Expanding the Plant GSTome Through Directed Evolution: DNA Shuffling for the Generation of New Synthetic Enzymes With Engineered Catalytic and Binding Properties

Evangelia G. Chronopoulou¹, Anastassios C. Papageorgiou², Farid Ataya³, Irini Nianiou-Obeidat⁴, Panagiotis Madesis⁵ and Nikolaos E. Labrou*¹

¹ Laboratory of Enzyme Technology, Department of Biotechnology, School of Food, Biotechnology and Development, Agricultural University of Athens, Athens, Greece, ² Turku Centre for Biotechnology, University of Turku and Abo Akademi University, Turku, Finland, ³ Department of Biochemistry, College of Science, King Saud University, Riyadh, Saudi Arabia, ⁴ Laboratory of Genetics and Plant Breeding, School of Agriculture, Forestry and Natural Environment, Aristotle University of Thessaloniki, Thessaloniki, Greece, ⁵ Institute of Applied Biosciences, Centre for Research and Technology Hellas (CERTH), Thessaloniki, Greece

Glutathione transferases (GSTs, EC. 2.5.1.18) are inducible multifunctional enzymes that are essential in the detoxification and degradation of toxic compounds. GSTs have considerable biotechnological potential. In the present work, a new method for the generation of synthetic GSTs was developed. Abiotic stress treatment of Phaseolus vulgaris and Glycine max plants led to the induction of total GST activity and allowed the creation of a GST-enriched cDNA library using degenerated GST-specific primers and reverse transcription-PCR. This library was further diversified by employing directed evolution through DNA shuffling. Activity screening of the evolved library led to the identification of a novel tau class GST enzyme (PvGmGSTUG). The enzyme was purified by affinity chromatography, characterized by kinetic analysis, and its structure was determined by X-ray crystallography. Interestingly, PvGmGSTUG displayed enhanced glutathione hydroperoxidase activity, which was significantly greater than that reported so far for natural tau class GSTs. In addition, the enzyme displayed unusual cooperative kinetics toward 1-chloro-2,4-dinitrochlorobenzene (CDNB) but not toward glutathione. The present work provides an easy approach for the simultaneous shuffling of GST genes from different plants, thus allowing the directed evolution of plants GSTome. This may permit the generation of new synthetic enzymes with interesting properties that are valuable in biotechnology.

Keywords: glutathione transferase, directed evolution, DNA shuffling, protein engineering, synthetic biotechnology

INTRODUCTION

GSTs are multifunctional enzymes that have evolved from a thioredoxin-like ancestor gene (Mannervik, 2012; Labrou et al., 2015). They are involved in different functions such as the detoxification, metabolism, and transport or sequestration of a wide range of endogenous or xenobiotic compounds. GSTs catalyze the nucleophilic attack of reduced GSH (γ-Glu–Cys–Gly)
on the electrophilic center of these compounds, leading to the formation of GSH conjugates that display higher solubility and reduced toxicity (Deponte, 2013; Labrou et al., 2015; Perperopoulou et al., 2018).

The majority of cytoplasmic GSTs forms dimers of two identical or different subunits of 23–30 kDa (Labrou et al., 2015; Pégéot et al., 2017). Each subunit displays two ligand-binding sites: a G-site and an H-site. The GSH binds with high specificity to the G-site, which is conserved and is located at the N-terminal domain of the polypeptide. The H-site is the binding site for the electrophilic substrate. It is less conserved and determines the affinity and specificity of GSTs toward the electrophile substrates (Labrou et al., 2015; Pégéot et al., 2017). An induced-fit mechanism has been proposed to facilitate the binding and accommodation of the substrates (GSH and electrophile substrate) to the G- and H-sites (Neuefeind et al., 1997; Axarli et al., 2009a,b).

GSTs are expressed both constitutively and in response to biotic and abiotic stresses including herbicides, herbicide safeners, temperature, chill, drought, light, heavy metals, pathogens, and others (Skipsy et al., 2011; Kissoudis et al., 2015; Islam et al., 2017, 2018; Nianiou-Obeidat et al., 2017; Skopelitou et al., 2017). GSTs are encoded by a large and diverse gene family in plants, which is termed the GSTome. The GSTome differs in the number of GSTs, herbicide specificity, and inducibility in different plants and stress conditions (Liu et al., 2013; Csiszár et al., 2014; Pégéot et al., 2014; Han et al., 2018). Heavy metals and high temperature are also considered as inductors of GST expression and function (Gajewska and Skłodowska, 2008; Wang et al., 2017).

The GSTome consists of the functional GSTs that are encoded and expressed by a genome (Mannervik, 2012). The GSTome comprises the cytosolic, mitochondrial, and microsomal superfamilies. Each superfamily composed by several diverse classes (Nianiou-Obeidat et al., 2017). For example, the cytosolic superfamily in plants has fourteen different classes: tau (U), phi (F), theta (T), zeta (Z), lambda (L), gamma (γ)-subunit of the eukaryotic translation elongation factor 1B (EF1B), dehydroascorbate reductase (DHAR), metaxin, tetrachlorohydroquinone dehalogenase (TCHQD), Ure2p, and microsomal prostaglandin E synthase type 2 (mPGES-2) (Liu et al., 2013; Lallement et al., 2014a,b). Recently, three new classes were identified in plants: hemerythin (GSTH), iota (GSTI), and glutathionyl-hydroquinone reductases (GHRs) (Yang et al., 2014). The tau and phi classes have the largest number of GSTs in plants (Liu et al., 2013; Csiszár et al., 2014; Pégéot et al., 2014; Han et al., 2018). Both classes contribute considerable and play key roles in the detoxification of several classes of herbicides (Edwards and Dixon, 2000; Chronopoulou and Labrou, 2009).

