Archives of Biochemistry and Biophysics xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Independent evolution of four heme peroxidase superfamilies

Marcel Zámocký [a,b,*], Stefan Hofbauer [a,c], Irene Schaffner [a], Bernhard Gasselhuber [a], Andrea Nicolussi [a], Monika Soudi [a], Katharina F. Pirker [a], Paul G. Furtmüller [a], Christian Obinger [a]

[a] Department of Chemistry, Division of Biochemistry, VIBT – Vienna Institute of BioTechnology, University of Natural Resources and Life Sciences, Muthgasse 18, A-1190 Vienna, Austria
[b] Institute of Molecular Biology, Slovak Academy of Sciences, Dúbravská cesta 21, SK-84551 Bratislava, Slovakia
[c] Department for Structural and Computational Biology, Max F. Perutz Laboratories, University of Vienna, A-1030 Vienna, Austria

ARTICLE INFO

Article history:
Received 26 November 2014
and in revised form 23 December 2014
Available online xxxx

Keywords:
Heme peroxidase
Peroxidase–catalase superfamily
Peroxidase–cyclooxygenase superfamily
Peroxidase–chlorite dismutase superfamily
Peroxidase–peroxygenase superfamily

ABSTRACT

Four heme peroxidase superfamilies (peroxidase–catalase, peroxidase–cyclooxygenase, peroxidase–chlorite dismutase and peroxidase–peroxygenase superfamily) arose independently during evolution, which differ in overall fold, active site architecture and enzymatic activities. The redox cofactor is heme *b* or posttranslationally modified heme that is ligated by either histidine or cysteine. Heme peroxidases are found in all kingdoms of life and typically catalyze the one- and two-electron oxidation of a myriad of organic and inorganic substrates. In addition to this *peroxidatic* activity distinct (sub)families show pronounced catalase, cyclooxygenase, chlorite dismutase or peroxygenase activities. Here we describe the phylogeny of these four superfamilies and present the most important sequence signatures and active site architectures. The classification of families is described as well as important turning points in evolution. We show that at least three heme peroxidase superfamilies have ancient prokaryotic roots with several alternative ways of divergent evolution. In later evolutionary steps, they almost always produced highly evolved and specialized clades of peroxidases in eukaryotic kingdoms with a significant portion of such genes involved in coding various fusion proteins with novel physiological functions.

Introduction

Heme peroxidases use heme *b* or posttranslationally modified heme as redox cofactor to catalyze the hydrogen peroxide-mediated one- and two-electron oxidation of a myriad of molecules including aromatic molecules (e.g., coniferyl alcohol or tyrosine), cations (e.g., Mn^{2+}), anions (e.g., ascorbate or halides) or even proteins (e.g., cytochrome *c*). During turnover H_2O_2 is reduced to water and one-electron donors (AH_2) are oxidized to the respective radicals ($^{\cdot}AH$) (Reaction 1) whereas two-electron donors like halides (X^-) are oxidized to the corresponding hypohalous acids (HOX) (Reaction 2). Besides these *peroxidatic* reactions very few heme peroxidases also show a reasonable *catalatic* reaction (Reaction 3) and use a second hydrogen peroxide molecule as two-electron donor thereby releasing dioxygen. One additional activity catalyzed by a special group of heme peroxidases is the peroxygenation reaction, i.e., the (selective) introduction of peroxide-derived oxygen functionalities into organic molecules (Reaction 4).

$$H_2O_2 + 2AH_2 \rightarrow H_2O + 2{\cdot}AH \quad \text{(Reaction 1)}$$

$$H_2O_2 + X^- + H^+ \rightarrow H_2O + HOX \quad \text{(Reaction 2)}$$

$$H_2O_2 + H_2O_2 \rightarrow 2H_2O + O_2 \quad \text{(Reaction 3)}$$

$$H_2O_2 + RH \rightarrow H_2O + ROH \quad \text{(Reaction 4)}$$

In the last decade – due to the application of powerful sequencing techniques – an ever increasing amount of protein sequences (including numerous heme peroxidases) were automatically assigned to related protein families due to typical conserved motifs [1]. However, critical analysis shows that these annotations often need corrections based on the knowledge of the relationship between sequence, structure and function of the respective protein

