

# Diversification in substrate usage by glutathione synthetases from soya bean (*Glycine max*), wheat (*Triticum aestivum*) and maize (*Zea mays*)

Mark SKIPSEY\*<sup>1</sup>, Benjamin G. DAVIS† and Robert EDWARDS\*

\*Crop Protection Group, School of Biological and Biomedical Sciences, University of Durham, South Road, Durham DH1 3LE, U.K., and †Department of Chemistry, University of Oxford, Mansfield Road, Oxford OX1 3TA, U.K.

Unlike animals which accumulate glutathione ( $\gamma$ -glutamyl-L-cysteinyl-glycine) alone as their major thiol antioxidant, several crops synthesize alternative forms of glutathione by varying the carboxy residue. The molecular basis of this variation is not well understood, but the substrate specificity of the respective GSs (glutathione synthetases) has been implicated. To investigate their substrate tolerance, five GS-like cDNAs have been cloned from plants that can accumulate alternative forms of glutathione, notably soya bean [hGSH (homoglutathione or  $\gamma$ -glutamyl-L-cysteinyl- $\beta$ -alanine)], wheat (hydroxymethylglutathione or  $\gamma$ -glutamyl-L-cysteinyl-serine) and maize ( $\gamma$ -Glu-Cys-Glu). The respective recombinant GSs were then assayed for the incorporation of differing C-termini into  $\gamma$ -Glu-Cys. The soya bean

enzyme primarily incorporated  $\beta$ -alanine to form hGSH, whereas the GS enzymes from cereals preferentially catalysed the formation of glutathione. However, when assayed with other substrates, several GSs and one wheat enzyme in particular were able to synthesize a diverse range of glutathione variants by incorporating unusual C-terminal moieties including D-serine, non-natural amino acids and  $\alpha$ -amino alcohols. Our results suggest that plant GSs are capable of producing a diverse range of glutathione homologues depending on the availability of the acyl acceptor.

**Key words:** glutathione synthetase, *Glycine max* (soya bean), homoglutathione synthetase, hydroxymethylglutathione, *Triticum aestivum* (wheat), *Zea mays* (maize).

## INTRODUCTION

The majority of higher plants contain the tripeptide glutathione ( $\gamma$ -glutamyl-L-cysteinyl-glycine), which serves important functions as an antioxidant, a scavenger of reactive chemical species and a signalling agent [1,2]. In several economically important crops, glutathione is replaced either completely or in part by more unusual thiols. For example, in legume species including soya bean (*Glycine max* L.), hGSH (homoglutathione or  $\gamma$ -glutamyl-L-cysteinyl- $\beta$ -alanine) is the dominant thiol [3]. In other legumes such as *Medicago truncatula*, both glutathione and hGSH are synthesized, accumulating in an organ-specific manner [4]. In cereals of the Poaceae family, such as wheat (*Triticum aestivum* L.), hmGSH (hydroxymethylglutathione or  $\gamma$ -glutamyl-L-cysteinyl-serine) co-accumulates with glutathione [5]. Maize (*Zea mays* L.) also has the capacity to accumulate  $\gamma$ -ECE ( $\gamma$ -glutamyl-L-cysteine-glutamic acid) on exposure to heavy metals [6].

Glutathione, hGSH and related variants are synthesized from  $\gamma$ -EC ( $\gamma$ -glutamyl-L-cysteine), which is formed from L-glutamate and L-cysteine in an ATP-dependent reaction catalysed by  $\gamma$ -ECS ( $\gamma$ -glutamyl-L-cysteine synthetase; EC 6.3.2.2). Glutathione is then synthesized from  $\gamma$ -EC and glycine by an ATP-dependent GS (glutathione synthetase; EC 6.3.2.3). Most GS enzymes selectively produce glutathione due to their tight acyl acceptor specificity for glycine as a substrate, but in legumes the enzyme has evolved to preferentially incorporate  $\beta$ -alanine, effectively becoming an hGS (homoglutathione synthetase). Thus the hGSs from *M. truncatula* and pea (*Pisum sativum* L.) selectively catalyse the synthesis of hGSH from  $\gamma$ -EC and  $\beta$ -alanine [7,8]. The

relative abundance of glutathione and hGSH in legumes is then determined by the expression of either GS or hGS enzymes [9].

In wheat, the source of hmGSH is less clear-cut. While  $\gamma$ -EC remains the direct precursor, the serine could either be incorporated directly by an ATP-dependent GS-like reaction or could be derived through a post-synthetic modification of the glutathione molecule [5]. Similarly, the source of the glutamate analogue,  $\gamma$ -ECE, in maize is not known, with the possibilities being that it is derived from an alternative GS activity or from the degradation of phytochelatin, which are polymers of ( $\gamma$ -Glu-Cys)<sub>n</sub>-Glu-Cys-Gly [6]. To address whether or not the GSs of maize and wheat could determine the synthesis of alternative forms of glutathione, we have cloned enzymes from the respective cereals and determined their substrate preference, comparing the results obtained with an hGS from soya bean.

## MATERIALS AND METHODS

### Cloning of thiol synthetase cDNAs

DNA probes were prepared by PCR using a specific primer (MS-3, gCgAAgCCHCARMgAgARggHggAgg) based on aligned GS sequences [10]. PCR amplification was performed on cDNA prepared from total RNA prepared from either 10-day-old soya bean cell-suspension cultures (cv. Mandarin) [11], 10-day-old wheat shoots [12] or on plasmid DNA prepared from a mass excised cDNA library prepared from etiolated 10-day-old maize roots [13]. MS-3 was used in combination with the non-specific primer

Abbreviations used: AABA, BABA and GABA,  $\alpha$ -,  $\beta$ - and  $\gamma$ -aminobutyric acid respectively; ACV synthetase, L- $\delta$ -( $\alpha$ -amino adipoyl)-L-cysteinyl-D-valine synthetase; BAIBA,  $\beta$ -aminoisobutyric acid;  $\gamma$ -EC,  $\gamma$ -glutamyl-L-cysteine;  $\gamma$ -ECE,  $\gamma$ -glutamyl-L-cysteine-glutamic acid;  $\gamma$ -ECS,  $\gamma$ -glutamyl-L-cysteine synthetase; ESI, electrospray ionization; GS, glutathione synthetase; hGS, homoglutathione synthetase; hGSH, homoglutathione; hmGSH, hydroxymethylglutathione; TOF, time-of-flight.

<sup>1</sup> To whom correspondence should be addressed (email mark.skipsey@durham.ac.uk).

The nucleotide sequence data reported will appear in DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under the accession numbers AJ272035 (*GmGS*), AJ579380 (*TaGS1*), AJ579381 (*TaGS2*), AJ579382 (*TaGS3*) and AJ579383 (*ZmGS*).

OG9 for soya bean [11], whereas, for maize and wheat, two specific primers were designed from GS sequences in the respective EST (expressed sequence tag) databases, namely *ZmGS3'* (gCgCTgAgCTTgCCAgTCAAgtATAggC) and *TaGS3'* (gCggTCgACCTCATCTgTgAggAgAATgC) respectively. Amplification products were cloned into pGEM-T Easy (Promega, Chilworth, Southampton, U.K.) and sequenced using an Applied Biosystems 373 DNA sequencer. For soya bean, the full-length *GmGS* cDNA was cloned as described previously [10]. The partial wheat (*TaMS3-1*) and maize (*ZmMS3-1*) clones were PCR-labelled with digoxigenin (Roche, Lewes, East Sussex, U.K.) and used to screen cDNA libraries prepared from shoots [14] and roots of etiolated seedlings [13] respectively. In each case, 200 000 plaque-forming units were screened and positive clones were purified and sequenced on both strands.