The wide catalytic capabilities of GSTs along with their ideal structural features, such as stability, efficient heterologous expression in E. coli and purification by a single-step affinity chromatography have encouraged their exploitation in different areas of biotechnology (Perperopoulou et al., 2018). For example, selected GST isoenzymes are being exploited for the assembly of enzyme biosensors, which can find application in the measurements of xenobiotics, such as drugs, toxins, and herbicides (Kapoli et al., 2008; Chronopoulou et al., 2012b; Oliveira et al., 2013; Materon et al., 2014). Furthermore, GSTs have been used in nanobiotechnology for the construction of biochips (Voelker and Viswanathan, 2013; Zhang et al., 2013; Zhou et al., 2014), nanowires and nanorings (Bai et al., 2013; Hou et al., 2013). In plant biotechnology, GSTs are useful tools in plant breeding programs for the development of plant varieties with multiple stresses resistant traits. Alternatively, the use of genetic engineering allows the development of transgenic plants with traits beyond the limitation of the existing genetic variability (Kissoudis et al., 2015; Nianiou-Obeidat et al., 2017). There is, therefore, an urgent need to discover new GST isoenzymes with desired properties for the development of new or novel applications. Protein engineering efforts for the design of new enzymes with improved catalytic and structural properties are required (Broo et al., 2002; Kurtovic et al., 2008; Runarsdottir and Mannervik, 2010).

In the present work, DNA shuffling was employed for the design and creation of a library of tau class GSTs (GSTUs) from abiotic stress-treated Phaseolus vulgaris and Glycine max. Screening of the library led to the selection of a new GST variant. The new enzyme was characterized by kinetic analysis and X-ray crystallography. The results demonstrated that random recombination of fragments from homologous GSTUs from different plants can give rise to new functionally synthetic GST enzyme.

**RESULTS AND DISCUSSION**

**Analysis of the Catalytic Diversity of GSTome From P. vulgaris and G. max Under Control and Abiotic Stress Treatments**

Transcriptomics and genomics projects have showed that plants have multiple genes coding for GSTs (Nianiou-Obeidat et al., 2017; Han et al., 2018). For example, in the Glycine max var. Williams 82 genome, 101 gene loci encode putative GSTs (Liu et al., 2015). The analysis of P. vulgaris trascriptomic and genomic data (available at https://phytozome.jgi.doe.gov) reveal the presence of at least 52 transcripts that encode putative GSTs (unpublished results). Plant GSTs are inducible enzymes that respond to biotic and abiotic stresses (Chronopoulou et al., 2012a; Csiszár et al., 2014; Pégéot et al., 2014; Han et al., 2018). In the present study, the induction of total GST activity in P. vulgaris and G. max tissues was evaluated in response to different chemical and physical stress agents to expand the repertoire of differentially expressed GST isoenzymes with diverse catalytic and functional properties.
FIGURE 1 | GST-specific activities of extracts from tissues (leaves, shoots, roots) of control and stressed plants after 24 and 48 h treatment for (A,B, respectively) and Phaseolus vulgaris (C,D, respectively) total activity was measured using CDNB, CuO OH, ethacrynic acid, fluorodifen, and NBD-Cl as substrates. Results represent the means of triplicate determinations, with variation less than 5% in all cases.

FIGURE 2 | Schematic diagram of the experimental approach used for the generation of the synthetic GST gene and its corresponding enzyme PvGmGSTUG (i) Abiotic stress treatments of Phaseolus vulgaris and Glycine max plants lead to induction of total GST activity and allowed the creation of a GST-enriched cDNA library using degenerated GST-specific primers and reverse transcription-PCR; (ii) The GST-enriched library was further diversified employing directed evolution through DNA shuffling; (iii) Activity screening of the evolved library led to the isolation of a novel tau class GST enzyme (PvGmGST), which was purified and characterized.
Given the inducible expression of GSTs under different abiotic stress conditions, young *P. vulgaris* and *G. max* plants were exposed to different abiotic stressors, such as a mixture of different herbicides (atrazine, alachlor, fluazifop-p-butyl), heavy metals (nickel, zinc, and chromium) as well as heat-shock (37°C). The purpose of these combined stress treatments was to invoke the expression of GST activities that are induced only following exposure to abiotic stresses (Kissoudis et al., 2015). Following the treatments, plants were harvested, homogenized, and crude extracts were assayed for GST activities using spectrophotometric assays and a range of different model substrates. Total GST activity was extracted from different plant tissues (leaf, root, and shoot) of both control plants and treated plants and measured using five different substrates: 1-chloro-2,4-dinitrobenzene (CDNB), cumene hydroperoxide (CuOHO), the herbicide fluorodifen, ethacrynic acid, and p-nitrobenzyl chloride (pNBC). The choice of these substrates was based on different chemistries involved in catalytic reactions (e.g., nucleophile substitution, addition, hydroperoxidation) to expand the possibilities of obtaining a broad range of catalytic functionalities. Prior experience has demonstrated that a high proportion of functional GSTs can be obtained by this approach (Chronopoulou et al., 2012a; Li et al., 2017a,b). The results (Figure 1) showed that the application of multiple stress conditions resulted in a large increase in total GST activity. For example, using CDNB as a substrate, a 1.4–2.3-fold increase was observed in different tissues, compared to the control plants in the 48 h treatment. CuOOH and ethacrynic acid produced a 1.1–3.2-fold and 1.1–5.6-fold increase in total GST activity, respectively. The results indicated that following abiotic stress treatment different GST isoenzymes were upregulated in different *P. vulgaris* and *G. max* tissues, suggesting an increased diversity in catalytic activities.

