Cloning and Characterization of Polyphenol Oxidase cDNAs of *Phytolacca americana*

Richard W. Joy IV²*, Munetaka Sugiyama, Hiroo Fukuda, and Atsushi Komamine³

Biological Institute, Faculty of Science, Tohoku University, Aoba-yama, Aoba-ku, Sendai 980, Japan

Two cDNA clones encoding polyphenol oxidases were isolated from a cDNA library constructed from a log-phase suspension culture of *Phytolacca americana* (pokeweed) producing betalains. The clones exhibit 93 and 86% sequence identity at the nucleotide and deduced amino acid levels, respectively. Both clones contain two copper-binding domains characterized by histidine-rich regions, which are found ubiquitously in all polyphenol oxidases/tyrosinases, and a putative third histidine-rich, copper-binding region, which is common to all plant polyphenol oxidases. One of the *Phytolacca* cDNA deduced amino acid sequences contains the ubiquitous transit peptide for all proteins targeted to the internal lumen of thylakoid membranes of plastids and is considered to be 98 residues in length based on a proposed sequence cleavage site motif. This would produce a processed peptide of approximately 54 kDa. In addition to common features of transit peptides, it was found that an additional conserved region for polyphenol oxidases was located between the hydroxy amino acid-rich region and the thylakoid transfer domain. Spatial and temporal expression was investigated by northern blot analysis of total RNA from various organs of *Phytolacca* plants. Transcripts of the two clones were found to be 2.1 and 2.3 kb, respectively. Both transcripts were present only at substantial levels in ripening, betalain-containing fruit.

PPO (also known as tyrosinase in some species) has been shown to be almost universally present in a diverse and varied number of species such as insects, fungi, vertebrates, and plants (Mayer, 1987). Studies with vertebrates have shown that there is a direct correlation between pigmentation and a functional tyrinosinase gene (Yamamoto et al., 1989; Shibahara et al., 1990; Tanaka et al., 1990). However, to date the physiological function of PPOs in plants has yet to be elucidated.

These nuclear-encoded enzymes are located in the plastids of plant cells and catalyze the hydroxylation of monophenols, such as Tyr, and the dehydrogenation of ortho-diphenols, such as 1-dopa, to form ortho-quinones. In biological systems, the ortho-quinones go on to form the cyclical intermediate cyclodopa, also called leucodopachrome, in vertebrate pathways. These reactions are known to typically form the beginning of the melanin and betalain pathways (Steglich and Strack, 1990; Pawelek, 1991; Riley, 1993). Specifically, tyrosinase catalyzes the first two reactions in the melanin pathway, namely the production of 1-dopa from Tyr and the production of dopaquinone from 1-dopa. These reactions are also typical of the betalain pathway in the members of the Centrospermae, which produce the prominent red betacyanin pigmentation and yellow betaxanthin pigmentation; however, the enzymology for these initial steps still needs to be determined conclusively (Endress, 1979; Rink and Bohm 1985, 1991; Bohm and Rink, 1988; Steglich and Strack, 1990; Kishima et al., 1991). Tyr, as in melanin production, is the sole precursor to these secondary metabolites.

Recently there have been several reports of the isolation of plant PPO genes and cDNAs (Cary et al., 1992; Shahar et al., 1992; Hunt et al., 1993; Newman et al., 1993). Of the species described in these reports, tomato (*Lycopersicon esculentum*) PPOs have been most extensively characterized, with seven individual genomic clones isolated (Shahar et al., 1992; Newman et al., 1993). Additional PPOs from *Vicia faba* (Cary et al., 1992) and *Solanum tuberosum* (Hunt et al., 1993) have begun to add to the information on this gene. In tomato it was found that PPO expression was highly expressed in floral tissues (Shahar et al., 1992). Expression in potato was found in younger developing tissues but also in fully developed inflorescences at low levels (Hunt et al., 1993).

In *Phytolacca americana* (also known as pokeweed), red pigmentation (betalains) occurs naturally in various organs of the plant. Cell cultures derived from these plants have also been found to contain high concentrations of betalains in rapidly proliferating cell cultures (Sakuta et al., 1986). Furthermore, betalain accumulation was shown to be regulated at the step of Tyr hydroxylation in association with cell proliferation (Hirano and Komamine, 1994). As such, this material has provided a valuable system for studying the regulatory mechanisms of betalain synthesis. As the first part of a molecular investigation into the role of PPO in plant growth and development and its possible role in the betalain pathway, we report the cloning and characterization of PPO cDNAs from *P. americana*.