### Expression and purification of recombinant enzymes

PCR was used to engineer restriction sites for direct cloning of the coding sequences into pET-24a or pET-24d (Novagen, Madison, WI, U.S.A.) essentially as described for *GmGS* [10]. For the maize and wheat clones, the following primer pairs were used in the presence of 1 M betaine: *ZmGS*, *ZmGS5'* Nde (CCAATACCTCATATgAgTgCCgCCATgCCg) and *ZmGS3'* (gCgCTCgAgCTTgCCAgTCAAgtATAggC); *TaGS1* either *TaGSBNdensp* (CCCTgAggCATATgggAgCCgAggCgC) and *TaGS3'* (gCggTCgACCTCATCTgTgAggAgAATgC) or *TaGSSigNde* (gCgCATA-TgTCCTCTTgCgTCTCTCTCTCC) and *TaGS3'*; *TaGS2*, *TaGSCNde* (CCACTgCCgCATATgAgCACC) and *TaGSC3'* (gCgCCTCgAgCTTgTCggTCAggTATAAgC). The pET-based plasmids were resequenced and used to transform Rosetta™ DE3 (Novagen) bacteria. The production and purification of the recombinant GSs was then performed as described for *GmGS* [10]. In each case, the purity of the His-tagged proteins was confirmed by SDS/PAGE (10% polyacrylamide) before use.

### Enzyme assays

Protein concentration was determined using the Bio-Rad dye-binding reagent, with  $\gamma$ -globulin standards, as recommended by the manufacturer. Enzyme assays were prepared in 250 mM Tris/HCl (pH 8.0) containing 20 mM MgCl<sub>2</sub>, 5 mM dithiothreitol, 10 mM ATP, 1 mM  $\gamma$ -EC and 100 mM (unless otherwise stated) of the co-substrate to be assayed, in a total volume of 100  $\mu$ l. The reaction mixtures were incubated at 30 °C for up to 120 min, with enzyme or substrate omitted for controls. At time intervals, 20  $\mu$ l samples were derivatized for 15 min at room temperature (21 °C) using 200  $\mu$ l of 0.2 M Tris/HCl (pH 8.0) containing 0.24  $\mu$ mol of monobromobimane. After adding 780  $\mu$ l of 5% (v/v) acetic acid, 50  $\mu$ l samples were analysed for fluorescent *S*-bimane derivatives by HPLC, with authentic glutathione and hGSH used to quantify tripeptide products [15]. After ensuring time and protein dependence for the assay, enzyme activity was expressed in pkat (= 1 pmol of product formed s<sup>-1</sup>).

Reaction products were also analysed directly before bimane derivatization using a TOF (time-of-flight) mass spectrometer (Micromass LCT; Micromass, Manchester, U.K.) using ESI (electrospray ionization), in positive ion mode. The reaction mixture (20  $\mu$ l) was injected on to a Jupiter C18 150 mm  $\times$  2 mm 5  $\mu$ m column (Phenomenex, Macclesfield, U.K.) in water/acetonitrile/formic acid (180:20:1, by vol.) at a flow rate of 0.2 ml  $\cdot$  min<sup>-1</sup>. The eluate was analysed for the expected mass of ions in the range 100–800 Da using the supplied MassLynx software, after calibration with sodium iodide.

### Determination of amino acid configuration in hmGSH

Samples of *S*-bimane-derivatized hmGSH were purified by HPLC from crude wheat leaf extracts or enzyme preparations after derivatization with monobromobimane, freeze-dried, resuspended in 10 mM Hepes buffer (pH 7.9) containing 0.2 mM EDTA, and digested with 1 unit of bovine  $\gamma$ -glutamyltranspeptidase for 2 h at 25 °C. The reaction products were analysed by reverse-phase HPLC [15] or on a Chirobiotic V 250 mm  $\times$  4.6 mm column (Astec, Whippany, NJ, U.S.A.) using the same gradient conditions. The retention times of reaction products were compared with those for derivatized authentic L-Cys-L-Ser and L-Cys-D-Ser standards, prepared on an automated peptide synthesizer, and the identity of the peptides was confirmed by ESI-MS.

## RESULTS

### Cloning of GS-like sequences from soya bean, wheat and maize

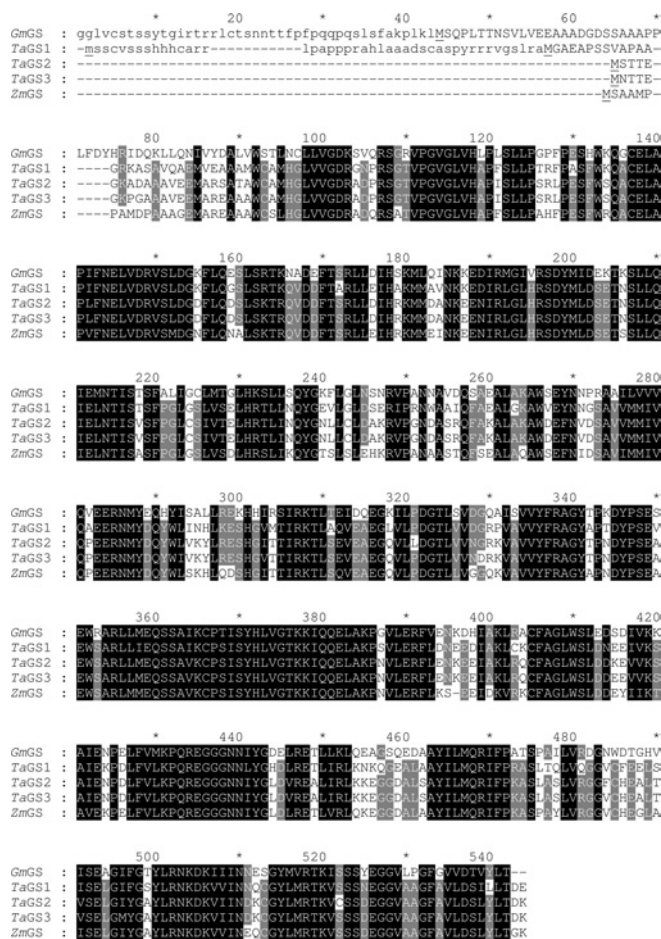
As described previously for the cloning of *GmGS* [10], PCR amplification generated a 354 bp sequence from wheat (*TaMS3-1*) and a 353 bp sequence from maize (*ZmMS3-1*), which were found to be closely related to other plant GS sequences when subjected to BLAST searches [16]. Following digoxigenin labelling of the PCR products, four positive clones (*TaGS1–TaGS4*) were obtained from a wheat library probed with *TaMS3-1*. When probed with *ZmMS3-1*, 17 positive clones, *ZmGS1–ZmGS17*, were obtained from the maize library. Following restriction enzyme analysis, the wheat clones were fully sequenced on both strands. A combination of restriction enzyme analysis and sequencing demonstrated all 17 maize clones to be essentially identical, except that some of the sequences were attenuated at the 5'-end. Three different wheat clone sequences were obtained, since *TaGS1* and *TaGS4* (= *TaGS1*) were identical. *TaGS2* and *TaGS3* differed in 13 out of 475 amino acids, but had homologous nucleotide sequences in both 5'- and 3'-untranslated regions. *TaGS1* was only 77 and 79% identical at the amino acid level with the *TaGS2* and *TaGS3* clones respectively. The amino acid sequences of the four GS clones isolated from wheat and maize are shown aligned with *GmGS* for reference in Figure 1.

