CYP98A3 from Arabidopsis thaliana is a 3’-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway

Guillaume Schoch, Simon Goepfert, Marc Morant, Alain Hehn, Denise Meyer, Pascaline Ullmann and Danièle Werck-Reichhart*

Department Stress Response, Institute of Plant Molecular Biology, CNRS UPR 2357, 28 rue Goethe, F-67083 Strasbourg Cedex, France

Running Title: CYP98A3, a 3-hydroxylase of p-coumaroyl esters in Arabidopsis

Corresponding author*:

Danièle Werck-Reichhart, Department Stress Response, Institute of Plant Molecular Biology, CNRS UPR 2357, 28 rue Goethe, F-67083 Strasbourg Cedex, France
Tel.: (33) 3 90 24 18 54
Fax.: (33) 3 90 24 18 84
e-mail: daniele.werck@ibmp-ulp.u-strasbg.fr
Summary

The 4 and 5-hydroxylations of phenolic compounds in plants are catalyzed by cytochrome P450 enzymes. The 3-hydroxylation step, leading to the formation of caffeic acid from p-coumaric acid, remained however elusive, alternatively described as a phenol oxidase, a dioxygenase or a P450, with no decisive evidence for the involvement of either in the reaction in planta. In this paper we show that CYP98A3, which was the best possible P450 candidate for a 3-hydroxylase in the Arabidopsis genome, is highly expressed in inflorescence stems and wounded tissues. The recombinant CYP98A3 expressed in yeast does not metabolize free p-coumaric acid, nor its glucose or CoA esters, p-coumaraldehyde, or p-coumaryl alcohol, but very actively converts the 5-O-D-shikimate and quinate esters of trans-p-coumaric acid into the corresponding caffeic acid conjugates. The shikimate ester is converted 4 times faster than the quinate derivative. Antibodies directed against recombinant CYP98A3 specifically reveal differentiating vascular tissues in stem and root. Taken together, these data show that CYP98A3 catalyzes the synthesis of chlorogenic acid and very likely also the 3-hydroxylation of lignin monomers. This hydroxylation occurs on depsides the function of which was so far not understood, revealing an additional and unexpected level of networking in lignin biosynthesis.
INTRODUCTION

The systematic genome sequencing is revealing a large number of orphan genes and their phylogenetic relatedness to genes with characterized function. EST sequences, on the other hand, is providing preliminary information on levels, patterns of expression, and conservation of genes among species. Taken together, such information can be exploited as a clue to gene function and to track down missing steps in important pathways.

The sequencing of the whole genome of *Arabidopsis thaliana* has revealed 273 cytochrome P450 genes distributed into 45 families and subfamilies (http://drnelson.utmem.edu/CytochromeP450.html; http://www.biobase.dk/P450/). P450s thus form the largest superfamily of enzymes involved in plant metabolism, but the function of 80% of these enzymes is still unknown. Our attention was first drawn to the CYP98 family by its phylogeny and structure. An analysis of P450 phylogeny in *A. thaliana* (Figure 1) shows that the CYP98 family is most closely related to CYP73A5, coding for the cinnamic acid 4-hydroxylase, the second enzyme and first P450 in the phenylpropanoid pathway (1). CYP73A5 and CYP98s seem to have evolved from the same ancestor as CYP79s which also have in common with CYP73A5 to use aromatic substrates derived from the shikimate pathway (2,3). It was thus tempting to speculate that the substrate of CYP98s was derived from aromatic amino acids as well. The *Arabidopsis* CYP98 family is formed of only three genes. CYP98A3 is present in single copy, CYP98A8 and CYP98A9 are 69% identical to one another and only 52% identical to CYP98A3. All P450 genes in the phenylpropanoid pathway (CYP73A5, CYP84A1, CYP75B1) that have been characterized so far in *Arabidopsis* are present in single copy in the genome. Such a situation is unusual in other P450 families, most of them showing multiple duplications. Phylogenetic analysis thus pointed to CYP98A3 as an enzyme likely to be involved in the phenylpropanoid pathway.
Such an hypothesis was comforted by the high frequency of CYP98A3 ESTs reported from many *Arabidopsis* libraries (roots, rosette, inflorescence, siliques, seeds), but also by the high frequency of other CYP98 ESTs detected in a variety of plants species and tissues. Among tissues expressing high CYP98 message were poplar and pine xylem (4,5), soybean hypocotyl and stem, as well as cotton fibers. In support of the latter EST data, a CYP98 cDNA was PCR-isolated from sweetgum xylem together with those for CYP73 and CYP84 which catalyze cinnamic acid and coniferyl aldehyde hydroxylations in lignin biosynthesis (6). Message frequency, wide distribution and location thus suggested probable involvement of CYP98As in a high throughput pathway and a function in the formation of some structural element, possibly for the formation or reinforcement of the cell-wall. A good candidate function for CYP98A3 was 3-hydroxylation of the phenylpropanoid ring, a still elusive step in the phenylpropanoid pathway needed for the synthesis of lignin monomers and other abundant and widespread plant compounds such as chlorogenic acid.

In this paper we confirm by RNA blotting that the CYP98A3 is constitutively expressed in all plant tissues and show that its accumulation is increased in wounded leaves. The CYP98A3 protein expressed in yeast does not metabolize free *p*-coumaric acid nor its glucose or CoA esters, but hydroxylates the coumaroyl esters of shikimic and quinic acids with a high efficiency, higher than that previously reported for the 4-hydroxylation of cinnamic acid, the upstream P450 catalyzed step in the phenylpropanoid pathway. The enzyme selectively metabolizes the natural 5-*O*- and *trans*-isomers of the substrates. Polyclonal antibodies raised against the recombinant enzyme specifically label lignifying tissues in stem and roots. Taken together, these data demonstrate that CYP98A3 is involved in the biosynthesis of chlorogenic acid, and strongly suggest that it also catalyzes the 3-hydroxylation of the lignin monomers. Although previously foreseen by Heller and Kühnl (7),
this new development is rather unexpected and raises a new degree of complexity and additional gridding level in the already complex lignin biosynthesis pathway.

