Licorice β-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin

Hikaru Seki*†, Kiyoshi Ohyama†, Satoru Sawai‡§, Masaharu Mizutani¶, Toshiyuki Ohnishi¶‖, Hiroshi Sudo‡§, Tomoyoshi Akashi**, Toshio Aoki**, Kazuki Saito†§, and Toshiya Muranaka*†,††

*Kihara Institute for Biological Research, Yokohama City University, Maioka-cho 641–12, Totsuka-ku, Yokohama, Kanagawa 244-0813, Japan; †RIKEN Plant Science Center, Suehiro-cho 1–7-22, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan; ‡Tokiwa Phytochemical Co., Ltd. Kinoko 158, Sakura, Chiba 285-0801, Japan; §Graduate School of Pharmaceutical Sciences, Chiba University, Yayoi-cho 1–33, Inage-ku, Chiba 263-8522, Japan; ¶Institute for Chemical Research, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan; and **Department of Applied Biological Sciences, Nihon University, Fujisawa, Kanagawa 252-8510, Japan

Edited by Joe Chappell, University of Kentucky, Lexington, KY, and accepted by the Editorial Board July 14, 2008 (received for review April 24, 2008)

Glycyrrhizin, a major bioactive compound derived from the underground parts of *Glycyrrhiza* (licorice) plants, is a triterpene saponin that possesses a wide range of pharmacological properties and is used worldwide as a natural sweetener. Because of its economic value, the biosynthesis of glycyrrhizin has received considerable attention. Glycyrrhizin is most likely derived from the triterpene β-amyrin, an initial product of the cyclization of 2,3-oxidosqualene. The subsequent steps in glycyrrhizin biosynthesis are believed to involve a series of oxidative reactions at the C-11 and C-30 positions, followed by glycosyl transfers to the C-3 hydroxyl group; however, no genes encoding relevant oxidases or glycosyltransferases have been identified. Here we report the successful identification of *CYP88D6*, a cytochrome P450 monooxygenase (P450) gene, as a glycyrrhizin-biosynthetic gene, by transcript profiling-based selection from a collection of licorice expressed sequence tags (ESTs). CYP88D6 was characterized by *in vitro* enzymatic activity assays and shown to catalyze the sequential two-step oxidation of β-amyrin at C-11 to produce 11-oxo-β-amyrin, a possible biosynthetic intermediate between β-amyrin and glycyrrhizin. CYP88D6 coexpressed with β-amyrin synthase in yeast also catalyzed *in vivo* oxidation of β-amyrin to 11-oxo-β-amyrin. *CYP88D6* expression was detected in the roots and stolons by RT-PCR; however, no amplification was observed in the leaves or stems, which is consistent with the accumulation pattern of glycyrrhizin *in planta*. These results suggest a role for CYP88D6 as a β-amyrin 11-oxidase in the glycyrrhizin pathway.

expressed sequence tag | medicinal plant | secondary metabolite | monooxygenase | isoprenoid

Triterpene saponins belong to a class of natural plant products that includes various bioactive compounds found in medicinal plants (1). The roots and stolons of *Glycyrrhiza* plants (*G. uralensis* Fisch. and *G. glabra* L., Fabaceae), which are among the most important crude drugs in the world (2), and contain a large amount (2–8% of the dry weight) of glycyrrhizin, an oleanane-type triterpene saponin.

Because of its sweet taste, glycyrrhizin is used worldwide as a natural sweetener and flavoring additive. Moreover, various pharmacological activities of glycyrrhizin, including anti-inflammatory (3), immunomodulatory (4), antiulcer (5), and antiallergy activities (6), have been reported. Glycyrrhizin also has antiviral activity against various DNA and RNA viruses (reviewed in ref. 7), including HIV (8, 9) and severe acute respiratory syndrome (SARS)-associated *coronavirus* (10). Glycyrrhizin has been used in Japan for more than 20 years as a hepatoprotective agent for chronic hepatitis (11, 12). Therefore, a large amount of licorice and its extracts are on the world market as sweetening agents and medicinal materials; in fact, the annual value of global trade in licorice root is estimated at over US $42.1 million (13).

Although the various pharmacological properties of glycyrrhizin have been extensively studied, its biosynthesis remains poorly understood. The biosynthesis of glycyrrhizin (Fig. 1) involves the initial cyclization of 2,3-oxidosqualene, a common precursor of both triterpenes and phytosterols (14), to the triterpene β-amyrin, followed by a series of oxidative reactions at positions C-11 (two-step oxidation) and C-30 (three-step oxidation) and glucuronylation of the C-3 hydroxyl group, although the precise order of intermediate production is unknown.

Two genes that encode enzymes involved in the early stages of triterpene skeleton formation, namely *squalene synthase* (*SQS*) and *β-amyrin synthase* (*bAS*), have been functionally isolated from *G. glabra* (15, 16); however, no progress has been made toward the identification of genes involved in the multiple oxidations and glycosylations leading from β-amyrin to glycyrrhizin.

Cytochrome P450s play critical roles in oxidative reactions during the biosynthesis of diverse natural plant products, including terpenoids. Recent gene discovery efforts using functional genomic-based approaches identified several P450s responsible for the production of terpenoids with particular importance for human health, such as Taxol, a highly effective anticancer drug derived from *Taxus* species (17, 18), and artemisinin, an antimalarial sesquiterpene lactone from *Artemisia annua* (19, 20).

