Characterization of a multicopper oxidase gene cluster in *Phanerochaete chrysosporium* and evidence of altered splicing of the *mco* transcripts

Luis F. Larrondo, Bernardo González, Dan Cullen and Rafael Vicuña

INTRODUCTION

White rot fungi are the only organisms capable of efficiently degrading lignin to CO₂ and H₂O. Among these, the basidiomycete *Phanerochaete chrysosporium* has been the most extensively characterized (Gold & Alic, 1993; Kirk & Farrell, 1987). When growing on wood, *P. chrysosporium* is highly aggressive towards lignin. This property, together with its ability to grow at rather elevated temperatures (about 40°C) in compost, makes it a suitable micro-organism for biotechnological applications, such as biopulping and soil bioremediation. Its genome has been sequenced (Martinez et al., 2002). The genome of *P. chrysosporium* contains five and ten genes encoding MnP and LiP, respectively (Cullen, 1997; Martinez et al., 2004).

Laccase is another type of extracellular enzyme that has been implicated in lignin degradation. Laccases catalyse the one-electron oxidation of phenols, aromatic amines and other electron-rich substrates, with the concomitant four-electron reduction of O₂ to H₂O. This versatile phenol oxidase also participates in lignin biogenesis in higher plants and in a broad range of cellular processes such as sporulation (Leatham & Stahman, 1981), fruiting body formation and plant pathogenesis (Choi et al., 1992; Geiger et al., 1986; Marbach et al., 1985).

Laccases belong to the multicopper oxidase family, which also contains ascorbate oxidase, Fet3 ferroxidases and ceruloplasmin (Solomon et al., 1996). Of these so-called blue proteins, Fet3 (Askwith et al., 1994), ceruloplasmin (Williams et al., 1974) and *P. chrysosporium* MCO1 (Larrondo et al., 2003) are capable of efficiently oxidizing ferrous ion. In their active sites, multicopper oxidases feature copper centres that can be distinguished by their spectrophotometric properties. Type I copper (blue copper)
exhibits an intense absorption at about 600 nm, owing to the charge transfer between Cu(II) and a cysteine residue. Type II copper shows a very weak absorption, and functions as a one-electron acceptor. Type III copper contains two copper centres absorbing at 330 nm, and functions as a two-electron acceptor (Jonsson et al., 1995; Solomon et al., 1996).

Crystal structures of three fungal laccases have been determined, allowing identification of residues involved in copper binding at the three copper centres (Ducros et al., 1998; Piontek et al., 2002; Bertrand et al., 2002; Hakulinen et al., 2002). Multiple sequence alignments of fungal laccases available in public databases have identified a distinctive sequence signature that distinguishes fungal laccases from plant laccases and other multicopper oxidases. This signature consists of four unangled sequence segments, L1 to L4, which contain the 12 residues serving as copper ligands (Kumar et al., 2003).

For decades it was thought that P. chrysosporium does not produce laccase (Eriksson et al., 1990; Kirk & Farrell, 1987; Thurston, 1994; Hatakka, 1994). More recently, laccase-like activity was reported in P. chrysosporium cultures grown under certain conditions (Srinivasan et al., 1995; Dittmer et al., 1997; Rodriguez et al., 1997), although these results were not widely accepted (Podgornik et al., 2001). Analysis of the P. chrysosporium genome showed that it does not contain any genes encoding conventional laccases. Instead, this micro-organism produces a multicopper oxidase possessing strong ferroxidase activity with catalytic parameters similar to those of Fet3p (Larrondo et al., 2003). The physiological function of this protein (MCO1) is still unknown. It might be involved in modulating extracellular Fenton-based chemical reactions (Larrondo et al., 2003), which appears to be one of the roles of yeast Fet3p (Stoj & Kosman, 2003; Shi et al., 2003). However, most iron in wood would be expected to be in the ferric form, due to spontaneous oxidation.

Further inspection of the P. chrysosporium genome revealed three additional mco genes, whose characterization is reported here. Multiple alignment and analysis of the L1–L4 signature defined by Kumar et al. (2003) confirmed that these genes indeed code for a novel type of multicopper oxidase. Unexpectedly, cDNAs corresponding to the mco genes showed the presence of multiple splicing variants. The altered transcripts contained several introns, and in some cases lacked certain exons. In all instances, the presence of in-frame stop codons within introns would prevent synthesis of an active enzyme. A similar phenomenon has been recently observed in P.450 (Yadav et al., 2003) and peroxidase (M. Stuardo and others, unpublished results) transcripts of the same fungus. We refer to this seemingly unproductive phenomenon as ‘altered splicing’, as opposed to the well-established concept of ‘alternative splicing’, which generally connotes translation into functional protein variants.