EXPERIMENTAL PROCEDURES

Chemicals

Chlorogenic, shikimic, D-quinic, p-coumaric, caffeic and ferulic acids were from Sigma (l’Isle d’Abeau Chesnes, France). p-Coumaryl alcohol and p-coumaraldehyde were gifts from Dr. M. Barber (Southampton University). trans-5-O-Coumaroyl and trans-5-O-caffeoyl-shikimic acids were gifts from Dr. W. Heller (GSF, Munich). β-Megaspermin was provided by Dr. S. Kauffmann (IBMP, Strasbourg). [14C]-p-coumaric acid was synthesized enzymatically from trans-[3-14C]cinnamate (Isotopchim, Ganagobie, France) using microsomes from recombinant yeast expressing CYP73A1 (8). The 4- and 3- isomers of p-coumaroyl shikimic acid were generated by heating a solution of the trans-5-O-isomer in sodium phosphate buffer 0.1M pH 7.4 for 1 h at 90°C, and the cis isomer of trans-5-O-p-coumaroyl shikimic acid was obtained by irradiation for 10 min at 254 nm (9). p-Cinnamoyl-CoA and p-coumaroyl CoA were synthesized according to (10,11). Radiolabelled p-coumaroyl-1- and 4-glucosides were synthesized using recombinant tobacco glucosyl transferase (TOGT) incubated with and p-coumaric acid and UDP-[14C]-glucose as described in (12).

Cell culture and extraction

β-Megaspermin (50 nM) was added under sterile conditions to a flask containing 10 mL of a 6 day old culture of tobacco BY cell suspension. After 4h incubation in the dark, cells were harvested by filtration, frozen in liquid nitrogen, and stored at -80°C. Crude extract was
prepared using the protocol described by Heller and Kühnl (7), slightly modified as follows. Four grams of frozen cells were homogenized in a mortar with 0.2 g Dowex 1 x 2 and suspended in 0.1 M potassium phosphate pH 7 containing 1% polyvinylpolypyrrolidone, 28 mM 2-mercaptoethanol, 0.2 mM phenylmethylsulfonyl fluoride and 20 mM sodium metabisulfite. After centrifugation at 10,000g for 20 min, the supernatant was desalted on a Sephadex G-25 column (HiTrap Desalting, Pharmacia), equilibrated in 50 mM potassium phosphate pH 7.0, 10 mM DTT. The eluted fraction was concentrated on a Centricon 10 (Amicon), and either used directly as a source of hydroxycinnamoyl CoA :quinate/shikimate hydroxycinnamoyl transferases, or stored at -80°C in 20% glycerol.

Enzymatic preparation of coumaroyl-quinic acid

trans-5-O-p-Coumaroyl-D-quinic acid was synthesized enzymatically from trans-4-coumaroyl-CoA and D-quinic acid using the desalted and concentrated extract from tobacco BY cells. The incubation mixture, containing in a final volume of 500 µl 0.5 mM 4-coumaroyl-CoA, 4 mM D-quinic acid, 50 mM potassium phosphate pH 7.0 and 250 µL (0.55 mg protein) of the crude extract, was incubated for 2 hours at 28°C in the dark. After addition of 60 µl acetic acid, p-coumaroyl quinic acid was extracted three times with 1 volume of ethyl acetate and evaporated to dryness under argon. Identity of the product was checked by UV spectroscopy and negative electrospray mass spectrometry (m/z 337.3). It was further purified by HPLC for enzyme kinetic analysis.

cDNA isolation and expression in yeast

The CYP98A3 coding sequence was amplified from an A. thaliana Col-0 siliques cDNA library (13) by PCR using primers 5’-CGGGATCC ATGTCGTGGTTTCTAATAGC and 5’-GCGAATTCTTACATATCGTAAGGCACGC, designed according to the data from genome
sequencing (AC002409, T20B5.9). Bam HI and Eco RI restriction sites were added in 5’ and 3’, respectively, for cloning in the yeast expression vector pYeDP60 (14). The PCR mixture contained 10 ng of template, 20 pmol primers, 0.5 µM dNTPs, 3.5 mM MgCl₂ and 10% DMSO in a total volume of 50 µL. It was pre-heated for 2 min at 94°C before addition of 5 units of Pfu DNA polymerase (Stratagene). After 3 min additional heating at 94°C, 25 cycles of amplification were carried out as follows: 1 min denaturation at 92°C, 30 s annealing at 52°C, 5 min extension at 72°C. The reaction was completed by 10 min extension at 72°C, plus 30 additional seconds at 72°C after addition of Taq DNA polymerase (Gibco-BRL) to graft 3’-A overhangs for cloning into a T-tailed vector. Double-stranded pGEM-T (Promega) subclones were checked by sequencing using the prism Ready Reaction Dye Deoxy Terminator Cycle method of Applied Biosystems Inc, and transferred into expression vectors. The sequence data were analyzed using the GCG sequence analysis Software package, version 8.1 or Clustal X. Transformation of the Saccharomyces cerevisiae strain WAT11, engineered to inducibly express the NADPH-cytochrome P450 reductase from Arabidopsis thaliana ATR1 upon galactose induction, was performed as described in (14). Yeast were grown as previously described (14), and microsomes were isolated after 16 to 24 hours induction on 20 g L⁻¹ galactose at 20 or 30°C.

Another recombinant yeast strain was constructed, expressing a 4-His-tagged CYP98A3, using the 3’ primer 5’-CGGAATTCTTAATGATGATGATGCATATCGTAAGGCACGCGTT.

The CYP98A8 (AC011765, F1M20.22) and CYP98A9 (AC011765, F1M20.23) coding sequences were amplified by PCR using as template A. thaliana Col-0 genomic DNA with the primers 5’-GGAAGATCTATGATTATATCTAATTTC and 5’-GGGGTACCTTAATCTAAAGGTAAAGGTA, and 5’-GGAAGATCTATGGATTTATTACTCATAC and 5’-
CGGAATTCTTAAGGTATAACTTTGTG, respectively. CYP98A8 was cloned into the Bam HI and Kpn I sites, and CYP98A9 into the Bam HI and Eco RI sites of pYeDP60.

