

Acta Crystallographica Section D

**Biological
Crystallography**

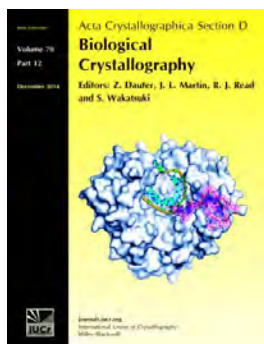
ISSN 1399-0047

Structural implications of the C-terminal tail in the catalytic and stability properties of manganese peroxidases from ligninolytic fungi

Elena Fernández-Fueyo, Sandra Acebes, Francisco J. Ruiz-Dueñas, María Jesús Martínez, Antonio Romero, Francisco Javier Medrano, Victor Guallar and Angel T. Martínez

Acta Cryst. (2014). **D70**, 3253–3265

This open-access article is distributed under the terms of the Creative Commons Attribution Licence <http://creativecommons.org/licenses/by/2.0/uk/legalcode>, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are cited.



Acta Crystallographica Section D: Biological Crystallography welcomes the submission of papers covering any aspect of structural biology, with a particular emphasis on the structures of biological macromolecules and the methods used to determine them. Reports on new protein structures are particularly encouraged, as are structure–function papers that could include crystallographic binding studies, or structural analysis of mutants or other modified forms of a known protein structure. The key criterion is that such papers should present new insights into biology, chemistry or structure. Papers on crystallographic methods should be oriented towards biological crystallography, and may include new approaches to any aspect of structure determination or analysis. Papers on the crystallization of biological molecules will be accepted providing that these focus on new methods or other features that are of general importance or applicability.

Crystallography Journals **Online** is available from journals.iucr.org

Structural implications of the C-terminal tail in the catalytic and stability properties of manganese peroxidases from ligninolytic fungi

Elena Fernández-Fueyo,^a
 Sandra Acebes,^b Francisco J.
 Ruiz-Dueñas,^a María Jesús
 Martínez,^a Antonio Romero,^a
 Francisco Javier Medrano,^{a*}
 Victor Guallar,^{b,c,*} and Angel T.
 Martínez^{a*}

^aCentro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain,
^bJoint BSC–CRG–IRB Research Program in Computational Biology, Barcelona Supercomputing Center, Jordi Girona 29, 08034 Barcelona, Spain, and ^cICREA, Passeig Lluís Companys 23, 08010 Barcelona, Spain

Correspondence e-mail:
 fjmedrano@cib.csic.es, victor.guallar@bsc.es,
 atmartinez@cib.csic.es

The genome of *Ceriporiopsis subvermispora* includes 13 manganese peroxidase (MnP) genes representative of the three subfamilies described in ligninolytic fungi, which share an Mn²⁺-oxidation site and have varying lengths of the C-terminal tail. Short, long and extralong MnPs were heterologously expressed and biochemically characterized, and the first structure of an extralong MnP was solved. Its C-terminal tail surrounds the haem-propionate access channel, contributing to Mn²⁺ oxidation by the internal propionate, but prevents the oxidation of 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonate) (ABTS), which is only oxidized by short MnPs and by shortened-tail variants from site-directed mutagenesis. The tail, which is anchored by numerous contacts, not only affects the catalytic properties of long/extralong MnPs but is also associated with their high acidic stability. Cd²⁺ binds at the Mn²⁺-oxidation site and competitively inhibits oxidation of both Mn²⁺ and ABTS. Moreover, mutations blocking the haem-propionate channel prevent substrate oxidation. This agrees with molecular simulations that position ABTS at an electron-transfer distance from the haem propionates of an *in silico* shortened-tail form, while it cannot reach this position in the extralong MnP crystal structure. Only small differences exist between the long and the extralong MnPs, which do not justify their classification as two different subfamilies, but they significantly differ from the short MnPs, with the presence/absence of the C-terminal tail extension being implicated in these differences.

Received 10 July 2014
 Accepted 16 October 2014

PDB references: MnP6, 4czn;
 MnP6–Mn²⁺, 4czo; 4czp;
 MnP6–Cd²⁺, 4czq; 4czr

1. Introduction

Fungal degradation of lignin, a key step in carbon recycling in land ecosystems, has been described as an 'enzymatic combustion' (Kirk & Farrell, 1987), in which high redox-potential haem peroxidases – the so-called lignin peroxidases (LiPs), manganese peroxidases (MnPs) and versatile peroxidases (VPs) – play a central role (Ruiz-Dueñas & Martínez, 2009). Recent evolutionary studies (Ruiz-Dueñas *et al.*, 2013) show that MnPs evolved from fungal generic peroxidases (similar to plant peroxidases) by developing an Mn-binding site formed by three acidic residues near the haem propionates, as shown in the first MnP crystal structure (Sundaramoorthy *et al.*, 1994). MnPs then gave rise to VPs by incorporating an exposed catalytic tryptophan and finally to LiPs by loss of the VP Mn-oxidation site, with the presence of the exposed tryptophan being characteristic of both LiP (Piontek *et al.*, 1993; Poulos *et al.*, 1993) and VP (Pérez-Boada *et al.*, 2005) crystal structures, and the latter also conserving the above Mn-binding site. Owing to their ancestral origin, MnPs have been reported from many lignin-degrading

basidiomycetes (the so-called white-rot fungi) in the orders Polyporales and Agaricales, a distribution that is being enlarged through genomic studies.

After deciphering the first basidiomycete genome (Martinez *et al.*, 2004), that of the model white-rot fungus *Phanerochaete chrysosporium*, a number of genomes of other wood-decaying basidiomycetes have been sequenced at the Joint Genome Institute (JGI; Floudas *et al.*, 2012; Ruiz-Dueñas *et al.*, 2013), looking for biotechnological tools that will allow the complete utilization of plant biomass in ligno-cellulose biorefineries (Ragauskas *et al.*, 2006). In these transformations, lignin removal or modification is a prerequisite for the production of cellulose, fermentable sugars and other products. Therefore, lignin-degrading organisms and enzymes are the biocatalysts of choice for the initial processing of plant feedstocks in biomass routes to substitute for the current petrochemical routes in the development of a sustainable bioeconomy (Martínez *et al.*, 2009).

A recent survey of over 30 fungal genomes by Floudas *et al.* (2012) (i) provided evidence for the involvement of LiPs, VPs and MnPs in lignin degradation, since the corresponding genes are present in all sequenced white-rot fungi and are absent from all brown-rot fungi (that depolymerize cellulose without degrading lignin), and (ii) recognized three MnP subfamilies defined by the length of the predicted C-terminal tail (C-tail). The above correlation between the presence of ligninolytic peroxidase genes and the wood-rotting decay pattern is still valid, although Riley *et al.* (2014) have recently reported the absence of these genes in the genomes of two atypical white-rot fungal species. The first-described MnP from *P. chrysosporium* has largely been characterized (Gold *et al.*, 2000), being representative of the long MnPs (357-residue mature protein). This was the third peroxidase for which a crystal structure has been solved (after cytochrome *c* peroxidase and LiP), and several high-resolution structures are currently available (Sundaramoorthy *et al.*, 2005, 2010). Extralong MnPs (366 residues) were described as having a longer C-tail and a higher thermostability (Li *et al.*, 2001), but no crystal structures are available. Finally, Mn-independent activity on 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonate) (ABTS) and some phenols has been reported for some short MnPs (Giardina *et al.*, 2000; Steffen *et al.*, 2002), and the first crystal structure of a short MnP (the 337-residue *Pleurotus ostreatus* MnP) has recently been solved (Fernández-Fueyo *et al.*, 2014).

The genome of *Ceriporiopsis subvermispora* was sequenced as a selective lignin degrader of interest for industrial delignification (Scott *et al.*, 1998). A large expansion in the number of MnP genes was found, to give a total of 13, which is more than in any other known fungal genome (Fernández-Fueyo, Ruiz-Dueñas, Ferreira *et al.*, 2012). Interestingly, this genome includes representatives of the three MnP subfamilies (Floudas *et al.*, 2012), a fact that has never been reported for any other genome to date. This enabled comparison of the three subfamilies without the interference from the evolutionary history of the species that can be produced when genes from different species (occupying different habitats) are compared. With this purpose, *C. subvermispora* extra-long,

long and short MnP genes were heterologously expressed and the kinetic and stability properties of the enzymes were determined. At the same time, the first crystal structure of an extralong MnP was solved from this fungus and was compared with the known structures of other MnPs. This structure and an *in silico* shortened-tail form were used in substrate diffusion and electronic coupling simulations to explain the differences in the oxidized substrates. Finally, several directed variants were produced at the C-tail and the haem-propionate channel. The final goal is to obtain a structural and functional validation of the three MnP subfamilies and to study the involvement of the C-tail in the different enzyme properties that are observed.

2. Materials and methods

2.1. Genome sequencing and inventory of MnP genes

The genome of *C. subvermispora* was sequenced at the JGI in a project coordinated by Daniel Cullen (FPL, USDA, Madison, USA) and Rafael Vicuña (Pontificia Universidad Católica, Santiago, Chile). The results of gene annotation (a total of 12 125 models) are available at the JGI portal (<http://genome.jgi.doe.gov/Cersu1/Cersu1.home.html>). The inventory of *C. subvermispora* haem-peroxidase genes was obtained as described by Fernández-Fueyo, Ruiz-Dueñas, Miki *et al.* (2012). The manual annotation process included, among other steps, confirmation of the presence of characteristic residues at the haem pocket and putative substrate-oxidation sites after homology modelling.

2.2. Heterologous expression of coding sequences

The coding sequences of the *C. subvermispora* gene models 49863, 50686, 117436, 157986 and 124076 corresponding to the MnP5, MnP6, MnP10, MnP12 and MnP13 isoenzymes, respectively, were designed with *Escherichia coli* preferred codon usage and were synthesized by ATG:biosynthetics. The five genes were cloned in the pET-23a(+) vector (Novagen) for expression and directed mutagenesis. *E. coli* DH5 α was used for plasmid propagation, and peroxidases were produced in *E. coli* BL21 (DE3) pLysS cells, folded *in vitro* and purified (Fernández-Fueyo, Ruiz-Dueñas, Miki *et al.*, 2012). The molecular masses from SDS-PAGE matched those from the predicted protein sequences, and the electronic absorption spectra showed the Soret and other bands characteristic of the peroxidase resting state.

2.3. Directed mutagenesis

Stop mutations were introduced into the MnP6 (G348stop) and MnP10 (G344stop) coding sequences by polymerase chain reaction (PCR) using the expression plasmids pET-23a-50686 and pET-23a-117436 (see above) as templates (Fernández-Fueyo, Ruiz-Dueñas, Miki *et al.*, 2012). The 5'-CC CTC ATT CCT CAC TGC TCC GAC TAA CTC GAG AAC TGC-3' and 5'-CCC CAT TGC AAC TAA GGC CAA GAC TGC CC-3' direct primers (the mutated codons are shown in italics) and reverse primers bearing the complementary sequences were

Table 1
Crystallographic data-collection and refinement statistics.

Values in parentheses are for the last resolution shell.

	MnP6	MnP6 + Mn ²⁺	MnP6 + Mn ²⁺ (anomalous)	MnP6 + Cd ²⁺	MnP6 + Cd ²⁺ (anomalous)
Data collection					
Space group	<i>P</i> 4 ₃ 2 ₁ 2	<i>P</i> 4 ₃ 2 ₁ 2	<i>P</i> 4 ₃ 2 ₁ 2	<i>P</i> 4 ₃ 2 ₁ 2	<i>P</i> 4 ₃ 2 ₁ 2
Unit-cell parameters (Å)	<i>a</i> = <i>b</i> = 108.6, <i>c</i> = 68.8	<i>a</i> = <i>b</i> = 108.5, <i>c</i> = 68.5	<i>a</i> = <i>b</i> = 108.7, <i>c</i> = 68.2	<i>a</i> = <i>b</i> = 108.8, <i>c</i> = 68	<i>a</i> = <i>b</i> = 108.8, <i>c</i> = 68.4
Wavelength (Å)	0.933400	0.980110	1.741350	0.980110	1.922240
Resolution range (Å)	50.00–1.20 (1.27–1.20)	50.00–1.20 (1.27–1.20)	50.00–1.90 (2.01–1.90)	50.00–1.20 (1.27–1.20)	50.00–1.90 (2.10–1.98)
Total No. of reflections	1124979	1410599	576129	1055455	311398
No. of unique reflections	128076	127081	60358	128714	47444
R _{merge} (%)	8.2 (119.6)	8.3 (174.2)	5.5 (34.5)	4.7 (128.9)	2.9 (6.4)
Completeness (%)	99.9 (99.3)	99.5 (97.2)	97.9 (87.4)	99.8 (98.9)	87.4 (33.4)
<i>I</i> / <i>σ</i> (<i>I</i>)	18.7 (1.8)	14.2 (1.0)	25.2 (2.9)	19.2 (1.1)	42.6 (7.3)
Multiplicity	8.8 (8.6)	11.1 (7.1)	9.5 (3.5)	8.2 (4.9)	6.6 (1.4)
Solvent content (%)	50.77	50.77	50.77	50.77	50.77
Matthews coefficient (Å ³ Da ⁻¹)	2.50	2.50	2.50	2.50	2.50
Subunits in asymmetric unit	1	1	1	1	1
Wilson <i>B</i> factor (Å ²)	10.5	13.7	26.1	15.1	21.8
Refinement					
Resolution range (Å)	50.0–1.20	50.0–1.20	50.0–1.90	50.0–1.20	50.0–1.98
Working reflections	127105	126815	60357	127474	47437
R _{work} /R _{free} (%)	13.4/15.1	13.9/15.4	16.7/19.1	15.1/16.2	15.9/19.1
Protein atoms (non-H)	2701	2713	2701	2715	2701
Haem groups	1	1	1	1	1
Ca ²⁺ ions	2	2	2	2	2
Water molecules	595	491	296	445	333
Na ⁺ ions	1	0	0	0	0
Mn ²⁺ ions	0	2	2	0	0
Cd ²⁺ ions	0	0	0	3	3
Mean <i>B</i> factors (Å²)					
Protein atoms (non-H)	12.56	18.86	28.17	18.86	22.53
Haem group	8.89	11.49	23.42	14.75	18.76
Ca ²⁺	7.90	10.99	21.96	13.06	15.65
Water molecules	25.78	28.24	33.04	30.42	31.46
Na ⁺ ions	12.28	—	—	—	—
Mn ²⁺ ions	—	20.51	39.05	—	—
Cd ²⁺ ions	—	—	—	20.42	45.60
Deviations from ideality					
R.m.s.d., bond lengths (Å)	0.008	0.009	0.007	0.008	0.007
R.m.s.d., angles (°)	1.244	1.371	1.079	1.366	1.231
Ramachandran plot statistics (%)					
Preferred	98.08	98.08	97.53	97.26	96.70
Allowed	1.65	1.92	2.47	2.74	3.30
Outliers	0.27	0.00	0.00	0.00	0.00
PDB code	4czn	4czo	4czp	4czq	4czr

used, respectively. 11 additional variants (E35L, E39L, G82L, D85L, D179V, E35L/E39L, G82L/D85L, D85L/D179V, E35L/E39L/G82L, E35L/E39L/D85L and E35L/E39L/D179V) were obtained by mutating the following residues (located around the haem-propionate channel) in the previously obtained MnP6 short variant (with the G348stop mutation): E35L (GAA for CTA), E39L (GAG for CTG), G82L (GGT for CTT), D85L (GAC for CTC) and D179V (GAC for GTC).

