Cloning and Expression of Cytochrome P450 Enzymes Catalyzing the Conversion of Tyrosine to p-Hydroxyphenylacetaldoxime in the Biosynthesis of Cyanogenic Glucosides in Triglochin maritima

John Strikart Nielsen and Birger Lindberg Møller*

Plant Biochemistry Laboratory, Department of Plant Biology and Center for Molecular Plant Physiology (PlaCe), The Royal Veterinary and Agricultural University, 40 Thorvaldsensvej, DK–1871 Frederiksberg C, Copenhagen, Denmark (J.S.N., B.L.M.)

Two cDNA clones encoding cytochrome P450 enzymes belonging to the CYP79 family have been isolated from *Triglochin maritima*. The two proteins show 94% sequence identity and have been designated CYP79E1 and CYP79E2. Heterologous expression of the native and the truncated forms of the two clones in *Escherichia coli* demonstrated that both encode multifunctional N-hydroxylases catalyzing the conversion of tyrosine to p-hydroxyphenylacetaldoxime in the biosynthesis of the two cyanogenic glucosides taxiphyllin and triglochinin in *T. maritima*. This renders CYP79E functionally identical to CYP79A1 from Sorghum bicolor, and unambiguously demonstrates that cyanogenic glucoside biosynthesis in *T. maritima* and *S. bicolor* is catalyzed by analogous enzyme systems with p-hydroxyphenylacetaldoxime as a free intermediate. This is in contrast to earlier reports stipulating p-hydroxyphenylacetonitrile and triglochinin in *T. maritima*. This renders CYP79E functionally identical to CYP79A1 from Sorghum bicolor, and unambiguously demonstrates that cyanogenic glucoside biosynthesis in *T. maritima* and *S. bicolor* is catalyzed by analogous enzyme systems with p-hydroxyphenylacetaldoxime as the only free intermediate in *T. maritima*. L-3,4-Dihydroxyphenyl[1-14C]Ala (DOPA) was not metabolized by CYP79E1, indicating that hydroxylation of the phenol ring at the meta position, as required for triglochinin formation, takes place at a later stage. In *S. bicolor*, CYP71E1 catalyzes the subsequent conversion of p-hydroxyphenylacetaldoxime to p-hydroxymandelonitrile. When CYP79E1 from *T. maritima* was reconstituted with CYP71E1 and NADPH-cytochrome P450 oxidoreductase from *S. bicolor*, efficient conversion of tyrosine to p-hydroxymandelonitrile was observed.

Triglochin maritima (seaside arrow grass) contains two Tyr-derived cyanogenic glucosides, triglochinin (4-carboxy-methyl-5-cyano-5-β-d-glucopyranosyloxy-cis-penta-2, 4-dienoic acid) and taxiphyllin (β-d-glucopyranosyloxy-(R)-p-hydroxymandelonitrile) (Sharples et al., 1972; Conn, 1973). The biosynthesis of taxiphyllin in *T. maritima* has been shown to proceed with N-hydroxy-Tyr, (Z)-p-hydroxyphenylacetaldoxime, p-hydroxyphenylacetoneitrile, and p-hydroxymandelonitrile as intermediates (Hösel and Nahrstedt, 1980; Cutler et al., 1981; Nielsen and Møller, 1999) (Fig. 1) and to be catalyzed by cytochrome P450 (Cyt P450) enzymes (Nielsen and Møller, 1999). In Sorghum bicolor, the same intermediates had previously been shown to be involved in the synthesis of β-d-glucopyranosyloxy-(S)-p-hydroxymandelonitrile (dhurrin), the epimer of taxiphyllin, and to be catalyzed by two multifunctional Cyt P450 enzymes, CYP79A1 (P450tyr) (Sibbesen et al., 1994, 1995; Koch et al., 1995) and CYP71E1 (P450ox) (Kahn et al., 1997; Bak et al., 1998a).

Additional gene sequences encoding Cyt P450s belonging to the CYP79 family are available (http://drnelson.utmem.edu/CytochromeP450.html) from glucosinolate-producing plant species and are thought to catalyze a similar set of reactions in glucosinolate biosynthesis (Bak et al., 1998b). However, except for CYP79A1 from *S. bicolor*, the precise catalytic properties of the CYP79 members remain elusive, because neither reconstitution of isolated proteins nor functional expression has been achieved. In *S. bicolor*, CYP79A1 catalyzes the conversion of Tyr to p-hydroxyphenylacetaldoxime, whereas CYP71E1 catalyzes the conversion of (Z)-p-hydroxyphenylacetaldoxime to p-hydroxymandelonitrile. The involvement of two multifunctional enzymes explains why this pathway is highly channeled in *S. bicolor* with p-hydroxyphenylacetaldoxime as the only intermediate in equilibrium with the exogenously added substrate (Møller and Conn, 1980; Halkier et al., 1989).

Based on a common biosynthetic route for cyanogenic glucoside biosynthesis in a number of different plant species (Cutler and Conn, 1981; Collinge and Hughes, 1982; Koch et al., 1992), it is tempting to infer the involvement of analogous multifunctional Cyt P450s. However, this may not be so in *T. maritima*, since in this plant, p-hydroxyphenylacetaldoxime and p-hydroxyphenylacetonitrile are free to equilibrate (Cutler et al., 1981; Nielsen and Møller, 1999). The Cyt P450-catalyzed conversion of aldoxime to nitrile is a dehydration reaction and, as such, is non-classical. In *T. maritima* it could be carried out as an additional enzyme activity associated with the first multifunctional Cyt P450 enzyme instead of being the first catalytic event catalyzed by the second Cyt P450 involved. If so, the second Cyt P450 in *T. maritima* would be a normal C-hydroxylase.

To discriminate between these possibilities, cloning of a putative Cyt P450 enzyme from *T. maritima* that is analo-
gous to CYP79A1 from *S. bicolor* was initiated based on a PCR approach. We report the isolation and functional reconstitution of two cDNA clones from *T. maritima*, designated CYP79E1 and CYP79E2, encoding two multifunctional Cyt P450s catalyzing the conversion of L-Tyr to (Z)-p-hydroxyphenylacetaldoxime. A CYP71E1 homolog is thought to catalyze the conversion of (Z)-p-hydroxyphenylacetaldoxime to p-hydroxymandelonitrile. p-Hydroxyphenylacetonitrile is thought to constitute the branch point between taxiphyllin and triglochin synthesis. Dhurrin is shown for comparison.