---

¹ This work was supported by grants from the Ministry of Education, Science and Culture in Japan. R.W.J. was supported by a fellowship from the Japanese Society for the Promotion of Science while in Japan.

² Present address: Department of Biological Sciences, University of Calgary, 2500 University Drive N.W., Calgary, Alberta, Canada, T2N 1N4.

³ Present address: Department of Chemical and Biological Sciences, Faculty of Science, Japan Women's University, Mejirodai 2-8-1, Bunkyo-ku, Tokyo 112, Japan.

* Corresponding author; e-mail rich@acs.ucalgary.ca; fax 1–403–289–9311.
MATERIALS AND METHODS

Plant Materials

Suspension cultures were prepared from callus initiated from stem explants of *Phytolacca americana*. L. Cells were subcultured every 7 d in Murashige and Skoog medium (1962) containing 3% Suc and 5 μM 2,4-D as described previously (Sakuta et al., 1986). This cell line actively accumulates betalains during log-phase growth. For the analysis of RNA, wild-grown plants were collected from the Tohoku University campus area.

PCR Amplification of PPO cDNA Fragments

Total RNA was extracted from log-phase suspension cultures of *P. americana*, and poly(A)+ RNA was isolated by oligo(dT)-cellulose chromatography (Sambrook et al., 1989). cDNA was prepared by first-strand synthesis with reverse transcriptase and second-strand synthesis with the Klenow fragment of DNA polymerase (cDNA Synthesis Kit, Pharmacia). Consensus sequences of the copper-binding domains of published PPO/Tyr sequences were used to design the following oligonucleotides: (I) 5'-GCGGATCC(TCGAC)(AG)(AT)(AT)(AGCT)(AG)TG(AG)TG-3'; (II) 5'-GCGGATCC(CAGCT)TA(CT)TGG(AG)A(CT)TGG-3'; and (III) 5'-GCGGATCC(CAGCT)TA(CT)TGG(AG)A(CT)TGG-3'.

Oligonucleotides were derived from consensus regions from the comparison of copper-binding domains published by Shahar et al. (1992). Oligonucleotide I was designed from copper-binding domain A amino acids: Pro, Tyr, Trp, Glu/Gln, Trp; oligonucleotide II was designed from copper-binding domain B amino acids: His, His, Ala/Val, Asn/Phe/Tyr.

Using constructed cDNA as a template, oligonucleotide I as a forward primer, and oligonucleotide II as a reverse primer, DNA sequences were amplified by PCR (50 cycles denaturing, annealing, and polymerization were carried out at 93°C for 1 min, 45°C for 1 min, and 73°C for 2 min, respectively). PCR products were run on low-melting agarose and amplified DNA was purified with a Prep-A-Gene (United States Biochemical) and [α-32P]dCTP (Amersham). Automated nucleotide sequencing was also performed with the Taq Dye Primer Cycle Sequencing kit (Applied Biosystems, Foster City, CA) using the fluorescent dye-labeled M13 universal or reverse primers. The cycle-sequencing reaction was carried out in a Renovern thermal cycler (Renover Science Co., Ltd., Tokyo, Japan) and the nucleotide sequences were obtained on an automated DNA sequencer (model 373A, Applied Biosystems). Nucleotide and amino acid sequences were analyzed using the GENE-TYX software system (Software Development Co., Tokyo, Japan).

Isolation and Analysis of RNA

Total RNA was isolated from various plant organs of *Phytolacca* using the protocol of Sambrook et al. (1989). Total RNA was fractionated through Mops-formaldehyde 1% agarose gels and transferred to Hybond N* filters (Amersham) according to the methods of Sambrook et al. (1989). Filters were hybridized to PAP1 or PAP2 cDNA probes (32P-labeled by random-primed synthesis; Mega-prime, Amersham) in hybridization buffer containing 50% formamide at 42°C. Filters were washed under high stringency according to the methods of Sambrook et al. (1989). Washed filters were exposed to Fuji RX films (Fuji Photo Film Co., Ltd., Tokyo, Japan) under intensifying screens at −80°C.

