Published in final edited form as:

Biochem Soc Trans. 2009 August ; 37(Pt 4): 772–777. doi:10.1042/BST0370772.

Two distinct groups of fungal catalase/oxidases

Marcel Zámocký^{*,†,1}, Paul G. Furtmüller^{*}, and Christian Obinger^{*}

^{*}Metalloprotein Research Group, Division of Biochemistry, Department of Chemistry, BOKU – University of Natural Resources and Applied Life Sciences, Muthgasse 18, A-1190 Vienna Austria

[†]Institute of Molecular Biology, Slovak Academy of Sciences, Dúbravská cesta 21, SK-84551 Bratislava, Slovakia

Abstract

Catalase/oxidases (KatGs) are bifunctional haem *b*-containing (Class I) oxidases with overwhelming catalase activity and substantial oxidase activity with various one-electron donors. These unique oxidoreductases evolved in ancestral bacteria revealing a complex gene-duplicated structure. Besides being found in numerous bacteria of all phyla, *katG* genes were also detected in genomes of lower eukaryotes, most prominently of sac and club fungi. Phylogenetic analysis demonstrates the occurrence of two distinct groups of fungal KatGs that differ in localization, structural and functional properties. Analysis of lateral gene transfer of bacterial *katGs* into fungal genomes reveals that the most probable progenitor was a *katG* from a bacteroidetes predecessor. The putative physiological role(s) of both fungal KatG groups is discussed with respect to known structure–function relationships in bacterial KatGs and is related with the acquisition of (phyto)pathogenicity in fungi.

Keywords

ascomycete; bacteroidete; catalase/oxidase (KatG); fungus; lateral gene transfer; phylogenetic analysis

Introduction

H₂O₂ is a harmful metabolic by-product of aerobic life that also acts as second messenger in signal transduction pathways [1]. During cellular evolution, its rapid and effective removal by various oxidoreductases was of essential importance. Cells evolved not only enzymes capable of efficient dismutation of H₂O₂ [i.e. haem catalases, manganese catalases and KatGs (catalase/oxidases)], but also enzymes that reduce hydrogen peroxide with the help of various organic and inorganic one- and two-electron donors (haem oxidases and non-haem oxidases, e.g. peroxiredoxins).

KatGs represent one of the most abundant families of Class I of the non-animal haem oxidase superfamily [2]. They are unique in accomplishing efficiently both catalytic and oxidative activity with various substrates [3]. Most currently known KatG representatives (360 sequences at present) are encoded in bacterial genomes, and mechanistic knowledge about these peculiar bifunctional oxidases has derived from studies on bacterial and

archaeal species [4]. In contrast, eukaryotic KatGs, abundant mainly among fungi and protists, have hardly been described.

It is a well-known phenomenon that LGT (lateral gene transfer) has often occurred in the evolutionary history between all three domains of cellular life, i.e. Archaea, Bacteria and Eukarya [5]. We have reconstructed the LGT from ancient bacteria into fungal genomes and discuss the functional implications in the present paper. Phylogenetic and sequence analysis reveals the presence of two distinct groups of fungal representatives, namely of intracellular and extracellular KatGs. On the basis of both multiple sequence alignment and known structure–function relationships of bacterial enzymes, these two groups are critically analysed with respect to (putative) enzymatic and functional properties.

Data mining

For multiple sequence alignment and phylogenetic analysis, complete protein sequences of 49 haem KatGs were used, including all 36 currently known fungal KatG sequences (updated from all databases in March 2009; see Table 1). For comparison, 13 selected bacterial KatGs known already from previous work [2] were added. All used protein sequences and their corresponding ID-numbers and genes can be followed in PeroxiBase [6].

Multiple sequence alignment

Multiple sequence alignment was performed with Clustal X, version 2.0 [7]. The following optimized parameters were used: for pairwise alignments, gap opening penalty 9, gap extension penalty 0.1, and for multiple alignment, gap opening penalty 8, gap extension penalty 0.2. Residue-specific penalties and hydrophilic penalties were activated, and the gap separation distance was set to 4. Gonnet protein weight matrix was used and the delay-divergent cut off was optimized to 23%.

