Turning points in the evolution of peroxidase–catalase superfamily: molecular phylogeny of hybrid heme peroxidases

Marcel Zámocký · Bernhard Gasselhuber · Paul G. Furtmüller · Christian Obinger

Abstract Heme peroxidases and catalases are key enzymes of hydrogen peroxide metabolism and signaling. Here, the reconstruction of the molecular evolution of the peroxidase–catalase superfamily (annotated in pfam as PF00141) based on experimentally verified as well as numerous newly available genomic sequences is presented. The robust phylogenetic tree of this large enzyme superfamily was obtained from 490 full-length protein sequences. Besides already well-known families of heme b peroxidases arranged in three main structural classes, completely new (hybrid type) peroxidase families are described being located at the border of these classes as well as forming (so far missing) links between them. Hybrid-type A peroxidases represent a minor eukaryotic subfamily from Excavates, Stramenopiles and Rhizaria sharing enzymatic and structural features of ascorbate and cytochrome c peroxidases. Hybrid-type B peroxidases are shown to be spread exclusively among various fungi and evolved in parallel with peroxidases in land plants. In some ascomycetous hybrid-type B peroxidases, the peroxidase domain is fused to a carbohydrate binding (WSC) domain. Both here described hybrid-type peroxidase families represent important turning points in the complex evolution of the whole peroxidase–catalase superfamily. We present and discuss their phylogeny, sequence signatures and putative biological function.

Keywords Hybrid-type heme b peroxidase · Peroxidase–catalase superfamily · Molecular evolution · Multifunctional enzymes · WSC domain

Abbreviations
APx Ascorbate peroxidase
APx-CcP Hybrid-type A peroxidase
CcP Cytochrome c peroxidase
CDE Chlorite dismutase-dye decolorizing peroxidase-EfeB
EST Expressed sequence tags
HyBpox Hybrid-type B peroxidase
KatG Catalase–peroxidase
PDB Protein Data Bank
Pfam Database of protein families
WSC Cell-wall integrity and stress response component

Introduction

Heme peroxidases are ubiquitous oxidoreductases present in all kingdoms of life. They can be divided in three main structural superfamilies and two minor families (overview in [1]). The three main superfamilies are the (1) peroxidase–catalase superfamily, (2) the peroxidase–cyclooxygenase superfamily [2] and (3) the CDE superfamily [3]. The latter is comprised of so-called chlorite dismutases, dye-decolorizing peroxidases and EfeB (which are heme-binding/sensing proteins of unclear biological function). In PeroxiBase (http://peroxibase.toulouse.inra.fr).
fr [4]) currently over 9,580 heme peroxidases are annotated and 8,160 of them belong to the peroxidase–catalase superfamily.

Heme peroxidases catalyze the oxidation of one- or two-electron donors by hydrogen peroxide. There is a great variability regarding the nature of one electron donors ranging from metal cations, aliphatic or aromatic compounds to proteins (Reaction 1). Typical two-electron donors can be halides (like Cl\(^{-}\), Br\(^{-}\) or I\(^{-}\)) or thiocyanate (SCN\(^{-}\)) which are oxidized to the corresponding hypohalous acids (HOX) (Reaction 2). There are only few heme peroxidase (sub) families that are able to dismutate hydrogen peroxide in the absence of electron donors as typical (monofunctional) catalases (Reaction 3). However, the mechanism of hydrogen peroxide dismutation by these peroxidases is different from typical heme catalases and thus should be designated as (pseudo) catalatic [5].

\[
\begin{align*}
H_2O_2 + 2AH_2 & \rightarrow 2H_2O + 2HA^* \\
H_2O_2 + X^- + H^+ & \rightarrow H_2O + HOX \\
2H_2O_2 & \rightarrow 2H_2O + O_2
\end{align*}
\] (1) (2) (3)

This work focuses on the peroxidase–catalase superfamily (PF00141). Already in 1992, Welinder [6] for the first time recognized the phylogenetic and structural relationship between heme peroxidases from plants, fungi and bacteria. In the subsequent years, the denominations “plant-type peroxidase superfamily” or “superfamily of heme peroxidases from plants, fungi and bacteria” or “non-animal heme peroxidase superfamily” were used in literature. In the meanwhile numerous further representatives from the kingdoms of Archaea, Fungi, Protista and Plantae were detected. Importantly, even in some Metazoon genomes members of this superfamily were found. Thus, it is more appropriate to denominate the whole superfamily according to the typical reaction specificities of its members [1] and not according to the taxonomical origin of selected members (as usual in older literature, e.g., [7]. It will be outlined below, how bifunctional enzymes (i.e., catalase–peroxidases) form the basis of this superfamily and how during evolution the (pseudo-) catalatic activity was lost and monofunctional peroxidases evolved. The denomination peroxidase–catalase superfamily reflects this development. This superfamily represents a good example for the divergent evolution of a mid-size gene that acquired diverse functions in different genomes. The usage of manifold electron donors by the different organisms in Reactions 1 and 2 mirrors the physiological necessities. The particular peroxidase gene evolved either by vertical descent or by horizontal gene transfer (HGT) as observed in several clades within the evolutionary history of this superfamily [8].

All members of the peroxidase–catalase superfamily that were investigated at the protein level contain non-covalently bound heme \(b\) and have a histidine as proximal ligand as well as a catalytic distal histidine [9]. For all other residues in the heme cavity some variability is observed and described below. Three main classes are distinguished (Class I, II and III) of which many biochemical and physical data have been collected in the last decades (e.g., [10–12]). Only recently, a few of the evolutionary missing links between these classes started to be systematically investigated (e.g., [13]).

Class I of this superfamily (1,852 annotated members, classified according to [6] is apparently the most divergent one. It comprises sequences of various catalase–peroxidases, ascorbate peroxidases, cytochrome \(c\) peroxidases and also hybrid-type subfamilies discovered recently. Class II (609 annotated members) is constituted mainly by manganese, lignin and versatile peroxidases from lignin-degrading fungi [14], but also numerous novel heme peroxidase sequences from fungi, that are not able to decay wood, which were annotated recently [15]. Class III (5,701 annotated members) is formed by large multigenic families of plant secretory peroxidases [16] including the well-known horseradish peroxidase (HRP).

