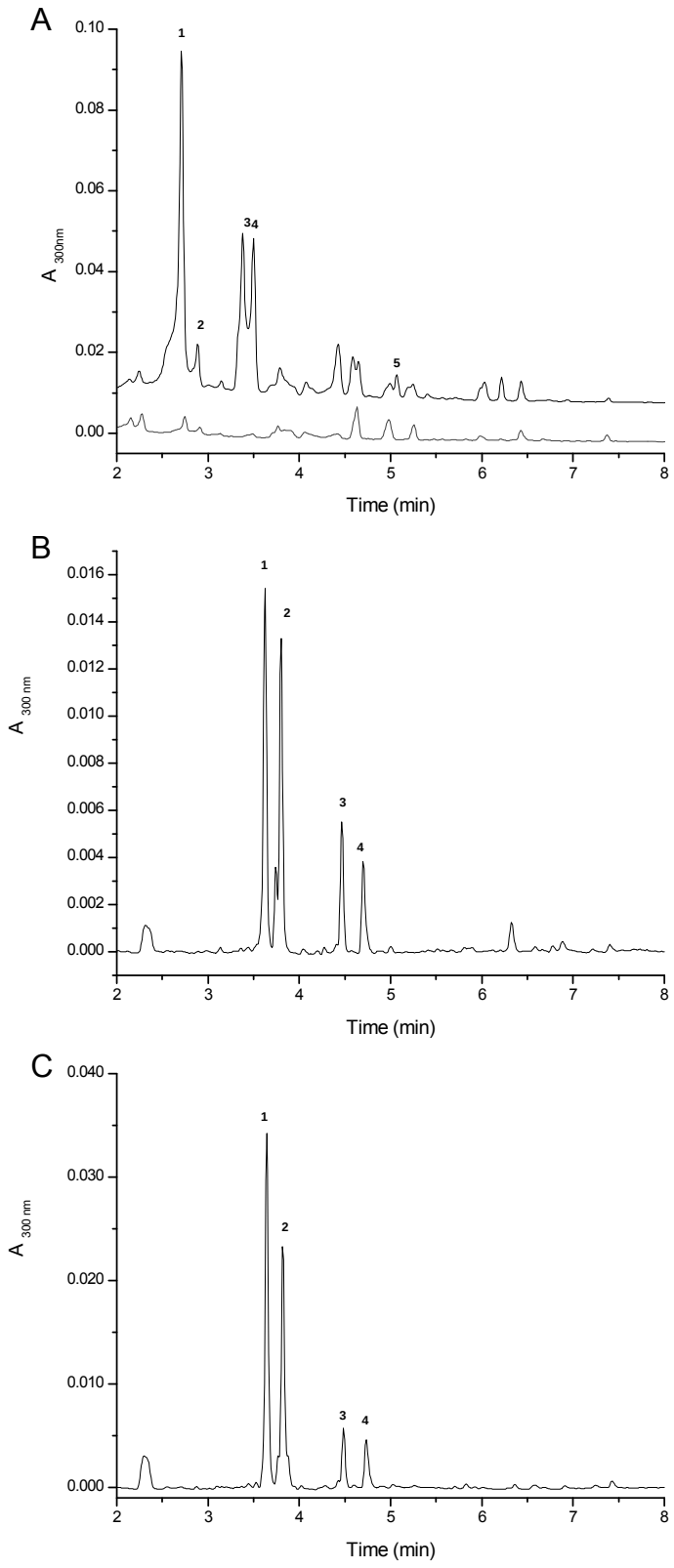


Figure 2.