**METHODS**

**Fungal strains and culture conditions.** P. chrysosporium homokaryotic strain RP-78 (Stewart et al., 2000) and dikaryotic strain BKM-F-1767 were obtained from the Center for Mycology Research, Forest Products Laboratory, Madison, WI. The fungus was grown in defined media containing wood-derived crystalline cellulose (Avicel) as sole carbon source for six days, as previously described (Wymelenberg et al., 2002). Mycelia were harvested by filtration through miracloth (Calbiochem), followed by immersion in liquid nitrogen, and stored at −90 °C.

**cDNA cloning and analysis.** RNA was extracted from frozen mycelia as described by Manubens et al. (2003). Poly(A) RNA was isolated using a magnetic-capture technique involving oligo(dT)25 Dynabeads (Dynal), following the manufacturer’s instructions. cDNAs were obtained by RT-PCR of the purified poly(A) RNA, as described by Wymelenberg et al. (2002), using the primers indicated in Table 1. The RT-PCR products were cloned into pGEM-T easy vector (Promega) and introduced into *Escherichia coli* DH5α cells (Stratagene), following standard procedures (Ausubel et al., 1992). Nucleotide sequences were determined with the ABI Prism Big Dye Terminator Cycle Sequencing kit (Perkin-Elmer Applied Biosystems) on ABI automated sequencers. Genomic nucleotide sequences for the *Pc*-mco4A, *Pc*-mco1A and *Pc*-mco2A genes are available at www.jgi.doe.gov/whiterot, and were produced by the US Department of Energy Joint Genome Institute (www.jgi.doe.gov). GenBank accession numbers assigned to the cDNAs are indicated below. Nucleotide and amino-acid-sequence similarity searches used the BLAST method (Altschul et al., 1990) on the National Center for Biotechnology Information databases. Nucleotide and amino-acid sequences were analysed by using the DNASTAR software (DNASTAR, Madison, WI).

**RESULTS**

**Characterization of the mco cluster in P. chrysosporium**

Careful analysis of the *P. chrysosporium* genome revealed four putative MCO-encoding sequences on scaffold 56 between coordinates 138628 and 163323. Of these, only

<table>
<thead>
<tr>
<th>Target</th>
<th>Primer sequence</th>
<th>Product size (nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mco2</td>
<td>FW ACT TAC GAC CCC CCC ACA</td>
<td>2728 1931</td>
</tr>
<tr>
<td>RV TCG CTG AAT CTC ACC GGC CA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mco3</td>
<td>FW ATG GCC GCC GCT GAG GAC AT</td>
<td>2671 1871</td>
</tr>
<tr>
<td>RV AAT ATG CCC TCG TCT TCG GT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mco4</td>
<td>FW ATG GAC CAG AGC GCA TAT TC</td>
<td>2768 1776</td>
</tr>
<tr>
<td>RV TCA CGC GCT GCA CTG CCT CA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FW, Forward; RV, reverse.
one (mco1) features a defined secretion signal and has had its protein product further characterized (Larrondo et al., 2003). As depicted in Fig. 1A, the four mco sequences are clustered in a 25 kb segment, where mco1, mco2 and mco3 are tightly grouped, while mco4 is located 11 kb upstream of mco1. In the latter intergenic region, only one predicted gene model (pc.56.75.1), encoding a putative acid phosphatase, is observed. Due to the complex intron–exon composition of these sequences, particularly the presence of short exons, the automatic annotation process generated inaccurate models for all four genes. Therefore, models for mco1 (Larrondo et al., 2003) and the three new mco sequences were manually constructed and confirmed by cDNA cloning and sequencing (see below). Comparison of the genomic sequences with those of the corresponding cDNAs showed the presence of 14 introns in mco2 and mco3, and 18 in mco4 (Fig. 1B). As previously reported, mco1 contains 19 introns (Larrondo et al., 2003). The mco3 gene is interrupted by an 8·1 kb retroelement inserted in intron 12 of the A allele, while allele B is intact (L. F. Larrondo and others, unpublished results). Therefore, mco3 cDNA was isolated from dikaryotic P. chrysosporium strain BKM-1767. The nucleotide identity between the four mco sequences ranged from 54·6 % to 69·2 % (Fig. 1C). Identities between mco cDNAs and deduced amino-acid sequences are tabulated in Figs 1D and 1E, respectively. The highest levels of similarity were observed between mco2 and mco3.