Enzyme purification, production of polyclonal antibodies

CYP98A3, 4-His-tagged at the C-terminus, was solubilized in 0.9 % Emulgen 911 (Kao Atlas) and purified on a Ni2+-loaded HiTrap Chelating column (Pharmacia Biotech) using the procedure recommended by the manufacturer, with elution in sodium phosphate 50 mM pH 7.4 containing 10% glycerol, 0.02% Emulgen 911, 0.5 M NaCl and 60 mM imidazole. Polyclonal antibodies were raised in rabbits by successive injections of once 16 µg and 5 times 8 µg purified protein emulsified in Freund's complete and incomplete adjuvants, respectively, and used for western blot analysis as in (15).

Standard assay of the 3'-hydroxylase (C3'H) activities

The standard assay for C3'H contained in a total volume of 200 µL 100 mM sodium phosphate buffer pH 7.4, 3 µg microsomal protein from yeast (i.e. 0.6 pmol of CP98A3), and 4 to 145 µM substrate and 600 µM of NADPH. The reaction was incubated at 28°C for 5 min and terminated by the addition of 20 µl acetic acid. The products were extracted three times with 2 volumes of ethyl acetate, the organic phase was pooled, evaporated under argon, and the residue dissolved in 300 µL of 10% acetonitrile, 90% water, 0.2% acetic acid (v/v/v) was analyzed by reverse-phase HPLC (LiChrosorb RP-18 Merck, 4 x 125 mm, 5 µm; flow rate 1 ml min⁻¹; 5 min isocratic 10% acetonitrile, then 20 min linear gradient from 10 to 52% acetonitrile in water containing 0.2 % acetic acid). For an accurate determination of kinetic constants with the quinate esters which are not efficiently extracted with ethyl acetate, the reaction was stopped with 40 % acetonitrile and 0.2% acetic acid before centrifugation 15 min at 18,000g, and 4 fold dilution for HPLC analysis.
Absorbance of the eluant was monitored with a diode array detector, the retention times were 6.5 min for caffeoyl-quinate, 10.9 min for caffeoyl-shikimate, 10.5 min for coumaroyl-quinate and 12.6 min for coumaroyl-shikimate. Substrate conversion was calculated from peak areas at 320 nm by comparison with injected standards. Amounts of substrates and products in incubation media and pooled extracts were calculated using the extinction coefficient at 340 nm: 6,200 M\(^{-1}\)cm\(^{-1}\) for substrates and 15,700 M\(^{-1}\)cm\(^{-1}\) for products (9). For characterization of the reaction products, HPLC elutions corresponding of the peaks of products were pooled, evaporated and submitted to mass spectrometry analysis on a BioQ triple quadrupole (Micromass).

Kinetic data were fitted using the nonlinear regression program DNRPEASY derived by Duggleby and Leonard from DNRP53 (16).

Assay of CoA and glucose esters hydroxylation

Incubations with cinnamoyl CoA, \(p\)-coumaroyl CoA, \(p\)-coumaroyl glucose ester and 4-glucoside were performed as for shikimate and quinate esters, except that up to 10 fold higher concentrations of yeast microsomes and longer incubation times were also assayed to exclude any possibility of low rate metabolism. Products, extracted 3 fold in two volumes of ethyl acetate, were analyzed both directly or after one hour hydrolysis in 1N HCl at 90°C. Products of hydrolysis and glucose conjugates were analyzed by HPLC as described for shikimate and quinate esters, using diode array and radio-detection (Packard Flow Scintillation Analyzer 500TR). HPLC analysis of intact CoA conjugates was performed without prior extraction using a gradient of acetonitrile in 15 mM (NH\(_4\))\(_2\)HPO\(_4\), 15 mM HCl, pH 5.5 (flow rate 1 mL min\(^{-1}\); 3 min isocratic 10% acetonitrile, then 13 min linear gradient from 10 to 50% acetonitrile).
**Measurement of O-methyl transferase activity**

Methylation of the C3’H reaction products was assayed using recombinant caffeoyl-CoA-O-methyl transferase (CCoAOMT1 from tobacco) expressed in *E. coli* and purified as described in (17), alone or combined to hydroxycinnamoyl-CoA transferase prepared from tobacco cells as described above. The assay contained in 100 µL Na phosphate buffer 40 µM pH 7.4, 40 µM *S*-adenosyl methionine, 0.2 mM MgCl₂, 2 mM dithiothreitol, 40 µM 5-O-caffeoyl-shikimate or quinate (or caffeoyl CoA) and 20 µg purified protein. In some cases, 22 µg of tobacco BY cells crude extract were added as a source of hydroxycinnamoyl-CoA transferase. After 2 hours incubation at 30°C, the reaction was stopped by addition of 25 µL 4N HCl. Products were hydrolyzed 30 min at 90°C and extracted 3 times with 2 volumes ethyl acetate. The organic phase was pooled, evaporated, and the residue dissolved in 300 µL of 10% acetonitrile, 90% water, 0.2% acetic acid (v/v/v) was analyzed by reverse-HPLC as described above.

**Spectrophotometric Measurements**

Spectrophotometric measurements of total P450 content and evaluation of substrate binding were performed according to (18) and (19), respectively. Substrate-binding spectra were recorded using double cuvettes.

