FEBS Letters 589 (2015) 1778–1786

journal homepage: www.FEBSLetters.org

Identification and functional analysis of 2-hydroxyflavanone *C*-glucosyltransferase in soybean (*Glycine max*)

Yoshihiro Hirade [a,*], Naoyuki Kotoku [b], Kazuyoshi Terasaka [c], Yumiko Saijo-Hamano [a], Akemi Fukumoto [a], Hajime Mizukami [c,d]

[a] *Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamada-oka, Suita City, Osaka 565-0871, Japan*
[b] *Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamada-oka, Suita City, Osaka 565-0871, Japan*
[c] *Graduate School of Pharmaceutical Sciences, Nagoya City University, 3-1 Tanabe-dori, Mizuho-ku, Nagoya City, Aichi 467-8603, Japan*
[d] *The Kochi Prefectural Makino Botanical Garden, 4200-6 Godaisan, Kochi City, Kochi 781-8125, Japan*

ARTICLE INFO

Article history:
Received 25 February 2015
Revised 28 April 2015
Accepted 1 May 2015
Available online 12 May 2015

Edited by Ulf-Ingo Flügge

Keywords:
2-Hydroxyflavanone *C*-glucosyltransferase
Enzymatic aryl-*C*-glucosylation
C-Glucoside biosynthesis
Catalytic residues
Structural modeling
Secondary metabolites

ABSTRACT

_C_-Glucosyltransferase is an enzyme that mediates carbon–carbon bond formation to generate _C_-glucoside metabolites. Although it has been identified in several plant species, the catalytic amino acid residues required for _C_-glucosylation activity remain obscure. Here, we identified a 2-hydroxyflavanone _C_-glucosyltransferase (UGT708D1) in soybean. We found that three residues, His20, Asp85, and Arg292, of UGT708D1 were located at the predicted active site and evolutionarily conserved. The substitution of Asp85 or Arg292 with alanine destroyed _C_-glucosyltransferase activity, whereas the substitution of His20 with alanine abolished _C_-glucosyltransferase activity but enabled _O_-glucosyltransferase activity. The catalytic mechanism is discussed on the basis of the findings.

1. Introduction

Plants produce various flavonoid derivatives that are often *C*-glycosylated [1–3], as exemplified by *C*-glucosylflavones, vitexin and isovitexin, wherein the anomeric carbon of the glucose moiety is directly linked to the aromatic carbon of the flavone skeleton (see Fig. 1 for chemical structures). *C*-Glucosylflavones are prevalent in both monocots and dicots and are attracting attention because they exhibit various pharmacological activities [4–6].

C-Glucosylflavones are biosynthesized from flavanone via *C*-glucosylation of 2-hydroxyflavanone or flavone (Fig. 1). *C*-Glucosyltransferases (CGTs), which *C*-glucosylate either the C-8 or -6 position of 2-hydroxyflavanone, have been identified in the monocots rice (*Oryza sativa*) and maize (*Zea mays*) and the dicot buckwheat (*Fagopyrum esculentum*) [7–9]. A CGT that *C*-glucosylates the C-6 position of flavone has recently been identified in the dicot *Gentiana triflora* [10]. These enzymes use UDP-glucose as a sugar donor and catalyze the transfer of the glucose moiety to the aromatic carbon of the acceptor substrate.

Plant CGTs identified to date belong to family 1 glycosyltransferases (http://www.cazy.org/GlycosylTransferases.html), which are widely distributed among organisms and are known as UDP-sugar-dependent glycosyltransferases (UGTs). In plants, UGTs comprise a highly divergent multigene family and are involved in the glycosylation of a wide range of small lipophilic molecules [11,12]. Most plant UGTs isolated to date show *O*-glycosylation activity; however, UGTs with *S*- or *N*-glycosylation activities have also been reported [13,14]. The crystal structures of several plant UGTs have been revealed to contain two structural domains: N-terminal acceptor-binding and

Abbreviations: CGT, *C*-glucosyltransferase; EAS, electrophilic aromatic substitution; EST, expressed sequence tag; HPLC, high-performance liquid chromatography; MS, mass spectrometry; OGT, *O*-glucosyltransferase; *Os*CGT, *Oryza sativa* *C*-glucosyltransferase; PSPG, plant secondary product glycosyltransferase; UDP, uridine diphosphate; UGT, UDP-sugar-dependent glycosyltransferase

Author contributions: HM and KT conceived the study; YH designed the study and conducted major parts of the study; KT contributed to NMR analysis; NK contributed to synthesis of 2-hydroxynaringenin and mass spectrometry analysis; YS-H and AF contributed to DNA manipulations, protein expression, and protein purification; YH wrote the manuscript; HM helped revising the manuscript.

