



BIOLOGICAL  
CRYSTALLOGRAPHY

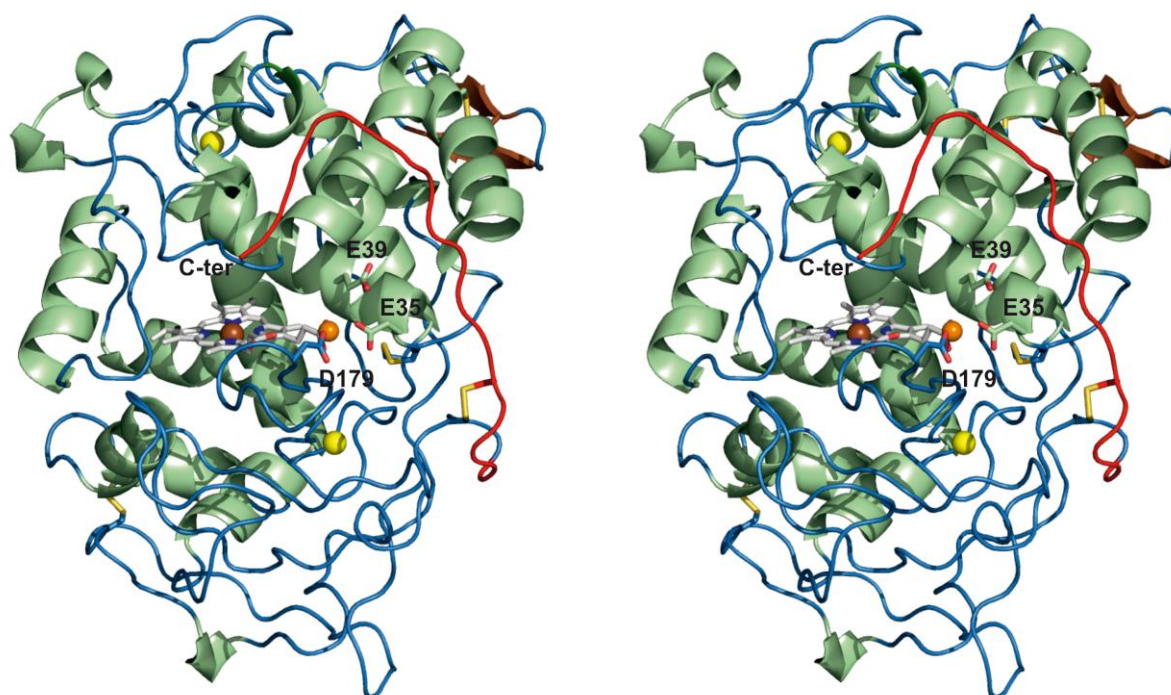
**Volume 70 (2014)**

**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<sup>2+</sup>-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).



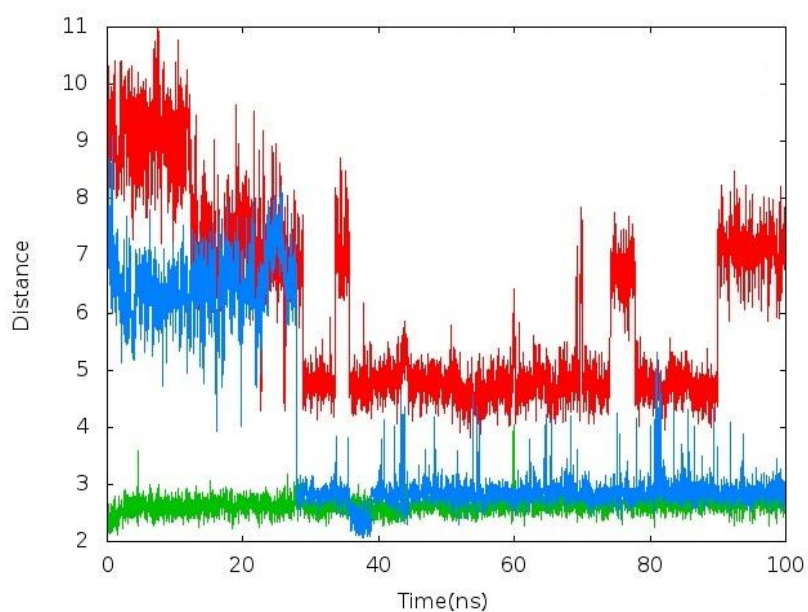
**Figure S1.** Stereo view of general structure of extralong MnP. A ribbon model of *C. subvermispora* extralong MnP6 (4CZN) is shown with helices in green, short strands in dark red, and **C-tail extension (Gly348-Ser364 being absent from short MnP) in red** (Pro365 is not included in 4CZN, but it is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

```

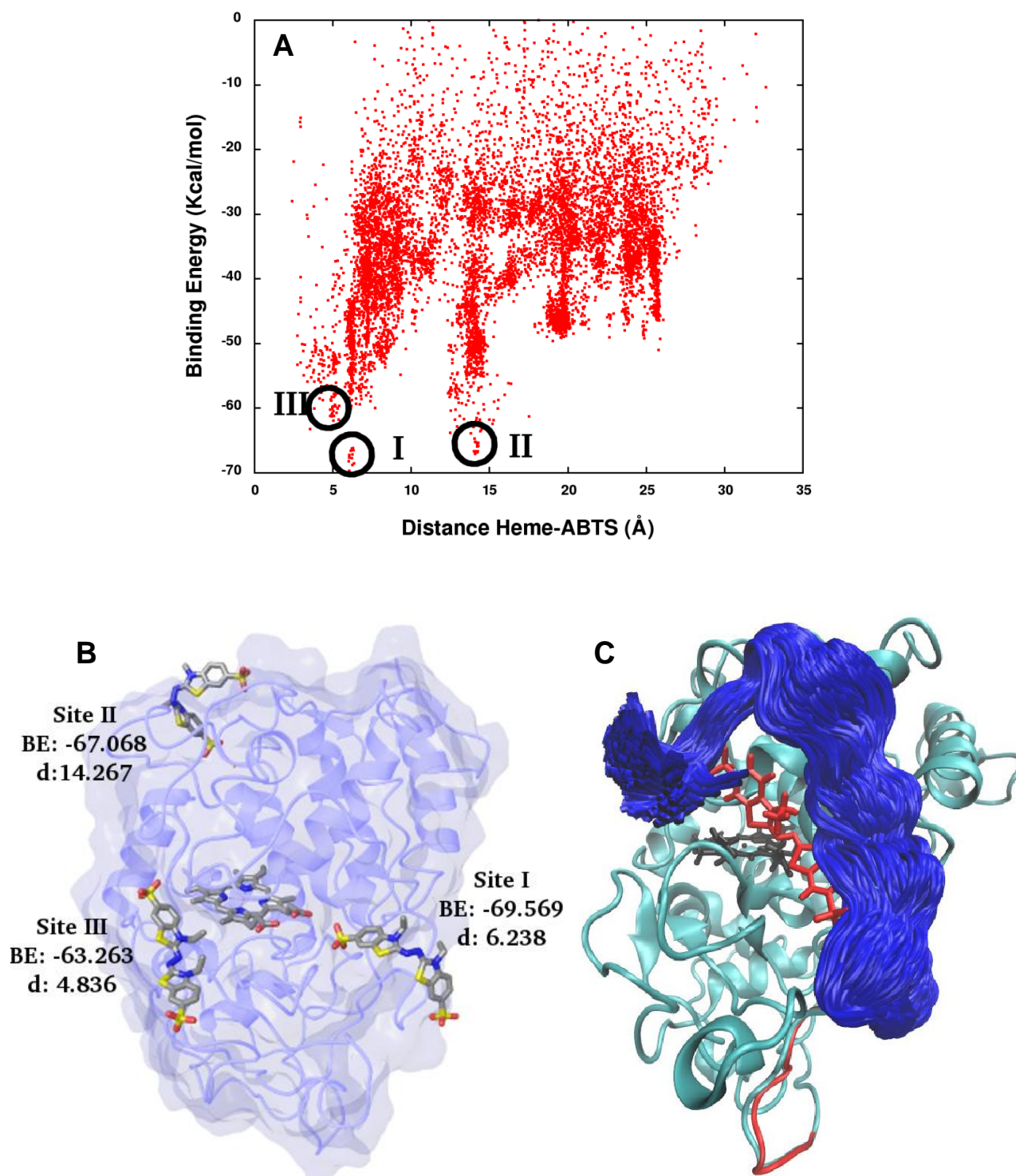
MnP-12 (157986)      --MAFPALLALVALAASVRAAPAS---SSAVCS*SDGT-IVSNAVCCDFIPLAQDLQSMVLQN 31
DICSQ MnP3 (169843)  MAFKLLSIVSLVALATVASAAPS----RTVCS*SDGT-VVPDSCCEFLPLAEALQTQVLMG 31
PHLRA MnP2 (CAC85963) MAFNFAAILAFVSLAAVTSAAAPS----KTTC*SNV-VVPDAVCCDFVPLASALQSEVLMG 31
MnP-6 (50686)       --MSFATLLAIVSLAAIATAAPT----AVCS*SDGT-RVSNVAVCCDFVSLGQDLQSMVLQG 31
MnP-10 (117436)     --MAFTSFVALAALVGIASAAPT----TIC*PDGT-RVSNHACC*AFIPLAEDLQKTFPMN 31
PHACH MnP1 (66512)   --MAFKSLIAFVALAAVRAAPT----AVCPDGT-RVSHAA*CC*AFIPLAQDLQETIFQN 31
MnP-5 (49683)       --MAFTSLLALSALVAVSRAPT----AVCS*SDGT-RVSNACC*AFIPLAQDLQETLFMN 31
DICSQ MnP6 (169524) --MVFKALLIS-VLAAFOITK*GAL-IRRATCS*SDGT-VVANSACC*VLPVPIQDIQENLFDG 31