### Shuffling of cDNAs Encoding GSTs From *P. vulgaris* and *G. max* and Selection of a New Variant

The method of DNA shuffling is an effective strategy for generating genetic diversity and for identifying protein variants with improved or altered functional or structural properties. The DNA shuffling protocol consists of the following steps: (i) selection and preparation of genes to be shuffled, (ii) digestion of the selected genes with DNase I for generation of a mixture of DNA fragments (size 50–100 bp), (iii) reassembly of DNA fragments with PCR without primers, and (iv) amplification of reassembled products by a conventional PCR. During the PCR reactions, point mutations may be generated. Abiotic stress treatment of *P. vulgaris* and *G. max* makes them perfect starting materials for producing a cDNA library enriched with GSTs (Figure 2). Thus, RNA from stressed tissues (leaf, root, and shoot treated for 48 h) was reverse transcribed and the GST genes were amplified using PCRs and degenerate primers. The PCR amplicons of putative GST genes were cloned and the resulted recombinant plasmids were isolated, mixed, and used for *in vitro* recombination by DNA shuffling (Zhao and Arnold, 1997; Axarli et al., 2016, 2017). Following *in vitro* recombination, a single PCR product was cloned into the pEXP5-CT/TOPO®-TA plasmid. Different colonies (180 in total) were screened for GST activity. Approximately 46% of the picked colonies exhibited GST activity, suggesting that the recombination produced a library with high proportion of catalytically active GSTs (Figure 3). Interestingly, the mean specific activity was 0.17 U/mg and 52% of the active colonies displayed specific activity higher than 0.1 U/mg. The GST variant that displayed the highest activity was selected for further characterization. This clone was sequenced (Figure 4A) and revealed a 672 bp open reading frame encoding a protein of 224 amino acid residues with a molecular mass of 26,088.08 Da and a theoretical pI of 5.80. BLAST searches showed that both its nucleotide (BLASTN) as well as amino acid sequences (BLASTP) were novel and absent from all public databases (Tables 1, 2, Supplementary Figures 1, 2). The phylogenetic relationship of this new enzyme with other GSTs from all known classes was investigated by the construction of a phylogenetic tree that was generated by multiple amino acid sequence alignment (Figure 4B). The alignment was created using representative members of all classes of the *Glycine max* GST family (*Gm*GSTs) (McGonigle et al., 2000; Liu et al., 2015). The enzyme that resulted from DNA shuffling, which was denoted as *PvGmGSTUG* in accordance with the nomenclature proposed by Edwards et al. (2000), clustered together with the tau class GSTs. Of note, as evident from the data provided in Tables 1, 2, *PvGmGSTUG* displayed the highest homology...
FIGURE 4 | Sequence and phylogenetic analysis (A). Nucleotide and amino acid sequence of the PvGmGSTUG. (B) Phylogenetic analysis of PvGmGSTUG with representative members from all classes of the Glycine max GST family. Sequences were aligned with the CLUSTAL Omega sequence alignment program (Sievers et al., 2011) and the phylogenetic tree was constructed using Geneious 9.1.2 software (http://www.geneious.com; Kearse et al., 2012) with the UPGMA tree building method and iTOL v1.0 software (Ciccarelli et al., 2006). Various classes can be distinguished: Phi (GSTF), Tau (GSTU), Lambda (GSTL), Theta (GSTT), Dehydroascorbate reductase (DHAR), Elongation factor 1Bγ (EF1Bγ), Zeta (GSTZ), and Tetrachloro-hydroquinone dehalogenase (TCHQD). The accession numbers of Glycine max GSTs that were used for this phylogenetic tree are: Phi class: GmGSTF1 (AJE59615.1), GmGSTF2 (AJE59616.1), GmGSTF3 (AJE59618.1), Tau class: GmGSTU1 (AJE59648.1), GmGSTU2 (AJE59647.1), GmGSTU3 (AJE59651.1), Lambda class: GmGSTL1 (AJE59633.1), GmGSTL2 (AJE59634.1), GmGSTL3 (AJE59635.1), Theta class: GmGSTT1 (AJE59641.1), GmGSTT2 (AJE59642.1), GmGSTT3 (AJE59643.1), DHAR class: GmGSTDHAR1 (AJE59631.1), GmGSTDHAR2 (AJE59630.1), GmGSTDHAR3 (AJE59629.1), EF1Bγ class: GmGSTEF1Bgamma1 (AJE59625.1), GmGSTEF1Bgamma2 (AJE59626.1), GmGSTEF1Bgamma3 (AJE59627.1), Zeta class: GmGSTZ2 (AJE59689.1), GmGSTZ1 (AJE59691.1) TCHQD class: GmGSTTCHQD1 (AJE59638.1), GmGSTTCHQD2 (AJE59639.1), and GmGSTTCHQD3 (AJE59640.1).

(87 and 86 % homology at the nucleotide and amino acid level, respectively,) with a GST from Medicago truncatula (nucleotide and amino acid accession codes XM_003623148.2 and XP_003623196.1, respectively), rather than the GSTs from Phaseolus vulgaris and Glycine max, in excellent agreement with their evolution history. This important observation further supports the evolution theory of legume plants (Cronk et al., 2006). Amino acid sequence alignments and phylogenetic analysis of PvGmGSTUG with the tau class GSTs from G. max and P. vulgaris revealed that the PvGmGSTUG displayed higher
identity with the PvtGSTU2-2 and GmGSTU8-8 isoenzymes and, from the evolutionary point of view, formed a separate clade (Figure 5). Although the accurate prediction of the parent sequences was impossible, we could nevertheless speculate that most of the PvtGSTUG sequence was derived from PvtGSTU2-2 (Chronopoulou et al., 2012a) and GmGSTU8-8 (Pouliou et al., 2017).