Phylogenetic analysis

Phylogenetic reconstruction with the distance method was performed with the Mega package, version 4.1 [8]. The following optimized parameters were used: JTT matrix, pairwise deletion of gaps, gamma distribution of mutation rates with gamma optimized to 1.85. As a test of inferred phylogeny, 1000 bootstrap replicates were used. The reconstructed majority rule consensus tree was presented with the Tree Explorer of the same Mega package.

Phylogenetic reconstruction using the MP (maximum parsimony) method was accomplished also within the Mega package [8]. Parameters were optimized as follows: CNI (close-neighbour-interchange) with search level 2 was selected with the initial tree formed by random addition with 200 replicates, and all alignment sites were used for this method.

Phylogenetic reconstruction with the ML (maximum-likelihood) method was performed within the PHYLIP package, version 3.68 [9]. The ProML program with JTT matrix and gamma distribution was applied (gamma was optimized to 1.85). For statistical purposes, 100 bootstrap replicates were applied, and the consensus tree was formed with the Tree Consense program of the same PHYLIP package.

Signal sequence prediction

Putative signal sequences were detected using the predictive algorithm of the program SignalP at <http://www.cbs.dtu.dk/services/SignalP>. The available eukaryotic signal sequence database was chosen for this prediction [10].

Evolutionary relationships

The molecular evolutionary relationships within a group of 36 fungal and 13 bacterial *katG* genes (Table 1) have been analysed by application of three distinct phylogenetic methods, namely (i) NJ (neighbour-joining) distance method, (ii) MP method, and (iii) ML method. The resulting reconstructed unrooted tree obtained by using the whole coding regions is presented in Figure 1. Very similar tree topologies were obtained with the three approaches. The condensed NJ-tree is displayed with statistical support from 1000 bootstrap replications. Figure 1 clearly demonstrates that fungal *katG* genes have segregated in two distinct groups.

Group 1 comprises more representatives (denominated KatG1) and includes both pathogenic and non-pathogenic fungi. The only currently known basidiomycete representative, a plant pathogen (i.e. *Ustilago maydis*), falls in this group. The reconstructed phylogenetic tree suggests that Group 1 can be divided into three major clades. The first clade contains genes from phytopathogenic and saprobic ascomycetes. Recently, the successful heterologous expression and preliminary biochemical characterization of *Magnaporthe grisea* KatG1, a representative of clade 1, has been reported [11]. The second clade is also represented by mainly phytopathogenic fungi and includes several *Gibberella* and *Fusarium* species that even have two KatG1-encoding genes. These *katG1* sequences are paralogues, i.e. they derive from a recent intraspecies gene duplication event. Finally, aspergilli sequences of both pathogenic and non-pathogenic species dominate the third clade of Group 1. This clade includes the gene encoding *Penicillium marneffei* KatG that has been demonstrated to be induced in the virulent yeast phase of this dangerous human pathogen [12].

Group 2 (denominated KatG2) can be divided into two subclades. The two *Trichoderma* sequences were segregated rather early, whereas all six remaining KatG2 representatives are phytopathogenic sac fungi. It is interesting to see that all members of Group 2 have a N-terminal sequence for protein secretion (Table 2), whereas all members of Group 1 are intracellular proteins with high structural and functional similarities to their bacterial counterparts [11]. It is important to note that signal sequences for secretion are found also in a few bacterial KatGs but not in bacteroidetes (Table 2), which are phylogenetically closely related to fungal KatGs (Figure 1) and might be at the origin of the eukaryotic enzymes (see below). This suggests that the ascomycete-specific signal sequence has been acquired later on by adaptive evolution of KatGs from Group 2. This adaptation step resulted in unique extracellular KatGs that could be effectively involved in fungal defence against oxidative burst during plant attack [13]. This is underlined by the fact that expression of KatG2, but not KatG1, is enhanced under oxidative stress conditions [11]. In any case, only pathogenic fungi encode both KatG1 and KatG2.