As a resource for gene discovery in glycyrrhizin biosynthesis, we have generated an EST library from the stolons of *G. uralensis* plants, comprising approximately 56,000 cDNAs, which were assembled into 10,372 unique sequences and annotated (H. Sudo, H. Seki, N. Sakurai, H. Suzuki, D. Shibata, A. Toyoda, Y. Totoki, Y. Sakaki, O. Iida, M. Kojoma, T. Shibata, T.M., and K.S., unpublished work). In this study, we used our collection of licorice ESTs to search for glycyrrhizin-biosynthetic genes. Mining of our EST dataset for putative P450 genes and subsequent

Author contributions: K.S. and T.M. designed research; H. Seki, K.O., S.S., M.M., T.O., and H. Sudo performed research; T. Akashi and T. Aoki contributed new reagents/analytic tools; H. Seki, K.O., and S.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. J.C. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database [AB433175 (CYP88D2), AB433176 (CYP88D3), AB433177 (CYP88D4), AB433178 (CYP88D5), AB433179 (CYP88D6), AB437320 (CYP93E3), and AB433810 (LjCPR1)].

††To whom correspondence should be addressed. E-mail: muranaka@yokohama-cu.ac.jp.

‖Present address: Michael Smith Laboratories, University of British Columbia, 2185 East Mall, Vancouver, BC, Canada V6T 1Z4.

This article contains supporting information online at www.pnas.org/cgi/content/full/0803876105/DCSupplemental.

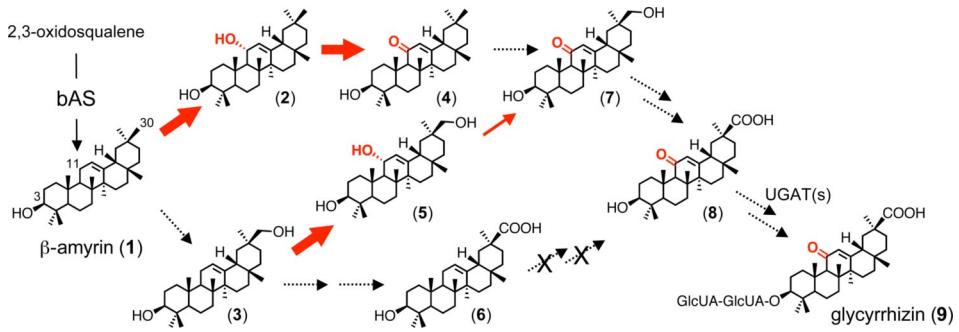

Fig. 1. Proposed pathway for the biosynthesis of glycyrrhizin. A cartoon depiction of the proposed pathway for the biosynthesis of glycyrrhizin. The structures of the possible biosynthetic intermediates between β-amyrin (1) and glycyrrhizin (9) are shown: (2) 11α-hydroxy-β-amyrin; (3) 30-hydroxy-β-amyrin; (4) 11-oxo-β-amyrin; (5) 11α,30-dihydroxy-β-amyrin; (6) 11-deoxoglycyrrhetinic acid; (7) 30-hydroxy-11-oxo-β-amyrin; and (8) glycyrrhetinic acid. In addition to β-amyrin (1) and glycyrrhizin (9), the occurrence of compounds (4), (6), (7), and (8) has been confirmed *in planta* by GC-MS (K.O., K.S., and T.M., unpublished results). In addition to β-amyrin (1), compounds (3) and (6) were tested as substrates in *in vitro* enzyme assays (Figs. 2 and 3). Arrows indicate a single oxidation reaction, glycosylation step, or the initial cyclization step catalyzed by β-amyrin synthase, *bAS*. Red arrows indicate reactions catalyzed by the CYP88D6 enzyme as described herein, broken arrows signify undefined steps, and Xs through arrows indicate that the CYP88D6 enzyme was not able to use 11-deoxoglycyrrhetinic acid (6) as a substrate in *in vitro* assays. UDP-glucuronosyltransferases (UGATs).

transcript profiling-based selection of the candidate P450s led to the identification of *CYP88D6*. CYP88D6 was shown *in vitro* and *in vivo* (in *bAS*-expressing yeast) to catalyze two sequential oxidation steps in the glycyrrhizin pathway: the conversion of β-amyrin (compound **1** in Fig. 1) to 11-oxo-β-amyrin (compound **4** in Fig. 1) via 11α-hydroxy-β-amyrin (compound **2** in Fig. 1). CYP88D6 represents a unique subfamily of CYP88 enzymes that appears to be restricted to the Fabaceae. The results obtained in this study not only provide insight into the production of a wide variety of triterpene saponins within the Fabaceae, they also reveal the potential of using yeast cells in the production of valuable triterpenoids.

Results

Selection of Candidate P450s. Keyword searches of our annotated ESTs identified 37 unique P450 contigs (H. Sudo, H. Seki, N. Sakurai, H. Suzuki, D. Shibata, A. Toyoda, Y. Totoki, Y. Sakaki, O. Iida, M. Kojoma, T. Shibata, T.M., and K.S., unpublished work). To narrow down the candidate P450s, a transcript profiling approach was taken. The expression of each P450 gene was monitored both in the glycyrrhizin-accumulating underground organs (stolons and roots) and the aboveground organs (leaves and stems) where no glycyrrhizin can be detected. Five candidate P450s representing four distinct families (CYP72, CYP83, CYP88, and CYP714) were initially selected because their expression profiles were consistent with the organ-specific accumulation pattern of glycyrrhizin.

***In Vitro* CYP88D6 Enzymatic Activity Assay.** We first examined the potential β-amyrin oxidase activity of all five candidate P450s using *in vitro* assays. For this purpose, microsomes from *Spodoptera frugiperda* 9 (*Sf*9) insect cells expressing each of the candidate P450s were assayed *in vitro* with β-amyrin (**1**) followed by gas chromatography–mass spectrometry (GC-MS). Of the five candidates, only CYP88D6 (according to the recommendation of the Cytochrome P450 Nomenclature Committee) showed clear β-amyrin-oxidizing activity.