2.4. Kinetics of Mn²⁺ and ABTS oxidation and of inhibition by Cd²⁺

Absorbance changes during substrate oxidation in 0.1 M tartrate (at different pH values) were recorded using a UV-160 spectrophotometer (Shimadzu). The oxidation of Mn²⁺ was followed at pH 5 by the formation of the Mn³⁺-tartrate complex ($\epsilon_{238} = 6.5 \text{ mM}^{-1} \text{ cm}^{-1}$). ABTS oxidation was

followed at pH 3.5 by the formation of the ABTS cation radical ($\epsilon_{436} = 29.3 \text{ mM}^{-1} \text{ cm}^{-1}$). All reactions were carried out at 25°C (using ~0.01 μM enzyme) and were initiated by the addition of 0.1 mM H₂O₂. Means and standard errors for the apparent affinity constant (*K*_m) and enzyme turnover (*k*_{cat}) were obtained by nonlinear least-squares fitting to the Michaelis–Menten model. Constant fitting to the normalized equation $v = (V_{\text{max}}/K_m)[S]/(1 + [S]/K_m)$, where *v* is the reaction rate and [S] is the substrate concentration, yielded the catalytic efficiency as *V*_{max}/*K*_m values, which were converted into *k*_{cat}/*K*_m, and corresponding standard errors.

The apparent kinetic constants of MnP6 short-variant oxidation of (i) Mn²⁺ at pH 5.0 in the presence of 0–1 mM Cd²⁺ and (ii) ABTS at pH 4.5 in the presence of 0–30 mM Cd²⁺ were determined as above to investigate inhibitory effects. Inhibition constants (*K*_i) were obtained from inverse plots.

2.5. pH and temperature stabilities and the optimal pH

To study the effect of pH on stability, the enzymes were dissolved in Britton–Robinson (B&R) buffer pH 2–9 and kept at 4°C for 24 h. Activity was estimated by the oxidation of Mn^{2+} (1.5 mM) in 0.1 M tartrate pH 5 at 25°C. The reaction was initiated by the addition of 0.1 mM H_2O_2 as described above. The highest activity immediately after mixing (at any pH) was taken as 100%, and the percentage of residual activity at the different pH conditions was referred to this value.

To study the effect of temperature on enzyme stability, the enzymes in 10 mM tartrate pH 5 were incubated at 5°C intervals in the range 25–70°C. Residual activity was estimated as described above. Activity immediately after mixing at 25°C was considered as 100%, and the percentage of residual activity was estimated referring to this value. Temperature stability was presented as the 10 min T_{50} , *i.e.* the temperature at which 50% of the activity was lost after a 10 min incubation.

The same experiments were performed in the presence of 1 mM CaCl_2 , 1 mM MnSO_4 , 1 mM CdSO_4 and 0.5 mM MnSO_4 plus 0.5 mM CdSO_4 , using the extralong MnP6 and its short variant, to evaluate the effect of these cations on the pH and thermal stabilities.

The optimal pH values for substrate oxidation were determined by measuring the oxidation of Mn^{2+} (1.5 mM) and ABTS (5 mM) in 0.1 M tartrate pH 3–6 and B&R buffer pH 2–6, respectively, as described above.

2.6. Crystallization and refinement

Crystallization trials were carried out at 22°C by the sitting-drop vapour-diffusion method. 96-well MRC2 plates (with 50 μl reservoir solution) and commercially available screens from Emerald Bio and Jena Biosciences were used. The drops consisted of 0.2 μl protein solution (10 mg ml^{-1} in 10 mM sodium tartrate pH 5.0) and 0.2 μl reservoir solution. Crystals suitable for X-ray data collection were obtained in 0.1 M sodium acetate pH 4.6, 18% PEG 4000. Crystals were soaked with 5 mM MnCl_2 or CdCl_2 in the crystallization solution to obtain the protein–metal complexes.

For X-ray data collection, crystals were transferred into a cryosolution containing the well condition and 20% (*w/v*) glycerol and were cooled in liquid nitrogen prior to data collection. X-ray diffraction intensities were collected on the X06DA beamline at the Swiss Light Source, Villigen, Switzerland, the BL21-XALOC beamline at ALBA, Barcelona, Spain and the PROXIMA1 beamline at SOLEIL, Gif-sur-Yvette, France at 100 K. Diffraction data were indexed, integrated, merged and scaled using *XSCALE* and *XDS* (Kabsch, 2010). Data-collection statistics are shown in Table 1.

The structure was solved by molecular replacement using the crystal structure of *P. chrysosporium* MnP (PDB entry 3m5q; Sundaramoorthy *et al.*, 2010) as the search model and the *AutoMR* program implemented in the *PHENIX* package (Adams *et al.*, 2010). The final models were obtained by consecutive rounds of refinement performed with the *PHENIX* package followed by manual building performed

with *Coot* (Emsley & Cowtan, 2004) using σ_A -weighted $2F_o - F_c$ and $F_o - F_c$ electron-density maps. Solvent molecules were introduced into the structure automatically in the refinement as implemented in the *PHENIX* package and were visually inspected. A total of 5% of the reflections was used to calculate the R_{free} value throughout the refinement process. The root-mean square (r.m.s.) deviation for the common C^α residues of the extralong MnP when compared with the long MnP used as a model (PDB entry 3m5q) is 0.62 Å for 352 residues, with 76% sequence identity for a total of 361 residues. The structures were validated using *MolProbity* (Chen *et al.*, 2010). Data-collection, refinement and final model statistics are summarized in Table 1. Figures were produced with *PyMOL* (Schrödinger) and *DeepView/Swiss-PdbViewer* (<http://www.expasy.org/spdbv>). The coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 4czn (MnP6), 4czo (MnP6– Mn^{2+}), 4czp (MnP6– Mn^{2+} , anomalous), 4czq (MnP– Cd^{2+}) and 4czr (MnP– Cd^{2+} , anomalous).

The final model of the noncomplexed protein consisted of 366 amino acids, one haem group, two Ca^{2+} ions and one Na^+ ion. This includes two extra amino acids at the N-terminus of the protein that came from the cloning stages; the last C-terminal residue (Pro365) was not visible in the electron-density map. The final models for the Mn and Cd complexes consisted of 367 amino acids, one haem group, two Ca^{2+} ions and two Mn^{2+} ions (Mn complex) or three Cd^{2+} ions (Cd complex). The extra residue in the latter models corresponded to the last Pro365, which became visible in the electron-density maps with the metals bound to the protein.

2.7. Computational simulations

The coordinates for the *C. subvermispora* extralong MnP6 were taken from PDB entry 4czn. The terminal proline (which was not visible in this crystal structure) was added. A truncated MnP6 structure was obtained by removing the last 18 residues. The protonation states were assigned assuming that the system was at pH 3.0 (to take into account the shift in the pK_a of acidic residues owing to the presence of the dianionic substrate ABTS) with the *Protein Preparation Wizard* from Schrödinger. Therefore, the haem propionates and some residues in its access channel (Glu35, Glu39 and Asp179) were protonated. Partial charges for the haem and ABTS moieties were obtained by quantum-mechanics/molecular-mechanics (QM/MM) calculations at the B3LYP(lacvp*)/OPLS2005 level of theory with *Qsite* (v.5.7; Schrödinger). Electronic couplings were obtained at the same level of QM/MM theory by the fragment charge difference (FCD) approach (Voityuk, 2012) using the e-coupling server (<http://ecouplingserver.bsc.es>). The FCD method calculates the electronic coupling of adiabatic states by diagonalization of the Hamiltonian using the charge difference operator. More details of the implementation of the FCD approach in proteins have recently been given by Wallrapp *et al.* (2013).

A 100 ns molecular-dynamics (MD) simulation was carried out with *AMBER* for both extralong MnP6 and its *in silico*

shortened-tail form. The proteins were placed in an orthorhombic box of TIP3P water with an ionic strength of 0.15 M. A time step of 2.0 fs was used throughout the simulations in combination with the *SHAKE* algorithm to constrain bond lengths involving H atoms. Nonbonded interactions were explicitly evaluated for distances below 9 Å. The particle mesh Ewald method was employed to treat long-range electrostatic interactions. Constant pressure and temperature (NPT ensemble) were maintained by weakly coupling the system to an external bath at 100 kPa and 300 K using the Martyna–Tobias–Klein barostat and the Nose–Hoover thermostat, respectively.

The *PELE* (*Protein Energy Landscape Exploration*) software (Borrelli *et al.*, 2005) was used to study ABTS diffusion.

This software uses a Monte Carlo scheme to describe the protein–ligand conformational dynamics. It is capable of reproducing their induced fit, including ligand migration and cavity search, in a computationally efficient manner (typically requiring no more than 24 h). At each iteration the algorithm performs (i) translation and rotation of the ligand, (ii) protein perturbation following a combination of normal modes, (iii) side-chain prediction and (iv) overall minimization. Subsequently, the final structure is accepted or rejected based on a Metropolis criterion (for more details, see <https://pele.bsc.es>). In this study, the ligand first freely explored the whole protein surface and was then forced to explore the area close to the haem propionates by using a cutoff distance of 20 Å between the haem CHA atom and the ABTS centre of mass.

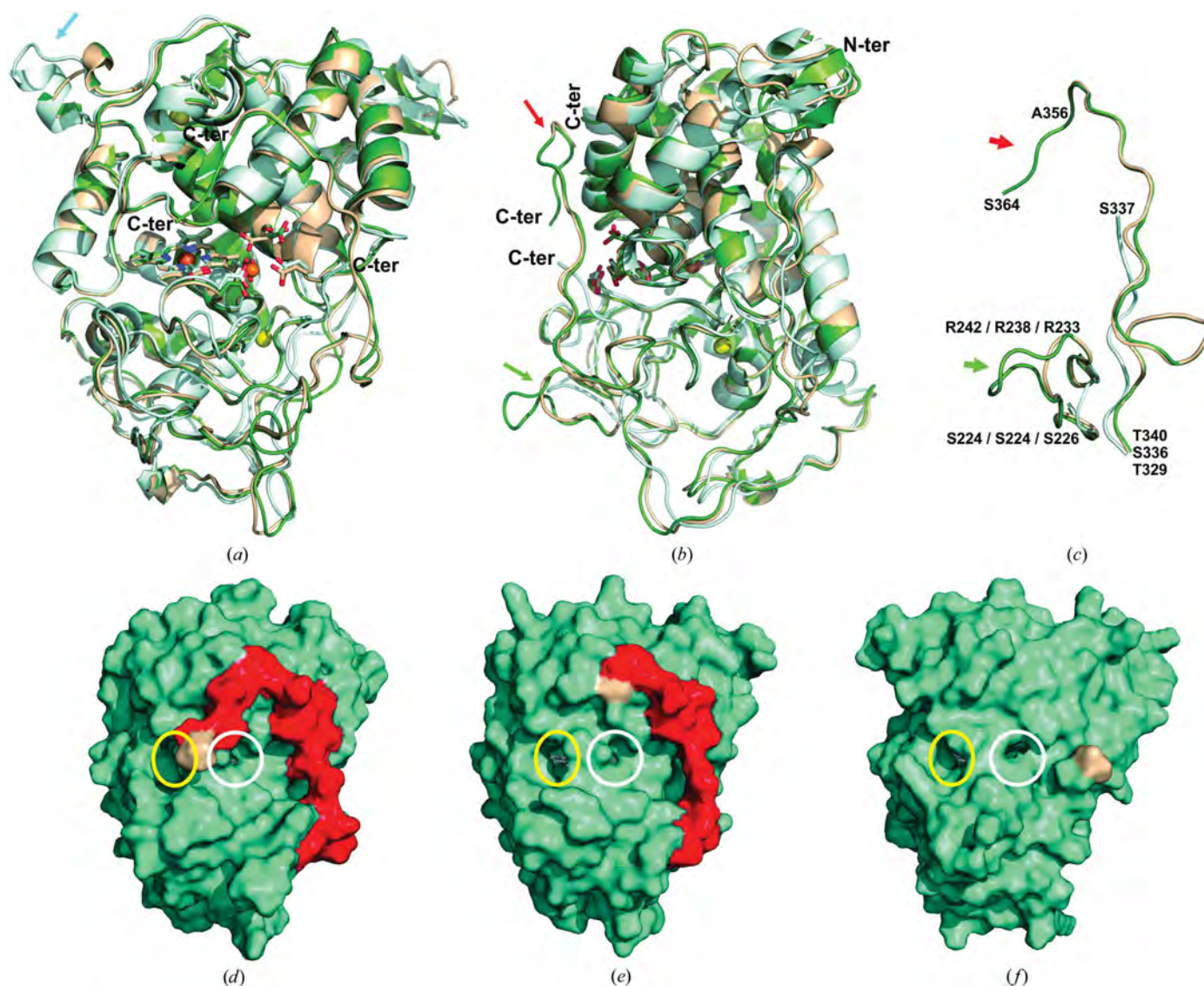


Figure 1

Crystal structure of extralong MnP compared with short and long MnPs. (*a*, *b*) Superimposition of extralong MnP6 from *C. subvermisporea* (PDB entry 4czn; green), long MnP1 from *P. chrysosporium* (PDB entry 3m5q; light brown) and short MnP4 from *P. ostreatus* (PDB entry 4bm1; light blue) with the three Mn²⁺-binding residues as CPK-coloured sticks, Na⁺ and Fe³⁺ as orange spheres and the two Ca²⁺ ions as yellow spheres (see Fig. 3*a* for the Mn²⁺-binding site). (*c*) Superimposition of the C-tail extension and neighbouring loop in extralong MnP with the same regions of long and short MnPs. (*d–f*) Surfaces of *C. subvermisporea* extralong MnP (PDB entry 4czo with Mn²⁺) and the above long and short MnPs, respectively, showing the main (yellow circles) and propionate (white circles) haem-access channels and the long and extralong C-tails in red [the final residue, in light brown, corresponds to Pro365 in (*d*), Ala357 in (*e*) and Ser337 in (*f*)].