**Figure 1.** Putative biosynthetic pathway for taxiphyllin and triglochin, the two cyanogenic glucosides in *T. maritima*. In the present study, CYP79E1 and CYP79E2 were isolated and shown to catalyze the conversion of L-Tyr to (Z)-p-hydroxyphenylacetaldoxime. A CYP71E1 homolog is thought to catalyze the conversion of (Z)-p-hydroxyphenylacetaldoxime to p-hydroxymandelonitrile. p-Hydroxyphenylacetonitrile is thought to constitute the branch point between taxiphyllin and triglochin synthesis. Dhurrin is shown for comparison.

**MATERIALS AND METHODS**

**PCR Approach to Generate cDNA Fragments of a CYP79 Homolog in *Triglochin maritima***

A unidirectional plasmid cDNA library was made by Invitrogen (Carlsbad, CA) from flowers and fruits (schizocarp) of *T. maritima*, using the expression vector pcDNA2.1, which contains the *lacZ* promoter. Plant material was collected at Aflandshage on Southern Amager, at the coast of Øresund, Denmark, frozen directly in liquid N₂, and stored at −80°C.
Degenerate PCR primers were designed based on conserved amino acid sequences in CYP79A1 derived from *Sorghum bicolor* (GenEMBL no. U32624; Koch et al., 1995), CYP79B1 from *Sinapis alba* (GenEMBL no. AF069494; Bak et al., 1998b), CYP79B2 from Arabidopsis (GenEMBL no. AF069495; Bak et al., 1998b), and PCR fragment of CYP79D1 from *Manihot esculenta* (GenEMBL no. AF140613; Andersen et al., 2000) (Fig. 2). Two rounds of PCR amplification reactions (total volume: 50 μL) were carried out using 100 pmol of each primer, 5% (v/v) dimethyl sulfoxide, 200 μM dNTPs, and 2.5 units of Taq DNA polymerase in PCR buffer (50 mM KCl, 10 mM Tris-HCl, pH 8.8, 1.5 mM MgCl₂, and 0.1% [v/v] Triton X-100). Thermal cycling parameters were 2 min at 95°C, 30 cycles (5 s at 95°C, 30 s at 45°C, and 45 s at 72°C), and finally 5 min at 72°C. The first PCR reaction was performed using primer 1F and 1R (Table I) on 100 ng of template DNA prepared from the cDNA library, or genomic DNA prepared using an extraction kit (Nucleon Phytopure Plant DNA Extraction Kit, Amersham, Buckinghamshire, UK).

The PCR products were purified using a kit (QIAquick, Qiagen, Valencia, CA), eluted in 30 μL of 10 mM Tris-HCl, pH 8.5, and used as template (1 μL) for the second round of PCR reactions carried out using PCR fragments derived from both cDNA and genomic DNA and using the two degenerate primers 2F and 2R (Table I). An aliquot (5 μL) of the PCR reaction was applied to a 1.5% (w/v) agarose/TBE gel, and a band of the expected size (approximately 200 bp) was observed using both cDNA and genomic DNA as template. The rest of the PCR reaction was purified using the QIAquick kit, and eluted in 30 μL Lo bet. The purified PCR fragments (5 μL) were digested with EcoRI and BamHI, excised from a 1.5% (w/v) agarose/TBE gel, purified using an agarose gel extraction kit (QIAEX II, Qiagen), and ligated into an EcoRI- and BamHI-digested pBluescript II SK vector (Stratagene, La Jolla, CA). Seven clones derived from the cDNA library and three clones derived from genomic DNA were sequenced (ALF Express, Pharmacia, Piscataway, NJ) using the fluorescent-labeled primer cycle sequencing kit with 7-deaza dGTP (Thermo Sequenase, Amersham). All sequence analyses were performed using programs in the GCG Wisconsin Sequence Analysis package (Genetics Computer Group, Madison, WI).

**Screening of a Plasmid cDNA Library Made from Flowers and Fruits of *T. maritima***

Both cDNA and genomic DNA produced an identical PCR fragment with high sequence resemblance to the other known CYP79s. The PCR fragment was used as template to generate a 350-bp digoxigenin-11-dUTP-labeled (Boehringer Mannheim, Basel) probe (TRI1) by PCR using T3 and T7 primers. The labeled probe was employed to screen 660,000 colonies of the pcDNA2.1 cDNA library. Hybridizations were carried out overnight at 68°C in 5× SSC (0.75 M NaCl and 75 mM sodium citrate, pH 7.0), 0.1% (w/v) N-lauroylsarcosine, 0.02% (w/v) SDS, and 1% (w/v) blocking reagent (Boehringer Mannheim). Membranes were washed twice under high-stringency conditions (65°C, 0.1× SSC, and 0.1% [w/v] SDS), incubated with anti-digoxigenin-AP, and developed using 5-bromo-4-chloro-3-indolylphosphate and nitroblue tetrazolium according to the manufacturer's (Boehringer Mannheim) instructions. Positive colonies were rescreened under the same conditions, and single positive colonies were sequenced and analyzed.

**PCR Approach to Design 5′ End Probes to Screen for Full-Length Clones***

Library screening resulted in two very similar partial clones designated #1 and #2, that particularly differ in their...
N-terminal sequence (Fig. 2). To isolate the corresponding full-length clones from the pcDNA2.1 library, two consecutive PCR reactions were performed using the same PCR conditions as above, with the exception that the annealing temperature was set at 55°C. The first PCR reaction was performed with primers 3F and 3R (Table I) using 100 ng of DNA from the cDNA library as template. The purified PCR products from the first PCR reaction were used as template (1 μL) for a second round of PCR reactions using primer 4R#1 or 4R#2 against primer 3F (Table I). The PCR fragments from the second reaction were excised from the gel, purified on the agarose gel (w/v) agarose/TBE gel, and the slowest migrating bands were used to cut out a fragment (1,045 bp) of the middle portion of the constructs that derive from PCR were analogous fragment from the cDNA clone. The remaining portions of the PCR clones, which was then replaced with the full-length construct was synthesized using primers 4R#1 and 4R#2 against primer 3F (Table I). The PCR fragments were digested with NdeI and HindIII and ligated into NdeI- and HindIII-digested pSP19G10L expression vector (Barnes, 1996). The unique restriction sites NcoI and PmlI were used to cut out a fragment (1,045 bp) of the middle part of the PCR clones, which was then replaced with the analogous fragment from the cDNA clone. The remaining portions of the constructs that derived from PCR were sequenced to exclude PCR errors. Clone #2 was found to be in frame with the first 24 amino acids of the lacZ gene. A similar clone of #1, in frame with lacZ at the same nucleotide position as in #2, was also expressed in pcDNA2.1 for comparison.

