Cloning and characterization of AmphiPrxV, a new member of peroxiredoxin family from the amphioxus *Branchiostoma belcheri tsingtauense*

S. Zhang[†,*], Z. Liu[†], J. Yuan[†], M. Liu[†], M. S. Sawant[†] and A. Xu[#]

[†]Department of Marine Biology, Ocean University of Qingdao, Qingdao 266003, P.R. China
[#]College of Life Science, Zhongshan University, Guangzhou, 510275, P.R. China

An amphioxus cDNA, *AmphiPrxV*, encoding peroxiredoxin V, was isolated from a gut cDNA library of *Branchiostoma belcheri tsingtauense*. It contained 187 amino acid residues deduced from the nucleotide sequence, and showed 65% amino acid sequence identity with human AOEB166. The analysis of N- and C-terminal domains revealed amino acid sequences characteristic of features of mitochondrial and peroxisomal targeting sequences. Alignment of the amino acid sequences of different subtypes of peroxiredoxin proteins indicated that the amino acids surrounding the conserved Cys78 in AmphiPrxV were PGCS, which was neither similar to PVCT in 1-Cys nor to FVCP in 2-Cys. Also, AmphiPrxV contained a third conserved cysteine, Cys103. Phylogenetic analysis showed that PrxV members, including AmphiPrxV, were clustered with neither 1-Cys group nor 2-Cys group, forming a distinct group. It appears that PrxV represents a new group of the peroxiredoxin family.

PEROXIREDOXINS (Prxs) are a newly recognized family of antioxidant proteins that protect organisms against cellular damages caused by the oxygen radicals such as alkyl hydroperoxides, hydroperoxides and hydroxyls generated by oxidation processes. All the Prx proteins studied so far contain a highly conserved structural motif that surrounds a cysteine residue, which corresponds to Cys47 in yeast Prx, forming the peroxidatic centre[1]. They are widely expressed in animal tissues and well conserved throughout evolution[2–5], suggesting a fundamental and essential oxidative homeostatic function for Prxs in living organisms. Prxs have been implicated via their antioxidant activity, in a number of cellular functions including cell proliferation and differentiation, protection of other proteins from oxidative damage and intracellular signalling[2,6–8].

The family of Prxs was initially identified in yeast as a 27 kDa protein[9], and was subsequently demonstrated to be a ubiquitous protein in organisms from bacteria to human[1,10]. Recently, new members of the Prx family with distinct primary sequences, and activities and functions have been discovered, and multiple subtypes of Prx are frequently found in one species: for example, three subtypes of Prx exist in *Escherichia coli*, five in budding yeast, and six (PrxI–PrxVI) in mammals to date[1–3,10–14]. Amphioxus or lancelet, a cephalochordate, has long been regarded as the living invertebrate most closely related to the proximate invertebrate ancestor of vertebrates. It has been recognized as the most important model animal to analyse the origin and evolution of vertebrates[15,16]. However, the study of Prxs in this evolutionarily important animal remains untouched hitherto. In the present study, we report the cloning and characterization of amphioxus peroxiredoxin V, AmphiPrxV, with mitochondrial and peroxisomal sorting signals. We also provide evidences that PrxV represents a new group of the Prx family.

Table 1. Representative members of the peroxiredoxin protein family

Protein	Organism of source	Accession number in GenBank	Amino acid
PhPrxV	*Papio hamadryas*	AAG13451	215
BtPrxV	*Bos taurus*	AF305564	220
DmPrxV	*Drosophila melanogaster*	AAF55497	157
hAOEB166	*Homo sapiens*	AF106944	214
MmPrxV	*Mus musculus*	NP_036151	210
hTPX	*H. sapiens*	CAB62210	162
CaPrxV	*Cercopithecus aethiops*	AAG13453	215
SsPrxV	*Sus scrofa*	AAG13452	162
rAOEB166	*Rattus norvegicus*	AF110732	214
hAOE37-2	*H. sapiens*	U25182	271
hAop1	*H. sapiens*	D49396	256
hNKEFB	*H. sapiens*	L19185	198
hORF06	*H. sapiens*	D14662	224
hPAG	*H. sapiens*	X67951	199
hTSA	*H. sapiens*	Z22548	198
mAOP1	*M. musculus*	M28723	257
mGPx	*M. musculus*	Y12883	224
mOSF-3	*M. musculus*	D21252	199
mTSA	*M. musculus*	X82067	198
raiPLA2	*R. norvegicus*	AF014009	224
rHBP23	*R. norvegicus*	D30035	199
rPrxIII	*R. norvegicus*	AF106944	257
rPrxIV	*R. norvegicus*	AF106945	273
rTSA	*R. norvegicus*	U06099	198
AtPrx	*Arabidopsis thaliana*	Y12089	216
ScPrx1	*Saccharomyces cerevisiae*	L14640	196
ScPrx2	*S. cerevisiae*	P34227	261
Ov-tpx-2	*Onchocerca volvulus*	AF029247	199
BtaiPLA2	*B. taurus*	AF090194	224
DPx-6005	*D. melanogaster*	AF311878	222
DmSTP	*D. melanogaster*	AF321614	242
ABP-25	*Cynops pyrrhogaster*	D37808	200
RAB24	*Oryza sativa*	D63917	220
RBT-NKEF	*Oncorhynchus mykiss*	U27125	200
OvTSA	*O. volvulus*	U31052	232
SsPrx	*Sulfolobus* sp.	U36479	215
Bm-TPx-2	*Brugia malayi*	U47100	199
HvPrx	*Hordeum vulgare*	X96551	218
DiPxn	*Dirofilaria immitis*	AF027387	235
DiTSA	*D. immitis*	AF001007	199

*For correspondence. (e-mail: zscshab@public.qd.sd.cn)

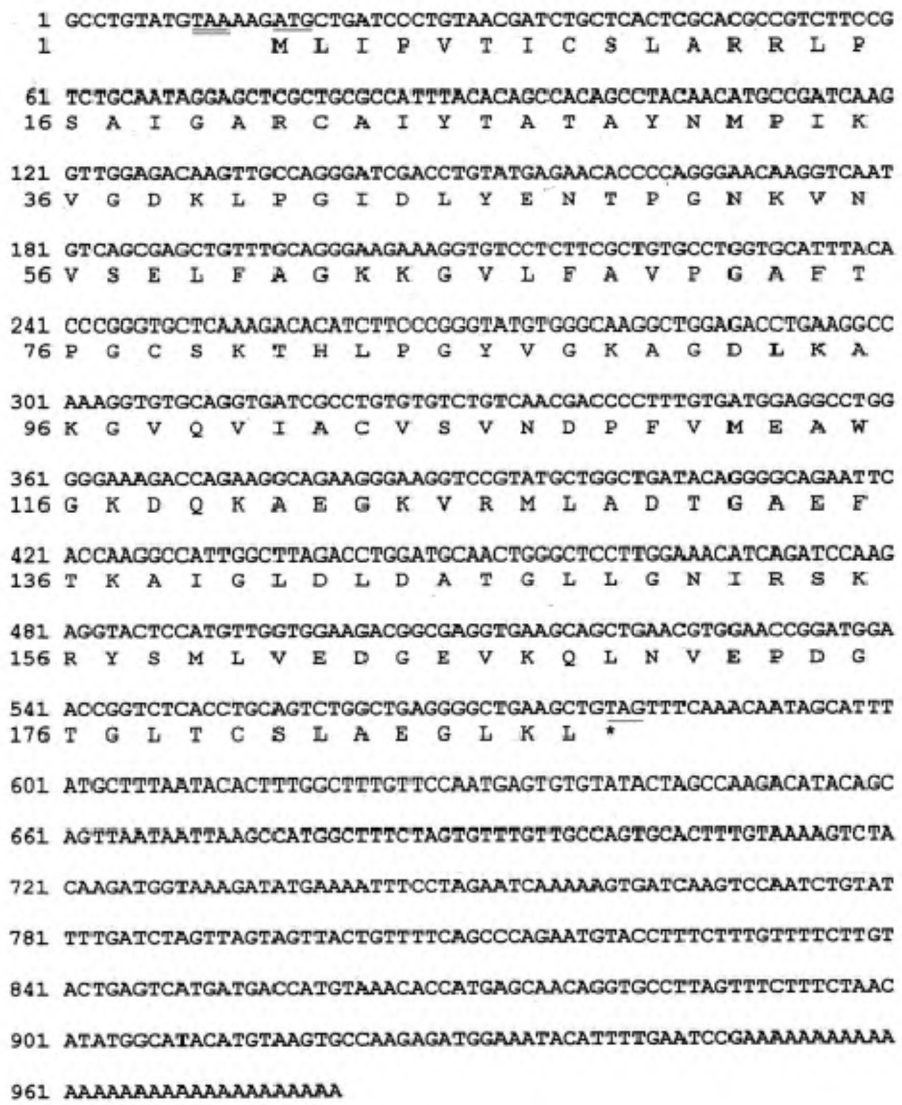

Figure 1. Nucleotide and deduced amino acid sequences of *AmphiPrxV* (accession number in GenBank: AF498232). Translational start and terminal sites are underlined, and the asterisk represents the stop codon. In-frame stop codon within the 5′ UTR is double underlined.

Adult amphioxus *Branchiostoma belcheri tsingtauense* were collected from the sandy bottom of the sea near Shazikou, Qingdao, China and cultured in containers with continuous aeration. They were starved for two days in sterilized, filtered sea water before dissection to remove all food in the gut. The guts were dissected out and rinsed three times with 50 mM Tris-HCl containing 100 mM NaCl, pH 7.5 and then frozen immediately in liquid nitrogen until use.

Total RNA was isolated from the guts of amphioxus with TRIZOL reagent (GIBCO-BRL, Gaithersburg, MD). Gut cDNA library of adult amphioxus was constructed with SMART cDNA Library Construction Kit (CLONTECH, Palo Alto, CA, USA), according to the method described previously[17]. Synthesized cDNA was ligated into pcDNA3-sfiI vector which was slightly modified from pcDNA3 vector (Invitrogen Inc.), and transformed into *E. coli* DH5αcells.

cDNA clones were randomly selected for sequencing. The insert length of each selected clone was examined by PCR with primers T7 (5′-TAATACGACTCACTATAGGGA-3′) and SP6 (5′-ATTTAGGTGACACTATAGAA-3′), prior to plasmid DNA preparation. Both strands of all selected clones were sequenced with ABI PRISM 377XL DNA Sequencer, and all sequences were then analysed for coding probability with DNATools program[18].

Initial comparison against the GenBank protein database was performed using the BLAST network server at the National Center for Biotechnology Information[19,20]. Multiple protein sequences were aligned and the phylogenetic tree was constructed using the MegAlign program by the CLUSTAL method in DNASTAR[21,22].

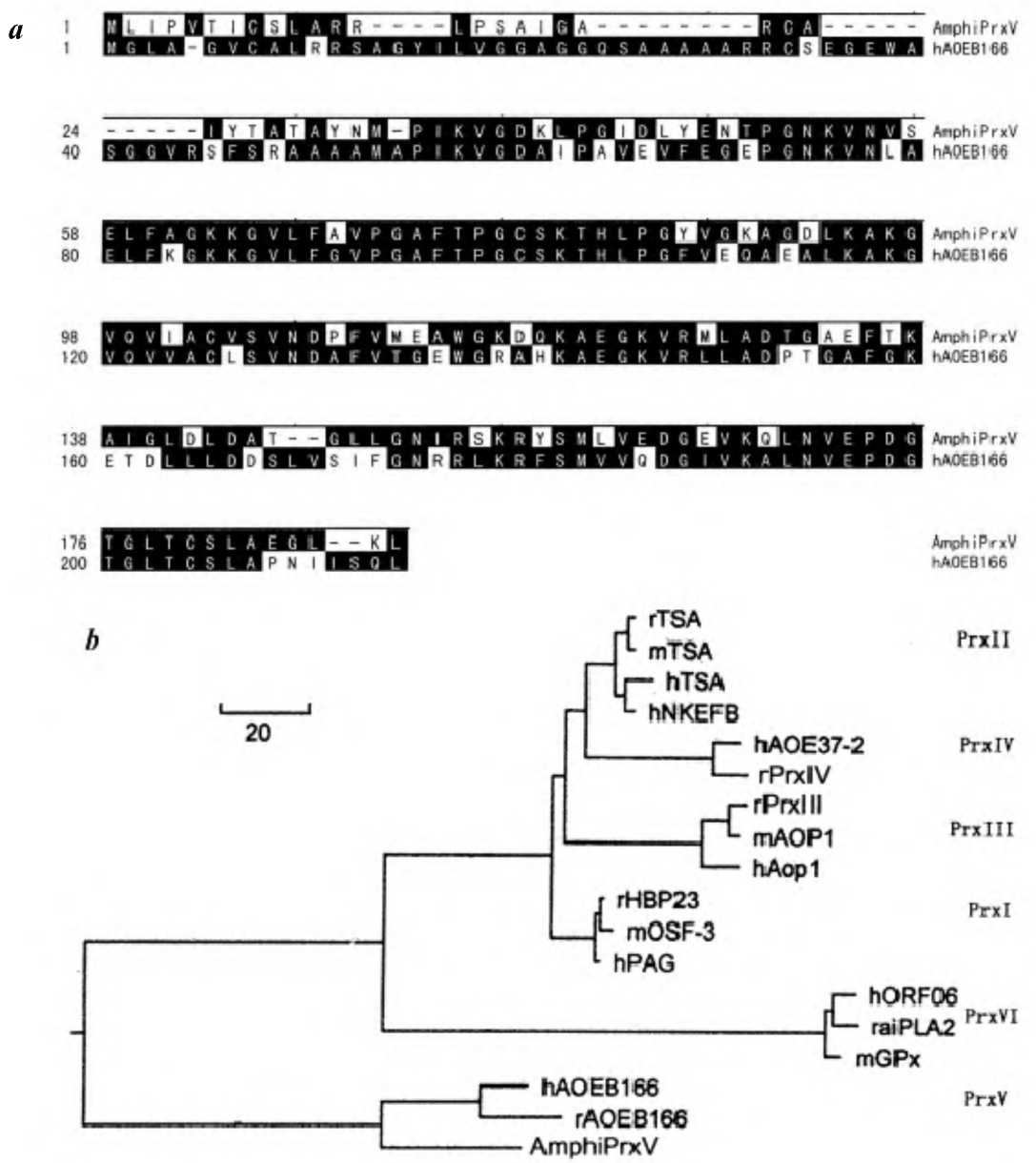

Figure 2. Amino acid sequence alignment of AmphiPrxV and human AOEB166 using the MegAlign program (DNASTAR) by the CLUSTAL method (***a***), and phylogenetic analysis of AmphiPrxV and identified different subtypes of mammalian Prxs (***b***). Shaded (with solid black) amino acids residues are those that match the consensus. Gaps introduced into sequences to optimize alignments are represented by (–). The phylogenetic tree was generated using the MegAlign program (DNASTAR) by the CLUSTAL method, and the branch length represents the evolutionary distance. See Table 1 for sequence references.

Accession numbers of the Prx protein sequences in the GenBank database used for comparison are listed in Table 1.

The cDNA sequence obtained from the clone 089 selected randomly was 980 nucleotides in length, and encoded a protein of 187 amino acid residues with a predicted molecular weight of about 19.8 kDa (Figure 1). There exists an initiator methionine codon at the 5′ end and a stop codon at the 3′ end, followed by a long 3′ untranslated region (UTR) with a polyadenine stretch. Besides, the 5′ UTR contained an in-frame stop codon upstream from the first start codon ATG. The cDNA thus contains a complete open reading frame.

The initial database searches by BLAST suggested an affinity between the amphioxus protein and human AOEB166 protein with 65% identity and 76% similarity (Figure 2 *a*). The phylogenetic tree constructed from the complete sequences of the amphioxus protein and identified members of mammalian Prxs indicated that the amphioxus protein was also in the PrxV clade along with

human AOEB166 and rat AOEB166 (Figure 2 *b*); the amphioxus protein was accordingly designated AmphiPrxV (GenBank accession number: AF498232).

Human AOEB166 and rat AOEB166 possess mitochondrial presequence features[3]. Analysis of the N-terminal 30 amino acid residues of AmphiPrxV using motif search (PSORT, Version 6.3, World Wide Web) revealed that it was a hydrophobic domain, which equalled a signal localized to mitochondria. This is in line with the fact that mitochondria are the sites of high level of oxidative activity, and antioxidants such as SOD, GST, catalase and members of the PrxIII subtype have been identified in mitochondria matrix[23]. In addition, a peroxisomal targeting sequence, LKL, of peroxisomal targeting signal 1 (PTS1) family was detected at the C-terminus in AmphiPrxV. Interestingly, a comparable PTS1 sequence (LKL) has been identified at the C-terminus in yeast peroxisomal matrix protein, but not in mammals[24]. PTS1 was initially defined in mammalian cells using various tripeptides extending from heterologous reporter proteins, and has now been found in a majority of peroxisomal matrix proteins. It is likely that AmphiPrxV represents a member of the Prx family with mitochondrial and peroxisomal sorting signals. Further studies are needed to define the intracellular location and functional significance of AmphiPrxV in organelles.

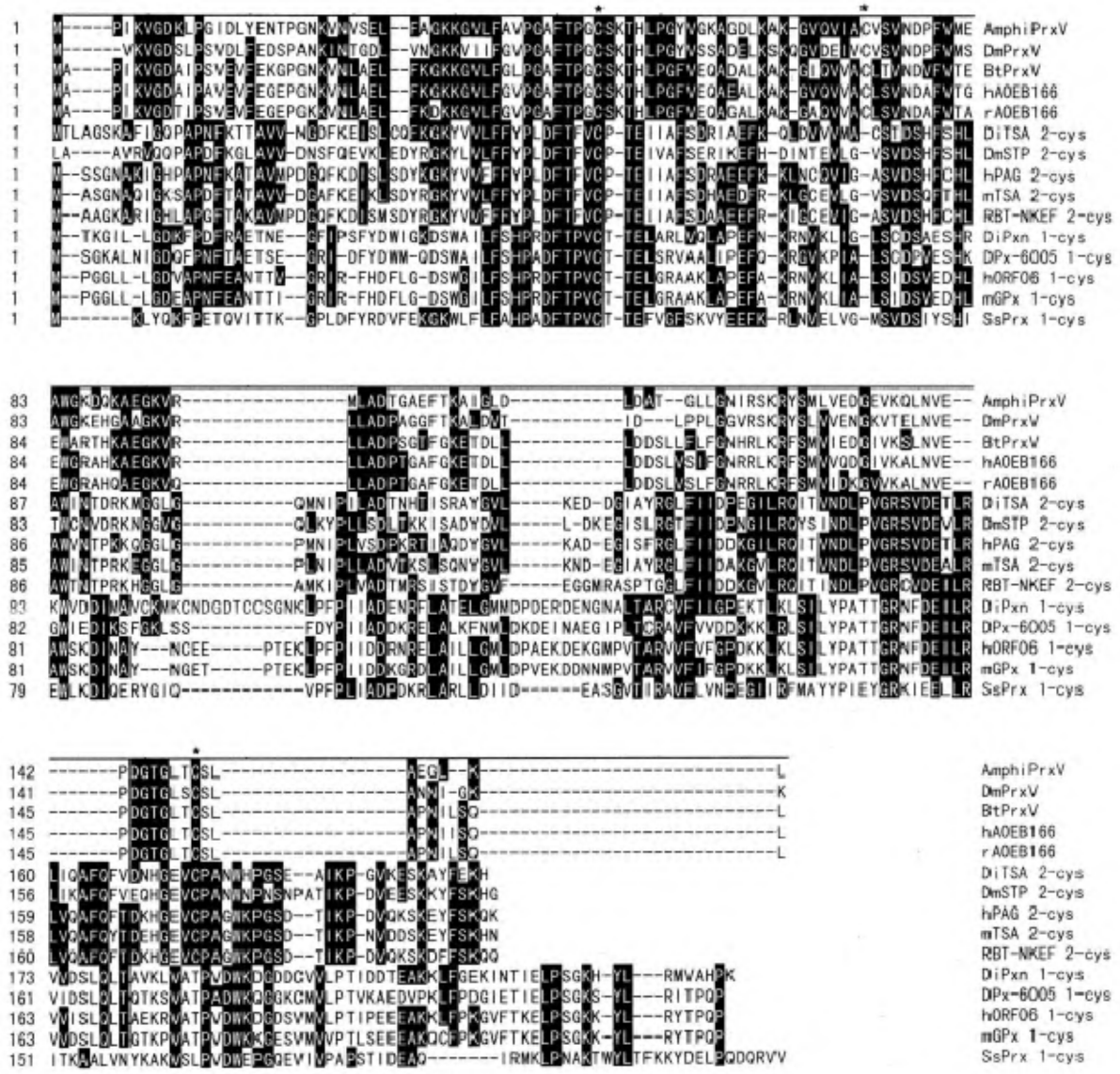

Figure 3 *a*. Amino acid sequence alignment of representative members of both 1-Cys and 2-Cys as well as PrxV proteins without their predicated mitochondrial presequences using the MegAlign program (DNASTAR) by the CLUSTAL method. Shaded (with solid black) amino acids residues are those that match the consensus, and the asterisk represents the conserved Cys. Gaps introduced into sequences to optimize alignments are represented by (–). See Table 1 for sequence references.

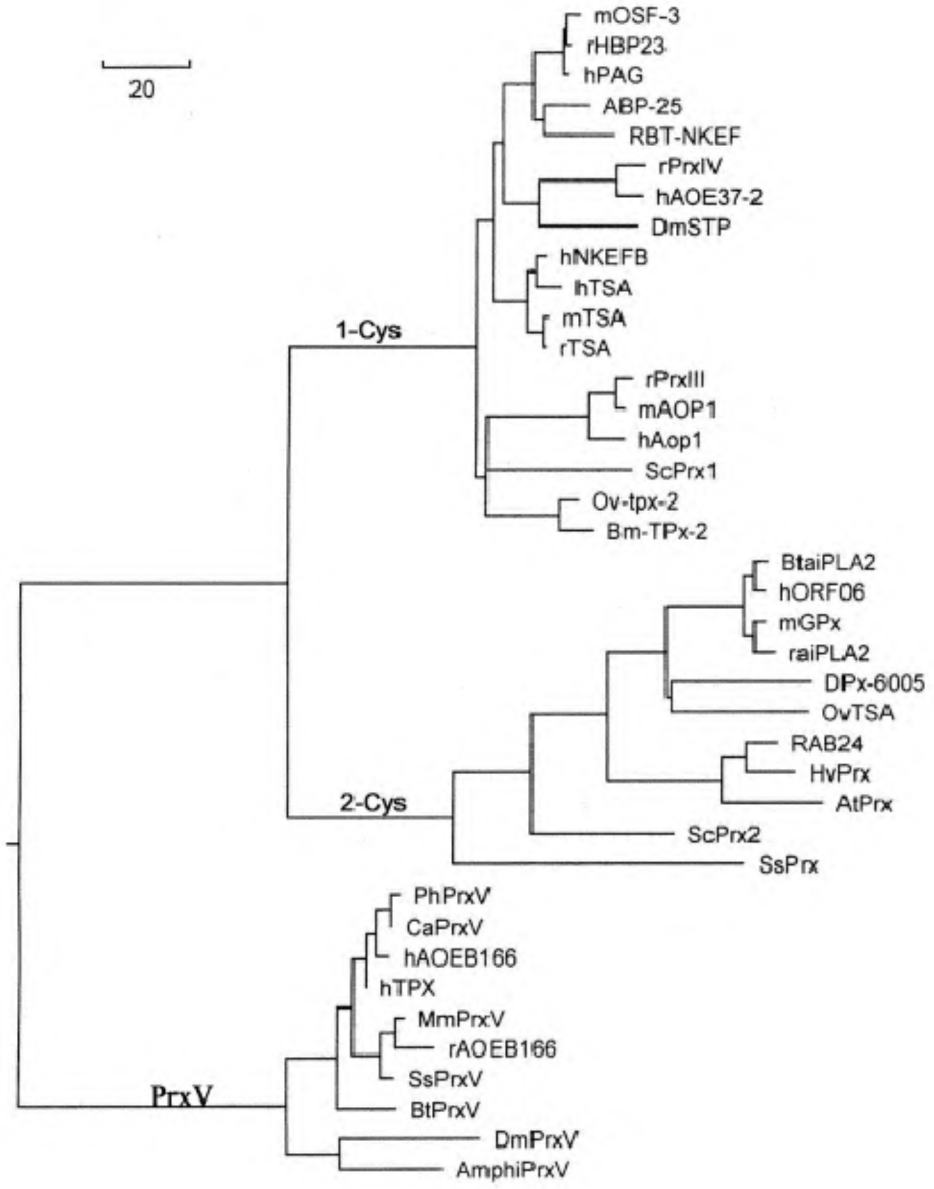

Figure 3 *b*. Phylogenetic tree constructed from the complete amino acid sequences of peroxiredoxin family. The tree branch length represents the evolutionary distance. See Table 1 for sequence references.