The predicted amino acid sequences of the four cDNAs from the cereal species were more similar to GS rather than hGS sequences. *ZmGS* was most identical with a recently deposited maize cDNA putatively assigned to be a GS (EMBL accession no. AJ302784), which was predicted to encode a 45.9 kDa protein. However, *ZmGS* was 84 amino acids longer at its N-terminus than AJ302784 and contained four in frame methionine residues upstream, all of which scored highly as predicted translation start sites using NetStart 1.0 [17]. *TaGS1*, *TaGS2* and *TaGS3* were most identical with an unassigned rice cDNA clone (EMBL accession no. AK068792).

In common with *GmGS*, the *TaGS1* protein contained an extended N-terminus in comparison with the other GS sequences cloned in the present study (Figure 1), and was predicted to be targeted to the chloroplast by both PSORT [18] and ChloroP [19]. Similarly, the rice cDNA AK068792, which most closely resembled *TaGS1* in the database, also contained a homologous N-terminal extension predicted to target the putative peptide to the chloroplast. The other GS sequences from the cereals were not predicted to be targeted to subcellular compartments and therefore would be deposited in the cytosol.

### Expression of recombinant GS polypeptides

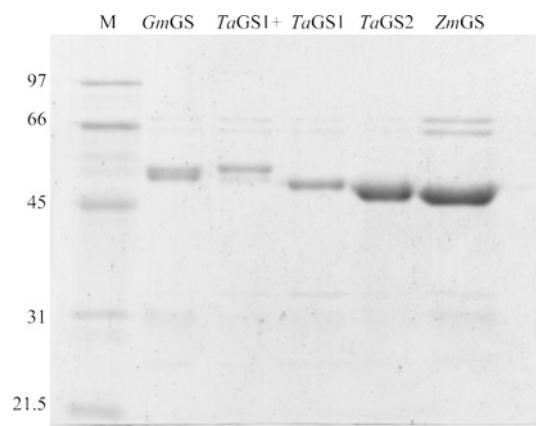
The plant GS-coding sequences were subcloned into pET24 for the expression of the C-terminal His-tagged fusion proteins,



**Figure 1** CLUSTAL W alignment of the predicted peptide sequences obtained for the GS cDNAs cloned from soya bean, wheat and maize

Residues shaded black are identical or conserved in all five sequences and those shaded in grey, identical in only four of the five sequences. The methionine residues used as translation start codons in pET constructs have been underlined. The conserved alanine residues found within GS, but not within hGS protein sequences [7], can be seen at position 530–531 in the cereal crop sequences but not in the soya bean sequence. Putative N-terminal polypeptide extensions encoding cellular targeting signals are shown in lower-case letters.

using the first methionine residue in each case to initiate translation. *TaGS3* could not be subcloned due to the difficulty in amplifying this cDNA with the required restriction enzyme sites. Similarly, difficulty was initially experienced in PCR amplification of the other wheat and maize clones, although the correct products were eventually obtained after adding betaine to the PCR. Since *TaGS1* was predicted to contain an N-terminal signal peptide, two pET constructs were employed that either encoded the protein with the signal peptide or utilized an endogenous methionine residue at position 45, two amino acids downstream of the predicted cleavage site to initiate translation. When analysed by SDS/PAGE, the lysate from those induced bacteria harbouring the pET constructs were found to contain recombinant polypeptides with molecular masses between 50 and 60 kDa that were absent from control bacteria. As reported previously [8], considerable difficulty was experienced in expressing the recombinant plant GS proteins. This expression problem was overcome using Rosetta<sup>TM</sup> (DE3) cells grown in a medium containing 1% glucose with an overnight induction at 15°C that improved both the solubility and yield of the recombinant GS proteins to between 5 and 20% of the total soluble protein extract.



**Figure 2** Analysis of purified recombinant GS polypeptides

His-tagged proteins were purified on Ni-chelate affinity columns and analysed by Coomassie Blue staining after SDS/PAGE (10% polyacrylamide). From left to right: M, reference molecular-mass standards; *GmGS* (minus leader sequence); *TaGS1+* (plus leader sequence); *TaGS1* (minus leader sequence); *TaGS2* (clone has no leader sequence); and *ZmGS* (clone has no leader sequence). Contaminating polypeptides around 60 and 30 kDa were present in all preparations and were presumably from the *E. coli* extract.

### Substrate selectivity of plant GSs

The His-tagged enzymes were purified using Ni-chelate affinity chromatography and purity was assessed by SDS/PAGE (Figure 2). When these purification runs were repeated with *Escherichia coli* extracts that did not contain recombinant GS proteins, no thiol synthetase activity was determined, confirming that the affinity-purified preparations were free of any contaminating bacterial GS enzyme. In the case of *GmGS*, depending on extraction conditions, the respective affinity-purified preparation could be resolved into a doublet of polypeptides. We presume that this was due to proteolytic processing of the recombinant *GmGS*, but on the basis of specific activity determinations between batches, this truncation did not affect activity. Purified protein (1 µg) was incubated with  $\gamma$ -EC in the presence of a representative range of acyl donor substrates including L-amino acids, D-amino acids and  $\alpha$ -amino alcohols in the presence of ATP. Enzyme activity was determined by quantifying the amount of product formed after derivatization to the fluorescent S-bimane adducts and analysis by HPLC, with the respective molecular masses of the underivatized tripeptides confirmed by ESI-TOF MS (Table 1). Under all assay conditions, GS-catalysed product formation was directly proportional to enzyme concentration and incubation time over 120 min for all purified enzymes. Under standard assay conditions, incubations were carried out for 60 min. In the case of the *TaGS1* clone, no significant differences were observed in the specific activities displayed when the recombinant protein was expressed with, or without, the signal peptide sequence. All activity data presented with respect to *TaGS1* was obtained with recombinant enzyme expressed without the signal peptide sequence.

Despite the presence of hGSH and hmGSH in place of glutathione in several plant species, the substrate selectivity of the respective GSs has received only limited attention. Initial experiments focused on the use of glycine and  $\beta$ -alanine as C-terminal substrates using a fixed concentration of 1 mM  $\gamma$ -EC. With the soya bean GS, the recombinant enzymes displayed a 6-fold higher specific activity with  $\beta$ -alanine as substrate compared with glycine (Table 1). With *TaGS1*, *TaGS2* and *ZmGS*, glycine was the preferred substrate, although significant activity was also observed with  $\beta$ -alanine in each case. The purified recombinant

**Table 1 Specific activities of purified recombinant thiol synthetases towards a range of substrates**

All assays were performed in triplicate with 1 mM  $\gamma$ -EC under standard assay conditions with 100 mM co-substrate, except for L-cysteine, which was used at 10 mM so as not to saturate the monobromobimane derivatization. Only those substrates with which an activity could be measured after correcting for enzyme-free controls are shown in the Table and the mass of the predicted reaction products,  $(M + H)^+$ , was confirmed by MS in each case using NaI as calibrant. Other substrates assayed for which no enzyme activity could be detected include L-serine, D- and L-glutamic acid, L-valine, L-threonine, L-lysine, L-cysteine, L-arginine, L-asparagine, DL-BABA, 6-aminohexanoic acid and n-butylamine. Each value represents the mean  $\pm$  S.D. for three replicate experiments. NA, no detectable activity.

