Glutathione Transferases Are Structural and Functional Outliers in the Thioredoxin Fold[†]

Holly J. Atkinson[‡] and Patricia C. Babbitt*[,§]

[‡]*Program in Biological and Medical Informatics and* [§]*Departments of Bioengineering and Therapeutic Sciences and Pharmaceutical Chemistry, University of California, San Francisco, California 94158-2330*

ABSTRACT: Glutathione transferases (GSTs) are ubiquitous scavengers of toxic compounds that fall, structurally and functionally, within the thioredoxin fold suprafamily. The fundamental catalytic capability of GSTs is catalysis of the nucleophilic addition or substitution of glutathione at electrophilic centers in a wide range of small electrophilic compounds. While specific GSTs have been studied in detail, little else is known about the structural and functional relationships between different groupings of GSTs. Through a global analysis of sequence and structural similarity, it was determined that variation in the binding of glutathione between the two major subgroups of cytosolic (soluble) GSTs results in a different mode of glutathione activation. Additionally, the convergent features of glutathione binding between cytosolic GSTs and mitochondrial GST kappa are described. The identification of these structural and functional themes helps to illuminate some of the fundamental contributions of the thioredoxin fold to catalysis in the GSTs and clarify how the thioredoxin fold can be modified to enable new functions.

The glutathione transferases (GSTs)[1] make up a critically important group of enzymes, found in all classes of eukarya and bacteria (*1−3*), including at least 18 GSTs expressed in humans (*4*). GSTs catalyze a broad range of reactions that involve the addition of glutathione (GSH) to substrate compounds, but their archetypal functional role is in enzymatic detoxification of xenobiotics (*5*, *6*). GSTs have additional important roles in cell signaling and other cellular processes. GST pi has been shown to regulate JNK signaling (*7*), and GST mu from mice forms inhibitory complexes with ASK1, another member of the MAP kinase pathway (*8*). Members of the alpha and sigma classes are involved in the biosynthesis of sex steroids and prostaglandins, respectively (*9*, *10*), and mutations or the absence of specific GSTs is associated with numerous human diseases, including Parkinson's and Alzheimer's (*11*), and increased risk of cardiovascular disease (*12*). GSTs are thought to be responsible for resistance to some chemotherapeutic compounds in addition to carcinogenic compounds (*13*).

On the basis of historical conventions, GSTs are grouped into classes designated by Greek letters (e.g., mu, omega, and sigma) that usually have different general substrate profiles, while members of the same class have more subtle differences in substrate recognition (*14*). There are typically many GSTs transcribed within the same organism, and individual GSTs tend to be promiscuous in transforming a set of related compounds (*14*). Structurally, cytosolic GSTs (also known as soluble GSTs and distinguished from the membrane-associated GSTs and the metallo-GSTs) function as dimers; each monomer is composed of a conserved thioredoxin domain containing the glutathione (GSH) binding site followed by a more variable α-helical domain (the GST C domain) containing the binding site for the GSH acceptor substrate. The fundamental theme in GST catalysis is the activation of GSH for transfer to a substrate by the stabilization of the GSH thiolate (*15*).

While specific GSTs have been extensively characterized, it is difficult to compare members of different classes, and there has been little investigation in this area. A clarification of the relationships between classes could help to delineate major functional shifts, yielding useful information for predicting the catalytic capabilities of uncharacterized GSTs. As canonical GSTs incorporate a thioredoxin fold, associated structural and functional themes relate and distinguish classes beyond the basic capability of adding glutathione to electrophilic substrates. Further, we find commonalities between catalysis in the GSTs and the much larger class of all thioredoxin fold enzymes. New insight regarding the contribution of the thioredoxin fold itself was obtained from examination of how mitochondrial GST kappa (*16*), an enzyme from a different superfamily within the thioredoxin fold class, is able to catalyze the same basic molecular function as cytosolic GSTs. When the kappa enzyme was first discovered, it was seen to catalyze the primary diagnostic reaction associated with cytosolic GSTs (*17*), and it was consequently labeled a glutathione transferase. However, later sequence and structural analysis demonstrated that members of the kappa class are far more similar to another superfamily within the thioredoxin fold class, the protein disulfide oxidoreductase DsbA-like enzymes, than to the cytosolic GSTs (*16*, *18*).

Here we analyze the GSTs, using information from sequence and structure, to improve our understanding of how the different classes are related. In addition, we describe two major subgroups

*To whom correspondence should be addressed: UCSF MC 2550, 1700 4th St., San Francisco, CA 94158-2330. Telephone: (415) 476-3784. Fax: (415) 514-9656. E-mail: babbitt@cgl.ucsf.edu.

[1]Abbreviations: GST, glutathione transferase; GSH, glutathione; DsbA, disulfide bond isomerase A; QSOX, quiescin-sulfhydryl oxidase; Y-GST, tyrosine-type GST; S/C-GST, serine/cysteine-type GST; B1, *β*-strand 1; H1, α-helix 1; CDNB, 1-choloro-2,4-dinitrobenzene; HCCA, 2-hydroxychromene-2-carboxylic acid; PDB, Protein Data Bank; rmsd, root-mean-square deviation.
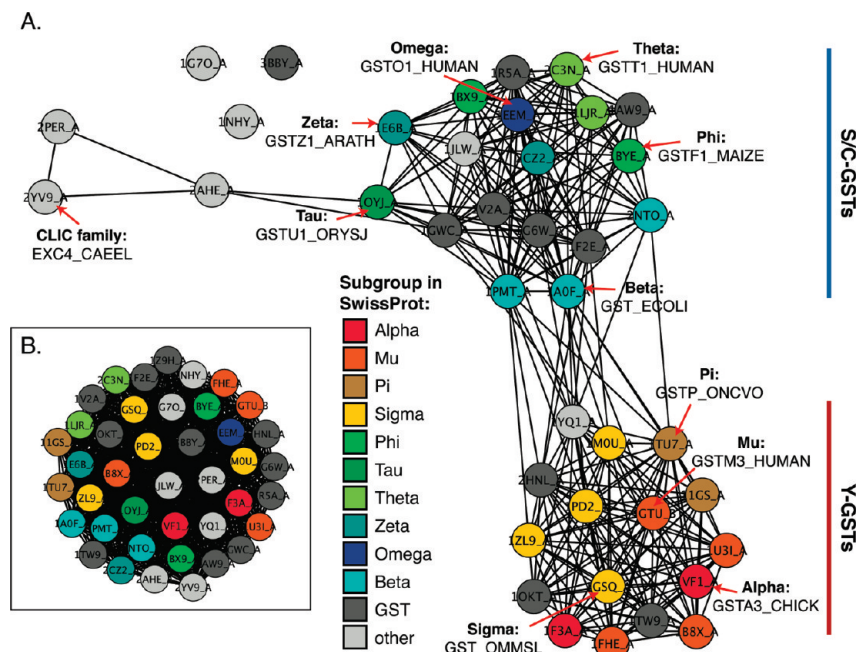
FIGURE 1: Cytosolic GSTs fall into two major subgroups based on structural similarity. (A) Structure similarity network, containing 40 structures that are a maximum of 60% identical (by sequence) that span the cytosolic GSTs and some other proteins with similar structures (including *Escherichia coli* Grx 2, PDB entry 1G7O). Similarity is defined by FAST scores (*25*) better than a score of 16.0; edges at this threshold represent alignments with a median rmsd of 2.5 Å across 177 aligned positions, while the rest of the edges represent better alignments. Each node is colored by the class of the associated sequence in the SwissProt database (as designated in the key, Swissprot uses the term subgroup to designate GSTs normally designated in the literature as classes). Some sequences are only classified to the GST superfamily ("GST"), and others have no classification or one outside of the GST superfamily ("other"). Each structure is labeled with its PDB entry and chain, and examples from each class are labeled with the entry for the associated sequence in SwissProt. (B) Structure similarity network containing the same structures as in panel A, shown at the less stringent threshold of 7.5. Edges at this threshold represent alignments with a median rmsd of 4.5 Å across 153 aligned positions. Nodes are colored as in panel A.

