Plant glutathione S-transferase classification, structure and evolution

Sasan Mohsenzadeh¹*, Maryam Esmaeili¹, Fateme Moosavi¹, Maryam Shahrtash¹, Babak Saffari¹ and Hassan Mohabatkar¹,²

¹Department of Biology, Faculty of Sciences, Shiraz University, Shiraz 71454, Iran.
²Department of Biotechnology, Faculty of Advanced Sciences and Technologies, Isfahan University, Isfahan, Iran.

Accepted 15 June, 2011

Glutathione S-transferases are multifunctional proteins involved in diverse intracellular events such as primary and secondary metabolisms, stress metabolism, herbicide detoxification and plant protection against ozone damages, heavy metals and xenobiotics. The plant glutathione S-transferase superfamily have been subdivided into eight classes. Phi, tau, zeta, theta, lambda, dehydroascorbate reductase and tetrachlorohydroquinone dehalogenase classes are soluble and one class is microsomal. Glutathione S-transferases are mostly soluble cytoplasmic enzymes. To date, the crystal structures of over 200 soluble glutathione S-transferases, present in plants, animals and bacteria have been resolved. The structures of glutathione S-transferase influence its function. Phylogenetic analysis suggests that all soluble glutathione S-transferases have arisen from an ancient progenitor gene, through both convergent and divergent pathways.

Key words: Glutathione S-transferases (GST), classification, structure, evolution, phylogenetic analysis, xenobiotics.

INTRODUCTION

Glutathione S-transferases (GSTs, E.C.2.5.1.18), or glutathione transferases, are a heterogeneous group of cell detoxifying enzymes, which catalyse the conjugation of tripeptide glutathione (GSH) to electrophilic sites on a wide range of phytotoxic substrates (Kampranis et al., 2000). This group of enzymes is a diverse protein family encoded from multi gene families (Wilce and Parker, 1994; Edwards et al., 2000; Dixon et al., 2002). GST activities have been identified in eukaryotes and prokaryotes analyzed to date (Wilce and Parker, 1994; Wongsantichon and Ketterman, 2005). Some functions of GSTs are herbicide detoxification, signal transduction, plant protection against ozone damages, heavy metals and xenobiotics (Esmaeili et al., 2009). Cytosolic GSTs have been classified in classes ranging from mammals, plants, insects, parasites, fungus, to bacteria (Chelvanayagam et al., 2001; Ketterer, 2001; Wongsantichon and Ketterman, 2005).

Soluble GST is a dimer hydrophobic protein of 50-kDa with isoelectric point of 4-5 (Dixon et al., 2002). Phi, tau, theta and zeta classes of GSTs are dimeric proteins and possess a serine residue in their active sites; dehydroascorbate reductase (DHAR) and lambda classes of GSTs differ from these enzymes in being monomers and having a catalytic cysteine in their active sites. The tetrachloro hydroquinone dehalogenase (TCHQD) protein may also contain a serine in its active site. However, the structural details of plant microsomal GST proteins are not yet available (Basantani and Srivastava, 2007). In the case of phi and tau GSTs, only those subunits from the same class will dimerize. Within a class, however, the subunits can dimerize even if they are quite different in their amino-acids sequences. As determined for the GSTs that are active in the mechanisms of herbicide metabolism in maize and wheat, the ability to form...
heterodimers greatly increases the diversity of the GSTs in plants but the functional significance of these subunits mixing and matching is yet to be determined (Dixon et al., 2002).

Gene analysis and genomics projects indicate that plants have more than 40 genes coding for GSTs and that the proteins share as little as 10% amino acid identity (Wilce and Parker, 1994; McGonigle et al., 2000). The gene organizations of plant GSTs have been very well established and the chromosomal locations of many GST genes have now been identified. The number of exons is different for each GST class. For example, phi class of GST genes contains three exons and the tau class contains two, whereas zeta class, which catalyze isomerase reactions, have three exons in their genes. Many of the GST genes are present in repeating units on plant chromosomes (Edwards et al., 2000; Dean et al., 2003; Basantani and Srivastava, 2007).

**PLANT GST CLASSIFICATION**

Using a classification system based on immunological cross-reactivity and sequence relatedness, soluble mammalian GSTs have been divided into eight classes including alpha, mu, pi, sigma, theta, kappa and zeta (Dixon et al., 1998; Hayes et al., 2005). Following, the purification and cloning of GSTs are active in herbicide detoxification in maize in the 1980s (Timmerman, 1989), it quickly became apparent that the plant enzymes differed significantly in sequence from their mammalian counterparts (Droog, 1997). A new classification system was proposed (Droog et al., 1995; Dixon et al., 1998; Edwards et al., 2000). This classification is based on sequence similarity and has been confirmed by the crystal structures of GSTs (Wilce and Parker, 1994; McGonigle et al., 2000). The GSTs have been classified recently into four classes: Type I (phi), type II (zeta), type III (tau) and type IV (theta). Type I GSTs that comprise the entire class of plant GSTs with herbicide-detoxifying activity, contain three exons and one intron. The other large group, type III, consists mainly of auxin-induced GSTs, with genes containing two exons and two introns. The GSTs have highly divergent type I isoenzymes and have now been placed in a separate class (Droog et al., 1995; Droog, 1997; Hayes and Mellelan, 1999). Type II GSTs have ten exons and nine introns (Droog et al., 1995; Dixon et al., 1998). Recently, a type IV group has been proposed for several Arabidopsis genes with five introns (Dixon et al., 1998; Edwards et al., 2000).

Interestingly, the theta and zeta classes of GST share a high degree of sequence similarity even over long evolutionary periods (Board et al., 1997; Frova, 2003). A monophyletic origin of zeta and theta GSTs have been suggested and is in good agreement with the conservation of intron number, active site residue and with the function of these enzymes in all eukaryotes (Dixon et al., 1998; Frova, 2003). In addition to the existing classes of GSTs (phi, tau, zeta and theta), Arabidopsis also contains outlying members of the superfamily that falls into two distinct groups based on sequence similarity. One group, DHARs, was recently identified in other plants (Jakobsson et al., 1999) and the other group was classified as the new lambda class of GSTs (Droog, 1997; Wongsantichon and Ketterman, 2005). These classes are plant specific and differ from other plant GSTs in being monomeric like glutathione-dependent oxidoreductases rather than as conjugating enzymes (Jakobsson et al., 1999). Although, lambda GSTs do not show any DHAR or other activity normally associated with GSTs, these enzymes do have GSH-dependent thiol transferase activity, as do DHARs and are known to be co-induced with phi and tau GSTs in cereals following exposure to herbicide safeners (Dixon et al., 2002). Finally, plants also contain genes encoding microsomal GSTs, which although unrelated to the main GST superfamily, have similar glutathione-dependent activities. Microsomal GSTs are membrane-associated proteins in glutathione metabolism (Dixon et al., 2002).

As an example, a phylogenetic tree (Figure 1) shows relationships between selective sequences of GSTs in gramineae, an important plant family. The Iranian GST sequences in wheat submitted to NCBI GenBank (Saffari et al., 2007) are included.

**PLANT GST STRUCTURE**

**Soluble GSTs**

To date, the crystal structures of over 200 soluble GSTs, present in the main classes of plants, animals and bacteria have been resolved. At present, structural information about plant GSTs is available for phi GSTs from Arabidopsis (Reinemer et al., 1996), a tau class from wheat (Thom et al., 2002), rice GST1 (OsGSTU1) (Dixon et al., 2003) and maize (Neufeind et al., 1997a, b), and for a zeta-class GST from Arabidopsis (Thom et al., 2001).

Most cytosolic GSTs are enzymatically active as dimers, homo- or heterodimers of subunits ranging from 23 to 30 kDa in size (Frova, 2003). Monomeric forms of cytosolic GST have been demonstrated convincingly in non-mammalian species (Cromer et al., 2002). The only GSTs in plant shown to be active as monomers are Arabidopsis lambda and DHARs (Frova, 2003).

Studies with co-expressed GST subunits in recombinant bacteria suggest that dimerization is spontaneous but restricted to subunits of the same class. In maize, the heterodimerization of tau GSTs, zmGST V and zmGST VI, is similar with their homodimerization, however, heterodimer formation is favoured with the maize phi GSTs zmGST I and zmGST II (Sommer and Böger, 1999).

Despite the low overall level of sequence identity
Figure 1. Phylogenetic tree illustrating relationships between selective DNA sequences of GSTs in gramineae. Our submitted GST sequences in wheat are shown with *. The numbers in parenthesis are NCBI GenBank accession numbers.
across the classes, all the structures have the same basic protein fold (Xiao et al., 1996), which consists of two domains. It is suggestive of a strong evolutionary pressure for conservation of some structural motifs involved at the active site (Sheehan et al., 2001) and a dimeric composition (Öztetik, 2008). The N-terminal domain 1 (approximately residues 1 to 80) is classified as part of the thioredoxin superfamily fold, which also includes glutaredoxin, disulfide-bond formation facilitator and glutathione peroxidase. This domain consists of a βαβαβα structural motif. The core of the domain is composed of three layers with the β-sheet sandwiched between α-helices (α/β/α) (Sheehan et al., 2001). A more variable C-terminal domain (domain II, approximately two-thirds of the protein) that is entirely helical suggests that differences in the C-terminal domain may be responsible for differences in substrate specificity between the GSTs classes (Wilce and Parker, 1994). Moreover, there is a short (5-10 residues) linker region which connects the N-terminal and the C-terminal domains (Frova, 2003).

In the dimeric enzymes, the two subunits are related by a two-fold axis, the N-terminal domain of one subunit facing and interacting with the C-terminal domain of the partner. The protein surfaces engaged in dimerization may be hydrophobic, as in the alpha, mu, pi and phi classes, or hydrophilic as in the theta, sigma, beta or tau classes. Salt bridges between specific residues may additionally help in the stabilization of the quaternary structure (Frova, 2003).

Domain I is a thioredoxin-like fold (βαβαβαβα), which consists of two motifs, the N-terminal and the C-terminal. The former begins with an N-terminal β-strand (β-1), and with α-helix (α-1) and then a second β-strand (β-2) which is parallel to β-1. A loop region leads into a second α-helix (α-2), which is connected to the C-terminal motif. This motif consists of two sequential β-strands (β-3 and β-4), which are antiparallel and which are followed by a third α-helix (α-3) at the C-terminus of the fold. The four β-sheets are essentially in the same plane, with two helices below this plane (α-1 and α-3) and α-2 above it, facing the solvent. The loop that connects α-2 to β-3 features a characteristic proline residue, which is in the less favoured cis conformation and is highly conserved in all GSTs. This is referred to as the cis-Pro loop, which, while playing no direct role in catalysis, appears to be important in maintaining the protein in a catalytically competent structure (Allocati et al., 1999). Domain II is a variable number (four to seven) of α-helices linked to domain I by a short linker sequence (Sheehan et al., 2001).

There are two ligand-binding sites per subunit, which are a specific glutathione-binding site (G-site) constructed mainly from residues of domain I, and the hydrophobic substrate-binding site (H-site), which is formed primarily by residues with non-polar side chains lying in domain II. The two sites together constitute the catalytically active across the classes, all the structures have the same basic protein fold (Xiao et al., 1996), which consists of site. The N-terminal domain is quite conserved, and contains specific residues critical for GSH binding and catalytic activity. Specifically, the conserved Tyr7 of the mammalian alpha, mu and pi classes, and Ser17 of the ubiquitous theta and zeta, of the plant specific phi and tau and of insect delta classes, have a crucial role in the catalytic activation of GSH. The Tyr/Ser hydroxyl group acts as hydrogen bond donor to the thiol group of GSH, promoting the formation and stabilization of the highly reactive thiolate anion, which is the target for nucleophilic attack of an electrophilic substrate (Frova, 2006). By site-directed mutagenesis, the Ser (or Tyr) residues have proven catalytically essential in GST catalysis in different organisms (Öztetik, 2008).

As mentioned earlier, lambda and DHAR classes are the only GSTs shown to be active as monomers. These plant-specific classes, instead of a Ser or a Tyr, have a cysteine at their usual active site positions, a residue that promotes the formation of mixed disulphides with glutathione rather than the formation of the thiolate anion (Frova, 2006).

The G and H sites of the enzyme are quite mobile when the crystal structure is determined, suggesting that the GST subunits undergo significant conformational changes when binding with the substrates. This is demonstrated by the difference in the structure of the apoenzyme of phi zmGSTF as compared with the ternary complexes (containing GSH and substrate) of other phi class of GSTs (Neufeind et al., 1997). A significant induced-fit mechanism for GSH binding has been suggested for other classes of GSTs (Sheehan et al., 2001).

Microsomal GSTs

Most MAPEG proteins are involved in the synthesis of eicosanoids, leukotrienes and prostglandins, catalyzing GSH-dependent transferase or isomerase reactions. Microsomal GSTs have less than 10% sequence identity with the cytosolic GSTs (Frova, 2006). Usually, their subunits are shorter, with an average length of 150 amino acids. Crystallization experiments have been reported for three members of the MAPEG family. These are microsomal glutathione S-transferase 1, microsomal prostaglandin E synthase 1 and leukotriene C4 synthase (Hebert et al., 2005).