### Substrate Specificity and Kinetic Analysis of PvtGSTUG Enzyme

Recombinant PvtGSTUG was purified to homogeneity by affinity chromatography on S-hexyl-GSHagarose adsorbent (Figure 6). The substrate specificity of PvtGSTUG was evaluated using a broad range of substrates. The results (Table 3) showed that PvtGSTUGs could catalyze a broad range of reactions. Several halogenated aromatic compounds were acceptable substrates. They included CDNB and its analogs: 1-bromo-2,4-dinitrobenzene (BDNB), 1-iodo-2,4-dinitrobenzene (IDNB), and 4-chloro-7-nitrobenzofurazan. PvtGSTUG was also examined for GST-dependent peroxidase activity (GPOX) using CuOOH, tert-butyldihydroperoxide, and benzoyl peroxide as substrates. Among all the peroxides tested, CuOOH and lauroyl peroxide were the best substrates. PvtGSTUG also catalyzed the conjugation of GSH with isothiocyanates.
FIGURE 5 | Sequence and phylogenetic analysis (A). Amino acid sequence alignments of PvGmGSTUG with the tau class GSTs from Glycine max and Phaseolus vulgaris (B). Phylogenetic analysis of GmPvGSTUG with the tau class GSTs from Glycine max and Phaseolus vulgaris. Phylogenetic tree was constructed by the (Continued)
**FIGURE 5** | neighbor joining method using Geneious v9.1.2 software (Kearse et al., 2012) after alignment of the protein sequences using the Clustal Omega sequence alignment program (Sievers et al., 2011). The figures were created using Geneious v9.1.2 software (Kearse et al., 2012). Conserved areas are shown shaded: ▣ 100% identity, □ 80–100% identity, □ □ 60–80% identity, □ □ □ <60% identity. The accession numbers and gene codes of the GST sequences that were used were: PvGSTU1-1 (AEX38000.1); PvGSTU2-2 (AEX38001.1); PvGSTU3-3 (NP_171792); Phvul.1 (006G023500.1|PACid:27165305); Phvul.2 (008G195500.1|PACid:27155547); Phvul.3 (008G195600.1|PACid:27151113); Phvul.4 (002G080200.1|PACid:27169916); Phvul.5 (005G053000.1|PACid:27149482); Phvul.6 (005G053200.1|PACid:27149239); Phvul.7 (005G054000.1|PACid:27150418); Phvul.8 (code 005G054100.1|PACid:27148744); and Phvul.9 (005G054200.1|PACid:27149131). The accession numbers of Glycine max GST sequences that were used were: GmGSTU1-1, AAA33973; GmGSTU2-2, CA71784; GmGSTU3-3, CA848717; GmGSTU4-4, AAC18566; GmGSTU5-5, AA34795; GmGSTU6-6, AA34796; GmGSTU7-7, AA34797; GmGSTU8-8, AA34798; GmGSTU9-9, AA34799; GmGSTU10-10, AA34800; GmGSTU11-11, AA34801; GmGSTU12-12, AA34802; GmGSTU13-13, AA34803; GmGSTU14-14, AA34804; GmGSTU15-15, AA34805; GmGSTU16-16, AA34806; GmGSTU17-17, AA34807; GmGSTU18-18, AA34808; GmGSTU19-19, AA34809; and GmGSTU20-20, AA34810.

PvGmGST, displayed high catalytic activity toward the aliphatic allyl-isothiocyanate, compared to the aromatic phenethyl-isothiocyanate.

The dependence of catalytic activity of *PvGmGSTUG* enzyme was investigated using steady-state kinetic analysis. The analysis was performed by employing two different model reaction systems: the GSH/CDNB and the GSH/CuOOH (Figure 7). The GSH/CDNB is a typical SN2 nucleophilic substitution reaction whereas the GSH/CuOOH reaction is an oxidative reaction (e.g., hydroperoxidase activity). The results are summarized in Table 4. *PvGmGSTUG* obeyed normal Michaelis-Menten kinetics when GSH was used as a variable substrate in both types of reactions. The unusual low K_m value (K_m 17 ± 1 μM) obtained for GSH in its reaction with CuOOH suggested that the enzyme is able to perform efficient catalysis under physiological conditions where the concentration of GSH is reduced, as for example under oxidative stress (Skopelitiou et al., 2017).

When CDNB was used as the variable substrate, the enzyme showed cooperative allosteric kinetics. A Hill coefficient (n_H) of 1.8 ± 0.1 was measured with CDNB. Previous studies have established that in several *tau* class GSTs, although the H-site of neighboring subunits is remote, a reasonable communication between them exists. For example, in the case of a mutant form of *GmGST4-4* (Axarli et al., 2016) structural examination revealed that Lys104, which is located at the dimer interface, plays a key role in inter-subunit communication as well as in the cooperative allosteric kinetics observed with this enzyme.

**Thermal Stability**

To evaluate whether the simultaneous shuffling of GST genes from different plants allowed the generation of structurally stable GST fold, thermal inactivation and unfolding measurements were achieved as illustrated in Figure 8A. The half-inactivation temperature (T_m) was 45.9 ± 0.2°C, which lies within the expected range for mesophilic enzyme and is close to that determined for other native GST isoenzymes (Skopelitiou et al., 2017; Perperopoulou et al., 2018). This suggests that the *PvGmGSTUG* structure displays normal stability and that no detrimental mutations or insertions were introduced during the shuffling of GST genes.

Differential scanning fluorometry (DSF) was also performed to assess the temperature-induced unfolding of the enzyme. DSF was carried out in the absence (Figure 8B) or presence of different concentrations of the substrate (GSH) and the reaction product [S-(p-nitrobenzyl)-GSH] (Figures 8C,D). The unfolding profile of the free enzyme as well as of the enzyme-GSH complex exhibited a single transition with a symmetric peak, with the maximum fluorescence intensity, corresponding to T_m, at 55 ± 0.1°C (n = 4) (Supplementary Figures 3A-C). On the other hand, in the presence of S-(p-nitrobenzyl)-GSH, an increase of the protein Gibbs free energy of unfolding was observed, which usually is depicted as a T_m shift at higher temperatures (Supplementary Figure 3C) (Lea and Simeonov, 2012). This T_m shift suggested a more stable structure with a closed, compact conformation, compared to that of the free enzyme or the enzyme-GSH complex, an indication of an induced-fit mechanism of *PvGmGSTUG* catalysis (Axarli et al., 2009a; Figure 8B).

**Crystalllographic Analysis and Structural Characterization of *PvGmGSTUG***

To better understand its properties, *PvGmGSTUG* was subjected to structural determination by X-ray crystallography (Figure 9). *PvGmGSTUG* was crystallized with two molecules in the crystallographic asymmetric unit that followed the typical dimer formation found in other GSTs (Axarli et al., 2009a; Pégeot...
### TABLE 3 | Substrate specificity for purified recombinant PrGmGSTUG.

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Structure</th>
<th>Specific activity (U/mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Chloro-2,4-dinitrobenzene</td>
<td><img src="image1.png" alt="Structure" /></td>
<td>14.6</td>
</tr>
<tr>
<td>1-Bromo-2,4-dinitrobenzene</td>
<td><img src="image2.png" alt="Structure" /></td>
<td>6.9</td>
</tr>
<tr>
<td>1-Fluoro-2,4-dinitrobenzene</td>
<td><img src="image3.png" alt="Structure" /></td>
<td>ND</td>
</tr>
<tr>
<td>1-Iodo-2,4-dinitrobenzene</td>
<td><img src="image4.png" alt="Structure" /></td>
<td>0.8</td>
</tr>
<tr>
<td>p-Nitrobenzyl chloride</td>
<td><img src="image5.png" alt="Structure" /></td>
<td>ND</td>
</tr>
<tr>
<td>4-Chloro-7-nitrobenzofurazan</td>
<td><img src="image6.png" alt="Structure" /></td>
<td>4.5</td>
</tr>
<tr>
<td>Cumene hydroperoxide</td>
<td><img src="image7.png" alt="Structure" /></td>
<td>6.64</td>
</tr>
<tr>
<td>t-Butyl hydroperoxide</td>
<td><img src="image8.png" alt="Structure" /></td>
<td>0.5</td>
</tr>
<tr>
<td>Benzoyl peroxide</td>
<td><img src="image9.png" alt="Structure" /></td>
<td>ND</td>
</tr>
<tr>
<td>Trans-2-Nonenal</td>
<td><img src="image10.png" alt="Structure" /></td>
<td>0.07</td>
</tr>
<tr>
<td>Trans-4-Phenyl-3-buten-2-one</td>
<td><img src="image11.png" alt="Structure" /></td>
<td>ND</td>
</tr>
</tbody>
</table>

(Continued)
TABLE 3 | Continued

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Structure</th>
<th>Specific activity (U/mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethacrynic acid</td>
<td><img src="image" alt="Ethacrynic acid" /></td>
<td>1.2</td>
</tr>
<tr>
<td>Fluorodifen</td>
<td><img src="image" alt="Fluorodifen" /></td>
<td>ND</td>
</tr>
<tr>
<td>Allylisothiocyanate</td>
<td><img src="image" alt="Allylisothiocyanate" /></td>
<td>7.3</td>
</tr>
<tr>
<td>Phenethylisothiocyanate</td>
<td><img src="image" alt="Phenethylisothiocyanate" /></td>
<td>1.6</td>
</tr>
<tr>
<td>2-Hydroxyethyl disulfide (2-2-dithiodiethanol)</td>
<td><img src="image" alt="2-Hydroxyethyl disulfide" /></td>
<td>ND</td>
</tr>
<tr>
<td>Dehydroascorbate</td>
<td><img src="image" alt="Dehydroascorbate" /></td>
<td>ND</td>
</tr>
<tr>
<td>Bromosulfophthalein</td>
<td><img src="image" alt="Bromosulfophthalein" /></td>
<td>ND</td>
</tr>
</tbody>
</table>

ND: Non-detectable enzyme activity
Results represent the means of triplicate determinations, with variation less than 5% in all cases.

et al., 2014; Skopelitou et al., 2017). The final structure (Table 5) displayed good geometry with 93.7% of the residues in the preferred and accepted regions of the Ramachandran plot and 6.3% in the disallowed regions. Residues 1–5 in both chains, and the fragments 214–224 (chain B), and 216–224 (chain A) were not included in the structure owing to high disorder. The root mean square deviation in bond length and angle was 0.010 Å and 1.52°, respectively. The analysis revealed that each monomer of PtGSTU consists of two distinct domains: at the N-terminal region a small α/β thioredoxin-like domain with βαβαβαβαβα folding topology is formed. The topology is arranged in the order β2, β1, β3, and β4. At the C-terminal region a large helical domain is formed (Figure 8B). At the end of helix H3 a short linker (residues 79–91) begins that joins the N- and C-terminal domains.

Coulombic surface analysis has previously shown that the G-site exhibits positive electrostatic potential, which may play a key role in -SH ionization of the bound GSH (Labrou et al., 2001). Similarly, the contribution of positively-charged residues in the adjustment of the electrostatic field has also been found in other GSTs (Patskovsky et al., 2000; Chronopoulou et al., 2012a). It is widely accepted that a Ser residue is the catalytic amino acid in GSTs of tau and phi classes (Labrou et al., 2001; Chronopoulou et al., 2012a), and that it stabilizes the
deprotonated form (GS\textsuperscript{−}) of bound GSH (Lo Piero et al., 2009). Structure superposition of \textit{PvGmGSTUG} with \textit{G. max GSTU4-4} (PDB id 2vo4) revealed an rms deviation of 0.92 \textdegree\ A in Cα positions for 131 aligned residues and identified Ser16 as the catalytic residue. However, several changes were found in the vicinity of the active site (Figure 9). The conserved Glu69 and Ser70 correspond to Glu66 and Ser67 that form hydrogen bonds with the γ-Glu moiety of GSH (Figure 10A). The glycyl moiety of GSH interacts with Lys40 in \textit{GmGSTU4-4}. In \textit{PvGmGSTUG}, a Phe residue replaces Lys, a change that could affect the orientation of GSH. Arg18, another conserved residue among tau GST sequences corresponds to Arg21 in \textit{PvGmGSTUG}. Arg18 has been suggested to stabilize the interactions between helices H1 and H4 through a strong electrostatic interaction with Asp103. A similar interaction appears also in \textit{PvGmGSTUG} with Asp105, the structural equivalent residue of Asp103 in \textit{GmGSTU4-4}. Tyr107, a key residue at the H-site of \textit{GmGSTU4-4}, which forms aromatic interactions with the benzyl group of Nb-GSH in the \textit{GmGSTU4-4-Nb-GSH} complex (Axarli et al., 2009a). In \textit{PvGmGSTUG}, it is replaced by a Cys residue, a change that

**FIGURE 7** | Steady-state kinetic analysis (A). Steady-state kinetic analysis of \textit{PvGmGSTUG} using GSH as a variable substrate (I) and CDNB at a fixed concentration. Steady-state kinetic analysis of \textit{PvGmGSTUG} using the CDNB as a variable substrate (II) and GSH at a fixed concentration (B). Steady-state kinetic analysis of \textit{PvGmGSTUG} using GSH as a variable substrate (I) and CuOOH at a fixed concentration. Steady-state kinetic analysis of \textit{PvGmGSTUG} using the CuOOH as a variable substrate (II) and GSH at a fixed concentration. Experiments were performed in triplicate.

**TABLE 4** | Steady-state kinetic parameters of \textit{PvGmGSTUG} for the CDNB/GSH substrate system (A) and for the CuOOH/GSH substrate system (B).

<table>
<thead>
<tr>
<th>Substrate system</th>
<th>K\textsubscript{m} (mM) (GSH)</th>
<th>S\textsubscript{0.5} (mM) (CDNB)</th>
<th>k\textsubscript{cat} (min\textsuperscript{−1}) (GSH)</th>
<th>n\textsubscript{H} (CDNB)</th>
<th>k\textsubscript{cat}/K\textsubscript{m} (mM\textsuperscript{−1} min\textsuperscript{−1}) (GSH)</th>
<th>k\textsubscript{cat}/S\textsubscript{0.5} (mM\textsuperscript{−1} min\textsuperscript{−1}) (CDNB)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong> CDNB/GSH</td>
<td>1.17 ± 0.09</td>
<td>0.88 ± 0.05</td>
<td>194.1 ± 4.85</td>
<td>1.77 ± 0.14</td>
<td>165.9 ± 0.14</td>
<td>217.5</td>
</tr>
<tr>
<td><strong>B</strong> CuOOH/GSH</td>
<td>0.017 ± 0.001</td>
<td>0.34 ± 0.02</td>
<td>29.61 ± 0.32</td>
<td>1.741.5 ± 127.9</td>
<td>87.1 ± 3.08</td>
<td></td>
</tr>
</tbody>
</table>
We report here the first directed-evolution study of GST genes from different plants and provide the first crystal structure of a synthetic GST. The data demonstrate the power of protein engineering and DNA shuffling in developing enzymes with engineered catalytic activities. From the evolutionary point of view, the results show that the recombination of segments from homologous GSTs from different plants can generate synthetic enzymes of practical significance that can be exploited for the creation of more sustainable and environmentally-friendly biocatalysts. The unusual low $K_m$ value obtained for GSH with CuOOH suggests that the enzyme is able to perform efficient catalysis under conditions where the concentration of GSH is low, such as in the case of oxidative stress. This supports the potential for the future application of this enzyme as a genetic tool in agricultural biotechnology for the development of genetically engineered plants with high resistant to stress conditions.

### MATERIALS AND METHODS

#### Materials

All enzyme substrates and antibiotics were obtained from Sigma-Aldrich (USA). The pCR T7/CT-TOPO kit, pEXP5-CT TOPO TA Cloning Kit, DNase I, and SuperScript II reverse transcriptase were purchased from Invitrogen (USA). KAPA Taq and KAPA High fidelity DNA polymerase were purchased from KAPA Biosystems (USA). The miniplasmid isolation kit was purchased from Macherey–Nagel, (Germany). The QIAquick Gel Extraction kit was purchased from Qiagen (USA).

#### Methods

**Plant Growth and Stress Conditions**

*P. vulgaris* and *G. max* seeds were pre-germinated (72 h at 30°C) on distilled water-moistened Whatman 2MM filter paper. After germination, they were transferred to plastic pots containing soil. The plants were grown in a controlled environment (12-h day/12-h night cycle, at 25°C during the day and 21°C during the night at 65% humidity) and watered with deionized water every 4 days. Plants (3–4 weeks after germination with three or four pairs of leaves) were stressed using a three-step protocol. In the first step, plants were sprayed with a mixture of heavy metals consisting of nickel (150 µM), zinc (200 µM), and chromium (50 µM) and left for 24 h. In the second step, a herbicide mixture composed of fluazifop-p-butyl (diluted 1:250), atrazine (0.2 mM), and alachlor (0.2 mM) in ethanol solution (20% v/v) was used to treat plants. In the third step, plants were subjected to heat stress at 37°C for 24 h. Control plants did not receive any treatment. Tissue samples (leaves, shoots, and roots) from treated and control plants were collected after 24 and 48 h.

**GST Activity Measurements in *P. vulgaris* and *G. max* Extracts in Response to Multiple Stresses**

For protein and GST enzyme assays, plant tissues (roots, shoots, leaves) of treated and control plants were ground to a fine powder using a mortar and liquid nitrogen. The ground material was extracted with potassium phosphate buffer (50 mM, pH 6) containing 0.1 mM EDTA and 1% w/v polyvinylpyrrolidone (3:1 buffer volume/fresh weight). The homogenate was subsequently centrifuged at 13,000 × g for

---

**TABLE 5 | Data collection and refinement statistics.**

<table>
<thead>
<tr>
<th>Beamline</th>
<th>ESRF ID23-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>0.9730</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>50.0–3.5 (3.6–3.5)</td>
</tr>
<tr>
<td>Space group</td>
<td>P4$_1$2$_1$2$_1$</td>
</tr>
<tr>
<td>Cell parameters</td>
<td>a, b, c (Å) = 51.51, 227.5, 90</td>
</tr>
<tr>
<td>Total observations/Unique</td>
<td>34,044/7,132</td>
</tr>
<tr>
<td>Completeness</td>
<td>98.6 (99.0)</td>
</tr>
<tr>
<td>$R_{meas}$</td>
<td>0.099 (1.38)</td>
</tr>
<tr>
<td>$CC_{1/2}$</td>
<td>0.998 (0.656)</td>
</tr>
<tr>
<td>Reflections used in refinement (work/free)</td>
<td>6,380/709</td>
</tr>
<tr>
<td>$R_{work}/R_{free}$</td>
<td>0.28/0.34</td>
</tr>
<tr>
<td>Number of non-hydrogen atoms</td>
<td>3,431</td>
</tr>
<tr>
<td>RMS bonds (Å)</td>
<td>0.010</td>
</tr>
<tr>
<td>RMS angles (°)</td>
<td>1.52</td>
</tr>
<tr>
<td>Clashscore</td>
<td>14.4</td>
</tr>
<tr>
<td>B-factor (Å$^2$)</td>
<td>80.5</td>
</tr>
<tr>
<td>PDB id</td>
<td>6GHF</td>
</tr>
</tbody>
</table>

*Numbers in parenthesis refer to the outermost resolution shell.*
10 min (4°C) and the supernatant was used for enzyme activity and protein determinations (Bradford, 1976), using bovine serum albumin as the standard. Enzyme activity was estimated toward CDNB, CuOOH, fluorodifen, ethacrynic acid, and p-nitrobenzyl chloride (Tappel, 1978; Sato, 1995; Dixon et al., 2003; Axarli et al., 2009a).

**Molecular Cloning**

Total RNA from leaves, shoots, and roots was isolated as previously described (Brusslan and Tobin, 1992) and checked by agarose electrophoresis for its integrity. Total RNA was subjected to DNase treatment with the RNase-free DNase. cDNA synthesis was achieved in a total volume of 20 µL using 1–5 µg of total RNA, 0.5 µg oligo(dT)12–18, 10 mM of each dNTP, and sterile water to a final volume of 12 µL. After incubation at 65°C for 5 min, 5 x superscript buffer, 10 mM dithiothreitol, 40 Units RNaseOUT™, and 200 Units reverse transcriptase Superscript II (Invitrogen) were added in a thermocycler, which was operated at 42°C for 50 min and then at 70°C for 15 min.

Amplification of the GST genes by gradient PCR was performed using KapaTaq DNA polymerase and degenerate primers. Degenerated primers (Supplementary Table 1) were used in order to recover known and probably unknown GST sequences from *P. vulgaris* and *G. max*. The degenerated primers were designed based on nucleotide and aminoacid sequence alignments (Lang and Orgogozo, 2012) of theta class GST genes, derived from multiple related species (Axarli et al., 2009a; Han et al., 2018). The primers were designed based on similarities of the nucleotides at the 5’ and 3’ end sequences.

The following conditions were used for all sets of primers (see below) in a PCR volume of 50 µL: 1 µg cDNA, 10 pmol of forward primer, 30 pmol of reverse primer, 100 µM of each dNTP, 5 x KapaTaq buffer, and 1 Unit KAPA Taq DNA polymerase. The program used in the thermocycler was the same for all sets of primers: 94°C for 60 s, Tm annealing 37°C for 90 s (the first 7–10 cycles), 44°C for 90 s (the next 7–10 cycles), 53°C for 90 s (the last 30–40 cycles), and 72°C for 50 s.

The PCR products were analyzed on a 1% (w/w) agarose gel and the corresponding bands were cut out and cleaned using the QIAquick™ Gel Extraction kit (Qiagen), according to the manufacturer’s instructions. The clean PCR products were A-tailed using Taq polymerase before being ligated to the pEXPS-CT vector using the TOPO®-TA Kit (Invitrogen, USA). The recombinant plasmids (pEXPS-CT-GSTs) were used to transform *Escherichia coli* TOP10 cells.

**Preparation of DNA for Shuffling and Construction of GST Gene Library**

Recombinant plasmids (pEXPS-CT-GSTs) were mixed 1:1 in a final volume of 24 µL. The mixture was equilibrated at 15°C and supplemented with 5 µL of DNase buffer (400 mM Tris-HCl pH 8.0, 100 mM MgSO4, and 10 mM CaCl2), 21 µL of TE buffer (10 mM Tris pH 8.3 and 1 mM EDTA), and DNase (0.7 Units). At different times, aliquots of 6 µL were obtained and stop solution (20 mM EGTA, pH 8.0) was added and heated at 65°C for 10 min. Agarose gel electrophoresis 2% (w/w) of the DNase products was performed to check for digestion. Random fragments of 50–100 bp obtained after 8–15 min were selected for the shuffling procedure.

Reassembly of DNA fragments was carried out. The DNA fragments were used in PCR in the presence of 10 x Pfu buffer, 100 µM of each dNTP, and 1.25 units Pfu polymerase. The PCR cycle consisted of denaturation at 94°C for 0.5 min, annealing at 55°C for 1 min, and polymerization at 72°C for 1.65 s per cycle, with 40 repeats of the cycle to amplify the reassembled products. PCR reassembly product (1 µg) was used as template in a second PCR with the degenerate primers (Supplementary Table 1). This PCR contained 10 pmol of each forward primer, 30 pmol of each reverse primer, 10 x Pfu buffer, 100 µM of each dNTP, and 1.25 Units Taq/Pfu DNA polymerases. The reaction consisted of 11 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 30 s, and polymerization at 72°C for 45 s, with as well as of 14 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 30 s, and polymerization at 72°C for 45 s and 25 s per cycle, followed by...
Chronopoulou et al. DNA Shuffling of Plant GSTome

FIGURE 9 | Ribbon representation and Coulombic surface analysis of the PvGmGSTUG dimer. (A) Ribbon representation of the PvGmGSTUG dimer. Each subunit is shown in a different color. The two-fold axis that relates the two subunits is perpendicular to the plane of the page. (B) Ribbon representation of PvGmGSTUG subunit. Helices and strands are shown in orange and cyan, respectively. The helices are labeled. The figures were using CHIMERA (Pettersen et al., 2004). (C) Coulombic surface analysis of PvGmGSTUG dimer. The analysis was carried out using UCSF Chimera (http://www.cgl.ucsf.edu/chimera). The Coulomb electrostatic surface shows regions of neutral (white), positive (blue), and negative (red) charge.

FIGURE 10 | Active site comparison and residue interaction (A). Comparison of PvGmGSTUG and GmGSTU4-4 active sites after superposition. PvGmGSTUG and GmGSTU4-4 residues are depicted and labeled in cyan and gray, respectively. Nb-GSH bound to GmGSTU4-4 is shown (B). Representation of the interaction between Asp105 and Lys102 in the PvGmGSTUG dimer. Asp105 and Lys102 are shown as spheres and are colored according to the atom type. The picture also depicts the location of the H-site in each subunit.

final extension of 10 min at 72°C. The product of this reaction was run on a 1% (w/v) agarose gel, excised, and purified using a QIAquick™ Gel Extraction kit (Qiagen). The extracted product was ligated to a T7 expression vector (pEXP5-CT/TOPO®-TA). The resulting plasmid library was transformed into E. coli TOP10 and E. coli BL21(DE3) cells.

Screening of Library, and Expression and Purification of Recombinant Enzymes

Screening of the library and expression of the recombinant enzymes were carried out as described by Axarli et al. (2016). Enzyme purification was carried out using affinity chromatography on S-hexyl-GSH-Agarose as previously described (Axarli et al., 2009a). Protein purity was judged by SDS PAGE.
**Assay of Enzyme Activity and Kinetic Analysis**

Enzyme assays were carried out as previously described (Axarli et al., 2009a; Skopelitou et al., 2012). The Bradford assay was used for protein determination. Kinetic analysis was performed as described by Axarli et al. (2016).

**Thermal Stability and Inactivation**

Thermal inactivation of purified PrGmGSTUG was performed in potassium phosphate buffer (20 mM, pH 7) for 5 min at different temperatures (15–65°C). The enzyme was subsequently assayed for residual activity (enzyme activity at 4°C was considered 100%). Melting temperatures (T_m) were determined from the plot of relative inactivation (%) vs. temperature (°C). The T_m value corresponds to the temperature at which 50% of the initial enzyme activity is lost after heat treatment.

The thermal stability of PrGmGSTUG was also investigated using DSF on a Real-time PCR StepOne™ instrument (Applied Biosystems, USA). The thermal stability was measured in potassium phosphate buffer (20 mM, pH 7) using the Protein Thermal Shift™ Dye (Applied Biosystems, USA). Fluorescence monitoring was carried out at 10–95°C in increments of 1°C with a ramping rate of 2%. Melting temperatures (T_m) were estimated using the Protein Thermal Shift™ Analysis Software (Applied Biosystems). Ligand-binding analysis was also achieved with DSF in the presence of different concentrations (0.1, 0.5, 1.0 mM) of GSH and S-(p-nitrobenzyl)-GSH under the same heating and buffer conditions.

**Crystalization and Data Analyses**

The protein was crystallized with the hanging drop vapor diffusion method using 2 µL of protein mixed with 2 µL of reservoir solution containing polyethylene glycol 4000, 20% (w/v), sodium succinate 0.2 M, and HEPES-NaOH (0.1 M, pH 7.0). X-ray diffraction data (Table 5) were collected on the ID23-1 beamline at the European Synchrotron Radiation Facility (France) under cryogenic conditions (100 K). Crystals were initially transferred to a reservoir solution containing 20% v/v glycerol for 2 s and then flash-cooled in liquid nitrogen. Diffraction data were processed with XDS (Kabsch, 2010) and scaled with AIMLESS (Evans and Murshudov, 2013).

**Structure Determination, Refinement, and Analysis**

Structure determination was pursued with the molecular replacement method using Phaser (McCoy et al., 2007). The structure of a Ricinus communis GST (PDB ID 4J2F; sequence identity 46.8% with PrGmGSTUG) was employed as the search model after modification with SCULPTOR (Bunkócz and Read, 2011) that truncated side-chains from non-identical residues. Refinement was carried out initially with PHENIX (Adams et al., 2010) and subsequently with REFMAC (Murshudov et al., 2011). The low-resolution refinement options in REFMAC (Kovalskiy et al., 2016) were utilized owing to the limited resolution of the structure. Structure-based sequence alignment was performed with Secondary Structure Matching (Krissinel and Henrick, 2004). The structure was validated using validation tools in COOT and PHENIX. Figures were created with CHIMERA (Pettersen et al., 2004). The structure has been deposited in the Protein Data Bank (PDB id 6GHF).

**AUTHOR CONTRIBUTIONS**

NEL, EGC, and ACP conceived and designed the experiments; EGC and ACP performed the experiments; FA, PM, and IN-O analyzed the data; NEL, EGC, ACP, PM, and IN-O wrote the paper.

**ACKNOWLEDGMENTS**

NEL and FA acknowledge the International Scientific Partnership Program at King Saud University, Saudi Arabia, for funding this research work through ISPPP# 0071.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.01737/full#supplementary-material

**REFERENCES**


Copyright © 2018 Chronopoulou, Papageorgiou, Ataya, Nianiou-Obeidat, Madesis and Labrou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AH is currently co-organizing a Research Topic with one of the authors NEL and confirms the absence of any other collaboration.