of cytosolic GSTs and reveal an important global theme that characterizes the difference between them. (The designation of "major subgroup" used here should be distinguished from the "subgroup" term that is sometimes used in the literature to reference the Greek letter-labeled classes.) Second, a structural analysis reveals the contributions from convergent evolution and from the thioredoxin fold itself in the similar mode of binding of GSH between cytosolic GSTs and mitochondrial GST kappa.

EXPERIMENTAL PROCEDURES

Data Set Sources and Curation; Annotations. All sequences and structures representing the proteins containing a thioredoxin fold were assembled from the union of the PFAM Thioredoxin-like Clan (CL0172; PFAM release 22.0) (*19*) and all sequences classified into relevant Trx fold superfamilies in SwissProt (*20*). Sequences (drawn from the UniProt Knowledgebase Release 14.0) are members of the Thioredoxin-like Clan if they align with any member HMMs with a score better than the gathering threshold. The 20 relevant SwissProt superfamilies are as follows: FMP46 family, GST superfamily, OST3/OST6 family, SCO1/2 family, SH3BGR family, UPF0413 family, ahpC/TSA family, arsC family, calsequestrin family, chloride channel CLIC family, glutaredoxin family, glutathione peroxidase family, hupG/hyaE family, iodothyronine deiodinase family, nucleoredoxin family, peroxiredoxin 2 family, phosducin family, protein disulfide isomerase family, quiescin-sulfhydryl oxidase (QSOX) family, and thioredoxin family. This union set of all thioredoxin (Trx) fold sequences contains 29206 sequences. The structures in the structure similarity network in Figure 6B are the 159 chains associated with the 29206 sequences mentioned above that were

not theoretical models and had chain sequences that were at most 60% identical to any other chain as determined by cd-hit (*21*). They are annotated with a PFAM classification if their chain sequence aligns with a PFAM family model with a score better than the gathering threshold. The structures in the networks in Figure 1 are the structures that clustered together with the cytosolic GSTs in the network in Figure 6B.

The Trx fold sequences were filtered to a maximum of 40% sequence identity using cd-hit, and sequences shorter than 60 amino acids were discarded, resulting in a data set of 4082 representative sequences. The 622 sequences in the sequence network in Figure 2 are those that clustered with the cytosolic GSTs in a sequence similarity network of the 4082 representative Trx fold sequences thresholded at a BLAST (*22*) E value of 1×10^{-12}.

Catalytic motifs (displayed in Figure 6C) were calculated for the sequence clusters in the 1×10^{-12} Trx fold network containing the GST kappa sequences, the nearest DsbA-like sequences, and the S/C-GSTs. The motifs were calculated by tabulating the amino acids aligning to the "CxxC" motif within the PFAM DSBA and GST_N models for each member of the cluster.

Construction of Networks: Sequence and Structure. The sequence similarity networks were constructed as described previously (*23*), with pairwise similarities between proteins determined using pairwise BLAST alignments (*22*) and resulting networks visualized in Cytoscape 2.6 using the Organic layout (*24*). These BLAST E values are calculated from alignments with a database of homologous sequences, violating the expected background model and rendering these "E value" scores rather than true expected values (*23*). The structure similarity networks were constructed and visualized in the same way, except pairwise
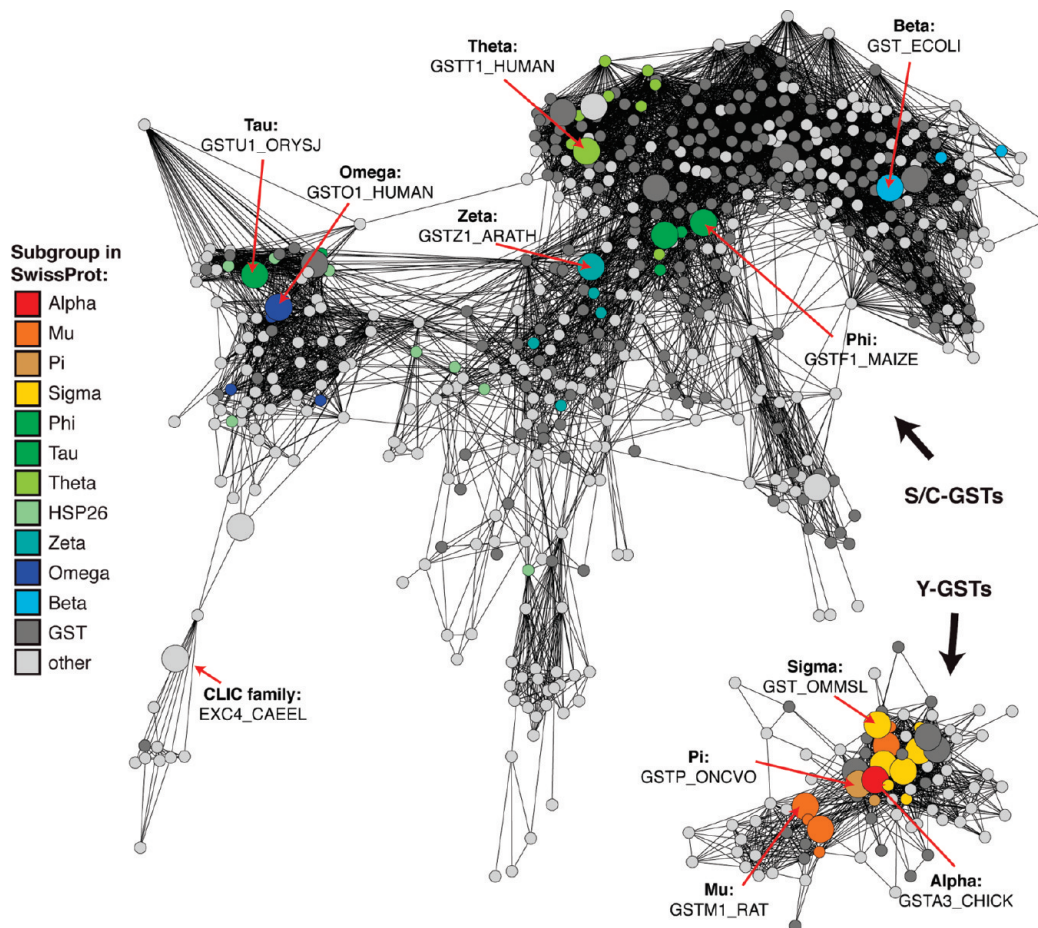
FIGURE 2: Tyrosine-type GSTs are less populated and less diverse than the other major subgroup of GSTs. Sequence similarity network, containing 622 sequences that are a maximum of 40% identical and span the cytosolic GSTs. Similarity is defined by pairwise BLAST alignments better than an E value of 1×10^{-12}; edges at this threshold represent alignments with a median 27% identity over 200 residues, while the rest of the edges represent better alignments. Each node is colored by the classification of the sequence in SwissProt, if available. Large nodes represent sequences that are associated with the 40 structures in Figure 1.

similarity between structure chains was determined using FAST (25).