Fig. 2*A* shows the CYP88D6-dependent formation of two GC-MS-detectable compounds. The retention time (Rt) and mass spectrum of each major (peak 2: Rt = 18.8 min; Fig. 2*D*) and minor product (peak 1: Rt = 15.7 min; Fig. 2*C*) showed excellent matches to those of authentic 11-oxo-β-amyrin (**4**) and 11α-hydroxy-β-amyrin (**2**) [supporting information (SI) Fig. S1], respectively. In contrast, the peaks were not detected in assays using microsomes from empty vector control *Sf*9 cells (Fig. 2*B*). These results indicate that CYP88D6 mainly catalyzes the sequential two-step oxidation of β-amyrin (**1**) at the C-11 position to yield 11-oxo-β-amyrin (**4**), a possible biosynthetic intermediate of glycyrrhizin.

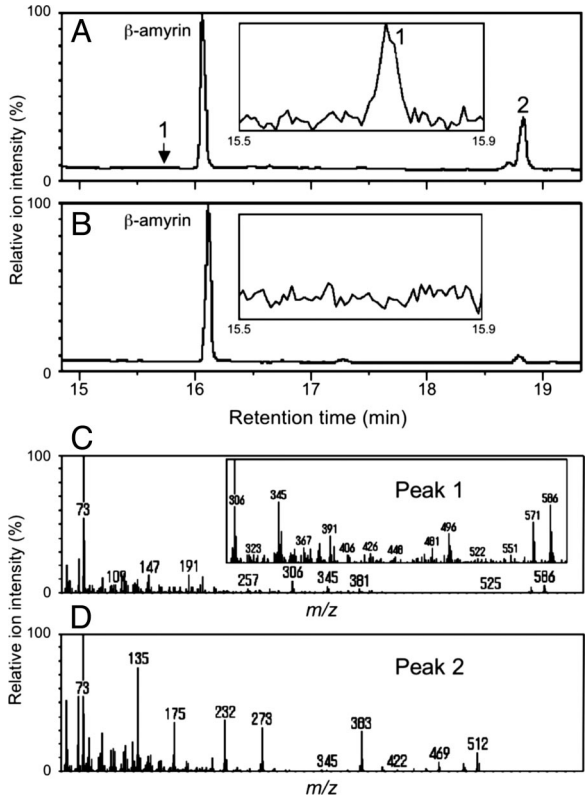

Fig. 2. *In vitro* oxidation of β-amyrin at C-11 by CYP88D6. GC-MS analysis (total ion chromatograms) of the reaction products resulting from *in vitro* assays containing β-amyrin (1) as the substrate and microsomal fractions isolated from (*A*) CYP88D6-expressing *Sf*9 cells or (*B*) empty vector control *Sf*9 cells. Insets show enlargements of the extracted ion chromatograms at *m/z* 586 corresponding to retention times of 15.5–15.9 min. (*C*) Mass spectrum of peak 1 from the GC profile shown in (*A*), with an enlargement of the spectrum corresponding to a *m/z* of 300–600 shown in the inset. (*D*) Mass spectrum of peak 2 from the GC profile shown in (*A*). The retention times and mass spectra of peaks 1 and 2 compare well with those of authentic 11α-hydroxy-β-amyrin (2) (Rt = 15.7) and 11-oxo-β-amyrin (4) (Rt = 18.8) (Fig. S1), respectively. GC was performed with an HP-5 column.

PLANT BIOLOGY

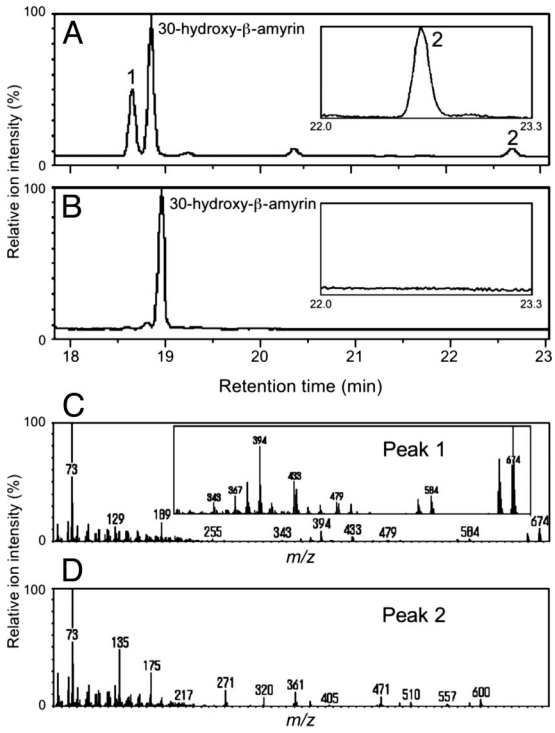

Fig. 3. *In vitro* oxidation of 30-hydroxy-β-amyrin at C-11 by CYP88D6. GC-MS analysis (total ion chromatograms) of the reaction products resulting from *in vitro* assays containing 30-hydroxy-β-amyrin (3) as the substrate and microsomal fractions isolated from (*A*) CYP88D6-expressing *Sf*9 cells or (*B*) empty vector control *Sf*9 cells. Enlargements of the chromatograms corresponding to retention times of 22.0–23.3 min are shown as insets. Mass spectra of peak 1 (*C*) and peak 2 (*D*) from the GC profile shown in (*A*) compare well with those of authentic 11α,30-dihydroxy-β-amyrin (5) (Rt = 18.6) and 30-hydroxy-11-oxo-β-amyrin (7) (Rt = 22.6) (Fig. S1), respectively. GC was performed with an HP-5 column.