3. Results

3.1. The first crystal structure of an extralong MnP (compared with short and long MnPs)

We have solved the crystal structure of the extralong MnP6 from *C. subvermispora* after its heterologous expression (Table 1). The overall folding is similar to that of other fungal peroxidases, with the structure divided into two domains by the haem group (Fig. 1*a*, green; Supplementary Fig. S1¹). A metal-binding site is located near the haem propionates, being occupied by Na⁺ in the recombinant enzyme (PDB entry 4czn) and coordinating one Mn²⁺ ion in PDB entry 4czo (from crystals soaked in metal-ion solution). The upper domain is mainly helical with a first structural Ca²⁺ ion, whereas the lower domain contains α -helices and a non-ordered region stabilized by a second Ca²⁺ ion. The most significant feature is the extralong tail described below in the C-terminal region (Fig. 1*b*). Most of the tail is well defined (Fig. 2*a*) and only the last residue (Pro365) that was not visible in the electron-density map and the one before (Ser364) showed significant disorder. The *B* factors for the residues forming the C-tail (Gly348–Pro365) are comparable to the average for the protein, except for Ser364, which showed a higher value, and the last Pro365, which did not show any electron density in the metal-free form of the protein (Fig. 2*b*). The presence of Mn²⁺ or Cd²⁺ bound to the protein seems to stabilize the conformation of the last two residues of the C-tail, suggesting that the metal ions might stabilize the enzyme (as described below). A comparison of the experimental *B* factors of the C-tail residues with those obtained in the MD calculations is shown in Fig. 2*b*). It can be observed that those obtained in the aqueous-medium simulation show higher values than those from the crystal medium. The extra residues that are only present in the extralong forms show a higher mobility in both approaches.

The first difference from other MnP crystal structures (Fig. 1*a*, arrow) concerns the surroundings of the distal Ca²⁺, where an exposed loop exists in short MnP (light blue) compared with long (light brown) and extralong (green) MnPs. The four extra residues (Gly32 and Pro54–Leu56 in *P. ostreatus* MnP4) are conserved in other short MnPs but are absent in all long and extralong forms examined (Supplementary Fig. S2). A second difference lies in the loop below the Mn²⁺-binding site (Figs. 1*b* and 1*c*, green arrows). In short MnPs this loop is

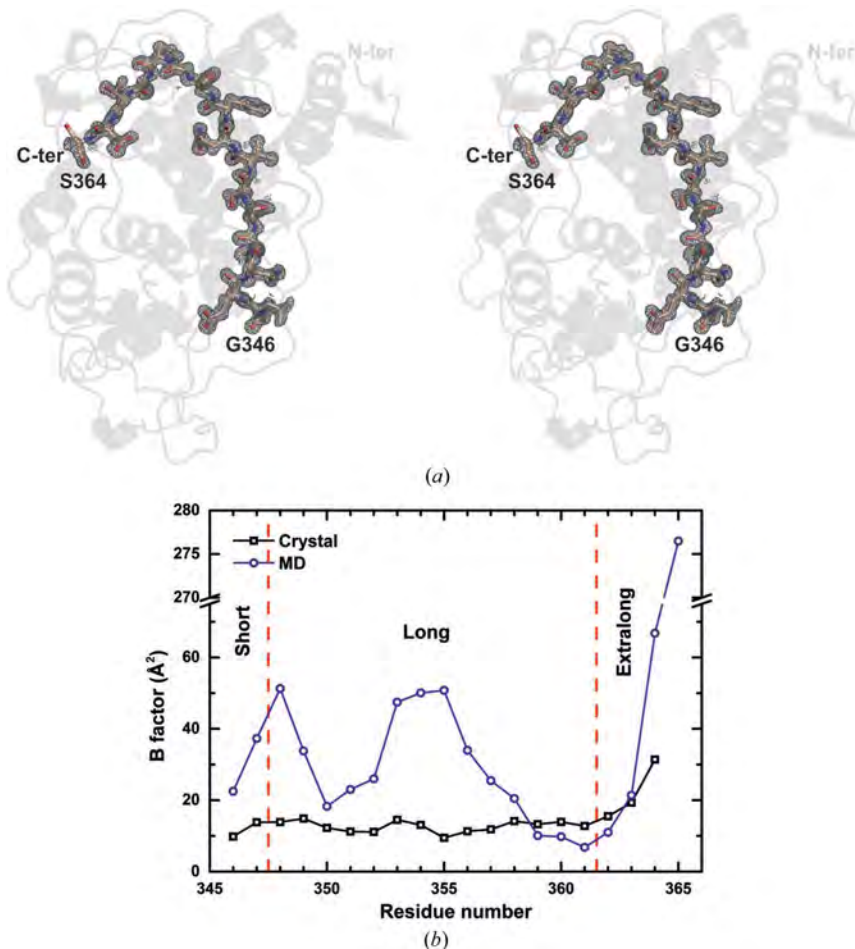


Figure 2

MnP C-terminal tail. (a) Omit map of the C-tail of *C. subvermispora* extralong MnP6. The $2F_o - F_c$ map was calculated by deleting residues Gly346–Pro365 and is shown at 1σ . (b) Comparison of the *B* factors from the crystal structure with those obtained in MD calculations (for water medium).

small and does not protrude from the surface of the protein, while it includes seven extra residues in some extralong and all of the long MnPs and four more residues in all of the other extralong MnPs (Supplementary Fig. S2). The last difference concerns the C-tail itself (Figs. 1*b* and 1*c*, red arrows), with 14 extra residues in long MnPs with respect to short MnPs and with 18–23 extra residues in extralong MnPs (Supplementary Fig. S2), including *C. subvermispora* MnP6 (the Gly348–Ser364 tail extension of which with respect to short MnPs is indicated in red in Supplementary Fig. S1; Pro365 is not present in PDB entry 4czn but is present in PDB entries 4czo and 4czq that include Mn²⁺ and Cd²⁺ ions, respectively). At the beginning of the tail, the long/extralong MnPs present one extra disulfide bridge that forms a turn pointing towards the surface. The tail goes up to near the top of the protein and then turns down towards the region between the Mn²⁺-binding site (haem-propionate channel) and the main haem-access channel. The last two differences result in a protrusion below the Mn²⁺-binding site that causes some access restrictions in long and extralong MnPs. The differences in the C-tail position

¹ Supporting information has been deposited in the IUCr electronic archive (Reference: MN5072).

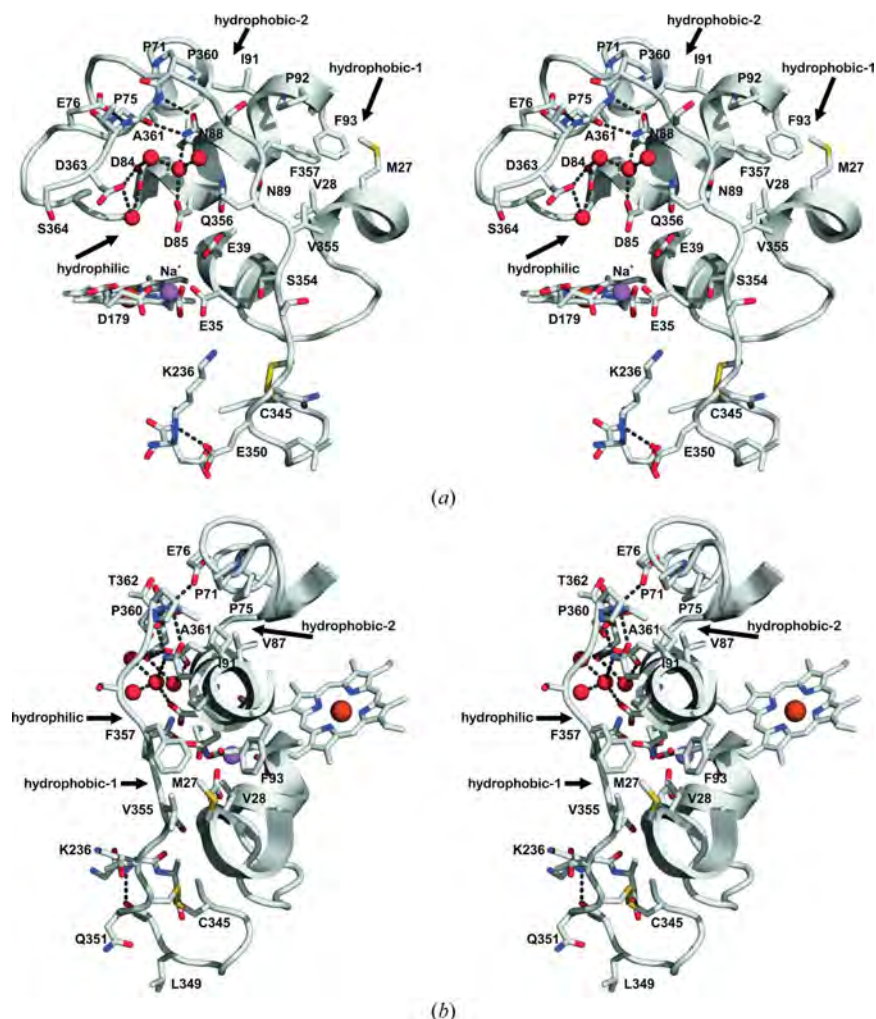


Figure 3

Extralong C-terminal tail (Cys345–Ser364). (a, b) Two stereoviews of the *C. subvermisporea* MnP6 (PDB entry 4czn) C-terminal tail with waters as red spheres and dashed lines for interactions smaller than 3.5 Å, showing (i) the first hydrophobic patch involving Val355 and Phe357 at the extralong tail, together with Met27, Val28, Pro92 and Phe93 in the protein body, (ii) the second hydrophobic patch involving Pro360 and Ala361, together with Pro71, Pro75, Val87 and Iso91, and (iii) the hydrophilic patch involving Gln356, Pro360, Ala361 and Asp363, together with Glu76, Asp84, Asp85, Asn88 and Asn89 (although they are not in close contact, the side chain of Gln356 is wedged in between the side chains of Asp85 and Asn88).

in extralong, long and short MnPs are clearly shown in Figs. 1(d), 1(e) and 1(f).

3.2. Detailed description of the extralong C-tail

Fig. 3 provides a close-up view of the C-tail in the extralong MnP crystal structure, including the interactions that lock it in place. At the base, Glu350 makes a hydrogen bond to the main-chain N atom of Lys236, and the extra disulfide bridge between Cys345 and Cys352 forces the main chain to go around in a loop. Then, several solvent molecules (not shown in the figure) connect Ser354 to the main-chain N atom of Glu35 involved in Mn^{2+} cation binding. The remaining residues form three interaction patches (two hydrophobic and one hydrophilic). The hydrophobic patches involve Val355/Pro357 and Pro360/Ala361 in the tail and four residues of the protein body in each case, while the hydrophilic patch involves

Gln356, Pro360, Ala361 and Asp363 in the tail and five residues in the protein body (Fig. 3). The main chain of Pro360 and Ala361 makes contacts with the side chains of Glu76 and Asp363, which also interact with three more residues directly or through a network of four solvent molecules.

All long and extralong MnPs (Supplementary Fig. S2) present a highly conserved sequence in the common part of the tail, while the last part does not show this conservation. This also happens with the residues that interact with the tail. Thus, six of the eight protein-body residues involved in the hydrophobic patches in *C. subvermisporea* MnP6 are conserved, one (Val28) is substituted by isoleucine or leucine, conserving the hydrophobic character, and only the position of Met27 presents a hydrophilic residue. The residues involved in the hydrophilic patches are also highly conserved. Thus, three of the four tail residues participating in these interactions in *C. subvermisporea* MnP6 are conserved and only one (Asp363) varies, although in most cases residues with side chains capable of hydrogen bonding are present. Among the five protein-body residues forming part of this interaction, three are strictly conserved, while alanine can substitute for two of them (Glu76 and Asp84). In summary, there is a high degree of conservation among all long and extralong MnPs concerning the residues that interact to lock the C-tail in place.

3.3. Catalytic properties of native MnPs and their short directed mutants

All *C. subvermisporea* native (i.e. wild-type recombinant) MnPs oxidize Mn^{2+} , but the long and extralong MnPs showed a higher catalytic efficiency ($>5000 \text{ s}^{-1} \text{ mM}^{-1}$) than the short MnPs ($\sim 750 \text{ s}^{-1} \text{ mM}^{-1}$) (Table 2). The difference resulted in lower K_m values (from 116 to 9–11 μM) when the length of the C-tail increases, suggesting a higher affinity for Mn^{2+} . The Mn^{3+} generated by MnPs can oxidize ABTS. However, the short MnP also exhibits Mn-independent activity on ABTS, as advanced in other studies, which is absent for the long and extralong forms (Table 2).