Interactive versions of the networks provided in the figures are available from http://babbittlab.compbio.ucsf.edu/resources/GST/ and allow visualization of the networks at threshold cutoffs specified by the user. Software for visualization of the networks is freely available from the Cytoscape Consortium (http://www.cytoscape.org/) and can be easily installed and used on personal computers. We note that the thresholds used in visualizing the static networks shown in the figures in this paper were chosen to best illustrate major subgroups and classes from each other. Additional explanation regarding the effects of threshold choice on the detail at which clustering is viewed is provided in ref 23. Additional networks designed to illustrate the relationship of the GSTs with other superfamilies among the thioredoxin fold proteins are provided in ref 26.

Structural Alignments. Representative structures from the GST superfamily were aligned and used to generate a structure-based sequence alignment using the Chimera MatchMaker and Match→Align commands (27) (see Figures 3 and 5). The structures present in the alignment in Figure 5 with a displayed sulfur from GSH or a GSH analogue are as follows: Y-GSTs, alpha (1K3L, 1F3A, 1EV9, 1ML6, 1VF1, 1TDI, and 1B48), mu (1M9A, 1FHE, 2FHE, 6GST, 1GSU, 2C4J, and 1B4P), pi (1PGT), sigma (1M0U, 1ZL9, 2GSQ, 2CVD, and 1PD2), and other (1Q4J); S/C-GSTs, beta (1N2A and 1PMT), omega

(1EEM), phi (1BYE and 1GNW), tau (1OYJ and 1GWC), theta (1PN9 and 1LJR), and zeta (1FW1 and 2CZ2); and glutaredoxin, 3GRX. The GSH binding site alignment between tau 1OYJ and kappa 1R4W (Figure 6) was achieved using the Chimera 'match' command to align the structures according to the bound GSH molecule (27). The predicted H-bonds shown in Figure 6D and all figures depicting structures were calculated or created with UCSF Chimera (28).

All data files generated in the analysis, including sequence files, alignments, and networks, are available (http://babbittlab.compbio.ucsf.edu/resources/GST).

RESULTS AND DISCUSSION

In the following sections, we describe the criteria that support classification of cytosolic GSTs into two major subgroups, first because of overall sequence and structural similarity and second because of the differences in the organization and composition of their active sites. Next is a discussion of how these active site differences affect general aspects of catalysis. Finally, there is a description of how mitochondrial GST kappa binds glutathione using universal aspects of the thioredoxin fold as well as cytosolic GST-like interactions.

Cytosolic GSTs Are Most Appropriately Classified into Two Major Groups. While GSTs have long been classified into a collection of family-like classes, another level of classification
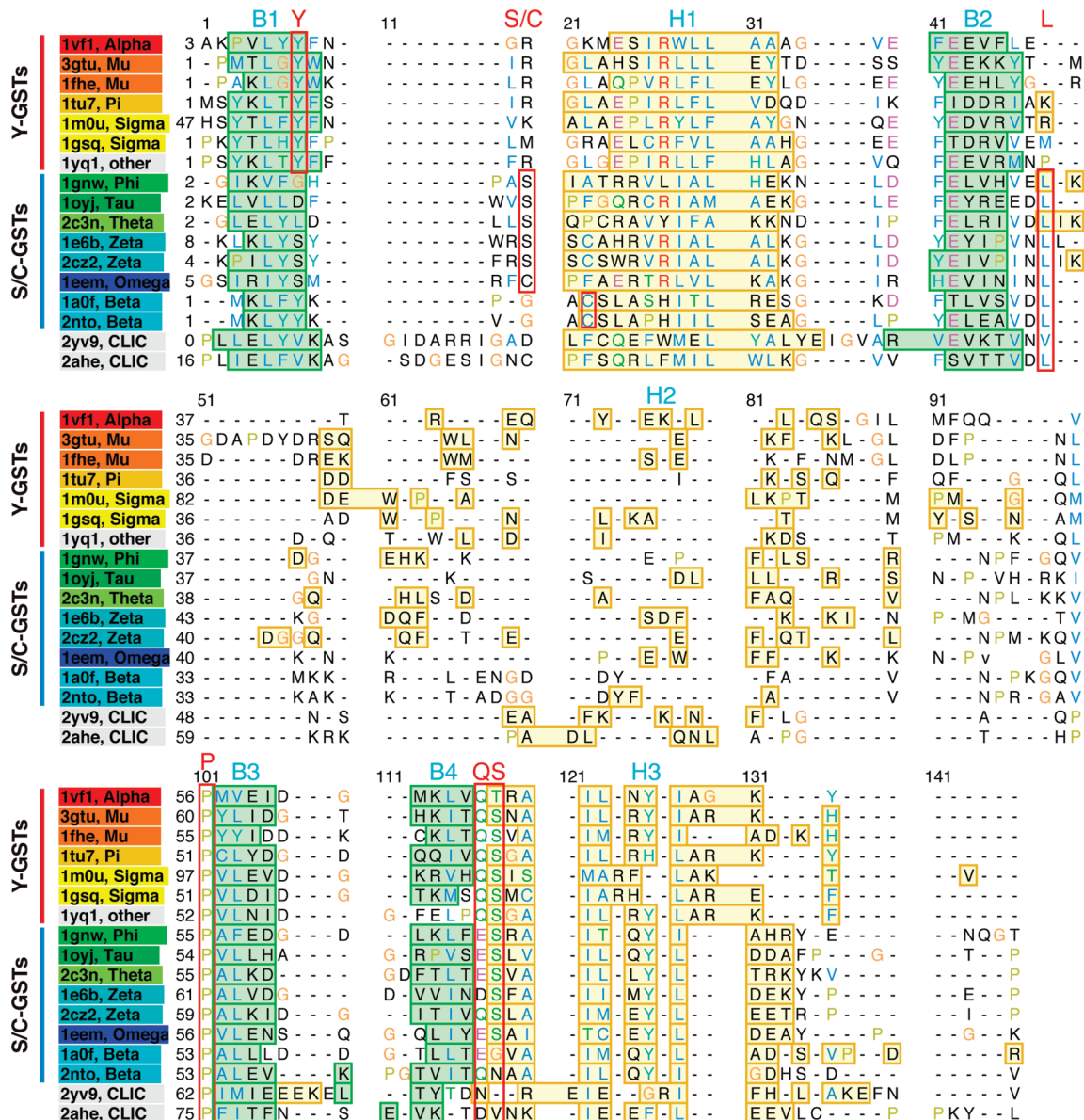
FIGURE 3: Structure-based sequence alignment highlights differences between tyrosine-type and serine/cysteine-type cytosolic GSTs. This alignment of the thioredoxin-like domain from representatives spanning the cytosolic GSTs lists both the PDB entry and GST class label for each structure. Amino acid positions of interest are denoted by red boxes and labels that correspond to the locations marked on the diagram of the thioredoxin fold in Figure 4. Residues in α-helices have a yellow background, and those in β-strands have a green background. This alignment was prepared and displayed using UCSF Chimera (*27*).