RESULTS

Isolation of a cDNA Fragment for *Phytolacca* PPO

Using double-stranded cDNA produced from poly(A)+ RNA from log-phase suspension cultures of *P. americana* as a template, a 337-bp fragment was amplified by PCR with primers designed on the basis of conserved sequences for copper A- and B-binding domains of PPO. The fragment was excised, blunt-end ligated into pBluescript and sequenced. The amino acid sequence deduced from the determined nucleotide sequence was found to have high similarity to the copper B-binding domain of known PPOs (data not shown). The consensus sequences used to amplify *Phytolacca* PPO DNA were expected to produce a 500- to 600-bp fragment that was anticipated to comprise both copper-binding domains. The resulting fragment (PCR337), however, contained only one of the copper-binding domains and only one of the primer sequences. The reason for the possible truncation of the expected product is unknown.

Isolation and Characterization of PPO cDNA Clones

The resulting amplified product, PCR337, was used as a probe to screen a AZAP II cDNA library produced from log-phase suspension cultures of *P. americana*. Thirty plaques remained positive after tertiary screening of approximately 1.5 × 10⁶ primary plaques. Of these, two

DNA Sequence Analysis

The cDNA inserts of the PCR-amplified PPO probe and PPO clones from the AZAP II library were sequenced in both directions using the dideoxy chain termination method with a Sequenase, version 2.0, DNA-sequencing kit (United States Biochemical) and [α-32P]dCTP (Amersham).
unique clones were isolated and further analyzed at the nucleotide and deduced amino acid sequence levels. Clone 7 is 1749 bp in length with an uninterrupted open reading frame of 1425 bp encoding a deduced sequence that is 475 amino acids in length. Clone 7 does not contain a start ATG codon; thus, it is likely a truncated clone. It contains a stop codon; (TGA) at position 1426 in the DNA codon (data not shown). Clone 32 is 2099 bp in length with an uninterrupted open reading frame of 1758 bp encoding a deduced sequence that is 586 amino acids in length.

Clone 32 contains a stop codon (TGA) at position 1426 in the DNA codon (data not shown). Clone 32 is 2099 bp in length with an uninterrupted open reading frame of 1758 bp encoding a deduced sequence that is 586 amino acids in length. Clone 32 contains a start ATG codon at position 11 in the 5' end of the clone. It contains a stop codon (TGA) at position 1769 in the DNA sequence and a polyadenylation signal (AAATAAA) at position 1872, 103 nucleotides downstream from the stop codon (data not shown). Little sequence divergence was found between the clones in the copper-binding domains; however, dissimilarities in DNA and deduced amino acid sequences were found 3' of the copper B-binding domain. PPOs encoded by clone 32 and clone 7 were designated PAPl and PAP2, respectively.

Both clones contain putative copper-binding domains characteristic of known PPOs. Copper-binding domains reside between positions 334 and 383 (domain A, 21 amino acid residues being identical) and between positions 1523, 97 nucleotides downstream from the stop codon (data not shown). Clone 32 is 2099 bp in length with an uninterrupted open reading frame of 1758 bp encoding a deduced sequence that is 586 amino acids in length. Clone 7 contains a start ATG codon at position 11 in the 5' end of the clone. It contains a stop codon (TGA) at position 1769 in the DNA sequence and a polyadenylation signal (AAATAAA) at position 1872, 103 nucleotides downstream from the stop codon (data not shown). Little sequence divergence was found between the clones in the copper-binding domains; however, dissimilarities in DNA and deduced amino acid sequences were found 3' of the copper B-binding domain. PPOs encoded by clone 32 and clone 7 were designated PAPl and PAP2, respectively.

Figure 1 shows the deduced amino acid sequences of the two PPO clones from *P. americana*. Both clones contain putative copper-binding domains characteristic of known PPOs. Copper-binding domains reside between positions 334 and 383 (domain A, 21 amino acid residues being identical) and between positions 1523, 97 nucleotides downstream from the stop codon (data not shown). Clone 32 is 2099 bp in length with an uninterrupted open reading frame of 1758 bp encoding a deduced sequence that is 586 amino acids in length. Clone 7 contains a stop codon (TGA) at position 1426 in the DNA codon (data not shown). Clone 32 is 2099 bp in length with an uninterrupted open reading frame of 1758 bp encoding a deduced sequence that is 586 amino acids in length. Clone 32 contains a start ATG codon at position 11 in the 5' end of the clone. It contains a stop codon (TGA) at position 1769 in the DNA sequence and a polyadenylation signal (AAATAAA) at position 1872, 103 nucleotides downstream from the stop codon (data not shown). Little sequence divergence was found between the clones in the copper-binding domains; however, dissimilarities in DNA and deduced amino acid sequences were found 3' of the copper B-binding domain. PPOs encoded by clone 32 and clone 7 were designated PAPl and PAP2, respectively.