The intention of this paper is to highlight features of the missing links between these three main classes. Hybrid-type A peroxidases (also named ascorbate-cytochrome \(c\) peroxidases, APx-CcP) are shown to be positioned between the classical families of ascorbate and cytochrome \(c\) peroxidases, whereas hybrid-type B enzymes (originally also abbreviated as APx-CcPs) are quite different from hybrid-type A proteins. A distinct group of ascomycetous hybrid-type B enzymes are shown to be fusion proteins containing a N-terminal peroxidase domain and C-terminal WSC-carbohydrate binding domain(s). We discuss the evolution and typical sequence signatures for both hybrid peroxidases and speculate about their physiological function. It will be demonstrated that hybrid-type peroxidases represent real turning points in the robust evolution of three main structural classes of the peroxidase–catalase superfamily. Their future biochemical analysis will further contribute to the understanding of the evolution of structure–function relationships of these abundant oxidoreductases.

Materials and methods

Sequence data mining and multiple sequence alignment

All protein sequences used in this study were collected from PeroxiBase (http://peroxibase.toulouse.inra.fr [4]) where they were previously verified and annotated. Multiple sequence alignment of full-length protein sequences
was performed with Muscle program [17] implemented in the MEGA 5 package. Optimized parameters were: gap open −2.9 gap extend 0, hydrophobicity multiplier 2. Maximum of alignment iterations was set to 1,000. Used clustering method was UPGMB for the first two iterations, for other iterations Neighbor-Joining (NJ) and minimal diagonal length 28. Obtained alignment was inspected mainly for the presence of conserved catalytic residues on both distal and proximal sides of the heme prosthetic group and further refined in GeneDoc [18]. Those sequences that did not possess the conserved essential residues in the heme cavity [i.e., distal Trp(Phe, Tyr)-His-Arg(Lys) and proximal His-Asp-Trp(Phe)] were segregated from the main alignment. After inspection and refinements the alignment used for phylogeny contained 490 full-length sequences. Sequences of observed peroxidase pseudogenes were analyzed separately. For this purpose a second independent alignment focused mainly on the non-functional ascorbate peroxidases was prepared with Muscle using the same parameters as above.

Structural alignment

Structural neighbors within the peroxidase–catalase superfamily were identified with the Dali Database (http://ekhidna.biocenter.helsinki.fi/dali) where the PDB structural hits were sorted according to obtained Z-score [19]. Structural alignment was performed with ESPript program suite http://esprit.ibcp.fr/ESPript/ESPript [20] where the input was the previously refined Muscle-alignment file. As top secondary structure of this alignment the sequence of *Burkholderia* KatG with known 3D structure was selected (PDB code 1MWV, Table 1). Parameters for similarity calculations were Risler, global score of 0.7, and consensus over 50 % were displayed. Obtained structural alignment was edited in GeneDoc [18].

Identification of introns and exons in putative peroxidase genes

For identification of donor splice sites and acceptor splice sites in various peroxidase genes the program suite NetAspGene 1.0 of the CBS server was used (http://www.cbs.dtu.dk/services/NetAspGene/). GT-AG consensus sequence for the borders between exons and introns was present in most, but not all hybrid-type peroxidase genes.

Molecular phylogeny

Molecular phylogeny within the whole peroxidase–catalase superfamily was reconstructed using the MEGA package, version 5 [21]. Muscle-aligned protein sequences including sequences of all 22 proteins of this superfamily with known 3D structure were subjected to Neighbor-Joining (NJ) or Maximum-Likelihood (ML) methods. For NJ 1,000 bootstraps, Jones-Taylor-Thornton (JTT) model of distribution and γ parameter optimized to 0.88 was used. For ML 100 bootstraps, Whelan and Goldman (WAG) model of amino acid substitutions with 3 γ categories was applied. The branch swap filter was set to very strong and the number of threads was set to 1. Reconstructed rooted phylogenetic tree was depicted with the program FigTree (http://tree.bio.ed.ac.uk/) in a circular polar form with branches transformed as cladograms. The branching details for particular (sub) families were presented with the Tree Explorer program of the MEGA package in the rectangular form. Ancestral protein sequences were inferred from the ML-reconstructed tree using the ancestors option of the MEGA suite [21] and exported as current site ancestors and corresponding most probable sequences.
Prediction of signal sequences and glycosylation in peroxidases

Putative signal sequences for selected peroxidase sequences were analyzed using the predictive algorithm of the program SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/) [22]. The appropriate prediction database was chosen according to determined phylogenetic relationship. Intracellular sequences were further subjected to subcellular localization analysis using TargetP 1.1 from the same online suite [22]. The N- and O-glycosylation sites were predicted with NetNGlyc1.0 and NetO Glyc 4.0 servers, respectively [23].

Homology modeling

Homology modeling of putative peroxidases from various subfamilies was performed with I-Tasser [24]. Obtained structural models were rendered with PyMOL (http://www.pymol.org).

Results and discussion

The peroxidase–catalase superfamily: division in three main classes

We have reconstructed a robust phylogenetic tree of the whole peroxidase–catalase superfamily comprising up to 490 full-length protein sequences of members from all known subfamilies. The general presentation of this tree (Fig. 1) clearly distinguishes all three main structural classes already defined by Welinder in 1992 [6]. However, upon closer inspection the occurrence of many so far undescribed clades as well as missing links is obvious.