Substrate	Recombinant enzyme*				Tripeptide product	
	<i>GmGS</i>	<i>TaGS1</i>	<i>TaGS2</i>	<i>ZmGS</i>	HPLC retention time (min)	$(M + H)^+$
Glycine	501 $\pm$ 75	1620 $\pm$ 13	2895 $\pm$ 345	1781 $\pm$ 416	8.5	308.09
$\beta$ -Alanine	3306 $\pm$ 469	56 $\pm$ 5	403 $\pm$ 77	591 $\pm$ 29	10.9	322.11
L-Alanine	NA	23 $\pm$ 5	564 $\pm$ 169	177 $\pm$ 12	10.3	322.11
D-Alanine	NA	NA	118 $\pm$ 5	168 $\pm$ 17	16.7	322.11
AABA	NA	NA	29 $\pm$ 6	NA	16.9	336.12
GABA	2287 $\pm$ 107	107 $\pm$ 34	584 $\pm$ 130	138 $\pm$ 20	13.9	336.12
BAIBA	1641 $\pm$ 106	10 $\pm$ 2	16 $\pm$ 3	NA	20.1	336.12
L-2,3-Diaminopropionic acid	1100 $\pm$ 182	33 $\pm$ 3	58 $\pm$ 10	65 $\pm$ 22	8.5	337.12
Ethanolamine	150 $\pm$ 11	NA	17 $\pm$ 1	NA	18.2	294.11
3-Amino-1-propanol	88 $\pm$ 8	NA	29 $\pm$ 7	NA	23.0	308.13
D-Serine	35 $\pm$ 2	NA	64 $\pm$ 13	NA	8.6	338.10
L-Ornithine	21 $\pm$ 1	NA	7 $\pm$ 3	NA	10.1	365.15

\*Activities are given in pkat (pmol of product formed  $\cdot$  s $^{-1}$ ).

**Table 2 Kinetic characterization of purified recombinant GS enzymes**

$K_m$  data for both  $\beta$ -alanine and glycine were assayed with a constant  $\gamma$ -EC concentration (1 mM).  $K_m$  data for  $\gamma$ -glutamylcysteine were assayed with a  $\beta$ -alanine concentration of 10 mM for *GmGS* and a glycine concentration of 10 mM for all other enzymes. For  $V_{max}$ , activities are expressed as pkat  $\cdot$  (mg of pure protein) $^{-1}$ .  $k_{cat}/K_m$  values presented in the Table are with respect to the acyl acceptor. ND, kinetic constants could not be determined due to the very low activities determined.  $V_{max}$  values are given in pkat  $\cdot$  (mg of purified GS) $^{-1}$ .

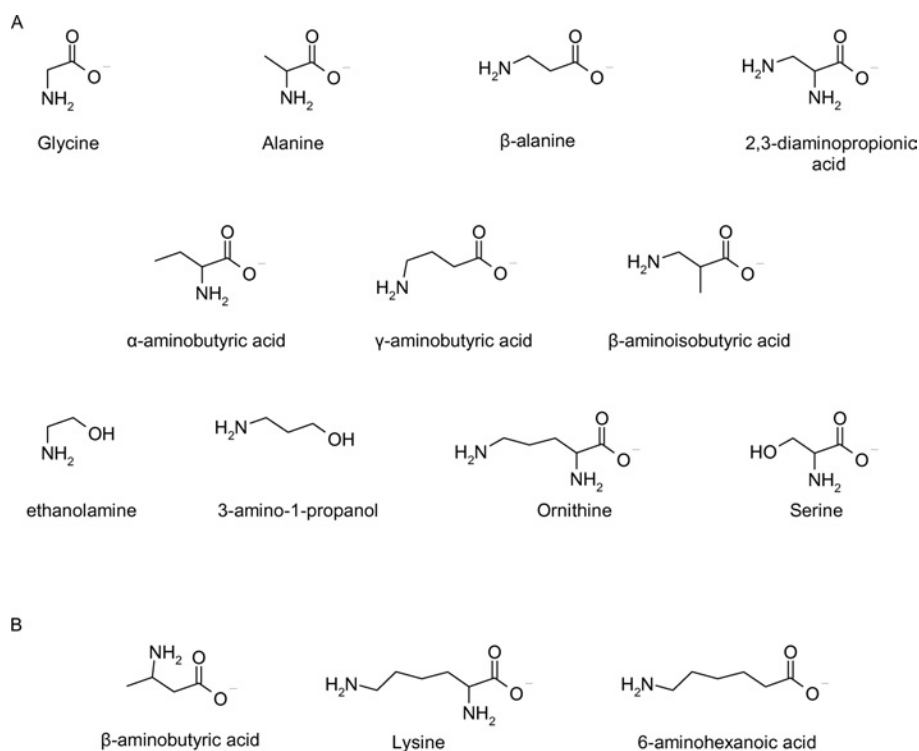
Kinetic constant	Co-substrate		
	Glycine	$\beta$ -Alanine	
<i>GmGS</i>	$V_{max}$ (pkat $\cdot$ mg $^{-1}$ )	364 $\pm$ 36	2132 $\pm$ 152
	$K_m$ (mM)	19 $\pm$ 6	0.32 $\pm$ 0.08
	$k_{cat}/K_m$ (M $^{-1}$ $\cdot$ s $^{-1}$ )	107	37083
	$K_m$ for $\gamma$ -EC (mM)	ND	0.06 $\pm$ 0.04
<i>TaGS1</i>	$V_{max}$ (pkat $\cdot$ mg $^{-1}$ )	663 $\pm$ 35	ND
	$K_m$ (mM)	0.16 $\pm$ 0.04	ND
	$k_{cat}/K_m$ (M $^{-1}$ $\cdot$ s $^{-1}$ )	24003	ND
	$K_m$ for $\gamma$ -EC (mM)	0.18 $\pm$ 0.07	ND
<i>TaGS2</i>	$V_{max}$ (pkat $\cdot$ mg $^{-1}$ )	2817 $\pm$ 172	5660 $\pm$ 3153
	$K_m$ (mM)	0.07 $\pm$ 0.02	170 $\pm$ 136
	$k_{cat}/K_m$ (M $^{-1}$ $\cdot$ s $^{-1}$ )	212404	175
	$K_m$ for $\gamma$ -EC (mM)	0.10 $\pm$ 0.02	ND
<i>ZmGS</i>	$V_{max}$ (pkat $\cdot$ mg $^{-1}$ )	1304 $\pm$ 179	372.4 $\pm$ 597
	$K_m$ (mM)	0.18 $\pm$ 0.08	592 $\pm$ 1077
	$k_{cat}/K_m$ (M $^{-1}$ $\cdot$ s $^{-1}$ )	38060	3.3
	$K_m$ for $\gamma$ -EC (mM)	0.03 $\pm$ 0.02	ND

enzymes were used for more detailed kinetic analysis to define the observed substrate preference between glycine and  $\beta$ -alanine (Table 2). The purified *GmGS* showed a significant specificity for  $\beta$ -alanine compared with glycine, as shown by an approx. 350-fold higher  $k_{cat}/K_m$  that is largely a consequence of an approx. 60-fold lower  $K_m$ , indicating a strong ground state binding affinity discrimination. The kinetic constants determined were very similar to those reported for the respective enzyme partially purified from soya bean and runner bean [20]. In contrast, glycine was the preferred substrate with the *TaGS* and *ZmGS* recombinant enzymes. The >1000-fold specificity discrimination, as indicated by  $k_{cat}/K_m$ , for the maize and wheat GS enzymes was again largely

a consequence of striking differences in  $K_m$ . However, on the basis of  $k_{cat}$ , a different pattern in their utilization of these two amino acids emerges (Table 2). While *TaGS1* showed negligible turnover with  $\beta$ -alanine, *TaGS2* had, in fact, a much greater  $V_{max}$  with this substrate than with glycine.