* Corresponding author. Fax: +81 72 730 7478.
E-mail address: hiradeyoshihiro@gmail.com (Y. Hirade).

http://dx.doi.org/10.1016/j.febslet.2015.05.010
0014-5793/

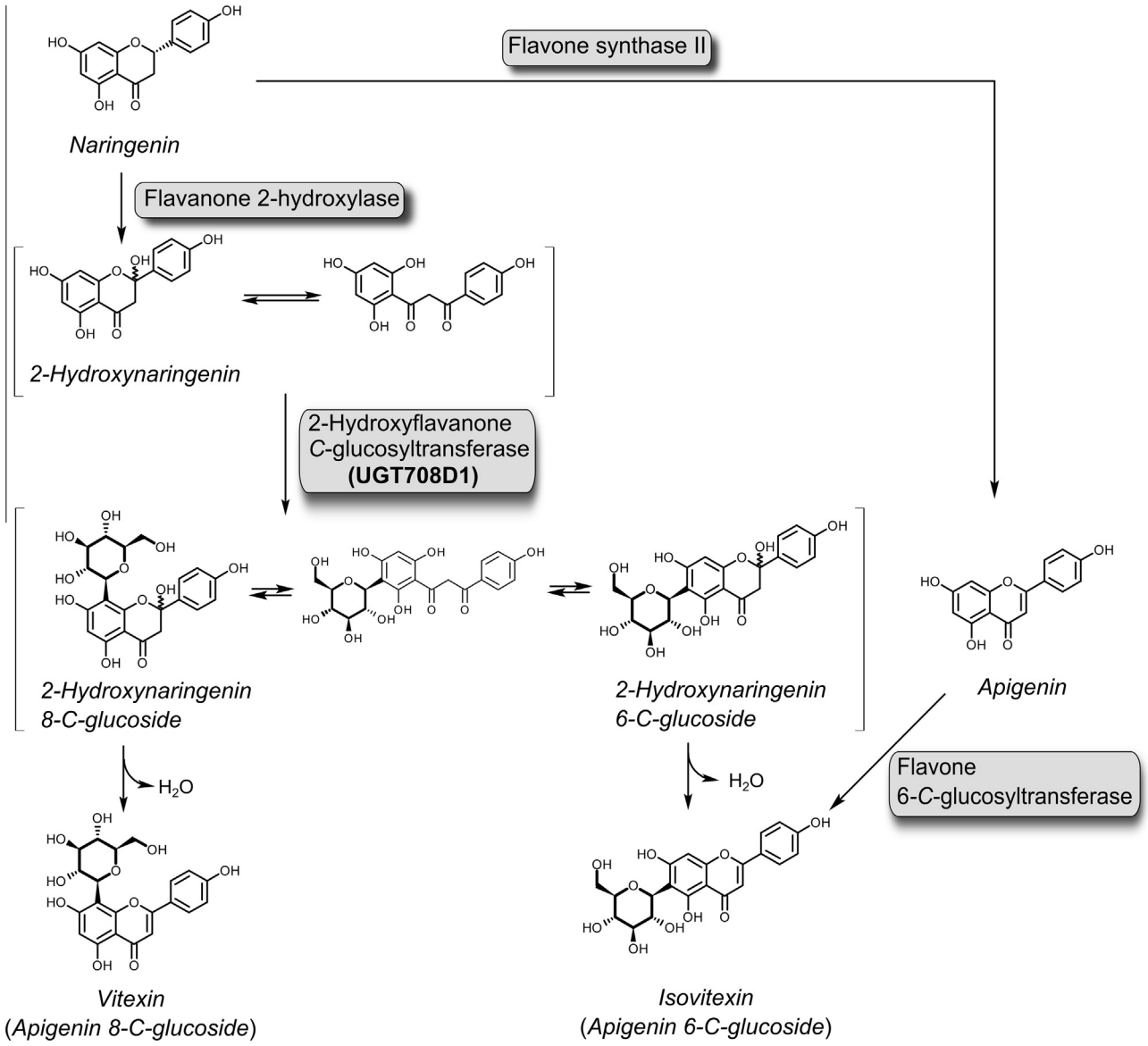

Fig. 1. Proposed biosynthetic pathway for *C*-glucosylflavones. *C*-Glucosylflavones (vitexin and isovitexin) are biosynthesized from flavanone (naringenin) via *C*-glucosylation of 2-hydroxyflavanone (2-hydroxynaringenin) [7–9] or flavone (apigenin) [10]. 2-Hydroxynaringenin is present in equilibrium between open-chain and ring-closed forms, and either or both of them are *C*-glucosylated by 2-hydroxyflavanone CGT. *C*-Glucosylated-2-hydroxyflavanone can be present in equilibrium among the ring-closed forms (6-*C*-glucoside and 8-*C*-glucoside) and open-chain form, and the ring-closed forms are spontaneously or enzymatically dehydrated to yield corresponding *C*-glucosylflavones. Glucose moieties are indicated by bold lines.

C-terminal sugar-donor-binding domains, which form a catalytic pocket, and a highly conserved plant secondary product glycosyltransferase (PSPG) motif in the C-terminal domain recognizes the UDP-sugar [14–18].

In contrast to *O*-glucosyltransferases (OGTs), which catalyze O–C bond formation to produce *O*-glucoside metabolites, relatively little is known about CGTs. A few reports based on the analysis of the rice 2-hydroxyflavanone CGT (*Os*CGT) have proposed that enzymatic aryl-*C*-glucosylation proceeds by an electrophilic aromatic substitution (EAS) mechanism [7,19]. However, the catalytic amino acid residues that can properly account for the reaction remain unknown owing to the absence of structural information.