The intron composition of the four genes is similar, especially in conserved areas such as the exons encoding the histidines of the three copper centres (in mco1, exons V, VI, VII, XVI, XIX and XX). Conservation of the splicing sites is also observed in several positions.

**Fig. 1.** The *P. chrysosporium* mco gene cluster. (A) Genomic organization of the mco genes present in scaffold 56. (B) Intron/exon structure of the mco gene family members. Introns are represented as small white rectangles; exons correspond to the larger grey boxes. Exon numbers are indicated. (C), (D) and (E) show the nucleotide identity between the mco genes, the cDNAs and the deduced amino-acid sequences, respectively.
Altered splicing of *P. chrysosporium* mco transcripts

In order to obtain the cDNAs of the mco genes, RT-PCRs were conducted with poly(A) RNA derived from Avicel cultures grown for 6 days (Wymelenberg *et al.*, 2002). The primers employed in these experiments are shown in Table 1. The RT-PCR products were cloned in pGEM-T, and after transformation in *E. coli*, ten clones were analysed by PCR for each product. Clones from *mco1* appeared to be full length, based on their sizes and partial sequencing. Unexpectedly, however, length differences were observed among the different clones of the other mco cDNAs (data not shown). Those showing major variations were sequenced. Comparison of the genomic and cDNA sequences allowed the confirmation of our manual annotation of the mco genes, but at the same time revealed the presence of several introns in some of the cDNAs.

Analysis of the mco2 clones showed the presence of at least three different types of cDNAs. One of them corresponded to the full-length coding sequence, where all the predicted introns had been spliced (*mco2A*). In turn, *mco2A-11* corresponded to a full-length cDNA plus intron I, whereas

---

**Fig. 2.** Schematic representation of mco genes and characterized cDNAs. The intron/exon composition of mco2 (A), mco3 (B) and mco4 (C) are indicated to scale. The different cDNAs identified for each gene are also represented; their names indicate whether or not they correspond to alleles A and B. Exons encoding some of the histidines involved in the coordination of T3 copper are indicated in red and blue, whereas the exon encoding the ferroxidase domain is indicated in brown. Exons VI and XI of mco3, which appear to be erroneously processed in transcript mco3B-E6/11, are indicated in light brown and cyan, respectively. The lengths of the predicted proteins, as well as the positions up to which homology is maintained with respect to the full-length protein, are indicated in the columns on the right.
mco2A-I8 was the full-length cDNA with intron VIII (Fig. 2A).

Inspection of the mco3 cDNAs revealed seven different types of transcripts. Only one (mco3B) had complete splicing of all introns and encoded a full-length multicopper oxidase (Fig. 2B). Transcript mco3B-I4/I11 lacked exon VII, but possessed introns IV and XI. Transcript mco3B-I5/I10 had introns V and X, whereas transcripts mco3B-I6 and mco3B-I10 had introns VI and X, respectively. Transcript mco3B-IM still contained a considerable number of introns: II, IV, V, VI, VIII and IX. Finally, transcript mco3B-E6/I11 was devoid of all introns, but lacked a region expanding part of exon VI (light brown) to exon XI (cyan).

Assessment of the transcripts derived from mco4 showed the presence of four different types of product (Fig. 2C). Two corresponded to full-length cDNA from alleles A and B, where all the introns had been correctly spliced out. Transcript mco4B-I5 contained part of intron V and lacked exon VI, whereas transcript mco4B-I13 possessed intron XIII.

mco-encoded sequences

Assuming splicing of all predicted introns, the four mco genes encode for multicopper oxidases containing all residues required to assemble the copper centres characteristic of this family (Solomon et al., 1996). However, as described above, most transcripts characterized here contained one or more introns. The altered transcripts generated premature stop codons within the introns which would terminate translation (Fig. 2). Such truncated proteins, if synthesized, would lack enzymic activity due to the absence of key amino-acid residues. Although mco3B-I5/I10 might encode a truncated protein, reading through the intron would lead to the presence of 90 residues with no homology to other MCOs, and to the appearance of a stop codon in exon VI.

The cDNA mco3B-I4/I11 presented an interesting feature: the absence of exon VII. This exon contained the residue equivalent to Glu-185 of the Saccharomyces cerevisiae Fet3 protein, which has been shown to be essential for ferroxidase activity (Bonaccorsi di Patti et al. 2000; Wang et al., 2003). Fig. 3 shows a multiple alignment of some fungal laccases, Fet3 proteins and P. chrysosporium MCO sequences in the area of this residue. If the entire exon VII is spliced from mco3 (exons IX in mco1, VII in mco2 and IX in mco4), the resulting proteins do not have the aforementioned residue, but they gain homology to conventional laccases (Fig. 3). A hypothetical transcript from mco3 lacking all the introns plus exon VII would possess an ORF for a putative MCO that would be unable to oxidize Fe^{3+} due to the absence of the 17–18 aa segment where the Glu residue is located.