In soya bean, the presence of a  $\beta$ -alanine-selective GS correlated with the preferential accumulation of hGSH in this legume. It was therefore of interest to determine whether or not the accumulation of hmGSH in wheat and  $\gamma$ -ECE in maize could be similarly explained by an unusual amino acid selectivity of the respective GS enzymes. Therefore all enzymes were probed with the diverse array of C-terminal substrates shown in Figure 3. None among *TaGS1*, *TaGS2* or *ZmGS* would accept L-serine or L-glutamate as co-substrate with  $\gamma$ -EC. When assayed with the other amino acid co-substrates, enzyme activity was only seen with L-alanine, notably with *TaGS2*. The utilization of L-alanine by the cereal GS enzymes then prompted the question as to what other amino derivatives (Figure 3) could serve as GS substrates for the wheat, maize and soya bean enzymes (Table 1). Putative reaction products were analysed by a combination of HPLC and MS (Figure 4).

*TaGS1* showed a broadly similar selectivity for non-physiological amine substrates to *GmGS*, albeit with lower activities determined. *TaGS2* proved to be the most versatile enzyme, acting on most of the substrates tested including D-alanine and D-serine, and being the only GS to utilize AABA ( $\alpha$ -aminobutyric acid). *ZmGS* also catalysed the incorporation of L- and D-alanine but otherwise resembled *TaGS1* in its selectivity and did not utilize D-glutamic acid. For reference, none of the recombinant GSs could utilize DL-BABA (DL- $\beta$ -aminobutyric acid), 6-aminohexanoic acid or n-butylamine as substrates (Figure 3). A clear difference in the selectivity of *GmGS* and the wheat and cereal enzymes was the relative usage of GABA ( $\gamma$ -aminobutyric acid) and BAIBA ( $\beta$ -aminoisobutyric acid). All enzymes used GABA efficiently, but the introduction of a methyl group on the equivalent  $\alpha$ -carbon in BAIBA provoked at least a 10-fold decrease in activity with the cereal GS, but only a 30% reduction in *GmGS*. These results are in contradiction to a previous report utilizing partially purified hGS from *Phaseolus coccineus*, which was 5-fold more active in utilizing BAIBA than GABA [20]. Although GABA has been



**Figure 3** Chemical structures of some potential substrates

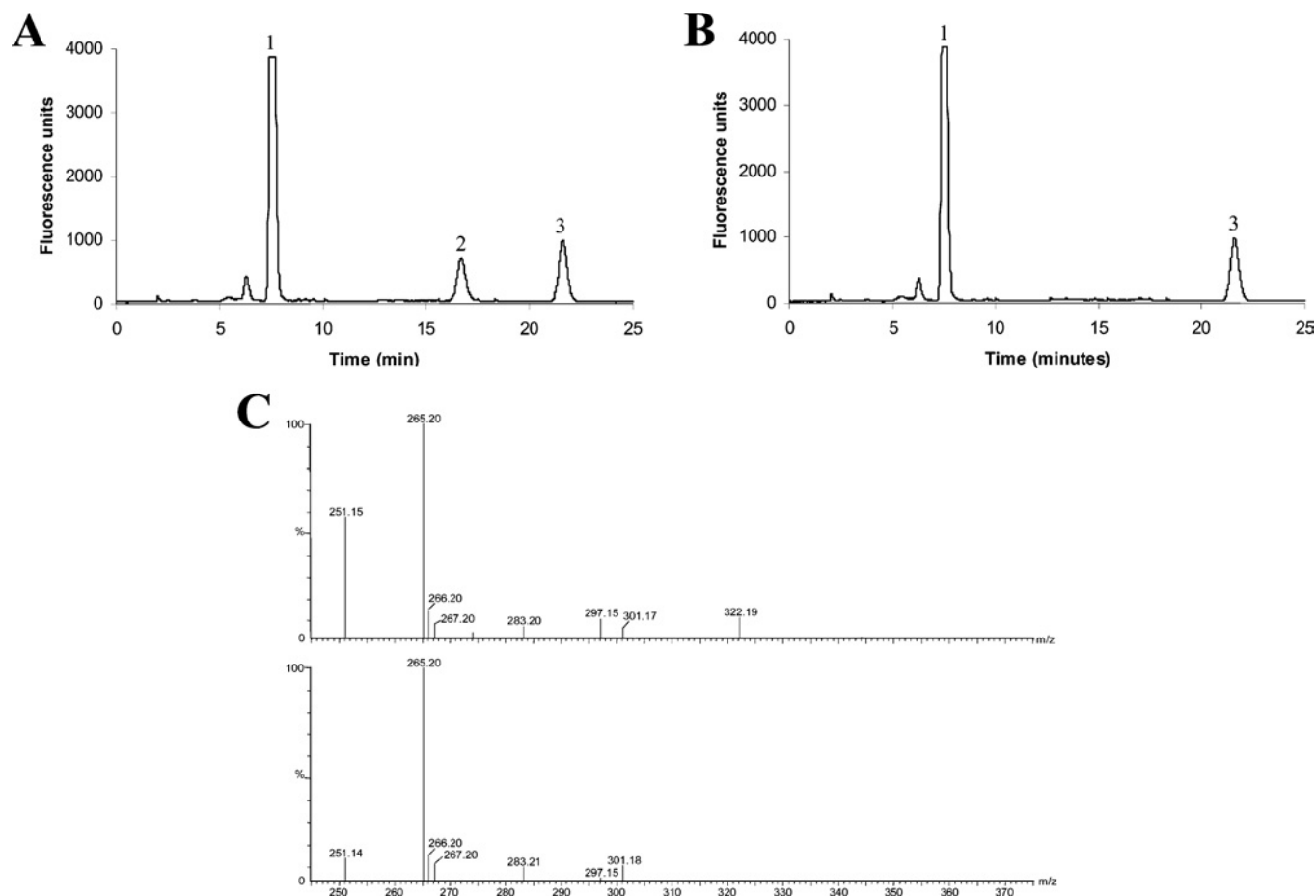
(A) Structures of substrates utilized *in vitro* by recombinant GS proteins from either *G. max* L., *T. aestivum* L. or *Z. mays* L. (B) Examples of substrates not utilized *in vitro* by any recombinant GS protein.

shown to be a substrate for several GS enzymes, it may not be a universally acceptable substrate as previous reports showed that this acyl acceptor was tolerated by a partially purified GS from mung bean (*Vigna radiata*) but was not utilized by the respective enzyme from peas [21] or tobacco (*Nicotiana tabacum*) [22]. With the diamino acceptor 2,3-diaminopropionic acid, the site of peptide bond formation was not unequivocally demonstrated. However, based on the activity seen with BAIBA, it would seem most likely that where incorporation occurred the  $\beta$ -amino group was used. The utilization of  $\beta$ - and  $\gamma$ -amino groups was most markedly observed with *GmGS*, but in all cases further extension of the amino-group-bearing side chain beyond C-3 (ornithine) resulted in a total loss of activity [i.e. extension to C-4 (lysine) or C-5 (6-aminohexanoic acid)]. Replacement of the  $\alpha$ -carboxy residue with an alcohol function (e.g. substitution of  $\beta$ -alanine with 3-amino-1-propanol and glycine with ethanolamine) also reduced incorporation, with only *GmGS* and *TaGS2* capable of utilizing these substrates.