In the present study, a 2-hydroxyflavanone CGT was identified in the dicot soybean (*Glycine max*). Catalytic residues required for CGT activity were investigated by structural modeling and site-directed mutagenesis.

2. Materials and methods

2.1. Chemicals

2-Hydroxynaringenin was synthesized according to an earlier report [20]. UDP-glucose and naringenin were purchased from Wako. Apigenin, isovitexin, phloretin, and vitexin were purchased from Sigma–Aldrich. All other chemicals were of commercial reagent-grade quality.

2.2. Plasmids

The soybean cDNA clone (GMFL02-51-N04) was purchased from the National Bioresource Project (http://www.legumebase.brc.miyazaki-u.ac.jp/top.jsp) [21]. Plasmid construction and

site-directed mutagenesis were performed using standard procedures (see Supplementary methods for details).

2.3. Heterologous expression

Escherichia coli strain BL21 (DE3) and vector pET-15b (Novagen) were used for heterologous expression. Cells carrying the recombinant plasmid were incubated at 30 °C in Luria–Bertani medium containing 50 μg/mL ampicillin until OD600 reached 0.5–0.6. After adding isopropyl 1-thio-β-D-galactoside to a final concentration of 0.4 mM, the cells were further incubated at 20 °C for 16 h to induce expression. Protein expression was confirmed by Western blot analysis using anti-His-tag antibody (QIAGEN).

2.4. Crude enzyme preparation

Cells expressing the recombinant protein were resuspended in 50 mM Tris–HCl buffer (pH 8.5) (500 μL/1 g cells) and disrupted by sonication. After centrifugation at 15000×*g* for 20 min, the supernatant was used as a crude enzyme.

2.5. Purification of recombinant enzyme

The recombinant protein was purified by affinity chromatography on a His-TRAP column (GE Healthcare) with a linear gradient of imidazole (pH 8.0), followed by anion-exchange chromatography on a Resource Q column (GE Healthcare) with a linear gradient of NaCl (pH 8.5), using an ÄKTA purifier system (GE Healthcare), at a temperature below 4 °C (Supplementary Fig. S1). The protein concentration was estimated using the extinction coefficient at 280 nm of 1.123 mg/mL/cm (http://web.expasy.org/protparam/).

2.6. Glucosyltransferase assay

Enzymatic activity was assayed with 0.8 mM UDP-glucose, 0.8 mM acceptor substrate, and recombinant protein (crude enzyme or purified enzyme) in 50 mM Tris–HCl buffer (pH 8.5) in a total volume of 50 μL at 30 °C for 60 min. The reaction was stopped by the addition of twice the volume of methanol. After centrifugation, the supernatant was used for high-performance liquid chromatography (HPLC) analysis.

2.7. Acid treatment

Acid treatment was applied as described in the Supplementary methods.

2.8. HPLC

HPLC was performed on an ODS column (CAPCELL PAK C18 MG III, 2.0 × 150 mm, 5 μm, Shiseido) with a NanoSpace SI-2 instrument (Shiseido) equipped with a photodiode array detector (Thermo Scientific). HPLC conditions are as follows: flow rate, 200 μL/min; detection, absorbance of 287 nm; solvent A, 0.5% formic acid; solvent B, acetonitrile; gradient, 10–80% B in 20 min.

2.9. Mass spectrometry (MS)

MS analysis was performed with a Q-Tof Ultima API mass spectrometer (Waters Corporation). The MS conditions are in Supplementary Table S1. Detailed results of MS are shown in Supplementary Table S2.

2.10. Structural modeling

The structural model of UGT708D1 (Supplementary model) was built using the program MODELLER version 9.13 [22] based on the crystal structure of UGT71G1 (PDB ID: 2ACW) [15]. Pairwise alignment for modeling was performed with SSEARCH [23,24] (Supplementary Fig. S2). The structure of UDP-glucose molecule in the template was transferred into the generated model as a rigid body using the MODELLER program.

3. Results

3.1. Rice C-glucosyltransferase homologs in soybean

Soybean produces vitexin in its roots [25], suggesting that the genes responsible for vitexin biosynthesis, including CGT, are present in this plant and expressed in the roots. Because phylogenetic studies with functionally characterized UGTs have found a cross-species correlation between phylogenetic clades and substrate specificity [26–28], phylogenetic analysis was performed to determine whether closely related homologs clustering with *Os*CGT, which was the first characterized CGT in plants [7], are present in the soybean genome [29]. Protein sequences of translated open reading frames (ORFs) predicted from the soybean genome were obtained from Phytozome v9.0 (http://www.phytozome.net/), and a BLAST search using the protein sequence of *Os*CGT as a query was performed against the data set. The protein sequences of the top 100 ORFs from the BLAST results (Supplementary Table S3) together with *Os*CGT were aligned and subjected to phylogenetic tree construction. Two ORFs *Glyma09g29160.1* and *Glyma16g33750.1* were found to fall into a single cluster with *Os*CGT (Supplementary Fig. S3, Table 1). To narrow the candidate set, the transcriptional profiles of the ORFs were investigated by searching the NCBI EST database. Of the two genes, only *Glyma09g29160.1* was expressed in vitexin-producing roots (Supplementary Table S4), suggesting its possible involvement in vitexin biosynthesis. *Glyma09g29160.1* was designated as UGT708D1 by the UGT Nomenclature Committee (http://www.flinders.edu.au/medicine/sites/clinical-pharmacology/ugt-homepage.cfm). The nucleotide sequence of UGT708D1 was deposited in the DDBJ/EMBL-Bank/GenBank database (accession number LC003312).