hierarchy is necessary to accurately describe the interrelationships of the different major subgroups. All cytosolic GSTs have a similar structure, but as illustrated using a structure similarity network, the classes that were first discovered and which are found almost exclusively in eukaryotes, the alpha, mu, pi, and sigma classes of the Y-type major subgroup, are significantly more similar to one another than to the classes in the other major GST subgroup (Figure 1A). Despite important differences that distinguish these groups, such as the extra active site helix in the alpha class (*29*) and the hydrophilic dimer interface in the sigma class (*30*), when viewed within the context of the larger GST superfamily, these proteins are grouped tightly together. In protein similarity networks, proteins are represented as points and similarity relationships (based on pairwise structure or sequence alignments that are better than a threshold) are represented as edges connecting the points. When displayed, similar proteins tend to cluster together within the networks,

revealing the interrelationships between groups of proteins (*23*). The similar eukaryotic classes will be termed the tyrosine-type GSTs (Y-GSTs), as they are associated with an interaction between a tyrosine and glutathione as an aspect of their mechanisms, insofar as this mechanism is known. The other major subgroup of cytosolic GSTs will be termed the serine/cysteine-type GSTs (S/C-GSTs), described in detail the next section. This separation of GSTs into two major subgroups is also evident from the perspective of protein sequence, as shown by a sequence similarity network (Figure 2).

Structure-based similarity networks are most useful for interrogating GST relationships at the level of gross domain variation, while sequence similarity networks expose clustering by sequence motifs within the GST domain. In terms of structural comparisons, both Y- and S/C-GSTs are extremely similar; structural superpositions, even between the most distant pairs of GSTs, are highly significant. [FAST, the structural alignment program used
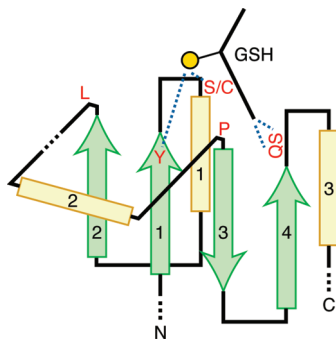
FIGURE 4: Diagram illustrating the secondary structure elements of the thioredoxin-like domain (Trx) found in GSTs. α-Helices are shown as yellow blocks, while β-strands are shown as green arrows. Each element is numbered for reference in the text. The Trx domain is extended in GSTs at the C-terminal end, which continues to complete the overall GST structure. The sulfhydryl group of glutathione (GSH) is displayed as a yellow ball, and key interactions between GSH and residues from the Trx domain are shown as dashed blue lines. Amino acid positions of interest are shown in red type and correspond to the red boxes in the alignment in Figure 3.

to calculate the edges in Figure 1, denotes alignments with scores between 0 and 1.5 as insignificant, and scores greater than 1.5 as significant (*25*). All structures are nearly completely connected at the threshold of 7.5 in Figure 1B, emphasizing the similarity in overall fold between the two major subgroups.] The structural similarity network in Figure 1B clusters together all GSTs because the typical count and three-dimensional arrangement of secondary structure elements are fairly consistent across all cytosolic GSTs. However, calculating sequence similarity between these two main groups of GSTs tests the limits of sequence alignment algorithms; this distant and generally insignificant level of sequence similarity is not included in the sequence similarity network in Figure 2, resulting in no displayed connection between the two groups. Even the most critical catalytic residues, those that interact with the sulfhydryl of glutathione, vary between GST classes, and the locations of these residues change between S/C-GSTs and Y-GSTs, confounding sequence alignments (to be discussed in the following section). The number of residues that are completely conserved across both major subgroups is small (Figure 3).

Although many more enzymes from the tyrosine-type classes have been characterized, the other classes of the S/C major subgroup are actually far more populated and represent a broader range of structural and taxonomic diversity. This disparity in population is conspicuous in the sequence network in Figure 2. Unlike the eukaryotic Y-GSTs, the other classes in the S/C GST major subgroup are found in all kingdoms of life outside of archaea, which lack the machinery for the synthesis of glutathione (*31*). It has been suggested that the Y-GSTs have evolved more recently (*15*), and this is supported by their taxonomic distribution. The better-known roles of certain Y-GSTs in human disease and their longer history of study are the primary reasons for their more thorough structural sampling (*32*).

Tyrosine-Type versus Serine/Cysteine-Type GSTs. The overall differences in sequence and structure between the two major subgroups of GSTs are partly due to changes at specific positions within the fold. The residue that interacts directly with the nucleophilic sulfur of GSH changes both in character and in location relative to the fold between the two major subgroups; this has important implications for changes in their catalytic

mechanism, as far as they are known. A number of residues located within the thioredoxin-like domain found in all GSTs have key roles, and a structure-based sequence alignment illustrates where these residues change between the two major subgroups (Figure 3). In the first and most important point difference, where Y-GSTs have a tyrosine positioned at the end of the first β-strand (B1), the second major subgroup has a serine or cysteine several positions later at the amino terminus of helix 1 (H1); these residues are conserved and key to catalysis in most characterized enzymes. Following this, the second major subgroup of GSTs will be termed serine/cysteine-type GSTs (S/C-GSTs). In characterized Y-GSTs, the tyrosine hydroxyl group is thought to act as a hydrogen bond donor to the sulfur of GSH, lowering its pK_a to stabilize a nucleophilic thiolate (*15*).

Representatives of the S/C-GSTs use their active site residues in an analogous manner, with some exceptions. In theta class GSTs, the hydroxyl group of serine is used to activate the bound GSH (*33, 34*); this role is also ascribed to the serine of the phi and tau classes. In the omega class, however, an active site cysteine forms a disulfide bond with glutathione, and it has the "thiol-transferase" activity associated with glutaredoxins; it is invisible to conventional biochemical assays for GST function (*35*). The zeta class enzymes have a serine that is critical to their functions but also have a reactive active site cysteine that plays a role in binding GSH, although this cysteine is not required for catalysis [see the **SSC** motif in Figure 3 (*36, 37*)]. The last unusual major S/C-GST class is the beta class, which has been seen with GSH bound via a mixed disulfide to a conserved cysteine. However, mutagenesis studies have shown that some beta class enzymes can also transform substrates in the absence of this cysteine (*3*), including one that maintains near-wild-type levels of GST activity with the model substrate CDNB (1-chloro-2,4-dinitrobenzene) (*38*). (These catalytic residues are shown boxed and labeled "Y" or "S/C" in Figures 3 and 4.) *E. coli* yfcG, a GST superfamily member most similar to the members of the beta, phi, and theta classes, has recently been shown to efficiently catalyze the reduction of a model glutaredoxin substrate (*39*). Importantly, yfcG has no active site cysteine, and the yfcG crystal structure shows a threonine in the position of the S/C-GST catalytic residue with the threonine hydroxyl group within hydrogen bonding distance of the sulfhydryl of GSH (*39*). YfcG provides additional evidence that the S/C-GST group is more variable than the Y-GSTs, and as more GSTs are characterized, the picture will likely become better resolved.