**RNA blot analysis**

RNA was isolated from three months old plants. For wounding experiments, leaves were lacerated with a razor blade and aged for 24 hour under continuous light in standard Murashige and Skoog medium. Control leaves were aged without laceration. Total RNA was prepared using the RNAeasy® Plant Mini Kit (Qiagen GmbH, Germany), quantified and concentrations were adjusted to 8 µg µL⁻¹. RNA blot analysis was performed using 16 µg of
total RNA separated on a 1.2 % denaturing formaldehyde/agarose gel and blotted onto BrightStar-Plus™ membrane (Ambion). After RNA fixation 1 h at 80°C, the membrane was stained with methylene blue to check integrity and equal loading of RNA. The $^{32}$P-labelled probe corresponding to the entire coding region of CYP98A3 was synthesized by random priming using Ready-To-Go™ DNA labeling Beads (Amersham Pharmacia Biotech Inc). The membrane was hybridized in 5 x SSC, 5 x Denhardt’s solution, 0.5% SDS, 2 mM EDTA, 100 µg µL$^{-1}$ sonicated salmon sperm at 65°C, then washed twice at 60°C in 0.2 x SSC, 0.1% SDS and signal was recorded by autoradiography.

**Tissue print hybridization and histochemical detection of lignin**

Stems and roots transversal hand cuts were printed onto nitrocellulose Schleicher and Schuell 0.2 µm, washed twice 20 min in PBS containing 0.4% Tween 20, then blocked and revealed as a standard immunoblot using pre-immune or anti-CYP98A3 polyclonal serum diluted 1/1,000. Before dilution, the crude serum was incubated 5 min with an equal volume of microsomes of recombinant yeast over-expressing CYP73A1 (8) to minimize background staining and possible cross-recognition of CYP73 epitopes. Protein antibody complexes were detected using alkaline phosphatase-conjugated goat anti-rabbit IgG with 5-bromo-4 chloro-3-indolyl phosphate and nitroblue tetrazolium as substrates, in the presence of 640 mg L$^{-1}$ levamisole (Sigma) to inhibit plant phosphatases.

Hand cut transverse stem and root sections were also stained with phloroglucinol-HCl for lignin (C$_6$-C$_3$ cinnamaldehydes and C$_6$-C$_1$ benzaldehydes) staining.

**Phylogenetic analysis**

An Arabidopsis P450 database was constructed using information available at http://drnelson.utmem.edu/BiblioD.html and http://www.biobase.dk/P450/p450list.shtml. For
each family, a consensus sequence was generated using Clustal X 1.8 (available at http://www-igbmc.u-strasbg.fr/BioInfo/) and Genedoc (available at http://www.psc.edu/biomed/genedoc). Sequences truncated from the hypervariable membrane anchor up to the proline-rich hinge region were used to generate the alignments and consensus. The final alignment of the consensus sequences and the phylogenetic tree were generated with Clustal X and Treeview 1.5.2 (available at http://taxonomy.zoology.gla.ac.uk/rod/rod.html).

RESULTS

CYP98A3 gene expression in Arabidopsis

The scanning of the numerous CYP98 ESTs available in data banks suggests a significant level of constitutive gene expression in many plant tissues, in particular expression in lignin-rich tissues such as stems, xylem and fibers. In Arabidopsis, CYP98A3 appears as one of the constitutively expressed P450 genes, but ESTs do not give any precise idea of its tissue- or organ-specific expression. RNA blot analysis was thus performed with total RNA from the different plant organs, and from leaves lacerated and aged on growth medium to activate genes of phenylpropanoid metabolism involved in repair and defense mechanisms (20) (Fig. 2). This analysis shows that CYP98A3 message is present in all plant tissues but is by far highest in stems, then in roots and siliques. In leaves, message accumulation is induced by wounding. Expression of CYP98A3 in Arabidopsis is thus high in lignin synthesizing tissues.

Isolation of the CYP98A3 cDNA and expression of the protein in yeast
The coding sequence from CYP98A3, available from genome sequencing, was used to design PCR primers for amplification of the complete cDNA. Restriction sites allowing insertion into the yeast expression vector pYeDP60 (14) were added at both ends. The amplicon was first cloned into a pGemT vector for complete sequencing before transfer to the expression vector. Galactose-induced expression in the WAT11 yeast strain, co-expressing the *A. thaliana* P450 reductase ATR1, under standard conditions routinely led to the obtention of about 150 pmol of P450 per mg of yeast microsomes, i.e. 15 nmol L⁻¹ culture (Fig. 3A). CYP98A8 and CYP98A9, expressed under similar conditions, were usually produced at lower levels. The best preparation contained 107 pmol per mg microsomal protein and 5.3 nmol L⁻¹ culture for CYP98A8, and 118 pmol per mg protein and 7 nmol L⁻¹ culture for CYP98A9). Addition of an 4-His tag at the C-terminus of CYP98A3 did not significantly alter its level of expression and allowed protein purification for the production of rabbit polyclonal antibodies. These antibodies were specific for CYP98A3 and did not cross-react with yeast-expressed CYP73A1, CYP98A8 or CYP98A9 (Fig. 3B).

