Chemistry & Biology
Article

In Planta Biocatalysis Screen of P450s Identifies 8-Methoxypsoralen as a Substrate for the CYP82C Subfamily, Yielding Original Chemical Structures

Tanya Kruse,[1,*] Kwongling Ho,[1] Hye-Dong Yoo,[1,2] Thomas Johnson,[1] Matt Hippely,[1] Joon-Hyun Park,[1] Richard Flavell,[1] and Steve Bobzin[1]
[1]Ceres, Inc., 1535 Rancho Conejo Boulevard, Thousand Oaks, CA 91320, USA
[2]Present address: NCE Sciences, 29540 Kohoutek Way, Union City, CA 94587, USA.
*Correspondence: tkruse@ceres-inc.com
DOI 10.1016/j.chembiol.2008.01.008

SUMMARY

An in vivo plant screen that allows for the analysis of exogenously applied substrates against transgenic *Arabidopsis* lines overexpressing individual cytochrome P450s has been developed. By deploying this screen with a subset of 91 P450s, we have identified an original substrate for members of the CYP82C subfamily. The therapeutic compound 8-methoxypsoralen was hydroxylated by plants overexpressing *CYP82C2* or *CYP82C4*, forming 5-hydroxy-8-methoxypsoralen. Additionally, plants further modified this product to create a glycosylated compound, likely the compound 5-*O*-β-D-glucopyranosyl-8-methoxypsoralen. The discovery of adducts of therapeutic compounds demonstrates the potential of this biocatalysis screening approach to create compounds that may be of pharmacological value. Additionally, this platform provides a means to expand the general knowledge base of P450 enzyme/substrate combinations and may provide valuable tools for a vast array of biocatalytic and bioremediation processes.

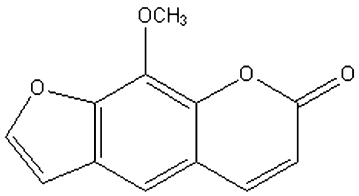
INTRODUCTION

Cytochrome P450s are heme-containing proteins responsible for a vast array of oxidative reactions involved in detoxification and biosynthetic pathways. Comprising one of the largest superfamilies of enzymes, over 7700 unique, named P450 sequences have been identified (http://drnelson.utmem.edu/CytochromeP450.html), and the numbers are likely to continue to climb rapidly as more genomes are sequenced. About one-third of the P450s identified to date have been found in plants, indicating much greater diversity than that seen in vertebrates or microorganisms. In fact, a typical plant genome contains several hundred P450s (*Arabidopsis* has 246 functional genes; rice contains 356 real genes [Nelson et al., 2004]), whereas a typical vertebrate genome possesses less than 100 (humans have 57 functional P450s [Nelson et al., 2004]), and a typical microorganism genome may contain less than 10 (*Saccharomyces cerevisiae* has 3 [Nelson et al., 2004]).

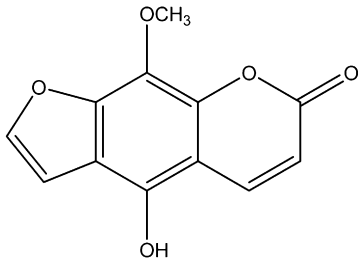
P450s catalyze most of the oxidation reactions in plant metabolism, including hydroxylation, epoxidation, dealkylation, dehydration, isomerization, and carbon-carbon bond cleavage (Kahn and Durst, 2000; Werck-Reichhart et al., 2002). Many of these P450-catalyzed reactions are carried out in a stereospecific and regiospecific manner, which contributes to the enormous structural diversity seen in plant secondary metabolism. Although many P450s have been identified, there are relatively few that have been functionally characterized and whose native substrates are known. Therefore, it is of great interest to determine what substrates these enzymes can modify in order to annotate genes and to more clearly elucidate the biosynthetic pathways of phenylpropanoids, terpenoids, alkaloids, and other secondary metabolic pathways in plants.

Biocatalytic processes in the pharmaceutical and chemical industries have exploited the stereospecific and regiospecific capabilities of enzymes to overcome limitations in modern synthetic chemistry methods (Duport et al., 1998; Picataggio et al., 1992; Bezalel et al., 1996). The number and diversity of plant P450s will expand the set of tools available for biocatalytic processes for the production of chemicals, intermediates, and products in the pharmaceutical and chemical industries. In addition, plant P450s may have utility in the production of new compounds from known chemical scaffolds, in bioremediation, or in improved plant defense. In order to gain a better understanding of this potential, it is necessary to establish a knowledge base of successful P450/substrate combinations, and to include both native and non-native substrates in this evaluation.

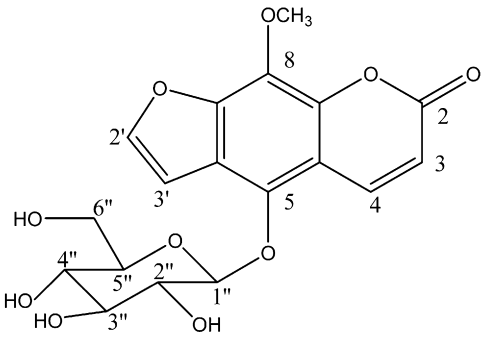
It is a daunting task to determine the substrates and products of these enzymes. Much work has been done to identify specific P450s in targeted reaction pathways (Chau and Croteau, 2004; St-Pierre and De Luca, 1995; Tian et al., 2004). However, in *Arabidopsis* alone, there are still only 41 of 246 P450s (17%) with information on their known biochemical functions (Schuler et al., 2006). Adding to the challenge is the fact that some P450s act on a single substrate at alternate positions, whereas some P450s can hydroxylate several different substrates (Schuler and Werck-Reichhart, 2003). We have available a collection of over 15,000 *Arabidopsis* lines, each overexpressing a single gene. Using this proprietary resource, we made a selection of 91 P450s representing 47 subfamilies to be utilized in our experiments. Here, we report an in vivo screening method using a set of *Arabidopsis* seedlings overexpressing individual P450s to determine potential substrates for P450s. The plant in vivo



1



2



3

Figure 1. The Structures of 8-Methoxypsoralen, 1; 5-Hydroxy-8-Methoxypsoralen, 2; and 5-*O*-β-D-Glucopyranosyl-8-Methoxypsoralen, 3

system described here has the advantage of not only identifying substrates previously unknown to be modified by plant P450s, but also captures further modifications by endogenous secondary metabolic pathway genes, which could create even further novel chemistry.

Using this plant in vivo screen, a set of 50 compounds was tested against one panel of 45 P450s, and an additional 51 chemicals, for a total of 101, were screened against a second panel of 46 P450s. The complete list of 91 P450s used in the screen is found in the Supplemental Data available with this article online. Using this plant in vivo system, we identified 8-methoxypsoralen (8-MOP) (Figure 1, compound **1**) as a substrate for members of the CYP82C subfamily. Further studies with a yeast expression system confirmed this activity. In both plants and yeast overexpressing *CYP82C2* or *CYP82C4*, a hydroxyl group is added to the 5 position of 8-MOP, creating 5-hydroxy-8-methoxypsoralen (Figure 1, compound **2**). In plants, the newly

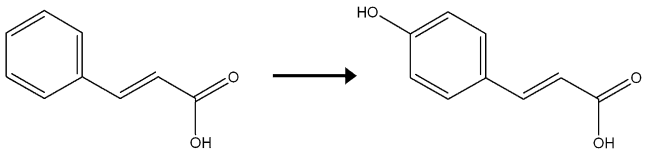


Figure 2. The Reaction Catalyzed by CYP73A5
This enzyme converts cinnamic acid to *p*-coumaric acid.

introduced hydroxyl group is the site of additional glycosylation, presumably via the action of a yet to be identified endogenous glycosyltransferase that subsequently adds a sugar moiety to the 5 position, thus producing the 8-methoxypsoralen glycoside that has been tentatively identified as the compound, 5-*O*-β-D-glucopyranosyl-8-methoxypsoralen (Figure 1, compound **3**). The biocatalytic conversion of 8-methoxypsoralen by CYP82C-subfamily enzymes is only one fully characterized example of many bioconversions identified in our screening panels encompassing 91 P450s. The diversity of P450s and other biosynthetic enzymes in the Plant Kingdom provide a rich supply of biocatalytic diversity. Such enzymes may provide the keys to the catalysis of a vast array of valuable chemical reactions and can provide new resources for biotechnological applications in the future.

RESULTS

Plant In Vivo Screen Development

An in vivo plant assay was developed to easily screen substrates of interest against 91 *Arabidopsis* transgenic lines, each overexpressing a different P450 gene. Initial method development was conducted by employing the well-characterized enzyme CYP73A5 (cinnamate-4-hydroxylase [C4H]), which is known to convert cinnamic acid to *p*-coumaric acid in one of the early steps of phenylpropanoid and lignin biosynthesis (Figure 2) (Urban et al., 1994; Pierrel et al., 1994). Under our experimental conditions, a transgenic line overexpressing *CYP73A5* converted cinnamic acid (m/z = 147 $[M+H]^+$) (Figure 3A) to coumaric acid (m/z = 163 $[M+H]^+$) (Figure 3C). This product was absent in a control sample that did not contain any substrate (Figure 3E). The assay system and associated LC/MS detection method was capable of detecting nanogram amounts of substrate and product, and both the media and the plant tissue contained the product (data not shown). Without the benefit of overexpression (OE), the wild-type level of endogenous CYP73A5 under these conditions was not sufficient to produce a detectable amount of product. This may be due to the fact that, with a few exceptions, plant P450s have low and/or cell-specific expression patterns that do not allow for accumulation of a significant amount of products (Duan and Schuler, 2006). Additionally, there is likely coordinate expression and tight control of metabolic precursors under wild-type conditions. A second model P450 line used in the method development was a transgenic plant line overexpressing *CYP84A1* (ferulate 5-hydroxylase [F5H]). This enzyme is known to hydroxylate a number of substrates, with coniferaldehyde believed to be the best substrate (Humphreys et al., 1999). In our assay system, plant lines overexpressing *CYP84A1* in the presence of 1 mM coniferaldehyde (m/z = 177 $[M+H]^+$) accumulated a new peak at m/z = 193 $[M+H]^+$, consistent with the addition of a hydroxyl group (Figure 3D). This

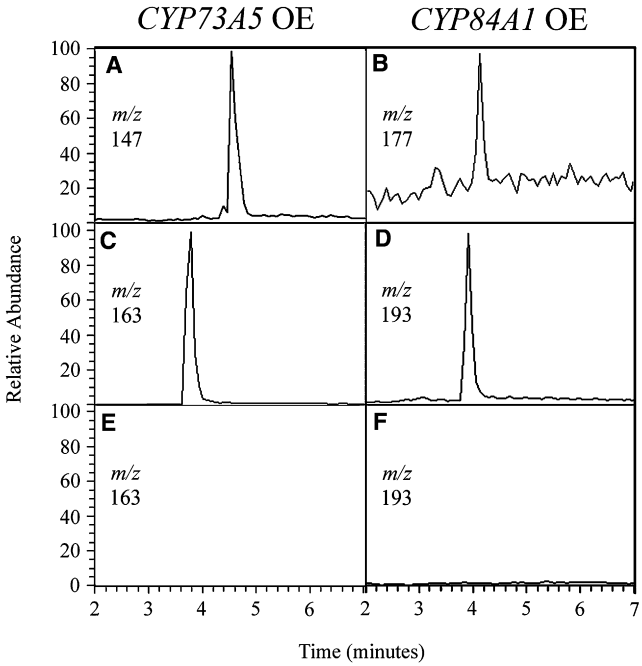


Figure 3. LC/MS Analysis of Model Experiments Demonstrating Successful Hydroxylation Using Characterized P450s

(A–F) The left column of chromatograms displays data from a *CYP73A5* OE line, whereas the right column displays data from a *CYP84A1* OE line.

(A and B) Single ion chromatograms of *m/z* = 147, the cinnamic acid substrate, and *m/z* = 177, the coniferaldehyde substrate, respectively.

(C and D) Single ion chromatograms of *m/z* = 163 and *m/z* = 193, the hydroxylated products.

(E and F) Single ion chromatograms of *m/z* = 163 and *m/z* = 193 of plants that were not exposed to substrate.

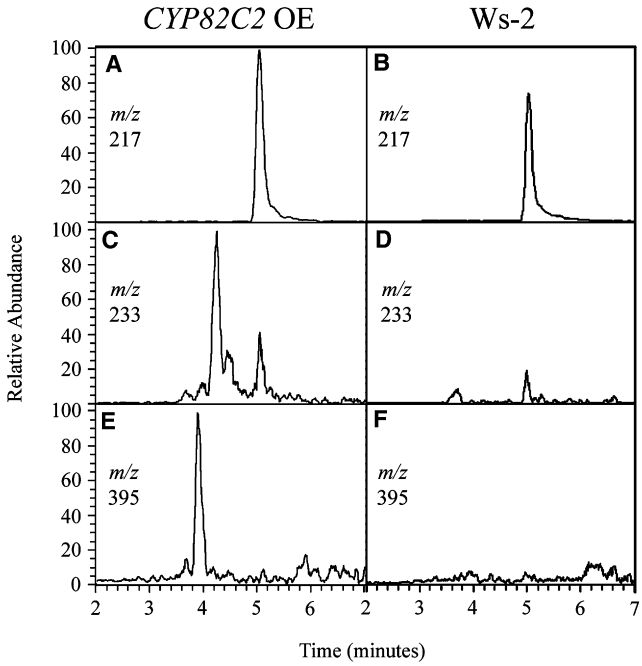


Figure 4. LC/MS Analysis from the Plant In Vivo Screen Using 8-MOP as the Substrate

(A–F) The left column of chromatograms displays data from the *CYP82C2* OE line, whereas the right column displays data from a Ws-2 control.

(A and B) Single ion chromatograms of *m/z* = 217, the 8-MOP substrate.

(C and D) Single ion chromatograms of *m/z* = 233, the appropriate mass for a hydroxylated product.

(E and F) Single ion chromatograms of *m/z* = 395, the appropriate mass for a hydroxylated product that has been subsequently glycosylated.

peak was absent in a control sample that did not contain any substrate (Figure 3F). Once experimental conditions were established and validated by employing these well-characterized P450s, the more extensive screen was established and run.

As the masses of each substrate were known, it was possible to develop rapid liquid chromatography/mass spectrometry (LC/MS) methods to specifically detect the masses of predicted products based on the known structural conversions of P450s. Additionally, with the wealth of secondary metabolic enzymes present in plants, it was possible and quite likely that further functionalization (such as methylation or glycosylation) could occur at the newly created hydroxyl group(s). In an attempt to capture and identify this additional diversity, selected ion chromatograms (SICs) were created for masses corresponding to the predicted products of 18 different chemical modifications, including such reactions as simple hydroxylation (+16 amu), demethylation (−14 amu), and more complicated reactions such as a hydroxylation followed by glycosylation (+178 amu). A data analysis tool developed in house allowed for quick analysis of all data, and for the comparison of each single ion chromatogram to that of the wild-type *Arabidopsis* Wassilewskija (Ws-2) control exposed to the same substrate. The screen was qualitative, seeking peaks above the level of the baseline established by the Ws-2 control. Many peaks of interest were identified in the screen; out of the 100 substrates tested, 8 compounds resulted in products that were present in a primary screen and in a subsequent confirmation screen. These 8 were in addition to those resulting from the successful tests done with the lines overexpressing *CYP73A5* and *CYP84A1* that are described above as part of the screening method development. One of the substrates that provided positive results was 8-methoxypsoralen (8-MOP, **1**).

Plant In Vivo Screen Using 8-MOP as the Substrate

After a 1 day incubation of the transgenic line bearing the *CYP82C2* gene in the presence of 1 mM 8-MOP, there was still a significant amount of substrate present in extracts of the seedlings and media, detected at *m/z* = 217 $[M+H]^+$ (Figures 4A and 4B). However, there was a significant new peak at *m/z* = 233 $[M+H]^+$, consistent with the addition of a hydroxyl group in the *CYP82C2* OE line when compared with the wild-type Ws-2 control seedlings (Figures 4C and 4D). This peak was unique to *Arabidopsis* bearing the *CYP82C2* gene; no other plant lines in the panel of 91 OE lines generated this peak. Furthermore, a second, new peak at *m/z* = 395 $[M+H]^+$ was identified in the OE line, consistent with the subsequent addition of a glycosyl group (Figures 4E and 4F). The relative elution positions of the new peaks were also consistent with the increases in polarity of the proposed compounds. These data were generated by using a mixture of seeds from five separate transformation events. Each of these events was then tested individually by using the

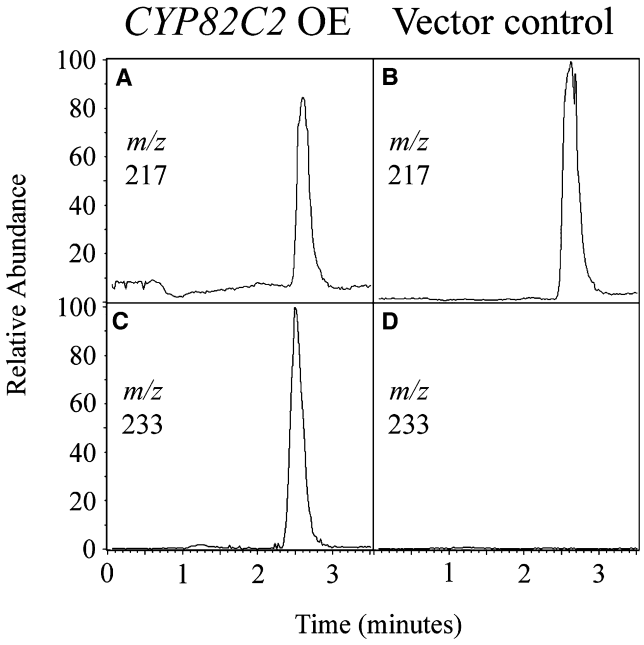


Figure 5. LC/MS Analysis from Yeast in Vivo Tests Using 8-MOP as the Substrate

(A–D) The left column of chromatograms displays data from yeast overexpressing *CYP82C2*, whereas the right column displays data from a CRS905 empty vector control.

(A and B) Single ion chromatograms of *m/z* = 217, the 8-MOP substrate.

(C and D) Single ion chromatograms of *m/z* = 233, the mass of the hydroxylated product.

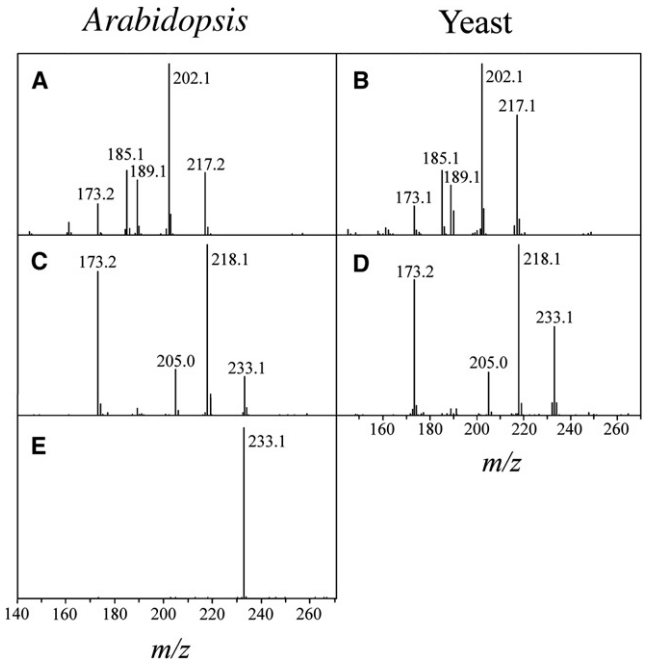


Figure 6. LC/MS/MS Analysis of the 8-MOP Substrate and Its Adducts

(A–E) The left column displays spectra from *Arabidopsis*, whereas the right column displays data from yeast, both overexpressing *CYP82C2*.

(A and B) LC/MS/MS fragmentation pattern of *m/z* = 217.

(C and D) LC/MS/MS fragmentation pattern of *m/z* = 233.

(E) LC/MS/MS fragmentation pattern of *m/z* = 395.

plant in vivo screen. Four out of the five events showed positive results, displaying the new peaks at *m/z* = 233 $[M+H]^+$ and *m/z* = 395 $[M+H]^+$, whereas event #1 failed to show any activity (data not shown).

Yeast In Vivo Tests Using 8-MOP as the Substrate

In order to be able to produce metabolites quickly, easily, and in large quantities for purification, we explored a yeast OE system. The *CYP82C2* cDNA construct was expressed in WAT11 yeast (Pompon et al., 1996), and a similar in vivo assay was conducted. Figure 5 displays the SICs for the substrate (Figures 5A and 5B) and the hydroxylated product (Figures 5C and 5D) in yeast bearing the *Arabidopsis CYP82C2* gene as compared with an empty vector control. Yeast expressing CYP82C2 also produce the peak at *m/z* = 233 $[M+H]^+$. Not surprisingly, subsequent glycosylation was not seen in the yeast expression system; the ion at *m/z* = 395 $[M+H]^+$ was not detected in either sample (data not shown). In both the plant and yeast in vivo systems, the product was found not only in the plant tissue or yeast pellet, but was identified in the media, indicating that the product was partially excreted. In vitro tests with membranes isolated from this yeast strain supported the in vivo data (data not shown).

LC/MS/MS and NMR Analyses of Products

To gain further information on the nature of the products, LC/MS/MS was performed on samples from both the plant and yeast *CYP82C2* OE lines to look for common fragmentation patterns in the substrates and products. Figures 6A and 6B show the fragmentation pattern of the MOP substrate (*m/z* = 217) found in the *Arabidopsis* and yeast samples, respectively. The most abundant fragment ions are at *m/z* = 202 (−15), at 189 (−28 amu), at 185 (−32 amu), and at 173 (−44 amu). Figures 6C and 6D display the fragmentation patterns of the product at *m/z* = 233 found in the plant and yeast samples. These masses show common patterns with that of the 8-MOP substrate, including ions at *m/z* = 218 (−15 amu), at *m/z* = 205 (−28 amu), and at *m/z* = 173. Only the plant sample contained the unique ion at *m/z* = 395, and its only fragment ion was at *m/z* = 233, consistent with the loss of the sugar moiety (Figure 6E). These data provided more conclusive evidence that the ion at *m/z* = 233 represents a hydroxylated form of 8-MOP and the ion at *m/z* = 395 is a glycosylated form of 8-MOP. Both the yeast and plant systems generated the same hydroxylated metabolite, providing further evidence that CYP82C2 is the sole enzyme responsible for the activity.

Nuclear magnetic resonance (NMR) data were acquired to conclusively determine the exact positions and structures of the newly added substituents. The proposed hydroxylated product (*m/z* = 233) was easily isolated from 1L of culture of the yeast strain bearing the *CYP82C2* gene. After methanol extraction and fractionation, the purified peak was analyzed by ^{1}H-NMR and compared with the results from the substrate (**1**). Five resonances were identified in the ^{1}H-NMR spectrum of the proposed hydroxylated product (four methine doublets at δ 8.32, δ 7.76,

δ 7.08, δ 6.20, and one methyl group at δ 4.05), thus verifying the addition of a hydroxyl group to compound **1**. 2D-NMR revealed two ^{1}H-^{1}H COSY relationships between resonances at δ 7.76 and δ 7.08 as well as at δ 8.32 and δ 6.20, confirming that the hydroxyl group was added at C-5, resulting in 5-hydroxy-8-methoxypsoralen (**2**). The glycosylated adduct was scaled up, extracted, and purified from transgenic *Arabidopsis* seedlings bearing the *CYP82C2* gene. The ^{1}H-NMR spectrum of this product revealed four methine doublets (δ 8.54, δ 7.80, δ 7.30, and δ 6.33) for the psoralen ring structure, seven protons (δ 4.50, δ 3.85, δ 3.73, δ 3.58, δ 3.46, δ 3.44, δ 3.28) for a sugar moiety, and one methoxy group at δ 4.19. Compared with the ^{1}H-NMR spectrum of compound **2**, two methine resonances at δ 7.80 and δ 7.30 in the furan ring and at δ 8.54 and δ 6.33 in the pyran lactone ring remained; therefore, the sugar moiety was present at C-5. The large coupling constants of the anomeric proton at δ 4.50 (J = 8.0 Hz, H-1″) and the C-2 proton at δ 3.58 (J = 8.0 Hz, 7.0 Hz, H-2″) indicated that the C-1, C-2, and C-3 protons were all axial. The NMR data, together with the fact that a majority of plant Family 1 glycosyltransferases, which modify small molecules, use UDP-glucose in the transfer reaction (Lim and Bowles, 2004), indicated that 5-*O*-β-D-glucopyranosyl-8-methoxypsoralen (**3**) is the most likely product of this secondary reaction in plants.

Analysis of Related P450s

Seedlings overexpressing *CYP82C4* and *CYP82F1*, two related P450s, were also available in our larger collection of *Arabidopsis* OE lines, and these were tested in the plant in vivo system. Plants overexpressing *CYP82C4*, an isoform of *CYP82C2*, were in fact able to hydroxylate and subsequently glycosylate 8-MOP, and the products accumulated at similar levels to that of the *CYP82C2* OE line (data not shown). However, extracts from seedlings overexpressing *CYP82F1*, a P450 from a different subfamily, did not contain a modified form of 8-MOP. Similar results were found when these same P450s were overexpressed in yeast WAT11; CYP82C4 hydroxylated 8-MOP, whereas CYP82F1 did not (data not shown). The related P450, *CYP82C3*, was represented in our original panel of 91 lines; however, it was found to have a non-natural point mutation resulting in an incomplete protein and thus did not show activity with the 8-MOP substrate. Based on the high homology between the CYP82C isoforms (91%–92% identical), it is likely that a functional CYP82C3 protein may be capable of hydroxylating 8-MOP as well.

Segregation analysis was performed on *Arabidopsis* lines overexpressing *CYP82C2* and *CYP82C4* to determine the copy number of the transgenes they contained. The resulting ratio of BASTA-resistant to BASTA-insensitive individuals clearly indicated that the transgenes were present in a single copy in all transgenic events. In addition, RT-PCR was performed in order to confirm the OE of the *CYP82C2* and *CYP82C4* genes (Supplemental Data).

DISCUSSION

An in planta method to identify putative substrates for P450 enzymes has been established and validated by screening a subset of 91 *Arabidopsis* P450s. This method can be applied to a broader set of P450 enzymes and potentially to other biosynthetic enzyme families to provide biocatalytic reagents that can be employed in a variety of pharmaceutical, chemical, bioremediation, and other industrial processes, and to expand the general knowledge base of plant secondary metabolism. Use of a plant-based screen has the advantage of exposing the introduced enzyme, substrate, and product(s) to related endogenous enzymes that would not be present in other model systems like yeast. The presence of endogenous enzymes may be valuable by further metabolizing the products of the introduced enzyme, as in the production of 5-*O*-β-D-glucopyranosyl-8-methoxypsoralen (**3**) described here, thus increasing the opportunities to identify novel structures and further elucidate biosynthetic pathways. There is increasing evidence that P450s are often associated with metabolons (Ralston and Yu, 2006), and, as such, the presence of endogenous enzymes may be necessary for catalytic efficiency of the introduced P450. Finally, on a more practical level, by using a plant system one avoids potential complications of heterologous protein OE like variations in codon usage.

One result from these initial experiments identified 8-MOP as a substrate for CYP82C2 and CYP82C4 from *Arabidopsis*. 8-MOP is a furanocoumarin, or psoralen, that is primarily found in plants belonging to the Apiaceae and Rutacea families. Furanocoumarins are likely part of the chemical defense systems of plants, functioning as phytoalexins, and are toxic to a wide variety of organisms, including fungi, bacteria, insects, and mammals (Berenbaum and Zangerl, 1996). The biosynthesis of furanocoumarins can be induced by various stresses, including fungal infection (Tietjen et al., 1983), insect herbivory (Zangerl and Berenbaum, 1990), and atmospheric pollution (Dercks et al., 1990).

Furanocoumarins are able to intercalate with DNA, and some members of this structure class have been found to have therapeutic effects, most notably 5-methoxypsoralen (bergapten, 5-MOP) and 8-MOP. 8-MOP and 5-MOP are photosensitizers (Pathak et al., 1981) and have been administered orally prior to UVA irradiation in order to augment the phototherapy treatment for psoriasis and vitiligo (Honingsmann et al., 1979; Hann et al., 1991). Furanocoumarins have also been implicated in food-drug interactions; for instance, furanocoumarins present in grapefruit juice effect the bioavailability of certain drugs by inhibiting human CYP3A4. In fact, 5-MOP was shown to be the most potent inhibitor of CYP3A4 (Ho et al., 2001).

It is intriguing that the CYP82C2 and CYP82C4 enzymes can use 8-MOP as a substrate since *Arabidopsis* has not been reported to produce furanocoumarins. The *CYP82* family has not yet been assigned to a metabolic pathway in *Arabidopsis*; their true endogenous substrates are currently unknown, although one might conclude from these data that their substrates may have some structural similarities to 8-MOP or that this pathway is only induced under environmental or stress conditions that have yet to be studied. Members of the *CYP82* family have only been identified in dicots thus far; the *CYP82C* subfamily is restricted to three species, *Glycine max* (soybean), *Populus trichocarpa* (black cottonwood), and *Arabidopsis thaliana* (http://drnelson.utmem.edu/CytochromeP450.html). Several CYP82-family members have been found to be expressed in response to wounding or other stress elicitors (Frank et al., 1996; Takemoto et al., 1999). Only one member of the CYP82 family has been characterized and acts on a known substrate.

CYP82E4v1, identified in tobacco, was found to function as a nicotine demethylase, converting nicotine to nornicotine (Siminszky et al., 2005).

The plant P450 screening platform described here has been used to identify 8-MOP as a substrate for *Arabidopsis* CYP82C2 and CYP82C4, leading to the formation of both 5-hydroxy-8-methoxypsoralen (**2**) and a glycosylated form of 5-hydroxy-8-methoxypsoralen (**3**). It is possible that adducts of pharmacologically active compounds, like 8-MOP, formed in this screen could lead to new therapies or those with greater efficacy. Although the application of P450s to industrial processes has been relatively limited thus far, there is no doubting their potential. There is significant and promising progress in the application of this impressive class of enzymes (Bernhardt, 2006). Even the identification of enzymes that function at low catalytic rates could be very valuable, as the identified enzyme activities could be optimized in whole-cell or cell-free systems, and the enzymes could be engineered for greater catalytic activity or greater specificity via directed evolution or chimeragenesis (Bernhardt, 2006; Gillam, 2005; Valetti and Gilardi, 2004). Recent advances in technology like the use of nanodiscs (Duan and Schuler, 2006) and modified oxygen biosensors (Olry et al., 2007) may be helpful in characterizing identified P450s and their variants.

Much still needs to be done in order to identify the endogenous and non-native substrates that the vast array of plant P450s are capable of acting on, but such work will provide a wealth of information on plant secondary metabolic pathways and will help expand the utility of this large superfamily of enzymes in many new biocatalytic processes for the production of chemicals, intermediates, and products in the pharmaceutical and chemical industries. The number and diversity of plant P450s and other biosynthetic enzymes combined with the screening platforms in plants and yeast described herein will expand the set of tools available for such work.

SIGNIFICANCE

Although some effort has gone into the identification and classification of the thousands of P450s known to exist in nature, much work still needs to be done to determine the potential substrates utilized and the products generated by this varied and versatile class of enzymes. An in vivo plant screening platform that allows for the analysis of potential substrate/P450 interactions has been developed. One example demonstrated here is the identification of a substrate, 8-MOP, for members of the CYP82C subfamily, enzymes whose natural substrates were unknown. In addition to the hydroxylation provided by the P450s themselves, the plants were found to further modify this product to create a glycosylated compound, likely the compound 5-*O*-β-D-glucopyranosyl-8-methoxypsoralen. This platform utilizing sensitive analytical chemistry techniques demonstrates the ability to identify novel adducts of therapeutic compounds, thus providing important information that could lead to the development of other novel compounds of pharmacological value. In addition, specific enzymes that prove to be useful biocatalysts in such processes as bioremediation and fine chemical synthesis may be discovered.

EXPERIMENTAL PROCEDURES

Transgenic Plant Construction

The *Arabidopsis* ecotype used throughout was Wassilewskija (Ws-2). General plant handling and transformation protocols followed standard procedures (Weigel and Glazebrook, 2002). CRS338, the vector used for creating the plant transgenic lines, was constructed by modifying CRS301 (Schneeberger et al., 2005). The BstXI sites of CRS301 were destroyed by digesting the vector with BstXI, treating the ends to create blunt ends, and ligating the vector back together. A 35S:hYFP:OCS modular cassette with BstXI tails surrounding hYFP was created in a second vector, A7SK, and was subsequently moved into the modified CRS301 construct. Finally, the hYFP was removed via digestion with BstXI, and a BstXI-tailed lacZ stuffer sequence was ligated in from pUC18. This final vector was digested with BstXI, releasing the LacZ insert, which was replaced with *Arabidopsis* P450 inserts.

Yeast Construction

A new vector, CRS905, was created to overexpress the *Arabidopsis* P450s in yeast. The first step in creating CRS905 was to replace the GAL promoter of pYES3 (Lu et al., 1997) with the PGK promoter; then, a stuffer sequence containing forward and reverse BstXI sites was inserted into the multicloning site. Lastly, the LacZ gene was subcloned between the two BstXI sites. This final vector was digested with BstXI, releasing the LacZ insert, which was replaced with *Arabidopsis* P450 inserts. The P450 inserts were generated via two methods. The first method involved digestion of the original plant CRS338 vectors containing the cDNAs of the P450s of interest with BstXI and ligation of this insert with cut CRS905. The second method involved amplification of the P450 of interest from either a wild-type *Arabidopsis* cDNA library or from cDNA generated via RT-PCR from RNA isolated from plants overexpressing the P450 of interest, using primers containing BstXI linkers for ligation into the cut CRS905 vector. The *Saccharomyces cerevisiae* strain WAT11, a derivative of W303-1B (*MATa; ade2-; his3-11, −15; leu2-3, −112; ura3-1; can^R; cyr$^+$*), was transformed with isolated plasmid (Elble, 1992) and was selected on AHC+Trp/Dextrose plates. This strain expresses ATR1, the *Arabidopsis* NADPH-P450 reductase (Pompon et al., 1996). AHC+Trp growth media consisted of 0.17% yeast nitrogen base, 0.5% ammonium sulfate, 1% casamino acids, 0.003% adenine sulfate, 0.002% tryptophan, 0.788% Tris-HCl, and either 2% dextrose or galactose, and included 2% Bactoagar for plate media. Yeast were grown at 28°C.

Arabidopsis In Vivo Assay

For each transgenic line in the primary screen, seeds from all available transformation events were mixed together to form a single sample pool representing one P450 OE line. In confirmation screens, the seeds from each individual event were tested separately. *Arabidopsis* seeds were surface sterilized with a 25% (v/v) bleach solution containing 0.1% (v/v) Triton X-100 for 5 min and stored at 4°C until use. Approximately 25 seeds were added per well, and each experiment included Ws-2 and no seedling (media only) controls. Plants were grown in 24-well microtiter plates containing 1.25 ml ½ x MS media with slow shaking in a 22°C chamber under long daylight (16 hr light, 8 hr dark) conditions. The seedlings were allowed to germinate and develop for 7 days. Substrate was prepared as a 100 mM stock in DMSO and was added to each well to a final concentration of 1 mM. The seedlings were incubated in the presence of substrate for 24 hr at 22°C, and the wet tissues were subsequently lysed via bead homogenization by using a QIAGEN TissueLyser and 1.5 ml tubes. The crushed tissue was extracted with 1 ml MeOH for 1 hr at 37°C with shaking. After centrifugation at 15,500 × g for 3 min, the supernatants were moved to glass tubes containing the media from the microtiter wells. A total of 1 ml hexane was added to the aqueous MeOH extracts, and the tubes were vortexed. The mixtures were equilibrated for 15 min, and then centrifuged at 3,200 × g for 15 min. The top hexane layers were discarded, and the remaining extracts were dried overnight in a Savant AES2010 SpeedVac. Each sample was resuspended in 150 μl MeOH, sonicated in a water bath to assist resuspension, and filtered through a 0.45 μm syringe filter (Alltech) before analysis by LC/MS.

Yeast In Vivo Assay

Yeast were inoculated into 500 μl AHC+Trp/Dextrose media in a 96-well deep-well plate and incubated for 3 days at 28°C before being diluted 1:50 into 500 μl

AHC+Trp/Galactose. After a 24 hr incubation, 5 μl 100 mM 8-MOP was added, and the cultures were allowed to incubate for another 24 hr. Yeast were lysed via glass bead lysis (425–600 μm, Sigma). The samples were extracted with the addition of 500 μl MeOH and incubation at 37°C for 1 hr with shaking. The plate was centrifuged at 800 × g for 5 min, and the supernatant was collected for analysis by LC/MS.

Analytical LC/MS Methods

The yeast samples were separated and analyzed on an Alliance HT 2795 HPLC system connected to a Micromass ZMD single quadrupole mass spectrometer (Waters) with an XTerra MSC_{18} 5 μm 2.1 × 20 mm IS column (Waters) and a gradient solvent system of H_2O + 0.1% formic acid in MeOH + 0.1% formic acid (95:5 held for 0.2 min, followed by a gradient from 95:5 to 5:95 over 1.3 min, followed by a return to initial conditions and a hold for 2 min; total run length of 3.5 min with a flow rate 0.5 ml/min). For method development of the plant screens using cinnamic acid and coniferaldehyde as substrates, the plant extracts were separated and analyzed on an Alliance HT 2795 HPLC system connected to a Micromass ZMD single quadrupole mass spectrometer (Waters) with an XTerra MS C_{18} 5 μm 2.1 × 50 mm column (Waters) and a gradient solvent system of H_2O + 0.1% formic acid (95%–5%) and acetonitrile + 0.1% formic acid (95:5 held for 1 min, followed by a gradient from 95:5 to 5:95 over 6 min, followed by a return to initial conditions over 2 min and a final hold at initial conditions for 1 min; total run length of 10 min with a flow rate 0.5 ml/min). In the initial plant screens, the plant extracts were separated and analyzed on a Finnigan Surveyor Plus HPLC system connected to a Finnigan LCQ Deca Max XP ion trap mass spectrometer (Thermo Electron Corp.), using the same column and gradient settings as described above. Subsequent analysis of plant tissue and of all plant and yeast fractions generated during purification procedures was done similarly, but with one exception: acetonitrile was replaced by MeOH.

NMR Analysis

NMR experiments were performed in CD_3OD and $CDCl_3$ by using the residual proton signals as reference standards (δ_H 3.31 for CD_3OD and δ_H 7.26 for $CDCl_3$). ^{1}H-NMR data were obtained at room temperature on a Bruker Avance 600 MHz spectrometer operating at 600.01 MHz.

Isolation of Hydroxylated Product and NMR Analysis

The hydroxylated product of 8-MOP (**2**) was easily scaled up in yeast. 1 L AHC+Trp/Galactose media was inoculated with WAT11 bearing *CYP82C2* that had been grown to saturation in AHC+Trp/Dextrose. The cultures were incubated at 28°C with shaking for 1 day. A total of 1 ml 500 mM 8-MOP prepared in DMSO was added to the flask, for a final substrate concentration of 500 μM. The culture was incubated for another 3 days. Yeast culture media was transferred to a round-bottom flask. The media were removed by rotoevaporation, and the resulting residue was resuspended in 20 ml MeOH. The MeOH suspension was transferred to a 50 ml centrifuge tube, and the sample was centrifuged for 1 hr at 1,800 × g. The supernatant was transferred to a round-bottom flask and removed by rotoevaporation, yielding 6.5 g crude dry extract. Solid-phase extraction of this material on a C_{18} Extract-Clean column (Alltech) with a block gradient of H_2O/MeOH resulted in 150 mg in the 25% H_2O/MeOH fraction. This fraction was further separated by using a C_{18} 5 μm 250 × 10 mm column (Phenomenex) with a gradient solvent system of H_2O + 0.1% formic acid (65%–0% in 28 min) in MeOH + 0.1% formic acid (flow rate 2.0 ml/min, detection at 300 nm) to yield three major peaks. Subsequent LC/MS analysis confirmed that peak 2, eluting at 22.2 min, was the target compound (**2**) and that peak 3, eluting at 22.5 min, was the substrate (**1**). Compound **2** was obtained as a white, amorphorous solid; UV (MeOH) λ max: 315 nm; LRAPCIMS *m/z* = 232.19 ([M+H]$^+$); ^{1}H-NMR δ 8.32 (1H, d, *J* = 10 Hz, H-4), δ 7.76 (1H, d, *J* = 2Hz, H-2′), δ 7.08 (1H, d, *J* = 2 Hz, H-3′), δ 6.2 (1H, d, *J* = 10 Hz, H-3), δ 4.05 (3H, s, -OCH_3).

Isolation of Glycosylated Product and NMR Analysis

Four 24-well microtiter plates of *Arabidopsis* seedlings overexpressing *CYP82C2* were grown, and substrate was added as per the standard plant in vivo assay described above. Seedlings from each 24-well microtiter plate were transferred into a 145 ml mortar, ground thoroughly with a pestle, and transferred into a 50 ml centrifuge tube; the substance was rinsed with MeOH for final volume of 35 ml. Tubes containing crushed seedlings were placed in a sonication bath (Branson 2510) for 30 min at room temperature, then chilled in ice-water for 5 min; this cycle was repeated six times. Tubes were then shaken at ambient temperature overnight. The extraction mixture was vacuum filtered through a coarse sintered glass funnel and filter paper (Whatman); the filtrate was reduced to <2 ml by rotary evaporation, resuspended with 15% acetonitrile, frozen, and lyophilized to dryness. The dry extract was dissolved in MeOH for semipreparative HPLC (Waters 600), by using the following chromatographic conditions: a C_{18} 5 μm 150 × 10 mm column (Alltech Alltima); solvent A, H_2O + 0.1% formic acid; solvent B, MeOH + 0.1% formic acid; flow rate of 1.5 ml/min; monitoring at 308 nm. Three major peaks eluted with a gradient of 20%–80% B in 30 min. The presence of the glycosylated product in the first two peaks was verified by analytical LC/MS. Fractions for peaks 1 and 2, eluting at 26.0 min and 27.2 min, respectively, were pooled and reduced to <5 ml by rotary evaporation. The concentrated peaks were mixed with 15% acetonitrile, frozen, and lyophilized to dryness. Peak 2 was redissolved in MeOH and further separated by using a gradient of 40%–50% B in 35 min, yielding the purified glycosylated product at 20.7 min. Compound **3** was obtained as a white, amorphorous solid; UV (MeOH) λ max: 308 nm; LRAPCIMS *m/z* = 395.07 ([M+H]$^+$); ^{1}H-NMR δ 8.54 (1H, d, *J* = 10 Hz, H-4), δ 7.80 (1H, d, *J* = 2Hz, H-2′), δ 7.30 (1H, d, *J* = 2.3 Hz, H-3′), δ 6.33 (1H, d, *J* = 10 Hz, H-3), δ 4.50 (1H, d, *J* = 8.0 Hz, H-1′), δ 4.19 (3H, s, -OCH_3), δ 3.85 (1H, dd, *J* = 12 Hz, 2.1 Hz, H-6″), δ 3.73 (1H, dd, *J* = 12 Hz, 4.0 Hz, H-6″), δ 3.58 (1H, dd, *J* = 8.0 Hz, 7.0 Hz, H-2″), δ 3.46 (1H, m, H-4″), δ 3.44 (1H, m, H-3″), δ 3.28 (1H, m, H-5″).

Segregation Analysis

Arabidopsis lines overexpressing *CYP82C2* or *CYP82C4* were subjected to segregation analysis to determine the copy number of the transgenes they contained. Approximately 200 seeds from 5 separate events from each line, along with a Ws-2 control, were sterilized and plated onto ½ × MS + 200 μl/l 5.78% BASTA (glufosinate ammonium, Farnam Companies) media. The plates were incubated at 22°C in long daylight conditions for 11 days. The seedlings were scored, and the ratio of BASTA-resistant to BASTA-sensitive strains was determined.

SUPPLEMENTAL DATA

Supplemental Data include the complete list of 91 P450s used in the plant in vivo screen as well as details on RT-PCR performed in order to confirm the overexpression of the *CYP82C2* and *CYP82C4* genes and are available at http://www.chembiol.com/cgi/content/full/15/2/149/DC1/.

ACKNOWLEDGMENTS

We thank the Cloning, Transformation, and Culture Collection teams at Ceres for generating the *Arabidopsis* and yeast overexpression lines and for maintaining the plants and seeds. We also thank Erich Grotewold for his comments on the manuscript.

Received: January 12, 2007
Revised: October 30, 2007
Accepted: January 24, 2008
Published: February 22, 2008

REFERENCES

Berenbaum, M.R., and Zangerl, A.R. (1996). Phytochemical Diversity and Redundancy in Ecological Interactions, *Volume 30* (New York: Elsevier Press).

Bernhardt, R. (2006). Cytochromes P450 as versatile biocatalysts. J. Biotechnol. *124*, 128–145.

Bezalel, L., Hadar, Y., Fu, P.P., Freeman, J.P., and Cerniglia, C.F. (1996). Initial oxidation products in the metabolism of pyrene, anthracene, fluorene, and dibenzothiophene by white rot fungus *Pleurotus ostreatus*. Appl. Environ. Microbiol. *62*, 2554–2559.

Chau, M., and Croteau, R. (2004). Molecular cloning and characterization of a cytochrome P450 taxoid 2α-hydroxylase involved in Taxol biosynthesis. Arch. Biochem. Biophys. *427*, 48–57.

Dercks, W., Trumble, J., and Winter, C. (1990). Impact of atmospheric pollution alters linear furanocoumarin content in celery. J. Chem. Ecol. *16*, 443–454.

Duan, H., and Schuler, M.A. (2006). Heterologous expression and strategies for encapsulation of membrane-localized plant P450s. Phytochem. Rev. *5*, 507–523.

Duport, C., Spagnoli, R., Degryse, E., and Pompon, D. (1998). Self-sufficient biosynthesis of pregnenolone and progesterone in engineered yeast. Nat. Biotechnol. *16*, 186–189.

Elble, R. (1992). A simple and efficient procedure for transformation of yeasts. Biotechniques *13*, 18–20.

Frank, M.R., Deyneka, J.M., and Schuler, M.A. (1996). Cloning of wound-induced cytochrome P450 monooxygenase expressed in pea. Plant Physiol. *110*, 1035–1046.

Gillam, E.M.J. (2005). Exploring the potential of xenobiotic-metabolising enzymes as biocatalysts: evolving designer catalysts from polyfunctional cytochrome P450 enzymes. Clin. Exp. Pharmacol. Physiol. *32*, 147–152.

Hann, S.K., Cho, M.Y., Im, S., and Park, Y.K. (1991). Treatment of vitiligo with oral 5-methoxypsoralen. J. Dermatol. *18*, 324–329.

Ho, P.C., Saville, D.J., and Wanwimolruk, S. (2001). Inhibition of human CYP3A4 activity by grapefruit flavonoids, furanocoumarins and related compounds. J. Pharm. Pharm. Sci. *4*, 217–227.

Honingsmann, H., Jaschke, E., Gschnait, F., Brenner, W., Fritsch, P., and Wolff, K. (1979). 5-methoxypsoralen (bergapten) in photochemotherapy of psoriasis. Br. J. Dermatol. *101*, 369–378.

Humphreys, J.M., Hemm, M.R., and Chapple, C. (1999). New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. Proc. Natl. Acad. Sci. USA *96*, 10045–10050.

Kahn, R., and Durst, F. (2000). Function and evolution of plant cytochrome P450. Recent Adv. Phytochem. *34*, 151–189.

Lim, E.-K., and Bowles, D.J. (2004). A class of plant glycosyltransferases involved in cellular homeostasis. EMBO J. *23*, 2915–2922.

Lu, Y.-P., Li, Z.-S., and Rea, P.A. (1997). AtMRP1 gene of *Arabidopsis* encodes a glutathione S-conjugate pump: isolation and functional definition of a plant ATP-binding cassette transporter gene. Proc. Natl. Acad. Sci. USA *94*, 8243–8248.

Nelson, D.R., Schuler, M.A., Paquette, S.M., Werck-Reichhart, D., and Bak, S. (2004). Comparative genomics of rice and *Arabidopsis*. Analysis of 727 cytochrome P450 genes and pseudogenes from a monocot and a dicot. Plant Physiol. *135*, 756–772.

Olry, A., Schneider-Belhaddad, F., Heintz, D., and Werck-Reichhart, D. (2007). A medium-throughput screening assay to determine catalytic activities of oxygen-consuming enzymes: a new tool for functional characterization of cytochrome P450 and other oxygenases. Plant J. *51*, 331–340.

Pathak, M.A., Parrish, F.A., and Fitzpatrick, T.B. (1981). Psoralens in photochemotherapy of skin diseases. Farmaco *36*, 479–491.

Picataggio, S., Rohrer, T., Deanda, K., Lanning, D., Reynolds, R., Mielenz, J., and Eirich, L.D. (1992). Metabolic engineering of *Candida tropicalis* for the production of long-chain dicarboxylic acids. Biotechnology (N. Y.) *10*, 894–898.

Pierrel, M.A., Batard, Y., Kazmaier, M., Mignotte-Vieux, C., Durst, F., and Werck-Reichhart, D. (1994). Catalytic properties of the plant cytochrome P450 CYP73 expressed in yeast. Substrate specificity of a cinnamate hydroxylase. Eur. J. Biochem. *224*, 835–844.

Pompon, D., Louerat, B., Bronine, A., and Urban, P. (1996). Yeast expression of animal and plant P450s in optimized redox environments. Methods Enzymol. *272*, 51–64.

Ralston, L., and Yu, O. (2006). Metabolons involving plant cytochrome P450s. Phytochem. Rev. *5*, 459–472.

Schneeberger, R.G., Zhang, K., Tatarinova, T., Troukhan, M., Kwok, S.F., Drais, J., Klinger, K., Orejudos, F., Macy, K., Bhakta, A., et al. (2005). Agrobacterium T-DNA integration in *Arabidopsis* is correlated with DNA sequence compositions that occur frequently in gene promoter regions. Funct. Integr. Genomics *5*, 240–253.

Schuler, M.A., and Werck-Reichhart, D. (2003). Functional genomics of P450s. Annu. Rev. Plant Biol. *54*, 629–667.

Schuler, M.A., Duan, H., Bilgin, M., and Ali, S. (2006). Arabidopsis cytochrome P450s through the looking glass: a window on plant biochemistry. Phytochem. Rev. *5*, 205–237.

Siminszky, B., Gavilano, L., Bowen, S.W., and Dewey, R.E. (2005). Conversion of nicotine to nornicotine in *Nicotiana tabacum* is mediated by CYP82E4, a cytochrome P450 monooxygenase. Proc. Natl. Acad. Sci. USA *102*, 14919–14924.

St-Pierre, B., and De Luca, V. (1995). A cytochrome P-450 monooxygenase catalyzes the first step in the conversion of tabersonin to vindoline in *Catharanthus roseus*. Plant Physiol. *109*, 131–139.

Takemoto, D., Hayashi, M., Doke, N., Nishimura, M., and Kawakita, K. (1999). Molecular cloning of a defense-response-related cytochrome P450 gene from tobacco. Plant Cell Physiol. *40*, 1232–1242.

Tian, L., Musetti, V., Kim, J., Magallanes-Lundback, M., and DellaPenna, D. (2004). The *Arabidopsis* LUT1 locus encodes a member of the cytochrome p450 family that is required for carotenoid ε-ring hydroxylation activity. Proc. Natl. Acad. Sci. USA *101*, 402–407.

Tietjen, K.G., Hunkler, D., and Matern, U. (1983). Differential respone of cultured parsley cells to elicitors from two non-pathogenic strains of fungi. Eur. J. Biochem. *131*, 401–407.

Urban, P., Werck-Reichhart, D., Teutsch, H.G., Durst, F., Regnier, S., Kazmaier, M., and Pompon, D. (1994). Characterization of recombinant plant cinnamate 4-hydroxylase produced in yeast. Eur. J. Biochem. *222*, 843–850.

Valetti, F., and Gilardi, G. (2004). Directed evolution of enzymes for product chemistry. Nat. Prod. Rep. *21*, 490–511.

Weigel, D., and Glazebrook, J. (2002). Arabidopsis: A Laboratory Manual (Plainview, NY: Cold Spring Harbor Laboratory Press).

Werck-Reichhart, D., Bak, S., and Paquette, S. (2002). Cytochromes. In The Arabidopsis Book, C.R. Somerville and E.M. Meyerowitz, eds. (Rockville, MD: American Society of Plant Biologists), p. 450.

Zangerl, A.R., and Berenbaum, M.R. (1990). Furanocoumarin induction in wild parsnip: genetics and populational variation. Ecology *71*, 1933–1940.