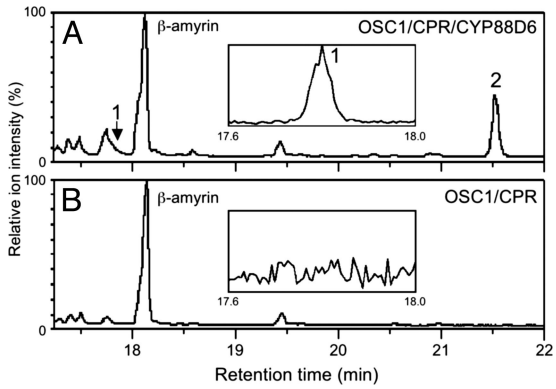

Fig. 4. *In vivo* formation of 11-oxo-β-amyrin in yeast cells co-expressing β-amyrin synthase (LjOSC1) and CYP88D6. GC-MS analysis (total ion chromatograms) of ethyl acetate extracts from yeast strains expressing (*A*) *LjOSC1*, *LjCPR1*, and *CYP88D6* (OSC1/CPR/CYP88D6), or (*B*) *LjOSC1* and *LjCPR1* (OSC1/CPR). Insets show enlargements of the extracted ion chromatograms at *m/z* 586 corresponding to retention times of 17.6–18.0 min. Retention times and mass spectra of peaks 1 and 2 in compare well with those of authentic 11α-hydroxy-β-amyrin (2) (Rt = 17.8) and 11-oxo-β-amyrin (4) (Rt = 21.5) (Fig. S1), respectively. GC analyses were performed with a DB-1 column.

The enzymatic activity of CYP88D6 was examined further using 30-hydroxy-β-amyrin (**3**) and 11-deoxoglycyrrhetinic acid (**6**) as potential substrates. When 30-hydroxy-β-amyrin (**3**) was used as the substrate, two compounds were formed within the CYP88D6-containing microsomes (Fig. 3*A*). The Rt and mass spectrum of each major (peak 1: Rt = 18.6 min; Fig. 3*C*) and minor product (peak 2: Rt = 22.6 min; Fig. 3*D*) showed excellent matches to those of authentic 11α,30-dihydroxy-β-amyrin (**5**) and 30-hydroxy-11-oxo-β-amyrin (**7**) (Fig. S1), respectively. In contrast, the peaks were not detected in assays using microsomes from empty vector control *Sf*9 cells (Fig. 3*B*). These results indicate that *in vitro*, CYP88D6 primarily catalyzes a single oxidation event when 30-hydroxy-β-amyrin (**3**) is used as the substrate. Moreover, no enzymatic activity was found when CYP88D6-containing microsomes were assayed with 11-deoxoglycyrrhetinic acid (**6**) as the substrate (data not shown).

Because some members of the CYP88A subfamily, including those from *Arabidopsis*, barley, and pea, have been biochemically characterized as *ent*-kaurenoic acid (diterpenoid) oxidases that function in gibberellin biosynthesis (21, 22), we also examined the potential *ent*-kaurenoic acid oxidase activity of CYP88D6. However, CYP88D6-mediated conversion of *ent*-kaurenoic acid to further oxidized diterpene products was not observed in our *in vitro* enzymatic activity assays (data not shown).

***In Vivo* CYP88D6 Enzymatic Activity Assay.** To verify the results of our *in vitro* assays (Fig. 2), we examined the β-amyrin oxidase activity of CYP88D6 in an engineered yeast strain that produces β-amyrin endogenously. For this purpose, *bAS* from *Lotus japonicus* (*LjOSC1*) (23) was expressed under the control of a constitutive promoter (*ADH1*) in the wild-type yeast strain BJ2168. Subsequently, CYP88D6 was co-expressed with cytochrome P450 reductase (CPR) from *L. japonicus* (named *LjCPR1*, GenBank accession no. AB433810; T. Akashi, T. Nakao, T. Aoki, and S. Ayabe, unpublished work) as a redox partner in *LjOSC1*-transformed yeast. The expression of *LjCPR1* and *CYP88D6* was under the control of the galactose-inducible promoters *GAL10* and *GAL1*, respectively. The overexpression of a plant-derived CPR is important for efficient electron transfer to plant P450s in recombinant yeast (24).

Following the culture of the transgenic yeast strain in medium containing galactose to induce the expression of *LjCPR1* and *CYP88D6*, ethyl acetate extracts of the cells were analyzed.

As shown in Fig. 4*A*, the *LjOSC1*/*LjCPR1*/*CYP88D6*-expressing yeast strain (OSC1/CPR/CYP88D6) converted β-amyrin to 11-oxo-β-amyrin (**4**) as the major product (peak 2: Rt = 21.5 min) and 11α-hydroxy-β-amyrin (**2**) as the minor product (peak 1: Rt = 17.8 min). Both compounds were unique to the yeast strain expressing *LjOSC1*, *LjCPR1*, and *CYP88D6*; the expression of *LjOSC1* and *LjCPR1* (OSC1/CPR) alone did not yield detectable levels of either compound (Fig. 4*B*).

We also analyzed the culture medium of the *LjOSC1*/*LjCPR1*/*CYP88D6*-expressing cells (Fig. S2*A*), and found that nearly 40% of the 11-oxo-β-amyrin synthesized was secreted into the medium, whereas only 4–8% of β-amyrin was present in the culture medium. The final yields of 11-oxo-β-amyrin and 11α-hydroxy-β-amyrin at 2 d of culture after galactose-induction were approximately 1.6 and 0.2 mg/l, respectively.

The yeast derived 11-oxo-β-amyrin (**4**) and 11α-hydroxy-β-amyrin (**2**) were purified on a silica gel column, and their identities were confirmed by NMR analyses (see *SI Materials*).

The identification of the *LjOSC1*/*LjCPR1*/*CYP88D6*-expressing yeast-derived products as 11-oxo-β-amyrin (**4**) and 11α-hydroxy-β-amyrin (**2**) indicates that CYP88D6 is capable, in two steps, of oxidizing β-amyrin to 11-oxo-β-amyrin (**4**), with 11α-hydroxy-β-amyrin (**2**) as a probable reaction intermediate. Small amounts of the reaction intermediate might be released as a general consequence of catalysis.