Expression in Escherichia coli

All expression constructs were transformed into the E. coli strains JM109 (Stratagene) and XL-1 Blue (Stratagene). The native and Δ(1–52)2E1(10aa) constructs of clone #1 and the construct of clone #2 were co-transformed with pSBET (Schenk et al., 1995), which encodes a tRNA gene for rare Arg codons, into JM109. Single colonies were grown overnight in Luria-Bertani medium (50 μg/mL ampicillin, 37°C, 225 rpm) and used to inoculate 100X volume of modified TB medium (50 μg/mL ampicillin, 1 mM thiamine, 75 μg/mL δ-amino-levulinic acid, 1 mM isopropyl β-D-thiogalactopyranoside [IPTG], 28°C, 125 rpm, 48 h).

Expression Constructs

Three different constructs of clone #1 were generated with PCR, using Pwo polymerase (Boehringer Mannheim) to introduce a NdeI restriction site at the start codon and a HindIII restriction site immediately after the stop codon. A full-length construct was synthesized using primers 6F#1(na) and 6R#1. Two truncated constructs were made using primers 6F#1[Δ(1–31)17,na(8aa)] and 6R#1 or primers 6F#1[Δ(1–52)2E1(10aa)] and 6R#1 (Table I). The PCR fragments were digested with NdeI and HindIII and ligated into NdeI- and HindIII-digested pSP19g10L expression vector (Barnes, 1996). The unique restriction sites NcoI and PmlI were used to cut out a fragment (1,045 bp) of the middle part of the PCR clones, which was then replaced with the analogous fragment from the cDNA clone. The remaining portions of the constructs that derived from PCR were sequenced to exclude PCR errors. Clone #2 was found to be in frame with the first 24 amino acids of the lacZ gene. A similar clone of #1, in frame with lacZ at the same nucleotide position as in #2, was also expressed in pcDNA2.1 for comparison.

Table 1. Compilation of primers used

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Amino Acid Sequence</th>
<th>Nucleotide Sequence*</th>
<th>Restriction Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Fb</td>
<td>DPNPSNA</td>
<td>GCGGAATTCGAYAAYCWCWSIAAAGGC</td>
<td>EcoRI</td>
</tr>
<tr>
<td>1Rb</td>
<td>FNV/LPHVA</td>
<td>GCGGATCCGCAACTGGTGGAAHTRRAA</td>
<td>BamHI</td>
</tr>
<tr>
<td>2F</td>
<td>SNAVEW</td>
<td>GCGGAATTCGCAAGYCTGGTGGATGG</td>
<td>EcoRI</td>
</tr>
<tr>
<td>2R</td>
<td>HPVAXFEN</td>
<td>GCGGATCCTRARAAAGYCTGGTGG</td>
<td>BamHI</td>
</tr>
<tr>
<td>3F</td>
<td></td>
<td>GCGGAATTCGACAGAAGAAAAAGCTGGAC</td>
<td>EcoRI</td>
</tr>
<tr>
<td>3R</td>
<td>VVTYRSS</td>
<td>GCGGATCCAGAGGATTGACCGAACAAC</td>
<td>BamHI</td>
</tr>
<tr>
<td>4R#1f</td>
<td>TVLFLL</td>
<td>GCGGATCCAGAGGACAGTACT</td>
<td>BamHI</td>
</tr>
<tr>
<td>4R#2f</td>
<td>ATLFLL</td>
<td>GCGGATCCAGAGGACAGTACT</td>
<td>BamHI</td>
</tr>
<tr>
<td>5F#1</td>
<td></td>
<td>GCGGATCCAGAGGACAGTACT</td>
<td>BamHI</td>
</tr>
<tr>
<td>5R#1</td>
<td>MELITI</td>
<td>GCGGATCCAGAGGACAGTACT</td>
<td>BamHI</td>
</tr>
<tr>
<td>6F#1</td>
<td>MELLITIL</td>
<td>GCGGATCCAGAGGACAGTACT</td>
<td>NdeI</td>
</tr>
<tr>
<td>6R</td>
<td>LLQSF</td>
<td>GCGGATCCAGAGGACAGTACT</td>
<td>HindIII</td>
</tr>
<tr>
<td>6F#1[Δ(1–31)17,na(8aa)]</td>
<td></td>
<td>MALLAVFLFLFKQ</td>
<td>NdeI</td>
</tr>
<tr>
<td>6F#1[Δ(1–52)2E1(10aa)]</td>
<td></td>
<td>MARQVHSSWNLPPGP</td>
<td>NdeI</td>
</tr>
</tbody>
</table>

* The sequence is shown from 5’ end to 3’ end.  F, Forward primer; R, reverse primer.  Amino acid consensus sequence based on alignment in Fig. 2 and used for primer design.  A specific primer for pcDNA2.1 placed just upstream the insertion site of the 5’ end of the cDNA library.  Covers a sequence that is identical in the two clones #1 and #2.  Covers a sequence that is specific for either of the two clones #1 and #2.  The asterisk indicates a stop codon.