3.2. CGT activity of UGT708D1

Based on the in silico analysis, the enzymatic activity of UGT708D1 was assayed using 2-hydroxynaringenin, a possible precursor in vitexin biosynthesis, as a substrate. A soybean cDNA clone encoding the full-length UGT708D1 was obtained from the National Bioresource Project (http://www.legumebase.brc.miyazaki-u.ac.jp/top.jsp), and heterologous expression of UGT708D1 was achieved in *E. coli*. HPLC analysis indicated that UGT708D1 had UDP-glucose-dependent enzymatic activity toward 2-hydroxynaringenin (Fig. 2). MS analysis of *peak 2* of Fig. 2 showed a deprotonated molecular ion of *m*/*z* 449.109, which was assigned as 2-hydroxynaringenin monoglucoside. To establish the type of linkage between the aglycone and glucose moiety, tandem

Table 1
OsCGT homologs in soybean.

Gene locus ID[a]	Length[b] (aa)[c]	BLAST results			
		% Identity	% Similarity	*E*-Value	Score (bit)
Glyma09g29160.1	480	41	58	4.00e−82	305
Glyma16g33750.1	480	42	60	6.00e−104	323

[a] Gene locus abbreviations are as follows: Glyma, *Glycine max*.
[b] Putative protein length.
[c] Amino acids.

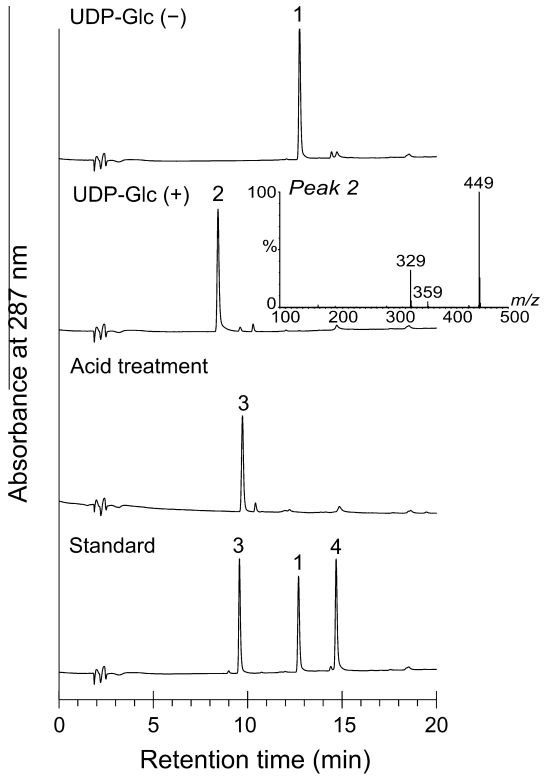

Fig. 2. HPLC analysis of UGT708D1 reaction products. Enzymatic activity of recombinant UGT708D1 protein toward 2-hydroxynaringenin was assayed with (second panel) and without (top panel) UDP-glucose. The reaction products were analyzed by HPLC. The MS/MS spectrum of the parent ion at *m*/*z* 449 from *peak 2* is shown in inset. HPLC profile of the UGT708D1 reaction product after acid treatment is shown in third panel. Reference standards are as follows: 1, 2-hydroxynaringenin; 3, vitexin; 4, apigenin.

mass spectrometry (MS/MS) analysis was performed. The product ion spectrum of the precursor ion at *m*/*z* 449 showed ions of $[M-H-90]^-$ at *m*/*z* 359 and $[M-H-120]^-$ at *m*/*z* 329 (Fig. 2, inset) corresponding to the cross-ring cleavage of an aryl-*C*-glucosyl moiety (Supplementary Fig. S4), indicating that the glucose moiety was attached to 2-hydroxynaringenin by a C–C bond [30]. To reveal the linkage position of the glucosyl moiety, acquisition of a nuclear magnetic resonance (NMR) spectrum of *peak 2* was attempted, but unsuccessful owing to instability of this compound. However, when the reaction product was treated with acid for product identification, *peak 3* was generated with disappearance of *peak 2* (Fig. 2). The retention time of *peak 3* was consistent with those of vitexin and isovitexin, the dehydration products of 8- or 6-*C*-glucosylated 2-hydroxynaringenin, respectively. These compounds could not be separated under these HPLC conditions. However, HPLC analysis using different conditions resolved *peak 3* into two compounds whose retention times coincided with those of vitexin and isovitexin (Supplementary Fig. S5). Thus, recombinant UGT708D1 catalyzed *C*-glucosylation of the A ring of 2-hydroxynaringenin in the presence of UDP-glucose. The catalytic activity toward 2-hydroxynaringenin was approximately 11 nkat/mg.