#### Active-site model based on chemical mapping

The use of a broad range of acyl acceptors allowed an assessment of the substrate specificity of the S1' binding subsite in these peptide synthetases. For *GmGS*, three broad levels of selectivity were seen. The five substrates showing the highest activity (Table 1), namely glycine,  $\beta$ -alanine, GABA, BAIBA and L-2,3-diaminopropanoic acid, all have an amine substitution on the primary carbon. Unlike all the other GS enzymes examined, while *GmGS* was intolerant to  $\alpha$ -substitution,  $\beta$ -substitution was acceptable as seen with BAIBA and L-2,3-diaminopropanoic acid. A secondary level of activity (21–150 pkat  $\cdot$  mg<sup>-1</sup>), roughly an order of magnitude lower, was also observed for the simple

substituted ethylamines [X-(CH<sub>2</sub>)<sub>2</sub>-NH<sub>2</sub>, where X = OH, CH<sub>2</sub>OH or CH<sub>2</sub>CH(NH<sub>2</sub>)CO<sub>2</sub>H], indicating that a certain amount of structural diversity and/or branching can be tolerated at the  $\gamma$ -position relative to the NH<sub>2</sub> group. However, acceptors beyond a C-4 chain length (e.g. 6-aminohexanoic acid, Lys) are too large to occupy the *GmGS* subsite. For *TaGS1*, only one high-level acyl acceptor was identified in the screen, namely glycine (1620 pkat  $\cdot$  mg<sup>-1</sup>). However, several moderate level acceptor substrates were turned over at a rate that was one order of magnitude lower (10–107 pkat  $\cdot$  mg<sup>-1</sup>). From this pattern of substrate usage, it was concluded that *TaGS1* must have a relatively short and narrow binding site since it could tolerate only minor  $\alpha$ -branching tolerance (L-Ala), was stereoselective for the L-configuration (L-Ala not D-Ala) and had weak  $\beta$ -branching tolerance (BAIBA and L-2,3-diaminopropanoic acid), but had acceptance of unbranched  $\omega$ -aminocarboxylic acids up to C-3 in length (Gly,  $\beta$ -Ala and GABA). For *TaGS2*, a broad substrate tolerance was seen. The optimal substrates (Table 1) Gly,  $\beta$ -Ala, L-Ala and GABA demonstrated tolerance of  $\alpha$ -branching, even for the bulkier ethyl substituent, as seen in the activity of this enzyme towards AABA. Although an accompanying  $\alpha$ -stereoselectivity was observed, this was relatively weak, with the ratio of L-Ala turnover to D-Ala turnover being 5:1. Limited activity (7–118 pkat  $\cdot$  mg<sup>-1</sup>) was seen towards all other amines tested except those beyond a C-4 chain length, which were not tolerated, as was the case with all GS enzymes tested. Moderate  $\beta$ -branching tolerance was observed although this was substantially less than that shown by *GmGS*. *ZmGS* displayed high-level activity for just Gly and  $\beta$ -Ala (Table 1) with lesser activity (65–177 pkat  $\cdot$  mg<sup>-1</sup>) observed for  $\alpha$ -branched L- and D-Ala. Moderate activity towards GABA but not towards longer carbon chain nucleophiles indicated a shorter C-3-chain-binding subsite. Some weak tolerance of



**Figure 4** HPLC and MS analyses of a GS assay with a non-natural co-substrate

(A) HPLC fluorescence trace of *S*-bimane-derivatized reaction products of purified recombinant *TaGS2* assay using  $\gamma$ -GC and D-alanine as substrates. Peak 1, excess  $\gamma$ -EC in the assay. Peak 2, novel reaction product ( $\gamma$ -Glu-Cys-D-Ala), which was not observed when *TaGS2* was incubated with  $\gamma$ -EC minus D-alanine. Peak 3, bromobimane degradation product. (B) Control incubation of (A), with *TaGS2* incubated with  $\gamma$ -GC only. (C) Upper and lower traces: mass spectral analysis of total reaction mixture shown in (A) and (B) respectively before derivatization with monobromobimane. The ( $M + H$ )<sup>+</sup> at 251.1 corresponds to unchanged  $\gamma$ -GC, whereas the novel peak at 322.1 present in the upper trace but not in the lower trace corresponds to  $\gamma$ -Glu-L-Cys-D-Ala in trace (A).

$\beta$ -branching was also observed with L-2,3-diaminopropanoic acid, but not with BAIBA.

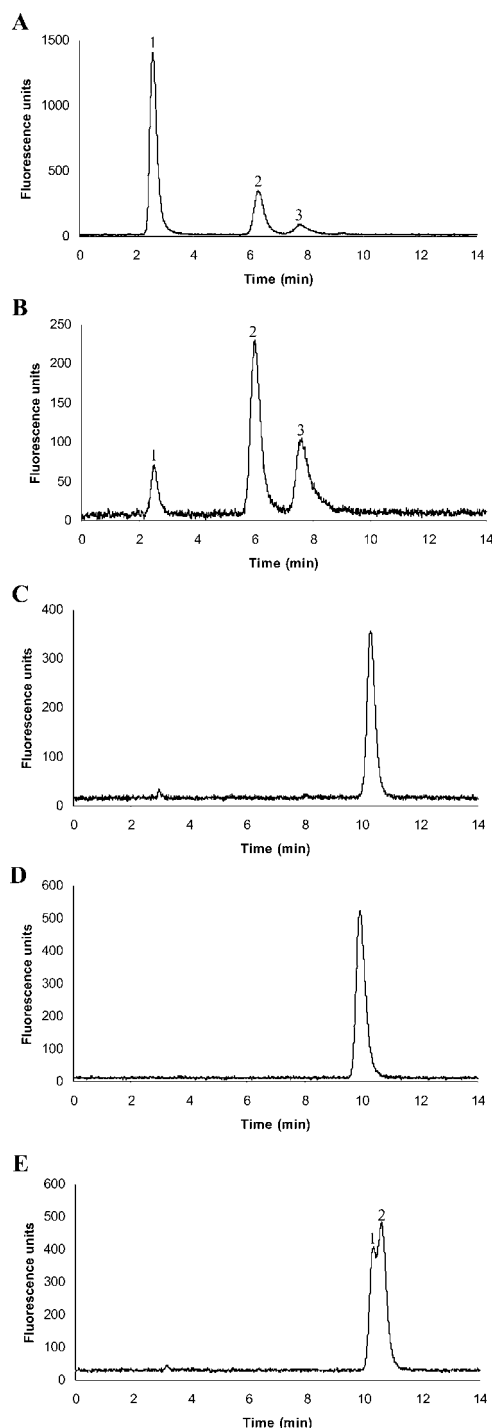
#### D-Serine as a substrate for *TaGS*

The lack of activity of *TaGS1* and *TaGS2* with L-serine suggested that neither enzyme was responsible for catalysing the direct formation of the all L-configuration hmGSH (i.e.  $\gamma$ -L-Glu-L-Cys-L-Ser). However, intriguingly, the activity seen with D-serine in *TaGS2* demonstrated that it would be possible to synthesize the D-serine variant of hmGSH. Such a possibility could not be discounted, since the configuration of serine in hmGSH was not reported during its original identification in wheat plants [5]. Alternatively,  $\gamma$ -L-Glu-L-Cys-L-Ser may be synthesized in wheat by the *TaGS* after inversion of the D-serine to L-serine during or after ligation. An analogous co-ligation/epimerization is seen during tripeptide synthesis catalysed by ACV synthetase [L- $\delta$ -( $\alpha$ -amino adipoyl)-L-cysteinyl-D-valine synthetase], a multifunctional enzyme, involved in penicillin biosynthesis [23]. By analogy to the synthesis of glutathione, ACV synthetase combines both  $\gamma$ -ECS and GS activities by sequentially synthesizing L- $\delta$ -( $\alpha$ -amino adipoyl)-L-cysteine and then incorporating valine via a peptide bond. During this latter incorporation, L-valine is inverted to the D-configuration. The possibility that the incorpor-