Measurements of Expression Levels and Biosynthetic Activities

Expression levels of the different constructs were determined by CO difference spectroscopy and quantified using
an extinction coefficient $\epsilon_{450-490}$ of 91 mm$^{-1}$ cm$^{-1}$ (Omura and Sato, 1964). Spectra were made from 100 or 500 µL of whole E. coli cells or using the rich phases from Triton X-114 phase partitioning solubilized in 50 mm KH$_2$PO$_4$/K$_2$HPO$_4$, pH 7.5, 2 mM EDTA, 20% (v/v) glycerol, and 0.2% (v/v) Triton X-100 (total volume: 1 mL). E. coli cells for in vivo studies were prepared by centrifugation (2 min and 30 s at 7,000 g) of 1 mL of cell culture and resuspension in 100 µL of 50 mM N-[2-hydroxy-1,1-Bis(hydroxymethyl)ethyl]glycine (Tricine), pH 7.9, and 1% L-phenylmethylsulfonyl fluoride. For in vitro studies, spheroplasts were made from E. coli (JM109) cells expressing native or $\Delta$(1–52)$_{2E1}$ (I0ao) constructs of clone #1 or the construct of clone #2, followed by temperature-induced phase partitioning (0.6% [v/v] Triton X-114 and 30% [v/v] glycerol) as previously described (Halkier et al., 1995).

Measurements of in vivo catalytic activity were carried out by the administration of [U-14C]Tyr (0.35 µCi, 7.39 µM), p-hydroxyphenylacetaldoxime (0 or 0.1 mM), or p-hydroxyphenylacetonitrile (0 or 0.1 mM) to resuspended E. coli cells (100 µL). In vitro activities were measured in reconstitution experiments using the rich phase from phase partitioning. A standard reaction mixture (total volume: 50 µL) contained 5 µL of rich phase, 0.375 unit of S. bicolor NADPH-Cyt P450 oxidoreductase, 5 µL of L-α-dilauroyl phosphatidylcholine (DLPC), 0.6 mM NADPH, and 14 mM KH$_2$PO$_4$/K$_2$HPO$_4$, pH 7.9. The following substrates were tested: L-[U-14C]Tyr (0.20 µCi, 9.04 µM), L-[U-14C]Phe (0.20 µCi, 8.8 µM), and L-3,4-dihydroxyphenyl[3-14C]Ala (0.20 µCi, 400 µM). L-[U-14C]Tyr (0.20 µCi, 9.04 µM) was also tested in reconstitution experiments that included purified CYP71E1 (Kahn et al., 1997; Bak et al., 1998a) from S. bicolor.

Incubations (1 h, 30°C, shaking water bath) were started by the addition of substrate (in vivo experiments) or NADPH (in vitro experiments), and were stopped by the addition of ethyl acetate. Biosynthetic activity was monitored by the formation of radioactive products using thin-layer chromatography (TLC) analysis, as previously described (Møller and Conn, 1979), and detection and quantification using a phosphor imager (Storm 840, Molecular Dynamics, Sunnyvale, CA). Before TLC application, the sample was extracted with ethyl acetate. During this step, the surplus of radiolabeled Tyr remains in the aqueous phase, thus preventing overexposure at the origin. The total ethyl acetate phase was applied to the TLC. In some experiments, inevitable carry-over of small amounts of the aqueous phase results in the appearance of a Tyr band at the origin. Unlabeled reference compounds (p-hydroxyphenylacetaldoxime, p-hydroxyphenylacetonitrile, and p-hydroxybenzaldehyde) were prestreaked on the TLC plates to permit visual detection under UV light.

### RESULTS

#### Cloning of CYP79E1 and CYP79E2

A sequence alignment of CYP79A1 and putative N-hydroxylases belonging to the CYP79 family is shown on Figure 2. Based on this alignment, four degenerate oligonucleotide primers covering two CYP79-specific regions were designed (1F, 2F, 1R, and 2R) (Fig. 2) and employed in nested PCR reactions using genomic DNA and cDNA made from flowers and fruits of T. maritima as templates. A PCR fragment of the expected size (approximately 200 bp) and showing 62% to 70% identity to CYP79 sequences at the amino acid level was isolated using both templates and employed to screen the cDNA library. Two clones, denoted #1 and #2, were isolated and verified by sequence comparison to share high sequence identity with the CYP79 family. Using clone-specific PCR primers, a full-length clone corresponding to #1 was isolated. The open reading frame encodes a protein with a molecular mass of 60.8 kDa. A comparison of the full-length sequence of clone #1 with that of clone #2 demonstrated that clone #2 is 6 bp shorter at the 5’ end, but contains a Met codon not found in clone #1 at a position corresponding to amino acid residue 26 specified by clone #1. The surrounding sequence of this Met codon did not fit the general context sequence for a start codon in a monocotyledonous plant (Joshi et al., 1997). Most likely, clone #2 lacks 6 bp to be full-length.

Cyt P450 enzymes are heme-containing enzymes constituting a supergene family. In plants, they are divided into two distinct groups (Durst and Nelson, 1995): The A-group probably derived from a common ancestor and is involved in the biosynthesis of secondary plant products such as cyanogenic glucosides and glucosinolates. The non-A-group is heterogeneous and clusters near to animal, fungal, and microbial Cyt P450s. Cyt P450s showing amino acid sequence identities above 40% are grouped within the same family (Nelson et al., 1993). Cyt P450s showing more than 55% identity belong to the same subfamily. The Cyt P450s encoded by clones #1 and #2 showed 45% to 49% identity, respectively, to the previously known members of the CYP79 family (Table II) and accordingly were identified as the first two members of a new subfamily and assigned as CYP79E1 (accession no. AF140609) and CYP79E2 (accession no. AF140610), respectively, by the International Cy-

| Table II. Percent identity and similarity between six members of the CYP79 family |
|-----------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Identity                  | CYP79E1 | CYP79E2 | CYP79A1 | CYP79B1 | CYP79B2 | CYP79D1 |
| CYP79E1                   | 95.2    | 61.7    | 58.1    | 58.9    | 60.0    |
| CYP79E2                   | 94.1    | 61.5    | 57.6    | 58.5    | 59.2    |
| CYP79A1                   | 48.8    | 48.8    | 65.5    | 67.1    | 65.8    |
| CYP79B1                   | 44.9    | 44.9    | 51.3    | 92.3    | 65.1    |
| CYP79B2                   | 44.5    | 44.6    | 52.6    | 89.3    | 67.3    |
| CYP79D1                   | 46.4    | 46.5    | 51.5    | 49.1    | 50.7    |
Expression of CYP79E1 and CYP79E2 in E. coli