Expression of CYP88D6. The organ specificity of *CYP88D6* expression was analyzed by RT-PCR (Fig. 5). Transcripts of *CYP88D6*

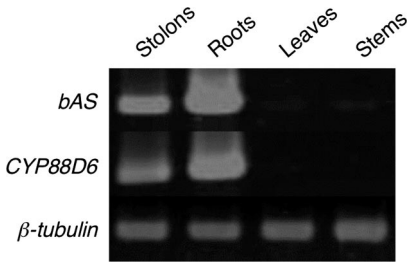

Fig. 5. Expression pattern of *CYP88D6*. RT-PCR analysis of β-amyrin synthase (*bAS*) and *CYP88D6* mRNA levels in the stolons, roots, leaves, and stems. *β-tubulin* was included as an internal control.

were detected in the roots and stolons, whereas no transcripts were observed in the leaves or stems. These results are similar to those for *bAS*, and are consistent with the fact that glycyrrhizin accumulates predominantly in the underground parts of *Glycyrrhiza* plants. These observations, combined with the data from our functional analyses, strongly suggest a key role for CYP88D6 as a β-amyrin 11-oxidase in glycyrrhizin biosynthesis.

CYP88D Is a Fabaceae-Specific Subfamily. Although the 493-residue sequence of CYP88D6 showed the highest degree of similarity to an unnamed protein product predicted from the genome of *Vitis vinifera* (53% identity; CAO70436) in a BLAST search, the P450 nomenclature committee (c/o Dr. D.R. Nelson) informed us that CYP88D6 showed 60% identity to CYP88D1, which corresponds to a predicted P450 from *Medicago truncatula* (ABE91093), although the record of ABE91093 was recently withdrawn from GenBank for unknown reasons.

In terms of enzymes of known function, CYP88D6 is most similar to the *ent*-kaurenoic acid oxidases CYP88A6 (50% identity) from pea (named PsKAO1) (22) and CYP88A4 (49% identity) from *Arabidopsis* (named AtKAO2) (21), both of which function in the biosynthesis of gibberellins.

Different from the CYP88A subfamily of enzymes, which are distributed widely among plants, the CYP88D subfamily appears to be restricted to the Fabaceae and is not found in *Arabidopsis* or rice (25). Extensive searches of plant EST databases (see *SI Methods*) identified several partial cDNA sequences potentially encoding additional CYP88D subfamily members from *M. truncatula* and *L. japonicus*; however, no ESTs encoding CYP88D1 were identified. These partial sequences were extended to full-length by RACE PCR, which resulted in the identification of four ORFs encoding additional CYP88D subfamily members: CYP88D2 (62% sequence identity with CYP88D6) and CYP88D3 (79% identity) of *M. truncatula*, and CYP88D4 (71% identity) and CYP88D5 (72% identity) of *L. japonicus*. Phylogenetic analysis (Fig. 6) showed that CYP88D1-D6 are grouped separately from the CYP88A enzymes and other CYP88 family proteins of unknown function, including CYP88B1 from tomato and CYP88E1 from *Selaginella moellendorffii* (the primary sequences of CYP88B1 and CYP88E1 may be found at http://drnelson.utmem.edu/BiblioD.html).

Discussion

In this study, we used an EST-based approach to identify *CYP88D6*, a P450 family gene that encodes β-amyrin 11-oxidase, which functions in the glycyrrhizin pathway.

Glycyrrhizin is biosynthesized via a series of oxidative reactions at the C-11 and C-30 positions of β-amyrin. The co-occurrence of two possible biosynthetic intermediates *in planta*, 11-oxo-β-amyrin (**4**) and 11-deoxoglycyrrhetinic acid (**6**) (K.O., K.S., and T.M., unpublished results), indicates that more than one biosynthetic route likely exists from β-amyrin to glycyrrhizin in terms of the order of oxidation at C-11 and C-30 (Fig. 1).

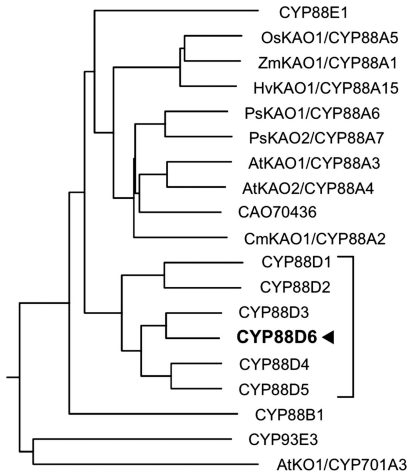

Fig. 6. Phylogenetic relationships among the CYP88 proteins. A phylogenetic tree was generated based on a comparison of the CYP88 primary sequences using the ClustalW program (26). The arrowhead indicates the P450 functionally defined in this study, and the bracket indicates the CYP88D subfamily. CYP88A family members include *ent*-kaurenoic acid oxidases (KAOs) functionally identified in *Arabidopsis* (AtKAO1 and AtKAO2; 21), pea (PsKAO1 and PsKAO2; 22), pumpkin (CmKAO1; 27), rice (OsKAO1, AP000616), maize (ZmKAO1; 28), and barley (HvKAO1; 21). The outgrouping CYPs consist of the *ent*-kaurene oxidase, CYP701A3 (AtKO1), from *Arabidopsis* (29) and a β-amyrin 24-hydoxylase, CYP93E3, from *G. uralensis*, identified as part of this work (see Fig. S3 and *SI Methods*).

Biochemical analyses suggest that CYP88D6 primarily catalyzes the sequential two-step oxidation of β-amyrin (**1**) at C-11 to yield 11-oxo-β-amyrin (**4**) (Figs. 2 and 4). Note that CYP88D6 readily catalyzed the hydroxylation of 30-hydroxy-β-amyrin (**3**) at C-11, but failed to convert the resultant compound, 11α,30-dihydroxy-β-amyrin (**5**), to the ketone 30-hydroxy-11-oxo-β-amyrin (**7**) with appreciable efficiency (Fig. 3). Moreover, CYP88D6 exhibited no enzymatic activity when 11-deoxoglycyrrhetinic acid (**6**) was used as the substrate (data not shown). These results suggest that the preferred route of glycyrrhizin biosynthesis involves oxidation at C-11 first; however, a more detailed and comparative biochemical characterization of CYP88D6 and the as yet unidentified enzyme(s) that catalyzes the oxidation at C-30 are necessary to verify this speculation.