As expected, Class II and Class III (comprising only eukaryotic proteins) are evolutionary descendants of Class I, which contains both prokaryotic and eukaryotic representatives. The phylogenetic origin of Class I is (according to the ML method) positioned among predecessors of katG genes coding for catalase–peroxidases (KatGs) from planktonic aerobic and heterotrophic bacteria [25].
Catalase–peroxidases are homodimeric oxidoreductases having a two-domain monomeric structure with a functional N-terminal heme-containing domain and a C-terminal domain without prosthetic group. KatGs are bifunctional having both a (pseudo-) catalase activity [5] and a broad peroxidase activity [26, 27]. From the plant bacterial branch, a major KatG clade [28] and minor KatG branches evolved (Fig. 1). The major KatG clade includes sequences from all bacterial phyla but also fungal katG genes. Among fungal KatGs the majority is located intracellularly (either in the cytosol or in peroxisomes) [29]. A smaller second group contains signal sequences for secretion [8]. The further evolutionary development of the catalase–peroxidase superfamily led towards (so far) putative KatGs from eukaryotic photosynthetic Algae and heterotrophic Stramenopiles positioned still within the minor KatG branches.

The next important evolutionary branching beyond bifunctional KatGs proceeded first in direction of hybrid-type A1 peroxidases that formed the basis for the evolution of three important groups within Class I, namely minor hybrid-type clade A2, ascorbate peroxidases from Algae and green plants and cytochrome c peroxidases from Opisthokonts phyla Chaoanoflagellida and Fungi (Fig. 1). Importantly, very soon in the evolution of the catalase–peroxidase superfamily (but already in the early eukaryotic world) at the beginning of the evolution of hybrid-type A peroxidase a common ancestor (labeled as * in Fig. 1) of fungal Class II peroxidases (including manganese and lignin peroxidases) and recently discovered and rare metazoan ascorbate peroxidases as well as of Class III peroxidases segregated. Most probably at this stage the two-domain structure typical for ancestral and modern KatGs disappeared by losing the (heme-free) domain (Fig. 2).

Besides KatGs only very few representatives of Class I are composed of two domains (Fig. 2c). Inspection of sequence suggests that in these rare cases both domains most probably are able to bind the prosthetic group (see following discussion of hybrid-type A peroxidases). There are also some examples of the occurrence of peroxidase-like domains lacking both the distal and the proximal histidines and thus (most probably) the heme group (Fig. 2e).

More typically, most known ascorbate peroxidases (APxs) (Fig. 2b, d) and all known cytochrome c peroxidases (CcPs) (Fig. 2f) are single domain proteins having lost the ancestral C-terminal domain. This is also true for all currently known members of Class II and Class III peroxidases (Fig. 2g–j) that reveal a high overall sequence similarity with the N-terminal domain of the ancestral KatG. The high level of evolutionary conservation of the heme peroxidase domain is obvious by comparison of the available X-ray structures (Table 1) of selected members as depicted in Fig. 3. The typical conserved 12-α-helical bundle [30, 31] with low content of β-strands did not change significantly during evolution.

Peculiarities of hybrid-type A peroxidases

The investigation of hybrid-type A peroxidases (abbreviated as APx-CcP in PeroxiBase) started recently since it was found that they represent the missing link between ascorbate and cytochrome c peroxidases (Fig. 1). The presented detailed phylogenetic reconstruction (Fig. 4) additionally underlines that hybrid-type A peroxidases are also among first descendants of bifunctional catalase–peroxidases. Upon losing the (KatG-typical) C-terminal domain as well as the ability to dismutate hydrogen peroxide they became monofunctional peroxidases. It has been
demonstrated frequently that ancient enzymes were promiscuous and thus multifunctional before evolving more specialized catalytic functions later during evolution [32].

The multiple structural sequence alignment depicted in Fig. 5 suggests that the bifunctional activity of KatGs [5] has been lost stepwise. This can be exemplified by inspection of two representative hybrid-type A peroxidases, namely EgrAPx-CcP and LmAPx-CcP (Fig. 5). The (pseudo-) catalase activity of KatGs is based on the presence of a redox active cofactor, a post-translationally formed Trp-Tyr-Met adduct in close proximity to the heme group (BpKatG numbering: Trp111-Tyr238-Met264, Fig. 3a) [33]. Tyrosine 238, which is essential for the H2O2-degrading activity of KatG [34, 35], is located on the KatG-typical large loop LL1 that also contributes to the architecture of the substrate channel [5]. Both hybrid-type A peroxidases, EgrAPx-CcP and LmAPx-CcP, have lost this large loop including Tyr238 (Fig. 5b) as well as the C-terminal part of large loop LL2 (Fig. 5c). Thus, they are not able to form the covalent adduct [10]. In LmAPx-CcP Met264 is substituted by a leucine, whereas in all hybrid-type A peroxidases Trp111 is fully conserved together with the catalytic residues His112 and Arg108 (BpKatG numbering Fig. 5). The histidine–arginine pair is found in all (mono) functional peroxidases of the catalase–peroxidase superfamily and is important for the heterolytic cleavage of H2O2 in compound I formation [36]. The proximal heme architecture including the triad His279–Asp389–Trp330 (and the H-bonding network between these residues) of hybrid-type A peroxidases is still very similar to that of KatGs [37].

Hybrid-type A1 peroxidase from Leishmania major has been investigated in detail demonstrating that it can use both ascorbate (as monofunctional APx) and cytochrome c (as monofunctional CcP) as electron donors (Reaction 2) [38]. The elucidation of its X-ray structure [30] confirmed its intermediate position between ascorbate and cytochrome c peroxidases (see Fig. 3b and discussion below).

Hybrid-type A heme peroxidases form a small group within lower single cell eukaryotes (Fig. 1). Phylogenetically, they can be subdivided in the basal subfamily A1 from Trypanosomatids and subfamily A2 from Rhizaria and Excavates (Fig. 4). The latter evolved in parallel with monofunctional ascorbate peroxidases. Sequence analysis suggests that hybrid-type A peroxidases are mostly non-secretory oxidoreductases which is reminiscent of most