Figure 1. Deduced amino acid sequences of *P. americana* PPO clones PAPl (top sequence) and PAP2 (bottom sequence). A vertical line indicates a putative cleavage site of the transit peptide. Identical amino acid residues are indicated by colons. Dashes are gaps introduced to maximize identity.

Expression of PPO in *Phytolacca* Plants

Spatial and temporal expression was investigated by the analysis of total RNA from various organs of *Phytolacca* plants. Although roots, leaves, and floral organ stalk tissues at different developmental stages were examined for the presence of PPO transcripts by northern hybridization with PAPl and PAP2 cDNAs as probes, PPO transcripts were not detectable in any of these organs (data not shown). The presence of PPO transcripts was examined in fruits at different developmental stages (Fig. 4). No hybridization was detected with developing fruit that remained green; however, strong signals were detected in fruit that had turned red with betacyanin accumulation (Fig. 5).
Figure 2. Conserved sequences for PPO copper-binding domains A and B from various plant species and for a putative third His-rich copper-binding domain. V. faba, Broad bean (Cary et al., 1992); S. tuberosum, potato (Hunt et al., 1993); L. esculentum, tomato (Shahar et al., 1992; Newman et al., 1993); P. americana, pokeweed.

Figure 3. A. Generalized schematic diagram of transit peptides that target proteins to the thylakoid lumen of chloroplasts. B. Conserved region I (n-region) and the thylakoid transfer domain of various plant PPO transit peptides. V. faba, Broad bean (Cary et al., 1992); S. tuberosum, potato (Hunt et al., 1993); L. esculentum, tomato (Shahar et al., 1992; Newman et al., 1993); P. americana, pokeweed.
DISCUSSION

There are several lines of evidence that the two Phytolacca cDNA clones isolated in the present study encode PPOs. Both clones have high similarities at both the nucleotide and amino acid levels with known plant PPOs. When comparing the full-length clone (PAP1) with published PPO sequences from other species, it was found that phylogenetically distant species had less similarity than nearer species.

PPOs are characterized at the amino acid level specifically by containing at least two copper-binding domains that have high homology among divergent species (Shahar et al., 1992). The two clones isolated from Phytolacca contain both conserved sequences for these copper-binding domains (Fig. 2). In addition to the two known copper-binding domains among all species, it appears that plant species may contain a third His-rich region (Fig. 2) at the C terminus of each sequence that was originally pointed out by Hunt et al. (1993). As mentioned by those authors, however, it is not yet known if this His-rich region has any catalytic significance in PPO function.

The calculated molecular mass of the unprocessed PAP1 protein would be approximately 64.7 kD, with the processed protein being approximately 54 kD (assuming that the transit peptide cleavage site is between Ala and Ala in agreement with cleavage site motifs proposed by Gavel and von Heijne, 1990). These molecular masses are well within the accepted ranges of those published in the literature (Cary et al., 1992; Shahar et al., 1992; Yu et al., 1992; Hunt et al., 1993; Newman et al., 1993).

Most published sequences of plant PPOs contain a highly hydrophilic region at the N terminus. This region is characteristic of transit peptides known to be targeted to the internal lumen of thylakoid membranes in plastids (Keegstra et al., 1989; de Boer and Weisbeek, 1991). Phytolacca PPO PAP1 has a highly similar transit peptide sequence compared to all known plant PPO transit peptide sequences (Figs. 1 and 3). The cDNA clone of PAP2 was truncated at the 5' end and no sequence information for its transit peptide was available.

The transit peptide of PAP1 was estimated to be 98 amino acid residues and had several unique features that are common to all plant PPO transit peptides. First, N termini all have a high content of hydroxy amino acids at the beginning of the sequence. In Phytolacca, the hydroxy amino acid content for its entire transit peptide is 25%, with 56% hydroxylated amino acids within the first 23 residues. A second region that contains a loosely conserved region within the subsequent 30 amino acids exists and is common to all PPO transit peptides except V. faba (Fig. 3). The third region, called the thylakoid transfer domain, has substantial similarity within a 14-amino acid region represented principally by hydrophobic amino acids (Fig. 3). This region is nearly identical in all plant PPO transit peptides.