In this study, we revealed a function of CYP88 family proteins as triterpene-oxidizing enzymes. Some members of the CYP88A subfamily, including those from *Arabidopsis*, barley, and pea, have been biochemically characterized as *ent*-kaurenoic acid (diterpenoid) oxidases that function in gibberellin biosynthesis (26–29, 22). Different from the CYP88A subfamily enzymes, which are found in a variety of plants, the CYP88D subfamily appears to be restricted to the Fabaceae. Based on extensive searches of plant EST databases and subsequent RACE PCR analyses, we identified two additional CYP88D subfamily members each from *M. truncatula* (CYP88D2 and 88D3) and *L. japonicus* (CYP88D4 and 88D5). The functional identification of CYP88D6 hints at the involvement of other P450s in the CYP88D subfamily in the biosynthesis of triterpene saponins. Because, to our knowledge, no oleanane-type triterpene aglycones containing a C-11 carbonyl group have been identified from *M. truncatula* or *L. japonicus*, these additional CYP88D subfamily members may each act on β-amyrin or related triterpene substrates with unique reaction specificity. Recently, Shibuya *et al.* (30) identified CYP93E1, a cytochrome P450 from soybean that participates in soyasaponin biosynthesis as a β-amyrin and sophoradiol 24-hydroxylase. By using a PCR-based method, we isolated a cDNA from *G. uralensis* encoding a

PLANT BIOLOGY

homolog (named CYP93E3; 82% identity) of CYP93E1. Moreover, our *in vitro* enzymatic activity assay showed that CYP93E3 possesses β-amyrin 24-hydroxylase activity (Fig. S3). These observations clearly indicate that in *G. uralensis* and most probably in other legumes as well, the P450 enzymes involved in triterpene saponin biosynthesis are recruited in at least two very distant CYP families, CYP93 (CYP71 clan) of the A-type P450s and CYP88 (CYP85 clan) of the non-A type P450s. Plant P450s are generally assigned to two major clades: A-type and non-A type (31). Another example of the convergent evolution of P450s comes from the biosynthesis of gibberellins in higher plants. In *Arabidopsis*, CYP701A3 and CYP88As (CYP88A3 and A4) have been shown to catalyze very similar enzymatic reactions and two successive steps in the gibberellin biosynthetic pathway (21, 29, 32); however, CYP701A (CYP71 clan) is an A-type P450 whereas the CYP88As (CYP85 clan) are non-A type P450s. These examples clearly show that P450s involved in the same pathway or that catalyze similar reactions are not always phylogenetically related.

Licorice production depends on the collection of wild *Glycyrrhiza* plants, especially in China, and this has caused a decrease in licorice reserves and an increase in desertification (33). Moreover, the chemical synthesis of glycyrrhizin is difficult both in terms of quantity and cost. Therefore, the production of glycyrrhizin through plant tissue culture, hairy root culture (34), and cell culture (35) has been attempted, although the efforts were all unsuccessful. The possibility exists that *CYP88D6* and additional genes could be used to engineer the production of glycyrrhizin in a plant and/or microbial host. In this study, we described the successful production of a glycyrrhizin pathway intermediate, 11-oxo-β-amyrin (1.6 mg/l), in a wild-type yeast background by the coexpression of *bAS* (LjOSC1), CYP88D6, and LjCPR1 as a redox partner for CYP88D6 (Fig. 4 and Fig. S4). Further genetic optimization to enhance the availability of 2,3-oxidosqualene as described by Kirby *et al.* (36), by overexpression of a truncated form of 3-hydroxy-3-methylglutaryl CoA (HMG-CoA) reductase and down-regulation of *ERG7* (lanosterol synthase), may improve production. The conversion of 11-oxo-β-amyrin to glycyrrhizin will require further oxidation at the C-30 position, and this conversion could potentially involve other P450 enzymes. Functional analyses of the other CYP88D subfamily members that may exist in *G. uralensis* but which may be missing from our EST collection are thus a high priority.

Materials and Methods

Chemicals. The NMR spectra were recorded on a DSX-300 (Bruker) and ECA-500 (JEOL) spectrometer in $CDCl_3$. β-Amyrin (**1**) was synthesized by a previously described method (37). Glycyrrhetinic acid (**8**) was purchased from Sigma. Other standard samples were synthesized from either β-amyrin or glycyrrhetinic acid as described in the *SI Methods* and Fig. S5. ^1H and ^{13}C NMR chemical shifts are reported as δ-values based on the internal TMS ($\delta_H = 0$) and reference solvent signal ($CDCl_3$ $\delta_C = 77.0$).

***In Vitro* Enzymatic Activity Assay.** Heterologous expression of CYP88D6 using a *baculovirus*–insect cell system was performed as described in the *SI Methods*. The microsomal fraction of insect cells expressing CYP88D6 was prepared as previously described (38). The enzymatic activity of CYP88D6 was reconstituted by mixing CYP88D6-containing microsomes with purified *Arabidopsis* NADPH-P450 reductase (39). The assay was conducted in a total volume of 500 μl of 50 mM potassium phosphate buffer (pH 7.25) containing 1 mM NADPH, 0.1 U/ml NADPH-P450 reductase, 100 μg of microsomal protein, and 20 μM substrate. The samples were incubated for 2 h at 30 °C, and the reaction products were extracted three times with 0.5 volume of ethyl acetate. The organic phase was evaporated, trimethylsilylated, and analyzed by GC-MS.