---

**Fig. 3** Three-dimensional structures of representatives from all classes of the peroxidase–catalase superfamily. **a**, **b** Class I peroxidases, **c** Class II and **d** Class III peroxidases. **a** Two-domain monomeric structure of KatG from Burkholderia pseudomallei (PDB accession code: 1MWV, **b** hybrid-type A1 peroxidase from Leishmania major (PDB code: 3RIV), **c** manganese peroxidase from Phanerochaete chrysosporium (PDB code: 1MNP) and **d** Armoracia rusticana (horseradish) peroxidase (PDB code: 1ATJ). Color code: N-terminus is depicted in blue, C-terminus in red. In addition, the heme b group and essential distal and proximal residues are depicted. The violet sphere in (e) represents a bound manganese cation. Abbreviations of peroxidase names are taken from PeroxiBase.
Fig. 4 Details of the phylogenetic tree (Fig. 1) focusing on the hybrid-type A peroxidase clades. Obtained bootstrap values are presented for the ML method. Abbreviations of peroxidase names and corresponding ID numbers are taken from PeroxiBase. Subcellular location of hybrid A peroxidases is indicated.
bacterial KatGs which are their predecessors. Only the circozoan *Bigelowiella natans* reveals two paralogs with one being presumably secreted to cope with environmental stress (highlighted in blue in Fig. 4). This marine alga of secondary endosymbiotic origin contains a battery of oxidoreductases all possessing signal/leader peptides [39].

Among hybrid-type A2 peroxidases a unique dimeric two-domain structure is found in the peroxidase of *Euglena gracilis* (EgrAPx-CcP1). At first sight, this resembles the two-domain structure of KatGs (Figs. 2a and 3a), but it has been demonstrated that both domains bind heme and exhibit peroxidase activity [40]. This two-domain peroxidase is localized in the cytosol of *Euglena gracilis* [40]. All other currently known hybrid-type A peroxidases do not contain a second heme-binding peroxidase domain.

### Large diversity within clades of ascorbate and cytochrome c peroxidases

As outlined above, ascorbate peroxidases are evolutionary descendents of hybrid-type A1 peroxidases (Fig. 4). The main physiological role of modern APxs is hydrogen peroxide reduction by concomitant oxidation of ascorbate to monodehydroascorbate according to Reaction 1. The early segregating APx clade containing only microalgal representatives is strictly intracellular as are their direct predecessors. In the descendant clades from green and red alga and higher plants most representatives are located either in the cytosol or in the plastids, but few members reveal also a peroxisomal location (e.g., AcAPx4 and ZmApx4) [41]. As monodehydroascorbate reductase is also partially localized in protistan and plant peroxisomes [42] this organelle could also possess an efficient ascorbate-recycling system.

---

**Fig. 5** Selected parts of structural multiple sequence alignment of 48 members of the peroxidase–catalase superfamily. This alignment demonstrates both high conservation of the active site residues as well as some variability. a Region including residues at the distal heme side, b region of the large loop, c region including residues at the proximal heme side. Secondary structural elements taken from the 3D structure of KatG from *Burkholderia pseudomallei* (BpKatG, PDB code 1MWV) are depicted (h helix, s strand, t turn). Essential residues involved in catalysis are labeled as *"* and those residues that were involved in catalysis but later during the evolution mutated as ." residues discussed in the text are labeled with arrows. Sequences with known 3D structures are in bold. Parameters for the alignment are described in the Sect. “Materials and methods”. Abbreviations of peroxidase names are taken from PeroxiBase.
The physiological role of ascorbate from cyanobacteria through algae towards higher plants has undergone an interesting stepwise evolution serving not only as a major cellular water-soluble antioxidant and cofactor for many enzymes but also as a growth regulator and signal transducer [43]. The significant higher concentration of ascorbate in plants compared to cyanobacteria [44] was directly connected with the amplification of genes encoding both APxs (cf. PeroxiBase) and dehydroascorbate reductases [42]. Ascorbate peroxidases have neither a reasonable \( \text{H}_2\text{O}_2 \) dismutating activity (like KatG) nor can they use cytochrome \( c \) as electron donor (as CcP). The substrate ascorbate binds at the \( \gamma \)-heme edge through hydrogen bonds to conserved basic residues (typically lysine and arginine) and a heme propionate [45]. The proximal tryptophan (of the His-Asp-Trp triad) is (in contrast to CcP) not redox active [46]. During evolution of APxs the distal tryptophan was changed to phenylalanine (as found in most Class II and Class III peroxidases).

The present analysis has also shown that several APx genes in Chlorophyta and higher plants lack most of the essential residues for heme binding and/or catalysis (Fig. 6a, b). This is in contrast to the hybrid-type A2 peroxidase from \textit{Euglena gracilis} (EgrAPx-CcP1) with its unique structure of two functional domains [40]. This leads us to the hypothesis that right at the level of segregation...

Fig. 5 continued
from bifunctional KatG5s to monofunctional peroxidases (labeled with * in Fig. 1) the two-domain version might have existed for some time before being separated and losing functionality (Fig. 6). The occurrence of separated (HEME-free) KatG-like C-terminal domains still present in some genomes (cf. PeroxiBase) supports this hypothesis.
Also monofunctional cytochrome c peroxidases are direct descendants of hybrid-type A peroxidases (Fig. 1). Phylogenetic analysis suggested the presence of two CcP subfamilies. At the beginning of CcP evolution choanoflagellidal enzymes are found (Figs. 1 and 4). As expected from their direct predecessors they are intracellular enzymes and some of them are predicted (with high probability) to be targeted to mitochondria. Also the newly discovered Stramenopiles CcP representatives like most of the fungal CcPs of subfamily II [47] could be targeted to mitochondria. By contrast, CcP subfamily I [47] was predicted as non-mitochondrial, probably cytosolic. As obvious from here presented phylogenetic reconstruction this younger subfamily could have evolved through the loss or modification of the ancestral leader sequence.

In both CcP subfamilies the architecture of the heme cavity seems to be well conserved (i.e., distal triad Trp-His-Arg and proximal triad His-Asp-Trp). Cytochrome c peroxidases can neither dismutate H₂O₂ nor use ascorbate as electron donor. In contrast to APx the proximal Trp (Trp191 in SceCcP) is the site of entry for both cytochrome c electrons when CcP follows Reaction 1. The physiological role(s) of CcP is still under discussion. The enzyme from S. cerevisiae seems to be a protective enzyme for aerobic metabolism and is localized in the intermembrane space of mitochondria where it consumes hydrogen peroxide generated in the respiratory electron transport chain [48]. The physiological role of cytosolic CcPs of subfamily I remains unclear.