The expression vector pSP19g10L was used for expression of CYP79E1 and CYP79E2 constructs in E. coli. This expression vector contains the lacZ promoter fused with the short leader sequence of gene 10 from T7 bacteriophage (g10L) and has been shown effective for heterologous protein expression in E. coli (Olins and Rangwala, 1990). In Cyt P450s, increased expression levels have been obtained by modifying the 5′ end of the open reading frame to increase the content of A’s and T’s (Stormo et al., 1982; Schauder and McCarthy, 1989; Barnes et al., 1991) and by replacement of a number of codons at the 5′ end with codons specifying the N-terminal sequence of bovine P45017α (Barnes et al., 1991) or human P4502E1 or 2D6 (Gillum et al., 1994, 1995). To take advantage of this knowledge, a number of different constructs were made (Fig. 3). The first construct, CYP79E1na, encoded native CYP79E1 with silent mutations introduced at codons 3 and 5 to increase the AT content. The second construct, CYP79E1Δ (1–31)17α(8aa), encoded a truncated form of CYP79E1 in which 31 codons of the native 5′ sequence had been replaced by eight AT-enriched codons of P45017α (Barnes et al., 1991; Halkier et al., 1995). In the third construct, CYP79E1Δ (1–52)2E1(10aa), the first 52 codons of the native 5′ sequence were replaced by 10 AT-enriched codons of P4502E1, and silent mutations were introduced in codons 53 and 55. Because the CYP79E2 clone was isolated in frame with the first 24 codons of the lacZ gene in the vector pcdNA2.1, this was tested as a fourth expression construct designated CYP79E2lacZ(24aa) (Fig. 3). For comparison, an equivalent fifth construct, CYP79E1Δ (1–2)lacZ(24aa), was also prepared.

All constructs contained the original stop sequence, TAAT, which is found in most highly expressed E. coli genes. A strong positive correlation between the content of A and T (Stormo et al., 1982; Schauder et al., 1982; Schauder et al., 1983) and expression levels has been documented. To further optimize the expression, all constructs were tested in two strains, XL-1 Blue and JM109. The JM109 strain was most efficient. CYP79E1 and CYP79E2 contain 19 and 17 AGA or AGG Arg codons, which are rare in E. coli genes. A strong positive correlation between the occurrence of codons and rRNA content has been established (Ikemura, 1981). Accordingly, co-transformation with pSBET, which encodes a tRNA for rare Arg, was also pursued.

Carbon monoxide binding spectra using intact E. coli cells showed the absorption maximum at 450 nm, which is diagnostic for the formation of functional Cyt P450 with the following three constructs: CYP79E1na, CYP79E1Δ (1–52)2E1(10aa), and CYP79E2lacZ(24aa). The spectra were obtained without and with co-transformation with pSBET, but in all cases the Cyt P450 content was too low to permit quantification. To obtain an accurate determination, the Cyt P450s were enriched by isolation of E. coli spheroplasts, followed by temperature-induced Triton X-114 phase partitioning (Werck-Reichart et al., 1991; Halkier et al., 1995). The highest expression level (JM109 cells, 48 h) was 56 nmol/L culture and was obtained using CYP79E2lacZ(24aa) (Fig. 4). This level is comparable to the expression level of 62 nmol/L culture obtained with the S. biolor construct CYP79A1Δ (1–33)17α(8aa) (Halkier et al., 1995), which was included as a positive control. CYP79E1Δ (1–31)17α(8aa) with a modified P45017α N terminus and the empty vector control never revealed any detectable spectrum.

Determination of the Enzymatic Activity of CYP79E1 and CYP79E2

Radiolabeled Tyr was administered to E. coli cells containing all five CYP79E1 and CYP79E2 constructs described above as well as the S. biolor construct CYP79A1Δ (1–33)17α(8aa) as a positive control. In all cases, except for CYP79E1Δ (1–31)17α(8aa), p-hydroxyphenylacetaldoxime was used as the sole product (data not shown). This demonstrates that CYP79E1 and CYP79E2 are functionally equivalent to CYP79A1 isolated from S. biolor, in spite of earlier reports of p-hydroxyphenylacetanitirile as the only free intermediate in T. maritima (Hösel and Nahrstedt, 1980; Cutler et al., 1981). A detailed analysis of the expression levels obtained with the different constructs provided the following information: (a) The CYP79E1Δ (1–2)lacZ(24aa) and CYP79E2 lacZ(24aa) Constructs provided the highest activity; (b) for all constructs, JM109 cells were superior to XL-1 Blue; (c) in time course studies (0–72 h), the highest activity was obtained at 24 to 48 h; (d) omission of IPTG reduced the expression level to approximately 50%, as documented using the CYP79E1na construct; (e) co-transformation with pSBET further increased the high expression level obtained with CYP79E2lacZ(24aa), whereas the expression level obtained with all other constructs diminished.

For in vitro studies, the recombinant Cyt P450s obtained by expression of CYP79E1na, CYP79E1Δ (1–52)2E1(10aa), and CYP79E2lacZ(24aa) were partially purified by temperature-induced Triton X-114 phase partitioning (Werck-Reichart et al., 1991; Halkier et al., 1995) and reconstituted using S. biolor NADPH-Cyt P450 oxidoreductase and DLPC (Fig. 5). As demonstrated above using intact E. coli cells, Tyr is oxidized to phenolic acid, and the oxidation of Tyr was monitored by a diagnostic for the formation of functional Cyt P450 with the following three constructs: CYP79E1na, CYP79E1Δ (1–52)2E1(10aa), and CYP79E2lacZ(24aa). The spectra were obtained without and with co-transformation with pSBET, but in all cases the Cyt P450 content was too low to permit quantification. To obtain an accurate determination, the Cyt P450s were enriched by isolation of E. coli spheroplasts, followed by temperature-induced Triton X-114 phase partitioning (Werck-Reichart et al., 1991; Halkier et al., 1995). The highest expression level (JM109 cells, 48 h) was 56 nmol/L culture and was obtained using CYP79E2lacZ(24aa) (Fig. 4). This level is comparable to the expression level of 62 nmol/L culture obtained with the S. biolor construct CYP79A1Δ (1–33)17α(8aa) (Halkier et al., 1995), which was included as a positive control. CYP79E1Δ (1–31)17α(8aa) with a modified P45017α N terminus and the empty vector control never revealed any detectable spectrum.