To investigate the connection between the C-tail length and the catalytic properties of MnP, short variants of extralong MnP6 (with an 18-residue shortened tail) and long MnP10 (with a 14-residue shortened tail) were produced by directed mutagenesis. The two variants possessed the ability to oxidize ABTS that was absent from the parent enzymes, with the MnP6 short variant showing the highest catalytic efficiency (Table 2). On the other hand, the efficiency in oxidizing Mn^{2+}

Table 2

Kinetics of substrate oxidation by short, long and extralong MnPs from the *C. subvermispota* genome, two directed variants at the C-tail and 11 additional directed variants at the haem-propionate channel of the MnP short variant.

	Mn ²⁺			ABTS		
	<i>K_m</i> (μM)	<i>k_{cat}</i> (s ⁻¹)	<i>k_{cat}/K_m</i> (s ⁻¹ mM ⁻¹)	<i>K_m</i> (μM)	<i>k_{cat}</i> (s ⁻¹)	<i>k_{cat}/K_m</i> (s ⁻¹ mM ⁻¹)
Native MnPs						
MnP13 (short)	116 ± 12	88 ± 2	756 ± 65	2440 ± 200	93 ± 3	38 ± 2
MnP5 (long)	50 ± 7	327 ± 15	6600 ± 700	0	0	—
MnP10 (long)	59 ± 9	331 ± 20	5600 ± 500	0	0	—
MnP6 (extralong)	9 ± 2	83 ± 5	9540 ± 653	0	0	—
MnP12 (extralong)	11 ± 1	61 ± 0.8	5560 ± 325	0	0	—
Short-tail variants†						
MnP6 short variant	16 ± 2	50 ± 1	3030 ± 390	116 ± 6	29 ± 1	251 ± 11
MnP10 short variant	34 ± 3	118 ± 4	3500 ± 360	3430 ± 370	20 ± 1	5.8 ± 0.4
Propionate-channel variants‡						
E35L	0	0	—	118 ± 13	64 ± 2	546 ± 51
E39L	0	0	—	250 ± 12	30 ± 1	120 ± 16
G82L	46 ± 5	30 ± 1	651 ± 46	1290 ± 150	100 ± 4	77 ± 5
D85L	107 ± 14	19 ± 1	179 ± 18	45 ± 8	4 ± 0	91 ± 15
D179V	0	0	—	288 ± 12	76 ± 1	263 ± 13
E35L/E39L	0	0	—	80 ± 9	34 ± 1	417 ± 38
G82L/D85L	0	0	—	0	0	—
D85L/D179V	0	0	—	35 ± 3	35 ± 1	977 ± 60
E35L/E39L/G82L	0	0	—	844 ± 121	98 ± 5	116 ± 12
E35L/E39L/D85L	0	0	—	243 ± 17	14 ± 3	59 ± 3
E35L/E39L/D179V	0	0	—	146 ± 17	2 ± 0	15 ± 1

† Variants with an 18-residue shortened C-tail. ‡ Variants at the haem-propionate channel of the MnP6 short variant.

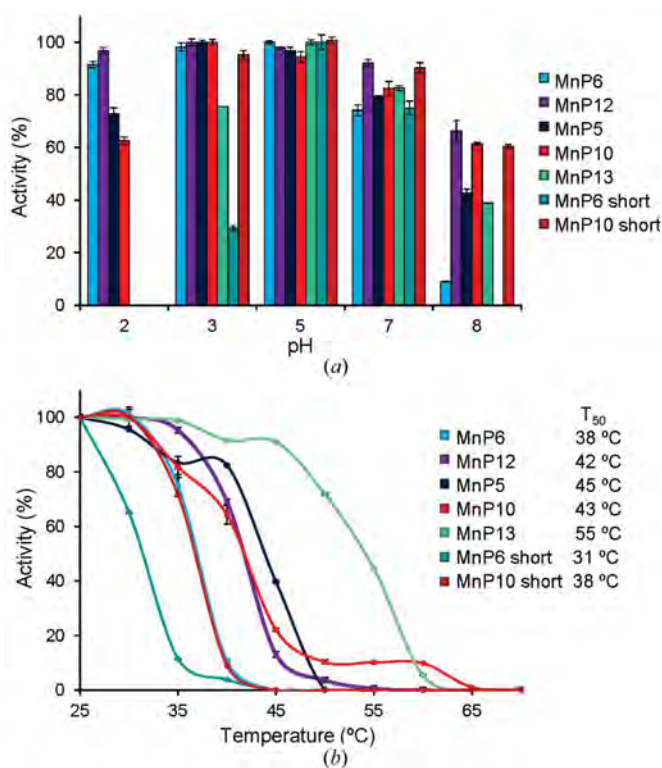


Figure 4

pH and temperature stabilities. (a) pH stability after incubating the *C. subvermispota* extralong, long and short MnPs and their short variants (obtained by partial removal of the C-tail) in the pH range 2–8 for 24 h at 4 °C (the results at pH 4 and pH 6 were similar to those obtained at pH 5). (b) Temperature stability after incubating the same MnPs and variants in the 25–55 °C range for 10 min at pH 5 (the corresponding *T*₅₀ values are included).

decreased when the C-tail was shortened, in agreement with the result obtained for the short MnP13.

Optimal oxidation of Mn²⁺ is produced at pH 5 for short MnP13, pH 4.5 for short-tail variants and pH 4 for native long and extralong MnPs. However, ABTS oxidation by the short MnP13 and the short variants was maximal at pH 4.0 and 3.5, respectively (long and extralong MnPs had no detectable activity at these pH values).

3.4. pH and thermal stabilities

All of the *C. subvermispota* native MnPs are relatively stable at pH 3, retaining over 75% activity after 24 h (Fig. 4a). However, strong differences were observed at pH 2, with the long and extralong MnPs maintaining over 60 and 90% activity, respectively, while the short MnP was very quickly (1 min) fully inactivated. Shortening the C-tail affects the pH stability, and the two short variants lost the stability at pH 2 characteristic of the native forms. With regard to temperature, a rough inverse correlation was found between the length of the C-tail and the thermal stability of the native forms (Fig. 4b), with the short MnP being more thermostable (*T*₅₀ ≈ 55 °C) than the other native MnPs (*T*₅₀ ≈ 38–45 °C). Shortening the C-tails resulted in variants with lower *T*₅₀ values.

3.5. Cation binding and its effect on stability and ABTS oxidation

The recombinant MnPs include two structural Ca²⁺ ions that are incorporated during *in vitro* folding, but have never been in contact with Mn²⁺ or Cd²⁺. Addition of any of these cations confers alkaline stability to both the native MnP6 and

its short variant (Figs. 5*a* and 5*c*, respectively), and a T_{50} higher by over 10°C was also obtained (Figs. 5*b* and 5*d*, respectively).

Binding of Mn^{2+} and Cd^{2+} (with only the first being a MnP substrate) was confirmed by anomalous difference electron-

density maps of the MnP6-Mn^{2+} and MnP6-Cd^{2+} complexes. The first map (Fig. 6*a*) showed two Mn^{2+} ions, one at the predicted Mn-oxidation site, coordinated by Glu35, Glu39, Asp179, the internal (with respect to the main haem-access

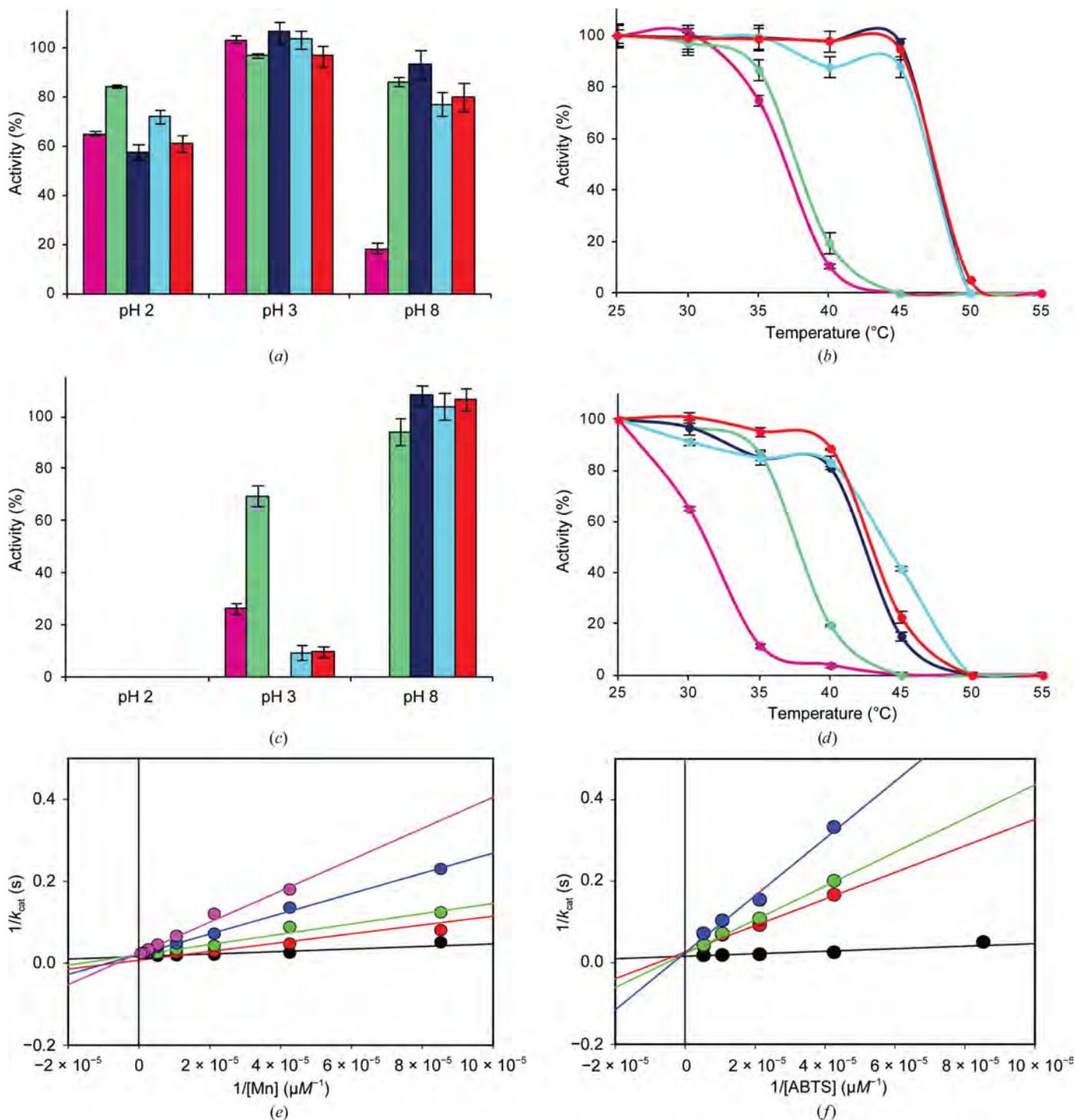


Figure 5

Effect of cations on the stability and activity of extralong MnP6 and its short variant. (*a, b*) Effect of 1 mM Ca^{2+} (green), Mn^{2+} (dark blue) or Cd^{2+} (cyan) and 0.5 mM Mn^{2+} + 0.5 mM Cd^{2+} (red) compared with a control without the above cations (magenta) on the stability of MnP6 after incubation in the pH range 2–8 for 24 h at 4°C (*a*) and in the 25–55°C range for 10 min at pH 5 (*b*). (*c, d*) Effect of cations on the pH (*c*) and thermal (*d*) stability of the MnP6 short variant assayed as described for (*a*) and (*b*). (*e, f*) Inverse plots of Mn^{2+} oxidation at pH 5.0 by the MnP6 short variant in the presence of 0 mM (black), 0.1 mM (red), 0.25 mM (green), 0.5 mM (blue) and 1.0 mM Cd^{2+} (magenta) (*e*) and ABTS oxidation at pH 4.5 in the presence of 0 mM (black), 5.5 mM (red), 15 mM (green) and 30 mM Cd^{2+} (blue) (*f*), revealing competitive inhibition in both cases.

channel) propionate of haem and two (not shown) water molecules, and the second in a neighbouring position at 6.7 Å from the external propionate. Cd²⁺ binding was observed at the same two locations, plus a third position at 10.7 Å from the external propionate (Fig. 6b).

Cd²⁺ inhibition of Mn²⁺ and ABTS oxidation by the MnP6 short variant was investigated. The inverse plots show a common ordinate intercept, indicating that Cd²⁺ is a competitive inhibitor not only of Mn²⁺ oxidation (Fig. 5e) but also of ABTS oxidation (Fig. 5f).

3.6. Directed mutagenesis at the haem-propionate channel

To confirm the ABTS oxidation site, a battery of mutants were prepared at the haem-propionate channel of the MnP6 short variant that (i) reduced the negative charge and (ii) introduced bulkier residues (Table 2). Substituting the above Mn²⁺-binding residues in single, double, triple or 'mixed' (with Gly82 and/or Asp85) variants always resulted in the complete loss of Mn-oxidation activity. Interestingly, the single and

double mutations did not reduce, but increased, the ability to oxidize ABTS.

However, the most interesting mutations were those in the G82L/D85L variant, a combination that blocks the haem-propionate channel (Fig. 7). The individual variants already showed a decrease in the ABTS oxidation efficiency, but the double mutation resulted in complete loss of activity towards both ABTS and Mn²⁺ (although the Mn²⁺-binding residues remained unchanged). Finally, a near-fourfold increase in ABTS oxidation efficiency was produced when the D85L and D179V mutations were combined in a double variant.