PHLRA MnP3 (Q96TS6)  --MAFKQLLT--AISIVSANAAL-TRRVA*CPDGVNTATNAVCCSLFAVRDLIQDLQDFD 32
MnP-13 (124076)     --MAFKPLAALVALLSVSIAHGAI-TRRVTC*PDGVNTVANACC*PLFAVRDDIQESLFDG 32
PLEOS MnP4 (1099081) MVNSFHSLLSTIALALLVPSVLAVPAHRAK*SKGR-TASNDACC*VWFVLDLDDIQENLFDG 31
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
MnP-12 (157986)     -DCGDAH*IIRLTFHDAI*PIS---RSLGPSAGG*GAG*MLIFPLVEPEFFASNGIDDSV 87
DICSQ MnP3 (169843)  -DCG*DAH*IIRLTFHDAI*AIS---RSN-ASAGG*GAG*MLIFPTVEPAFFANLGIADSV 87
PHLRA MnP2 (CAC85963) -DCGDAH*IIRLTFHDAI*AIS---QSMG*PSAGG*GAG*MLIFPTVEPAFFPNLGIADSV 87
MnP-6 (50686)       -DCGDAH*IIRLTFHDAVAIS---RKLGPSAGG*GAG*MLIFPLVEPEFAASNGIDDSV 87
MnP-10 (117436)     -DCGDAH*IIRLTFHDAVAIS---RKLGP*KGAG*GAG*MLIFPTVEPNFSA*NGIDDSV 87
PHACHMnP1 (66512)   -ECCGDAH*IIRLTFHDAI*AIS---RSQG*PKAGG*GAG*MLIFPTVEPNFSA*NGIDDSV 87
MnP-5 (49683)       -DCGDAH*IIRLTFHDAVAIS---RSQG*PSAGG*GAG*MLIFPTVEPNFSA*NGIDDSV 87
DICSQ MnP6 (169524)  GEC*GSEVH*SLRLTFHDAIGIS*PAI*AA*TVG*FVGG*GAG*SI*ILFEDIE*PNFHANN*GVDE*II 92
PHLRA MnP3 (Q96TS6)  GEC*GSEVH*SLRLTFHDAIGIS*PTI*AA*TVG*FVGG*GAG*SI*IA*FAE*IE*TN*FHANN*GVDE*II 92
MnP-13 (124076)     GK*GSEVH*SLRLTFHDAIGF*SP*AL*TA*Q*G*F*V*G*G*G*G*G*SI*IA*F*E*I*E*TG*Y*HAN*L*G*I*D*E*II 92
PLEOS MnP4 (1099081) GEC*GSEVH*SLRLTFHDAIGF*SP*AL*TR*Q*G*F*V*G*G*G*G*G*SI*IM*LF*SD*IE*TN*FA*ANN*GV*DD*IV 91