UGT708D1 showed enzymatic activity toward phloretin but not toward naringenin and apigenin (data not shown). This substrate specificity is similar to that of buckwheat CGT or *Os*CGT [7,9].

3.3. Predicted active site of UGT708D1

To investigate the catalytic residues for CGT activity, a structural model of UGT708D1 was constructed using the crystal structure of *Medicago truncatula* UGT71G1 [15], which showed the highest sequence homology with UGT708D1 among UGTs with available crystal structure, as a template. The predicted structure revealed the N-terminal acceptor-binding and C-terminal sugar-donor-binding pockets (Supplementary Fig. S6). Because both CGTs and OGTs use UDP-glucose as a sugar donor, the key factors that determine the *C*-glucosylation activity are likely to lie in the acceptor-binding pocket rather than in the sugar-donor-binding pocket. Also, enzymatic *C*-glucosylation implies the involvement of charged amino acid residues [31,32]. Based on these considerations, our investigation focused on the charged residues around the putative acceptor-binding pocket. In the structural model, four charged residues, namely His20, Asp85, Arg202, and Arg292, were found to be located around the acceptor-binding pocket (Fig. 3).

3.4. Sequence comparison of UGT708D1

To assess the evolutionary conservation of these residues, the protein sequence of UGT708D1 was compared with that of functionally characterized flavonoid UGTs. Sequence alignment showed that His20 of UGT708D1 was commonly conserved among the UGTs, irrespective of the difference between CGTs and OGTs (Fig. 4). Interestingly, Asp85 and Arg292 of UGT708D1 were specifically conserved in the 2-hydroxyflavanone CGTs characterized to date (Fig. 4), implying the functional importance of these residues in *C*-glucosylation reaction. No marked difference was observed in the highly conserved PSPG motif between CGTs and OGTs (Fig. 4), suggesting that the determinant factors for *C*-glucosylation activity were not present in the motif.

3.5. Mutation analysis of UGT708D1

To investigate the functional role of these residues, four alanine substitution mutant proteins (H20A, D85A, R202A, and R292A) were generated and assayed for their in vitro activity toward 2-hydroxynaringenin. HPLC analysis indicated that the D85A mutant and the R292A mutant lost catalytic activity, whereas the R202A mutant showed CGT activity, as did the wild-type UGT708D1 enzyme (Fig. 5A). The H20A mutant produced no *C*-glucoside product (*peak 2*); however, another small peak (*peak*

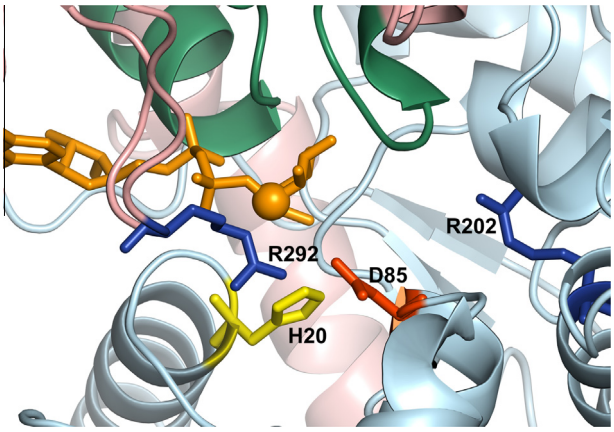

Fig. 3. Putative active site of UGT708D1. A cartoon representation of the predicted active site of UGT708D1 is shown. The N-terminal domain (residues 1–254), C-terminal domain (residues 255–480), and conserved PSPG motif (residues 353–396) are colored in light blue, pink, and green, respectively. UDP-glucose is indicated by the orange stick model, and the anomeric carbon is modeled as a sphere. Four charged residues around the acceptor binding site are drawn as a stick model and colored as follows: His20, yellow; Asp85, red; Arg202, blue; Arg292, blue. The image was drawn using PyMOL (http://www.pymol.org/).