3.7. Molecular simulation of ABTS access to haem

Additional evidence for ABTS oxidation at the haem-propionate channel was provided by dynamic ligand diffusion using *PELE*, and other molecular-simulation tools, on the crystal structure of extralong MnP6 and an *in silico* shortened-tail form. Firstly, a 100 ns MD simulation was performed to map the conformational plasticity in the haem-propionate channel. We observed that Glu39 and Asp179 (protonated

under ABTS-oxidation pH conditions) adopt different conformations, resulting in opening/closing of the channel (Supplementary Fig. S3).

A first free exploration of ABTS diffusion on the whole enzyme was performed with *PELE* (Supplementary Figs. S4a and S4b). The main energy minimum was identified at the entrance to the haem-propionate channel (site I), where more detailed local explorations were performed (Fig. 8a). For the shortened-tail form (red dots), the structures with the best interaction energy have ABTS closer to the active site. In contrast, ABTS diffusion on extralong MnP (blue dots) failed to identify a structure with good interaction energy, and ABTS always occupied much more distant positions. Using QM/MM techniques, the electronic coupling was calculated at the best positions predicted, being one order of magnitude higher for the shortened-tail MnP than for the extralong MnP (Fig. 8a, red and blue text, respectively).

The presence and dynamics of the C-tail hinders the entrance of the ligand into the haem-propionate channel (Supplementary Figs. S4c). The tail also limits the mobility of the residues at the entrance of the channel in the extralong enzyme. The *PELE* results showed that the better binding energies in the shortened-tail form correspond to longer Glu35–Asp179 distances,

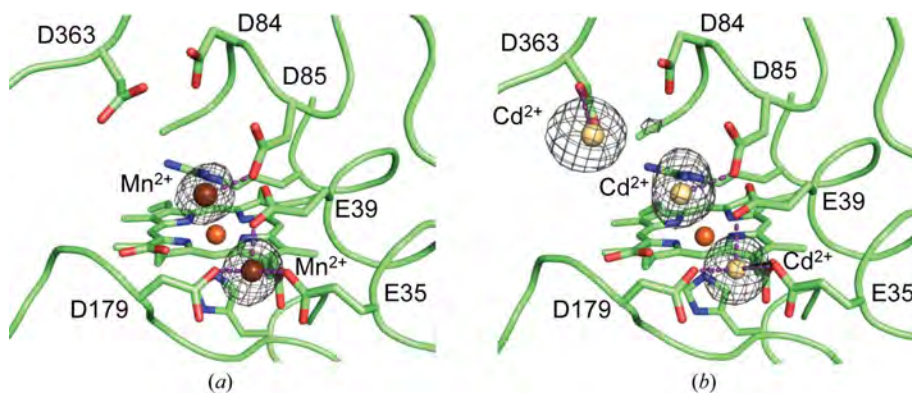


Figure 6
Cation binding on extralong MnP. (a, b) Anomalous difference electron-density maps showing two Mn²⁺ (a) and three Cd²⁺ (b) ions in the proximity of haem (and six acidic residues) in the crystal structures of metal complexes of extralong MnP6 from *C. subvermispora* (PDB entries 4czp and 4czr, respectively).

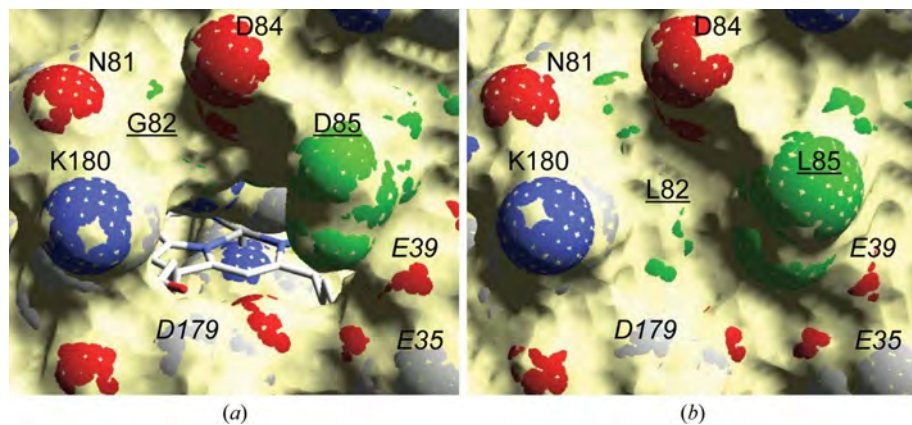


Figure 7
Haem-propionate channel in the *C. subvermispora* MnP6 short variant. (a) Channel entrance showing the haem cofactor (semitransparent surface, channel residues as van der Waals spheres and haem as CPK-coloured sticks; PDB entry 4czn). (b) Double G82L/D85L substitution closing the channel (*in silico* mutations).

enabling the approach of the ligand (Fig. 8*b*). The above is illustrated in Fig. 8(*d*), which shows ABTS approaching the haem propionate at the same time as AspH179 (and Gluh35) are displaced, widening the channel entrance in the shortened-tail form. In contrast, AspH179 and Gluh35 still block access of ABTS in the best ligand position in the extralong MnP6 (Fig. 8*c*).

4. Discussion

Recent inventories of basidiomycete peroxidases from genomes sequenced in the search for lignocellulose-degrading enzymes and other sources include growing numbers of short (up to 68), long (up to 40) and extralong (up to 13) MnP genes (Fernández-Fueyo *et al.*, 2014; Floudas *et al.*, 2012; Levasseur *et al.*, 2014; Ruiz-Dueñas *et al.*, 2013). In these studies, the three subfamilies are identified by the length of the C-tail, but evidence for biochemical and other differences is fragmentary, and their structural basis remains largely unknown.

The present comparison of short, long and extralong MnPs provides answers to some of these questions.

4.1. Structural–functional evaluation of the variable catalytic properties of MnPs

The long and extralong MnPs from *C. subvermispora* can only oxidize Mn^{2+} , as reported for *P. chrysosporium* MnP (Gold *et al.*, 2000). Their catalytic efficiency is higher than that estimated for the short MnP, but the latter is able to oxidize other substrates (such as ABTS and different phenols) in the absence of Mn^{2+} , as we previously reported for the short MnPs from the *P. ostreatus* genome (Fernández-Fueyo *et al.*, 2014). The first crystal structure of a short MnP has recently been solved by Fernández-Fueyo *et al.* (2014), and its comparison with the first extralong MnP structure solved here shows a C-tail extension (Gly348–Pro365 in *C. subvermispora* MnP6) surrounding the entrance of the haem-propionate channel.

Such a disposition limits the access of some substrates through this channel, as revealed by diffusion simulations. Indeed, shortening of the extralong tail decreases the estimated haem–ligand distance (from 6.3 to 4.0 Å), with a significant improvement in binding energy. Moreover, easier oxidation of ABTS by the short MnPs was predicted by electron-coupling calculations, with the *in silico* shortened-tail form yielding values that were an order of magnitude higher. Experimental evidence for the above predictions was obtained by reducing the C-tail length in extralong and long MnPs by directed mutagenesis, which resulted in the introduction of the ability to oxidize ABTS.

The above results, together with the observed competitive inhibition by Cd^{2+} of both Mn^{2+} and ABTS oxidation, strongly pointed to ABTS oxidation at the haem-propionate channel of short MnPs. However, an effect of the C-tail on oxidation of ABTS at the main haem-access channel, where oxidation of dyes and phenols has been shown in *P. eryngii* VP (Morales *et al.*, 2012), could not be discarded. Therefore, additional mutations at the haem-propionate channel were performed on an ABTS-oxidizing short variant. In this way, complete loss of ABTS and Mn^{2+} oxidation was obtained by a double mutation (G82L/D85L) that completely blocked the haem-propionate channel, demonstrating that the two substrates are oxidized at the same site. Conversely, improved ABTS oxidation was

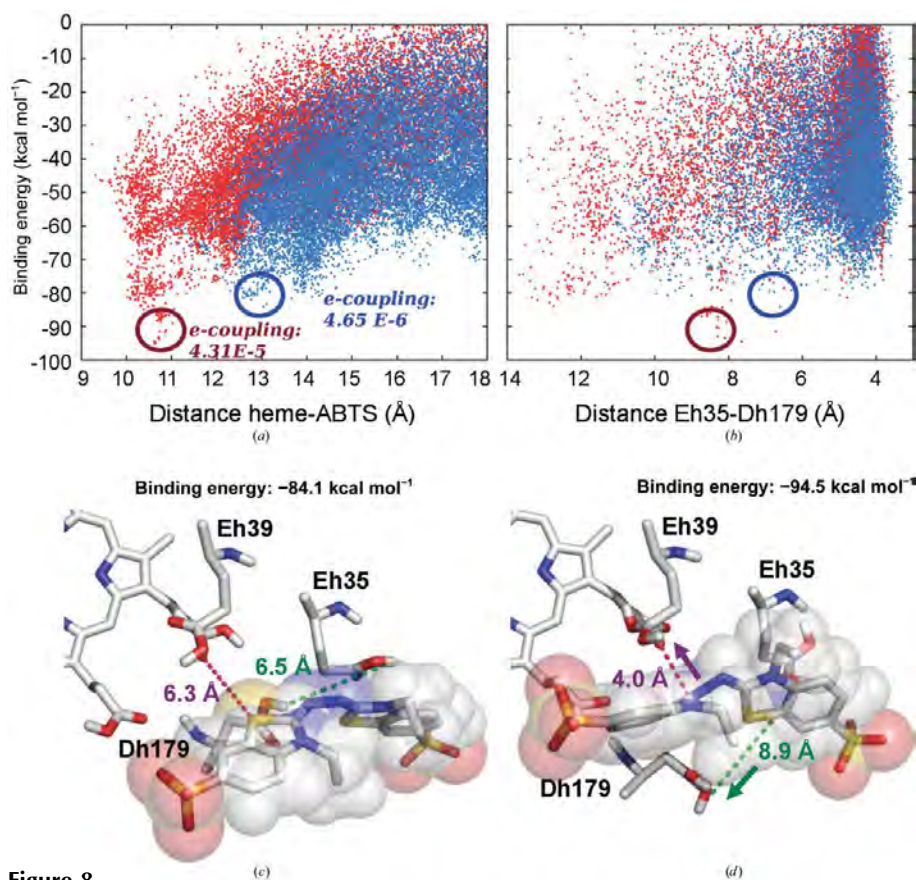


Figure 8

ABTS diffusion by *PELE* in the vicinity of the haem propionates. (*a*, *b*) Interaction energy versus distance between the haem CHA and the ABTS mass centre (*a*) and between protonated Gluh35 CG and AspH179 CD (*b*) during substrate diffusion at the haem-propionate region of *C. subvermispora* extralong MnP6 (blue) and its *in silico* shortened-tail form (red), with the haem–ABTS electron-coupling constants at the most favourable positions indicated (circles). Diffusion simulations were performed with *PELE* (Borrelli *et al.*, 2005), fixing a 20 Å cutoff distance between the haem CHA atom and the ABTS centre of mass. (*c*, *d*) Detail of the ABTS location and the positions of Glu35, Glu39 and Asp179 in two snapshots from the above substrate diffusion on the extralong MnP6 (*c*) and the shortened-tail form (*d*). Residues are shown as CPK-coloured sticks, with ABTS also shown with semitransparent van der Waals spheres. Distances between Dh179 and Eh35, and between ABTS and the most internal propionate of heme are indicated.

obtained by reducing the negative charge at the haem-propionate channel owing to reduced electrostatic repulsion with the anionic ABTS. Interestingly, the oxidation of a substrate other than Mn^{2+} in the haem-propionate channel has been described in ascorbate peroxidase (Mandelman *et al.*, 1998). Mn^{2+} oxidation by MnPs (Sundaramoorthy *et al.*, 2010) and VPs (Ruiz-Dueñas *et al.*, 2007) implies the closure and opening of the haem-propionate channel (by reorientation of the ion-binding side chains) for Mn^{2+} entering near the haem propionate and release of the Mn^{3+} formed, respectively. However, *PELE* simulations show that ABTS oxidation by short MnP requires much more significant reorganizations enabling the approach and oxidation of this bulky substrate.

4.2. Structural–functional evaluation of the variable stability properties of MnPs

The *C. subvermispora* long and especially extralong MnPs are exceptionally stable at acidic pH, while the short MnP is quickly inactivated, in agreement with the results for other short MnPs (Fernández-Fueyo *et al.*, 2014). On the other hand, only slight differences in thermal stability were observed between long and extralong MnPs, contrasting with the higher thermostability reported for the *Dichomitus squalens* extralong MnPs compared with the *P. chrysosporium* long MnPs (Li *et al.*, 2001). In fact, we found the highest thermal stability for the *C. subvermispora* short MnP, with a T_{50} value in the same range as those of the *P. ostreatus* short MnPs (43–57°C; Fernández-Fueyo *et al.*, 2014). Interestingly, removal of the C-tail does not result in improved thermostability; on the contrary, a 5–7°C decrease in T_{50} was observed, revealing that the presence of the tail does not always result in decreased MnP stability. This is explained by the numerous interactions with the protein body observed in the extralong MnP crystal structure, including three hydrophobic/hydrophilic patches. Therefore, more factors than just the tail length affect the thermal stability of MnPs.

Up to three Mn^{2+} or Cd^{2+} ions were identified in the anomalous difference electron-density maps of extralong MnP6–metal complexes. One of them occupied the Mn-oxidation site, in agreement with directed-mutagenesis results, and the others were located at neighbouring positions. Mn^{2+} binding at the oxidation site closes the lower part of the haem-propionate access channel owing to the reorientation of the Glu35 and Glu39 side chains, while the upper part of this channel remains open, being the site for binding of the second metal ion (near Asp85). Finally, the third metal ion in the MnP6– Cd^{2+} complex binds to one of the last residues of the C-tail (Asp363). Similar binding has been described for *P. chrysosporium* MnP (Sundaramoorthy *et al.*, 2005), although the third site does not exist owing to the shorter C-tail. Metal binding would contribute to haem fixation on the protein, resulting in improved enzyme stability, as reported for the *P. chrysosporium* and *D. squalens* MnPs (Mauk *et al.*, 1998; Youngs, Sundaramoorthy *et al.*, 2000) and as confirmed here for both short and long MnPs. Moreover, Cd^{2+} binding at the Mn-oxidation site enabled us to describe ABTS oxidation at

the same site, as shown by competitive inhibition results. On the other hand, MnP stabilization by Ca^{2+} agrees with the general effect found in eukaryotic peroxidases (from the superfamily of plant–fungal–prokaryotic peroxidases), where addition of this cation reverts previous unfolding (for example as caused by alkaline conditions; George *et al.*, 1999; Youngs, Moënne-Loccoz *et al.*, 2000).

4.3. Evolutionary relationships of basidiomycete peroxidases

The relationships between *C. subvermispora* haem peroxidases, a total of 26 proteins including one cytochrome *c* peroxidase, and those from other basidiomycetes, a total of up to 376 sequences, have been analyzed by Fernández-Fueyo, Ruiz-Dueñas, Miki *et al.* (2012). The phylogram obtained (see Supplementary Fig. S5) shows a cluster for the basidiomycete members of the superfamily of plant–fungal–prokaryotic peroxidases (almost 200 sequences). In this cluster, three main groups, corresponding to (i) short-MnPs, VPs and LiPs (with the latter forming a separate subgroup), (ii) long and extralong MnPs and (iii) generic peroxidases, are defined (together with several unclustered sequences). The separate clustering of short MnPs, and the intermixed presence of long and extralong MnPs, reveals a separate origin for the short MnP subfamily, while the long and extralong MnP types would not constitute two separate groups from an evolutionary point of view. The evolutionary relatedness between long and extralong MnPs would be on the basis of the previously discussed biochemical and structural similarities between these two peroxidase types, while the short MnPs would represent a distant branch more related to VPs, with which they share biochemical and structural characteristics.

5. Conclusions

The evolutionary analysis, together with the different biochemical and structural properties, including the presence of a 14–22-residue shorter C-tail in short MnPs, supports their classification as a different subfamily. In contrast, the long and extralong forms cluster together in the peroxidase phylogram, with their catalytic and stability properties showing only slight differences, and present similar structures with conserved C-tails (in spite of their different lengths). These modest differences do not justify maintaining them as two separate subfamilies in genome-annotation and other peroxidase studies. Interestingly, in the above comparison we demonstrate for the first time that substrates other than Mn^{2+} can be oxidized at the Mn-oxidation site of some MnPs. The present work provides a structural basis for the different catalytic and stability properties characterizing different MnPs and the involvement of the C-tail extension in some of them. In this way, we contribute to better structural and functional understanding of the enzymes involved in lignin attack, which is a key issue for the industrial use of plant biomass in ligno-cellulose biorefineries.

This work was supported by Spanish projects BIO2011-26694, CTQ2010-18123, BFU2011-24615 and CSD2009-00088

and the EU projects KBBE-2013-3-613549 and ERC-2009-Adg25027. Sequencing of the *C. subvermispota* genome by JGI was supported by the US DOE under Contract DE-AC02-05CH11231. FJR-D acknowledges a Ramón y Cajal contract. The authors thank the staff of Synchrotron SOLEIL (France), ALBA (Spain) and Swiss Light Source (Switzerland).

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Borrelli, K. W., Vitalis, A., Alcantara, R. & Guallar, V. (2005). *J. Chem. Theory Comput.* **1**, 1304–1311.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Fernández-Fueyo, E., Ruiz-Dueñas, F. J., Ferreira, P. *et al.* (2012). *Proc. Natl Acad. Sci. USA*, **109**, 5458–5463.
- Fernández-Fueyo, E., Ruiz-Dueñas, F. J., Martínez, M. J., Romero, A., Hammel, K. E., Medrano, F. J. & Martínez, A. T. (2014). *Biotechnol. Biofuels*, **7**, 2.
- Fernández-Fueyo, E., Ruiz-Dueñas, F. J., Miki, Y., Martínez, M. J., Hammel, K. E. & Martínez, A. T. (2012). *J. Biol. Chem.* **287**, 16903–16916.
- Floudas, D. *et al.* (2012). *Science*, **336**, 1715–1719.
- George, S. J., Kvaratskhelia, M., Dilworth, M. J. & Thorneley, R. N. F. (1999). *Biochem. J.* **344**, 237–244.
- Giardina, P., Palmieri, G., Fontanella, B., Rivieccio, V. & Sannia, G. (2000). *Arch. Biochem. Biophys.* **376**, 171–179.
- Gold, M. H., Youngs, H. L. & Gelpke, M. D. (2000). *Met. Ions Biol. Syst.* **37**, 559–586.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kirk, T. K. & Farrell, R. L. (1987). *Annu. Rev. Microbiol.* **41**, 465–501.
- Levasseur, A. *et al.* (2014). *BMC Genomics*, **15**, 486.
- Li, D., Youngs, H. L. & Gold, M. H. (2001). *Arch. Biochem. Biophys.* **385**, 348–356.
- Mandelman, D., Jamal, J. & Poulos, T. L. (1998). *Biochemistry*, **37**, 17610–17617.
- Martínez, Á. T., Ruiz-Dueñas, F. J., Martínez, M. J., del Río, J. C. & Gutiérrez, A. (2009). *Curr. Opin. Biotechnol.* **20**, 348–357.
- Martínez, D., Larrondo, L. F., Putnam, N., Gelpke, M. D., Huang, K., Chapman, J., Helfenbein, K. G., Ramaiya, P., Detter, J. C., Larimer, F., Coutinho, P. M., Henrissat, B., Berka, R., Cullen, D. & Rokhsar, D. (2004). *Nature Biotechnol.* **22**, 695–700.
- Mauk, M. R., Kishi, K., Gold, M. H. & Mauk, A. G. (1998). *Biochemistry*, **37**, 6767–6771.
- Morales, M., Mate, M. J., Romero, A., Martínez, M. J., Martínez, Á. T. & Ruiz-Dueñas, F. J. (2012). *J. Biol. Chem.* **287**, 41053–41067.
- Pérez-Boada, M., Ruiz-Dueñas, F. J., Pogni, R., Basosi, R., Choinowski, T., Martínez, M. J., Piontek, K. & Martínez, A. T. (2005). *J. Mol. Biol.* **354**, 385–402.
- Piontek, K., Glumoff, T. & Winterhalter, K. (1993). *FEBS Lett.* **315**, 119–124.
- Poulos, T. L., Edwards, S. L., Wariishi, H. & Gold, M. H. (1993). *J. Biol. Chem.* **268**, 4429–4440.
- Ragauskas, A. J., Williams, C. K., Davison, B. H., Britovsek, G., Cairney, J., Eckert, C. A., Frederick, W. J., Hallett, J. P., Leak, D. J., Liotta, C. L., Mielenz, J. R., Murphy, R., Templer, R. & Tschaplinski, T. (2006). *Science*, **311**, 484–489.
- Riley, R. *et al.* (2014). *Proc. Natl Acad. Sci. USA*, **111**, 9923–9928.
- Ruiz-Dueñas, F. J., Lundell, T., Floudas, D., Nagy, L. G., Barrasa, J. M., Hibbett, D. S. & Martínez, A. T. (2013). *Mycologia*, **105**, 1428–1444.
- Ruiz-Dueñas, F. J. & Martínez, Á. T. (2009). *Microb. Biotechnol.* **2**, 164–177.
- Ruiz-Dueñas, F. J., Morales, M., Pérez-Boada, M., Choinowski, T., Martínez, M. J., Piontek, K. & Martínez, Á. T. (2007). *Biochemistry*, **46**, 66–77.
- Scott, G. M., Akhtar, M., Lenz, M. J. & Swaney, R. E. (1998). *Environmentally Friendly Technologies for the Pulp and Paper Industry*, edited by R. A. Young & M. Akhtar, pp. 341–384. New York: Wiley.
- Steffen, K. T., Hofrichter, M. & Hatakka, A. (2002). *Enzyme Microb. Technol.* **30**, 550–555.
- Sundaramoorthy, M., Gold, M. H. & Poulos, T. L. (2010). *J. Inorg. Biochem.* **104**, 683–690.
- Sundaramoorthy, M., Kishi, K., Gold, M. H. & Poulos, T. L. (1994). *J. Biol. Chem.* **269**, 32759–32767.
- Sundaramoorthy, M., Youngs, H. L., Gold, M. H. & Poulos, T. L. (2005). *Biochemistry*, **44**, 6463–6470.
- Voityuk, A. A. (2012). *Phys. Chem. Chem. Phys.* **14**, 13789–13793.
- Wallrapp, F. H., Voityuk, A. A. & Guallar, V. (2013). *PLoS Comput. Biol.* **9**, e1002990.
- Youngs, H. L., Moënné-Loccoz, P., Loehr, T. M. & Gold, M. H. (2000). *Biochemistry*, **39**, 9994–10000.
- Youngs, H. L., Sundaramoorthy, M. & Gold, M. H. (2000). *Eur. J. Biochem.* **267**, 1761–1769.



Volume 70 (2014)

Supporting information for article:

Structural implications of the C-terminal tail in the catalytic and stability properties of manganese peroxidases from ligninolytic fungi

Elena Fernández-Fueyo, Sandra Acebes, Francisco J. Ruiz-Dueñas, María Jesús Martínez, Antonio Romero, Francisco Javier Medrano, Victor Guallar and Angel T. Martínez

The Supporting Information consists of five supplementary figures showing stereo view of the general structure of extralong MnP6 (Figure S1), multiple alignment of extralong, long and short MnP sequences from different basidiomycetes (Figure S2), mobility of Mn^{2+} -binding residues in the *in silico* shortened-tail form of *C. subvermispora* MnP6 as shown by MD (Figure S3), ABTS binding and C-tail and mobility as shown by PELE (Figure S4), and phylogram of *C. subvermispora* and other basidiomycete peroxidases (Figure S5).