It was proposed previously [2] that eukaryotic *katG* genes arose via LGT from bacterial genomes. Sequence analysis of the *katG1* promoter region as well as of the adjacent non-coding DNA region preceding the *M. grisea katG1* gene gave further evidence for LGT of only the KatG-coding region [11]. Phylogenetic reconstruction described in the present paper has revealed details of LGT from bacteria towards ancient fungi. This work demonstrates that two *katG* genes (both yet putative) from flavobacteria and sphingobacteria are the closest phylogenetic neighbours of all fungal KatGs (Figure 1, bacteroidetes clade). The data presented might indicate that LGT of *katG* genes has contributed to acquisition of pathogenicity in fungi during their adaptive evolution, consistent with previous

investigations that demonstrated the role of horizontal (lateral) gene transfer in the evolution of pathogenicity in various bacteria [14]. It has to be noted that bacteroidetes are strictly anaerobic bacteria nevertheless possessing enzymes for effective degradation of reactive oxygen species.

Essential residues and motifs

Multiple sequence alignment reveals a high level of conservation along the whole coding region (see Supplementary Figure S1 at <http://www.biochemsoctrans.org/bst/037/bst0370772add.htm>). On the distal side of the prosthetic haem group (Supplementary Figure S1A), the catalytic triad found in all Class I peroxidases (that include in addition ascorbate and cytochrome *c* peroxidases) is strictly conserved: Arg⁸⁷-Trp⁹⁰-His⁹¹ (*M. grisea* KatG1 numbering) [3]. Moreover, all complete sequences quoted in Table 1 show the presence of Asn¹²¹ (hydrogen-bonding partner of distal His⁹¹), Asp¹²⁰ (controls access of H₂O₂ to the active site) and the peculiar KatG-typical covalent adduct Trp⁹⁰-Tyr²³⁸-Met²⁶⁴ that is essential for the catalase, but not the peroxidase, activity of this unique oxidoreductase [15,16]. Regarding the proximal haem side (Supplementary Figure S1B), the triad His²⁷⁹-Trp³³⁰-Asp³⁸⁹ is also fully conserved in both groups of fungal KatGs [3]. These findings suggest that the principal mechanism of bifunctional activity as well as enzymatic parameters might be similar in prokaryotic and eukaryotic enzymes. So far, this has been demonstrated only for *M. grisea* KatG1 [11], but needs to be demonstrated for representatives of Group 2.

katG genes are the result of an ancestral gene duplication event [17] comprising an N-terminal haem-containing domain and a C-terminal domain without prosthetic group. Supplementary Figures S1C and S1D present highly conserved motifs from the C-terminal domain that resemble the distal and proximal side of the (N-terminal) haem cavity. Essential distal and proximal histidine residues are replaced by Ala⁴⁸⁶ and Leu⁶³⁰ respectively, thereby preventing haem binding, although other typical distal and proximal sequence motifs are highly conserved in the duplicated domain both in archaeal/bacterial and eukaryotic KatGs. The functional consequences of this two-domain structure are not fully understood. The C-terminal domain has been described to be essential for restructuring the active centre in the N-terminal domain [18], thereby promoting the correct architecture for the bifunctional activity of KatG [19]. This has been shown for bacterial KatGs [18,19], and it is reasonable to assume that identical structure–function relationships are valid in the (two-domain) eukaryotic enzymes.

In both fungal KatG groups, a putative gene fusion of *katG* with another coding region was detected (*Aspergillus flavus* KatG1 and *Gibberella zeae* KatG2; see Supplementary Table S1 at <http://www.biochemsoctrans.org/bst/037/bst0370772add.htm>). In the case of intracellular KatG from *Aspergillus flavus* (known to produce aflatoxin), *katG1* is fused with a gene encoding a (putative) protein of the major facilitator protein family known to mediate transport of substances across membranes [20]. In the case of extracellular *G. zeae* KatG2, the situation is even more complicated. According to gene prediction, extracellular (soluble) KatG2 might be anchored by a transmembrane anchor that is fused with an intracellular cytochrome P450 domain with homology with family 7-A (see Supplementary Figure S2 at <http://www.biochemsoctrans.org/bst/037/bst0370772add.htm>). Clearly, it has to be investigated whether these multidomain proteins are expressed completely and are (multi)functional enzymes *in vivo*. The inducible expression of sole KatG2 has already been described [21], and the spliced mRNA sequence obtained has been submitted to the EST (expressed sequence tag) database (accession number FD528596). Heterologous expression and characterization of recombinant *G. zeae* KatG2 is currently underway (P.G. Furtmüller, C. Obinger and M. Zámocký, unpublished work).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding This work was supported by the Austrian Science Fund FWF Project No. P20996.