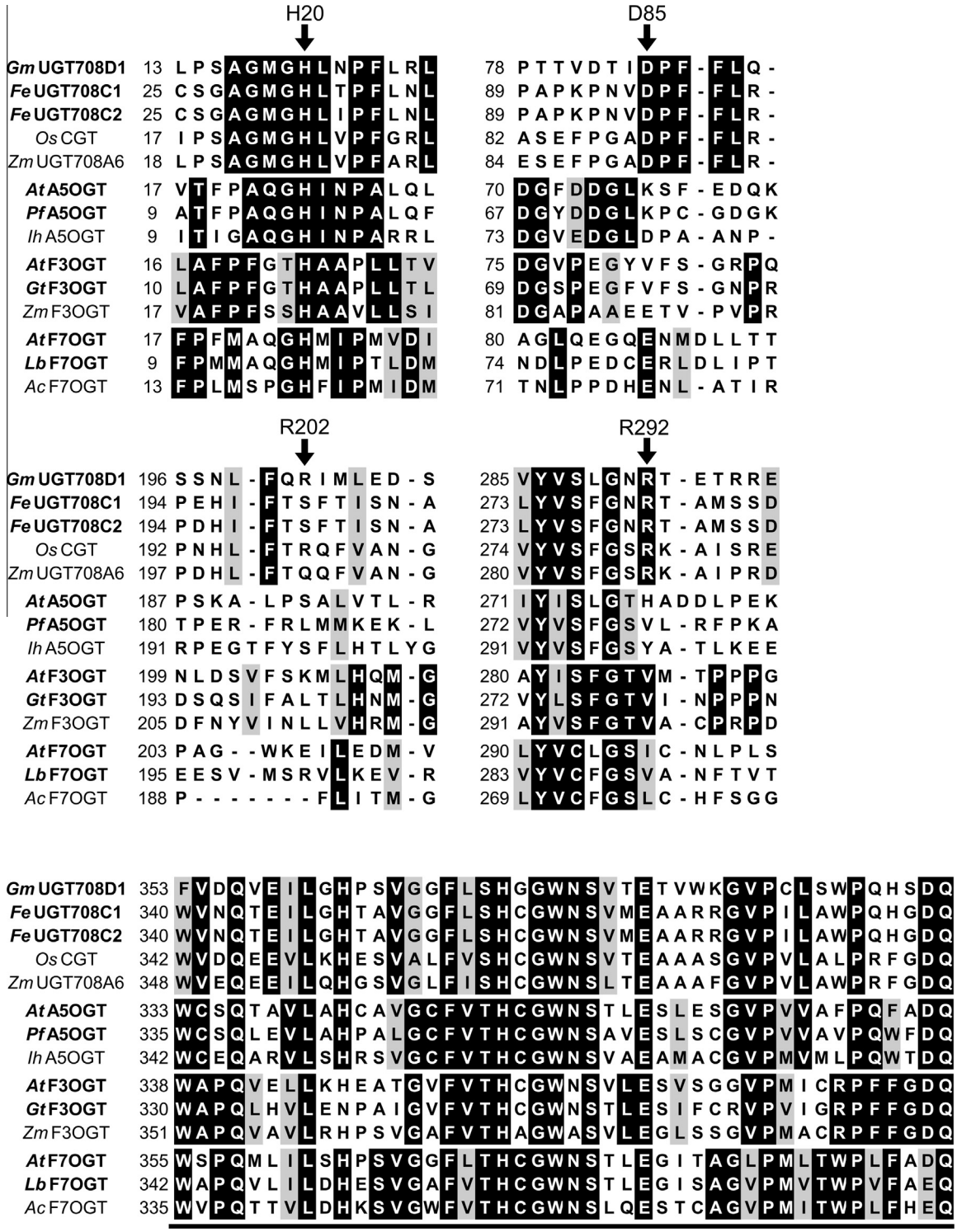

Fig. 4. Parts of a multiple sequence alignment of flavonoid CGTs and OGTs. Multiple sequence alignment was performed with MUSCLE [41]. Completely conserved sites among UGTs that share the same activity were shaded black (identical residues) or gray (similar residues) with BoxShade 3.31. Gene names of dicot UGTs are indicated in boldface. The sites of His20, Asp85, Arg202, and Arg292 of UGT708D1 are indicated by arrows. The conserved PSPG motif is indicated by a solid line below the alignment. The following abbreviations are used: F3OGT, flavonoid 3-*O*-glucosyltransferase; F5OGT, flavonoid 5-*O*-glucosyltransferase; F7OGT, flavonoid 7-*O*-glucosyltransferase; UGT708A6, UGT708C1, and UGT708C2, 2-hydroxyflavanone *C*-glucosyltransferase. Species abbreviations are as follows: Ac, *Allium cepa* (onion); At, *Arabidopsis thaliana* (thale cress); Fe, *Fagopyrum esculentum* (buckwheat); Gm, *Glycine max* (soybean); Gt, *Gentiana triflora* (*Ezorindo*); Ih, *Iris hollandica* (Dutch iris); Lb, *Lycium barbarum* (Chinese wolfberry); Pf, *Perilla frutescens* (beefsteak mint); Zm, *Zea mays* (maize). Accession numbers are shown in Supplementary Table S5.

5), which was not detected in the reaction product of the wild-type enzyme, was observed (Fig. 5A). When *peak 5* of Fig. 5A was analyzed by MS, an ion at *m*/*z* 449.109 was observed, indicating that *peak 5* was a monoglucoside of 2-hydroxynaringenin. The MS/MS spectrum of the parent ion at *m*/*z* 449 indicated the presence of a fragment ion of $[M-H-162]^-$ at *m*/*z* 287 (Fig. 5A, inset) arising from the neutral loss of a whole glucose moiety (Supplementary Fig. S7), suggesting that the glucosyl moiety was conjugated to