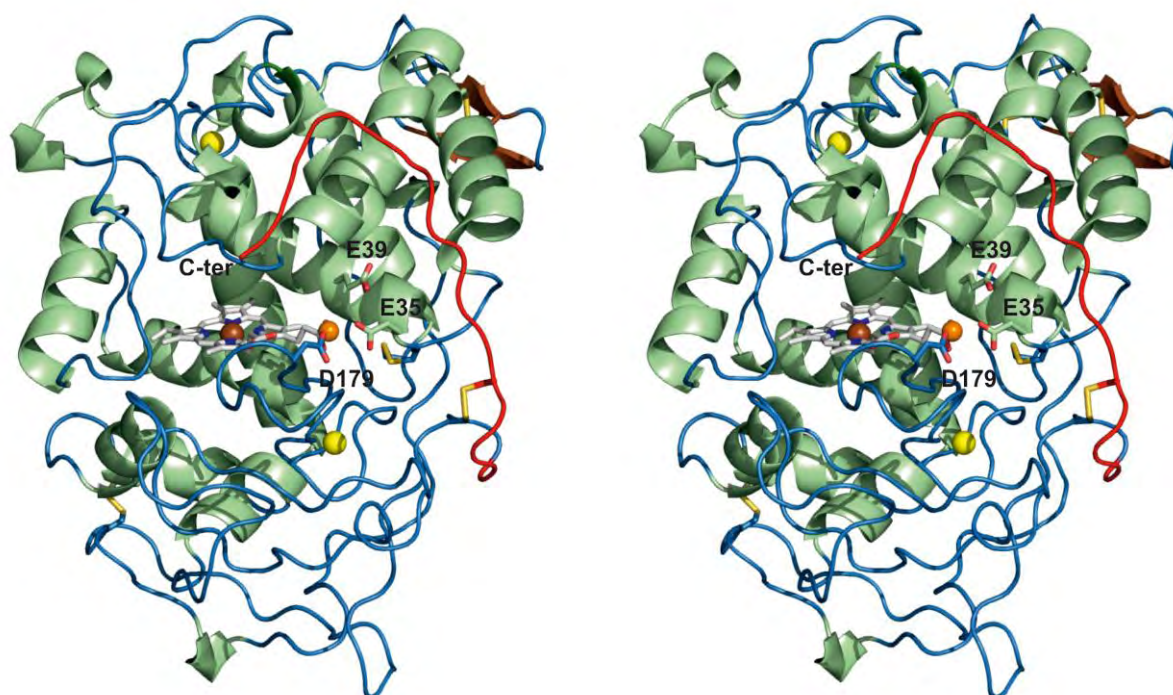


Figure S1. Stereo view of general structure of extralong MnP. A ribbon model of *C. subvermispora* extralong MnP6 (4CZN) is shown with helices in green, short strands in dark red, and C-tail extension (Gly348-Ser364 being absent from short MnP) in red (Pro365 is not included in 4CZN, but it is present in structures including a metal ion). Residues involved in the binding of Mn^{2+} (Glu35, Glu39 and Asp179) and forming disulphide bridges are shown as CPK sticks, Ca^{2+} ions as yellow spheres, and Na^{+} ion at the metal-binding site as orange sphere (see Figure 3A for Mn^{2+} binding).


```

MnP-12 (157986)      --MAFPALLALVALAASVRAAPAS---SSAVCS*SDGT-IVSNAVCCDFIPLAQDLQSMVLQN 31
DICSQ MnP3 (169843) MAFKLLSIVSLVALATVASAAPS----RTVCS*SDGT-VVPDSCCEFLPLAEALQTQVLMG 31
PHLRA MnP2 (CAC85963) MAFNFAAILAFVSLAAVTSAAAPS----KTTCS*NGV-VVPDAVCCDFVPLASALQSEVLMG 31
MnP-6 (50686)       --MSFTALLAIVSLAAIATAAPT----AVCS*SDGT-RVSNAVCCDFVSLGQDLQSMVLQG 31
MnP-10 (117436)    --MAFTSFVALAALVGIASAAPT----TIC*PDGT-RVSNHACCAFIPLAEDLQKTFPMN 31
PHACH MnP1 (66512)  --MAFKSLIAFVALAAVRAAPT----AVCS*PDGT-RVSHAACCAFIPLAQDLQETIFQN 31
MnP-5 (49683)      --MAFTSLLALSALVAVSRAAPT----AVCS*SDGT-RVSNHACCAFIPLAQDLQETIFMN 31
DICSQ MnP6 (169524) --MVFKALLIS-VLAAEQITK GAL-IRRATCS*SDGT-VVANSACCVLIPVIQDIQENLFDG 31
PHLRA MnP3 (Q96TS6) --MAFKQLLT--AIVSVANAAL-TRRAVCPDGVNTATNAVCCSFLFAVRDLIQDQLFDG 32
MnP-13 (124076)   --MAFKPLAALVALLSVSIAHGAI-TRRVTC*PDGVNTVANACCVLFAVRDDIQESLFDG 32
PLEOS MnP4 (1099081) MVNSFHSLLSTIALALLVPSVLA VPAHRAKCSKGR-TASNDACCVVFDVLDLDDIQENLFDG 31
: : : : : : : * . * . . . * * . . : : * : :

MnP-12 (157986)      -DCGD*DAH*IIRLTFH*IAIPIS---RSLG*PSAGG*E*G*MLIFPLVEPEFFASNGIDDSV 87
DICSQ MnP3 (169843) -DCGD*H*IIRLTFH*IAIAIS---RSN-ASAGG*AD*G*MLIFPTVEPAFFANLGIADSV 87
PHLRA MnP2 (CAC85963) -DCGD*DAH*LVRLIFH*IAIAIS---QSMG*PSAGG*AD*G*MLIFPTVEPAFFPNLGIADSV 87
MnP-6 (50686)       -DCGD*DAH*IIRLTFH*IAVAIS---RKLG*PSAGG*AD*G*MLIFPLVEPEFAASNGIDDSV 87
MnP-10 (117436)    -DCGD*DAH*VIRLTFH*IAVAIS---RKLGP*KGAGG*AD*G*MLIFPTVEPNFSANNGIDDSV 87
PHACHMnP1 (66512)  -DCGD*DAH*VIRLTFH*IAIAIS---RSQG*PKAGG*AD*G*MLIFPTVEPNFSANNGIDDSV 87
MnP-5 (49683)      -DCGD*DAH*VIRLTFH*IAVAIS---RSQG*PSAGG*AD*G*MLIFPTVEPNFSANNGIDDSV 87
DICSQ MnP6 (169524) GEGC*GEVH*SLRLTFH*IAIGIS*PA*IAATGVFGGG*AD*G*SIILFEDIEPNFHANNVDEII 91
PHLRA MnP3 (Q96TS6) GEGC*GEVH*SLRLTFH*IAIGIS*P*IASTGVFGGG*AD*G*IAIPAEIETNFHANNVDEII 92
MnP-13 (124076)   GKCG*GEVH*SLRLTFH*IAIGF*SP*ALTAQGF*GGG*AD*G*IAIESIETGYHANLGIDEII 92
PLEOS MnP4 (1099081) GEGC*GEVH*SLRLTFH*IAIGF*SP*ALTRQG*FGG*AD*G*IMLFSDIETNFANNVDDIV 91
. . . . . * * * . . . * * * . . . * * * . . . * * . . . : :

MnP-12 (157986)      NNLI*PFLSSHPTISAGDLVQFAGAVALSNC*PGAPRATFFAGRPNATAPAIDGLIPEPQDN 147
DICSQ MnP3 (169843) NNLI*PFLSQPPTISAGDLVQFAGAVAITNCPGAPQLEFLAGRPNATAPAIDGLIPEPQDN 147
PHLRA MnP2 (CAC85963) NNLI*PFLSQPPTISAGDLVHFAGAVAINSC*PGAPQLEFLAGRPNATAPAIDGLIPEPQDN 147
MnP-6 (50686)       NNLI*PFLSLHPTISAGDLVQFAGAVALSNC*PGAPRVQFLAGRPNHTIAIDGLIPEPQDN 147
MnP-10 (117436)    NNLI*PFMARHPTVSGADLVQFAGAVALSNC*PGAPRLEFLAGRPNHTIAIDGLIPEPQDN 147
PHACHMnP1 (66512)  NNLI*PFMQKHNTISAADLVQFAGAVALSNC*PGAPRLEFLAGRPNKHTIAAVDGLIPEPQDS 147
MnP-5 (49683)      NNLI*PFLAKHP-VSAGDLVQFAGAVALSNC*PGAPQLEFLAGRPNHTIAAVDGLIPEPQDN 146
DICSQ MnP6 (169524) DEQKPI*IAKHN-ITTAADFIQLAGAI*GVSNCPGAPQLN*VFIGR*PDATQ*PAPDKTVPEPDD 150
PHLRA MnP3 (Q96TS6) GEQAPFIQMTN-MTTADFIQFAGAVGVSNCPGAPALPVFVGRPDATQ*PAPDKTVPEPDD 151
MnP-13 (124076)   NEQATFI*IKHN-MTAGDFIQFAGAVGVSNCPGAPQLE*FLLGR*PAATA*PAPDKTVPEPDD 151
PLEOS MnP4 (1099081) EQQKPI*IAKHQ-VSFGDFIQFAGAVGVSNCPAGGPRIQF*LAGRSNVTKSPDHLVPEPDD 150
: : : : : * . . . . . * . . . . . * . . . . . * . . . . . * * * *

MnP-12 (157986)      VTSILAR---FDDAGG*P*PFEVVSLLASH*IAARA*KVDPTLDAAPF*ST*P*F*FTQVFL 203
DICSQ MnP3 (169843) ITKILAR---FDDAGG*P*PFEVVSLLASH*IAARA*KVDPTLDAAPF*ST*P*F*FTQVFL 203
PHLRA MnP2 (CAC85963) VTKILAR---FKDAGN*F*SPAEV*VALLASH*IAARA*HVDPTLDAAPF*ST*P*F*FTQVFL 203
MnP-6 (50686)       VTSILAR---FDDAGG*P*PFEVVSLLASH*IAARA*KVDPTLDAAPF*ST*P*F*FTQVFL 203
MnP-10 (117436)    VTKILAR---FDDAGG*P*PFEVVSLLASH*IAARA*KVDETI*DAAPF*ST*P*F*FTQVFL 203
PHACHMnP1 (66512)  VTKILAR---FDDAGG*P*PFEVVSLLASH*IAARA*KVDPTLDAAPF*ST*P*F*FTQVFL 203
MnP-5 (49683)      VTKILAR---FDDAGG*P*PFEVVSLLASH*IAARA*KVDETI*DAAPF*ST*P*F*FTQVFL 202
DICSQ MnP6 (169524) VDSILAR*QDA*FSD*VGG*P*PAE*V*V*LLASH*IAAAA*HVDPSI*PGT*P*F*ST*P*F*FTQVFFI 207
PHLRA MnP3 (Q96TS6) VDSILAR---FADAG*G*P*SSAE*V*V*LLASH*IAAAA*HVDPSI*PGT*P*F*ST*P*F*FTQVFFI 207
MnP-13 (124076)   VDSILAR---FGDAG-F*SPQ*E*V*IAL*LSH*IAAAA*HVDPAI*PGT*P*F*ST*P*F*PQIFI 206
PLEOS MnP4 (1099081) VTSILAR---MGDAG-F*PKD*E*V*V*LLASH*IAAAG*HVDPAI*PGT*P*F*ST*P*F*SQVFF 199
: . * * * : * * * . * * . * * . * * * * * . * * : . . * * * * * * * * . :

MnP-12 (157986)      ETLLKGVGFP*GTD*DN*V*GEVAS*PLP*L*G*D*T*S*G*G*N*D*G*M*MLR*Q*S*D*F*L*A*R*D*E*R*T*A*C*F*W*Q*S*F*V 263
DICSQ MnP3 (169843) DVLLKGVGFP*G*LN*NT*G*E*V*S*S*P*L*P*L*S*N*---G*TD*V*G*E*L*R*L*Q*S*D*F*L*A*H*D*P*R*T*A*C*F*W*Q*G*F*V 259
PHLRA MnP2 (CAC85963) EVLLKGVGFP*G*LAN*NT*G*E*V*S*S*P*L*P*V*D*---G*TD*V*G*E*L*R*L*Q*S*D*F*L*A*H*D*P*R*T*A*C*F*W*Q*G*F*V 259
MnP-6 (50686)       EVLLKGVGFP*G*LN*NT*G*E*V*S*S*P*L*P*L*G*D*T*S*G*G*K*D*T*G*M*MLR*Q*S*D*F*L*A*H*D*P*R*T*A*C*F*W*Q*G*F*V 263
MnP-10 (117436)    EVLLKGVGFP*G*TD*DN*NT*G*E*V*S*P*L*P*L*K*G*S*---G*N*D*T*G*E*M*R*L*Q*S*D*F*L*A*H*D*P*R*T*A*C*F*W*Q*G*F*V 259
PHACHMnP1 (66512)  EVLLKGVGFP*G*SAN*NT*G*E*V*S*P*L*P*L*G*S*---G*S*D*T*G*E*M*R*L*Q*S*D*F*L*A*H*D*P*R*T*A*C*F*W*Q*G*F*V 259
MnP-5 (49683)      EVLLKGVGFP*G*TD*DN*NT*G*E*V*S*P*L*P*L*S*T*S*---G*N*D*T*G*E*M*R*L*Q*S*D*F*L*A*H*D*P*R*T*A*C*F*W*Q*G*F*V 258
DICSQ MnP6 (169524) ETQLR*GT*L*FP*GT*G*GN*Q*G*E*V*S*P*L*H*---G*E*I*R*L*Q*S*D*S*E*L*A*R*D*S*R*T*A*C*E*W*Q*S*F*V 259
PHLRA MnP3 (Q96TS6) ETQLR*GT*L*FP*GT*G*GN*Q*G*E*V*S*P*L*H*---G*E*I*R*L*Q*S*D*S*E*L*A*R*D*S*R*T*A*C*E*W*Q*S*F*V 256
MnP-13 (124076)   EVQLR*GT*L*FP*GT*G*GN*Q*G*E*V*S*P*P*---G*E*I*R*L*Q*S*D*H*N*L*A*R*D*S*R*T*A*C*F*W*Q*S*L*A 255
PLEOS MnP4 (1099081) ETLLKGT*L*P*G*D*S*L*H*K*Q*V*K*S*P*L*P*---G*E*F*R*L*Q*S*D*E*L*L*A*R*D*S*R*T*S*E*W*Q*S*F*I 254
: . * * * : * * * . * * . * * . * * * * * . * * : . . * * * * * * * * . :

MnP-12 (157986)      NQQLMAESFKA*AFFK*LA*V*L*G*H*N*Q*AD*L*V*D*CS*E*V*V*P*I*P*V*P*G*D*G*K*P*AT*F*P*AT*G*P*Q*D*L*Q*L*T* 323
DICSQ MnP3 (169843) NEQE*FMA*Q*S*FKA*AMAK*LA*V*L*G*H*N*AD*L*V*D*CS*AV*V*P*K*P*LA*V*AV*AA*F*P*AT*K*G*P*D*DL*E*L*S*C 319
PHLRA MnP2 (CAC85963) NEQE*AM*AT*AFK*NA*V*K*L*AV*L*G*H*N*R*ND*L*V*D*CS*AV*V*P*K*P*AT*G*P*AT*F*P*AT*G*P*Q*D*DL*E*L*T*C 319
MnP-6 (50686)       DQQE*FMS*Q*S*FAS*AFK*LA*V*L*G*H*N*DD*L*D*CS*E*V*V*P*K*P*AV*D*K*P*AT*F*P*AT*G*P*Q*D*DL*E*L*S*C 323
MnP-10 (117436)    DEQE*FMA*ES*FKA*AMAK*LA*V*L*G*H*N*AS*L*TD*CS*DV*V*P*I*P*P*V*K*P*AS*F*P*AT*G*P*K*D*DL*E*L*T*C 319
PHACHMnP1 (66512)  NEQA*FMA*AS*FRA*AMS*KL*AV*L*G*H*N*RS*L*ID*CS*DV*V*P*K*P*AT*G*P*AM*F*P*AT*G*P*Q*D*DL*E*L*S*C 319
MnP-5 (49683)      NEQE*FMA*Q*S*FKA*AMS*KL*AV*L*G*H*S*RS*DL*ID*CS*DV*V*P*P*P*AP*AS*DA*H*F*P*AG*Q*TD*DD*V*EQ*AC 318
DICSQ MnP6 (169524) NNQAKLQ*SA*FKA*FR*RS*IL*GH*DE*SS*L*ID*CS*DV*V*P*P*P*AP*AS*DA*H*F*P*AG*Q*TD*DD*V*EQ*AC 318
PHLRA MnP3 (Q96TS6) NNQAKTQ*SA*FKA*FR*K*MT*IL*GH*SE*SS*L*IE*CS*E*V*I*Q*TP*-PA*LE*G*NA*H*L*P*AG*Q*TM*ND*IE*Q*AC 315
MnP-13 (124076)   NNQQM*QA*Q*FKA*AMAK*LA*V*L*G*Q*D*V*S*Q*M*V*D*CS*DV*I*P*V*P*-LP*P*AT*P*P*H*I*P*AG*LS*Q*NN*IE*Q*AC 314
PLEOS MnP4 (1099081) SNPNSMV*PK*FER*AMAK*MT*L*Q*N*PK*L*ID*CS*E*V*I*P*V*P*-R*G*RV*K*Q*P*L*P*AG*K*I*K*IE*AS*C 313
: . : * * . : : * * . : : * * * : * * * * * . * * : : : * * : :

MnP-12 (157986)      TAERFP*TL*SD*V*P*G*AT*E*TL*V*P*H*CS*DG*G*EN*CS*P*V*Q*F*AG*P*AT*G*F*NG*TD*-- 368
DICSQ MnP3 (169843) NTSRFP*NL*P*ID*H*GT*Q*E*AL*I*P*H*CS*DG*MS*CT*TV*Q*FD*G*P*ALS*F*GD*N*NS* 366
PHLRA MnP2 (CAC85963) TTEFP*TL*ST*P*GA*Q*Q*TL*I*P*H*CS*DG*TM*TS*NS*V*Q*FD*G*P*AT*F*NG*G*AD*DS 366
MnP-6 (50686)       LAERFP*TL*SD*V*P*GA*Q*ET*L*I*P*H*CS*DG*LE*NC*TS*V*Q*F*SG*P*AT*DS*P*---- 365
MnP-10 (117436)    RAERFP*TL*TV*DR*G*AV*Q*AL*I*P*H*CS*NG*GD*CS*P*V*Q*FD*G*P*A*----- 357
PHACHMnP1 (66512)  PSERF*P*TL*TV*Q*P*AS*Q*SL*IA*H*PD*G*MS*CS*P*V*Q*F*NG*P*A*----- 357
MnP-5 (49683)      FAERFP*TL*P*V*TP*G*AT*Q*TL*I*P*H*CS*NG*GED*CT*V*Q*F*T*G*P*A*----- 356
DICSQ MnP6 (169524) ASTFP*P*TL*P*DP*G*P*AS*V*AP*V*P*S*----- 342
PHLRA MnP3 (Q96TS6) ATTP*F*P*SL*SD*P*G*P*AT*S*V*AP*V*P*S*----- 339
MnP-13 (124076)   ATA*F*P*TL*P*ID*V*G*P*Q*TS*I*P*V*P*V*G*S*----- 338
PLEOS MnP4 (1099081) RKAF*F*P*RL*P*DK*GT*F*TS*IL*P*V*P*S*----- 337
* * * . * . : : .

```

Figure S2. Multiple alignment of extralong, long and short MnPs sequences. *C. subvermisporea* extralong (MnP6/MnP12), long (MnP5/MnP10) and short (MnP13) MnPs, extralong MnPs from *Phlebia radiata* (GenBank CAC85963) and *D. squalens* (JGI 169843), long MnP from *P. chrysosporium* (JGI 66512), and short MnPs from *P. radiata* (GenBank Q96TS6), *D. squalens* (JGI 169526) and *P. ostreatus* (JGI 1099081) are included. Extra residues in long and extralong MnPs are in magenta and red, respectively. Conserved residues include: **i**) eight cysteines (cyan); **ii**) nine Ca^{2+} ligands (green); **iii**) two active site histidines (dark gray); and **iv**) three acidic residues (red) forming the Mn^{2+} -oxidation site. Numbering starts at the first residue of the mature protein (red asterisk). Symbols below indicate full conservation of the same (*) or equivalent residues (:) and partial residue conservation (.) .

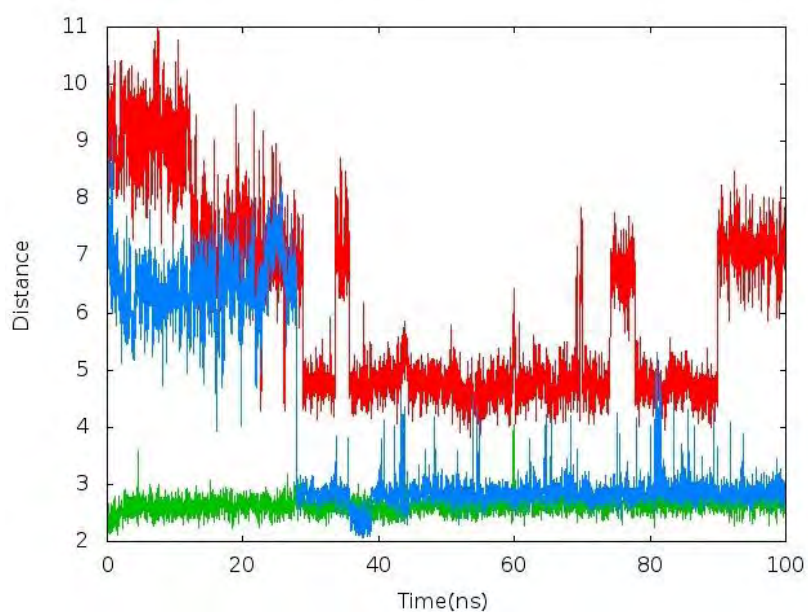


Figure S3. Mobility of Mn^{2+} -binding residues in the MnP6 *in silico* shortened-tail form as shown by MD. Distances (in Å) between heme-CG and protonated E35-HE2 (green), E39-HE2 (blue) and D179-HD2 (red) during 100 ns MD simulation (300 K) on the shortened-tail form of *C. subvermispora* MnP6.

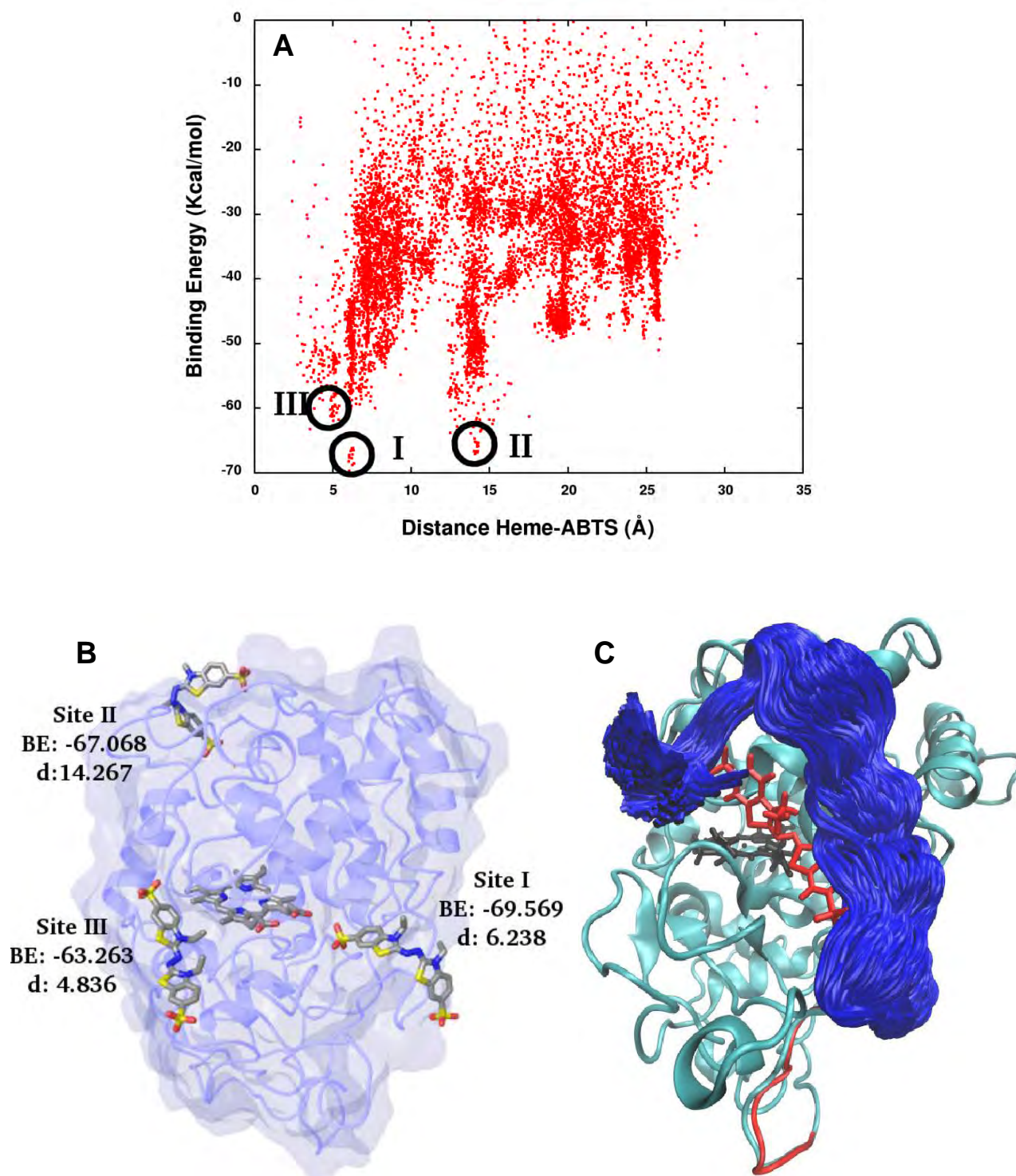


Figure S4. PELE molecular simulations on *C. subvermispora* MnP6. (A, B) Preliminary exploration on ABTS diffusion showing three minima in heme-ABTS distance vs binding energy (A), and location of the three sites at the entrance of the heme-propionate channel at 6.2 Å from the heme (I), near the distal Ca²⁺ binding site at 14.3 Å from the heme (II), and the main heme access channel at only 4.8 Å but with the worst binding energy (III). (C) Mobility of the extralong tail preventing ABTS approach, as shown by superimposition of extralong MnP6 (blue ribbon, and heme) and one ABTS position (orange sticks) selected from the PELE diffusion on the *in silico* shortened-tail form, the latter being in collision with the extralong tail whose mobility is indicated in dark blue (the end of the shortened-tail variant is in orange).

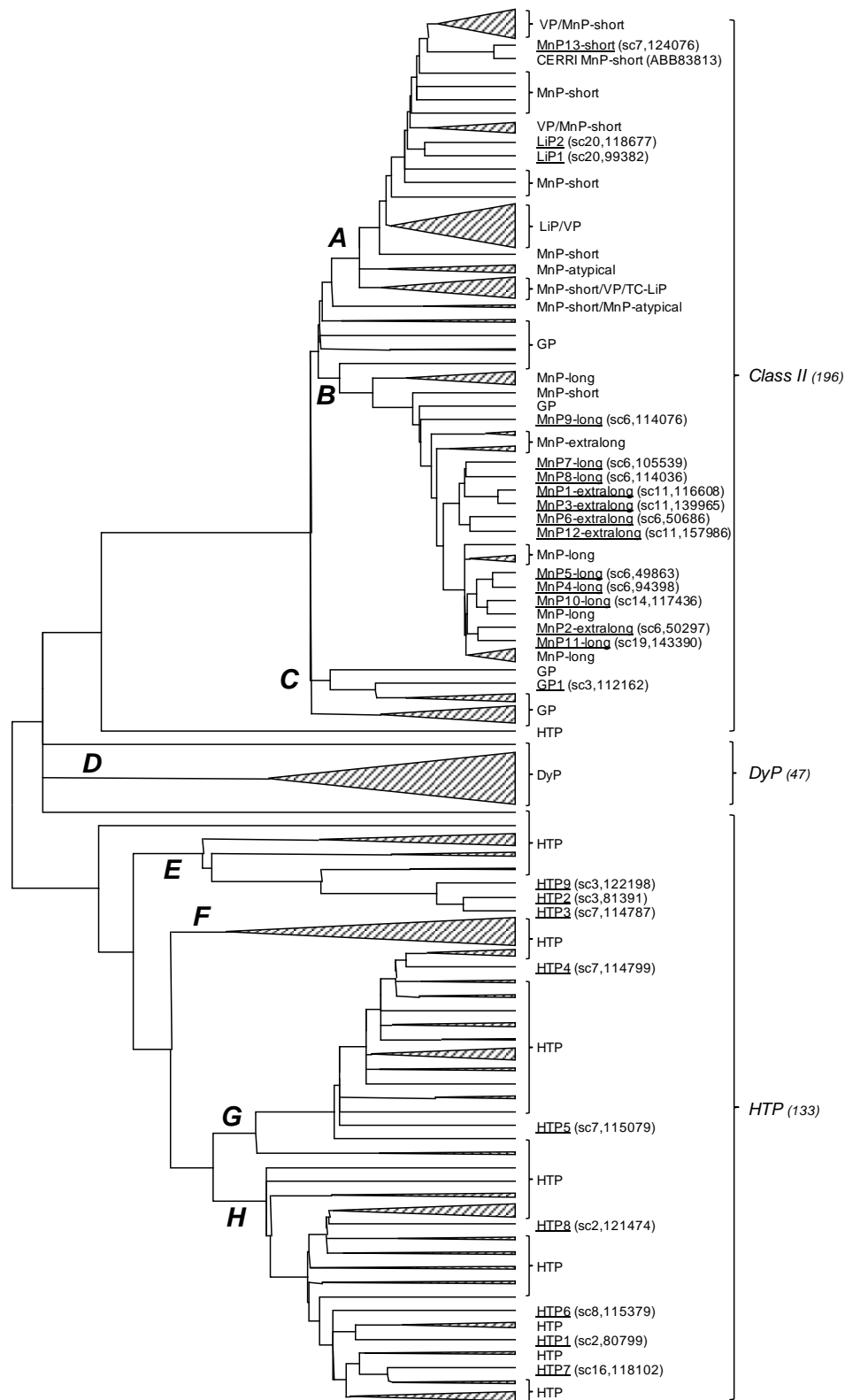


Figure S5. Phylogram of basidiomycete peroxidases (GenBank and genomes). The position of 25 sequences from *C. subvermispora* genome is shown (underlined) including one short, seven long and five extralong MnPs, two LiPs and one generic peroxidase (all in Class II of plant-fungal-prokaryotic peroxidase superfamily), and nine members of the heme-thiolate peroxidase (HTP) superfamily. Clusters where *C. subvermispora* sequences are not included were collapsed, and prokaryotic peroxidases are not included. The three main peroxidase groups, Class II, HTP and dye-decolorizing peroxidases (DyP), and the number of sequences in each of them (parentheses) are indicated, together with the scaffold (sc) and JGI references of the *C. subvermispora* gene models (parentheses). Additionally, MnP-atypical corresponds to a MnP with only two acidic residues at the oxidation site; TC-LiP corresponds to an unusual LiP from *T. cervina*; and CERRI MnP-short corresponds to a *Ceriporiopsis rivulosa* MnP (GenBank BBB83813) clustering together with *C. subvermispora* short MnP13. Adapted from Fernandez-Fueyo et al. (2012).