Abbreviations used

KatG	catalase/peroxidase
LGT	lateral gene transfer
ML	maximum likelihood
MP	maximum parsimony
NJ	neighbour-joining

References

1. Zámocký M, Furtmüller PG, Obinger C. The evolution of catalases from bacteria to man. *Antioxid. Redox Signaling*. 2008; 10:1527–1548.
2. Passardi F, Zámocký M, Favet J, Jakopitsch C, Penel C, Obinger C, Dunand C. Phylogenetic distribution of catalase-peroxidases: are there patches of order in chaos? *Gene*. 2007; 397:101–113. [PubMed: 17561356]
3. Smulevich G, Jakopitsch C, Droghetti E, Obinger C. Probing the structure and bifunctionality of catalase-peroxidase (KatG). *J. Inorg. Biochem.* 2006; 100:568–585. [PubMed: 16516299]
4. Singh R, Wiseman B, Deemagarn T, Jha V, Switala J, Loewen PC. Comparative study of catalase-peroxidases (KatGs). *Arch. Biochem. Biophys.* 2008; 471:207–214. [PubMed: 18178143]
5. Syvanen M. On the occurrence of horizontal gene transfer among an arbitrarily chosen group of 26 genes. *J. Mol. Evol.* 2002; 54:258–266. [PubMed: 11821918]
6. Passardi F, Theiler G, Zámocký M, Cosio C, Rouhier N, Teixeira F, Margis-Pinheiro M, Ioannidis V, Penel C, Falquet L, Dunand C. PeroxiBase: the peroxidase database. *Phytochemistry*. 2007; 68:534–539.
7. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007; 23:2947–2948. [PubMed: 17846036]
8. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Mol. Biol. Evol.* 2007; 24:1596–1599. [PubMed: 17488738]
9. Felsenstein J. Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. *Methods Enzymol.* 1996; 266:418–427. [PubMed: 8743697]
10. Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 1997; 10:1–6. [PubMed: 9051728]
11. Zámocký M, Furtmüller PG, Bellei M, Battistuzzi G, Stadlmann J, Vlasits J, Obinger C. Intracellular catalase/peroxidase from the phytopathogenic rice blast fungus *Magnaporthe grisea*: expression analysis and biochemical characterization of the recombinant protein. *Biochem. J.* 2009; 418:443–451. [PubMed: 19000033]
12. Xi L, Liu W, Li X, Liu Y, Li M, Zhang J, Li M. Differentially expressed proteins of pathogenic *Penicillium marneffei* in yeast and mycelial phases. *J. Med. Microbiol.* 2007; 56:298–304. [PubMed: 17314357]
13. Doke N, Miura Y, Sanchez LM, Park H-J, Noritake H, Yoshioka H, Kawakita K. The oxidative burst protects plants against pathogen attack: mechanism and role as an emergency signal for plant bio-defence: a review. *Gene*. 1996; 179:45–51. [PubMed: 8955628]