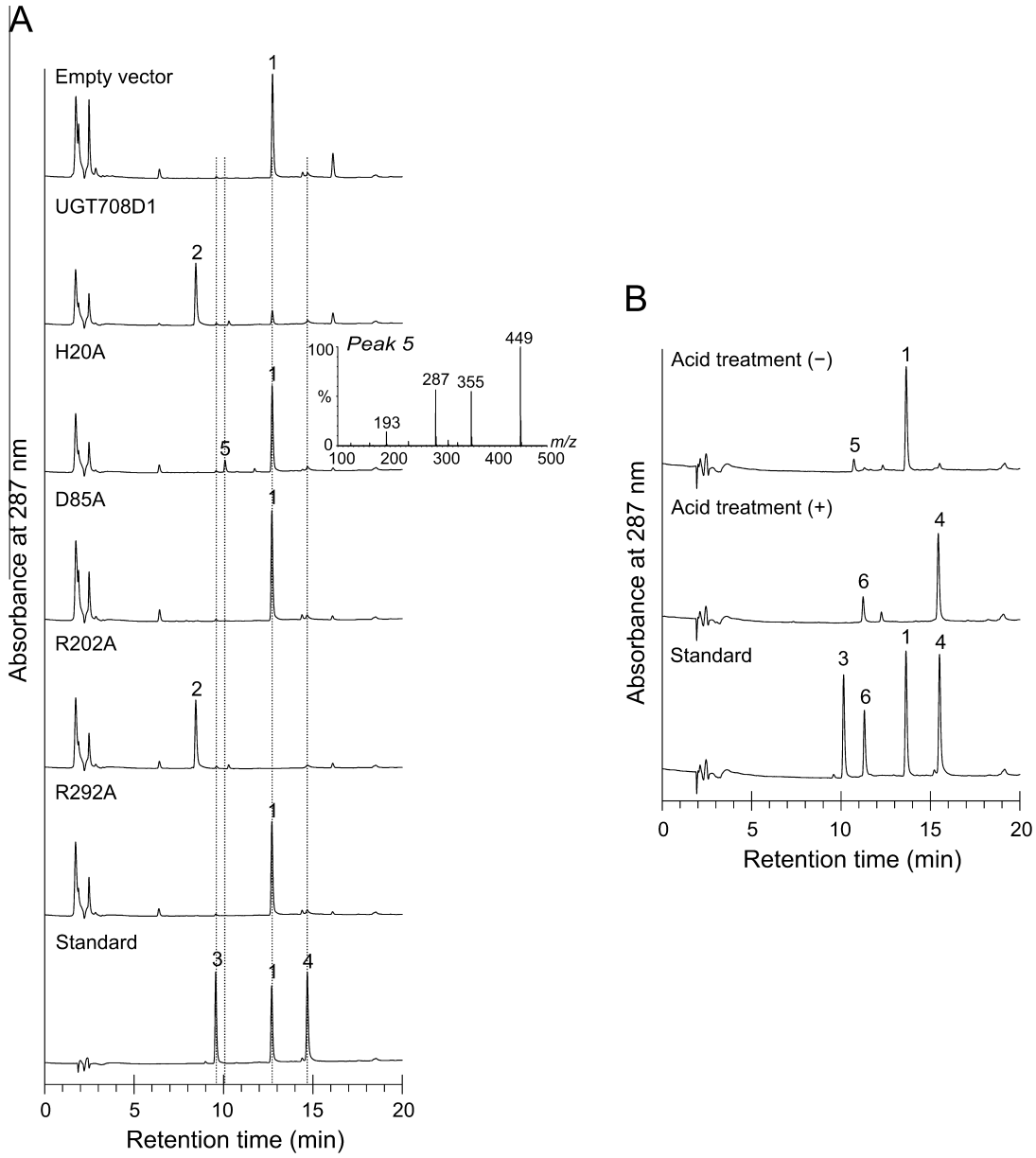

Fig. 5. (A) HPLC analyses of mutant UGT708D1 activity toward 2-hydroxynaringenin. Enzymatic activities of wild-type and mutant proteins (H20A, D85A, R202A, and R292A) were assayed in the presence of 2-hydroxynaringenin and UDP-glucose using the crude enzymes and the reaction products were analyzed by HPLC. A crude solution prepared from empty vector (pET-15b) transformant was used as negative control. The MS/MS spectrum of the parent ion at *m*/*z* 449 from *peak 5* is shown in inset. Reference standards are as follows: 1, 2-hydroxynaringenin; 3, vitexin; 4, apigenin. (B) The HPLC profile of the reaction product of H20A mutant before and after acid treatment. To infer the attached position of the glucosyl moiety, the reaction product of the H20A mutant (top panel) was treated with acid to facilitate dehydration (second panel). Reference standards are as follows: 1, 2-hydroxynaringenin; 3, vitexin; 4, apigenin; 6, apigenin 7-*O*-glucoside.

2-hydroxynaringenin by an *O*-glucoside linkage [30]. When the reaction product of the H20A mutant was treated with acid for product identification, two compounds, *peak 6* and *peak 4*, were generated with disappearance of *peak 5* and *peak 1* (Fig. 5B). The retention times of *peak 6* and *peak 4* coincided with those of apigenin 7-*O*-glucoside and apigenin, respectively (Fig. 5B). Considering that the dehydration of 2-hydroxynaringenin 7-*O*-glucoside and 2-hydroxynaringenin produced apigenin 7-*O*-glucoside and apigenin, respectively, the glucose moiety of *peak 5* was inferred to be attached to the 7-OH position of 2-hydroxynaringenin. Thus, H20A mutant protein lost CGT activity but retained low OGT activity. These results indicated that Asp85 and Arg292 of UGT708D1 were critical for enzymatic activity, whereas His20 of UGT708D controlled enzymatic regioselectivity.