ation of D-serine was associated with its epimerization was examined by preparing 1  $\mu$ mol of the putative tripeptide reaction product. After *S*-derivatization with bromobimane, the conjugate was digested with  $\gamma$ -glutamyltranspeptidase and the product co-chromatographed with *S*-bimane derivatives of authentic L-Cys-L-Ser and L-Cys-D-Ser. The digested Cys-Ser-bimane product formed by *TaGS2* co-chromatographed with the corresponding L-Cys-D-Ser-bimane standard rather than the L-Cys-L-Ser derivative, confirming that the serine had not undergone epimerization during ligation (Figure 5). hmGSH present in wheat plants was similarly *S*-bimane-derivatized and treated with  $\gamma$ -glutamyltranspeptidase after its purification by HPLC and was found to contain L-serine rather than D-serine.

#### DISCUSSION

GSs are members of the ATP-grasp superfamily of ligases forming amide bonds in post-translationally synthesized peptides [24,25]. Family members include D-alanine:D-alanine ligase, ribosomal protein S6:glutamate ligase and an  $\alpha$ -L-glutamate ligase involved in the biosynthesis of the coenzyme tetrahydroscarinapterin [25]. The usage of differing amino acids in peptide formation has arisen through gene evolution, but our studies suggest that there



**Figure 5** HPLC analysis of wheat hmGSH and *TaGS2* D-serine reaction product

(A) HPLC trace of  $\gamma$ -glutamyltranspeptidase digest, run on a Chirobiotic V column, of the purified product of *TaGS2* D-serine assay. MS identified peak 1 as undigested parent compound (S-bimane-derivatized  $\gamma$ -Glu-Cys-Ser:  $m/z^+$  528.18), peak 2 as the S-bimane-derivatized Cys-Ser dipeptide product of the digest ( $m/z^+$  399.13) and peak 3 as S-bimane-derivatized cysteine ( $m/z^+$  312.10). (B) HPLC trace of  $\gamma$ -glutamyltranspeptidase digest, run on a Chirobiotic V column, of hmGSH purified from wheat. MS confirmed the identity of each peak to be of identical mass to that shown in (A). (C) HPLC trace of purified peak 2 in (A), co-injected with S-bimane-derivatized L-Cys-D-Ser standard, on a Kingsorb C18 column. (D) HPLC trace of purified peak 2 in (B), co-injected with S-bimane-derivatized L-Cys-L-Ser standard, on a Kingsorb C18 column. (E) HPLC trace of S-bimane-derivatized L-Cys-L-Ser standard and S-bimane-derivatized L-Cys-D-Ser standard co-injected on to a Kingsorb C18 column. Peak 1 is S-bimane-derivatized L-Cys-L-Ser and peak 2 is S-bimane-derivatized L-Cys-D-Ser.

is already latent diversity in the ability of the GS enzymes to synthesize diverse tripeptides. Thus, by expressing and assaying four recombinant GSs from three crop species, we have demonstrated differing substrate specificities and determined that these enzymes can catalyse the synthesis of a broad range of novel tripeptide products. This suggests that the nature of the thiol tripeptides formed *in planta* must be determined primarily by acyl acceptor availability in specific compartments, namely a readily available supply of glycine in maize and wheat and  $\beta$ -alanine in soya bean respectively. This is particularly relevant, since several of the amine acceptors tested in the present study are known to exist in plants and could therefore be incorporated into novel tripeptides. Thus GABA is a significant component of the free amino acid pool in plant species [26] and has been reported at concentrations as high as  $1\text{--}2\ \mu\text{mol}\cdot(\text{g of FW})^{-1}$ , in soya bean leaves after biotic stress [27]. Similarly, AABA occurs in legumes, notably *Lens* sp. [28], and ethanolamine is known to accumulate in plants as an intermediate in choline synthesis after serine decarboxylation [29]. It is therefore possible that, if present in sufficiently high concentrations in a compartment along with a GS, these alternative amine acceptors could be incorporated into glutathione homologues in selected plant species. In the case of hGS, this diversification in the use of the C-terminus has resulted in the preferential incorporation of  $\beta$ -alanine and the resultant accumulation of hGSH in many legume species. The reasons for the selective use of  $\beta$ -alanine over glycine in hGS have received some attention. All cloned hGS enzymes to date, including the GS cloned from soya bean in the present study (Figure 1), have a substitution of leucine and proline for two alanine residues found in the putative glycyl binding domain of conventional GSs [30]. It has recently been demonstrated that mutation of the leucine alone, or both the leucine and proline residues, to the respective alanine residues (Figure 1) enhances the ability of an *M. truncatula* hGS to use glycine and synthesize glutathione [7]. However, as determined by comparison of the kinetic constants, *GmGS* exhibits a much greater preference for utilizing  $\beta$ -alanine over glycine as compared with the hGS of *M. truncatula* [7], suggesting that these two residues in the putative active site cannot be the only factors in determining hGS rather than GS specificity. Interestingly, when assayed with four structural isomers of aminobutyric acid, *GmGS* showed a marked preference for those substrates in which the amino group is on a terminal carbon rather than the  $\alpha$ -carbon next to the carboxy group, which helps explain the preference for  $\beta$ -alanine over L-alanine. The kinetic mechanism of the *Arabidopsis thaliana* GS has recently been determined and shown to have a random Ter-reactant mechanism [31]. Studies with the *Arabidopsis* GS demonstrated that the order of binding of the first pair of substrates determined the affinity of binding to the third substrate. It would now be of interest to extend these studies to other acyl acceptors using different plant GSs.

Although it is possible that additional *TaGS* and *ZmGS* clones, encoding enzymes that could catalyse the incorporation of serine and glutamic acid respectively could have been overlooked in our screen, the available evidence increasingly suggests that hmGSH and  $\gamma$ -ECE arise from post-synthetic modifications of glutathione rather than from unique GSs that catalyse the ligation of serine rather than glycine at the C-termini. In the case of wheat, it has been suggested that hmGSH could arise by transpeptidation on the basis of the demonstration of hmGSH formation when a yeast carboxypeptidase was incubated with glutathione and L-serine [32]. Similarly,  $\gamma$ -ECE may well arise from the proteolytic transpeptidation of phytochelatin, which are polymers of  $(\gamma\text{-Glu-Cys})_n$  formed by plants on exposure to heavy metals [6,33].

Our studies demonstrate that, among the ATP-grasp ligase superfamily [24,25], the plant GSs are flexible peptide synthetases.

While these enzymes are limited to synthesizing  $\gamma$ -glutamyl linked glutathione analogues, this is of interest in regulating the activity of glutathione-dependent enzymes. For example, the overexpression of glutathione S-transferases in tumours is associated with the onset of multidrug resistance, and as such there has been considerable research effort to prepare inhibitory glutathione analogues for therapy [34]. It will now be of interest to screen the library of novel tripeptides produced by plant GSs for their ability to inhibit glutathione S-transferases and other glutathione-dependent proteins.