14. Ehrlich GD, Hiller NL, Hu FZ. What makes pathogens pathogenic. *Genome Biol.* 2008; 9:225. [PubMed: 18598378]
15. Jakopitsch C, Kolarich D, Petutschnig G, Furtmüller PG, Obinger C. Distal side tryptophan, tyrosine and methionine in catalase-peroxidases are covalently linked in solution. *FEBS Lett.* 2003; 552:135–140. [PubMed: 14527675]
16. Regelsberger G, Jakopitsch C, Furtmüller PG, Rueker F, Switala J, Loewen PC, Obinger C. The role of distal tryptophan in the bifunctional activity of catalase-peroxidases. *Biochem. Soc. Trans.* 2001; 29:99–105. [PubMed: 11356135]
17. Zamocky M. Phylogenetic relationships in class I of the superfamily of bacterial, fungal, and plant peroxidases. *Eur. J. Biochem.* 2004; 271:3297–3309. [PubMed: 15291807]
18. Li Y, Goodwin DC. Vital roles of an interhelical insertion in catalase-peroxidase bifunctionality. *Biochem. Biophys. Res. Commun.* 2004; 318:970–976. [PubMed: 15147967]
19. Baker RD, Cook CO, Goodwin DC. Catalase-peroxidase active site restructuring by a distant and “inactive” domain. *Biochemistry.* 2006; 45:7113–7121. [PubMed: 16752901]
20. Pao SS, Paulsen IT, Saier MH Jr. Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* 1998; 62:1–34. [PubMed: 9529885]
21. Zamocky M, Jakopitsch C, Vlasits J, Obinger C. Fungal catalase peroxidases: a novel group of bifunctional oxidoreductases. *J. Biol. Inorg. Chem.* 2007; 12:S97.

The condensed NJ tree is shown, but very similar topologies were obtained using the MP method and the ML method. Bootstrap values in the nodes were obtained from NJ/MP/ML methods respectively. Alternative names and PDB codes of known structures are given in parentheses.

Table 1
Alphabetical list of 49 protein sequences used for phylogenetic analyses

Members of Group 1 of (intracellular) fungal peroxidases are designated KatG1, whereas extracellular KatGs are designated KatG2. Sequences deposited in databases as 'KatG3 (*)' are actually KatG1 paralogues, as shown in the text. Alternative names for some sequences are also shown.

Abbreviation	PeroxiBase ID	Organism	Phylum
AclKatG1	3409	<i>Aspergillus clavatus</i>	Ascomycota
AflKatG1	3394	<i>Aspergillus flavus</i>	Ascomycota
AfumKatG1	1881	<i>Aspergillus fumigatus</i>	Ascomycota
AniKatG1 (AntCpeA)	1905	<i>Aspergillus nidulans</i>	Ascomycota
AnKatG1 (AnCpeB)	5224	<i>Aspergillus niger</i>	Ascomycota
AorKatG1	3448	<i>Aspergillus oryzae</i>	Ascomycota
AteKatG1	3413	<i>Aspergillus terreus</i>	Ascomycota
BBFLKatG1	3590	<i>Flavobacterium bacterium BBFL7</i>	Bacteroidetes
BBFLKatG2	3610	<i>Flavobacterium bacterium BBFL7</i>	Bacteroidetes
BfkKatG	2290	<i>Burkholderia xenovorans</i> (formerly <i>Burkholderia fungorum</i>)	Proteobacteria
BgKatG1 (BgCPX)	1960	<i>Blumeria graminis</i>	Ascomycota
BpKatG	2303	<i>Burkholderia pseudomallei</i>	Proteobacteria
CgKatG1	3408	<i>Chaetomium globosum</i>	Ascomycota
EcoHPI	2394	<i>Escherichia coli</i>	Proteobacteria
EcoKatP	2386	<i>Escherichia coli</i>	Proteobacteria
FjKatG	3617	<i>Flavobacterium johnsoniae</i>	Bacteroidetes
FoKatG1	5228	<i>Fusarium oxysporum</i>	Ascomycota
FoKatG2	5229	<i>Fusarium oxysporum</i>	Ascomycota
FoKatG3*	5353	<i>Fusarium oxysporum</i>	Ascomycota
FtKatG	2678	<i>Francisella tularensis</i>	Proteobacteria
GmoKatG1	3415	<i>Gibberella moniliformis</i>	Ascomycota
GmoKatG2	3414	<i>Gibberella moniliformis</i>	Ascomycota
GzKatG1	2477	<i>Gibberella zeae</i>	Ascomycota
GzKatG2	2338	<i>Gibberella zeae</i>	Ascomycota
GzKatG3*	5354	<i>Gibberella zeae</i>	Ascomycota
HjKatG1	2539	<i>Hypocrea jecorina</i> (<i>Trichoderma reesei</i>)	Ascomycota
HjKatG2	5459	<i>Hypocrea jecorina</i> (<i>Trichoderma reesei</i>)	Ascomycota