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
MnP-12 (157986)     NNLI*PFL*SSH*PTISAGDLVQFAGAVALS*NC*PGAPRATFFAGRPNATA*PAIDGLI*PEPQDN 147
DICSQ MnP3 (169843)  NNLI*PFLSQ*PPTISAGDLVQFAGAVAI*TN*CPGAPQLEFLAGRPNATA*PAIDGLI*PEPQDN 147
PHLRA MnP2 (CAC85963) NNLI*PFLSQ*PPTISAGDLVHFAGAVAI*SN*CPGAPQLEFLAGRPNATA*PAIDGLI*PEPQDD 147
MnP-6 (50686)       NNLI*PFLSLH*PTISAGDLVQFAGAVALS*NC*PGAPRVQFLAGRPNH*TAIDGLI*PEPQDN 147
MnP-10 (117436)     NNLI*PFMAR*HPTV*SGADLVQFAGAVALS*NC*PGAPRLEFLAGRPNH*TAIDGLI*PEPQDD 147
PHACHMnP1 (66512)   NNLI*PFMQK*HNTISAADLVQFAGAVALS*NC*PGAPRLEFLAGRPNK*TAIDVGLI*PEPQDS 147
MnP-5 (49683)       NNLI*PFLAK*HP-V*SGADLVQFAGAI*AL*TN*CPGAPQLEFLAGRPNH*TAIDVGLI*PEPQDD 146
DICSQ MnP6 (169524)  DEQK*PI*IAKHN-I*TTAD*FI*QLAGAI*GV*SN*CPGAPQLN*V*F*GR*PDAT*Q*PAPDK*TV*PEP*DT 150
PHLRA MnP3 (Q96TS6)  GEQAP*FI*Q*MTN-M*TTAD*FI*Q*FAGAV*GV*SN*CPGAPAL*P*V*F*GR*PDAT*Q*PAPDK*TV*PEP*DT 151
MnP-13 (124076)     NEQAT*FI*LKH*N-M*TAGD*FI*Q*FAGAV*GV*SN*CPGAPQLE*F*LL*GR*PAAT*PAPDK*TV*PEP*DT 151
PLEOS MnP4 (1099081) EQQK*PIA*IKHQ-V*SF*GD*FI*Q*FAGAV*GV*SN*CGG*PR*IQ*F*LAGR*SN*V*TK*SP*DH*LV*PEP*DS 150
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
MnP-12 (157986)     VTSILAR----FDDAGG*F*TP*FEV*V*SL*LA*SH*IA*RA*KV*DP*TL*DA*AP*F*TS*PF*FT*Q*F*FL 203
DICSQ MnP3 (169843)  ITKILAR----FDDAGG*F*TP*FEV*V*SL*LA*SH*IA*RA*KV*DP*TL*DA*AP*F*TS*PF*FT*Q*F*FL 203