Metazoan representatives of the peroxidase–catalase superfamily

Originally this superfamily was always designated as “non animal” or as “superfamily of plant, bacterial and fungal peroxidases” [6, 7]. Our analysis clearly demonstrates that a distinct clade of (putative) ascorbate peroxidases was segregated very early in evolution (Fig. 1). It contains only sequences from Unikonts/Metazoan lineages. Already described symbiotic Hydra viridis ascorbate peroxidase [49] belongs to this clade in addition to other sequences of metazoan origin. Most of them are predicted as secretible proteins (Table 2). Since all descendant clades including Class II, Class III and hybrid-type B peroxidases (Fig. 1) also contain signal sequences for secretion, this might reflect an evolutionary turning point with respect to subcellular targeting. In case of the marine sponges from the phylum Porifera secretible peroxidases could help to combat marine microbes that live in the same niche [50]. In sea lice (i.e., copepod parasites of fish) these peroxidases might be involved in coping with the immune defence reactions of the host [51]. Whether ascorbate is the preferred electron donor of peroxidases from this minor clade remains to be verified experimentally as well as the eventual physiological role of ascorbate in ancestral metazoan lineages.

Peculiarities of hybrid-type B peroxidases

Until now hybrid-type B peroxidases have been abbreviated as “APx-CcP” similar to their hybrid-type A counterparts. However, it is obvious from sequence alignment (Fig. 5c) and first available experimental data (unpublished) that they are not able to use cytochrome c as electron donor. Their phylogenetic position is very far from both clades of APxs and CcPs (Fig. 1). Moreover, the basic residues responsible for ascorbate binding in APx at the γ-edge of heme [12] have partially been lost on the distal, but mainly on the proximal side (Fig. 5a, c). At the same time they have acidic Mn²⁺-binding residues on the distal

### Table 2 Prediction of signal and leader peptides in various (sub) classes of the peroxidase–catalase superfamily performed with SignalP (http://www.cbs.dtu.dk/services/SignalP/) or TargetP (http://www.cbs.dtu.dk/services/TargetP/)

| Peroxidase       | PeroxiBase ID | Taxonomy       | Signal or leader sequence length AA, (probability) |
|------------------|---------------|----------------|----------------------------------------------------|
| AqAPx            | 11256         | Porifera       | S 19 (0.959)                                       |
| AbHyBpox1        | 7607          | Basidiomycete  | 21 (0.592)                                         |
| BnaAPx-CcP1      | 2510          | Cercozoa       | S 22 (0.964)                                       |
| BnaAPx-CcP2      | 11106         | Cercozoa       | cytosolic (0.916)                                  |
| CaroAPx1         | 7150          | Ecdysozoa      | S 17 (0.959)                                       |
| CgHyBpox1        | 5357          | Ascomycete     | S 21 (0.646)                                       |
| CgHyBpox2        | 5361          | Ascomycete     | S 22 (0.872)                                       |
| CtheHyBpox       | 10136         | Ascomycete     | S 20 (0.560)                                       |
| GprHyBpox1       | 11620         | Chytridiomycete| S 29 (0.652)                                       |
| MagHyBpox1       | 2621          | Ascomycete     | S 19 (0.683)                                       |
| MagHyBpox2       | 5356          | Ascomycete     | S 22 (0.836)                                       |
| NcHyBpox1        | 5358          | Ascomycete     | S 20 (0.589)                                       |
| TasHyBpox1       | 11638         | Basidiomycete  | S 20 (0.810)                                       |
| TcrAPx-CcP1      | 2308          | Kinetoplastida | Mt (0.881)                                         |

Abbreviations of used sequence names and ID correspond to PeroxiBase

S secreted, Mt mitochondrial
side (Fig. 5a) which are conserved in all Class II manganese peroxidases [52]. The proximal heme cavity residues are identical to those from Class III peroxidases (Fig. 5c).

An interesting variability in the peroxidase active site among hybrid-type B occurs in the distal catalytic triad (Fig. 5a). Instead of the usual, highly conserved motif “Arg-X-X-Phe/Trp-His” “Arg-X-X-Tyr-His” is found in several proteins, e.g., in basidiomycetous DsHyBpox, HanHyBpox, MperHyBpox, ScomHyBpox1 and ascomycetous CgHyBpox2. Moreover, “Lys-X-X-Tyr-His” is found in phytopathogenic Sclerotinia sclerotiorum hybrid-type B peroxidase. Whether the modification of distal Trp/Phe towards Tyr affects the catalytic properties has to be tested. Whether a distal lysine might replace the arginine during heterolytic cleavage of hydrogen peroxide has to be tested too [53].

The present phylogenetic reconstruction demonstrates that hybrid-type B peroxidases are a sister clade to Class III and APx-related peroxidases (Fig. 1). We suggest to use the extension HyBpox for all members of this subfamily in future. The evolution of Class III peroxidases (including the well-known horseradish peroxidase, HRP) was reconstructed recently [54]. The main conclusion was that the emergence of Class III peroxidases was connected with the appearance of land plants as there were no members of this otherwise abundant Class found among Chlorophyta (green alga) so far. However, the present robust phylogenetic reconstruction points to a more complex scenario (Fig. 1). There was a common ancestor of hybrid-type B, APx-related and Class III peroxidases. Hybrid-type B peroxidases diverged early from the common node with Class III and APx-related (putative) peroxidases. The latter were separated later in evolution from the very abundant Class III clade (so far more than 5,700 sequences of Class III peroxidases are deposited in PeroxiBase). Interestingly, APx-related enzymes, which appear also in Batrachochytrium dendrobatidis. It is closely related to another chytrid peroxidase from the amphibian pathogen Batrachochytrium dendrobatidis. It is closely related to Cryptogam Class III peroxidases and basal for all APx-related genes. The possibility of horizontal gene transfer of the hybrid-type B peroxidase genes into and from Chytridiomycota genomes is probable. However, so far only a few genes from this ancestral (and rather rare) phylum are available to answer this question satisfactorily.