*CYP98A3 is a 3-hydroxylase of *p*-coumaric acid esters*

Substrate specificity was investigated using recombinant CYP98A3 in yeast microsomes, co-expressed with the *A. thaliana* P450 reductase ATR1. Free phenylpropanoids were shown to be the substrates of the P450s involved in 4 and 5-hydroxylation of the aromatic ring of the C₆-C₃ structure for lignin biosynthesis: cinnamic acid in the case of CYP73s (8,21), and coniferyl aldehyde, coniferyl alcohol, and to a lesser extend ferulic acid, in the case of CYP84s (6,22). The first tests were thus performed with free *p*-coumarate, *p*-coumaroyl aldehyde and *p*-coumaroyl alcohol which did not induce the shift in the CYP98A3 heme iron spin-state that would be expected upon binding of a P450 ligand (Type I ligand binding spectrum) (19,23), nor any trace of conversion into a more oxygenated molecule in
the presence of NADPH and CYP98A3. Recent data have indicated that the methylation step of caffeic acid into ferulic acid is likely to predominantly occur on a CoA conjugate (17,24-29). Glucose esters, on the other hand, were frequently described as alternative free energy-rich precursors of phenolic derivatives (30,31,32), while 4-O-β-D-glucosides are considered as detoxification, transport and storage forms of lignin precursors (33). The CoA ester, 1-O-glucoside (glucose ester), as well 4-O-glucoside of p-coumaric acid were thus assayed as substrates of CYP98A3. None of these compounds showed any sign of binding to CYP98A3, nor was converted into a more hydrophilic product. Two already old reports by Heller and Kühnl (7) and Kühnl et al., (9) described a P450-catalyzed 3'-hydroxylation of shikimate and quinate esters of p-coumaric acid by microsomal fractions of parsley and carrot cell cultures (Fig. 4). While the hydroxylation of the quinate ester was obviously linked to the biosynthesis of chlorogenic acid, it was postulated that shikimate esters were just metabolically transient intermediates for the formation of more oxygenated cinnamic acids, including lignin precursors. Competition and inhibition experiments suggested that a single P450 catalyzed both reactions. In order to investigate further what was so far considered as a rather odd hypothesis, we incubated recombinant CYP98A3 with 5-O-(4-coumaryl)-D-quinate or 5-O-(4-coumaryl)-shikimate in the presence of NADPH. Both the quinate and shikimate esters were very rapidly converted into a more hydrophilic product (Fig.5). The reaction was completely dependent on NADPH and CYP98A3, no conversion was obtained upon incubation with microsomes from yeast transformed with a void plasmid. Crude serum of rabbit immunized with purified 4His-CYP98A3 inhibited the reaction by 50 % compared to pre-immune serum. Comparison with standards of the HPLC retention times, UV absorption spectra and negative electrospray mass spectrometry analysis (m/z 335.3 for caffeoyl shikimate and m/z 353.3 for caffeoyl quinate) of the products indicated the formation of
caffeoyl derivatives. This was confirmed by acid hydrolysis of the products, leading to the formation of a product with the characteristics of caffeic acid.

As expected from their low sequence homology with CYP98A3, neither CYP98A8, nor CYP98A9 metabolized, even at low rates, the shikimate and quinate esters of p-coumaric acid.

*Characteristics of the reactions*

Catalytic parameters of the reactions were determined in Na phosphate 0.1 M pH 7.4 and at 28° C (Table I). Both \( K_m \) and \( K_{cat} \) favour the metabolism of the shikimate rather than that of the quinate ester, the catalytic efficiency of the enzyme being 4 fold higher with 5-O-(4-coumaryl)-shikimate. The \( K_{cat} \) of the C3’H is very high compared to that of other plant P450s expressed in yeast, in particular higher than that we routinely measure with the recombinant CYP73s (cinnamate 4-hydroxylases) under similar conditions. A high turnover for the 3’-hydroxylase was predicted by Heller and Kühnl (7) and was already suggested by Ulbricht and Zenk (34).

5-O-(4-coumaryl)-shikimate/quinate easily isomerize from *trans* to *cis* under UV-light or from 5-O- to the 3-O- or 4-O-isomers at physiologic pH (9). The latter process is accelerated at higher temperatures and results from a base-catalyzed intramolecular migration (35). Natural 3-O- or 4-O-isomers are naturally found in some plant tissues (36-38). As shown in Fig. 6 (A and B), recombinant CYP98A3 exclusively metabolizes the *trans* isomer of 5-O-(4-coumaryl)-shikimate. The *cis* form remains intact even after complete conversion of the *trans*. Microsomes from carrot cells cultures were previously reported to exclusively metabolize the 5-O-isomer of the quinate ester (9). Recombinant *Arabidopsis* CYP98A3 preferentially hydroxylates the 5-O-isomer, but also converts the 4-O- and 3-O- isomers although with a lower efficiency (Fig. 6, C and D). CYP98A3 thus shows a preference for the
isomer which is the most abundant under normal conditions and is formed by the $p$-
hydroxycinnamoyl CoA:shikimate-$p$-hydroxycinnamoyl transferase, but is also able to cope
with other isomers that may arise by isomerization in planta, for example under heat stress
conditions.

**Spectrophotometric detection of substrate binding**

Our initial screening for potential substrates was performed using spectrophotometric
methods for the detection of a shift in the maximum of absorbance of heme that is normally
expected upon binding of a substrate (39). This method which was very useful with other
P450s (19,40), detected very little change, if any, in P450 absorption upon addition of 5-$O$-(4-
coumaryl)-shikimate, and this despite a high expression of CYP98A3 in yeast microsomes,
low affinity and high rates of metabolism which implicate an optimal positioning in the active
site. Such an absence of low to high spin transition upon substrate binding seems to be shared
by other plant P450s metabolizing compounds with an hydroxyl group next to the position of
attack. It possibly means that heme coordination with the hydroxyl oxygen maintains such
oxidized P450s in a low-spin state, which would raise some questions concerning their redox
potential and interaction with P450 reductases. It may also mean that the active site is
naturally devoid of solvent and heme ligand in the resting state.

**The CYP98A3 protein is highly expressed in lignifying tissues.**

The expression pattern of the CYP98A3 gene and the high turnover of the 3’-
hydroxylation reaction favour the hypothesis that the bulk 3-hydroxylation of phenolics
occurs on the shikimate or quinate esterified forms of the phenylpropane structure. To test
further this working hypothesis, tissue-specific expression of the CYP98A3 protein was
visualized in the plant organ showing the highest gene expression. Stem and root transversal
sections were printed onto nitrocellulose and revealed with the polyclonal antibodies raised against recombinant CYP98A3 (Fig. 7). Hand sections of neighboring tissues were stained with phloroglucinol-HCl, to localize lignin accumulation. To follow xylem development in the mature inflorescence stem (41), prints were taken at different distance from the apical meristem. Expression of CYP98A3 correlated with active lignification, as was previously observed for the expression of CCoOMT in several dicot plants (25,26). Highest protein expression was detected in differentiating xylem, first confined in protoxylem from vascular bundles in the upper part of the stem, then in metaxylem and interfascicular region forming a continuous ring in the lower mature stem. In the mature root (Figure 7, G and F), some expression was observed in the cortical zone mostly constituted of secondary phloem, but CYP98A3 protein was mainly detected in the ring of differentiating xylem at the periphery of the steele which is largely formed of lignified secondary xylem (42).