Determination of the Enzymatic Activity of CYP79E1 and CYP79E2

Radiolabeled Tyr was administered to E. coli cells containing all five CYP79E1 and CYP79E2 constructs described above as well as the S. biolor construct CYP79A1Δ (1–33)17α(8aa) as a positive control. In all cases, except for CYP79E1Δ (1–31)17α(8aa), p-hydroxyphenylacetaldoxime was used as the sole product (data not shown). This demonstrates that CYP79E1 and CYP79E2 are functionally equivalent to CYP79A1 isolated from S. biolor, in spite of earlier reports of p-hydroxyphenylacetanitirile as the only free intermediate in T. maritima (Hösel and Nahrstedt, 1980; Cutler et al., 1981). A detailed analysis of the expression levels obtained with the different constructs provided the following information: (a) The CYP79E1Δ (1–2)lacZ(24aa) and CYP79E2 lacZ(24aa) Constructs provided the highest activity; (b) for all constructs, JM109 cells were superior to XL-1 Blue; (c) in time course studies (0–72 h), the highest activity was obtained at 24 to 48 h; (d) omission of IPTG reduced the expression level to approximately 50%, as documented using the CYP79E1na construct; (e) co-transformation with pSBET further increased the high expression level obtained with CYP79E2lacZ(24aa), whereas the expression level obtained with all other constructs diminished.

For in vitro studies, the recombinant Cyt P450s obtained by expression of CYP79E1na, CYP79E1Δ (1–52)2E1(10aa), and CYP79E2lacZ(24aa) were partially purified by temperature-induced Triton X-114 phase partitioning (Werck-Reichart et al., 1991; Halkier et al., 1995) and reconstituted using S. biolor NADPH-Cyt P450 oxidoreductase and DLPC (Fig. 5). As demonstrated above using intact E. coli cells, Tyr is oxidized to phenolic acid, and the oxidation of Tyr was monitored by

![Figure 3. N-Terminal amino acid sequences of the four different constructs of CYP79E1 and the construct of CYP79E2. Native sequence from CYP79E is underlined.](image-url)
metabolized to \(p\)-hydroxyphenylacetaldoxime by reconstituted recombinant CYP79E1 and CYP79E2. No \(p\)-hydroxyphenylacetaldoxime accumulates using an empty vector control or in reconstitutions without the addition of reductase. In the experiment without administration of NADPH, a small amount of \(p\)-hydroxyphenylacetaldoxime accumulated (Fig. 5, lane 1). Using the constructs CYP79E2\(_{\text{lacZ(24aa)}}\) and \(S. \text{bicolor}\) CYP79A1\(\Delta \ (1-33)_{17n(8aa)}\), the turnover of Tyr per nanogram of Cyt P450 obtained was about the same.

Because of the lack of accurate CO difference spectra of CYP79E1\(\text{na}\) and CYP79E1\(\ (1-52)_{2E1(10aa)}\) from the in vitro and in vivo experiments, it was not possible to measure the turnover of Tyr per nanogram of Cyt P450. The relative values for turnover of Tyr per milligram of protein are: CYP79E1\(\ (1-52)_{2E1(10aa)}\) = 1; CYP79E1\(\text{na}\) = 1.5; CYP79E2\(_{\text{lacZ(24aa)}}\) = 22; \(S. \text{bicolor}\) CYP79A1\(\Delta \ (1-33)_{17n(8aa)}\) = 46. As already stated above, CYP79E1\(\ (1-31)_{17n(8aa)}\) showed no activity at all. In CYP79E2\(_{\text{lacZ(24aa)}}\) and CYP79A1\(\Delta \ (1-33)_{17n(8aa)}\), accumulation of \(p\)-hydroxyphenylacetonitrile could be detected (Fig. 5). No catalytic activity could be detected using DOPA or \(l\)-[\(U-^{14}\text{C}\)]Phe as the substrate in reconstitution experiments (data not shown). Administration of DOPA to microsomal preparations from \(T. \text{maritima}\) flowers and fruits also did not show any measurable activity.

Reconstitution of CYP79E with CYP71E1

Reconstitution of the membrane-associated pathway of cyanogenic glucoside synthesis resulting in the formation of \(p\)-hydroxymandelonitrile (the aglycon of dhurrin, seen as \(p\)-hydroxybenzaldehyde in vitro) was achieved using enzymes from the two species \(S. \text{bicolor}\) and \(T. \text{maritima}\) (Fig. 6). In all reconstitutions, including Tyr, NADPH, NADPH-Cyt P450 oxidoreductase, CYP79E1 or CYP79E2, and CYP71E1, considerable amounts of \(p\)-hydroxyphenylacetonitrile and \(p\)-hydroxybenzaldehyde accumulated (Fig. 6).

**DISCUSSION**

The biosynthesis from amino acids of the aglycon part of cyanogenic glucosides has been shown to involve Cyt P450 enzymes in \(S. \text{bicolor}\) (Halkier and Møller, 1991), \(M. \text{esculenta}\) (Koch et al., 1992), and \(T. \text{maritima}\) (Nielsen and Møller, 1999) (Fig. 1). In \(S. \text{bicolor}\), two multifunctional Cyt P450 enzymes have been shown to catalyze all of the
membrane-associated steps leading to the aglycon of dhurrin. Tyr is converted to (Z)-p-hydroxyphenylacetaldoxime by CYP79A1 (Bibbesen et al., 1994, 1995; Koch et al., 1995), and (Z)-p-hydroxyphenylacetaldoxime is subsequently converted to p-hydroxymandelonitrile by CYP71E1 (Kahn et al., 1997; Bak et al., 1998a). In M. esculenta, a CYP79 homolog, CYP79D1, has recently been identified and characterized (Andersen et al., 2000). Also, the biosynthetic pathway of at least some glucosinolates involves Cyt P450 enzymes in the conversion of the parent amino acid to the corresponding aldoxime (Du et al., 1995). CYP79 homologs have been identified in the two glucosinolate-producing plants, S. alba (CYP79B1) and Arabidopsis (CYP79B2), but so far no successful expression of the isolated clones has been achieved (Bak et al., 1998b).