Bresell et al. (2005) have so far characterized four transmembrane domains, the amino and carboxyl termini of the protein protruding into the luminal side of the membrane, while putative sites for GSH and substrate binding are located in loops facing the cytosol. The 3D map of mGST1 and the projection structures of LTC4S and pGES1 shows the enzymes as homotrimers (Frova, 2006). Three-dimensional maps of mGST1 also
demonstrate that a large proportion of the protein monomer forms a left-handed four α-helix bundle (Holm et al., 2002).

PLANT GST EVOLUTION

Drug detoxification enzymes have existed in both prokaryotes and eukaryotes for more than 2.5 billion years (Nerbet, 1994; Nerbet and Dieter, 2000; Sheehan et al., 2001). GSTs, as detoxification enzymes, are widely distributed in aerobic organisms and are hypothesized to have evolved in aerobic bacteria for their ability to prevent oxygen toxicity (Mannervik and Danielson, 1988; Pemble and Taylor, 1992). GSTs constitute a very ancient protein superfamily that is thought to have evolved from a thioredoxin-like ancestor in response to the development of oxidative stress (Martin 1995; Sheehan et al., 2001). With the increasing amount of sequence/structure information available, it has become apparent that GSTs are related to other GSH- and cysteine-binding proteins, for example bacterial stringent starvation proteins, plant pathogen/stress resistance proteins, the URE2 protein from Saccharomyces cerevisiae (Rossjohn et al., 1996), eukaryotic translation elongation factor 1y (Sheehan et al., 2001) and macrophage inhibitory factor (Blocki et al., 1992; Sheehan et al., 2001). They share a thioredoxin-like fold, and are associated to stress-related proteins in a wide range of organisms (Board et al., 1997). Based on these evidences, it has been proposed that primordial stress proteins may be the ultimate ancestors of GSTs (Board et al., 1997; Sheehan et al., 2001).

Evolution of these cytosolic enzymes appears to have been through the addition of an all-helical domain after the thioredoxin βαβββα structure. In contrast, the crystal structure of the mitochondrial isoform, GSTK, provides clear evidence for a parallel evolutionary pathway, as the all-helical domain responsible for binding of the second, electrophilic substrate appears to have inserted within the βαβββα core after the βαβ motif (Robinson et al., 2004; Hayes et al., 2005). Moreover, the different mechanisms used to achieve the common N- and C-terminal domains of cytosolic GST illustrate two regions in the thioredoxin/glutaredoxin fold that are under less evolutionary constraint (Robinson et al., 2004).

Gene duplication, followed by exon shuffling, of an ancestral GSH binding protein may have been a mechanism that generated the different catalytic proper-ties of the members of the GST superfamily (Mannervik and Danielson, 1988; Armstrong, 1998; Sheehan et al., 2001). The functional diversification of GSTs in cell metabolism depends on their potential to respond to various xenobiotics; thus, in subfamilies of GSTs, selective pressure is very strong. Additional chromosome gene duplication, unequal crossing over, alternative splicing (around C-terminal region of GSTs) (Armstrong, 1998; Wongsantichon and Ketterman, 2005), swapping and mutagenesis (around N-terminal region of GSTs) lead to more gene distribution and functional heterogeneity of GSTs.

Finally, phylogenetic analysis would suggest that all soluble GSTs have arisen from an ancient progenitor gene, through both convergent and divergent pathways (Wilce and Parker, 1994).

CONCLUSION

As has been mentioned earlier, all soluble GSTs have arisen from an ancient progenitor gene. Nevertheless, it seems that exon shuffling, gene duplication, alternative splicing, swapping, mutagenesis and probably other unknown mechanisms have led to considerable sequence diversification, functional heterogeneity and finally evolution of GSTs. Although, most of the studies on polymorphisms of GSTs have focused on their polymorphisms in animals and particularly in humans (Saadat and Mohabatkar, 2004; Saadat et al., 2004), there are still many unanswered questions about plant GSTs (Scalla and Roulet, 2002): Why do plant species have different GST classes? Do these differences result in strength or weak points in dealing with stressful situations? Why do tau and phi classes outnumber other GST classes? How are plant GST genes regulated? Are there more than eight GST classes in plants? Are there more new members of GST classes in various plant species (Sheehan et al., 2001; Mohsenzadeh et al., 2009).

REFERENCES


Dean JD, Goodwin PH, Hsiang T (2003). Colletotrichum gloeosporioides infection induces differential expression of