Catalytic Consequences of the Shifted Active Site. While the Y-GST tyrosines and S/C-GST serines and cysteines have, to varying degrees, previously been identified as being important to catalysis, the conclusion that the major subgroup-specific locations of these residues have an impact on how glutathione is bound is a new observation with implications for specificity differences between the classes of each major subgroup. The structural alignment shows that these critical residues are found in two different locations, anchored to distinct elements of secondary structure and separated by a loop (Figures 3 and 4). This leads to a trend where GSH is bound relative to the fold; in particular, the location of the sulfur of GSH is positioned farther from the amino terminus of H1 if the enzyme is in the Y-GST group (Figure 5). Additionally, in many Y-GST structures, the N-terminal end of H1 is frayed, disrupting the aligned backbone groups that lead to the positive electrostatic environment at the end of conventional α-helices. It is unclear how this difference affects overall substrate preference or biological
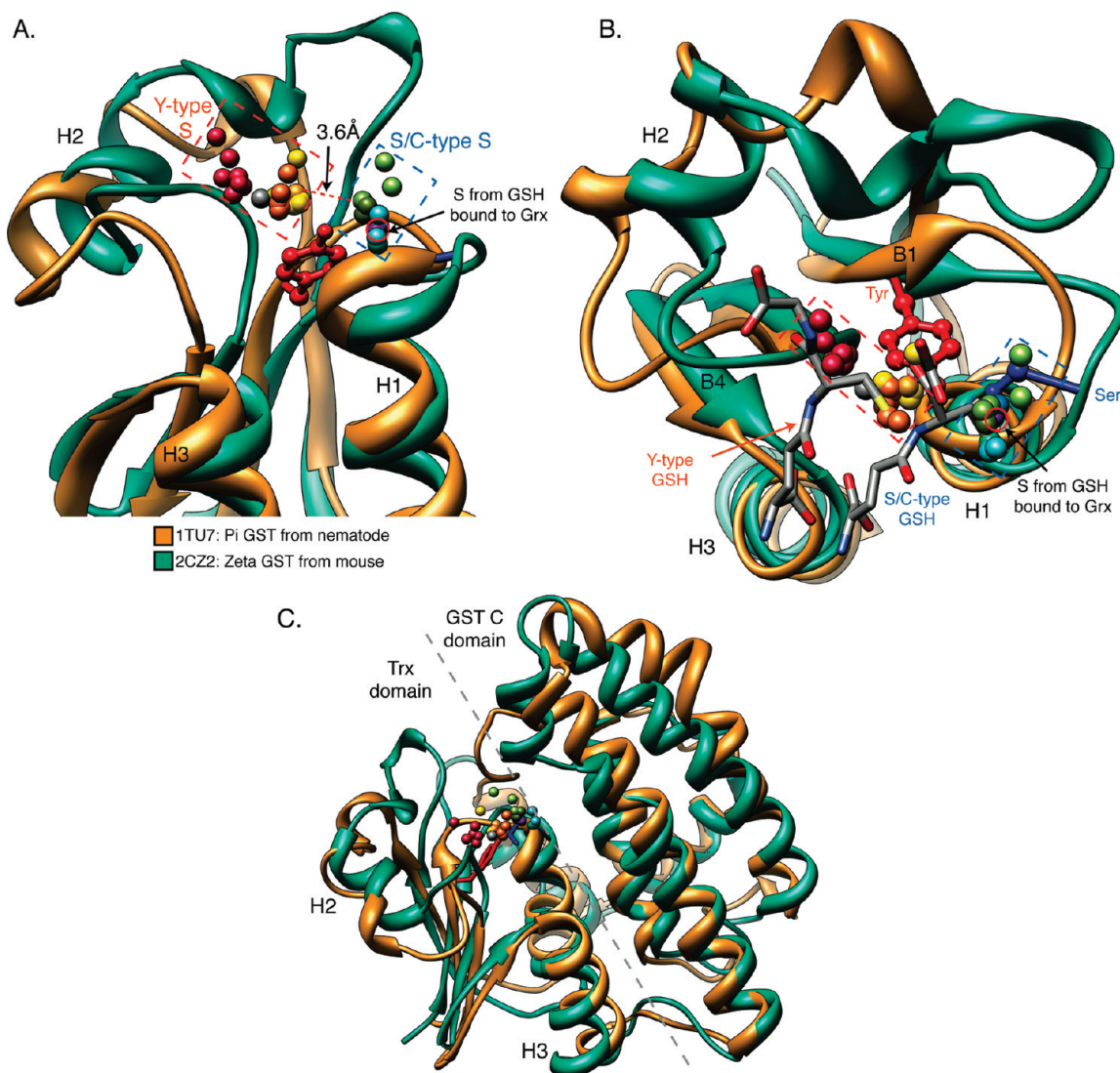
FIGURE 5: Two major subgroups of cytosolic GSTs bind glutathione differently. Structural alignments of representatives from each cytosolic GST class (e.g., alpha and omega) illustrate a difference in the location of the sulfhydryl sulfur of glutathione that is associated with the location of the catalytic residue relative to the fold. (A) A tyrosine-type GST, 1TU7, is shown aligned with a serine-type GST, 2CZ2. Only the Trx-like domains are displayed. Also shown as spheres are the locations of the sulfur from GSH in several examples (if available) from each GST class; the rest of GSH and enzyme ribbon diagrams are not shown. GSH sulfurs are colored according to the class of the parent enzyme based on the keys provided in Figures 1 and 2. The displayed alignment between the Trx-like domain of pi 1TU7 and zeta 2CZ2 has an rmsd of 1.29 Å across 96 atom pairs. The shortest distance between pairs of sulfurs from GSTs of the opposite major subgroup is 3.6 Å as marked on the figure. Also noted is the location of the sulfur of GSH bound to *E. coli* Grx 3 (3GRX). Sulfurs from a total of 33 liganded structures are shown (see Experimental Procedures). (B) Alternate view of the same alignment showing the location of the sulfurs relative to the amino terminus of H1. Here, the tyrosine from 1TU7 is colored red, and the serine from 2CZ2 is colored purple, as well as the full bound GSH from each displayed enzyme. Dashed boxes mark the locations of the GSH sulfurs in the two major subgroups of enzymes. (C) Full structural alignment between the two example GST monomers, with a dashed line marking the separation between the thioredoxin-like domain and the carboxyl-terminal GST domain.

function; different classes of GSTs (e.g., alpha, mu, and theta) have different substrate profiles, but the sparseness of activity profiles, the lack of knowledge about the physiological substrates, and limited class coverage make it difficult to predict commonalities in overall substrate preference within the Y-GSTs or S/C-GSTs. However, we speculate that this difference in the binding of GSH results in a change in how GSH is activated for transfer to a substrate. In other superfamilies found within the thioredoxin fold proteins, a nucleophilic thiolate provided by a cysteine is located in the same position relative to the thioredoxin fold as the critical residues of the S/C-GSTs. This thiolate is stabilized, in part, by the favorable interaction between this anion and the local electrostatic environment of the amino terminus of H1 (*40*). Whitbread and colleagues suggested that the proximity of the amino terminus of H1 could also favor formation of the

GSH thiolate in GST omega, particularly since certain GSH conjugation reactions are still catalyzed by this enzyme in the absence of its active site cysteine (*41*). As mentioned earlier, beta GSTs have also been seen to catalyze the addition of GSH in the absence of active site cysteines. Importantly, it is less likely that H1 of Y-GSTs contributes to the activation of GSH for transfer to a substrate because the sulfhydryl moiety is too distant to experience a significant effect from the electrostatic environment of the helix amino terminus.