The CYP79B1 clone was isolated by screening an S. alba cDNA library with an expressed sequence tag (CYP79B2) probe from Arabidopsis (Bak et al., 1998b). Both plants are members of the Brassicaceae family, but contain glucosinolates derived from different substrates (Hogge et al., 1988; Du et al., 1995). Despite this difference, the two full-length clones showed 89% identity at the amino acid level (Table II). Accordingly, the first attempts to isolate a CYP79 homolog from T. maritima was carried out using heterologous probes from S. bicolor to take advantage of the fact that both plants are monocotyledonous and use Tyr as substrate. However, no CYP79 homolog could be isolated using this strategy. Instead, we used a PCR approach, taking advantage of the presence of some highly conserved CYP79-specific regions covering part of the I-helix (primers 1F and 2F), β6-1 and β1-4 (primers 1R and 2R) (Fig. 2). These regions are involved in substrate recognition (SRS) (Gotth, 1992) as well as heme binding, and accordingly were thought to be less variable than other SRSs (Hasemann et al., 1995) and therefore suitable for the design of degenerate primers. Conserved regions within groups and families of Cyt P450s have previously been used successfully to design degenerate primers for the amplification of specific Cyt P450 clones (Holton et al., 1993; Meijer et al., 1993; Frank et al., 1996; Bak et al., 1998b).

The major PCR product obtained using cDNA and genomic DNA revealed 62% to 70% identity to other members of the CYP79 family. This confirmed that the primers were highly specific toward the CYP79 family, as illustrated with primer 1F (DNPSNA) covering a region that in the vast majority of other Cyt P450s contains a highly conserved Thr residue (Hasemann et al., 1995) instead of Asn. Only one of the degenerate primers (1R), covering FN(V/L)PHVA, did not anneal perfectly with the sequences subsequently obtained from T. maritima, having a single nucleotide difference at the 5’ end of the primer. Here the Ala codon in the primer was replaced by a Ser codon in the T. maritima sequences. The PCR fragment was used as a probe to screen the cDNA library made from T. maritima flowers and fruits. Microsomal preparations isolated from these tissues had previously been shown to have high catalytic activity compared with other tissues (Cutler et al., 1981; Nielsen and Møller, 1999), thus raising the abundance of the putative mRNA. Two partial CYP79 homologs were isolated and used to isolate a full-length and a nearly full-length clone designated CYP79E1 and CYP79E2, respectively, as the first members of a new subfamily (Fig. 2). These clones show 94% identity (Table II).

CYP79E1 showed 45% to 49% identity to other members of the CYP79 family (Table II). In sequence alignment to other members of the CYP79 family, CYP79E1 showed the highest identity and similarity to CYP79A1 (Table II). In contrast, CYP79A1 showed higher identity/similarity toward the three other sequences, CYP79B1, CYP79B2, and CYP79D1, that were all isolated from dicotyledonous plants (Table II). This is surprising considering that CYP79B1 and CYP79B2 are putatively involved in the biosynthesis of glucosinolates, whereas CYP79D1 catalyzes the conversion of aliphatic amino acids. The very high degree of identity between CYP79B1 and CYP79B2 (89%) is not seen between the three clones involved in the biosynthesis of cyanogenic glucosides, explaining the lack of success using a heterologous probe from S. bicolor for screening the T. maritima library.

Cyt P450s are thought to possess the same tertiary structure (Hasemann et al., 1995; Peterson and Graham, 1998); nevertheless, they only exhibit one highly conserved sequence, FFXGXRXXG (X being any amino acid), harboring the heme-binding Cys residue (Nelson et al., 1993). When restricted groups or families of Cyt P450s are aligned, additional conserved amino acid residues are found, such as the proposed highly conserved A-group heme-binding consensus sequence, PFGXXRXXG (Durst and Nelson, 1995; Bak et al., 1998b), and the CYP79-specific sequence, SFSTG(K/R)GC(A/I)A (Bak et al., 1998b). The CYP79 family diverges from the A-group heme-binding consensus sequence and within the CYP79...
family, CYP79E diverges in several positions from otherwise conserved CYP79 residues (Fig. 2). A notable difference is that CYP79E1, CYP79E2, and CYP79A2 (GenEMBL no. AB010692 comp [11,000–13,200 region]) contain the generally highly conserved Gly residue positioned two amino acids downstream from the heme-binding Cys residue (amino acid 479 in CYP79E1), whereas all the other CYP79s contain Ala at this position. This Gly residue is positioned closely to the heme plain and allows a sharp turn from the Cys-pocket into the L-helix (Hasemann et al., 1995).

In the highly conserved PERF region within microsomal Cyt P450s (448–451 in CYP79E1), the Phe residue has been replaced by His, as is also observed in other CYP79s (Bak et al., 1998b). The K-helix contains a partly conserved region, KETLR (392–396 in CYP79E1), in many Cyt P450s, with the Glu and the Arg residues being particularly conserved (Hasemann et al., 1995). In the CYP79 family, Lys has been replaced by Arg, Thr has been replaced by Ala, and Leu has been replaced with Phe, with the exception that CYP79A2 still contains Lys and CYP79E1 still contains Leu. Only minor differences are observed between the two different clones isolated from T. maritima. The region of positive charges preceding the Pro-rich region contains a repeat of 3× KS in CYP79E1, whereas CYP79E2 contains the sequence KPKS in the same region. Finally, the amino acids EGR (296–298 in CYP79E1) are deleted in CYP79E2, even though the Gly is conserved in all other CYP79s. The differences both within the CYP79 family and in relation to residues that are more or less conserved within the A-group or microsomal P450s in general illustrate how difficult it is, based on sequence alignments, to predict amino acids important for different functions.