PHLRA MnP2 (CAC85963) VTKILAR----FKDAGN*F*SP*AE*V*V*ALL*SH*IA*RA*HV*DP*TL*DA*AP*F*TS*PF*FT*Q*F*FL 203
MnP-6 (50686)       VTSILER----FDDAGG*F*TP*FEV*V*SL*LA*SH*IA*RA*KV*DP*TL*DA*AP*F*TS*PF*FT*Q*F*FL 203
MnP-10 (117436)     VTKILER----FDDAGG*F*TP*FEV*V*SL*LA*SH*IA*RA*KV*DE*TI*DA*AP*F*TS*PF*FT*Q*F*FL 203
PHACHMnP1 (66512)   VTKILQR----FEDAGG*F*TP*FEV*V*SL*LA*SH*IA*RA*KV*Q*TI*DA*AP*F*TS*PF*FT*Q*F*FL 203
MnP-5 (49683)       VTKILAR----FDDAGG*F*SP*FEV*V*SL*LA*SH*IA*RA*KV*DE*TI*DA*AP*F*TS*PF*FT*Q*F*FL 202
DICSQ MnP6 (169524)  VDSILAR*QD*AF*SD*V*G*G*F*TP*AE*V*V*ALL*SH*IA*AA*HV*DP*SI*PG*TP*F*TS*PE*FT*Q*F*FI 207
PHLRA MnP3 (Q96TS6)  VDSILAR----FADAG*G*F*SS*AE*V*V*ALL*SH*IA*AA*HV*DP*SI*PG*TP*F*TS*PE*FT*Q*F*FI 207
MnP-13 (124076)     VDSILAR----FGDAG-F*SP*Q*E*V*IA*LL*SH*IA*AA*HV*DP*AI*PG*TP*F*TS*PE*FT*Q*F*FI 206
PLEOS MnP4 (1099081) VTSILAR----MGDAG-F*PK*DE*V*V*ALL*SH*IA*AA*TI*DP*KL*AG*H*P*F*TS*PE*FT*Q*F*FI 199
: : * * * : : * * * * : : * * * * : : * * * * : : * * * * : : * * * * : : * * * * : : * * * * : : * * * * : :
MnP-12 (157986)     ETLLKGVG*F*PGT*DDNVGEVAS*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)  DVLLKGVG*F*PGL*NNNTGEV*SS*PLP*L*SN----G*TDV*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) EVLLKGVG*F*PGL*ANN*TEGV*SS*PLP*V*TD----G*TDV*G*E*L*R*L*Q*S*D*F*L*A*R*D*E*R*T*A*C*F*W*Q*G*F*V 259