Analysis of the L1–L4 signature sequences defined by Kumar et al. (2003) revealed that the four MCOs have differences within these consensus sequences (Table 2). This is especially evident in the case of L1 and L2. For example, in L1, Gln is replaced by Glu and Pro by Gly at positions nine and eleven, respectively, in all Pc-MCOs. In L2, the fifth position is Trp, instead of the Tyr present in

Fig. 3. Partial alignment of splicing variants at the ferroxidase domain. A partial alignment was constructed with CLUSTAL V using the Fet3 proteins from P. chrysosporium and S. cerevisiae, fungal laccases and the P. chrysosporium mco sequences harbouring or not harbouring the exons shown to be essential for ferroxidase activity (see text). The arrow indicates the position of Glu-185.
most laccases, while in the seventh position, three out of the four MCOs possess Ala instead of Ser. In the L2 QYCDGL subsignature, Leu has been replaced by Ile or Val in all four Pc-MCOs, whereas Cys has been replaced by Thr.

**DISCUSSION**

We previously described MCO1, a new multicopper oxidase from *P. chrysosporium* (Larrondo et al., 2003). Although MCO1 possesses residues characteristic of this protein family, it also has novel features that preclude its classification in any of the existing branches of multicopper oxidases. This observation was confirmed by its biochemical characterization, which showed that although MCO1 does not correspond to a canonical Fet3 protein, it behaves as a ferroxidase with catalytic parameters similar to those described for *S. cerevisiae* Fet3p. The activities of MCO1 are unlike conventional laccases.

In this work, we characterized three additional *P. chrysosporium* mco sequences. All four genes share the same transcriptional orientation and are organized within a 25 kb cluster. mco1, mco2 and mco3 are tightly grouped, with mco4 located 11 kb upstream of mco1. Based on sequence homology and intron composition, it can be presumed that some duplication events has taken place. In this context, it may be relevant that mco2 and mco3 exhibit the highest sequence identity and the same intron composition.

Gene clustering seems to be a common feature in *P. chrysosporium*. This phenomenon has been reported for ten LiP (Stewart & Cullen, 1999), three cellobiohydrolases (Covert et al., 1992), and two cytochrome P-450 genes (Yadav et al., 2003). Recent genome analysis has revealed additional clustering among P-450, MnP and copper radical oxidase genes (Martinez et al., 2004). Families of structurally related genes have also been found in *Pleurotus sajor-caju* (Soden & Dobson, 2001), the basidiomycete CECT 20197 (Mansur et al., 1997), *Pleurotus ostreatus* (Palmieri et al., 2000), *Trametes villosa* (Yaver & Golightly, 1996), *Agaricus bisporus* (Perry et al., 1993) and *Rhizoctonia solani* (Wahleithner et al., 1995).

Different splicing alterations were observed in the mco cDNAs. Most consisted of the presence of introns possessing stop codons. We refer to this phenomenon as 'altered splicing', to differentiate it from 'alternative' or 'differential' splicing. The latter gives rise to transcripts differing in the number of exons that lead to the production of active proteins. The fact that some transcripts exhibited modified
splicing sites, whereas others lacked some exons, makes us believe that the observed splicing variants do not correspond to intermediates captured in the splicing process. Thus, one of these messengers (mco4A-I5) included a modified splicing site in intron VI, 10 nt downstream from the original site. In this transcript, the 5 nt-long exon VII had been spliced, and consequently it was not present in the final messenger. In the same vein, transcript mco3B-E6/11 lacked a large segment comprising a portion of exon VI to the middle of exon XI. In this case, it is likely that several splicing events took place, since it has been observed that intron length in \textit{P. chrysosporium} is strongly constrained to a mean of about 54 nt (Martinez et al., 2004). The reasons for aberrant processing in \textit{P. chrysosporium}, which has also been observed in peroxidase (M. Stuardo and others, unpublished results) and P-450 (Yadav et al., 2003) transcripts, are presently unknown.

In comparison to higher eukaryotes, there are few examples of differential splicing reported in fungi. Some cases are the glucoamylase gene of \textit{Aspergillus niger} (Boel et al., 1984), the \textit{cox}-5 gene in \textit{Neurospora crassa} (Sachs & Yanofsky, 1991), the genes YKL186C/MTR2 and YML034W in \textit{S. cerevisiae} (Davis et al., 2000), the inosine-5’-monophosphate dehydrogenase gene in \textit{Pneumocystis carinii} f. sp. \textit{carinii} (Ye et al., 2001), and the \textit{ctl} and \textit{ctr}B genes in \textit{Xanthophyllomyces dendrorhous} (Lodato et al., 2003). Differential splicing has been reported in \textit{P. chrysosporium} exocellulbiohydrodrolases \textit{cbhl.1} and \textit{cbhl.2} (Sims et al., 1994; Birch et al., 1995), where the 3’ intron of both transcripts is not excised in a substrate-dependent manner. The two splicing variants appear to encode functional proteins, differing only in the presence or absence of a cellulose-binding domain.

Differential or altered splicing in \textit{P. chrysosporium} seems to be a common feature among highly conserved families of clustered genes. Thus, \textit{lipA}, \textit{lipB} and \textit{lipB} are members of a large cluster (Stewart & Cullen 1999), while \textit{mnp1} and \textit{mnp4} are only 5-7 kb apart (M. Stuardo and others, unpublished results). In turn, \textit{cbhl.1} and \textit{cbhl.2} are separated by only 750 bp, while \textit{pc-1} lies 322 bp apart from the \textit{pc-2} P-450 gene (Yadav et al., 2003). As reported in this work, the \textit{mco} genes are also clustered. Whether this altered splicing is somehow related to genomic multiplicity and/or genome organization remains to be determined. Assuming family members have overlapping or redundant roles, incorrect processing is less likely to have a deleterious impact relative to single genes. This might constitute a still-imperfect version of the differential splicing that is commonly observed in higher eukaryotes.

The MCO sequences diverge from the L1–L4 signature sequences defined for laccases (Kumar et al., 2003) in several positions (Table 2). In L1 for example, Glu replaces Gln and Gly replaces Pro in all four MCOs. Interestingly, Glu is also found in some laccases, such as those from \textit{Cryptococcus neoformans} (GenBank A36962), \textit{Cryptocentria parasiticus} (GenBank Q03966), \textit{Neurospora crassa} (GenBank P06811) and \textit{Podospora anserina} (GenBank P78722). Also of interest is the fact that three out of the four MCOs possess Ala instead of Ser at the seventh position of L2. This replacement is similar to that observed in most ascorbate oxidases and plant laccases. Substitution of Leu by Ile or Val in L2 is not unexpected, considering that this position is also modified in the laccases mentioned above. In addition, the Cys present in this signature has been replaced by a Thr in all Pc-MCO sequences. Once again, this position varies in the aforementioned laccases and also in laccases from \textit{Thanatephorus cucumeris} (GenBank P56193) and \textit{Mycelio- phthora thermophila} (Berka et al., 1997), where Cys has been replaced by Ala, Gly or Val.

The unusual activity attributed to MCO1, and the high similarity with MCO2, MCO3 and, to a lesser degree, with MCO4, suggests that this group constitutes a new branch of the fungal multicopper oxidase family. This is in agreement with the partial divergence observed in the L1–L4 sequences (Table 2). This new family grouping also correlates with the presence of some of the residues essential for ferroxidase activity (Fig. 3). Based on the strong similarities observed between MCO1, MCO2 and MCO3, it could be predicted that the last two proteins will also have ferroxidase activity. Although MCO4 also has high identity with MCO1, both the absence of a residue equivalent to Glu-185 and the partial divergence in that area may increase this prediction more tenuous.

Until our recent work describing MCO1 (Larrondo et al., 2003), the only fungal multicopper oxidases possessing ferroxidase activity were the members of the Fet3 family, which are type I transmembrane proteins. We have recently identified and characterized the Fet3 orthologue from \textit{P. chrysosporium}, which contains a typical ferroxidase transmembrane domain (L. F. Larrondo and others, unpublished results). MCO proteins lack this domain, thus supporting the idea that they constitute a new branch of the multicopper oxidase tree.

Of the four MCOs, only MCO1 harbours a secretion signal that is functional in \textit{Aspergillus nidulans} (Larrondo et al., 2003). The other three proteins have extended N-terminal domains that are longer than those of MCO1 and other members of the multicopper oxidase family, but lack defined secretion signals. Their cellular location and biological roles remain uncertain.

\textbf{ACKNOWLEDGEMENTS}

This work was financed by the Millenium Institute for Fundamental and Applied Biology, by grant 1030495 from FONDECYT and US Department of Energy grant DE-FG02-87ER13712.

\textbf{REFERENCES}