This shifted active site represents a pivotal break with features present in the rest of the thioredoxin fold proteins. Given that Y-GSTs are unlikely to use the amino terminus of H1 to stabilize a catalytic thiolate, they are distinct within the superfamilies that incorporate a thioredoxin fold. All other major superfamilies in the fold incorporate a nucleophilic thiolate provided by a cysteine
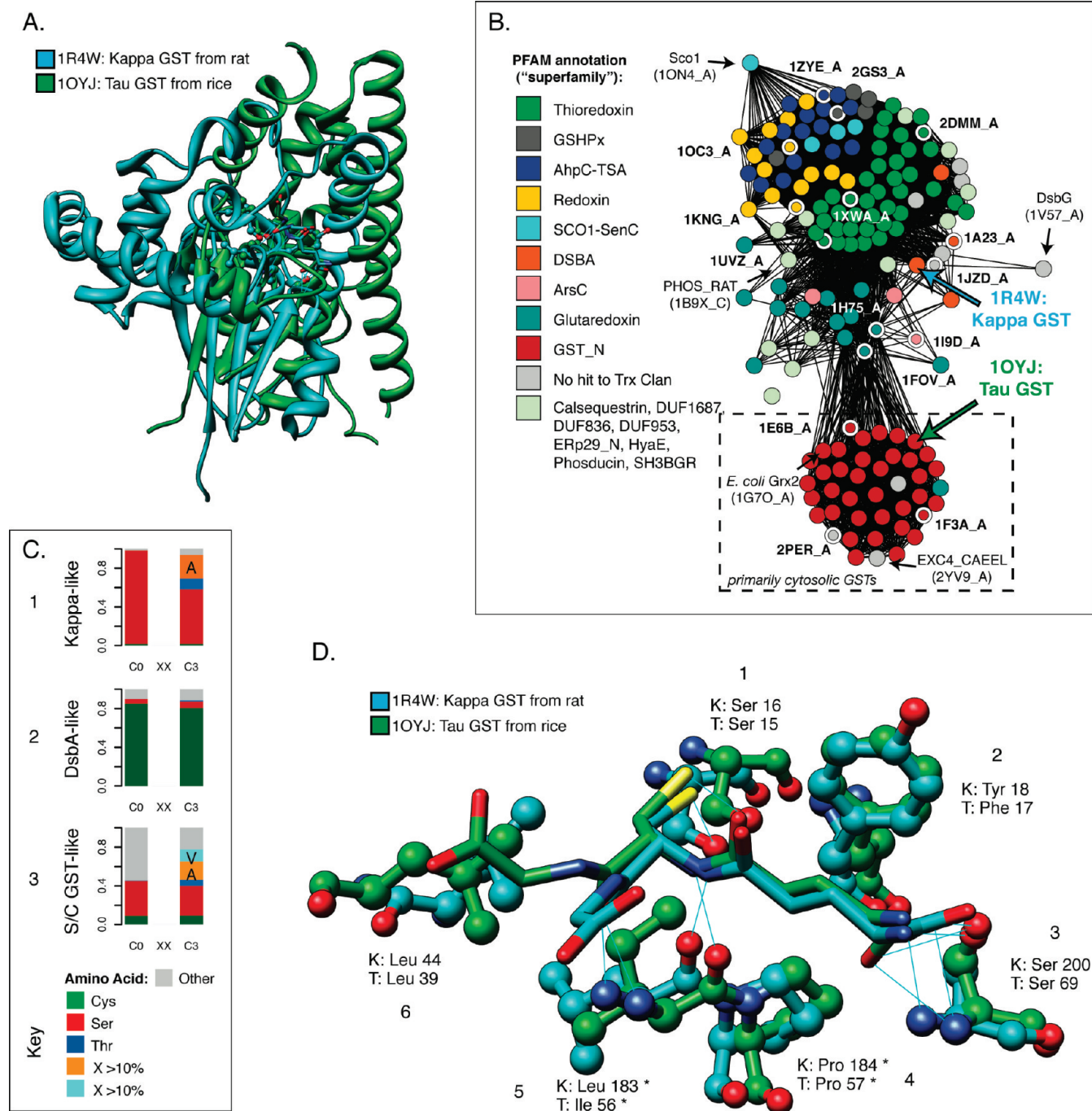
FIGURE 6: GST kappa glutathione binding site as an example of convergent evolution. (A) Superposition of the structures of mitochondrial GST kappa (1R4W) and S/C-GST tau (1OYJ) shows that they are very different. While both contain a form of the thioredoxin domain, in kappa, it is interrupted with a large insert before helix 2 (see Figure 4). The alignment superimposes 61 pairs of residues with a 2.17 Å rmsd, all within the Trx domain; six of these amino acids are identical, for a level of sequence identity of ~10%. The Trx domain in 1OYJ includes ~80 residues. (B) Within the thioredoxin fold class, GST kappa is structurally distant from the cytosolic GSTs. The structure similarity network shows 159 examples that span the thioredoxin fold. Structures are colored according to their PFAM annotation (see Experimental Procedures), which roughly corresponds to different superfamilies. The locations of 1R4W and 1OYJ are noted. The cytosolic GSTs in Figure 1B are excerpted from this network (dashed box), which is thresholded at a FAST score of 4.5. Edges at this threshold represent alignments with a median rmsd of 2.75 Å across 72 aligned positions. Selected structures (nodes with white borders) in the network are labeled with their PDB entries. (C) The kappa class active site residues are more like cytosolic GSTs than its nearest DsbA-like neighbor class. Bar charts 1−3 show the distributions of amino acids aligning with the catalytic Ser/Cys position in the cytosolic GSTs (C0) and the residue three positions later (C3), commonly termed the CxxC motif. The charts summarize 62, 87, and 538 diverse protein sequences, respectively, from each grouping. (D) Corresponding residues from kappa structure 1R4W and tau cytosolic GST 1OYJ that form binding interactions with glutathione. Predicted H-bonds are shown as thin blue lines. Interactions involving enzyme residues or substructures that are present in all or nearly all variants of the thioredoxin fold are marked with asterisks. Residue pairs are numbered consecutively and discussed in the text.

at the amino terminus of H1, the same position occupied by the critical residues of the S/C-GSTs (*26*). A number of these superfamilies also catalyze reactions involving GSH, including

the glutaredoxins (*42*). Figure 5 shows that in *E. coli* glutaredoxin 3, the sulfur of bound GSH is positioned near the positive end of H1, as in S/C-GSTs. Considering the ubiquity and diversity of the

thioredoxin fold, the withdrawal from such a fundamental feature of Trx fold catalysis by the Y-GSTs is remarkable.

The GSH Binding Site of Mitochondrial GST Kappa Is an Example of Convergent Evolution. Mitochondrial GST kappa and the cytosolic GSTs contain two divergent variants of the thioredoxin fold. Their common catalytic capabilities are due in part to the convergent evolution of a similar binding site for glutathione; their glutathione transferase function also relies on the presence of specific structural components that are common to nearly all variants of the thioredoxin fold. It is clear that GST kappa is more closely related to the DsbA-like superfamily of enzymes than to the cytosolic GSTs. Kappa and DsbA-like enzymes share both sequence similarity and a large insertion between β-strand 2 and H2 of the thioredoxin fold (*16*). (This insertion has few common elements between DsbA-like enzymes and kappa, however.) It has been reported that GST kappa has a more recent common ancestor with the DsbA-like enzymes than with any other thioredoxin fold member (*16, 43*). Structurally, GST kappa and cytosolic GSTs are quite distant (Figure 6A,B).