Abbreviation	PeroxisBase ID	Organism	Phylum
HmaKatG1	2440	<i>Haloarcula marismortui</i>	Euryarchaeota
MagKatG1	2288	<i>Magnaporthe grisea</i>	Ascomycota
MagKatG2	2337	<i>Magnaporthe grisea</i>	Ascomycota
MgKatG1	2538	<i>Mycosphaerella graminicola</i>	Ascomycota
MuuKatG	3571	<i>Mycobacterium tuberculosis</i>	Actinobacteria
NcKatG1 (NcCat2)	2181	<i>Neurospora crassa</i>	Ascomycota
NfKatG1	3412	<i>Neosartorya fischeri</i>	Ascomycota
NhaeKatG1	5480	<i>Nectria haematococca</i>	Ascomycota
NhaeKatG2	5374	<i>Nectria haematococca</i>	Ascomycota
PanKatG1	5373	<i>Podospora anserina</i>	Ascomycota
PenchrKatG1	6601	<i>Penicillium chrysogenum</i>	Ascomycota
PEspKatG	5801	<i>Pedobacter</i> sp.	Bacteroidetes
PmKatG1 (PmCPE1)	2182	<i>Penicillium mamefei</i>	Ascomycota
PnoKatG1	3449	<i>Phaeosphaeria nodorum</i>	Ascomycota
PtritKatG1	5798	<i>Pyrenophora tritici-repentis</i>	Ascomycota
SspKatG1	2714	<i>Synechococcus</i> sp.	Cyanobacteria
SyspKatG1	2479	<i>Synechocystis</i> sp.	Cyanobacteria
TatKatG1	5797	<i>Trichoderma atroviride</i>	Ascomycota
TatKatG2	5800	<i>Trichoderma atroviride</i>	Ascomycota
TstKatG1	5796	<i>Talaromyces stipitatus</i>	Ascomycota
TviKatG1	5799	<i>Trichoderma virens</i>	Ascomycota
UmKatG1	2327	<i>Ustilago maydis</i>	Basidiomycota

Table 2
Alphabetical list of protein sequences revealing signal sequences for secretion [10]

All members of Group 2 (KatG2) have an N-terminal signal sequence, whereas members of Group 1 (KatG1) do not have a signal sequence.

Abbreviation	Organism	Signal sequence length (amino acids); first native amino acid	Signal peptide probability (hidden Markov model)
AtKatG1	<i>Aspergillus flavus</i>	Non-secretory	0.000
BBFLKatG1	<i>Flavobacterium bacterium BBFL7</i>	Non-secretory	0.267
BBFLKatG2	<i>Flavobacterium bacterium BBFL7</i>	Non-secretory	0.000
EcoKatP	<i>Escherichia coli</i>	23; alanine	1.000
FjKatG	<i>Flavobacterium johnsoniae</i>	Non-secretory	0.000
FoKatG1	<i>Fusarium oxysporum</i>	18; glutamine	0.998
FoKatG2	<i>Fusarium oxysporum</i>	18; glutamine	0.999
FoKatG3	<i>Fusarium oxysporum</i>	Non-secretory	0.224
GmoKatG1	<i>Gibberella moniliformis</i>	Non-secretory	0.231
GmoKatG2	<i>Gibberella moniliformis</i>	18; glutamine	0.999
GzKatG2	<i>Gibberella zeae</i>	18; aspartic acid	1.000
HjKatG2	<i>Hypocrea jecorina (Trichoderma reesei)</i>	18; isoleucine	0.997
MagKatG2	<i>Magnaporthe grisea</i>	23; glutamine	1.000
NhaeKatG2	<i>Nectria haematococca</i>	21; aspartic acid	1.000
PEspKatG	<i>Pedobacter</i> sp.	Non-secretory	0.000
TatKatG2	<i>Trichoderma atroviride</i>	18; valine	0.999