The 6th clade of hybrid-type B peroxidases comprises dominantly ascomycetous representatives. Few basidiomycetous genes are present too. Most of the genes come from phytopathogenic Sordariomycetes and there is already evidence of their native expression from the EST database (Table 3) mainly in various mycelia grown in minimal media.

The last clade of the hybrid-type B peroxidase subfamily only contains ascomycetous proteins from Sordariomycetes; among them are also several variants from thermophilic fungi. Expression of peroxidases from this clade was already verified by several EST entries, most prominently in samples induced with oxidative stress (Table 3). The length of their overall open reading frame is at least twice as long as compared to hybrid-type B peroxidases described above due to the presence of a unique C-terminal fusion with WSC domain(s). The latter are classified as PF01822, i.e., a putative carbohydrate binding domain [58]. WSC domain contains up to eight conserved Cys residues that might form several disulfide bridges [59]. It is present only among eukaryotes, mainly in fungal exoglucanases but also among related proteins of higher metazoans. Genetic analysis revealed that WSC proteins are upstream regulators of the stress-activated PKC1-MAP kinase cascade [60]. Figure 8 depicts a multiple sequence alignment of WSC domains in fused hybrid-type B (HyB) peroxidases.
### Fig. 7 Details of phylogenetic tree (Fig. 1) of hybrid-type B peroxidases. Obtained bootstrap values are presented for ML method. Abbreviations of peroxidase names and corresponding ID numbers are taken from PeroxiBase. Those sequences where mRNA was detected (EST database) are labeled in bold.

| Basidiomycetes basal clade | Basidiomycetes clade | solely |
|----------------------------|----------------------|--------|
| **Ascomycetes mixed clade 1** | | Agaricomycetes |
| **Ascomycetes mixed clade 2** | | Agaricomycetes |

**Chytridiomycetes clade**

**Sordariomycetes**

| Basidiomycetes clade | Basidiomycetes clade | solely |
|----------------------|----------------------|--------|
| **Ascomycetes mixed clade 1** | | Agaricomycetes |
| **Ascomycetes mixed clade 2** | | Agaricomycetes |

**phytopathogenic clade of short variants (dominantly Ascomycetes)**

**Sordariomycetes**

| Basidiomycetes clade | Basidiomycetes clade | solely |
|----------------------|----------------------|--------|
| **Ascomycetes mixed clade 1** | | Agaricomycetes |
| **Ascomycetes mixed clade 2** | | Agaricomycetes |

**Sordariomycetes**

**Basidiomycetes clade**

**Eurotiomycetes**

**Ascomycetes clade with C-terminal WSC fusions**

**Sordariomycetes**

| Basidiomycetes clade | Basidiomycetes clade | solely |
|----------------------|----------------------|--------|
| **Ascomycetes mixed clade 1** | | Agaricomycetes |
| **Ascomycetes mixed clade 2** | | Agaricomycetes |
together with proteins that do not contain a peroxidase domain. A high level of conservation mainly in the region around the eight cysteines is evident. The physiological function of hybrid-type B peroxidases remains completely unknown. There are currently several EST sequence entries found in non-induced fungi, in fungi induced with oxidative stress and also in fruiting bodies of ascomycetes and basidiomycetes. Sequence analysis suggests that all hybrid-type B peroxidases are secreted (Table 2) carrying a signal peptide of constant length. So far there is no X-ray structure available but in silico analysis predicts intensive N- and O-glycosylation and the occurrence of several disulfide bridges. Structural prediction with I-TASSER suggests that the closest structural homolog of ascomycetous hybrid-type B peroxidase domains is the Class III peroxidase from *Hordeum vulgare* (Table 1). It is important to note that the sequence similarity is rather low.

### Conclusion

The presented phylogenetic reconstruction has demonstrated the important role of hybrid-type heme peroxidases as turning points in the evolution of the complex and abundant peroxidase domain. Parameters for the alignment are described in the Sect. “Materials and methods”. Abbreviations of peroxidase names are taken from PeroxiBase.

### Table 3  Expressed sequence tags with fragments of hybrid-type B peroxidase genes

| EST #     | Peroxidase    | Source                  | Length [bp] | Remark                      |
|-----------|---------------|-------------------------|-------------|-----------------------------|
| FQ924977  | BfuHyBpox2    | *B. fuckeliana*         | 665         | Xylan as carbon source      |
| JZS40908  | CthHyBpox     | *C. thermophilum*       | 397         | Mycelium induced with H₂O₂  |
| CCOZ4053  | HanHyBpox2    | *H. annosus*            | 789         | Culture from liquid Hagem medium |
| CD037344  | MagHyBpox1    | *M. grisea*             | 642         | Subtracted mycelial library |
| DC977217  | MagHyBpox2    | *M. grisea*             | 366         | Growth in nitrogen limiting conditions |
| G1176P141RN11.T0 | NcHyBpox1    | *N. crassa*             | 732         | Oxidative stress 1 h        |
| G1176P11L21.T0 | NcHyBpox1    | *N. crassa*             | 754         | Oxidative stress 1 h        |
| G1176P141FN11.T0 | NcHyBpox1    | *N. crassa*             | 730         | Oxidative stress 1 h        |
| BG278614.1| NcHyBpox1    | *N. crassa*             | 488         | Sexual cDNA library         |
| FG363154  | TviHyBpox1    | *T. virens*             | 814         | Non-induced mycelium        |

### Table 3  Expressed sequence tags with fragments of hybrid-type B peroxidase genes

The WSC domain of large hybrid-type B peroxidases is compared with that of proteins that do not possess a peroxidase domain. A high level of conservation mainly in the region around the eight cysteines is evident. The physiological function of hybrid-type B peroxidases remains completely unknown. There are currently several EST sequence entries found in non-induced fungi, in fungi induced with oxidative stress and also in fruiting bodies of ascomycetes and basidiomycetes. Sequence analysis suggests that all hybrid-type B peroxidases are secreted (Table 2) carrying a signal peptide of constant length. So far there is no X-ray structure available but in silico analysis predicts intensive N- and O-glycosylation and the occurrence of several disulfide bridges. Structural prediction with I-TASSER suggests that the closest structural homolog of ascomycetous hybrid-type B peroxidase domains is the Class III peroxidase from *Hordeum vulgare* (Table 1). It is important to note that the sequence similarity is rather low.