M. S. acknowledges financial support from the University of Durham through an Addison Wheeler Fellowship. We thank Dr J. Sanderson (Department of Chemistry, University of Durham) for the synthesis of L-Cys-L-Ser and L-Cys-D-Ser standards.

## REFERENCES

- Noctor, G., Arisi, A. C. M., Jouanin, L., Valadier, M. H., Roux, Y. and Foyer, C. H. (1997) Light-dependent modulation of foliar glutathione synthesis and associated amino acid metabolism in poplar overexpressing gamma-glutamylcysteine synthetase. *Planta* **202**, 357–369
- May, M. J., Vernoux, T., Leaver, C., Van Montagu, M. and Inzé, D. (1998) Glutathione homeostasis in plants: implications for environmental sensing and plant development. *J. Exp. Bot.* **49**, 649–667
- Klapheck, S. (1988) Homoglutathione: isolation, quantification and occurrence in legumes. *Physiol. Plant.* **74**, 727–732
- Frendo, P., Gallesi, D., Turnbull, R., Van de Sype, G., Herouart, D. and Puppo, A. (1999) Localisation of glutathione and homoglutathione in *Medicago truncatula* is correlated to a differential expression of genes involved in their synthesis. *Plant J.* **17**, 215–219
- Klapheck, S., Chrost, B., Starke, J. and Zimmermann, H. (1992) Gamma-glutamylcysteinylserine – a new homologue of glutathione in plants of the family Poaceae. *Bot. Acta* **105**, 174–179
- Meuwly, P., Thibault, P. and Rauser, W. E. (1993) Gamma-glutamylcysteinylglutamic acid – a new homolog of glutathione in maize seedlings exposed to cadmium. *FEBS Lett.* **336**, 472–476
- Frendo, P., Jimenez, M. J. H., Mathieu, C., Duret, L., Gallesi, D., Van de Sype, G., Herouart, D. and Puppo, A. (2001) A *Medicago truncatula* homoglutathione synthetase is derived from glutathione synthetase by gene duplication. *Plant Physiol.* **126**, 1706–1715
- Iturbe-Ormaetxe, I., Heras, B., Matamoros, M. A., Ramos, J., Moran, J. F. and Becana, M. (2002) Cloning and functional characterization of a homoglutathione synthetase from pea nodules. *Physiol. Plant.* **115**, 69–73
- Matamoros, M. A., Moran, J. F., Iturbe-Ormaetxe, I., Rubio, M. C. and Becana, M. (1999) Glutathione and homoglutathione synthesis in legume root nodules. *Plant Physiol.* **121**, 879–888
- Skipsey, M., Cummins, I., Andrews, C. J., Jepson, I. and Edwards, R. (2005) Manipulation of plant tolerance to herbicides through co-ordinated metabolic engineering of a detoxifying glutathione transferase and thiol co-substrate. *Plant Biotech. J.* **3**, 409–420
- Skipsey, M., Andrews, C. J., Townson, J. K., Jepson, I. and Edwards, R. (1997) Substrate and thiol specificity of a stress-inducible glutathione transferase from soybean. *FEBS Lett.* **409**, 370–374
- Cummins, I., Cole, D. J. and Edwards, R. (1997) Purification of multiple glutathione transferases involved in herbicide detoxification from wheat (*Triticum aestivum* L.) treated with the safener fenchlorazole-ethyl. *Pestic. Biochem. Physiol.* **59**, 35–49
- 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
- Thorn, R., Cummins, I., Dixon, D. P., Edwards, R., Cole, D. J. and Laphorn, A. J. (2002) Structure of a tau class glutathione S-transferase from wheat active in herbicide detoxification. *Biochemistry* **41**, 7008–7020
- Edwards, R., Blount, J. W. and Dixon, R. A. (1991) Glutathione and elicitation of the phytoalexin response in legume cell-cultures. *Planta* **184**, 403–409
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402
- Pedersen, A. G. and Nielsen, H. (1997) Neural network prediction of translation initiation sites in eukaryotes: perspectives for EST and genome analysis. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **5**, 226–233
- Nakai, K. and Kanehisa, M. (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* **14**, 897–911
- Emanuelsson, O., Nielsen, H. and von Heijne, G. (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* **8**, 978–984
- Klapheck, S., Zopes, H., Levels, H.-G. and Bergmann, L. (1988) Properties and localization of the homoglutathione synthetase from *Phaseolus coccineus* leaves. *Physiol. Plant.* **74**, 733–739
- Macnicol, P. K. (1987) Homoglutathione and glutathione synthetases of legume seedlings: partial purification and substrate specificity. *Plant Sci.* **53**, 229–235
- Hell, R. and Bergmann, L. (1988) Glutathione synthetase in tobacco suspension cultures: catalytic properties and localization. *Physiol. Plant.* **72**, 70–76
- Byford, M. F., Baldwin, J. E., Shiau, C. Y. and Schofield, C. J. (1997) The mechanism of ACV synthetase. *Chem. Rev.* **97**, 2631–2649
- Dinescu, A., Cundari, T. R., Bhansali, V. S., Luo, J. L. and Anderson, M. E. (2004) Function of conserved residues of human glutathione synthetase – implications for the ATP-grasp enzymes. *J. Biol. Chem.* **279**, 22412–22421
- Li, H., Xu, H. M., Graham, D. E. and White, R. H. (2003) Glutathione synthetase homologs encode alpha-L-glutamate ligases for methanogenic coenzyme F-420 and tetrahydroarcanterin biosyntheses. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9785–9790
- Shelp, B. J., Bown, A. W. and McLean, M. D. (1999) Metabolism and functions of gamma-aminobutyric acid. *Trends Plant Sci.* **4**, 446–452
- Wallace, W., Secor, J. and Schrader, L. E. (1984) Rapid accumulation of gamma-aminobutyric acid and alanine in soybean leaves in response to an abrupt transfer to lower temperature, darkness, or mechanical manipulation. *Plant Physiol.* **75**, 170–175
- Rozan, P., Kuo, Y. H. and Lambein, F. (2001) Amino acids in seeds and seedlings of the genus *Lens*. *Phytochemistry* **58**, 281–289
- Rontein, D., Nishida, I., Tashiro, G., Yoshioka, K., Wu, W. I., Voelker, D. R., Basset, G. and Hanson, A. D. (2001) Plants synthesize ethanolamine by direct decarboxylation of serine using a pyridoxal phosphate enzyme. *J. Biol. Chem.* **276**, 35523–35529
- Polekhina, G., Board, P. G., Gali, R. R., Rossjohn, J. and Parker, M. W. (1999) Molecular basis of glutathione synthetase deficiency and a rare gene permutation event. *EMBO J.* **18**, 3204–3213
- Jez, J. M. and Cahoon, R. E. (2004) Kinetic mechanism of glutathione synthetase from *Arabidopsis thaliana*. *J. Biol. Chem.* **279**, 42726–42731
- Okumura, R., Koizumi, Y. and Sekiya, J. (2003) Synthesis of hydroxymethylglutathione from glutathione and L-serine catalysed by carboxypeptidase Y. *Biosci. Biotech. Biochem.* **67**, 434–437
- Meuwly, P., Thibault, P., Schwan, A. L. and Rauser, W. E. (1995) Three families of thiol peptides are induced by cadmium in maize. *Plant J.* **7**, 391–400
- Lucente, G., Luisi, G. and Pinnen, F. (1998) Design and synthesis of glutathione analogues. *Il Farmaco* **53**, 721–735

Received 3 May 2005/29 June 2005; accepted 11 July 2005

Published as BJ Immediate Publication 11 July 2005, doi:10.1042/BJ20050718