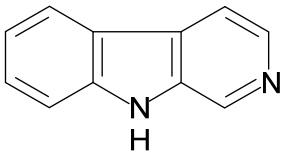
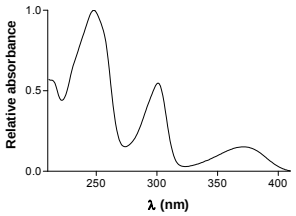
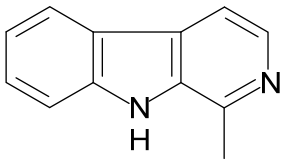
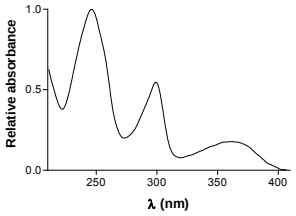
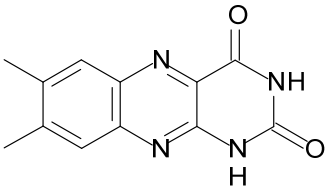
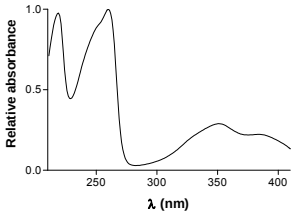
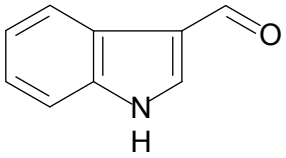
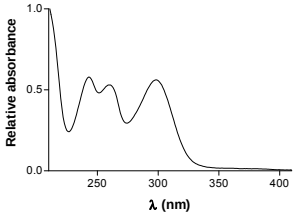
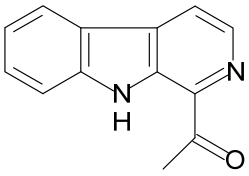
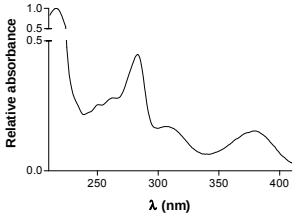
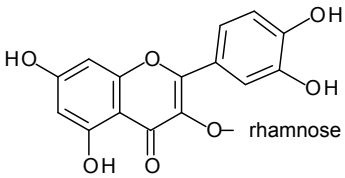
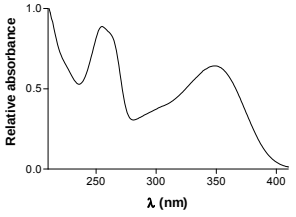
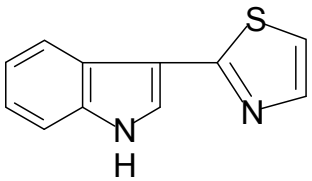
Peak	Compound	UV spectrum	Ions (*m/z*)
1	Norharmane 		**169.078** 142.069 115.057
2	Harmane 		**183.094** 168.072 142.068 115.057
3	Lumichrome 		**243.090** 216.080 198.070 172.091 145.079 130.068
4	Indole-3-aldehyde 		**146.063** 118.068 91.057
5	1-Acetyl-β-carboline 		**211.090** 193.080 169.080 142.068 140.053 115.058
6	Quercetin-3-*O*-rhamnoside 		**449.108** 303.053 229.055 129.052 85.023 71.044
7	Camalexin 	[Not resolved]	**201.053** 174.043 160.028 142.060 116.055

Figure 3.

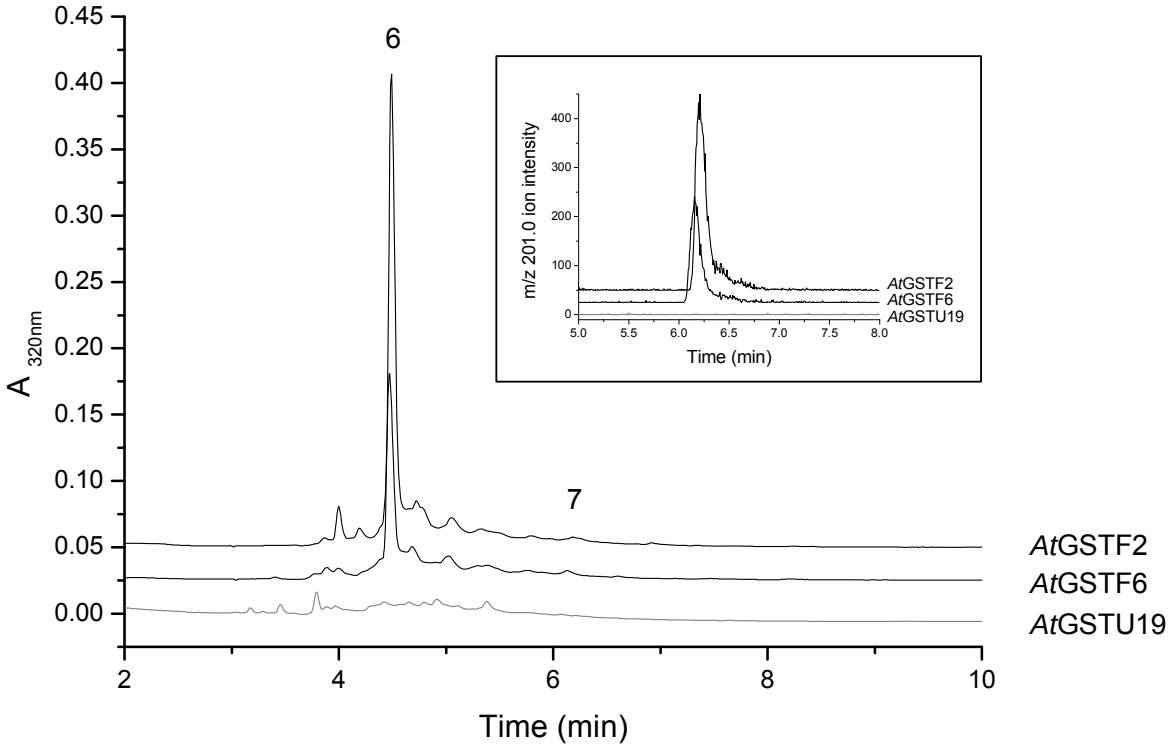

6
7
A 320nm
Time (min)
m/z 201.0 ion intensity
Time (min)
AtGSTF2
AtGSTF6
AtGSTU19
AtGSTF2
AtGSTF6
AtGSTU19

Figure 4.