Yeast *in Vivo* Assays. The plasmids used in our yeast *in vivo* assays, pYES3-ADH-OSC1 (constitutive expression of *bAS* driven by the *ADH1* promoter), pELC (galactose-inducible expression of *LjCPR1* driven by the *GAL10* promoter), and pDEST52-CYP88D6 (galactose-inducible expression of *CYP88D6* driven by the *GAL1* promoter) were constructed as described in the *SI Methods*. *Saccharomyces cerevisiae* BJ2168 (*MATa*, *prc1–407*, *prb1–1122*, *pep4–3*, *leu2*, *trp1*, *ura3–52*, and *gal2*; Nippon Gene) cells harboring both pYES3-ADH-OSC1 and pELC were transformed with pDEST52-CYP88D6 or pYES2 empty vector (Invitrogen) as a control. The recombinant cells were cultured in synthetic complete medium containing 2% glucose without uracil, tryptophan, or leucine (SC-U-W-L) for 2 d at 28 °C. The cells were then collected and resuspended in SC-U-W-L medium containing 13 μg/ml hemin and 2% galactose instead of glucose, and cultured at 28 °C for 2 d. The cells were then harvested and extracted with ethyl acetate, and portions of the extracts were analyzed by GC-MS after trimethylsilation.

GC-MS and NMR Analysis. See *SI Methods*.

RT-PCR Analysis. Total RNA was extracted from the roots and stolons using RNAwiz (Ambion Inc.), and from the leaves and stems with cetyltrimethylammonium bromide (CTAB)-containing extraction buffer essentially as described previously (40). Total RNA was treated with RNase-free DNase (TaKaRa Bio) and further purified using an RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's specifications. First-strand cDNA synthesis was carried out using a SMART RACE cDNA Amplification Kit (Clontech) with 1 μg of total RNA. PCR was performed using primers 1 and 2 for *bAS*, primers 3 and 4 for *CYP88D6*, and primers 5 and 6 for *β-tubulin* (for primer sequences see Table S1), with 33 cycles for *bAS* and *CYP88D6*, and 30 cycles for *β-tubulin*.

Isolation of CYP88D Subfamily Members from *M. truncatula* and *L. japonicus*. See *SI Methods*.

ACKNOWLEDGMENTS. We thank Drs. S. Ayabe (Nihon University) and Y. Ebizuka (University of Tokyo) for encouragement, and Drs. H. Hayashi (Iwate Medical University) and M. Kojoma (Health Sciences University of Hokkaido) for providing the plant materials and for valuable discussions. We thank Drs. N. Sakurai, H. Suzuki, and D. Shibata (Kazusa DNA Research Institute) and Drs. A. Toyoda, Y. Totoki, and Y. Sakaki (RIKEN GSC) for constructing the EST databases, and Drs. T. Shibata and O. Iida (Research Center for Medicinal Plant Resources) for providing the initial *G. uralensis* samples used to construct the cDNA library. We thank Dr. M. Shibuya (The University of Tokyo) for the gift of authentic 24-hydroxy-β-amyrin, and Dr. S. Yamaguchi (RIKEN PSC) for the gift of authentic *ent*-kaurenoic acid and GA12. We also thank Dr. S. Seo and Y. Toya (Tokiwa Phytochemical Co., Ltd.) for technical advice, and Drs. C. J. Arntzen (Arizona State Univ.) and M. Suzuki (Kihara Institute for Biological Research) for helpful discussions. T. Nakao (Nihon University) and T. Nishizawa (Kihara Institute for Biological Research) provided technical support. This study was supported in part by a grant from the New Energy and Industrial Technology Development Organization (NEDO), Japan, and by a Grant-in-Aid (No. 18780082) for Young Scientists (B) to H. S. from the Ministry of Education, Culture, Sports, Science and Technology.

1. Waller G-R, Yamasaki K eds. (1996) *Saponins Used in Food and Agriculture*: *Advances in Experimental Medicine and Biology*. (Plenum Press, NewYork), Vol. 405.
2. Gibson M-R (1978) Glycyrrhiza in old and new perspectives. *Lloydia* 41:348–354.
3. Matsui S, *et al.* (2004) Glycyrrhizin and related compounds down-regulate production of inflammatory chemokines IL-8 and eotaxin 1 in a human lung fibroblast cell line. *Int Immunopharmacol* 4:1633–1644.
4. Takahara T, Watanabe A, Shiraki K (1994) Effects of glycyrrhizin on hepatitis B surface antigen: A biochemical and morphological study. *J Hepatol* 21:601–609.
5. He J-X, Akao T, Nishino T, Tani T (2001) The influence of commonly prescribed synthetic drugs for peptic ulcer on the pharmacokinetic fate of glycyrrhizin from Shaoyao-Gancao-tang. *Biol Pharm Bull* 24:1395–1399.
6. Park H-Y, Park S-H, Yoon H-K, Han M-J, Kim D-H (2004) Anti-allergic activity of 18beta-glycyrrhetinic acid-3-*O*-beta-D-glucuronide. *Arch Pharm Res* 27:57–60.
7. Fiore C, *et al.* (2008) Antiviral effects of *Glycyrrhiza* species. *Phytother Res* 22:141–148.
8. Baba M, De Cleroq S, Nakashima H, Yamamoto N (1988) Mechanism of inhibitory effect of glycyrrhizin on replication of human immunodeficiency virus (HIV). *Antiviral Res* 10:289–298.
9. Ito M, *et al.* (1987) Inhibitory effect of glycyrrhizin on the in vitro infectivity and cytopathic activity of the human immunodeficiency virus [HIV (HTLV-III/LAV)]. *Antiviral Res* 7:127–137.
10. Cinatl J, *et al.* (2003) Glycyrrhizin, an active component of liquorice roots, and replication of SARS-associated coronavirus. *Lancet* 361:2045–2046.
11. Shibata S (2000) A drug over the millennia: pharmacognosy, chemistry, and pharmacology of licorice. *Yakugaku Zasshi* 120:849–862.
12. van Rossum T-G-J, Vulto A-G, de Man R-A, Brouwer J-T, Schalm S-W (1998) Glycyrrhizin as a potential treatment for chronic hepatitis C Aliment. *Pharmacol Ther* 12:199–205.
13. Parker P-M (2006) *The world market for licorice roots: A 2007 global trade perspective*. (ICON Group International Inc, San Diego).

14. Abe I, Rohmer M, Prestwich G-D (1993) Enzymatic cyclization of squalene and oxidosqualene to sterols and triterpenes. *Chem Rev* 93:2189–2206.
15. Hayashi H, Hirota A, Hiraoka N, Ikeshiro Y (1999) Molecular cloning and characterization of two cDNAs for *Glycyrrhiza glabra* squalene synthase. *Biol Pharm Bull* 22:947–950.
16. Hayashi H, *et al.* (2001) Cloning and characterization of a cDNA encoding β-amyrin synthase involved in glycyrrhizin and soyasaponin biosyntheses in licorice. *Biol Pharm Bull* 24:912–916.
17. Schoendorf A, Rithner C-D, Williams R-M, Croteau, R-B (2001) Molecular cloning of a cytochrome P450 taxane 10β-hydroxylase cDNA from *Taxus* and functional expression in yeast. *Proc Natl Acad Sci USA* 98:1501–1506.
18. Jennewein S, Rithner C-D, Williams R-M, Croteau R-B (2001) Taxol biosynthesis: taxane 13α-hydroxylase is a cytochrome P450-dependent monooxygenase. *Proc Natl Acad Sci USA* 98:13595–13600.
19. Teoh K-H, Polichuk D-R, Reed D-W, Nowak G, Covello P-S (2006) *Artemisia annua* L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. *FEBS Lett* 580:1411–1416.
20. Ro D-K, *et al.* (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440:940–943.
21. Helliwell C-A, Chandler P-M, Poole A, Dennis E-S, Peacock W-J (2001) The CYP88A cytochrome P450, *ent*-kaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proc Natl Acad Sci USA* 98:2065–2070.
22. Davidson S-E, Elliott R-C, Helliwell C-A, Poole A-T, Reid J-B (2003) The pea gene NA encodes *ent*-Kaurenoic acid oxidase. *Plant Physiol* 131:335–344.
23. Sawai S, *et al.* (2006) Functional and structural analysis of genes encoding oxidosqualene cyclases of *Lotus japonicus. Plant Sci* 170:247–257.
24. Pompon D, Louerat B, Bronine A, Urban P (1996) Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol* 272:51–64.
25. Nelson D-R, Schuler M-A, Paquette S-M, Werck-Reichhart D, 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.
26. Thompson J-D, Higgins D-G, Gibson T-J (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680.
27. Helliwell C-A, *et al.* (2000) Isolation of an *ent*-kaurene oxidase cDNA from *Cucurbita maxima*. *Aust. J Plant Physiol* 27:1141–1149.
28. Winkler R-G, Helentjaris T (1995) The maize Dwarf3 gene encodes a cytochrome P450-mediated early step in gibberellin biosynthesis. *Plant Cell* 7:1307–1317.
29. Helliwell C-A, *et al.* (1998) Cloning of the *Arabidopsis ent*-kaurene oxidase gene GA3. *Proc Natl Acad Sci USA* 95:9019–9024.
30. Shibuya M, *et al.* (2006) Identification of beta-amyrin and sophoradiol 24-hydroxylase by expressed sequence tag mining and functional expression assay. *FEBS J* 273:948–959.
31. Durst F, Nelson D-R (1995) Diversity and evolution of plant P450 and P450 reductases. *Drug Metabol Drug Interact* 12:189–206.
32. Helliwell C-A, Poole A, Peacock W-J, Dennis E-S (1999) Arabidopsis *ent*-kaurene oxidase catalyzes three steps of gibberellin biosynthesis. *Plant Physiol* 119:507–510.
33. Yamamoto Y, Tani T (2005) Field study and pharmaceutical evaluation of *Glycyrrhiza* uralensis roots cultivated in China (Material study). *J Trad Med* 22:86–97.
34. Saito K, Kaneko H, Yamazaki M, Yoshida M, Murakoshi I (1990) Stable transfer and expression of chimeric genes in licorice (*Glycyrrhiza uralensis*) using an Ri plasmid binary vector. *Plant Cell Rep* 8:718–721.
35. Hayashi H, Fukui H, Tabata M (1988) Examination of triterpenoids produced by callus and cell suspension cultures of *Glycyrrhiza glabra. Plant Cell Rep* 7:508–511.
36. Kirby J, Romanini D-W, Paradise E-M, Keasling J-D (2008) Engineering triterpene production in *Saccharomyces cerevisiae*–β-amyrin synthase from *Artemisia annua. FEBS J* 275:1852–1859.
37. Ohyama K, Suzuki M, Masuda K, Yoshida S, Muranaka T (2007) Chemical phenotypes of the *hmg1* and *hmg2* mutants of *Arabidopsis* demonstrate the *in planta* role of HMG-CoA reductase in triterpene biosynthesis. *Chem Pharm Bull* 55:1518–1521.
38. Ohnishi T, *et al.* (2006) C-23 hydroxylation by Arabidopsis CYP90C1 and CYP90D1 reveals a novel shortcut in brassinosteroid biosynthesis. *Plant Cell* 18:3275–3288.
39. Mizutani M, Ohta D (1998) Two isoforms of NADPH: Cytochrome P450 reductase in *Arabidopsis thaliana* Gene structure, heterologous expression in insect cells, and differential regulation. *Plant Physiol* 116:357–367.
40. Chang S, Puryear J, Cainey J (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Mol Biol Rep* 11:113–116.