MnP-6 (50686)       EVLLKGVG*F*PGL*DNNTGEV*SS*PLP*L*G*D*T*S*G*G*K*D*G*L*M*L*R*Q*S*D*F*L*A*H*D*P*R*T*A*C*F*W*Q*G*F*V 263
MnP-10 (117436)     EVLLKGVG*F*PGL*DNNTGEV*SS*PLP*L*K*G*S----G*NDT*G*E*M*L*R*Q*S*D*F*L*A*H*D*P*R*T*A*C*F*W*Q*G*F*V 259
PHACHMnP1 (66512)   EVLLKGVG*F*PGL*ANN*TEGV*SS*PLP*L*G*S----G*SDT*G*E*M*L*R*Q*S*D*F*L*A*H*D*P*R*T*A*C*F*W*Q*G*F*V 259
MnP-5 (49683)       EVLLKGVG*F*PGL*DNNTGEV*SS*PLP*L*TS----G*NDT*G*E*M*L*R*Q*S*D*F*L*A*H*D*P*R*T*A*C*F*W*Q*G*F*V 258
DICSQ MnP6 (169524)  ETQLR*G*TL*F*P*G*T*G*G*N*Q*G*E*V*E*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*G*TL*F*P*G*T*G*G*N*Q*G*E*V*E*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*G*TL*F*P*G*T*G*G*N*Q*G*E*V*E*S*P*Y-----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*LI*P*G*D*SL*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)     NQQLMAES*FKA*AFF*KL*SL*LG*HN*QAD*LVDC*SEV*V*P*V*P*G*DG*KP*AT*F*P*AT*G*P*QD*L*Q*L*T* 323
DICSQ MnP3 (169843)  NEQE*FMA*Q*SFKA*AMAKL*AVL*GH*ND*LDVDC*SAV*V*P*PK*P*LA*V*AV*SA*AF*P*AT*KG*PD*DL*E*L*S*C 319
PHLRA MnP2 (CAC85963) NEQE*AM*ATA*F*KN*AV*KKL*AVL*GH*NR*ND*LDVDC*SAV*V*P*PK*P*AT*G*P*AT*F*P*AT*G*P*QD*LE*L*T*C 319
MnP-6 (50686)       DQDE*FMS*Q*SFAS*AF*KL*AVL*GH*NT*DD*LDVDC*SEV*V*P*PK*P*AV*DK*P*AT*F*P*AT*G*P*QD*LE*L*S*C 323
MnP-10 (117436)     DEQE*FMA*ES*FKA*AMAKL*AI*LG*HN*RA*SL*TD*CS*DV*V*P*IP*P*AV*KK*P*AS*F*P*AT*G*P*K*D*LE*L*T*C 319