All eukaryotic Cyt P450 enzymes found so far, except CYP55A (Park et al., 1997), are membrane bound. Some are localized in the mitochondrial membranes, while most, including plant Cyt P450s, are localized in the endoplasmic reticulum. The N terminus of the microsomal Cyt P450s functions as a hydrophobic signal sequence directing the proteins to the endoplasmic reticulum (Sakaguchi et al., 1987). The CYP79Es from T. maritima contains the four structurally conserved domains in the N terminus (Halkier et al., 1995). CYP79E2s have a short stretch of four uncharged amino acids between the positively charged region and the Pro-rich region, as opposed to CYP79A1, which has a long stretch of 18 uncharged amino acids at this position. It has previously been suggested that a long stretch of uncharged residues serves to facilitate the access of the rather hydrophilic substrate Tyr to the active site (Koch et al., 1995). The fact that the cyanogenic glucosides in T. maritima are derived from Tyr contradicts this suggestion.

To verify the catalytic activity of the two CYP79 homologs from T. maritima, they were functionally expressed in E. coli. Codon usage in pro- and eukaryotic genes is different, and most highly expressed E. coli genes have a high AT content in the 5′ end of the gene. Therefore, modification of the 5′ end of most eukaryotic Cyt P450 genes is necessary to achieve high expression (Barnes et al., 1991; Halkier et al., 1995; Gillam, 1998). Chimeric CYP79E2, containing the first 24 amino acids of the bacterial lactZ gene, showed a dramatically increased expression level (Fig. 4) and turnover of Tyr per milligram of protein (Figs. 5 and 6) compared with CYP79E1 expressed with a few silent mutations or in a truncated form containing a modified P450E2 N terminus (Gillam et al., 1994, 1995) devoid of the membrane-spanning anchor (Fig. 4). A partial CYP79E1 clone in-frame with lacZ at the same nucleotide position as CYP79E2 showed the same high expression level as lacZ-modified CYP79E2, indicating the importance of the N-terminal sequence and choice of expression vector for efficient expression. Surprisingly, the CYP79E1 construct, containing a modified P45017α N terminus with a stretch of 13 hydrophobic amino acids for membrane insertion, resulted in no detectable expression and no activity. The same type of construct [TYRA(K1–25)] resulted in the highest level of expression of CYP79A1 (Halkier et al., 1995). All constructs showed higher expression level and activity when expressed in the E. coli strain JM109 compared with XL-1 Blue.

T. maritima contains the two cyanogenic glucosides tri-glochinin and taxiphillin, which are both derived from Tyr (Sharples et al., 1972; Conn, 1973). Tri-glochinin biosynthesis requires a ring cleavage step supposedly taking place after an additional hydroxylation of the phenolic side chain. To further study this pathway, the catalytic activity of the two isolated CYP79Es from T. maritima were tested using other putative substrates. Neither Phe nor DOPA were metabolized by the Cyt P450s, strongly suggesting that the ring cleavage reaction leading to triglochinin formation proceeds after p-hydroxyphenylacetone formation. The ring cleavage reaction could be dependent on hydroxylation carried out by a 2-oxoglutarate-dependent dioxygenase (Prescott and John, 1996).

(Z)-p-Hydroxyphenylacetaldoloxide is converted to p-hydroxybenzaldehyde and the presence of low amounts of p-hydroxybenzaldehde and the presence of low amounts of p-hydroxyphenylacetone (Fig. 6). CYP71E1 is a labile enzyme and, in its isolated state, releases considerable amounts of p-hydroxyphenylacetone from the active site before the subsequent C-hydroxylation reaction proceeds (Kahn et al., 1997; Bak et al., 1998a). Accordingly, p-hydroxyphenylacetone is also observed to accumulate in the reaction mixtures, including the reaction mixture solely composed of S. bioclor CYP79A1 and CYP71E1.

In biosynthetic experiments using S. bioclor microsomes, no p-hydroxyphenylacetone is observed (Møller and Conn, 1980) while p-hydroxyphenylacetone accumulates in T. maritima (Nielsen and Møller, 1999). In the reconstitution experiments (Fig. 6), the relative levels of p-hydroxyphenylacetone and p-hydroxybenzaldehyde were the same in all experiments, indicating that the activ-
ity of CYP71E1 is not differentially affected by the presence of CYP79A1 from S. bicolor or by native and truncated forms of CYP79E1 and CYP79E2 from T. maritima. This indicates that other membrane-associated regions of CYP79E1 are sufficient to retain correct orientation and association with CYP71E1. The efficient interaction between enzymes from the two species, T. maritima and S. bicolor, indicates a high conservation of the enzymes involved in the biosynthesis of cyanogenic glucosides, and suggest the presence of a CYP71E1 homolog in T. maritima. In vivo, it is conceivable that, in contrast to CYP71E1 from S. bicolor, the CYP71E1 homolog present in T. maritima releases sufficient amounts of p-hydroxyphenylacetoinitrile to permit the formation of triglochinin, the major cyanogenic glucoside in T. maritima.

ACKNOWLEDGMENTS

Dr. Søren Bak (Department of Entomology, University of Arizona, Forbes 410) is thanked for helpful discussions and Karina Juul Petersen for technical assistance. Dr. Peter Busk kindly provided a PCR fragment of CYP79D1, and Dr. Mette Dahl Andersen has kindly provided CYP79D1 for sequence comparison studies. Dr. Rachel Alice Kahn (Département d’Enzymologie Cellulaire et Moléculaire, Institut de Biologie Moléculaire des Plantes, Centre National de la Recherche Scientifique, Strasbourg, France) is acknowledged for kindly providing the NADPH-Cyt P450 oxidoreductase and CYP71E1 from S. bicolor. Dr. Hans-Henning Steinbiss (Max-Planck-Institut für Züchtungsforschung, Köln, Germany) is acknowledged for kindly providing the pSBET vector.

Received October 14, 1999; accepted December 28, 1999.

LITERATURE CITED


Bak S, Nielsen HL, Halkier BA (1998b) The presence of CYP79 homologous in glucosinolate-producing plants shows evolutionary conservation of the enzymes in the conversion of amino acid to aldoxime in the biosynthesis of cyanogenic glucosides and glucosinolates. Plant Mol Biol 38: 724–734


Sharples D, Spring MS, Stroker JR (1972) Biosynthesis of the major cyanogenic glycoside of Thalictrum aquilegifolium. Phytochemistry 11: 2999–3002


Tapper BA, Butler GW (1972) 2-Hydroxyaldoximes as possible precursors in the biosynthesis of cyanogenic glucosides. Phytochemistry 11: 1041–1046