Abbreviations: APx, ascorbate peroxidase; CcP, cytochrome *c* peroxidase; Cld, chlorite dismutase; DyP, dye-decolorizing peroxidase; EPO, eosinophil peroxidase; HGT, horizontal gene transfer; HRP, horseradish peroxidase; KatG, catalase–peroxidase; LDS, linoleate diol synthase; LPO, lactoperoxidase; LspPOX, *Lyngbya* peroxidase; ML, maximum likelihood method; MnP, manganese peroxidase; MPO, myeloperoxidase; PERCAL, calcium binding motif; PDB, Protein Data Bank; Pfam, protein families database; SCOP, structural classification of proteins; TPO, thyroid peroxidase; WSC, cell-wall integrity & stress response component.

* Corresponding author at: Department of Chemistry, Division of Biochemistry, VIBT – Vienna Institute of BioTechnology, University of Natural Resources and Life Sciences, Muthgasse 18, A-1190 Vienna, Austria. Fax: +43 1 47654 6059.
E-mail address: marcel.zamocky@boku.ac.at (M. Zámocký).

http://dx.doi.org/10.1016/j.abb.2014.12.025
0003-9861/

family. An important aspect of sorting structural superfamilies is to classify the numerous sequences according to their phylogenetic relationships [2,3] and to identify turning points and evolutionary hybrids. Moreover, interesting newly discovered evolutionary clades frequently with yet unknown function and mostly marginal but important sequence variations can be selected for further structural and kinetic studies.

This review demonstrates the independent evolution of four heme peroxidase superfamilies, namely of the (1) peroxidase–catalase superfamily, (2) peroxidase–cyclooxygenase superfamily, (3) peroxidase–chlorite dismutase superfamily and the (4) peroxidase–peroxygenase superfamily. We present an update of their reconstructed phylogeny and introduce their representative sequence signatures and essential amino acids in the heme cavity. From Fig. 1 it is obvious that each superfamily possesses a peculiar fold of the heme peroxidase domain that evolved independently with respect to other peroxidase superfamilies. It is further demonstrated that these ubiquitous oxidoreductases are present in all kingdoms of life. This strongly underlines the necessity to use the nomenclature suggested before (e.g., in [4] and not to talk about bacterial or plant or animal heme peroxidases as is still widespread in the literature and some databases). The denomination of the peroxidase superfamilies should reflect the characteristic enzymatic activities rather than their origin. The criteria for the definition of a peroxidase superfamily were presented already in [4] and here they are systematically applied on all four above listed heme peroxidase superfamilies that are further divided in particular families and subfamilies.

Peroxidase–catalase superfamily

The peroxidase–catalase superfamily previously known as the "superfamily of bacterial, fungal and plant heme peroxidases" [5] is currently the most abundant peroxidase superfamily present in various gene and protein databases. Representatives of this superfamily were detected not only in the domains of Bacteria, Archaea, eukaryotic kingdoms of Fungi and Plantae but besides numerous Protozoa and Chromista species also within Metazoan kingdom [6]. In Pfam database (http://pfam.xfam.org) this universally distributed superfamily is annotated as PF00141.

Here we have selected 500 representative full-length sequences that were used for an updated phylogenetic reconstruction based on the maximum likelihood (ML) method. Fig. 2 shows the occurrence of three well separated families already defined by K. Welinder in 1992 as distinct structural classes Class I, II and III [5]. We suggest to use for them the term *families* in analogy with all other heme peroxidase superfamilies (see below). Novel clades representing hybrid enzymes between the previously defined classes are well distinguishable in this reconstruction. Fig. 3 shows the typical sequence signatures of this superfamily that includes the distal amino acid triad Arg-Trp/Phe-His within the sequence –X-**R**-XX-**W**/**F**-**H**-X- and the proximal triad His-Trp/Phe-Asp (these amino acids are not neighbored in the primary sequence). The typical overall globular fold of representatives of this superfamily consists of twelve α-helices and was already acquired from the beginning. It was only slightly modified in the later steps of ongoing divergent evolution. Generally, the level of structural conservation is higher than the conservation of amino acid sequences [7].