Is 5-O-caffeoyl-shikimate a substrate for CCoAMT ?

Implication of CYP98A3 in lignification raises the problem of the next step in the lignin pathway. Does methylation also occur on a shikimate/quinate derivative or is the caffeic acid depside converted back to a CoA ester for methylation? To try answering this question we checked the substrate specificity of the recombinant caffeoyl-CoA O-methyl transferase (CCoAOMT1) which is expected, from its in vitro substrate specificity, in planta expression pattern (17), and downregulation impact on lignin synthesis (43) to be the best candidate for methylation of caffeoyl units in tobacco. This enzyme, and the other OMTs and CCoAOMTs from tobacco were already reported not to methylate chlorogenic acid (17). Chlorogenic acid is however considered as an accumulation product, while shikimate ester, which is the best substrate of the C3’H, is usually assumed to be a transient intermediate. We thus incubated 5-O-caffeoyl-shikimate with S-adenosyl methionine and CCoAOMT1. No
formation of ferulate ester was observed. This result is in agreement with the report of Kühnl et al. (9) that caffeoyl CoA-O-methyltransferase from carrot cells did not methylate chlorogenic acid nor 5-O-caffeoyl shikimate. Shikimate and quinate esters of caffeic acid thus do not seem to be substrates of OMTs or CCoAOMTs.

The hydroxycinnamoyl-CoA:shikimate/quinate-hydroxycinnamoyl transferases, which convert p-coumarate from the CoA to the quinate/shikimate esters in tomato or Cichorium endivia, were described as reversible enzymes (34,44). When an aliquot of the concentrated crude extract from tobacco cells containing the hydroxycinnamoyl-CoA transferase activity (see Material and Methods Enzymatic preparation of coumaroyl-quinic acid) was included together with CoA in the CCoACOMT assay, a ferulate derivative was obtained (data not shown). A lower conversion to ferulate, due to CCoAOMT activity of tobacco crude extract, was obtained in the absence of recombinant CCoAOMT1. It seems thus likely that the equilibrium between the CoA and shikimate ester pools in the plant cells allows for the methylation step to occur on the CoA ester.
DISCUSSION

The 3-hydroxylation of the hydroxycinnamoyl units so far remained the enigmatic step of the phenylpropanoid and lignification pathway. Initial attempts at characterization of a 3-hydroxylation of free caffeic acid often attributed an activity to soluble phenolases (for a review see (45)), but led to no conclusive identification of the enzyme involved. It was obvious that no P450 catalyzed this reaction since incubation of microsomes from various plants with radiolabeled cinnamic acid led to p-coumaric acid, but caffeic acid was never produced, even at low levels. Studies of the 3-hydroxylation reaction by several laboratories then led to the evidence that parallel pathways may exist, acting at the level of conjugated hydroxycinnamic acids, such as esters of CoA (46,47), shikimate and quinate (7,9), phenyllactate (48), or glucose (49). The CoA ester of p-coumaric acid was recently considered as the best potential substrate since the methylation of caffeate to ferulate, which is the next step in the pathway, was shown to mainly occur on the ester of CoA (17,24-26,28,43). Three enzymes were so far described to catalyze the 3-hydroxylation of coumaryl-CoA. One is a non-specific polyphenol oxidase (50), the second a soluble FAD-dependent hydroxylase (46), the third a Zn$^{2+}$-dependent dioxygenase which was described to be inactive at a normal cytoplasmic pH (47). None of them was characterized at the molecular level nor was regarded as a top candidate for catalyzing the reaction in planta.

Availability of complete genome information recently shed a new light on the problem. On the grounds of their phylogeny, high level and pattern of expression in a broad range of plant species, members of the CYP98 family of P450 enzymes emerged as a potential catalysts for the 3-hydroxylation of phenolic compounds. We show here that CYP98A3 from Arabidopsis is indeed a 3-hydroxylase of the hydroxycinnamoyl units, and that its expression is closely associated with lignification. This P450 does not take as a
substrate p-coumaric acid or its CoA ester, but the both the esters of shikimic and quinic acids. The involvement of shikimate and quinate ester in the biosynthesis of caffeoyl units used for lignification, that is suggested by these data, is in some way a surprise since it introduces a new level of complexity and gridding in the phenylpropanoid pathway. It was however very early suggested by the work of Ulbricht and Zenk (34) who demonstrated the existence of a p-hydroxycinnamoyl-CoA:shikimate p-hydroxycinnamoyl transferase in a broad range of plant species which never accumulate shikimate esters. This work was further supported by the biochemical characterization of a P450-dependent 3’-hydroxylation of shikimate and quinate esters of p-coumaric acid in the microsomal fraction of parsley and carrot cell cultures (7,9), and by the characterization of caffeoyl-CoA 3-O-methyltransferase activity in the soluble fraction of the same carrot cell cultures (51). Further investigations in this direction was however never pursued.

CYP98A3 catalyzes the hydroxylation of two structurally related substrates and their isomers. Such a relaxed substrate specificity is unusual among plant P450s, with the exception of the enzymes metabolizing fatty acids. It might be needed to compensate the easy interconversion of the isomers. In the case of CYP98A3, the shikimate ester is a better substrate than the quinate ester. This selectivity coincides with that of the p-hydroxycinnamoyl-CoA:shikimate p-hydroxycinnamoyl transferase previously characterized in radish (52) which also show a preference for shikimate over quinate for transfer of p-coumaroyl from CoA. The more efficient hydroxylation of the quinate ester in carrot, as well as the preference of some transferases for quinate in plants such as potato, C. endivia or apple (53) (37) (52) suggests that C3’H in other plant species might have different substrate preferences or specificities. Difference in substrate preference may also explain accumulation of specific esters in some plants taxa (36). Another remarkable characteristic of the C3’H is its high turnover number compared to other P450 oxygenases. Turnover data are not available
for the coniferaldehyde 5-hydroxylase (also sometimes called ferulate 5-hydroxylase or F5H),
but if the yeast-expressed cinnamate 4-hydroxylase (C4H) was initially reported to have a $K_{cat}$
of 400 min$^{-1}$ (21), in our hands it consistently turns over at 50 to 100 min$^{-1}$ (19). The C4H is
usually considered as a P450 with a very high turnover number, in agreement with its position
upstream of a high throughput pathway. The higher turnover (600 min$^{-1}$ with shikimate) of the
C3’H which is operating three steps downstream in the same pathway, probably explains why
shikimate esters were always described as transient intermediates never accumulating in plant
tissues, except for Palmae where they are considered as taxonomic markers (37).