Although DsbA-like enzymes are its closest neighbor in structure and sequence, GST kappa cannot catalyze DsbA-type reactions. [Note that HCCA isomerase, a bacterial enzyme that is sometimes described as similar to GST kappa, is a bona fide kappa class GST (*44*).] DsbA-like enzymes require a CxxC motif containing two cysteines as a central aspect of catalysis of the oxidation and isomerization of disulfide bonds in substrate proteins (*45*). GST kappa-like enzymes do not preserve these cysteines and in fact have a "CxxC" motif that is much more similar to that of S/C-GSTs; that is to say, for both the kappas and the S/C major subgroup, the first position of the motif is often occupied by a serine (Figure 6C). Characterized DsbA-like enzymes do not bind glutathione, unlike GST kappa.

In light of the overall differences between GST kappa and the cytosolic GSTs, the level of similarity at the GSH binding site is startling. The commonalities between GST kappa and the S/C-GSTs are key: the critical residues that interact with the sulfur of GSH are alike, and a significant number of additional residues that bind GSH have similar characters and orientations (Figure 6D). GST kappa uses an active site serine located at the amino terminus of H1 (Figure 6D, 1) to favor ionization of the GSH thiol (*16*), as does the S/C-GST tau from rice. Both enzymes bind GSH such that it is in steric contact with the ring from a tyrosine and phenylalanine from the first turn of H1 (Figure 6D, 2). Both enzymes contain a serine at the end of B4; the hydroxyl group of the serine interacts with the glutamyl group of GSH (Figure 6D, 3). This maps to the QS (Y-GST) or ES (S/C-GST) motif that is recognized as a conserved element of cytosolic GSTs that helps to bind the glutamyl moiety of GSH (*15*); the glutamine/glutamate residue from this motif is not present in kappa. A *cis*-proline from the beginning of B3 is in van der Waals contact with GSH in both enzymes (Figure 6D, 4); with the exception of the peroxiredoxin superfamily of enzymes, this proline is an absolutely conserved structural aspect of the Trx fold (*26*). In an additional contribution from the Trx fold, the bound GSH in both enzymes makes hydrogen bonding interactions with the backbone carbonyl of the preceding residue (Figure 6D, 5). Finally, although the glycinyl moiety of GSH has a different conformation between kappa and tau, the side chain of a leucine at the end of B2 makes contact with this end of GSH in both structures (Figure 6D, 6). This leucine is conserved in S/C-GSTs (Figure 3).

It is clear that GST kappa and the cytosolic GSTs represent fundamentally different versions of the Trx fold; consistent with the previous observations of Ladner et al., their mutual ability to add GSH to small electrophilic compounds appears to be convergently evolved (*16*). However, because of the large overall differences in sequence and structure between these enzymes, it would not be expected that GSH would be bound using a highly similar arrangement of residues. This highlights how the Trx fold provides a foundation for convergent evolution of a GSH binding site in contrast to the additional elements that are required for binding and activation of GSH.

CONCLUSIONS

Glutathione transferases have remodeled the thioredoxin fold as a unique collection of enzymes with complex and overlapping specificities. GSTs have abandoned the classic dithiol CxxC active site motif used by most other variations of the thioredoxin fold, exporting the essential catalytic residues to bound glutathione. Furthermore, in the cytosolic Y-GSTs, the glutathione binding site has shifted away from the amino terminus of the first α-helix, resulting in the withdrawal from an aspect of catalysis that is present in all other major superfamilies that incorporate a thioredoxin fold. The example of mitochondrial GST kappa provides further insight into how fundamental aspects of the thioredoxin fold were combined with novel modifications to enable the multiple reactions catalyzed by the GSTs; this enzyme class provides clues about how a DsbA-like enzyme can be modified to enable glutathione transferase activity and suggests how the thioredoxin fold serves as a foundation for the evolution of a GSH binding site. These distinctive characteristics of GSTs are evidence of an evolutionary vanguard, transforming the capabilities of the thioredoxin fold.

REFERENCES

1. Mannervik, B., Board, P. G., Hayes, J. D., Listowsky, I., and Pearson, W. R. (2005) Nomenclature for mammalian soluble glutathione transferases. *Methods Enzymol. 401*, 1–8.
2. Hayes, J. D., Flanagan, J. U., and Jowsey, I. R. (2005) Glutathione transferases. *Annu. Rev. Pharmacol. Toxicol. 45*, 51–88.
3. Allocati, N., Federici, L., Masulli, M., and Di Ilio, C. (2009) Glutathione transferases in bacteria. *FEBS J. 276*, 58–75.
4. Nebert, D. W., and Vasiliou, V. (2004) Analysis of the glutathione S-transferase (GST) gene family. *Hum. Genomics 1*, 460–464.
5. Armstrong, R. N. (1991) Glutathione S-transferases: Reaction mechanism, structure, and function. *Chem. Res. Toxicol. 4*, 131–140.
6. Josephy, P. D., and Mannervik, B. (2006) Molecular Toxicology , 2nd ed., Oxford University Press, New York.
7. Adler, V., Yin, Z., Fuchs, S. Y., Benezra, M., Rosario, L., Tew, K. D., Pincus, M. R., Sardana, M., Henderson, C. J., Wolf, C. R., Davis, R. J., and Ronai, Z. (1999) Regulation of JNK signaling by GSTp. *EMBO J. 18*, 1321–1334.
8. Cho, S. G., Lee, Y. H., Park, H. S., Ryoo, K., Kang, K. W., Park, J., Eom, S. J., Kim, M. J., Chang, T. S., Choi, S. Y., Shim, J., Kim, Y., Dong, M. S., Lee, M. J., Kim, S. G., Ichijo, H., and Choi, E. J. (2001) Glutathione S-transferase mu modulates the stress-activated signals by suppressing apoptosis signal-regulating kinase 1. *J. Biol. Chem. 276*, 12749–12755.
9. Johansson, A. S., and Mannervik, B. (2001) Human glutathione transferase A3-3, a highly efficient catalyst of double-bond isomerization in the biosynthetic pathway of steroid hormones. *J. Biol. Chem. 276*, 33061–33065.
10. Kanaoka, Y., Ago, H., Inagaki, E., Nanayama, T., Miyano, M., Kikuno, R., Fujii, Y., Eguchi, N., Toh, H., Urade, Y., and Hayaishi,

O. (1997) Cloning and crystal structure of hematopoietic prostaglandin D synthase. *Cell 90*, 1085–1095.

11. Li, Y. J., Oliveira, S. A., Xu, P., Martin, E. R., Stenger, J. E., Scherzer, C. R., Hauser, M. A., Scott, W. K., Small, G. W., Nance, M. A., Watts, R. L., Hubble, J. P., Koller, W. C., Pahwa, R., Stern, M. B., Hiner, B. C., Jankovic, J., Goetz, C. G., Mastaglia, F., Middleton, L. T., Roses, A. D., Saunders, A. M., Schmechel, D. E., Gullans, S. R., Haines, J. L., Gilbert, J. R., Vance, J. M., Pericak-Vance, M. A., Hulette, C., and Welsh-Bohmer, K. A. (2003) Glutathione S-transferase omega-1 modifies age-at-onset of Alzheimer disease and Parkinson disease. *Hum. Mol. Genet. 12*, 3259–3267.

12. Maciel, S. S., Pereira Ada, C., Silva, G. J., Rodrigues, M. V., Mill, J. G., and Krieger, J. E. (2009) Association between glutathione S-transferase polymorphisms and triglycerides and HDL-cholesterol. *Atherosclerosis 206*, 204–208.

13. Hayes, J. D., and Pulford, D. J. (1995) The glutathione S-transferase supergene family: Regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. *Crit. Rev. Biochem. Mol. Biol. 30*, 445–600.

14. Mannervik, B., and Danielson, U. H. (1988) Glutathione transferases: Structure and catalytic activity. *CRC Crit. Rev. Biochem. 23*, 283–337.

15. Armstrong, R. N. (1997) Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chem. Res. Toxicol. 10*, 2–18.

16. Ladner, J. E., Parsons, J. F., Rife, C. L., Gilliland, G. L., and Armstrong, R. N. (2004) Parallel evolutionary pathways for glutathione transferases: Structure and mechanism of the mitochondrial class kappa enzyme rGSTK1-1. *Biochemistry 43*, 352–361.

17. Harris, J. M., Meyer, D. J., Coles, B., and Ketterer, B. (1991) A novel glutathione transferase (13-13) isolated from the matrix of rat liver mitochondria having structural similarity to class theta enzymes. *Biochem. J. 278* (Part 1), 137–141.

18. Robinson, A., Huttley, G. A., Booth, H. S., and Board, P. G. (2004) Modelling and bioinformatics studies of the human kappa-class glutathione transferase predict a novel third glutathione transferase family with similarity to prokaryotic 2-hydroxychromene-2-carboxylate isomerases. *Biochem. J. 379*, 541–552.

19. Finn, R. D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S. R., Sonnhammer, E. L., and Bateman, A. (2006) Pfam: Clans, web tools and services. *Nucleic Acids Res. 34*, D247–D251.

20. Apweiler, R., Bairoch, A., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M. J., Natale, D. A., O'Donovan, C., Redaschi, N., and Yeh, L. S. (2004) UniProt: The Universal Protein knowledgebase. *Nucleic Acids Res. 32*, D115–D119.

21. Li, W., and Godzik, A. (2006) Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics 22*, 1658–1659.

22. Altschul, S. F., Madden, T. L., Schaffer, 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.

23. Atkinson, H. J., Morris, J. H., Ferrin, T. E., and Babbitt, P. C. (2009) Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLoS ONE 4*, e4345.

24. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res. 13*, 2498–2504.

25. Zhu, J., and Weng, Z. (2005) FAST: A novel protein structure alignment algorithm. *Proteins 58*, 618–627.

26. Atkinson, H. J., and Babbitt, P. C. (2009) An atlas of the thioredoxin fold class reveals the complexity of function-enabling adaptations. *PLoS Comput. Biol. 5*, e1000541.

27. Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C., and Ferrin, T. E. (2006) Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinf. 7*, 339.

28. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera: A visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605–1612.

29. Sinning, I., Kleywegt, G. J., Cowan, S. W., Reinemer, P., Dirr, H. W., Huber, R., Gilliland, G. L., Armstrong, R. N., Ji, X., and Board, P. G.; et al. (1993) Structure determination and refinement of human alpha class glutathione transferase A1-1, and a comparison with the mu and pi class enzymes. *J. Mol. Biol. 232*, 192–212.

30. Stevens, J. M., Armstrong, R. N., and Dirr, H. W. (2000) Electrostatic interactions affecting the active site of class sigma glutathione S-transferase. *Biochem. J. 347* (Part 1), 193–197.

31. Copley, S. D., and Dhillon, J. K. (2002) Lateral gene transfer and parallel evolution in the history of glutathione biosynthesis genes. *Genome Biol. 3*, research0025.

32. Sheehan, D., Meade, G., Foley, V. M., and Dowd, C. A. (2001) Structure, function and evolution of glutathione transferases: Implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem. J. 360*, 1–16.

33. Board, P. G., Coggan, M., Wilce, M. C., and Parker, M. W. (1995) Evidence for an essential serine residue in the active site of the theta class glutathione transferases. *Biochem. J. 311* (Part 1), 247–250.

34. Jemth, P., and Mannervik, B. (2000) Active site serine promotes stabilization of the reactive glutathione thiolate in rat glutathione transferase T2-2. Evidence against proposed sulfatase activity of the corresponding human enzyme. *J. Biol. Chem. 275*, 8618–8624.

35. Board, P. G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermiin, L. S., Schulte, G. K., Danley, D. E., Hoth, L. R., Griffor, M. C., Kamath, A. V., Rosner, M. H., Chrunyk, B. A., Perregaux, D. E., Gabel, C. A., Geoghegan, K. F., and Pandit, J. (2000) Identification, characterization, and crystal structure of the omega class glutathione transferases. *J. Biol. Chem. 275*, 24798–24806.

36. Board, P. G., Taylor, M. C., Coggan, M., Parker, M. W., Lantum, H. B., and Anders, M. W. (2003) Clarification of the role of key active site residues of glutathione transferase zeta/maleylacetoacetate isomerase by a new spectrophotometric technique. *Biochem. J. 374*, 731–737.

37. Board, P. G., and Anders, M. W. (2005) Human glutathione transferase zeta. *Methods Enzymol. 401*, 61–77.

38. Casalone, E., Allocati, N., Ceccarelli, I., Masulli, M., Rossjohn, J., Parker, M. W., and Di Ilio, C. (1998) Site-directed mutagenesis of the *Proteus mirabilis* glutathione transferase B1-1 G-site. *FEBS Lett. 423*, 122–124.

39. Wadington, M. C., Ladner, J. E., Stourman, N. V., Harp, J. M., and Armstrong, R. N. (2009) Analysis of the structure and function of YfcG from *Escherichia coli* reveals an efficient and unique disulfide bond reductase. *Biochemistry 48*, 6559–6561.

40. Kortemme, T., and Creighton, T. E. (1995) Ionisation of cysteine residues at the termini of model α-helical peptides. Relevance to unusual thiol pK_a values in proteins of the thioredoxin family. *J. Mol. Biol. 253*, 799–812.

41. Whitbread, A. K., Masoumi, A., Tetlow, N., Schmuck, E., Coggan, M., and Board, P. G. (2005) Characterization of the omega class of glutathione transferases. *Methods Enzymol. 401*, 78–99.

42. Fernandes, A. P., and Holmgren, A. (2004) Glutaredoxins: Glutathione-dependent redox enzymes with functions far beyond a simple thioredoxin backup system. *Antioxid. Redox Signaling 6*, 63–74.

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

44. Thompson, L. C., Ladner, J. E., Codreanu, S. G., Harp, J., Gilliland, G. L., and Armstrong, R. N. (2007) 2-Hydroxychromene-2-carboxylic acid isomerase: A kappa class glutathione transferase from *Pseudomonas putida*. *Biochemistry 46*, 6710–6722.

45. Ito, K., and Inaba, K. (2008) The disulfide bond formation (Dsb) system. *Curr. Opin. Struct. Biol. 18*, 450–458.