Family I was previously defined to contain intracellular bacterial catalase–peroxidases [8,9], cytosolic, chloroplastic and peroxisomal ascorbate peroxidases [10] and mitochondrial (membrane bound) cytochrome *c* peroxidases [11,12]. Catalase–peroxidases are the only representatives of this superfamily with a high *catalatic* activity according to Reaction 3 [13] besides a peroxidatic activity. Ascorbate and cytochrome *c* peroxidases are typical monofunctional peroxidases (i.e., cannot catalyze Reaction 3) that follow Reaction 1 with either ascorbate or cytochrome *c* as one electron donor (AH_2). The main function of Family I peroxidases seems to be scavenging of excess H_2O_2.

Family I is the most divergent one among the three structural classes. It contains eukaryotic intracellular [14,15] or even peroxisomal [16] as well as extracellular [17] catalase–peroxidases and also putative cytochrome *c* peroxidases from Choanoflagellida predecessors of fungi and animals [6]. Moreover, hybrid-type A heme peroxidases were shown to belong to Family I [7]. Their biochemical properties are between those of monofunctional APxs and CcPs [18,19], thus they represent an important turning point from ancient bifunctional catalase–peroxidases towards monofunctional specialization within this large superfamily (Fig. 2).

A peculiar minor descendant clade containing yet putative secretable ascorbate peroxidases from Unikonts/Metazoan lineages represents another evolutionary turning point (Fig. 2). A majority of further evolutionary descendants of this clade were verified as extracellular secreted proteins [6]. Among Family II sequences well known and intensively investigated lignin, manganese, versatile and generic peroxidases are found. Their main function seems to be the degradation of lignin-containing soil debris [20,21]. Recently, several novel clades with putative Family II members were detected also in various ascomycetous fungi unable to decay wood (Fig. 2). Thus, their physiological function and substrate specificity remain to be elucidated. Principally, they are able to catalyze Reactions 1 and 2.

The highest number of members are found in Family III that is comprised of numerous plant secretory peroxidases like horseradish peroxidase(s) (HRP) [e.g., 22] but also by a parallel abundant clade divided into "APx-related" yet putative algal and plant peroxidases [23] and hybrid-type B peroxidases that are present solely in the kingdom of fungi [6,7]. Sequence analysis has shown that most of these hybrid-type B peroxidases contain an additional C-terminal WSC sugar binding domain of unknown function. Family III peroxidases mainly follow Reaction 1. They are generally secreted into the cell wall or the apoplastic compartment and the vacuole and play a role in the oxidation of lignin precursors, auxin and secondary metabolites [22].

Peroxidase–cyclooxygenase superfamily

The peroxidase–cyclooxygenase superfamily has Pfam accession PF03098 and its members are widely distributed among all domains of life [4]. Therefore, its (old) denomination as "animal heme-dependent peroxidases" is misleading but still present in some public databases. We have selected 400 representative sequences for an updated ML phylogenetic reconstruction (Fig. 4). The seven main families are well conserved in this update but some of them are presented as more abundant if compared with previous reconstructed phylogenies [24–26] that focused on particular (sub) families.

In contrast to the peroxidase–catalase superfamily, in many cases members of the peroxidase–cyclooxygenase superfamily are multidomain proteins with one heme peroxidase domain of predominantly α-helical fold with a central heme-containing core of five α-helices. Moreover, this superfamily is unique in having the prosthetic heme group posttranslationally modified [26,27]. The heme is covalently bound to the protein via two ester linkages formed by conserved Asp and Glu residues (Fig. 5A). In one representative (i.e., myeloperoxidase) a third heme to protein linkage is formed [28]. As a consequence of these modifications the heme is distorted and these peroxidases exhibit unique spectral and redox properties [29]. All representatives catalyze both Reactions 1 and 2 but with most members studied so far halide oxidation seems to be the dominating physiological enzymatic activity.