A 3-hydroxylation step operating at the level of shikimate and quinate esters of $p$-
coumaric acid opens many possibilities and raises many questions (Fig. 8). The first is the
respective roles of shikimate and quinate esters. If shikimate esters are described as transient
intermediates, quinate derivatives such as chlorogenic acid commonly accumulate in some
plant species, and are alternatively described as growth regulators, disease resistance factors,
antioxidants and compounds affecting organoleptic quality of fruits (36,38,54,55). So, are the
shikimate and quinate esters equivalent in terms of metabolic flux or does it exist a
channeling to the synthesis of lignin and that of accumulated esters? Is the quinate ester
branch a dead-end or a bottleneck? Are other $p$-coumarate esters substrates of the C3’H? If a
channeling exists, how is it controlled: via 4-coumaroyl-CoA ligases, $p$-hydroxycinnamoyl-
CoA $p$-hydroxycinnamoyl transferases or other enzymes further converting the caffeoyl
esters?

This leads to another interesting question which is the branching of the pathway
downstream of the C3’H. Our data, and previous reports (17,51), seem to indicate that the
shikimate and quinate esters are not substrates of CCoAOMT transferases. This has to be
confirmed in other plant species and with other recombinant CCoAOMTs, but it seems that
the caffeate ester(s) have to be converted back to CoA esters for further methylation.
Hydroxycinnamoyl-CoA \textit{p}-hydroxycinnamoyl transferases are described as reversible enzymes (34,44,56). Since the exchange between shikimate/quinate and CoA is not energy consuming, the most likely scenario is that an equilibrium between the esters populations exists in plant cells. The fast and irreversible 3’-hydroxylation would then favour formation of caffeate derivatives and displace pools of conjugates towards more oxygenated structures.

The third question is the connection of the shikimate/quinate ester derivation to other wires of the metabolic grid. One of the possible connections is to the glucose esters which were reported to be converted to quinate esters via trans-esterification (57). If a 3-hydroxylase using a different substrate does not co-exist with the \textit{p}-coumaroyl shikimate/quinate 3’-hydroxylase in higher plants, the channeling of precursors along different wires of the metabolic grids may explain the independent pathways operating for the formation of guayacyl and syringyl precursors needed for lignin biosynthesis (20,28).

The last and most puzzling question is which kind of evolutionary pressure led plants to use shikimate/quinate esters rather than free acids, CoA or glucose esters for the 3-hydroxylation of hydroxy cinnamic acids. Was a P450 metabolizing or binding shikimate derivatives already present and recruited for the reaction? If so, what was the function of this ancestral P450? Or was it the combination of a need to stabilize the very autoxidizable caffeic acid and the extraordinary efficiency of shikimate ester conversion which drove CYP98 evolution? A very attractive hypothesis is that shikimate conjugation was selected since it provides a positive regulation mechanism and optimal tuning of lignin synthesis with the availability of precursors, leaving priority to the synthesis of aromatic amino acids for proteins, and other important compounds such as flavonoids.
Acknowledgements

The WAT11 yeast strain was provided by Aventis Agro (Lyon, France) and D. Pompon (Gif-sur-Yvette, France). We thank W. Heller (GSF Munich) for the generous gift of 5-O-(4-coumaroyl)-skikimate and 5-O-caffeoyl-shikimate and helpful discussion., M. Barber (Southampton University) for p-coumaraldehyde and p-coumaroyl alcohol. The synthesis of coumaroyl-CoA and caffeoyl-CoA by P. Goeffroy, and the gift of recombinant TOGT, 1-O-coumaroyl-glucoside and 4-O-coumaroyl-glucoside by P. Saindrenan, and that of recombinant CCoAMT by L. Hoffmann and M. Legrand (IBMP, Strasbourg) are very gratefully acknowledged. Technical help of M. LeRet and A. Lesot was greatly appreciated.
REFERENCES


Table I: Catalytic parameters of the 3’-hydroxylation catalyzed by recombinant CYP98A3

Catalytic parameters were determined for the recombinant enzyme co-expressed with the *A. thaliana* P450 reductase ATR1 in yeast microsomes.

<table>
<thead>
<tr>
<th>Substrate</th>
<th>$K_m$ (µM)</th>
<th>$K_{cat}$ (min$^{-1}$)</th>
<th>$K_{cat}/K_m$ (min$^{-1}$.µM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-coumaryl-shikimate</td>
<td>7 ± 1</td>
<td>612 ± 30</td>
<td>87</td>
</tr>
<tr>
<td>4-coumaryl-quininate</td>
<td>18 ± 2</td>
<td>399 ± 19</td>
<td>22</td>
</tr>
</tbody>
</table>
Figure legends

Figure 1: A phylogenetic tree of cytochromes P450 from *A. thaliana*

For simplification, the tree was built using consensus sequences representative of each family, except for CYP98s which are all represented. Families known to metabolize aromatic amino acids or simple phenolics are highlighted. The tree is unrooted and does not show distances.

Figure 2: Tissue distribution of CYP98A3 transcripts in the adult *A. thaliana* plant

Sixteen micrograms of total RNA were loaded in each lane. The full-length CYP98A3 sequence was used as a probe. L: leaves, AL: leaves detached and aged 24 hours on MS medium, WL: wounded leaves, i.e. detached, sliced with a razor blade and aged 24 hours on MS medium, St: inflorescence stems, F: flowers, Si: siliques, R: roots. Lower panel shows the methylene blue staining of the membrane to check for loading and transfer efficiency.