### Conclusion

The presented phylogenetic reconstruction has demonstrated the important role of hybrid-type heme peroxidases as turning points in the evolution of the complex and abundant peroxidase domain. Parameters for the alignment are described in the Sect. “Materials and methods”. Abbreviations of peroxidase names are taken from PeroxiBase.
peroxidase–catalase superfamily. This computational work provides an excellent basis for detailed spectroscopic and kinetic studies of these missing links and will help to understand the gradual development of structure–function relationships, substrate utilization and, in consequence, the physiological function. Knowledge about structure and catalytic properties will also answer the question whether these proteins are of interest for enzyme engineering and design.

Acknowledgments Our research was supported by the Austrian Science Fund (FWF projects P23855 B-11 and W1224, BioToP—Bio-molecular Technology of Proteins) as well as by the Slovak Grant Agency VEGA (Grant 2/0021/14).

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

1. Zamocky M, Obinger C (2010) Molecular phylogeny of heme peroxidases. In: Torres E, Ayala M (eds) Biocatalysis based on heme peroxidases. Springer, Heidelberg, pp 7–35. ISBN 978-3-642-12626-0
2. Zamocky M, Jakopitsch C, Furtmüller PG, Dunand C, Obinger C (2008) The peroxidase-cyclooxygenase superfamily: reconstructed evolution of critical enzymes of the innate immune system. Proteins 72:589–605
3. Gobirsch B, Kurker RC, Streit BR, Wilmot CM, DuBois JL (2011) Chlorite dismutases, DyoPs, and EfEβ: 3 microbial heme enzyme families comprise the CDE structural superfamily. J Mol Biol 408:379–398
4. Passardi F, Theiler G, Zamocky M, Cosio C, Rouhier N, Teixeira F, Margis-Pinheiro M, Ioannidis V, Penel C, Falquet L, Dunand C (2007) PeroxiBase: the peroxidase database. Physchimistry 68:1605–1611
5. Vlasits J, Jakopitsch C, Bernroitter M, Zámócky M, Furtmüller PG, Obinger C (2010) Mechanisms of catalase activity of heme peroxidases. Arch Biochem Biophys 500:74–81
6. Welinder KG (1992) Superfamily of plant, fungal, and bacterial peroxidases. Curr Opin Struct Biol 3:388–393
7. Passardi F, Bakalovice N, TeixeiraBK, Margis-Pinheiro M, Penel C, Dunand C (2007) Prenkaryotic origins of the non-animal peroxidase superfamily and organelle-mediated transmission to eukaryotes. Genomics 89:567–579
8. Zamocky M, Furtmüller PG, Obinger C (2010) Evolution of structure and function of Class I peroxidases. Arch Biochem Biophys 500:45–57
9. Dunford HB (2010) Peroxidases and catalases. Biochemistry, biophysics, biotechnology, and physiology, 2nd edn. John Wiley, Chichester
10. Jakopitsch C, Kolarich D, Petutschnig G, Furtmüller PG, Obinger C (2003) Distal side tryptophan, tyrosine and methionine in catalase-peroxidases are covalently linked in solution. FEBS Lett 552:135–140
11. Barrows T, Bhaskar B, Poulos TL (2004) Electrostatic control of the tryptophan radical in cytochrome c peroxidase. Biochemistry 43:8826–8834
12. Gumiero A, Murphy EJ, Metcalfe CL, Moody PCE, Raven EL (2010) An analysis of substrate binding interactions in the heme peroxidase enzymes: a structural perspective. Arch Biochem Biophys 500:13–20
13. Adak S, Pal S (2013) Ascorbate peroxidase acts as a novel deter-miner of redox homeostasis in Leishmania. Antiox Redox Signal 19:746–754
14. Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Hen-rissat B, Martinez AT, Utillar R, Spatafora JW, Yadav JS, Aerts A, Benoit I, Boyd A, Carlson A, Copeland A, Coutinho PM et al (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715–1719
15. Bödeker ITM, Nygren CMR, Taylor AF, Olson Å, Lindahl BD (2009) ClassII peroxidase-encoding genes are present in a phylogenetically wide range of ectomycorrhizal fungi. ISME J 3:1387–1395
16. Passardi F, Longet D, Penel C, Dunand C (2004) The class III peroxidase multigene family in rice and its evolution in land plants. Phytochemistry 65:1879–1893
17. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl Acids Res 32:1792–1797
18. Nicholas KB, Nicholas HB (1997) Genedoc: a tool for editing and annotating multiple sequence alignments. Distributed by the authors
19. Holm L, Rosenström P (2010) Dali server: conservation mapping in 3D. Nucl Acids Res 38:W545–W549
20. Gouet P, Robert X, Courcelle E (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucl Acids Res 31:3320–3323
21. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. Mol Biol Evol 28:2731–2739
22. Emanuelsen O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP, and related tools. Nat Protoc 2:953–971
23. Steentoft C, Vakhruhev SY, Joshi HJ, Kong Y, Vester-Chris-tensen MB, Schjoldager KT, Labvenen K, Dabelsteen S, Pedersen NB, Marcos-Silva L, Gupta R, Benetti EP, Mandel U, Brunak S, Kandall HH, Levery SB, Clausen H (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J 32:1478–1488
24. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified plat-form for automated protein structure and function prediction. Nat Protoc 5:725–738
25. Glockner FO, Kube M, Bauer M, Teeling H, Lombardot T, Ludwig W, Gade D, Beck A, Borzym K, Heitmann K, Rabus R, Schlesner H, Amann R, Reinahardt R (2003) Complete genome sequence of the marine planctomyctoe Pirellula sp. strain 1. Proc Natl Acad Sci USA 100:8298–8303
26. Jakopitsch C, Regelsberger G, Furtmüller PG, Rüker F, Peschek GA, Obinger C (2001) Catalase-peroxidase from Synechocystis is capable of chlorination and bromination reactions. Biochem Biophys Res Commun 287:682–687
27. Regelsberger G, Jakopitsch C, Rüker F, Peschek GA, Obinger C (1999) Spectral and kinetic studies on the oxidation of monosubstituted phenols and anilines by recombinant Synechocystis catalase-peroxidase complex. Biochemistry 38:10480–10488
28. Zamocky M, Gasselhuber B, Furtmüller PG, Obinger C (2012) Molecular evolution of hydrogen peroxide degrading enzymes. Arch Biochem Biophys 525:131–144
29. Zamocky M, Sekot G, Backova M, Godocikova J, Schäffer C, Farkasovsky M, Obinger C, Polek B (2013) Intracellular targeting of ascomycetous catalase-peroxidases (KatG1s). Arch Microbiol 195:393–402
30. Jasion VS, Polanco JA, Meharenna YL, Li H, Poulos TL (2011) Crystal structure of Leishmania major peroxidase and characterization of Compound I Tryptophan radical. J Biol Chem 286:24608–24615

 Springer
31. Patterson WR, Poulos TL (1995) Crystal structure of recombinant pea cytosolic ascorbate peroxidase. Biochemistry 34:4331–4341
32. Nam H, Lewis NE, Lerman JA, Lee DH, Chang RL, Kim D, Palsson BO (2012) Network context and selection in the evolution to enzyme specificity. Science 337:1101–1104
33. Jakopitsch C, Wiseman B, Vlasits J, Loewen P, Obinger C (2007) Redox intermediates in the catalase cycle of catalase-peroxidase from Synechocystis, Mycobacterium tuberculosis and Burkholderia pseudomallei. Biochemistry 46:1183–1193
34. Jakopitsch C, Auer M, Ivancich A, Rüker F, Furtmüller PG, Peschek GA, Obinger C (2002) Engineering the proximal heme cavity of catalase-peroxidase. J Inorg Biochem 91:78–86
35. Adak S, Data AK (2005) Leishmania major encodes an unusual peroxidase that is a close homologue of plant ascorbate peroxidase: a novel role of transmembrane domain. Biochem J 390:465–474
36. Hopkins JF, Spencer DF, Laboissiere S, Neilson JAD, Eveleigh RJM, Dunford DG, Gray MW, Archibald JM (2012) Proteomics reveals plastid- and periplastid-targeted proteins in the chlorarachniophyte alga Bigelowiella natans. Genome Biol Evol 4:1391–1406
37. Ishikawa T, Tajima N, Nishikawa H, Gao J, Rapolus M, Shibata H, Sawa Y, Shigeoka S (2010) Euglena gracilis ascorbate peroxidase forms an intramolecular dimeric structure: its unique molecular characterization. Biochem J 426:125–134
38. Mhadmi A, Noctor G, Baker A (2012) Plant catalases: peroxisomal, redox guardians. Arch Biochem Biophys 525:181–194
39. Gaffney DR (2013) The role of L-ascorbic acid recycling in responding to environmental stress and in promoting plant growth. J Exp Bot 64:33–53
40. Berrothiner M, Zamocky M, Furtmüller PG, Peschek GA, Obinger C (2009) Occurrence, phylogeny, structure and function of catalases and peroxidases in cyanobacteria. J Exp Bot 60:423–440
41. Sharp KH, Mewies M, Moody PCE, Raven EL (2003) Crystal structure of the ascorbate peroxidase-ascorbate structure. Nat Struct Mol Biol 10:303–307
42. Patterson WR, Poulos TL (2010) Thirty years of heme peroxidase structural biology. Arch Biochem Biophys 500:3–12
43. Zamocky M, Dunand C (2006) Divergent evolutionary lines of fungal cytochrome c peroxidases belonging to the superfamily of bacterial, fungal and plant heme peroxidases. FEBS Lett 580:6655–6664
44. Erman JE, Vitello LB (2002) Yeast cytochrome c peroxidase: mechanistic studies via protein engineering. Biochim Biophys Acta 1597:193–200
45. Habelthu M, Bosch TC (2005) Symbiotic Hydra express a plant-like peroxidase gene during oogenesis. J Exp Biol 208:2157–2165
46. Nigro AP, Goodin DB (2010) Reaction of N-hydroxyguanidine with the ferrous-oxy state of a heme peroxidase cavity mutant: a model for the reactions of nitric oxid synthase. Arch Biochem Biophys 500:66–73
47. Mathé C, Barre A, Jourda C, Dunand C (2010) Evolution and expression of class III peroxidases. Arch Biochem Biophys 500:58–65
48. Dunand C, Mathé C, Lazzarotto F, Margis R, Margis-Pinheiro M (2011) Ascorbate peroxidase-related is not a duplicate gene. Plant Signal Behav 6:1908–1913
49. Lazzarotto F, Teixeira FK, Siqueira JAG, Margis R, Margis-Pinheiro M (2011) Ascorbate peroxidase-related is not a duplicable gene. Plant Signal Behav 6:1908–1913
50. Wolf DG, Falk R, Hackam H, Theelen B, Boekhout T, Scorzetti G, Shapiro M, Block C, Salkin IF, Polacheck I (2001) Multidrug-resistant Trichosporon asahii infection of nongranulocytic patients in three intensive care units. J Clin Microbiol 39:4420–4425
51. Ponting CP, Hoffman K, Bork P (1999) A latrophilin/CL-1-like GPS domain in polycystin-1. Curr Biol 9:R585–R588
52. Verna J, Lodder A, Lee K, Vagts A, Ballester R (1997) A family of genes required for maintenance of cell wall integrity and for the stress response in Saccharomyces cerevisiae. Proc Natl Acad Sci USA 94:13804–13809
53. Lodder AL, Lee TK, Ballester R (1999) Characterization of the Wsc1 protein, a putative receptor in the stress response of Saccharomyces cerevisiae. Genetics 152:1487–1499