ANTIOXIDANTS & REDOX SIGNALING
Volume 10, Number 9, 2008

DOI: 10.1089/ars.2008.2057

Forum Original Research Communication

Evolutionary and Structural Insights Into the Multifaceted Glutathione Peroxidase (Gpx) Superfamily

Stefano Toppo,[1] Stefano Vanin,[2] Valentina Bosello,[1] and Silvio C.E. Tosatto[2]

Abstract

Glutathione peroxidase (GPx) is a widespread protein superfamily found in many organisms throughout all kingdoms of life. Although it was initially thought to use only glutathione as reductant, recent evidence suggests that the majority of GPxs have specificity for thioredoxin. We present a thorough *in silico* analysis performed on 724 sequences and 12 structures aimed to clarify the evolutionary, structural, and sequence determinants of GPx specificity. Structural variability was found to be limited to only two regions, termed oligomerization loop and functional helix, which modulate both reduced substrate specificity and oligomerization state. We show that mammalian GPx-1, the canonic selenocysteine-based tetrameric glutathione peroxidase, is a recent "invention" during evolution. Contrary to common belief, cysteine-based thioredoxin-specific GPx, which we propose the TGPx, are both more common and more ancient. This raises interesting evolutionary considerations regarding oligomerization and the use of active-site selenocysteine residue. In addition, phylogenetic analysis has revealed the presence of a novel member belonging to the GPx superfamily in Mammalia and Amphibia, for which we propose the name GPx-8, following the present numeric order of the mammalian GPxs. *Antioxid. Redox Signal*. 10, 1501–1513.

Introduction

GLUTATHIONE PEROXIDASES (GPxs) belong to a widespread family of proteins that, over the years, have been discovered in almost all kingdoms of life (51). They are included in the heme-free thiol peroxidase class together with peroxiredoxins and catalyze the reduction of H_2O_2 or organic hydroperoxides to water or corresponding alcohols, thus mitigating their toxicity (72). Novel functions have been recently ascertained for this protein superfamily, and their original role in the cell metabolism should be revisited and extended (6, 7). To date, seven members have been discovered in mammals, taking into account their high sequence similarity, conserved sequence patterns, known biochemical function, and characteristic catalytic triad formed by selenocysteine/cysteine, glutamine, and tryptophan (17). Most of the mammalian enzymes are selenoproteins using selenocysteine (Sec) in the catalytic site and glutathione (GSH) as reducing substrate. Sec, the 21st amino acid, is encoded by an in-frame UGA stop codon recognized by a specific Sec t-RNA when a particular stem loop, called selenocysteine insertion sequence (SECIS), forms immediately downstream or in the 3′ untranslated region (UTR) of the transcript (76).

Although GPx-like proteins have been reported in many species, and the importance of selenocysteine has been acknowledged, the expression of cysteine-based proteins is widespread throughout all kingdoms of life (27). Nonetheless, thanks to their importance and variety, Sec-based glutathione peroxidases have been extensively studied in mammals, and the first discovered member of this protein family was found acting as a glutathione-dependent antioxidant enzyme protecting hemoglobin from oxidative degradation (47). Previously known as cytosolic GPx (cGPx), because of its predominant subcellular localization, it is now referred to as GPx-1, and its activity was the first proven to depend on selenium in rats (55). This was confirmed by x-ray crystallography of the bovine GPx-1 orthologue, showing a selenocysteine residue in the active site (17). Both the catalytic triad and the suggested specificity for glutathione have been experimentally verified by site-directed mutagenesis (36). An

[1]Department of Biological Chemistry, and [2]Department of Biology, University of Padova, Italy.

in silico approach has revealed the importance of four arginine residues and a lysine residue directing the donor substrate toward the active site (3). Recently, GPx-1 was found to be regulated by a signaling pathway through the c-Abl–regulated and Arg-mediated phosphorylation of a tyrosine residue. This event seems to modulate the sensitivity of cells to oxidative stress (10).

With the present nomenclature, GPx-2, originally named GSHPx-GI, has been discovered in the gastrointestinal tract (11), whereas GPx-3 or plasma GPx (pGPx) is mainly a secreted protein (64). GPx-4, previously known as PHGPx (phospholipid hydroperoxide GPx), is the sole active on membrane-bound hydroperoxides (40, 73) and, being highly expressed in testis, plays a crucial structural role in spermatogenesis (43, 56). Presently, this is the only member known to behave as a moonlighting protein (71), found inactivated and polymerized in the mitochondrial capsule of the mature spermatozoa (39, 45). The subcellular localization of GPx-4 depends on distinct promoters (41) that control the specific expression of three distinct transcripts encoding for a cytosolic (8), mitochondrial (2), and nuclear form (37, 53).

GPx-5 is the first-discovered Cys-based member of the superfamily and has been found expressed in the epididymis as a secreted protein (22). Also named epididymal GPx (eGPx), its role has been partially proven to be part of the backup system protecting sperm from the toxicity of hydrogen peroxides in mice (75). In humans, instead, it has been found expressed at low levels, and most of the transcripts are incorrectly matured. This has raised questions about its real functional role, given that even its electron donor has not been determined (25).

The sixth member of the superfamily, GPx-6, was discovered as specifically expressed in the olfactory epithelium and previously named olfactory-metabolizing protein (OMP) (14). The peculiarity of this protein is that it is selenocysteine based in humans, whereas a cysteine is present in the active site of mouse and rat (33), which both possess a fossil inactive SECIS element in the 3′ UTR.

The last member discovered so far in mammals is GPx-7 (74). Being cysteine based, it has been found with low glutathione peroxidase activity, even though it has been confirmed to be involved in mitigating oxidative stress in breast cancer cells.

At the structural level, GPx-1, -2, -3, -5, and -6 are homotetrameric. The acknowledged dimer and tetramer interfaces (17, 72) are missing in GPx-4 and -7. This is evident at sequence-alignment level, where gaps are present between these groups, which may also account for the different substrate specificity. Recently, the structure of a poplar GPx-like protein was solved by x-ray crystallography and has been described as a homodimer, despite lack of the canonic oligomerization interfaces (31). Thanks to the completion of genomic projects, a vast number of sequences have been found to be putatively associated with the GPx superfamily, covering all kingdoms of life, especially bacteria. Novel evidence has highlighted different roles (7). A long record of traceable functions in different species confirms the crucial role of this widespread superfamily (6), although yielding the impression of being incomplete in its multifaceted aspects. In support of this assertion, well-documented studies have reported that GPx-1 is regulated by a signaling pathway mediated by the phosphorylation of a tyrosine by c-Abl and Arg tyrosine kinases (10). A yeast protein homologous to mammalian GPx-4, named yeast-GPx-3, is involved in the direct peroxide-dependent activation of the transcription factor Yap-1 (15), thus supporting the notion that glutathione peroxidases are more than simple antioxidants (27).

Given the multifaceted features of this superfamily, we investigated different aspects concerning the evolutionary, structural, and functional details with the intent to provide a thorough interpretation of the experimental evidence acquired so far. Phylogenic analysis has allowed us to trace the putative history of selenium utilization in the active site and evolution from a monomeric ancestor toward the oligomeric form predominant in vertebrates. In addition, phylogenic analysis has provided evidence that a novel member of this superfamily exists and, following the present classification of GPxs, we named it GPx-8. With recent x-ray structures and *in silico* modeling, it has been possible to study the superimposed structures and highlight the structural determinants of the oligomerization interfaces. At the functional level, we provide the strict sequence and structural features needed to discriminate real GPx members, by using glutathione as electron-donor substrate, from those with a marked specificity for thioredoxin as reductant. For the latter, the name is apparently a misnomer, insofar as they share a common phylogenic origin with canonical GPx, but do not use glutathione as electron donor.

Materials and Methods

Sequence retrieval and alignment

GPx-like sequences have been automatically extracted from the May 2007 release of UniProt (4) by using BLAST (1) searches, starting from the seven human GPx sequences. The relevant accession codes are summarized in Table 1. Highly similar sequences (e-value cutoff, 10e-5) have been retained for further analysis. Full-length amino acid sequences have been recovered from the corresponding nucleotide mRNA or genomic sequences. The use of either selenocysteine or cysteine in the active site has been confirmed by the presence of the TGA stop codon in the open reading frame. After automatic and manual checks, 724 unique full-length proteins have been obtained. Multiple alignment was constructed with MUSCLE (16) and CLUSTALW (67). The final alignment has been manually refined and used in the subsequent analysis. The multiple-sequence alignment has been obtained by using ESPript (23).

Phylogenic analysis

A preliminary quartet puzzling analysis has been performed with the Treepuzzle program (61, 62) to test whether a phylogenic approach could be applied to the original data set. Particular attention was paid to the vertebrate data set showing a resolution >94%. Phylogenic studies have been performed according to the maximum likelihood (ML), neighbor joining (NJ), and maximal parsimony (MP) methods (20). The ML analysis was performed with the PHYML 2.4 program (24). NJ and MP analyses were done by using PHYLIP 3.6 (19). The JTT substitution matrix (29) was used during reconstruction, whereas site heterogeneity was modeled with a four-category Γ distribution. Nonparametric bootstrap resampling (BT) (18) was performed with 1,000

Table 1. Overview of GPx Sequences and Structures*

Protein accession code	*PDB code*	*Quaternary structure*
GPX1_HUMAN	2F8A	Tetramer
GPX2_HUMAN	2HE3	Tetramer
GPX3_HUMAN	2R37	Monomer
GPX4_HUMAN	2GS3, 2OBI	Tetramer
GPX5_HUMAN	2I3Y	Tetramer
GPX6_HUMAN	n/a	Tetramer
GPX7_HUMAN	2P31	Tetramer
A3FNZ8_ROSI (popular GPx-5)	2P5Q (reduced), 2P5R (oxidized)	Dimer

*The table shows the accession numbers of the GPx sequences used to initiate the database search. The structures with listed PDB codes were used to derive the multiple structure alignment. Note that poplar GPx-5 was not used for database searches, and only the reduced form was included in the multiple structure alignment.

replicas to test the robustness of the tree topologies obtained from MP and ML analyses. The tree topologies were visualized with the Treeview 1.6.6 (50) and NJplot (52) programs. Distances were calculated by using PHYLIP 3.6, applying the JTT substitution matrix.

Structural analysis

Analysis of the structural variability of the GPx fold was based on the x-ray crystallographic structures with the following PDB codes summarized in Table 1. The oxidized poplar GPx structure (PDB code 2P5R) was excluded, as it represents a different functional state. Structures for the two missing human (GPx-6 and Q8TED1) and *Escherichia coli* GPx (accession code: BTUE_ECOLI) enzymes were built by comparative modeling from the template structures (PDB codes 2I3Y, 2P31 and 2P31, respectively) selected by highest MANIFOLD (5) sequence similarity by using the following procedure. The sequences were aligned with a profile–profile method generating a limited set of alternative alignments (60), from which the one with the lowest FRST energy (68) was automatically selected. The HOMER server (URL: http://protein.bio.unipd.it/homer/) was used to copy the conserved regions, with loops (69) and side chains (9) modeled subsequently. A multiple-structure alignment was constructed from pairwise structural alignments calculated with CE (58), taking the structure with PDB code 2GS3 (human GPx-4) as the reference on which the others are iteratively superimposed. All structures were visualized and analyzed with PyMol (DeLano Scientific, URL: http://www.pymol.org/).

Results

Sequence features

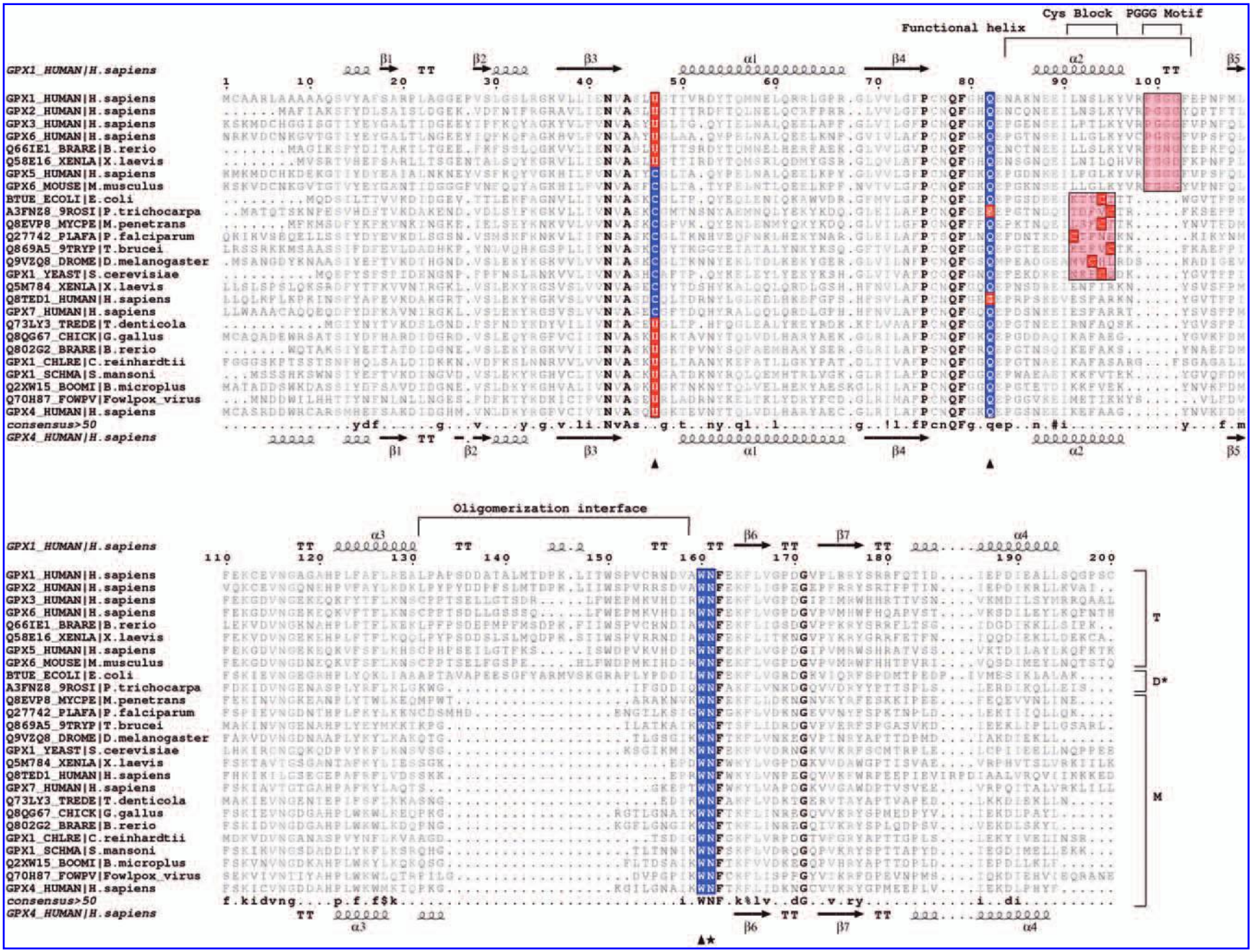
A representative subset of aligned GPx sequences is shown in Fig. 1, with the established catalytic triad and the fourth key residue asparagine close to the catalytic site (70) highlighted. Clusters of highly conserved amino acids and details concerning the oligomerization and functional interfaces that have been further confirmed at structural level also are shown (see structural considerations later). Sequences showing both the insertion of the motif "PGGG" in the functional helix and the extended central insertion (see Fig. 1) belong to the tetrameric form. The most significant signatures have been calculated on 724 aligned proteins, where the amino acids or variants in brackets are present at least in 75% of the proteins. The "L(V/I)VN(VT)ASx(**C/U**)G(L/F)TxxxYxxLxxL" motif surrounds the reactive cysteine/selenocysteine residue. The "G(L/F)x(V/I)L(G/A)FPCN**Q**FxxQEP" and "**WN**-FxxKFL(V/I)" patterns surround the glutamine, tryptophan, and asparagine involved in the catalytic site. The conserved "KxxVxGPx(Y/F)" motif resides between these two patterns. The corresponding oligomerization interface and functional helix are delimited by brackets in Fig. 1.

Phylogenic analysis

Preliminary analysis, performed on bacterial, archeae, and fungi sequences, revealed that the proteins of these taxa have a basal position to the metazoan sequences, as shown in Fig. 2. The protozoan sequences are grouped in two subclades that cluster in the same branches of the monomeric forms of human GPx-7, the novel GPx-8 (see text later), and GPx-4, respectively. No protozoan sequences are associated with the tetrameric human GPx-1, -2, -3, -5, and -6. It is worth noting how the invertebrate sequences are grouped into two subclades. The first one, including the sequences of Platyhelminthes, Arthropoda, and Nematoda, belongs to the human GPx-4 branch of the tree. The second has a basal position to the vertebrate sequences, including the human tetrameric GPx-1, -2, -3, -5, and -6. This suggests that the tetramerization event of GPX sequences has occurred recently during vertebrate radiation.

To investigate the evolutionary events that occurred in mammalian GPx sequences, a phylogeny including all the known vertebrate sequences has been constructed. The ML topology obtained with the amino acid data highlights two well-resolved main clades, called A and B in Fig. 3. The bacterial sequences have been used as an outgroup (data not shown). This revealed a strong statistical support in the node connecting clades A and B, which proved to be sister groups.

Clade B clusters the tetrameric human GPxs (GPx-1, -2, -3, -5, and -6). All these proteins, with the exception of GPx-5 and GPx-6 of mouse and rat, have a selenocysteine residue in the active site. This peculiarity suggests a common ancestor of this clade, with a tetrameric quaternary structure and active-site selenocysteine that has later reverted to cysteine in GPx-5 and mouse/rat GPx-6. Given that the outgroup sequences are cysteine-based bacterial GPx sequences, an ancestor with an active-site cysteine residue could be postulated and proposed. In addition, within clade


Cys Block
PGGG Motif
Functional helix
Oligomerization interface
GPX1_HUMAN|H.sapiens
GPX2_HUMAN|H.sapiens
GPX3_HUMAN|H.sapiens
GPX6_HUMAN|H.sapiens
Q66IE1_BRARE|B.rerio
Q58E16_XENLA|X.laevis
GPX5_HUMAN|H.sapiens
GPX6_MOUSE|M.musculus
BTUE_ECOLI|E.coli
A3FNZ8_9ROSI|P.trichocarpa
Q8EVP8_MYCPE|M.penetrans
Q27742_PLAFA|P.falciparum
Q869A5_9TRYP|T.brucei
Q9VZQ8_DROME|D.melanogaster
GPX1_YEAST|S.cerevisiae
Q5M784_XENLA|X.laevis
Q8TED1_HUMAN|H.sapiens
GPX7_HUMAN|H.sapiens
Q73LY3_TREDE|T.denticola
Q8QG67_CHICK|G.gallus
Q802G2_BRARE|B.rerio
GPX1_CHLRE|C.reinhardtii
GPX1_SCHMA|S.mansoni
Q2XW15_BOOMI|B.microplus
Q70H87_FOWPV|Fowlpox_virus
GPX4_HUMAN|H.sapiens
consensus>50
GPX4_HUMAN|H.sapiens
T
D*
M

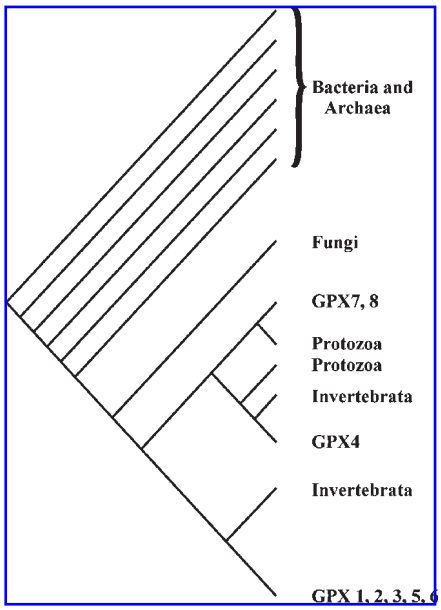


FIG. 2. Schematic representation of the phylogenic relation among 724 GPx sequences. The tree topology reflects the ML, MP, and NJ reconstructions (see text for details). The eukaryotic GPx sequences are grouped as a monophyletic clade in several reconstruction. The tree topology also shows the novel GPx-8 strictly connected to the known GPx-7 (see text). Plant GPx sequences have not been considered in this reconstruction.

B, two minor clades (named GPx-1, -2, and GPx-3, -5, and -6) are clearly isolated from each other.

Clade A shows two well-resolved subclades, including the monomeric human GPx-4 and GPx-7. Proteins grouped in this clade have both cysteine and selenocysteine in the active site and are monomeric. The basal node of these two groups is strongly supported and, as noted earlier, even in this case, a cysteine-based ancestor can be hypothesized, considering bacterial GPx sequences with cysteine as an outgroup. The GPx-4 subclade is particularly rich and contains sequences from fish, birds, and mammals that are all selenocysteine based in the active site, whereas only cysteine-based sequences belong to the GPx-7 subclade. Unexpectedly, a novel well-resolved and distinct clade has emerged from analysis of the GPx-7 subclade. The human Q8TED1 sequence has clustered with sequences from Mammalia and Amphibia. Its calculated distance from human GPx-7 amounts to 0.66 (calculated by using the JTT substitution matrix). This is higher than the calculated range of 0.335 to 0.467 found among different GPx proteins belonging to the subclades of the tetrameric forms (see the GPx-1, -2, and GPx-3, -5, and -6 clades of Fig. 3). In terms of sequence identity, human GPx-7 and Q8TED1 share 51% identical residues, whereas a higher identity is found in the GPx-1, -2 subclades (66%) and within members of the GPx-3, -5, and -6 subclades (67% on average). Given that the distance between human GPx-7 and Q8TED1 is twice the value obtained for the distances among other GPxs belonging to different clades, we propose to call the Q8TED1 cluster GPx-8. It forms a distant sister branch of the GPx-7 subclade and is not one of its members. Further analysis performed on the human sequences by using the MP approach agrees with the general topology obtained for the vertebrate sequences by applying the ML method. The relations among GPx-3, -5, and -6 are not well resolved in both methods, but their basal node and their relation with the subclades GPx-1, -2 result in strong statistical support (Fig. 3).

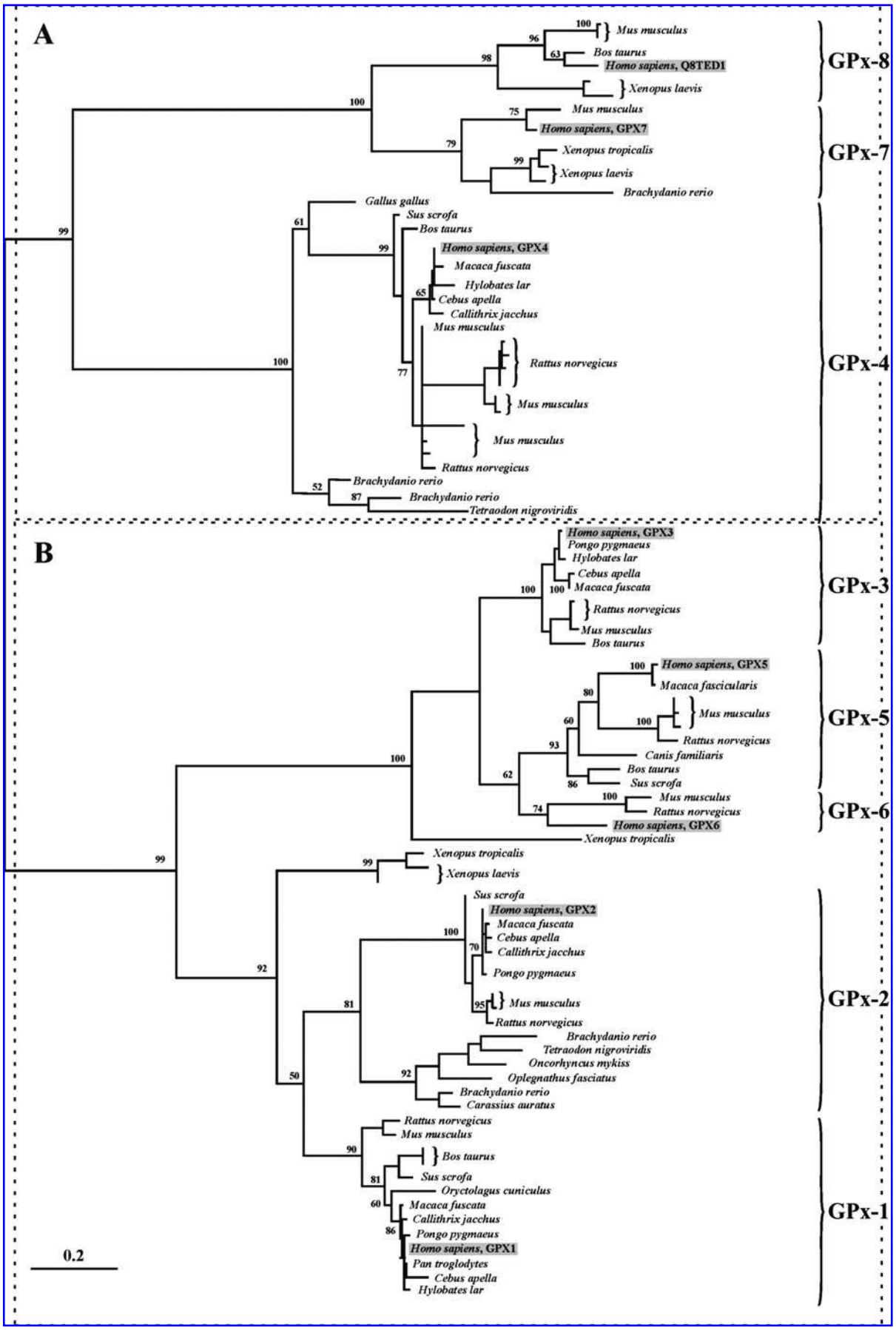
Structural considerations

Variability of the GPx fold was investigated through a multiple structure alignment of the known reduced GPx x-ray structures, as shown in Fig. 4. The overall structure is highly conserved, with the active-site residues, in particular, superimposing almost perfectly regardless of there being a cysteine or selenocysteine residue, despite pairwise sequence identities as low as 25%. The two notable exceptions are regions I and II in Fig. 4a, which we will name oligomerization loop and functional helix, respectively.

The oligomerization loop is located opposite the GPx active site between α-helix 3 and β-strand 6. It shows the highest structural variability of the entire fold, with three distinct variants (see Fig. 4b). The shortest variant is also the most abundant, with the loop turning the backbone from α-helix 3 through a corner into an irregular extended conformation. This corresponds to the monomeric GPx structures. A somewhat longer loop is seen in plant GPx, in which α-helix 3 is extended with a distinctive "GxxG" motif leading into a wobbled corner, returning into the irregular extended conformation. X-ray crystallographic data of poplar GPx-5 suggest this to be sufficient to form a dimeric GPx structure in solution (31, 49).

The third main variant is seen in the tetrameric GPx structures. Here, the loop is conserved and has grown to encompass two flat, irregular, extended conformations connected through a short α-helix. This peculiar structure allows two adjacent monomers to form a flat interaction interface. A single side chain protrudes into the central cavity of the oppo-

FIG. 1. Multiple sequence alignment of representative GPx superfamily members. The secondary structures from tetrameric human GPx-1 (pdb code: 2F8A, top) and monomeric human GPx-4 (pdb code: 2GS3, bottom) are representative for the superfamily. *Triangles*, The catalytic triad cysteine/selenocysteine, glutamine, and tryptophan. *The highly conserved asparagines. Note that in the poplar sequence (A3FNZ8_9ROSI) and the novel human GPx-8 (Q8TED1_HUMAN), the glutamine of the catalytic triad is replaced by glutamic acid and serine, respectively. The oligomerization interface and functional helix are delimited by *horizontal brackets*. *Shaded boxes*, Sequences containing the functionally relevant cysteine-block and the "PGGG" motif characteristic of the tetrameric form. The x-ray structure of poplar GPx-5 is homodimeric. The bacterial *E. coli* sequence (BTUE_COLI) has been predicted to be homodimeric, as it misses the "PGGG" motif responsible for tetramerization but possesses a putative oligomerization interface. These features are common to many bacterial sequences, and the *E. coli* protein has been chosen as representative. The asterisk in the dimeric group refers to the putative dimerization of the bacterial sequences. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article at www.liebertonline.com/ars).


A
Mus musculus
Bos taurus
Homo sapiens, Q8TED1
Xenopus laevis
Mus musculus
Homo sapiens, GPX7
Xenopus tropicalis
Xenopus laevis
Brachydanio rerio
Gallus gallus
Sus scrofa
Bos taurus
Homo sapiens, GPX4
Macaca fuscata
Hylobates lar
Cebus apella
Callithrix jacchus
Mus musculus
Rattus norvegicus
Mus musculus
Mus musculus
Rattus norvegicus
Brachydanio rerio
Brachydanio rerio
Tetraodon nigroviridis
GPx-8
GPx-7
GPx-4
B
Homo sapiens, GPX3
Pongo pygmaeus
Hylobates lar
Cebus apella
Macaca fuscata
Rattus norvegicus
Mus musculus
Bos taurus
Homo sapiens, GPX5
Macaca fascicularis
Mus musculus
Rattus norvegicus
Canis familiaris
Bos taurus
Sus scrofa
Mus musculus
Rattus norvegicus
Homo sapiens, GPX6
Xenopus tropicalis
Xenopus tropicalis
Xenopus laevis
Sus scrofa
Homo sapiens, GPX2
Macaca fuscata
Cebus apella
Callithrix jacchus
Pongo pygmaeus
Mus musculus
Rattus norvegicus
Brachydanio rerio
Tetraodon nigroviridis
Oncorhyncus mykiss
Oplegnathus fasciatus
Brachydanio rerio
Carassius auratus
Rattus norvegicus
Mus musculus
Bos taurus
Sus scrofa
Oryctolagus cuniculus
Macaca fuscata
Callithrix jacchus
Pongo pygmaeus
Homo sapiens, GPX1
Pan troglodytes
Cebus apella
Hylobates lar
0.2
GPx-3
GPx-5
GPx-6
GPx-2
GPx-1

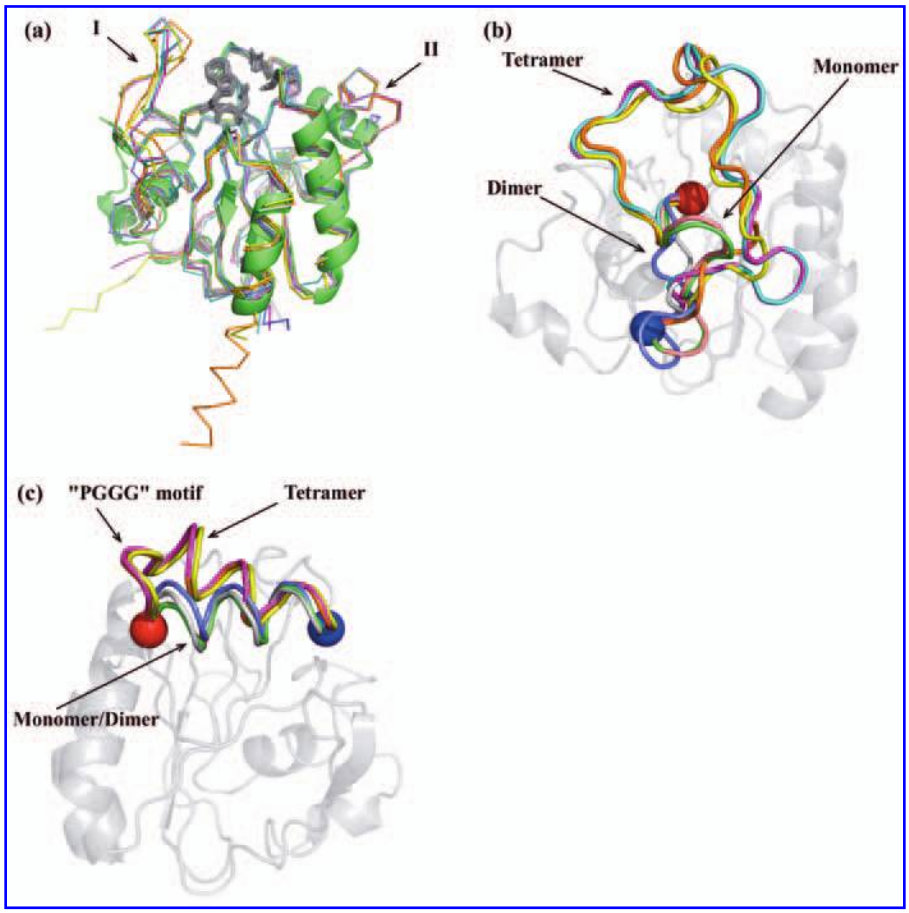


FIG. 4. GPx multiple-structure alignment. (**a**) The known x-ray GPx structures (see Table 1) are shown superimposed as colored wireframes with active site residues as grey sticks. The structure of human GPx-4, PDB code 2GS3, used as reference for the structural alignment, is shown in green as a cartoon. PDB codes, names, and colors for the remaining structures are as follows: 2F8A (human GPx-1, magenta), 2HE3 (human GPx-2, cyan), 2R37 (human GPx-3, orange), 2OBI (human GPx-4, light pink), 2I3Y (human GPx-5, yellow), 2P31 (human GPx-7, light grey), and 2P5Q (poplar GPx-5, blue). The two highlighted regions of structural variability correspond to the oligomerization loop (I) and functional helix (II). (**b**) Close-up of the oligomerization loop. Note the presence of three distinct orientations of increasing length, corresponding to monomeric, dimeric, and tetrameric structures, respectively. (**c**) Close-up of the functional helix. The two orientations correspond to monomeric/ dimeric (horizontal) and tetrameric (tilted upward) structures. Representative overall GPx structures are shown semitransparent in both (**b**) and (**c**), which are rotated for clarity compared with (**a**). The N- and C-terminal anchor points for the structurally variable region are shown with a blue and red sphere, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article at www. liebertonline.com/ars).

site monomer, locking both monomers into a tight connection. The central α-helix additionally stabilizes interactions with the remaining two monomers forming the tetrameric complex. Figure 5 compares the orientation of known dimeric and tetrameric GPx structures.

Some bacterial sequences, such as BTUE_ECOLI, hold an interesting variation in the oligomerization loop. These sequences have a longer loop than the monomeric or dimeric GPx proteins, without significant sequence similarity in this loop. Although the exact conformation can be only roughly predicted, it is conceivable that these sequences represent the link between present monomeric and tetrameric sequences. They would represent an attempt to form oligomeric structures, from which the characteristic tetrameric loop has later evolved. This view agrees well with the previously described phylogenic data.

The second site of structural variability is the functional helix shown as II in Fig. 4c. This helix has two different orientations. All monomeric and dimeric structures present an orientation in which the functional helix runs in parallel to the active site and is not significantly stabilized by hydrogen bonds. This is the orientation that was shown to harbor the resolving cysteine used by some GPx proteins to form an intrachain disulfide that allows reactivity with thioredoxin and preserves the enzyme from overoxidation (42). It is interesting to note how the locations of this resolving cysteine correspond to residues that point away from the active site, implying a significant structural rearrangement, as seen in the poplar x-ray structure (31). Because the resolving cysteine is not present in all monomeric GPx structures carrying active-site cysteine residues, it remains unclear how these enzymes avoid overoxidation (42).

The second structural variant of the functional loop is observed only in tetrameric GPx structures. Here, the helix axis is tilted toward the protein–protein interface and locked into

FIG. 3. Evolutionary relation among vertebrate GPx sequences. The figure shows the maximum likelihood tree ($-\ln L = -9946.31586$) obtained by using the PHYML program. MP tree topology was largely congruent. BT values (1,000 replicas) are reported at each node. For graphic reasons, the BT values at the terminal branch are omitted. *Curly brackets* are used to group the homologous sequences further separated into two main clades named A and B, corresponding to monomeric and tetrameric GPx sequences, respectively. The length of the branches reflects the relative evolutionary distance among the sequences (the bar represents 0.2 substitutions per site). Human GPxs are shaded in *grey*.

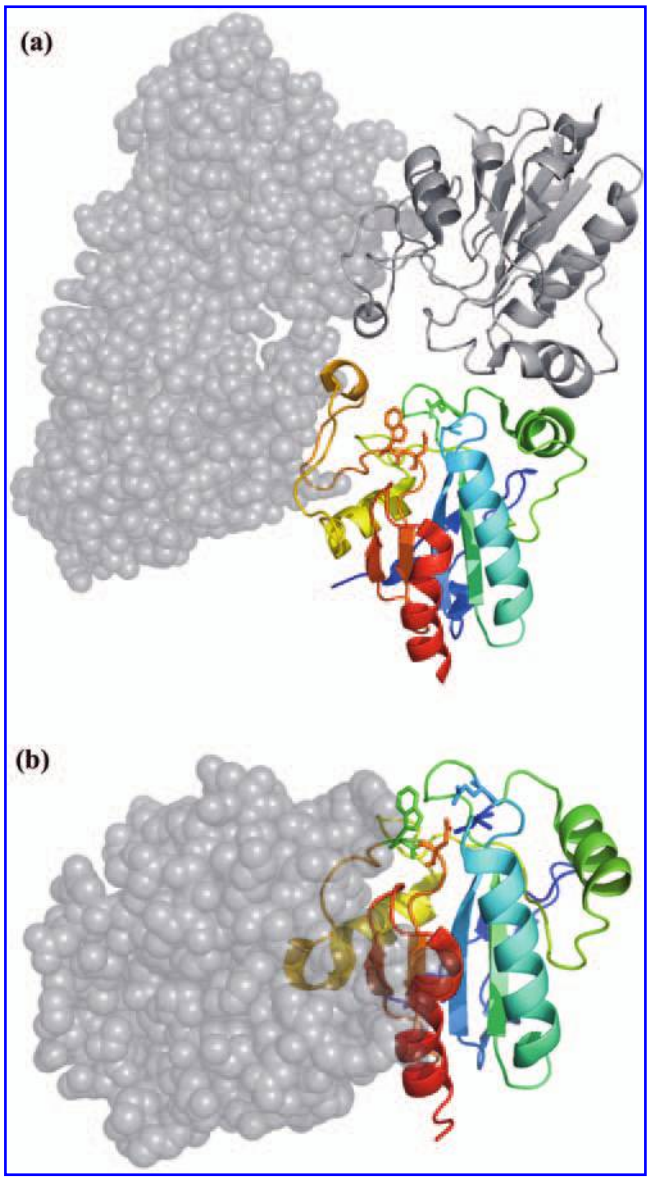


FIG. 5. Oligomeric GPx structures. The two structures of (**a**) tetrameric human GPx-1 (PDB code 2F8A) and (**b**) dimeric poplar GPx-5 (PDB code 2P5Q) are shown as cartoons and spheres in the same orientation. The chain in cartoon representation is colored along the main chain from N- (blue) to C-terminus (red). The dimeric interface for poplar GPx-5 is twisted and more compact than human GPx-1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article at www.liebertonline.com/ars).

position with a second GPx monomer by an intricate network of two interchain salt bridges (glutamic acid 89 with arginine 98) and a π interaction between tyrosine 96 residues. The backbone forms a short hairpin loop after the α-helix through the "PGGG" sequence motif, and is stabilized by a mixture of specific hydrophobic interactions and hydrogen bonds against the rest of the protein. This structural arrangement appears to have evolved as a subsequent feature of the GPx structure to force the formation of stable tetrameric structures, limiting steric accessibility of the active site and excluding thioredoxin as reducing substrate for the enzymatic reaction.

The phylogenic relation between different GPx subtypes and their structures is shown in Fig. 6.

Discussion

The vast amount of data ascribable to members of the GPx superfamily acquired so far has prompted us to investigate the intimate details of the evolutionary and functional history of GPxs. More than 700 full-length sequences distributed in all kingdoms of life could be retrieved from public databases. The recent x-ray structures covering most of the human GPxs, and in particular the GPx from poplar, have contributed to clarify different aspects of the function and the oligomerization state of these enzymes.

Selenium use

Phylogenic analysis confirmed the rather uncommon evolutionary history of selenium use in the active site of GPx. Recent genomic-scale approaches have already led to the suggestive hypothesis that the use of selenoproteins may fol-

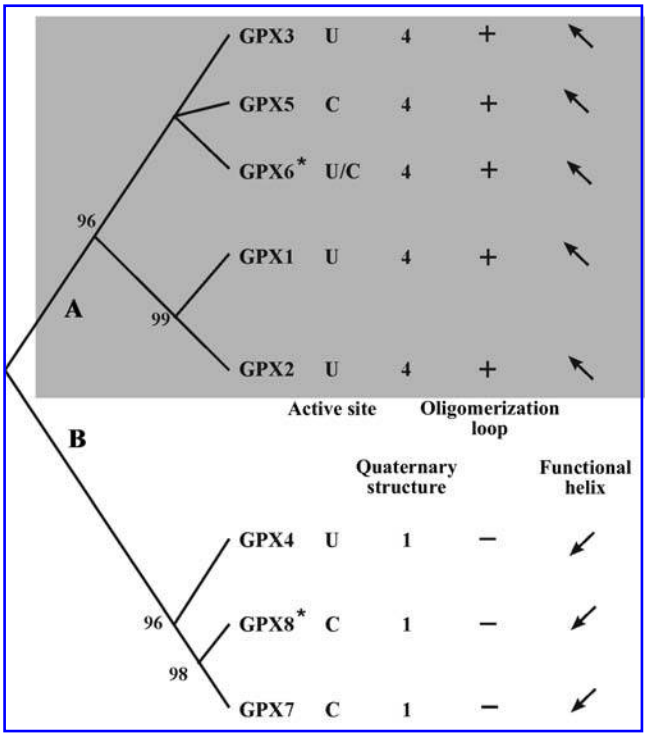


FIG. 6. Evolutionary and structural relations among orthologous vertebrate GPx sequences. The figure shows the maximal-likelihood tree ($-\ln L = -3{,}332.75007$) obtained by using the PHYML program. MP and NJ tree topologies are largely congruent. BT values (1,000 replicas) are reported at each node. The clade GPx-3, -5, -6 is monophyletic in all the reconstructions, but the relations among the sequences are not well resolved. For each sequence, the presence of a cysteine (C) or selenocysteine (U) residue in the active site is listed together with the quaternary structure. Two structural features, oligomerization loop and functional helix (see text for details), are also listed. The arrows reflect the spatial orientation of the functional helix (see Fig. 4c), whereas the asterisks for GPx-8 and GPx-6 refer to the *in silico* prediction of the corresponding structures.

low an uncommon evolutionary history (27), although a trend in the progressive acquisition of selenium may be outlined (77). Selenocysteine-based GPxs are extensively used in mammals, but have been reported in other vertebrates, such as *Gallus gallus* (32), the fish *Danio rerio* (66), the alga *Chlamydomonas reinhardtii* (21, 46), the flatworm *Schistosoma mansoni* (38), the nematode *Setaria cervi* (59), the arthropod *Boophilus microplus* (13), viruses (34, 78), and in the first bacterium ever found, the oral pathogen *Treponema denticola* (57). All this evidence and the present study support the idea of a rather unusual path for the evolution of cysteine and selenocysteine in GPxs. Their alternating nature raises questions about the intimate reasons for this swapping (54), although it may simply depend on the bioavailability of selenium in the different living conditions of the species throughout the past-to-present timeline (12, 35). Responding to the criteria of parsimony in evolution, the phylogenic analysis leads to the conclusion that the ancestral sequence would be cysteine based. Both fungi and bacteria, carrying an active-site cysteine, are basal to the metazoan sequences that have extensively adopted selenocysteine in the active site. A selenocysteine-based ancestor is predicted, as for Gpx-1, -2, -3, -5, and -6 (see Fig. 6).

Interestingly, a recent reversion has occurred in mammals, in which GPx-6 has selenocysteine in humans, but switched back to cysteine in rodents, which still carry a fossil SECIS element. Another reversion has determined the cysteine-based GPx-5, derived from a putative intermediate tetrameric selenocysteine-based ancestor basal to all tetrameric forms (see Fig. 6). In this case, a nonobvious path following the cysteine → selenocysteine → cysteine direction has been detected and is supported by the phylogenic data, suggesting the reversion to cysteine to be recent. This raises new questions, given that the cysteine-based enzymes are less efficient than the selenocysteine-based counterparts in countering oxidative stress.

The evolution of novel GPx forms suggests that these proteins may be used in different functional pathways in addition to the already ascertained canonic roles (27). The identification of a novel GPx-8, very close to GPx-7 (see Fig. 3), identified in mammals and amphibians, goes in this direction. Its peculiarity is the loss of the characteristics required to carry out its own enzymatic function efficiently. The key changes are the substitution of a catalytic-site glutamine with serine and the lack of a resolving cysteine in the cysteine block. It has been found expressed in a large-scale sequencing project of full-length transcripts in mouse and human and is mapped to the 5q11.2 human chromosome. GPx-8 clearly belongs to the GPx superfamily, but nothing is known about its functional role. This could be an example of a novel direction taken by the GPx superfamily during evolution, adopting a different function and concomitantly possibly lacking any peroxidase activity. This evidence is not a completely isolated phenomenon, as previously thought.

Conservation of the catalytic triad residues is strict for both the cysteine/selenocysteine residues (87% and 23%, respectively) and tryptophan (100%), but only 97% for the glutamine residue. Recent experimental evidence for poplar GPx-5 shows this enzyme to be active despite carrying a glutamic acid (see Fig. 1) in place of the glutamine (31, 49). This is the first evidence of a GPx-like protein without the canonic triad. Although this loss of conservation may appear marginal in absolute terms, it assumes relevance if compared with the strict conservation of other residues, such as the asparagine immediately following the conserved tryptophan of the triad. This residue plays an important functional role, and its conservation exceeds 99% over all sequences, redefining the triad as a catalytic tetrad (70).

Table 2. Distribution of Sequence Features for Different Taxa*

	Total	*CCM*	*CCD*	*CCT*	*C-M*	*C-D*	*C-T*	*UCM*	*UCD*	*UCT*	*U-M*	*U-D*	*U-T*
Mammalia	72	—	—	—	7	—	12	—	—	3	19	—	31
Vertebrata	20	—	—	—	6	—	—	—	—	—	4	—	10
Arthropoda	28	15	—	—	4	—	—	—	—	—	7	1	1
Nematoda	16	—	—	—	9	—	7	—	—	—	—	—	—
Trematoda	6	—	—	—	—	—	—	—	—	—	6	—	—
Alveolata	19	1	2	—	16	—	—	—	—	—	—	—	—
Euglenozoa	16	13	—	—	3	—	—	—	—	—	—	—	—
Fungi	36	35	—	—	1	—	—	—	—	—	—	—	—
Viridiplantae	63	62	—	—	—	—	—	—	—	—	1	—	—
Bacteria	440	372	53	—	14	—	—	—	—	—	1	—	—
Viruses	2	—	—	—	1	—	—	—	—	—	1	—	—
Others	6	1	—	—	2	—	—	—	—	—	2	1	—
Total	724	499	55	—	63	—	19	—	—	3	41	2	42

*The table heading reports an acronym summarizing all possible concurrent combinations of sequence characteristics described in the text. the first letter refers to the presence of either cysteine ("C") or selenocysteine ("U") in the active site. The second letter stands for the presence ("C") or lack ("—") of the resolving cysteine in the cysteine-block important for the thioredoxin specificity. The third letter refers to either the presence or lack of the oligomerization interface. Monomeric sequences ("M") do carry neither the "PGGG" motif in the functional helix nor the extended oligomerization loop. Dimeric sequences ("D") possess a long oligomerization loop but miss the "PGGG" motif. Tetrameric sequences ("T") have both the "PGGG" motif and the extended oligomerization loop. Note that only the extended loop of the oligomerization interface has been taken into account to evaluate novel dimeric sequences due to missing evidence at sequence level in case of the short loops. Such loops, indistinguishable from the monomeric pattern, are present in poplar GPx-5, for instance. The true extent of dimeric sequences may have therefore been underestimated.

Oligomerization state

Another interesting feature has allowed a putative evolutionary path toward the tetrameric form. It was previously thought that the central deletion, here called oligomerization loop, and the functional helix with the "PGGG" motif could be treated as tetrameric and dimeric interfaces, respectively (42). This arrangement suggested that during evolution, the aggregation event would first involve the functional-helix/dimeric interface and only later the oligomerization/tetrameric interface. Both poplar GPx-5 and analysis of the multiple-sequence alignment have prompted us to reconsider the sequence of oligomerization events. Among the features reported in Table 2, it is worth noting how 53 bacterial sequences (out of 440) show an extended loop in the oligomerization interface and lack the protruding "PGGG" motif. The extended oligomerization loop, shown in Fig. 1 for *E. coli*, is widespread in different taxa of the Eubacteria domain, such as Firmicutes, Proteobacteria, and Actinobacteria. Data from poplar confirms that a homodimer can be obtained starting from the oligomerization loop through small sequence variations. The evolutionary history of this trend to aggregate in the final tetrameric structure suggests an intermediate dimeric state also common to prokaryotic organisms. These would first involve the oligomerization loop, and only later, the functional helix that has developed the protruding "PGGG" motif to allow aggregation of the tetramer. An indirect support for this hypothesis is the absence of sequences concomitantly exhibiting a functional helix with the "PGGG" motif and missing the oligomerization loop. The definition of the previously known tetrameric interface (42) is replaced with the more general oligomerization loop, conveying the idea that this is the first region used to attempt aggregation. The term functional helix is likewise used in place of dimeric interface, as it does not only have a structural function during the second tetramerization step but rather serves as a fundamental functional element in GPx-like proteins using different reducing substrates, as demonstrated in poplar (31) and *D. melanogaster* (42).

Nomenclature clarification

Many sequences reveal a shift in specificity from glutathione to thioredoxin or related proteins characterized by a "CxxC" motif. This should be taken into account when defining the members of this superfamily as "glutathione peroxidases." Experimental evidence for this has been already reported in plants (26, 30, 31), insects (42, 48), yeasts (15, 65), and parasites such as *Plasmodium falciparum* (63) and *Trypanosoma brucei* (28), which all share common characteristics. In light of this different specificity, the definition of "glutathione peroxidases" appears a misnomer that refers to the donor specificity rather than a common evolutionary origin based on fold and sequence (42). For these proteins, we propose the novel functional class of "thioredoxin GPx-like peroxidases" (TGPx). This accounts for the common evolutionary origin, but highlights the preference for the donor substrate in the reaction. The minimal requisites that distinguish the thioredoxin-based peroxidases from true GPxs are as follows: (a) monomeric or dimeric structure; (b) an active site, peroxidatic, cysteine; and (c) an exposed, resolving, cysteine in the cysteine block delimiting the functional helix, which after oxidation arranges into an intramolecular disulfide bridge with the peroxidatic cysteine. These TGPx sequences are widespread in Arthropoda, Euglenozoa, Fungi, Viridiplantae, and Bacteria, as shown in columns CCM and CCD of Table 2. Although distantly related, these species seem to have adopted thioredoxin as reductant. Conversely, the use of both selenium and glutathione, characteristic of canonic GPx, seems to be a recent acquisition especially found in Vertebrata and Mammalia, which have, in addition, evolved into oligomeric structures.

Conclusions

The variety of forms present in higher eukaryotes is testimony to the acquired importance of this superfamily. GPxs do not limit their function to antioxidant defense of the cell, but rather participate in complex signaling cascades. For instance, the need for tetrameric structure may finely modulate the redox state of the cell in response to excess H_2O_2 or hydroperoxides or both, as previously demonstrated through the phosphorylation of a specific tyrosine (10). Yeast GPx-3 specifically activates the transcription factor Yap-1 (15). Similar functions may be ascertained for other members of this family. Belonging to the versatile thioredoxin fold known to suit many different functions (44), GPxs may have evolved to such an extent in function without perturbing their fold and keeping the active site strictly conserved. In this scenario, many questions may still need answers as novel challenges are offered by this multifaceted superfamily.

Acknowledgments

The first two authors contributed equally to this study. We thank Prof. Matilde Maiorino, Prof. Leopold Flohé, and Prof. Fulvio Ursini for insightful discussion and critical reading of the manuscript.

Abbreviations

Γ, gamma distribution; BT, bootstrap resampling; cGPx, cytosolic glutathione peroxidase; eGPx, epididymal glutathione peroxidase; GPx, glutathione peroxidase; GSHPx-GI, gastrointestinal selenium-dependent glutathione peroxidase; MP, maximal parsimony; NJ, neighbor joining; OMP, olfactory-metabolizing protein; pGPx, plasma glutathione peroxidase; PHGPX, phospholipid hydroperoxide glutathione peroxidase; Sec, selenocysteine; SECIS, selenocysteine insertion sequence; TGPx, thioredoxin GPx-like peroxidase; t-RNA, ribonucleic acid transfer; UTR, untranslated region.

References

1. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, and Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402, 1997.
2. Arai M, Imai H, Sumi D, Imanaka T, Takano T, Chiba N, and Nakagawa Y. Import into mitochondria of phospholipid hydroperoxide glutathione peroxidase requires a leader sequence. *Biochem Biophys Res Commun* 227: 433–439, 1996.
3. Aumann KD, Bedorf N, Brigelius-Flohé R, Schomburg D, and Flohé L. Glutathione peroxidase revisited: simulation of the catalytic cycle by computer-assisted molecular modelling. *Biomed Environ Sci* 10: 136–155, 1997.

4. Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, and Yeh LS. The Universal Protein Resource (UniProt). *Nucleic Acids Res* 33: D154–D159, 2005.
5. Bindewald E, Cestaro A, Hesser J, Heiler M, and Tosatto SC. MANIFOLD: protein fold recognition based on secondary structure, sequence similarity and enzyme classification. *Protein Eng* 16: 785–789, 2003.
6. Brigelius-Flohé R. Tissue-specific functions of individual glutathione peroxidases. *Free Radic Biol Med* 27: 951–965, 1999.
7. Brigelius-Flohé R. Glutathione peroxidases and redox-regulated transcription factors. *Biol Chem* 387: 1329–1335, 2006.
8. Brigelius-Flohé R, Aumann KD, Blöcker H, Gross G, Kiess M, Kloppel KD, Maiorino M, Roveri A, Schuckelt R, Ursini F, Wingender E, and Flohé L. Phospholipid-hydroperoxide glutathione peroxidase: genomic DNA, cDNA, and deduced amino acid sequence. *J Biol Chem* 269: 7342–7348, 1994.
9. Canutescu AA, Shelenkov AA, and Dunbrack RL. A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci* 12: 2001–2014, 2003.
10. Cao C, Leng Y, Huang W, Liu X, and Kufe D. Glutathione peroxidase 1 is regulated by the c-Abl and Arg tyrosine kinases. *J Biol Chem* 278: 39609–39614, 2003.
11. Chu FF, Doroshow JH, and Esworthy RS. Expression, characterization, and tissue distribution of a new cellular selenium-dependent glutathione peroxidase, GSHPx-GI. *J Biol Chem* 268: 2571–2576, 1993.
12. Copeland PR. Making sense of nonsense: the evolution of selenocysteine usage in proteins. *Genome Biol* 6: 221, 2005.
13. Cossio-Bayugar R, Miranda E, and Holman PJ. Molecular cloning of a phospholipid-hydroperoxide glutathione peroxidase gene from the tick, *Boophilus microplus* (Acari: Ixodidae). *Insect Biochem Mol Biol* 35: 1378–1387, 2005.
14. Dear TN, Campbell K, and Rabbitts TH. Molecular cloning of putative odorant-binding and odorant-metabolizing proteins. *Biochemistry* 30: 10376–10382, 1991.
15. Delaunay A, Pflieger D, Barrault MB, Vinh J, and Toledano MB. A thiol peroxidase is an H_2O_2 receptor and redox-transducer in gene activation. *Cell* 111: 471–481, 2002.
16. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform* 5: 113, 2004.
17. Epp O, Ladenstein R, and Wendel A. The refined structure of the selenoenzyme glutathione peroxidase at 0.2-nm resolution. *Eur J Biochem* 133: 51–69, 1983.
18. Felsenstein J. Phylogenies and the comparative method. *Am Naturalist* 125: 1–15, 1985.
19. Felsenstein J. PHYLIP (Phylogeny Inference Package) version 3.6. http://evolution.gs.washington.edu/phylip.html. 2002.
20. Felsenstein J. *Inferring phylogenies*. Sunderland, MA: Sinauer Associates, 2004:580.
21. Fu LH, Wang XF, Eyal Y, She YM, Donald LJ, Standing KG, and Ben-Hayyim G. A selenoprotein in the plant kingdom: mass spectrometry confirms that an opal codon (UGA) encodes selenocysteine in *Chlamydomonas reinhardtii* glutathione peroxidase. *J Biol Chem* 277: 25983–25991, 2002.
22. Ghyselinck NB, Jimenez C, Courty Y, and Dufaure JP. Androgen-dependent messenger RNA(s) related to secretory proteins in the mouse epididymis. *J Reprod Fertil* 85: 631–639, 1989.
23. Gouet P, Courcelle E, Stuart DI, and Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15: 305–308, 1999.
24. Guindon S and Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52: 696–704, 2003.
25. Hall L, Williams K, Perry AC, Frayne J, and Jury JA. The majority of human glutathione peroxidase type 5 (GPX5) transcripts are incorrectly spliced: implications for the role of GPX5 in the male reproductive tract. *Biochem J* 333: 5–9, 1998.
26. Herbette S, Lenne C, Leblanc N, Julien JL, Drevet JR, and Roeckel-Drevet P. Two GPX-like proteins from *Lycopersicon esculentum* and *Helianthus annuus* are antioxidant enzymes with phospholipid hydroperoxide glutathione peroxidase and thioredoxin peroxidase activities. *Eur J Biochem* 269: 2414–2420, 2002.
27. Herbette S, Roeckel-Drevet P, and Drevet JR. Seleno-independent glutathione peroxidases: more than simple antioxidant scavengers. *FEBS J* 274: 2163–2180, 2007.
28. Hillebrand H, Schmidt A, and Krauth-Siegel RL. A second class of peroxidases linked to the trypanothione metabolism. *J Biol Chem* 278: 6809–6815, 2003.
29. Jones DT, Taylor WR, and Thornton JM. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8: 275–282, 1992.
30. Jung BG, Lee KO, Lee SS, Chi YH, Jang HH, Kang SS, Lee K, Lim D, Yoon SC, Yun DJ, Inoue Y, Cho MJ, and Lee SY. A Chinese cabbage cDNA with high sequence identity to phospholipid hydroperoxide glutathione peroxidases encodes a novel isoform of thioredoxin-dependent peroxidase. *J Biol Chem* 277: 12572–12578, 2002.
31. Koh CS, Didierjean C, Navrot N, Panjikar S, Mulliert G, Rouhier N, Jacquot JP, Aubry A, Shawkataly O, and Corbier C. Crystal structures of a poplar thioredoxin peroxidase that exhibits the structure of glutathione peroxidases: insights into redox-driven conformational changes. *J Mol Biol* 370: 512–529, 2007.
32. Kong BW, Kim H, and Foster DN. Cloning and expression analysis of chicken phospholipid-hydroperoxide glutathione peroxidase. *Anim Biotechnol* 14: 19–29, 2003.
33. Kryukov GV, Castellano S, Novoselov SV, Lobanov AV, Zehtab O, Guigo R, and Gladyshev VN. Characterization of mammalian selenoproteomes. *Science* 300: 1439–1443, 2003.
34. Laidlaw SM and Skinner MA. Comparison of the genome sequence of FP9, an attenuated, tissue culture-adapted European strain of Fowlpox virus, with those of virulent American and European viruses. *J Gen Virol* 85: 305–322, 2004.
35. Lobanov AV, Fomenko DE, Zhang Y, Sengupta A, Hatfield DL, and Gladyshev VN. Evolutionary dynamics of eukaryotic selenoproteomes: large selenoproteomes may associate with aquatic life and small with terrestrial life. *Genome Biol* 8: R198, 2007.
36. Maiorino M, Aumann KD, Brigelius-Flohé R, Doria D, van den Heuvel J, McCarthy J, Roveri A, Ursini F, and Flohé L. Probing the presumed catalytic triad of selenium-containing peroxidases by mutational analysis of phospholipid hydroperoxide glutathione peroxidase (PHGPx). *Biol Chem Hoppe Seyler* 376: 651–660, 1995.
37. Maiorino M, Mauri P, Roveri A, Benazzi L, Toppo S, Bosello V, and Ursini F. Primary structure of the nuclear forms of phospholipid hydroperoxide glutathione peroxidase (PHGPx) in rat spermatozoa. *FEBS Lett* 579: 667–670, 2005.
38. Maiorino M, Roche C, Kiess M, Koenig K, Gawlik D, Matthes M, Naldini E, Pierce R, and Flohé L. A selenium-containing phospholipid-hydroperoxide glutathione peroxidase in *Schistosoma mansoni*. *Eur J Biochem* 238: 838–844, 1996.

39. Maiorino M, Roveri A, Benazzi L, Bosello V, Mauri P, Toppo S, Tosatto SC, and Ursini F. Functional interaction of phospholipid hydroperoxide glutathione peroxidase with sperm mitochondrion-associated cysteine-rich protein discloses the adjacent cysteine motif as a new substrate of the selenoperoxidase. *J Biol Chem* 280: 38395–38402, 2005.
40. Maiorino M, Roveri A, Ursini F, and Gregolin C. Enzymatic determination of membrane lipid peroxidation. *J Free Radic Biol Med* 1: 203–207, 1985.
41. Maiorino M, Scapin M, Ursini F, Biasolo M, Bosello V, and Flohé L. Distinct promoters determine alternative transcription of GPx-4 into phospholipid-hydroperoxide glutathione peroxidase variants. *J Biol Chem* 278: 34286–34290, 2003.
42. Maiorino M, Ursini F, Bosello V, Toppo S, Tosatto SC, Mauri P, Becker K, Roveri A, Bulato C, Benazzi L, De Palma A, and Flohé L. The thioredoxin specificity of *Drosophila* GPx: a paradigm for a peroxiredoxin-like mechanism of many glutathione peroxidases. *J Mol Biol* 365: 1033–1046, 2007.
43. Maiorino M, Wissing JB, Brigelius-Flohé R, Calabrese F, Roveri A, Steinert P, Ursini F, and Flohé L. Testosterone mediates expression of the selenoprotein PHGPx by induction of spermatogenesis and not by direct transcriptional gene activation. *FASEB J* 12: 1359–1370, 1998.
44. Martin JL. Thioredoxin: a fold for all reasons. *Structure* 3: 245–250, 1995.
45. Mauri P, Benazzi L, Flohé L, Maiorino M, Pietta PG, Pilawa S, Roveri A, and Ursini F. Versatility of selenium catalysis in PHGPx unraveled by LC/ESI-MS/MS. *Biol Chem* 384: 575–588, 2003.
46. Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, Witman GB, Terry A, Salamov A, Fritz-Laylin LK, Marechal-Drouard L, Marshall WF, Qu LH, Nelson DR, Sanderfoot AA, Spalding MH, Kapitonov VV, Ren Q, Ferris P, Lindquist E, Shapiro H, Lucas SM, Grimwood J, Schmutz J, Cardol P, Cerutti H, Chanfreau G, Chen CL, Cognat V, Croft MT, Dent R, Dutcher S, Fernandez E, Fukuzawa H, Gonzalez-Ballester D, Gonzalez-Halphen D, Hallmann A, Hanikenne M, Hippler M, Inwood W, Jabbari K, Kalanon M, Kuras R, Lefebvre PA, Lemaire SD, Lobanov AV, Lohr M, Manuell A, Meier I, Mets L, Mittag M, Mittelmeier T, Moroney JV, Moseley J, Napoli C, Nedelcu AM, Niyogi K, Novoselov SV, Paulsen IT, Pazour G, Purton S, Ral JP, Riano-Pachon DM, Riekhof W, Rymarquis L, Schroda M, Stern D, Umen J, Willows R, Wilson N, Zimmer SL, Allmer J, Balk J, Bisova K, Chen CJ, Elias M, Gendler K, Hauser C, Lamb MR, Ledford H, Long JC, Minagawa J, Page MD, Pan J, Pootakham W, Roje S, Rose A, Stahlberg E, Terauchi AM, Yang P, Ball S, Bowler C, Dieckmann CL, Gladyshev VN, Green P, Jorgensen R, Mayfield S, Mueller-Roeber B, Rajamani S, Sayre RT, Brokstein P, Dubchak I, Goodstein D, Hornick L, Huang YW, Jhaveri J, Luo Y, Martinez D, Ngau WC, Otillar B, Poliakov A, Porter A, Szajkowski L, Werner G, Zhou K, Grigoriev IV, Rokhsar DS, and Grossman AR. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318: 245–250, 2007.
47. Mills GC. Hemoglobin catabolism, I: glutathione peroxidase, an erythrocyte enzyme which protects hemoglobin from oxidative breakdown. *J Biol Chem* 229: 189–197, 1957.
48. Missirlis F, Rahlfs S, Dimopoulos N, Bauer H, Becker K, Hilliker A, Phillips JP, and Jaeckle H. A putative glutathione peroxidase of *Drosophila* encodes a thioredoxin peroxidase that provides resistance against oxidative stress but fails to complement a lack of catalase activity. *Biol Chem* 384: 463–472, 2003.
49. Navrot N, Collin V, Gualberto J, Gelhaye E, Hirasawa M, Rey P, Knaff DB, Issakidis E, Jacquot JP, and Rouhier N. Plant glutathione peroxidases are functional peroxiredoxins distributed in several subcellular compartments and regulated during biotic and abiotic stresses. *Plant Physiol* 142: 1364–1379, 2006.
50. Page RD. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12: 357–358, 1996.
51. Passardi F, Theiler G, Zamocky M, Cosio C, Rouhier N, Teixera F, Margis-Pinheiro M, Ioannidis V, Penel C, Falquet L, and Dunand C. PeroxiBase: the peroxidase database. *Phytochemistry* 68: 1605–1611, 2007.
52. Perriere G, and Gouy M. WWW-Query: an on-line retrieval system for biological sequence banks. *Biochimie* 78: 364–369, 1996.
53. Pfeifer H, Conrad M, Roethlein D, Kyriakopoulos A, Brielmeier M, Bornkamm GW, and Behne D. Identification of a specific sperm nuclei selenoenzyme necessary for protamine thiol cross-linking during sperm maturation. *FASEB J* 15: 1236–1238, 2001.
54. Romero H, Zhang Y, Gladyshev VN, and Salinas G. Evolution of selenium utilization traits. *Genome Biol* 6: R66, 2005.
55. Rotruck JT, Pope AL, Ganther HE, and Hoekstra WG. Prevention of oxidative damage to rat erythrocytes by dietary selenium. *J Nutr* 102: 689–696, 1972.
56. Roveri A, Casasco A, Maiorino M, Dalan P, Calligaro A, and Ursini F. Phospholipid hydroperoxide glutathione peroxidase of rat testis: gonadotropin dependence and immunocytochemical identification. *J Biol Chem* 267: 6142–6146, 1992.
57. Seshadri R, Myers GS, Tettelin H, Eisen JA, Heidelberg JF, Dodson RJ, Davidsen TM, DeBoy RT, Fouts DE, Haft DH, Selengut J, Ren Q, Brinkac LM, Madupu R, Kolonay J, Durkin SA, Daugherty SC, Shetty J, Shvartsbeyn A, Gebregeorgis E, Geer K, Tsegaye G, Malek J, Ayodeji B, Shatsman S, McLeod MP, Smajs D, Howell JK, Pal S, Amin A, Vashisth P, McNeill TZ, Xiang Q, Sodergren E, Baca E, Weinstock GM, Norris SJ, Fraser CM, and Paulsen IT. Comparison of the genome of the oral pathogen *Treponema denticola* with other spirochete genomes. *Proc Natl Acad Sci U S A* 101: 5646–5651, 2004.
58. Shindyalov IN and Bourne PE. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng* 11: 739–747, 1998.
59. Singh A and Rathaur S. Identification and characterization of a selenium-dependent glutathione peroxidase in *Setaria cervi*. *Biochem Biophys Res Commun* 331: 1069–1074, 2005.
60. Sommer I, Toppo S, Sander O, Lengauer T, and Tosatto SC. Improving the quality of protein structure models by selecting from alignment alternatives. *BMC Bioinform* 7: 364, 2006.
61. Strimmer K, Goldman N, and von Haeseler A. Bayesian probabilities and quartet puzzling. *Mol Biol Evolution* 14: 210–211, 1997.
62. Strimmer K and von Haeseler A. Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol Biol Evolution* 13: 964–969, 1996.
63. Sztajer H, Gamain B, Aumann KD, Slomianny C, Becker K, Brigelius-Flohé R, and Flohé L. The putative glutathione peroxidase gene of *Plasmodium falciparum* codes for a thioredoxin peroxidase. *J Biol Chem* 276: 7397–7403, 2001.
64. Takahashi K, Akasaka M, Yamamoto Y, Kobayashi C, Mizoguchi J, and Koyama J. Primary structure of human plasma glutathione peroxidase deduced from cDNA sequences. *J Biochem (Tokyo)* 108: 145–148, 1990.

65. Tanaka T, Izawa S, and Inoue Y. GPX2, encoding a phospholipid hydroperoxide glutathione peroxidase homologue, codes for an atypical 2-Cys peroxiredoxin in *Saccharomyces cerevisiae*. *J Biol Chem* 280: 42078–42087, 2005.
66. Thisse C, Degrave A, Kryukov GV, Gladyshev VN, Obrecht-Pflumio S, Krol A, Thisse B, and Lescure A. Spatial and temporal expression patterns of selenoprotein genes during embryogenesis in zebrafish. *Gene Expr Patterns* 3: 525–532, 2003.
67. Thompson JD, Higgins DG, and Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673–4680, 1994.
68. Tosatto SC. The Victor/FIRST. *J Comput Biol* 12: 1316–1327, 2005.
69. Tosatto SC, Bindewald E, Hesser J, and Manner R. A divide and conquer approach to fast loop modeling. *Protein Eng* 15: 279–286, 2002.
70. Tosatto SC, Bosello V, Fogolari F, Mauri P, Roveri A, Toppo S, Flohé L, Ursini F, and Maiorino M. The catalytic site of glutathione peroxidases. *Antioxid Redox Signal* (This issue) 2008.
71. Ursini F, Heim S, Kiess M, Maiorino M, Roveri A, Wissing J, and Flohé L. Dual function of the selenoprotein PHGPx during sperm maturation. *Science* 285: 1393–1396, 1999.
72. Ursini F, Maiorino M, Brigelius-Flohé R, Aumann KD, Roveri A, Schomburg D, and Flohé L. Diversity of glutathione peroxidases. *Methods Enzymol* 252: 38–53, 1995.
73. Ursini F, Maiorino M, and Gregolin C. The selenoenzyme phospholipid hydroperoxide glutathione peroxidase. *Biochim Biophys Acta* 839: 62–70, 1985.
74. Utomo A, Jiang X, Furuta S, Yun J, Levin DS, Wang YC, Desai KV, Green JE, Chen PL, and Lee WH. Identification of a novel putative non-selenocysteine containing phospholipid hydroperoxide glutathione peroxidase (NPGPx) essential for alleviating oxidative stress generated from polyunsaturated fatty acids in breast cancer cells. *J Biol Chem* 279: 43522–43529, 2004.
75. Vernet P, Rock E, Mazur A, Rayssiguier Y, Dufaure JP, and Drevet JR. Selenium-independent epididymis-restricted glutathione peroxidase 5 protein (GPX5) can back up failing Se-dependent GPXs in mice subjected to selenium deficiency. *Mol Reprod Dev* 54: 362–370, 1999.
76. Walczak R, Westhof E, Carbon P, and Krol A. A novel RNA structural motif in the selenocysteine insertion element of eukaryotic selenoprotein mRNAs. *RNA* 2: 367–379, 1996.
77. Zhang Y, Romero H, Salinas G, and Gladyshev VN. Dynamic evolution of selenocysteine utilization in bacteria: a balance between selenoprotein loss and evolution of selenocysteine from redox active cysteine residues. *Genome Biol* 7: R94, 2006.
78. Zhao L, Cox AG, Ruzicka JA, Bhat AA, Zhang W, and Taylor EW. Molecular modeling and in vitro activity of an HIV-1-encoded glutathione peroxidase. *Proc Natl Acad Sci U S A* 97: 6356–6361, 2000.

Address reprint requests to:
Stefano Toppo
Department of Biological Chemistry
Viale G. Colombo 3
I-35131 Padova, Italy

E-mail: stefano.toppo@unipd.it

Date of first submission to ARS Central, February 25, 2008; date of acceptance, March 6, 2008.

This article has been cited by:

1. Marcel Deponte. 2012. Glutathione catalysis and the reaction mechanisms of glutathione-dependent enzymes. *Biochimica et Biophysica Acta (BBA) - General Subjects* . [CrossRef]
2. Prakash M. Gopalakrishnan Nair, Sun Young Park, Jinhee Choi. 2012. Characterization and expression analysis of phospholipid hydroperoxide glutathione peroxidase cDNA from Chironomus riparius on exposure to cadmium. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **163**:1, 37-42. [CrossRef]
3. Ming J. Wu, Peter J. Rogers, Frank M. Clarke. 2012. 125th Anniversary Review: The role of proteins in beer redox stability. *Journal of the Institute of Brewing* **118**:1, 1-11. [CrossRef]
4. Taichi Kakihana , Kazuhiro Nagata , Roberto Sitia . 2012. Peroxides and Peroxidases in the Endoplasmic Reticulum: Integrating Redox Homeostasis and Oxidative Folding. *Antioxidants & Redox Signaling* **16**:8, 763-771. [Abstract] [Full Text HTML] [Full Text PDF] [Full Text PDF with Links]
5. Peter A. Bain, Kathryn A. Schuller. 2012. A glutathione peroxidase 4 (GPx4) homologue from southern bluefin tuna is a secreted protein: First report of a secreted GPx4 isoform in vertebrates. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* . [CrossRef]
6. Federico Fogolari, Alessandra Corazza, Stefano Toppo, Silvio C. E. Tosatto, Paolo Viglino, Fulvio Ursini, Gennaro Esposito. 2012. Studying Interactions by Molecular Dynamics Simulations at High Concentration. *Journal of Biomedicine and Biotechnology* **2012**, 1-9. [CrossRef]
7. Peter A. Bain, Kathryn A. Schuller. 2011. Molecular cloning of glutathione peroxidase cDNAs from Seriola lalandi and analysis of changes in expression in cultured fibroblast-like cells in response to tert-butyl hydroquinone. *Aquaculture* . [CrossRef]
8. Leopold Flohé , Stefano Toppo , Giorgio Cozza , Fulvio Ursini . 2011. A Comparison of Thiol Peroxidase Mechanisms. *Antioxidants & Redox Signaling* **15**:3, 763-780. [Abstract] [Full Text HTML] [Full Text PDF] [Full Text PDF with Links]
9. Heidi Marja Viitaniemi, Erica Helen Leder. 2011. Sex-Biased Protein Expression in Threespine Stickleback, Gasterosteus aculeatus. *Journal of Proteome Research* 110721130837015. [CrossRef]
10. A. J. M. Martin, M. Vidotto, F. Boscariol, T. Di Domenico, I. Walsh, S. C. E. Tosatto. 2011. RING: networking interacting residues, evolutionary information and energetics in protein structures. *Bioinformatics* **27**:14, 2003-2005. [CrossRef]
11. Naser A. Anjum, Iqbal Ahmad, Iram Mohmood, Mário Pacheco, Armando C. Duarte, Eduarda Pereira, Shahid Umar, Altaf Ahmad, Nafees A. Khan, Muhammad Iqbal, M.N.V. Prasad. 2011. Modulation of glutathione and its related enzymes in plants' responses to toxic metals and metalloids - a review. *Environmental and Experimental Botany* . [CrossRef]
12. Neil J. Bulleid, Lars Ellgaard. 2011. Multiple ways to make disulfides. *Trends in Biochemical Sciences* . [CrossRef]
13. Shi Kui Wang, Jeremy D. Weaver, Sheng Zhang, Xin Gen Lei. 2011. Knockout of SOD1 promotes conversion of selenocysteine to dehydroalanine in murine hepatic GPX1 protein. *Free Radical Biology and Medicine* **51**:1, 197-204. [CrossRef]
14. Maximo A. Benavides, Dong Hu, Marie Kristine Baraoidan, Annette Bruno, Pan Du, Simon Lin, Wancai Yang, Kirby I. Bland, William E. Grizzle, Maarten C. Bosland. 2011. L-methionine-induced alterations in molecular signatures in MCF-7 and LNCaP cancer cells. *Journal of Cancer Research and Clinical Oncology* **137**:3, 441-453. [CrossRef]
15. Qingning Liang, Yuchen Sheng, Ping Jiang, Lili Ji, Yuye Xia, Yang Min, Zhengtao Wang. 2011. The gender-dependent difference of liver GSH antioxidant system in mice and its influence on isoline-induced liver injury. *Toxicology* **280**:1-2, 61-69. [CrossRef]
16. Van Dat Nguyen, Mirva J. Saaranen, Anna-Riikka Karala, Anna-Kaisa Lappi, Lei Wang, Irina B. Raykhel, Heli I. Alanen, Kirsi E.H. Salo, Chih-chen Wang, Lloyd W. Ruddock. 2011. Two Endoplasmic Reticulum PDI Peroxidases Increase the Efficiency of the Use of Peroxide during Disulfide Bond Formation. *Journal of Molecular Biology* **406**:3, 503-515. [CrossRef]
17. Christina L. Takanishi, Li-Hua Ma, Matthew J. Wood. 2010. The role of active site residues in the oxidant specificity of the Orp1 thiol peroxidase#. *Biochemical and Biophysical Research Communications* **403**:1, 46-51. [CrossRef]
18. Astrid Borchert, Gabi Küttner, Elke Gießmann, Chi Chiu Wang, Helga Wessner, Rudolf Volkmer, Wolfgang Höhne, Hartmut Kuhn. 2010. Defining the immunoreactive epitope for the monoclonal anti-human glutathione peroxidase-4 antibody anti-hGPx4 Mab63-1. *Immunology Letters* **133**:2, 85-93. [CrossRef]
19. Felipe A. Arenas, Waldo A. Díaz, Carolina A. Leal, José M. Pérez-Donoso, James A. Imlay, Claudio C. Vásquez. 2010. The Escherichia coli btuE gene, encodes a glutathione peroxidase that is induced under oxidative stress conditions. *Biochemical and Biophysical Research Communications* **398**:4, 690-694. [CrossRef]

20. Vanja Radiši# Biljak, Lada Rumora, Ivana #epelak, Dolores Pancirov, Sanja Popovi#-Grle, Jasna Sori#, Tihana Žani# Grubiši#. 2010. Glutathione cycle in stable chronic obstructive pulmonary disease. *Cell Biochemistry and Function* **28**:6, 448-453. [CrossRef]
21. Mian Wang, Mingjiang Kang, Xingqi Guo, Baohua Xu. 2010. Identification and characterization of two phospholipid hydroperoxide glutathione peroxidase genes from Apis cerana cerana. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* **152**:1, 75-83. [CrossRef]
22. Xi-Lei LI, Gui-Ling WANG, Jia-Le LI. 2010. Full-length cDNA cloning and encoding protein structure analysis of GPX in Hyriopsis cumingii. *Hereditas (Beijing)* **32**:4, 360-368. [CrossRef]
23. Stefano Toppo, Leopold Flohé, Fulvio Ursini, Stefano Vanin, Matilde Maiorino. 2009. Catalytic mechanisms and specificities of glutathione peroxidases: Variations of a basic scheme. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1790**:11, 1486-1500. [CrossRef]
24. Sebastien A. Rider, Simon J. Davies, Awadhesh N. Jha, Andrew A. Fisher, Jan Knight, John W. Sweetman. 2009. Supra-nutritional dietary intake of selenite and selenium yeast in normal and stressed rainbow trout (Oncorhynchus mykiss): Implications on selenium status and health responses. *Aquaculture* **295**:3-4, 282-291. [CrossRef]
25. J. A. Bertout, A. J. Majmundar, J. D. Gordan, J. C. Lam, D. Ditsworth, B. Keith, E. J. Brown, K. L. Nathanson, M. C. Simon. 2009. HIF2 inhibition promotes p53 pathway activity, tumor cell death, and radiation responses. *Proceedings of the National Academy of Sciences* **106**:34, 14391-14396. [CrossRef]