Figure 3: Expression of CYP98A3, 4-His tagged CYP98A3, CYP98A8 and CYP98A9 in yeast

Upper panel: CO/CO-reduced difference spectra recorded with microsomes of recombinant yeast transformed with a void plasmid (a) or expressing CYP98A3 (b), 4-His tagged CYP98A3 (c), CYP98A8 (d) or CYP98A9 (e). Microsomes are 0.7 mg.mL\(^{-1}\) in the cuvettes.

Lower panel: Immunoblot analysis of the recombinant microsomes and of an *Arabidopsis* stem crude protein extract with polyclonal antiserum (1/10,000) raised against 4-His tagged CYP98A3. Six µg of protein were loaded in each lane. Void: microsomes of yeast transformed with a void plasmid, A3: recombinant CYP98A3, A3H: recombinant CYP98A3 with a 4-His tag, A9: recombinant CYP98A9, A8: recombinant CYP98A8, 73: recombinant CYP73A1 (C4H), crude extract from *Arabidopsis* stem, M: molecular mass markers.
Figure 4: Reactions catalyzed by the C3’H

Figure 5: HPLC analysis of the products of trans-5-O-(4-coumaroyl)-shikimate and trans-5-O-(4-coumaroyl)-D-quinate metabolism by recombinant CYP98A3

Absorbance was monitored at 320 nm. Conversion is shown after 5 min incubation of 2.5 µmol recombinant CYP98A3 in a 200 µL assay. Controls performed in the absence of NADPH or using microsomes of yeast transformed with a void plasmid gave similar results. A: conversion of trans-5-O-(4-coumaroyl)-shikimate (4 nmol in the assay). B: conversion of trans-5-O-(4-coumaroyl)-D-quinate (2 nmol in the assay). Peak 1 is the product, peak 2 the substrate: UV spectra are show on the right.

Figure 6: Substrate specificity of CYP98A3 for coumaroyl-shikimate isomers

The mix of cis and trans forms was generated by UV-irradiation (upper panel). The mix of 3-, 4- and 5- isomers was generated by incubation of the 5-isomer for 1 hour at 90°C (lower panel). Panels A and C: isomers used as substrates or incubated without NADPH. Panels B and D: metabolites obtained after 5 min (dashed line) or 70 min (solid line) incubation at 28°C with 1.2 pmol of recombinant CYP98A3. S: trans-5-O-(4-coumaroyl)-shikimate, P: trans-5-O-(4-caffeoyl)-shikimate, 1: trans-3-O-caffeoylshikimate, 2: trans-4-O-caffeoylshikimate, 3: cis-5-O-(4-coumaroyl)-shikimate, 4: trans-3-O-(4-coumaroyl)-shikimate, 5: trans-4-O-(4-coumaroyl)-shikimate. Absorbance was monitored at 320 nm. Estimated turnovers are around 600 min⁻¹ for S in both experiments, which shows that isomers are not strong competitors of the metabolism of trans-5-O-(4-coumaroyl)-shikimate.

Figure 7: Immunolocalization of CYP98A3 expression in stems and roots
Hand cut transversal sections of inflorescence stems and roots were stained with phloroglucinol-HCl, a red coloration reflecting lignin content. Adjacent sections were printed onto nitrocellulose and revealed using anti-CYP98A3 polyclonal antibodies. A blue staining is indicative of CYP98A3 expression. In stems, prints were taken at increasing distances from the apical meristem to monitor temporal and developmental expression of CYP98A3 in conjunction with the differentiation of lignified tissues. No blue staining was obtained with pre-immune antibodies. A, C, E, G: lignin staining with phloroglucinol. B, D, F, H: immunostaining of CYP98A3. A and B: upper segment of the stem, close to the flower bud. C and D: mid-stem. E and F: lower, well differentiated stem close to the rosette. G and H: root.

**Figure 8: The C3’H, a new dimension in the phenylpropanoid pathway**

The pathway that seems to be active in lignification is shown in black. Alternative pathways are shown in grey. Solid arrows indicate well characterized steps. Dashed arrows indicate other, putative, activities.

PAL: phenylalanine ammonia-lyase; C4H: cinnamate-4-hydroxylase; 4CL: 4-(hydroxy)cinnamoyl CoA ligase; CST: hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase; CQT: hydroxycinnamoyl-CoA:D-quinate hydroxycinnamoyl transferase; CCoA3H: p-coumaroyl CoA 3-hydroxylase; CCoAOMT: caffeoyl CoA O-methyltransferase; CCR: cinnamoyl CoA reductase; COMT: caffeic acid/5-hydroxyferulic acid O-methyltransferase; F5H: coniferaldehyde (ferulate) 5-hydroxylase; CAD: cinnamoyl alcohol dehydrogenase; C3’H: p-coumaroyl shikimate/quinate 3’-hydroxylase.
Figure 1:
Figure 2:

**CYP98A**

**control**
Figure 3:
Figure 4:

\[ \text{trans-(p-coumaroyl)-shikimic acid} \]

\[ \text{trans-(p-coumaroyl)-quinic acid} \]

\[ \text{NADPH} \quad \text{O}_2 \]

\[ \text{trans-(p-cafeoyl)-shikimic acid} \]

\[ \text{Chlorogenic acid} \]
Figure 5:

A

Retention time (min) 

Absorbance

CYP98A3

Control

B

Retention time (min)

Absorbance

CYP98A3

Control
Figure 6:

- Absorbance versus retention time (min) for S(E) and S(Z) compounds.
- Absorbance versus retention time (min) for cis/trans and ring-OH configurations.

Retention time (min): 9 to 15
Figure 7:
Figure 8:
CYP98A3 from Arabidopsis thaliana is a 3’-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway
Guillaume Schoch, Simon Goepfert, Marc Morant, Alain Hehn, Denise Meyer, Pascaline Ullmann and Daniele Werck-Reichhart

J. Biol. Chem. published online June 27, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M104047200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts