

Functional expression of cytochrome P450 in *Escherichia coli*: An approach to functional analysis of uncharacterized enzymes for flavonoid biosynthesis

Kai Uchida, Tomoyoshi Akashi, Toshio Aoki*

Department of Applied Biological Sciences, Nihon University, Fujisawa, Kanagawa 252-0880, Japan

*E-mail: aoki.toshio@nihon-u.ac.jp Tel: +81-466-84-3939 Fax: +81-466-84-3353

Received April 17, 2015; accepted June 5, 2015 (Edited by Y. Ito)

Abstract Biochemical analyses of metabolic enzymes are performed using in vitro assays with enzymes and substrates. Enzyme samples are generally prepared from heterologous cell expression systems, which are also used as tools to obtain substrates that are not commercially available or cannot be easily isolated from natural sources. Cytochrome P450 (CYP) comprises a widely distributed family of monooxygenases that are commonly used for biosynthesis of natural products. Since CYP activity requires an electron transport system with membrane bound CYP reductase (CPR), it has been believed that CYP is not easily expressed in *Escherichia coli*, despite multiple advantages as a heterologous system compared with other systems that employ eukaryotic yeast and insect cells. In this study, we demonstrated simple and efficient methods for functional expression of CYPs in *E. coli* using commercially available vectors in which the transmembrane-domain truncated CYP was co-expressed with CPR as a discrete polypeptide, and we also used them to identify CYP81E11, CYP81E12, and CYP81E18 of soybean as isoflavone 2'-hydroxylase. Culture conditions were optimized for the bioconversion of I2'H, and the highest production was 161 mg l⁻¹ of medium under optimized conditions. Subsequently, six other CYPs that are involved in flavonoid biosynthesis were tested for their applicability in the *E. coli* expression system. Establishment of the present method may facilitate functional expression of CYPs for preparation of CYP products, including substrates for enzymatic reactions, valuable natural products, and their unnatural derivatives.

Key words: Bioconversion, cytochrome P450, *Escherichia coli*, heterologous expression, isoflavone 2'-hydroxylase.

Functional characterization of genes is generally performed by two approaches: the biochemical and genetic analyses. Recent technological innovations have facilitated complete genome sequencing of many organisms and the accumulation of comprehensive “omics” information. The molecular genetic approach to functional identification of genes is increasingly feasible because of accumulating genomic and post-genomic information. However, the molecular genetic approach is not always available, necessitating continued use of complementary biochemical approaches. Biochemical functional analyses involve in vitro enzymatic assays using enzyme proteins and their substrates; heterologous cell expression systems are popular methods for obtaining these enzymes. Although target enzymes are sometimes purified directly from research materials, heterologous expression systems are useful for producing enzymes with cloned genes or cDNA encoding the target enzyme. Preparation of substrates, which is essential for the biochemical approach, is achieved by isolation

from biological materials, chemical synthesis, or by use of commercial products. If the substrate for an enzyme assay is not readily available, its preparation from a more available precursor using a heterologously expressed enzyme will be a powerful alternative.

Cytochrome P450 (CYP) is a heme-protein of the monooxygenase family, which is widely distributed in multiple organisms. CYPs of plants are involved in the biosynthesis of phytohormones and primary and secondary metabolites, and comprise electron transport systems with CYP reductase (CPR) in membranes such as those of the endoplasmic reticulum. Therefore, eukaryotic cells such as insect and yeast cells, are preferred over prokaryotic cells such as *Escherichia coli*, for heterologous expression of CYPs. However, eukaryotic cell systems have some limitations, and insect cells are not easy to culture and are unsuitable for production purposes because they die after baculovirus transformation. Although culturing yeasts is easier, low accumulation of the recombinant proteins and slow

Abbreviations: ALA, 5-aminolevulinic acid; CPR, cytochrome P450 reductase; CYP, cytochrome P450; F2H, flavanone 2-hydroxylase; FNS II, flavone synthase II; I2'H, isoflavone 2'-hydroxylase; IPTG, isopropyl- β -D-thiogalactopyranoside; IFS, 2-hydroxyisoflavanone synthase; MBR, membrane-binding region.

This article can be found at <http://www.jspcmb.jp/>

Published online July 16, 2015

growth speed are the drawbacks of yeast systems. Thus, an efficient functional CYP expression system in *E. coli* may facilitate biochemical functional characterization of CYPs and biosynthetic preparations of substrates that involve CYPs.

Almost all higher plants produce flavonoids as major secondary metabolites. Leguminous plants also produce a distinct class of flavonoids, isoflavonoids, and leguminous isoflavonoids account for more than 90% of the total isoflavonoids (Dewick 1993). Some isoflavonoids act as phytoalexins and some have beneficial activities for mammals. Among isoflavonoids, glyceollins have a characteristic tetracyclic structure and are known as pterocarpan. Glyceollins are synthesized via glycinol as the major phytoalexin in soybeans (*Glycine max*) and are known for their antibiotic and antitumor activities (Boué et al. 2009; Salvo et al. 2006; Weinstein et al. 1981). Medicarpin is another pterocarpan-type phytoalexin that is distributed in various legume lineages, including *Canavalia*, *Cicer*, *Glycyrrhiza*, *Medicago*, *Melilotus*, and *Trifolium*. Vestitol is biosynthesized by reductive cleavage of the dihydrofran ring of medicarpin (Akashi et al. 2006) and is distributed in various genera, including *Glycyrrhiza*, *Medicago*, and *Lotus*.

Several leguminous CYPs are involved in the biosynthesis of isoflavonoids, and among these 2-hydroxyisoflavanone synthase (IFS), a member of the CYP93C subfamily, plays particularly important roles (Akashi et al. 1999a; Steele et al. 1999). In addition, isoflavone 2'-hydroxylase (I2'H) belongs to the CYP81E subfamily and participates in the initial stage of pterocarpan biosynthesis (Figure 1). Isoflavonoids of legumes vary in the functional group at the 4' position and are classified as 4'-hydroxy and 4'-methoxy types; cDNA clones for I2'H have been isolated from the plants that predominantly accumulate 4'-methoxy type phytoalexin, such as *Glycyrrhiza echinata* (CYP81E1) (Akashi et al. 1998b), *Lotus japonicus* (CYP81E6) (Shimada et al. 2000), and *Medicago truncatula* (CYP81E7) (Liu et al. 2003). I2'H from these plants is highly active towards 4'-methoxy type isoflavones. On the other hand, although no cDNA or gene for I2'H has been identified from soybean, the specificity of its I2'H to 4'-hydroxy substrates is expected because of predominant accumulation of 4'-hydroxy type pterocarpan phytoalexins in soybean. In contrast with leguminous isoflavonoids, flavones are widely distributed in higher plants and are produced by the CYPs flavanone 2-hydroxylase (F2H) or flavone synthase II (FNS II) (Figure 1), which have been identified separately in multiple plant species (Akashi et al. 1999a, 1999b; Du et al. 2010a, 2010b; Fliegmann et al. 2010; Kitada et al. 2001; Lam et al. 2014; Martens and Forkmann 1999; Nakatsuka et al. 2005; Zhang et al. 2007).

In previous studies of plant CYP expression in *E.*

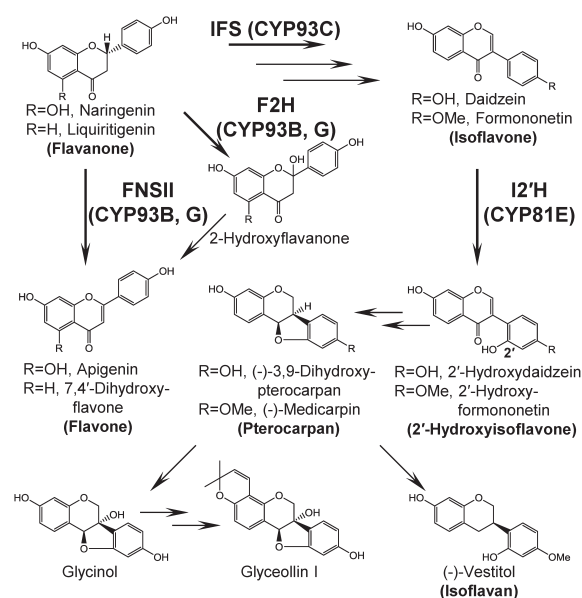


Figure 1. CYPs involved in flavonoid biosynthesis. IFS, 2-hydroxyisoflavanone synthase; F2H, flavanone 2-hydroxylase; FNSII, flavone synthase II; I2'H, isoflavone 2'-hydroxylase.

coli, functions of uncharacterized genes have been demonstrated, CYP proteins have been produced for crystal structure analyses, and metabolites of interest have been prepared using biotransformation methods. In particular, in vitro and in vivo assays with *E. coli* expression systems have been exploited to characterize the roles of CYP74 in allene oxide and aldehyde biosynthesis, CYP79 and CYP83 in glucosinolate biosynthesis, and CYP71 in the hydroxylation of sesquiterpenes (Brash 2009; Harada et al. 2011; Nafisi et al. 2006). *E. coli* expression systems have been used for large-scale preparation and purification of IFS and *p*-coumarate 3-hydroxylase (CYP98A). In these studies, the N-terminal region of plant CYPs was substituted with that of a mammalian CYP to maximize expression efficiency (Kim et al. 2011). In other studies, the fusion protein of IFS and CPR was expressed in *E. coli* to produce isoflavones by biotransformation (Kim et al. 2009; Leonard and Koffas 2007). Flavonol and flavanone have also been produced through biotransformation in *E. coli* expressing flavonoid 3'-hydroxylase (CYP75B) and flavonoid 3',5'-hydroxylase (CYP75A) fused with CPR, and various other flavonoid biosynthetic enzymes (Leonard et al. 2006; Zhu et al. 2014). In many cases, CYP was expressed as a fusion protein with CPR (Hotze et al. 1995; Kim et al. 2009; Leonard et al. 2006; Leonard and Koffas 2007; Zhu et al. 2014) because it has been believed that prokaryotic cells lack membranous scaffolding for the electron transport chain, and thus fusion with CPR has been believed to be essential for functional expression of CYP in *E. coli*.

In the present study, we report simple and efficient bioconversion methods for CYP-expression in *E. coli*

using commercially available vectors, which would expedite functional analyses of uncharacterized CYPs and preparation of enzyme assay substrates which are not available by other means. For convenience of vector construction, CYPs were expressed as distinct polypeptides from CPR with minimum modification, i.e., N-terminal truncation of the transmembrane domain. Accordingly, temperatures and culture supplements were optimized. Among supplements, isopropyl- β -D-thiogalactopyranoside (IPTG) was added to induce heterologous proteins, 5-aminolevulinic acid (ALA) or hemin were added to provide CYP prosthetic groups, ammonium iron (II) sulfate was added to provide iron, and ethanol was used to induce molecular chaperones (Thomas and Baneyx 1996, 1997). The efficiencies of various vector constructs were then compared, and the optimized method was used to analyze the function of soybean I2'H candidates, CYP81E11, CYP81E12, and CYP81E18. High yields of 2'-hydroxydaidzein were then achieved using bioconversion in *E. coli* expressing I2'H, demonstrating efficient synthesis of valuable CYP products.

Materials and methods

Chemicals

Daidzein, formononetin, and (\pm)-naringenin were purchased from LC Laboratories (Woburn, MA, USA), Toronto Research Chemicals (Toronto, Ontario, Canada), and Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan), respectively. (2S)-Liquiritigenin and 2'-hydroxydaidzein were obtained from laboratory stocks and 5-aminolevulinic acid hydrochloride was purchased from Cosmo Bio Co., Ltd. (Tokyo, Japan). Hemin was purchased from Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan). Isopropyl- β -D-thiogalactopyranoside, 99.5% ethanol, and ammonium iron (II) sulfate hexahydrate [$\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_6 \cdot 6\text{H}_2\text{O}$] were purchased from Wako Pure Chemical Industries, Ltd. (Osaka, Japan).

cDNA cloning

Total RNA was isolated using a SV Total RNA Isolation System (Promega Corporation, Madison, WI, USA), and templates for cDNA cloning were synthesized using a SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen, Carlsbad, CA, USA). Sources of mRNA were as follows: a callus culture of *G. max* cv. Tamba Kurodaizu (black soybean), UC-1 cell line, which was induced from a leaf explant, cultured in Murashige Skoog liquid medium supplemented with 0.45 μM 2,4-dichlorophenoxyacetic acid, 4.7 μM kinetin, and 3% sucrose, and treated with 0.3% yeast extract (BD Biosciences, CA, USA) for CYP81E11 (NCBI RefSeq accession number NP_001241120) and CYP81E18 (NCBI RefSeq accession number XP_003533574); seedlings of *L. japonicus* B-129 Gifu treated with 10 mM reduced glutathione (Shimada et al. 2000) for CYP81E6; and 10-day-old seedlings of *Oryza sativa*

cv. Nipponbare for CYP93G2. The cDNA clone GMFL01-20-B13 was provided by the National Bioresource Project (NBRP) and was used as a template for CYP81E12 (NCBI RefSeq accession number NP_001240000). The cDNA clones for CYP81E1, CYP93B1, CYP93B4, CYP93C2, and LjCPR1 were from laboratory stocks (Akashi et al. 1998a, 1998b, 1999a, 1999b; Seki et al. 2008). CYP nomenclature was based on that recommended on Dr. Nelson's web site (<http://drnelson.uthsc.edu/CytochromeP450.html>). Codon-optimized coding sequences of Δ 2-23-CYP81E1 and Δ 2-23-CYP81E6 (DDBJ accession numbers LC043097 and LC043098) were designed using GeneOptimizer[®] (Invitrogen) and the synthesized DNA samples (Invitrogen) were used to construct expression vectors.

Expression vectors for CYPs and LjCPR1

CYPs and LjCPR1 were co-expressed using the dual plasmid method and the tandem construct method. Membrane-binding regions (MBR) of CYPs and LjCPR1 were predicted using WoLFPSORT (<http://wolfsort.org/>) and SOSUI (<http://bp.nuap.nagoya-u.ac.jp/sosui/>), and MBR-truncated coding sequences were PCR amplified using KOD Plus Ver.2 (Toyobo, Tokyo, Japan), PrimeSTAR HS DNA Polymerase (Takara, Shiga, Japan), or PrimeSTAR MAX DNA Polymerase (Takara). Primer sets and DNA polymerases used for vector construction are listed in Table 1.

In the dual plasmid method, PCR products of CYPs and LjCPR1 were introduced into pET46 Ek/LIC (Novagen, Madison, WI, USA) and pCDF2 Ek/LIC vectors (Novagen), respectively (Figure 2A), using a ligation-independent cloning (LIC) method according to the manufacturer's instructions. In the tandem construct method, MBR-truncated LjCPR1 was initially introduced into the pET46 Ek/LIC vector, and CYPs were amplified with a T7 promoter using the CYP-introduced pET46 Ek/LIC vectors described above as templates. The amplified T7 promoter-CYP was integrated into the LjCPR1-containing pET46, which was treated with Xho I, using an In-fusion HD Cloning Kit (Clontech, Mountain View, CA, USA) (Figure 2B).

As an alternative method for tandem construction, a Gateway cassette was amplified using Gateway cassette-Fw and Gateway cassette-Rv primers and the Reading Frame B of a Gateway Vector Conversion System (Invitrogen) as a template; the cassette was then integrated into the pET46 and the cassette was amplified with the T7 promoter. The amplified T7 promoter-Gateway introduced into the LjCPR1-containing pET46 instead of the T7 promoter-CYP (Figure 2C). CYPs without MBR were cloned into the pCR8/GW/TOPO vector (Invitrogen) and were introduced into the Gateway cassette using a Gateway LR Clonase II Enzyme Mix (Invitrogen). The Gateway-based method was used to prepare the constructs of CYP93B1, CYP93B4, CYP93G2, and codon-optimized CYPs. To obtain the codon-optimized coding sequences for Δ 2-23-CYP81E1 and Δ 2-23-CYP81E6, two partial DNA fragments for each clone, which were separately synthesized (Invitrogen), were concatenated using an In-fusion HD Cloning Kit; these

Table 1. PCR Primers and DNA Polymerases for vector construction.

Primer name	Primer sequence (5'→3')	DNA Polymerase for PCR
Δ2-26-CYP81E1-Fw	<u>GACGACGACAAGAT</u> AT GAGAAAATTCAAAAACCTCCCAC	K
CYP81E1-Rv	<u>GAGGAGAAGCCCGGTTA</u> CTGCTTGAAAACCTTGTTGATG	
Δ2-23-CYP81E6-Fw	<u>GACGACGACAAGAT</u> AT GCTCGGCTCAAGAAAGTTCAAAAAAC	PH
CYP81E6-Rv	<u>GAGGAGAAGCCCGGTTA</u> CTTGAAAACATTGTTTATGACTG	
Δ2-24-CYP81E11-Fw	<u>GACGACGACAAGAT</u> AT GTTCCAAAGCAGAAAATTGAGAAAC	K
CYP81E11-Rv	<u>GAGGAGAAGCCCGGTTA</u> AAATTCCAATTTTAGTGGCAAG	
Δ2-26-CYP81E12-Fw	<u>GACGACGACAAGAT</u> AT GTTCCAAAGAAGCAGAAAAAGTAC	K
CYP81E12-Rv	<u>GAGGAGAAGCCCGGTTA</u> ATAACTTCCAACCTTTGCTGGC	
Δ2-24-CYP81E18-Fw	<u>GACGACGACAAGAT</u> AT GTTCCAAAGAAGAAAATTGAGAAAC	PH
CYP81E18-Rv	<u>GAGGAGAAGCCCGGTTA</u> AAATTTTAGTGGCAAGTGGGC	
Δ2-22-CYP93C2-Fw	<u>GACGACGACAAGAT</u> AT GAGTGCAAAATCAAAGTCCCTTC	K
CYP93C2-Rv	<u>GAGGAGAAGCCCGGTTA</u> CGACGAAAAGAGTTTGGGTAC	
Δ1-24-CYP93B1-Fw	TCCACCGTTACGGAAAAAAC	PM
CYP93B1-Rv	<u>TCAACATTCAATGATTTT</u> GAGGG	
Δ1-23-CYP93B4-Fw	CGCCGAAGGGGACCACC	PM
CYP93B4-Rv	<u>TCAAGCACCCGATATT</u> TGTGG	
Δ1-27-CYP93G2-Fw	TCAGTGGAAGCAGAAGCAG	PM
CYP93G2-Rv	<u>TCAGGAGTAGAAGGA</u> AGGGAG	
Δ2-57-LjCPR1-Fw	<u>GACGACGACAAGAT</u> AT GTGCGTCGTCGTTTTCATTTGGC	K
LjCPR1-Rv	<u>GAGGAGAAGCCCGGTTA</u> CACATACATCACGCAAATAC	
T7 promoter-CYP-Fw	<u>CTCTCAAATCTCGATAAT</u> ACGACTCACTATAGGG	PM
T7 promoter-CYP-Rw	<u>CTTACCAGACTCGAGATTT</u> GAG	
Gateway cassette-Fw	<u>GACGACGACAAGAT</u> AT TCAACAAGTTTGTACAAAAAAGC	PM
Gateway cassette-Rv	<u>GAGGAGAAGCCCGGTTA</u> CAACCACTTTGTACAAGAAAGC	

Δ, Region of deleted amino acids. Underlines indicate the sequence for the LIC method; boldface type and double-underlines indicate initiation and stop codons, respectively. Italic sequences are homologous to vectors for In-fusion. K, KOD Plus Ver.2; PH, PrimeSTAR HS DNA Polymerase; PM, PrimeSTAR MAX DNA Polymerase.

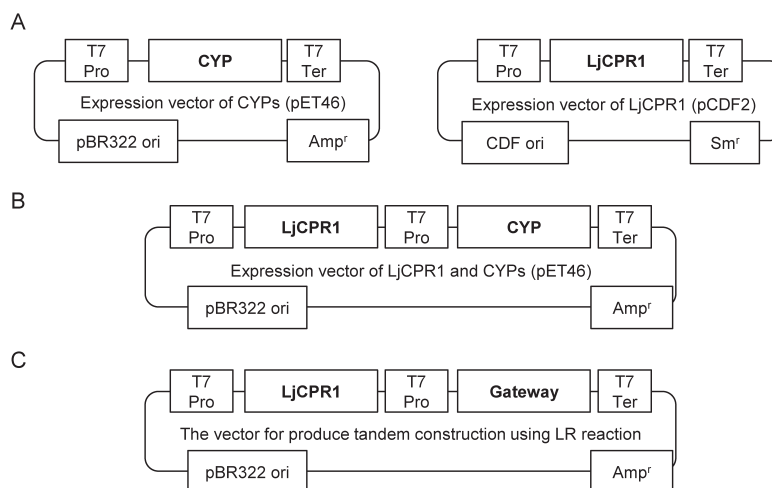


Figure 2. Schematic structure of expression vectors for dual plasmid (A) and tandem construct methods (B, C). The 5' regions of the coding sequences of CYP and LjCPR1 were truncated to remove the N-terminal membrane-binding region. T7 Pro, T7 promoter; T7 Ter, T7 terminator; Amp^r, ampicillin resistance gene; Sm^r, streptomycin resistance gene.

were introduced into the Gateway cassette in the LjCPR1-containing pET46 via pDONR221 (Invitrogen) using successive BP and LR reactions of the Gateway technology according to the manufacturer's instructions.

Fermentation

E. coli strain C41 (DE3) (Lucigen Corporation, Middleton, WI, USA) was transformed with the expression vectors mentioned above. A 10 ml of cold LB medium containing 50 mg l⁻¹

carbenicillin and 50 mg l⁻¹ spectinomycin for dual plasmid method, or 50 mg l⁻¹ carbenicillin for tandem construct method, was inoculated with a 1 ml overnight culture of recombinant *E. coli* in the same medium, and was cultured in a 50 ml CELLSTAR[®] filtertop spheroid cell culture tube (Greiner Bio-One, Tokyo, Japan) to an OD₆₀₀ of 0.4–0.5. IPTG and 2–4 mg of substrate were dissolved in 20 μl of a mixture of dimethyl sulfoxide and polyoxyethylene (80) sorbitan monooleate (1:1, v/v). These were then added to cultures with

other supplements and were incubated for 24 h. The tentative standard supplements used in optimization experiments (Figure 4) included 0.5 mM IPTG, 1 mM ALA, 1 mM thiamine, 50 mM FeCl₃, 1 mM MgCl₂, and 2.5 mM (NH₄)₂SO₄. See the Results for the optimized composition of supplements.

HPLC analysis

A 200 μ l aliquot of culture was mixed with an equal volume of ethyl acetate, was vortexed for a few seconds, and was then centrifuged at 13,300 g for 1 min at 4°C. Prior to IFS and F2H assays, cultures were treated with 1 M HCl in a boiling water bath for 10 min, and then mixed with ethyl acetate. An aliquot of organic phase (150 μ l) was then dried using blowing air and was dissolved in 100 μ l of methanol. Samples were analyzed using HPLC on a TSK-Gel ODS-80TM column (4.6 \times 150 mm; Tosoh, Tokyo, Japan) at a flow rate of 1.0 ml min⁻¹ at 40°C with a linear gradient elution for 20 min from 40 to 70% methanol. Eluates were monitored using a multiwavelength detector (MD-2010, JASCO, Tokyo, Japan). The reaction products were identified using comparisons of retention times and UV spectra with those from the authentic samples.

¹H-NMR analysis of 2'-hydroxydaidzein and 2'-hydroxygenistein

The reaction products of CYP81E18 using daidzein and genistein as substrates were isolated from the ethyl acetate extract of the culture medium by column chromatography on a silica-gel column (26 \times 100 mm) with chloroform:methanol (99:1–80:20, v/v) using a flash chromatograph system (YFLC AI-580, Yamazen, Osaka, Japan). 2'-Hydroxydaidzein was purified by thin-layer chromatography on a silica gel plate (Kieselgel F254, Merck, Kenilworth, NJ, USA) with the solvent toluene:ethyl acetate:methanol:light petroleum (6:4:1:3, v/v) and HPLC on a Shim-pack PREP-ODS column (20 \times 250 mm; Shimadzu, Kyoto, Japan) with 35% (v/v) acetonitrile at a flow rate of 10 ml min⁻¹. 2'-Hydroxygenistein was purified by silica-gel thin-layer chromatography with chloroform:methanol (9:1, v/v). The NMR spectra were recorded on a JMN ECA-500 system (JEOL, Tokyo, Japan). 2'-Hydroxydaidzein, ¹H-NMR (methanol-*d*₄) δ : 6.38 (1H, *dd*, *J*=2.5, 8.5 Hz, H-5'), 6.39 (1H, *d*, *J*=2.5 Hz, H-3'), 6.83 (1H, *d*, *J*=2.5 Hz, H-8), 6.93 (1H, *dd*, *J*=2.5, 9.0 Hz, H-6), 7.04 (1H, *d*, *J*=8.5 Hz, H-6'), 8.04 (1H, *d*, *J*=9.0 Hz, H-5), 8.11 (1H, *s*, H-2). 2'-Hydroxygenistein, ¹H-NMR (methanol-*d*₄) δ : 6.22 (1H, *d*, *J*=2.5 Hz, H-6), 6.35 (1H, *d*, *J*=2.5 Hz, H-8), 6.37 (1H, *dd*, *J*=2.5, 8.5 Hz, H-5'), 6.39 (1H, *d*, *J*=2.5 Hz, H-3'), 7.04 (1H, *d*, *J*=8.5 Hz, H-6'), 8.00 (1H, *s*, H-2).

Results

The optimum culture conditions

The basic principle of this method depends on expression of CYP as a polypeptide that is distinct from CPR in *E. coli* with minimum modification using commercially available expression vectors. To test this principle,

bioconversion was performed in recombinant *E. coli*. CYP81E12 was arbitrarily selected among the candidates for soybean I2'H. Coding sequences for transmembrane-domain-truncated CYP81E12 and LjCPR1 were integrated into pET46 and pCDF2 vectors, respectively, for use in the dual plasmid method (Figure 2A). Subsequently, recombinant *E. coli* were cultured in LB medium supplemented with IPTG, daidzein (substrate), ALA, ethanol, and ammonium iron (II) sulfate. After fermentation for 24 h, the reaction product was extracted with ethyl acetate and was analyzed using HPLC/PDA. CYP81E12 showed I2'H activity and converted daidzein into 2'-hydroxydaidzein (Figure 3).

Since the functional expression of CYP81E12 and LjCPR1 proved possible using the dual plasmid method, we further improved the conversion efficiency by optimizing temperature and culture supplements in the culture medium. Subsequently, 2'-hydroxydaidzein production was determined after 24 h fermentation and was expressed relative to the highest production for each condition. In these experiments, the optimal temperature was 25°C and production was reduced to 70% at 20°C, 22% at 30°C, and was undetectable at 37°C (Figure 4A). Addition of the heme precursor ALA was essential for functional expression of CYP81E12, and yields increased with increasing ALA concentrations up to 1.0 mM; however, further improvement of efficiency was not shown by increasing concentrations of ALA (Figure 4B). IPTG induces heterologous proteins via the lactose operon. Thus, we compared the two widely used concentrations of 0.5 mM and 1 mM, but no significant difference was observed (Figure 4C). Ethanol was added as an inducer of molecular chaperones, and it increased

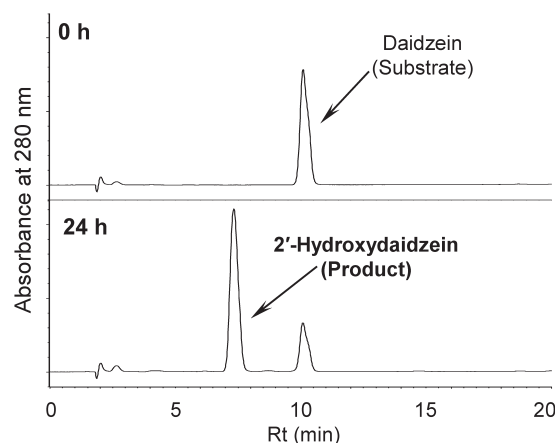


Figure 3. Bioconversion of daidzein into 2'-hydroxydaidzein by *E. coli* expressing CYP81E12. CYP81E12 and LjCPR1 were co-expressed using the dual plasmid method. Recombinant *E. coli* were fermented in culture medium supplemented with daidzein, and the culture was extracted with ethyl acetate and analyzed using HPLC after 24 h culture. A representative result under optimized fermentation conditions (see the text).

2'-hydroxydaidzein production by about two-fold at 1% and decreased production at higher concentrations (Figure 4D). Addition of hemin instead of ALA yielded lower production, and the use of nutrition-rich TB medium instead of LB medium decreased the production (data not shown). Ammonium iron (II) sulfate was used as an iron source for heme production in previous studies (Harada et al. 2011); therefore, we determined the effects of ammonium iron (II) sulfate in preliminary experiment. The supplement of 100 μ M ammonium iron (II) sulfate had a tendency to increase the productivity, although no statistically significant difference was observed (data not shown). According to these observations, standard conditions were established in LB medium supplemented with 0.5 mM IPTG, 1 mM ALA, 1% ethanol, and 100 μ M ammonium iron (II) sulfate at 25°C.

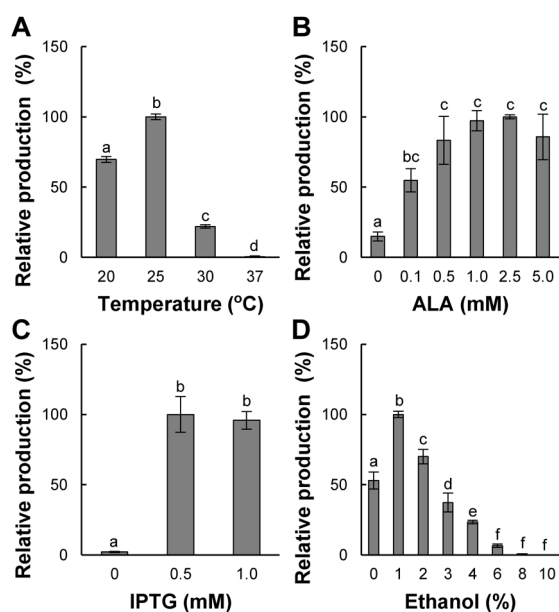


Figure 4. Effects of temperature (A), 5-aminolevulinic acid (ALA) (B), isopropyl- β -D-thiogalactopyranoside (IPTG) (C), and ethanol (D) on the efficiency of bioconversion by recombinant *E. coli* expressing CYP81E12. Data are presented as proportions of the highest production for each condition. Error bars indicate standard deviations (SD) ($n=3$). Different letters indicate significant differences according to Tukey-Kramer test ($p<0.05$).

Comparison of dual plasmid and tandem construct methods

In addition to the dual plasmid method, we tested a vector construction designated as the tandem construct method (Figure 2B, C). It enables the production of both polycistronic CPR-CYP and monocistronic CYP transcripts. Furthermore, CPR expression in the pET46 vector, whose copy number (more than 40) is larger than that of pCDF2 (20–40) according to the manufacturer's instructions, may result in its increased production compared to the dual plasmid method. Growth of *E. coli* and production of 2'-hydroxydaidzein was compared between the two methods for up to 48 h after induction with IPTG (Table 2). Although no differences in growth rates and final cell densities at the stationary phase were observed, the tandem construct method produced yields that were about 1.4-fold higher than those of the dual plasmid method at 48 h after induction. Moreover, production continued to increase during the stationary phase (more than 10 h after induction) and produced 161 mg l⁻¹ of culture medium at 48 h after induction (Table 2).

Bioconversion using *E. coli* expressing enzymes of CYP81 and CYP93 families

According to public databases, soybean expresses at least two additional members of the CYP81E subfamily (CYP81E11 and CYP81E18). We tested the activity of the CYP81E members of soybean, together with CYP81E1 from *G. echinata* and CYP81E6 from *L. japonicus*, using the present optimized *E. coli* expression method (Table 3). All soybean I2'Hs were active toward the 4'-hydroxy-type substrates daidzein and genistein but did not metabolize the 4'-methoxy-type substrate formononetin. 2'-Hydroxydaidzein and 2'-hydroxygenistein were identified by ¹H-NMR, as well as comparison of retention times and UV spectra with the authentic samples. On the other hand, CYP81E6 showed 4'-methoxy-type I2'H activity but CYP81E1 did not. No activity of CYP81E1 was observed even when the N-terminal truncation site was changed or the N-terminal region was substituted with that of CYP81E12. CYP81E1 and CYP81E6 were also expressed using the codon-optimized construct for

Table 2. Comparison of dual plasmid and tandem construct methods for co-expression of CYP81E12 and LjCPR1.

Fermentation time after IPTG induction (h)	2'-Hydroxydaidzein (mg l ⁻¹)		Cell density (OD ₆₀₀)	
	Tandem construct*	Dual plasmid*	Tandem construct	Dual plasmid
0	ND	ND	0.43±0.003	0.48±0.003
5	1.4±0.04	0.6±0.2	1.69±0.08	1.68±0.03
10	26±1.0	13±0.5	2.83±0.08	2.82±0.04
24	77±3.0	58±2.1	2.96±0.04	2.94±0.04
48	161±5.6	113±7.8	2.94±0.02	2.92±0.03

The production of 2'-hydroxydaidzein from daidzein and the growth of recombinant *E. coli* are shown. Each value is the mean±SD ($n=3$). *Productivity differed significantly between the two methods (two-way ANOVA, $p<10^{-10}$). ND, not detected.

E. coli, and the biotransformation efficacies were tested. The activity of CYP81E6 increased by about two fold, whereas the optimized CYP81E1 did not yield detectable amounts of product (Table 3). However, other flavonoid biosynthetic CYPs, including IFS (CYP93C2), F2H (CYP93B1 and 93G2), and FNSII (CYP93B4), showed functional expression in *E. coli* (Table 3).

Discussion

In this study, we established a series of efficient functional expression methods for plant CYPs in *E. coli* and performed functional analyses of previously uncharacterized CYPs of soybean. For simplicity and suitability of the present method, we avoided modifications of *E. coli* (Leonard et al. 2006; Zhu et al. 2014) and construction of CPR–CYP fusion proteins for purpose of its application to the preparation of various substrates for biochemical analysis of biosynthetic enzymes and to the production of various natural products of importance. Since optimal expression conditions generally vary depending on expression vectors and types of recombinant DNA expressed, fermentation temperatures and culture supplements were optimized for higher efficiency of bioconversion by soybean CYP81E12. The resulting optimal temperatures and concentrations of IPTG and ALA were similar to those in previous reports of *E. coli* expression systems for CYP (Figure 4A, B, C). The fermentation temperature should influence several aspects that affect the final productivity: total amount of translated enzyme, the proportion of active-form enzyme in the bacterial cells,

the stability of enzyme during the fermentation, and the reaction rate. The optimum productivity observed at 25°C can be due to the best balance among them. The addition of ethanol to the culture medium was tested for molecular chaperone induction and recombinant protein stabilization (Thomas and Baneyx 1996, 1997). In contrast with a previous study that showed an optimal ethanol concentration of 3% (Sahu et al. 2009), yields from the present system were increased 2-fold after addition of 1% ethanol (Figure 4D). To the best of our knowledge, this is the first report to demonstrate the effects of ethanol supplementation on CYP expression in *E. coli*. Furthermore, comparisons of dual plasmid and tandem constructs under optimal conditions (Figure 2) showed about 1.5-fold higher production with the tandem construct than with the dual plasmid construct. This observation may reflect efficient translation of CYP and CPR by the polycistronic transcript derived from the tandem construct or increased CPR copy number, and the highest production of 2'-hydroxydaidzein was 161 mg l⁻¹ of culture medium (Table 2). Previous reports show production of flavonoids such as daidzein, genistein, eriodictyol, and quercetin by bioconversion using *E. coli* expressing CYP of 1.1–107 mg l⁻¹ (Kim et al. 2009; Leonard et al. 2006; Leonard and Koffas 2007; Zhu et al. 2014). Hence, the present in vivo CYP expression system using *E. coli* offers a highly efficient method for production of substrates for enzyme assays and for synthesis of natural products of commercial value.

CPR is an essential redox partner for functional expression of CYP, and previous studies of CYP expression have employed endogenous CPR in

Table 3. Bioconversion by *E. coli* expressing selected CYPs that are involved in (iso)flavonoid biosynthesis.

Activity	Substrate (Product)	CYP	Species	Vector construction	Product*(mg l ⁻¹)	Functional identification
12'H	Daidzein (2'-Hydroxydaidzein)	CYP81E11	<i>G. max</i>	DP	47±5.4	This study
		CYP81E12	<i>G. max</i>	DP	58±2.1	This study
		CYP81E18	<i>G. max</i>	DP	60±11	This study
	Genistein (2'-Hydroxygenistein)	CYP81E11	<i>G. max</i>	DP	Trace	This study
		CYP81E12	<i>G. max</i>	DP	Trace	This study
		CYP81E18	<i>G. max</i>	DP	Trace	This study
Formononetin (2'-Hydroxyformononetin)		CYP81E11	<i>G. max</i>	DP	ND	—
		CYP81E12	<i>G. max</i>	DP	ND	—
		CYP81E18	<i>G. max</i>	DP	ND	—
		CYP81E1	<i>G. echinata</i>	TC	ND	Akashi et al. 1998b
		CYP81E1 optimized	<i>G. echinata</i>	TC	ND	This study
		CYP81E6	<i>L. japonicus</i>	TC	4.4±0.2	Shimada et al. 2000
		CYP81E6 optimized	<i>L. japonicus</i>	TC	8.4±0.05	This study
IFS	(±)-Naringenin (Genistein**) (2S)-Liquiritigenin (Daidzein**)	CYP93C2	<i>G. echinata</i>	TC	0.6±0.04	Akashi et al. 1999a
				TC	27±2	Akashi et al. 1999a
F2H	(±)-Naringenin (Apigenin**)	CYP93B1	<i>G. echinata</i>	TC	25±1.6	Akashi et al. 1998a
		CYP93G2	<i>O. sativa</i>	TC	3.7±0.2	Du et al. 2010b
FNS II	(±)-Naringenin (Apigenin)	CYP93B4	<i>T. hybrida</i>	TC	66±7.8	Akashi et al. 1999b

* Each value is the mean±SD (n=3). ** Converted from the direct product by treatment with HCl. Trace, Detectable (more than 300 μg l⁻¹) but not determined; ND, not detected (less than 300 μg l⁻¹); DP, Dual plasmid method; TC, Tandem construct method.

eukaryotic hosts such as yeast (Katsuyama et al. 2007), or have co-expressed heterologous (Harada et al. 2011; Kim et al. 2009; Leonard and Koffas 2007; Zhu et al. 2014) or homologous CPR (Harada et al. 2011; Hotze et al. 1995; Leonard et al. 2006). Moreover, genetic distances between the sources of the CPR and CYP reportedly influenced the enzyme activity (Kim et al. 2009). However, the LjCPR1 used in this study was selected from three paralogous CPRs of the leguminous plant *L. japonicus* for its ability to confer the highest IFS activities of *G. echinata* and was previously used as a redox partner for functional identification of licorice CYPs (Seki et al. 2008, 2011). Moreover, LjCPR1 was effective for legume CYPs (CYP81E6, CYP81E11, CYP81E12, CYP81E18, and CYP93C2), rice F2H (CYP93G2), and for *Torenia FNSII* (CYP93B4). Taken together, these studies of LjCPR1 indicate that heterologous plant CPR may be of versatile use for functional expression of plant CYPs.

Soybean CYP81E11, CYP81E12, and CYP81E18 show more than 80% amino acid sequence identity but only 60% identity to I2'H from *G. echinata* and *L. japonicus* (CYP81E1 and CYP81E6, respectively); therefore they have not been simply predicted to be orthologous to CYP81E1 and CYP81E6. The present characterization using the *E. coli* expression system demonstrated I2'H activities of these CYPs for the first time. Moreover, the present data show that these CYPs are active towards the 4'-hydroxy type substrates daidzein and genistein but not towards the 4'-methoxy substrate formononetin (Table 3). This substrate preference may reflect the production of 4'-hydroxy type phytoalexins by soybean and is in contrast with that of *G. echinata* I2'H (CYP81E1), which was shown to have selectivity for 4'-methoxy substrates over 4'-hydroxy substrates in an in vitro assay using a yeast expression system (Akashi et al. 1998b). The substrate specificity of I2'H so far characterized is in good agreement with the type of the end products, phytoalexin.

The activities of some CYPs tested in this study had also been investigated using a yeast expression system (see the references in Table 3), and thus the substrate preference and relative activity among the orthologous CYPs can be compared between the two expression systems. The IFS (CYP93C2) of *G. echinata*, which predominantly produces 5-deoxy isoflavonoids from liquiritigenin, showed preference for liquiritigenin over the 5-hydroxy type substrate naringenin (Table 3). Moreover, the activity of the CYP93B1 is quite higher than CYP93G2 in the present system (Table 3) and in a yeast expression system (unpublished data). Hence, these data indicate little differences in substrate specificity or relative CYP activities between heterologous expression systems in *E. coli* and yeast. However, CYP81E1 gave no detectable product in the *E. coli* system, even when the N-terminal truncation site was changed, the

N-terminal region was replaced with that of CYP81E12, or the codon usage was optimized. The reason for the unsuccessful bioconversion by CYP81E1, which should be its failure of expression or failure of enzymatic activity, remained unclear because the expression level of CYPs in *E. coli* was not enough to detect, even in the case that the CYP yielded a high amount of product. In contrast, CYP81E6 of *L. japonicus*, which has 83% amino acid sequence identity to CYP81E1, successfully converted formononetin with improved efficiency after optimization of codon usage (Table 3). Although this discrepancy remains unresolved, these data indicate that failure of the *E. coli* expression system can be accommodated by use of another orthologous CYP.

In conclusion, the present study demonstrates functional characterization of a CYP following biotransformation into an *E. coli* expression system using N-terminal-MBR truncated CYP and CPR expression vectors. Under optimized conditions, this system produced 161 mg l⁻¹ of product, indicating high efficiency for use in commercial and other applications that require products of CYP enzymes. Although some CYP enzymes may not be viable using this system, the present method may accelerate functional identification of uncharacterized CYPs and other biosynthetic enzymes, and it promises application to large-scale production of important natural products.

Acknowledgements

We wish to thank Mao Aketagawa (Nihon University) for technical assistance. This work was supported by Grant-in-Aid for Scientific Research on Innovation Areas (Biosynthetic Machinery), No. 25108723, by the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

- Akashi T, Aoki T, Ayabe S (1998a) Identification of a cytochrome P450 cDNA encoding (2S)-flavanone 2-hydroxylase of licorice (*Glycyrrhiza echinata* L.; Fabaceae) which represents licodione synthase and flavone synthase II. *FEBS Lett* 431: 287–290
- Akashi T, Aoki T, Ayabe S (1998b) CYP81E1, a cytochrome P450 cDNA of licorice (*Glycyrrhiza echinata* L.), encodes isoflavone 2'-hydroxylase. *Biochem Biophys Res Commun* 251: 67–70
- Akashi T, Aoki T, Ayabe S (1999a) Cloning and functional expression of a cytochrome P450 cDNA encoding 2-hydroxyisoflavanone synthase involved in biosynthesis of the isoflavonoid skeleton in licorice. *Plant Physiol* 121: 821–828
- Akashi T, Fukuchi-Mizutani M, Aoki T, Ueyama Y, Yonekura-Sakakibara K, Tanaka Y, Kusumi T, Ayabe S (1999b) Molecular cloning and biochemical characterization of a novel cytochrome 450, flavone synthase II, that catalyzes direct conversion of flavanones to flavones. *Plant Cell Physiol* 40: 1182–1186
- Akashi T, Koshimizu S, Aoki T, Ayabe S (2006) Identification of cDNAs encoding pterocarpan reductase involved in isoflavan phytoalexin biosynthesis in *Lotus japonicus* by EST mining. *FEBS Lett* 580: 5666–5670
- Boué SM, Tilghman SL, Elliott S, Zimmerman MC, Williams

- KY, Payton-Stewart F, Mirafior AP, Howell MH, Shih BY, Carter-Wientjes CH, et al. (2009) Identification of the potent phytoestrogen glycinol in elicited soybean (*Glycine max*). *Endocrinology* 150: 2446–2453
- Brash AR (2009) Mechanistic aspects of CYP74 allene oxide synthases and related cytochrome P450 enzymes. *Phytochemistry* 70: 1522–1531
- Dewick PM (1993) Isoflavonoids. In: Harborne JB (ed) *The Flavonoids. Advances in Research since 1986*. Chapman and Hall, London, pp 117–238
- Du Y, Chu H, Wang M, Chu IK, Lo C (2010a) Identification of flavone phytoalexins and a pathogen-inducible flavone synthase II gene (SbFNSII) in sorghum. *J Exp Bot* 61: 983–994
- Du Y, Chu H, Chu IK, Lo C (2010b) CYP93G2 is a flavanone 2-hydroxylase required for C-glycosylflavone biosynthesis in rice. *Plant Physiol* 154: 324–333
- Fliegmann J, Furtwängler K, Malterer G, Cantarello C, Schüler G, Ebel J, Mithöfer A (2010) Flavone synthase II (CYP93B16) from soybean (*Glycine max* L.). *Phytochemistry* 71: 508–514
- Harada H, Shindo K, Iki K, Teraoka A, Okamoto S, Yu F, Hattan J, Utsumi R, Misawa N (2011) Efficient functional analysis system for cyanobacterial or plant cytochromes P450 involved in sesquiterpene biosynthesis. *Appl Microbiol Biotechnol* 90: 467–476
- Hotze M, Schröder G, Schröder J (1995) Cinnamate 4-hydroxylase from *Catharanthus roseus*, and a strategy for the functional expression of plant cytochrome P450 proteins as translational fusions with P450 reductase in *Escherichia coli*. *FEBS Lett* 374: 345–350
- Katsuyama Y, Miyahisa I, Funa N, Horinouchi S (2007) One-pot synthesis of genistein from tyrosine by coincubation of genetically engineered *Escherichia coli* and *Saccharomyces cerevisiae* cells. *Appl Microbiol Biotechnol* 73: 1143–1149
- Kim DH, Kim BG, Jung NR, Ahn JH (2009) Production of genistein from naringenin using *Escherichia coli* containing isoflavone synthase–cytochrome P450 reductase fusion protein. *J Microbiol Biotechnol* 19: 1612–1616
- Kim YH, Kwon T, Yang HJ, Kim W, Youn H, Lee JY, Youn B (2011) Gene engineering, purification, crystallization and preliminary X-ray diffraction of cytochrome P450 *p*-coumarate-3-hydroxylase (C3H), the *Arabidopsis* membrane protein. *Protein Expr Purif* 79: 149–155
- Kitada C, Gong Z, Tanaka Y, Yamazaki M, Saito K (2001) Differential expression of two cytochrome P450s involved in the biosynthesis of flavones and anthocyanins in chemo-varietal forms of *Perilla frutescens*. *Plant Cell Physiol* 42: 1338–1344
- Lam PY, Zhu FY, Chan WL, Liu H, Lo C (2014) Cytochrome P450 93G1 is a flavone synthase II that channels flavanones to the biosynthesis of triclin *O*-linked conjugates in rice. *Plant Physiol* 165: 1315–1327
- Leonard E, Yan Y, Koffas MA (2006) Functional expression of a P450 flavonoid hydroxylase for the biosynthesis of plant-specific hydroxylated flavonols in *Escherichia coli*. *Metab Eng* 8: 172–181
- Leonard E, Koffas MA (2007) Engineering of artificial plant cytochrome P450 enzymes for synthesis of isoflavones by *Escherichia coli*. *Appl Environ Microbiol* 73: 7246–7251
- Liu CJ, Huhman D, Sumner LW, Dixon RA (2003) Regiospecific hydroxylation of isoflavones by cytochrome p450 81E enzymes from *Medicago truncatula*. *Plant J* 36: 471–484
- Martens S, Forkmann G (1999) Cloning and expression of flavone synthase II from *Gerbera hybrids*. *Plant J* 20: 611–618
- Nafisi M, Sønderby IE, Hansen BG, Geu-Flores F, Nour-Eldin HH, Nørholm MHH, Jensen NB, Li J, Halkier BA (2006) Cytochromes P450 in the biosynthesis of glucosinolates and indole alkaloids. *Phytochem Rev* 5: 331–346
- Nakatsuka T, Nishihara M, Mishiba K, Yamamura S (2005) Temporal expression of flavonoid biosynthesis-related genes regulates flower pigmentation in gentian plants. *Plant Sci* 168: 1309–1318
- Sahu SK, Rajasekharan A, Gummadi SN (2009) GroES and GroEL are essential chaperones for refolding of recombinant human phospholipid scramblase 1 in *E. coli*. *Biotechnol Lett* 31: 1745–1752
- Salvo VA, Boué SM, Fonseca JP, Elliott S, Corbitt C, Collins-Burrow BM, Curiel TJ, Srivastav SK, Shih BY, Carter-Wientjes C, et al. (2006) Antiestrogenic glyceollins suppress human breast and ovarian carcinoma tumorigenesis. *Clin Cancer Res* 12: 7159–7164
- Seki H, Ohyama K, Sawai S, Mizutani M, Ohnishi T, Sudo H, Akashi T, Aoki T, Saito K, Muranaka T (2008) Licorice β -amyryn 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proc Natl Acad Sci USA* 105: 14204–14209
- Seki H, Sawai S, Ohyama K, Mizutani M, Ohnishi T, Sudo H, Fukushima EO, Akashi T, Aoki T, Saito K, et al. (2011) Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. *Plant Cell* 23: 4412–4423
- Shimada N, Akashi T, Aoki T, Ayabe S (2000) Induction of isoflavonoid pathway in the model legume *Lotus japonicus*: molecular characterization of enzymes involved in phytoalexin biosynthesis. *Plant Sci* 160: 37–47
- Steele CL, Gijzen M, Qutob D, Dixon RA (1999) Molecular characterization of the enzyme catalyzing the aryl migration reaction of isoflavonoid biosynthesis in soybean. *Arch Biochem Biophys* 367: 146–150
- Thomas JG, Baneyx F (1996) Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing heat-shock proteins. *J Biol Chem* 271: 11141–11147
- Thomas JG, Baneyx F (1997) Divergent effects of chaperone overexpression and ethanol supplementation on inclusion body formation in recombinant *Escherichia coli*. *Protein Expr Purif* 11: 289–296
- Weinstein LI, Hahn MG, Albersheim P (1981) Host-pathogen interactions: XVIII. Isolation and biological activity of glycinol, a pterocarpan phytoalexin synthesized by soybeans. *Plant Physiol* 68: 358–363
- Zhang J, Subramanian S, Zhang Y, Yu O (2007) Flavone synthases from *Medicago truncatula* are flavanone-2-hydroxylases and are important for nodulation. *Plant Physiol* 144: 741–751
- Zhu S, Wu J, Du G, Zhou J, Chen J (2014) Efficient synthesis of eriodictyol from L-tyrosine in *Escherichia coli*. *Appl Environ Microbiol* 80: 3072–3080