Prxs are an emerging family of multifunctional enzymes and have been classified into two groups, 1-Cys and 2-Cys Prxs, on the basis of presence of either one or two highly conserved cysteine residues corresponding to Cys47 and Cys170 in yeast Prxs respectively[25]. 2-Cys Prxs contain both conserved cysteines, while members of the 1-Cys group lack the conserved Cys170 residue. The presence or absence of the second cysteine correlates with the amino acid sequence conservation in the neighbourhood of the first cysteine. The amino acid sequences surrounding Cys47 in 2-Cys Prxs are FVCP, whereas those in 1-Cys group are PVCT[25]. In a search of the GenBank database, the amino acid sequences of 15 representative members of 1-Cys, 2-Cys and PrxV proteins, excluding their predicated mitochondrial presequences were aligned (Figure 3 *a*). The amino acid residues surrounding Cys47 in both AmphiPrxV and its homologous proteins in other animals were PGCS, which was neither similar to FVCP nor to PVCT. Apart from the two conserved cysteines corresponding to Cys47 (AmphiPrxV Cys78) and Cys170 (AmphiPrxV Cys180) in yeast Prxs, AmphiPrxV also contained a third cysteine, Cys103 (Figure 3 *a*). Although some members of 1-Cys or 2-Cys Prxs often possess more cysteines in addition to the corresponding Cys47 and Cys170, neither the additional cysteines nor the amino acid sequences surrounding the additional cysteine are conserved among the members of 1-Cys or 2-Cys Prxs. In contrast, both the third cysteine

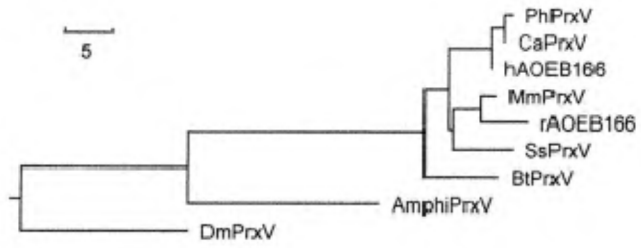

Figure 4. Phylogenetic relationship deduced from comparison of PrxV proteins without their predicated mitochondrial presequences from different animals. The phylogenetic tree was constructed using the MegAlign program (DNASTAR) by the CLUSTAL method, and the branch length represents the evolutionary distance. See Table 1 for sequence references.

and surrounding sequences were conserved in AmphiPrxV as well as the members of PrxV identified so far. The phylogenetic tree constructed from the complete sequences of the representative members of PrxV, 1-Cys and 2-Cys Prxs also showed that all 39 Prxs compared were classified into three groups, and members of PrxV including AmphiPrxV formed a distinct group in the phylogenetic tree (Figure 3 *b*). It appears that PrxV, including AmphiPrxV represents a new group of the Prx family. PrxV members have at present been identified only in seven mammals (including two possible isoforms in human), *Drosophila* and amphioxus.

Most Prxs exist as homodimers or heterodimers linked by disulfide bonds[26]. Site-directed mutagenesis of the conserved cysteines (Cys47 and Cys170) in yeast demonstrated that both the cysteines were essential for catalytic activity and/or formation of dimmers of 1-Cys and 2-Cys[27,28]. It is likely that the third conserved cysteine in PrxV also plays a role in the structure or function.

The phylogenetic tree constructed from the amino acid sequences of AmphiPrxV and its eight known counterparts without their predicated mitochondrial presequence from different animals indicated that AmphiPrxV was intermediate between the mammalian and fruit fly Prx proteins (Figure 4). This well reflects the established phylogeny of the chosen organisms, and agrees with the notion that the amphioxus is the basal lineage of chordates.

In conclusion, the amphioxus cDNA encodes the PrxV protein subtype with mitochondrial and peroxisomal sorting signals, and PrxV proteins form a new group of Prx family, since its members possess a third conserved Cys and specific amino acid sequences PGCS surrounding the first cystine, and are clustered with neither 1-Cys group nor 2-Cys group in the phylogenetic tree.

1. Chae, H. Z., Robison, K., Poole, L. B., Church, G., Storz, G. and Rhee, S. G., *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 7017–7021.
2. Lee, T. H., Yu, S. L., Kim, S. U., Lee, K. K., Rhee, S. G. and Yu, D. Y., *Gene*, 1999, **239**, 243–250.
3. Knoops, B. *et al.*, *J. Biol. Chem.*, 1999, **274**, 30451–30458.
4. Sarafian, T. A., Huang, C., Kim, A., de Vellis, J. and Shau, H., *Mol. Chem. Neuropathol.*, 1998, **34**, 39–51.
5. Sarafian, T. A. *et al.*, *J. Neurosci. Res.*, 1999, **56**, 206–212.
6. Araki, M., Nanri, H., Ejima, K., Murasato, Y., Fujiwwara, T., Nakashima, Y. and Ikeda, M., *J. Biol. Chem.*, 1999, **274**, 2271–2278.
7. Kang, S. W., Chae, H. Z., Seo, M. S., Kim, K., Baines, I. C. and Rhee, S. G., *ibid*, 1998, **273**, 6297–6302.
8. Lee, I., Spector, D., Godon, C., Labarre, J. and Toledano, M. B., *ibid*, 1999, **274**, 4537–4544.
9. Kim, K., Kim, I. H., Lee, K. Y., Rhee, S. G. and Stadtman, E. R., *ibid*, 1988, **263**, 4704–4711.
10. Zhou, Y. *et al.*, *Biochem. Biophys. Res. Commun.*, 2000, **268**, 921–927.
11. Jin, D.-Y., Chae, H. Z., Rhee, S. G. and Jeang, K. T., *J. Biol. Chem.*, 1997, **272**, 30952–30961.
12. Matsumoto, A., Okado, A., Fujii, T., Egashira, M., Niikawa, N., and Taniguchi, N., *FEBS Lett.*, 1999, **443**, 246–250.
13. Verdoucq, L., Vignols, F., Jacouot, J. P., Chartier, Y. and Meyer, Y., *J. Biol. Chem.*, 1999, **274**, 19714–19722.
14. Yamashita, H. *et al.*, *ibid*, 1999, **274**, 29897–29904.
15. Stokes, M. D. and Holland, N. D., *Am. Sci.*, 1998, **86**, 552–560.
16. Zhang, S. C., Yuan, J. D. and Li, H. Y., *Chin. Bull. Life Sci.*, 2001, **13**, 214–218 (in Chinese with English abstr.).
17. Liu, Z., Zhang, S., Yuan, J., Sawant, M. S., Wei, J. and Xu, A., 2002, *Curr. Sci.*, **83**, 101–104.
18. Rehm, B. H., *Appl. Microbiol. Biotechnol.*, 2001, **57**, 579–592.
19. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J., *J. Mol. Biol.*, 1990, **215**, 403–410.
20. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, Z., Miller, W. and Lipman, D. J., *Nucleic Acids Res.*, 1997, **25**, 3389–3402.
21. Clewley, J. P., *Mol. Biotechnol.*, 1995, **3**, 221–224.
22. Burland, T. G., *Methods Mol. Biol.*, 2000, **132**, 71–91.
23. Watabe, S., Kohno, H., Kouyama, H., Hiroi, T., Yago, N. and Nakazawa, T., *J. Biochem.* (*Tokyo*), 1994, **115**, 648–654.
24. Lametschwandtner, G., Brocard, C., Fransen, M., Veldhoven, P. V., Berger, J. and Hartig, A., *J. Biol. Chem.*, 1998, **273**, 33635–33643.
25. McGonigle, S., Dalton, J. P. and James, E. R., *Parasitol. Today*, 1998, **14**, 139–145.
26. Kang, S. W., Baines, I. C. and Rhee, S. G., *J. Biol. Chem.*, 1998, **273**, 6303–6311.
27. Chae, H. Z., Uhm, T. B. and Rhee, S. G., *Proc. Natl. Acad. Sci. USA*, 1994, **47**, 7022–7026.
28. Yim, M. B., Chae, H. Z., Rhee, S. G., Chock, P. B. and Stadtman, E. R., *J. Biol. Chem.*, 1994, **269**, 1621–1626.

ACKNOWLEDGEMENTS. We thank Dr Jianwen Wei and Chunyan Mou, Life Science College of Zhongshan University for help in cDNA library construction, and our colleagues for sample collection. The work was supported by a grant from the Ministry of Science and Technology of China awarded to S.Z.

Received 23 May 2002; revised accepted 21 October 2002