PHACHMnP1 (66512)   NEQE*FMA*AS*FRA*AMSKL*AVL*GH*NR*SL*LDVDC*SD*V*V*P*PK*P*AT*G*P*AM*F*P*AT*G*P*QD*LE*L*S*C 319
MnP-5 (49683)       NEQE*FMA*Q*SFKA*AMSKL*AVL*GH*SR*SD*LDVDC*SD*V*P*TP*P*AV*NK*P*AT*F*P*AT*G*P*K*D*LE*L*S*C 318
DICSQ MnP6 (169524)  NNQAKLQ*SA*FKA*FR*RR*SL*GH*DE*SS*LDVDC*SD*V*V*P*P*P*APAS*DA*HP*P*AG*Q*TD*DD*VE*Q*AC 318
PHLRA MnP3 (Q96TS6)  NNQAKTQ*SA*FKA*FR*RM*TL*GH*SE*SS*LI*E*C*SE*V*I*Q*TP--PALE*G*NA*HL*P*AG*Q*TM*ND*IE*Q*AC 315
MnP-13 (124076)     NNQQQ*MQA*Q*FKA*AMAKL*AVL*G*Q*D*V*S*Q*M*V*DC*SD*V*P*V*P--LP*P*AT*P*H*I*P*AGL*SN*NI*E*Q*AC 314
PLEOS MnP4 (1099081) SNPN*SM*V*PK*FER*AMAK*MATL*G*Q*N*PK*KL*ID*C*SE*V*I*P*V*P--R*GR*V*Q*P*TL*P*AG*K*TK*ID*E*AS*C 313
: : : * * * : : * * * * : : * * * * : : * * * * : : * * * * : : * * * * : : * * * * : : * * * * : : * * * * : :
MnP-12 (157986)     TAER*F*P*TL*SV*DP*G*AT*E*TL*V*PH*CS*DG*G*EN*CS*V*Q*F*AG*P*AT*G*F*NG*TD-- 368
DICSQ MnP3 (169843)  NTSR*F*P*NL*P*ID*H*GT*Q*E*AL*P*HC*SD*G*SM*S*CT*V*V*Q*F*DG*P*ALS*F*GD*N*NS* 366
PHLRA MnP2 (CAC85963) TTE*F*P*TL*ST*AP*GA*Q*TL*P*HC*SD*G*TM*TS*NS*V*Q*F*DG*P*AT*F*NG*G*AD*DS 366
MnP-6 (50686)       LAER*F*P*TL*SV*DP*GA*Q*E*TL*P*HC*SD*G*LEN*CS*V*Q*F*SG*P*AT*DS*P----- 365
MnP-10 (117436)     RAER*F*P*TL*TV*DR*G*AV*Q*AL*P*HC*SN*G*G*Q*CS*V*Q*F*DG*PA----- 357
PHACHMnP1 (66512)   PSE*F*P*TL*TV*Q*G*AS*Q*SL*IA*H*PD*G*SM*S*CS*V*Q*F*NG*GA----- 357
MnP-5 (49683)       FAER*F*P*TL*P*V*TP*G*AT*Q*TL*P*HC*SN*G*G*ED*CT*V*V*Q*F*TG*PA----- 356
DICSQ MnP6 (169524)  AST*F*P*TL*P*DP*G*PASS*V*AP*V*PPS----- 342
PHLRA MnP3 (Q96TS6)  AT*P*F*P*SL*SAD*P*G*P*AT*S*V*AP*V*PPS----- 339
MnP-13 (124076)     ATA*F*P*TL*P*ID*V*G*F*Q*TS*IP*V*P*V*GS----- 338
PLEOS MnP4 (1099081) RK*AF*F*P*RL*P*DK*G*TF*TS*IL*P*V*PS----- 337
* * * : : * * * * : : * * * * : : * * * * : : * * * * : : * * * * : : * * * * : : * * * * : : * * * * : :

```

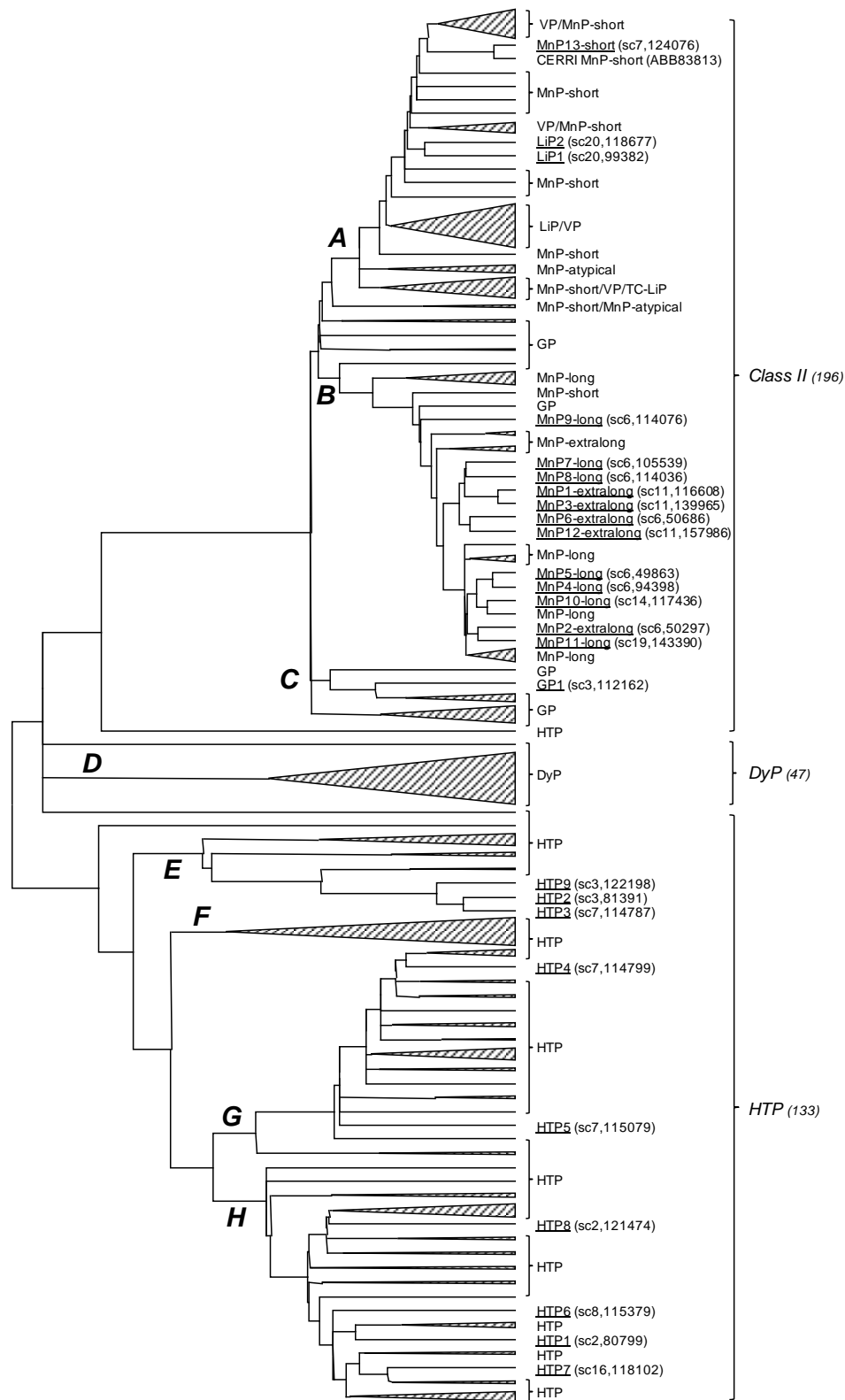
**Figure S2.** Multiple alignment of extralong, long and short MnPs sequences. *C. subvermisporea* extralong (MnP6/MnP12), long (MnP5/MnP10) and short (MnP13) MnPs, extralong MnPs from *Phlebia radiata* (GenBank CAC85963) and *D. squalens* (JGI 169843), long MnP from *P. chrysosporium* (JGI 66512), and 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<sup>2+</sup> ligands (green); **iii**) two active site histidines (dark gray); and **iv**) three acidic residues (red) forming the Mn<sup>2+</sup>-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  $\text{Mn}^{2+}$ -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. subvernispota* 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  $\text{Ca}^{2+}$  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).