Studies with different chloroplast-targeted proteins have several of these features in common. The high content of hydroxylated amino acids at the N terminus is a common and specific feature of chloroplast-targeted proteins. No conserved regions in any of these sequences have been found in the chloroplast-targeted proteins studied thus far (Keegstra et al., 1989; de Boer and Weisbeek, 1991). The second domain of approximately 30 amino acids after the hydroxilated amino acid-rich region in PPOs has not been observed to have conserved regions in proteins targeted to the thylakoid lumen. It does, however, contain several regions with highly similar amino acid sequences that are analogous to many lumen-targeted proteins including PPOs (cf. Keegstra et al., 1989; de Boer and Weisbeek, 1991; and refs. therein; and Fig. 3, region I). The similar sequences in the PPO n-region domain suggest that this region may have a function that is similar to that of the thylakoid transfer domain such as intraorganellar targetting, or that it comprises a recognition sequence for the stromal peptidase that apparently cleaves the precursor protein after region I to form an intermediate protein localized to the stroma (Somm er et al., 1994). The intermediate was subsequently shown to be processed to a mature polypeptide after translocation to the lumen. Thus, the PPO transit peptide contains regions for two-step processing to a mature protein (Somm er et al., 1994).

Based on the model by de Boer and Weisbeek (1991), it is thought that lumen-targeted proteins contain at least two...
main domains in the transit peptide that function in intra-organellar targeting. The first domain (the hydrophobic N terminus) targets the protein to the translocation complex where the protein enters the stroma and is subsequently cleaved. The thylakoid transfer domain, which also contains PPO region I in de Boer and Weisbeek’s model (also called the n-region), is considered in other species only to have similarly located charged amino acid residues, and targets the thylakoid membrane or lumen. Other than the thylakoid transfer domain and the transit peptide cleavage site, there appear to be no conserved regions among different lumen-targeted proteins reported. This, however, is not the case with plant PPOs when considering region I (the n-region; Fig. 3).

The C terminus of the transit peptide encodes amino acid sequences that allow the protein to be cleaved from the mature protein. A general conserved cleavage site motif has been proposed for all chloroplast-targeted proteins (Gavel and von Heijne, 1990). For plant PPO, this sequence only holds, in general, for the dipetide at the cleavage site (Fig. 1, residues 98 and 99). However, there appears to be a cleavage site motif specific for plant PPOs, which is dissimilar to other lumen-targeted proteins (cf. de Boer and Weisbeek, 1991, and refs. therein; Newman et al., 1993; Fig. 1). This suggests that cleavage site motifs may be specific to similar proteins rather than the more general “chloroplast motif.”

PPO transcripts of tomato were found to be highly expressed in floral tissues. Specifically, transcripts accumulated in arrested floral primordia and were highly expressed in subsequently developing floral organs (Shahar et al., 1992). In potato, however, the expression of PPO transcripts was localized to young foliage, and fully open flowers still contained detectable transcripts (Hunt et al., 1993). In the present study, the temporal and spatial expression patterns of Phytolacca PPO were analyzed by northern hybridization with PAP1 and PAP2 cDNAs as probes. Results indicated transcripts to be localized to non-vegetative regions of intact plants of Phytolacca (Figs. 4 and 5). Expression was specifically localized to ripened fruits that had turned red by a high accumulation of betalains. Betalain-containing tissues in Phytolacca are localized only to the few layers of the epidermal and subepidermal layers, with the exception of ripening fruit, in which all tissues except the seeds accumulate betalains (data not shown). Thus, transcript levels may be too low to detect in organs other than in ripening fruit.

The perpetual dilemma of work with plant PPO has always come down to the question, what is the physiological function? PPO is often referred to as an enzyme looking for a function. Possible physiological functions in a few species have been shown. The enzyme has been shown to be associated with glandular trichomes and to impart insect resistance in potato (Kowalski et al., 1992). PPO has also been shown to be associated with SI and PSII and is speculated to be involved in ATP production and O2 regulation (Tolbert, 1973; Lieberei and Biel, 1976; Lax and Vaughn, 1991). In addition to these suspected functions, results from the present study indicate that PPO is highly induced during the ripening phase of Phytolacca fruit, which has a high betalain content.

ACKNOWLEDGMENT

The authors would like to thank Dr. R. Kawaihara for his help and suggestions with the molecular cloning work.

Received July 27, 1994; accepted December 28, 1994.

Copyright Clearance Center: 0032-0889/95/107/1083/C 7.

LITERATURE CITED


Rink E, Bohm H (1985) Changed betaxanthin pattern in violet flowers of Portulaca grandiflora after the feeding of DOPA. Phytochemistry 24: 1475–1477


Tolbert N (1973) Activation of polyphenol oxidase of chloroplasts. Plant Physiol 51: 234-244