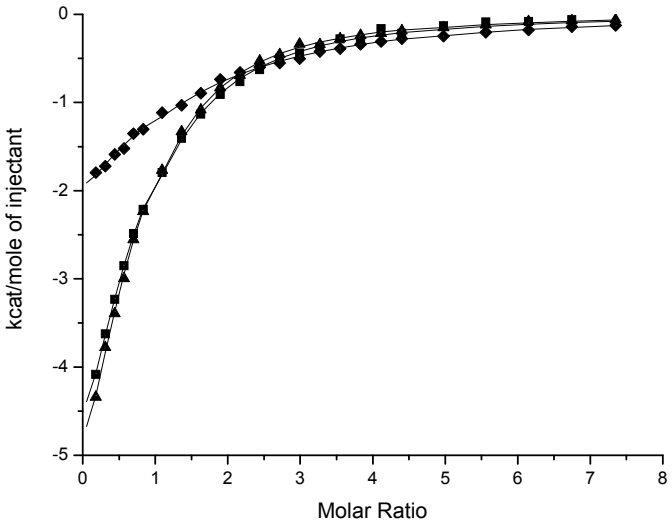


Figure 5.

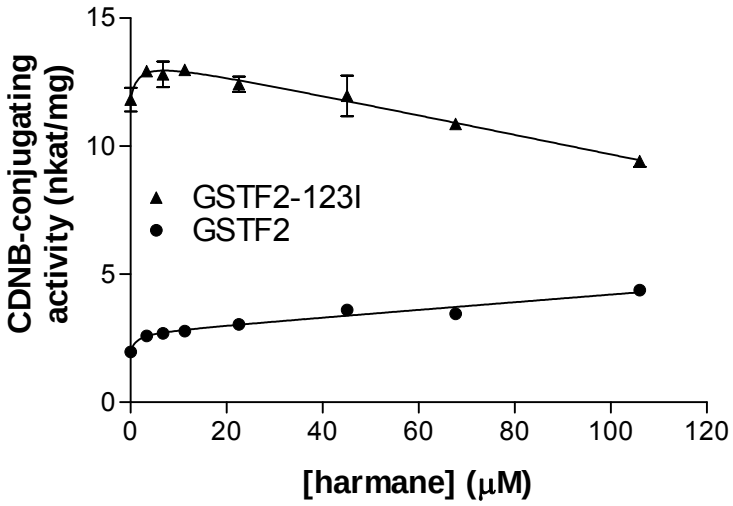
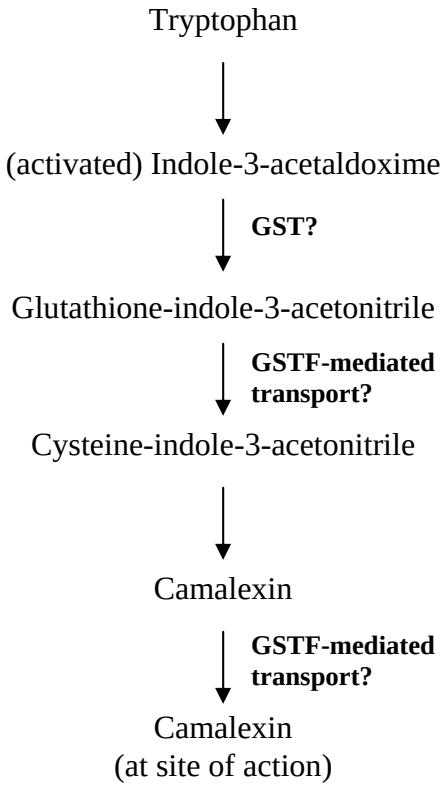

CDNB-conjugating activity (nkat/mg)
15
10
5
0
GSTF2-123I
GSTF2
0
20
40
60
80
100
120
[harmane] (μM)

Figure 6.


Tryptophan
(activated) Indole-3-acetaldoxime
GST?
Glutathione-indole-3-acetonitrile
GSTF-mediated transport?
Cysteine-indole-3-acetonitrile
Camalexin
GSTF-mediated transport?
Camalexin
(at site of action)