3.6. Phylogenetic analysis of UGT708D1

To establish an evolutionary relationship of UGT708D1 with other flavonoid UGTs, a phylogenetic tree was constructed (Fig. 6). The tree was linearized assuming equal evolutionary rates in all lineages to compare the relative gene divergence times of the interior nodes. The tree showed that four gene divergences (Fig. 6, nodes a–d) occurred almost simultaneously between the monocot and dicot flavonoid UGTs that share the same regioselectivity. The

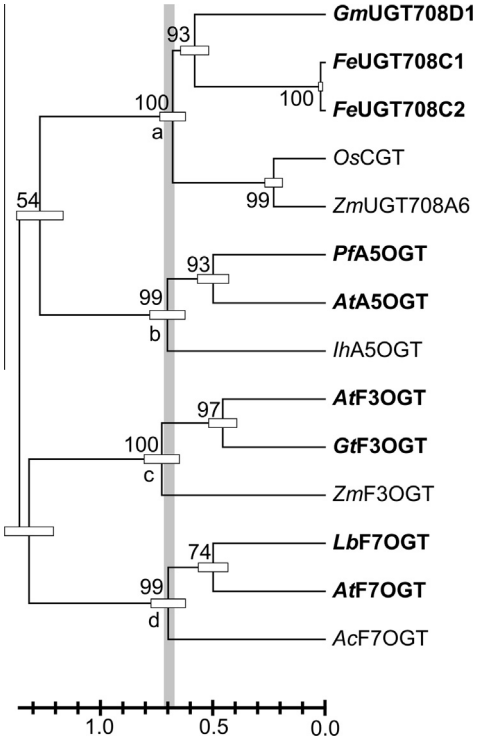

Fig. 6. Linearized phylogenetic tree of UGT708D1 and flavonoid glucosyltransferases. The evolutionary history was inferred using the maximum-likelihood method with the molecular clock. Bootstrap values (500 replicates) are indicated next to the branches. Confidence intervals (95%) are shown as white bars. The tree is drawn to scale, with branch lengths measured in the relative numbers of substitutions per site. Gene names of dicot UGTs are indicated in boldface. The phylogenetic analysis was performed with MEGA5 [42]. Abbreviations are shown in the legend of Fig. 4. Accession numbers are shown in Supplementary Table S5.

result suggests that plants acquired the ability to *C*-glucosylate 2-hydroxyflavanone before the monocot–dicot species split.

4. Discussion

In silico cloning using the protein sequence of a rice CGT (*Os*CGT) as a query led to isolation of a soybean CGT (UGT708D1), which was unambiguously shown to catalyze *C*-glucosylation of 2-hydroxynaringenin.

The enzymatic mechanism of plant CGTs remains to be elucidated. Previous study of a bacterial CGT, which *C*-olivosylates an aromatic polyketide, has suggested that an aspartate residue located at its active center is essential for CGT activity because the negative charge of this residue can activate the aromatic carbon of the acceptor substrate toward EAS reaction by resonance [31,32]. However, no key residue functionally equivalent to that in the bacterial CGT has been identified in plant CGTs. The present study suggests that Asp85 of UGT708D1 plays a role similar to that of the anionic residue in bacterial CGTs. The results of our study support this idea.

Although the highly conserved N-terminal histidine has been suggested to serve as a base to deprotonate the hydroxyl group of the acceptor substrate in OGT [15], the role of the corresponding residue in CGT is uncertain. Gutmann and Nidetzky have proposed that the N-terminal histidine of *Os*CGT functions to deprotonate the phenolic hydroxyl group of the aromatic substrate, leading to activation of the aromatic carbon of the substrate by resonance [33]; however, the histidine residue may be insufficient for activating the aromatic ring compared with the anionic residue. The present study suggests a role for the N-terminal histidine that is different from the one they proposed. Because the aromatic proton of the acceptor substrate is displaced by a glucosyl moiety as a consequence of the aryl-*C*-glucosylation reaction, efficient removal of the aromatic proton is expected to be required for catalysis. We accordingly propose that His20 of UGT708D1 functions as a base to accept the aromatic proton in a way similar to that in which the corresponding residue in OGT acts to accept the hydroxyl proton. Based on the observation that the N-terminal histidine is conserved among both CGTs and OGTs (Fig. 4), it appears that the histidine residue plays a common role in accepting the hydrogen atom that is substituted by the glucosyl moiety, irrespective of the difference between *C*- and *O*-glucosylation reactions. In a recently discovered flavone CGT [10], the corresponding residue was replaced by valine (data not shown), but the enzyme retained *C*-glucosylation activity. The molecular function of the N-terminus histidine may differ between 2-hydroxyflavanone CGT and flavone CGT.

The mechanism by which the H20A mutant exerts *O*-glucosylation activity remains elusive. Replacing the conserved N-terminal histidine with alanine usually abolishes OGT activity [14–18] although there is an exceptional case of soybean isoflavone OGT (*Gm*IF7GT) [34] in which the N-terminal histidine is dispensable for OGT activity. It is unknown whether the OGT activity of the H20A mutant may be relevant to the case of *Gm*IF7GT. A possible explanation is that appropriate positioning of the acceptor substrate and the glucose donor for *C*-glucoside formation is impaired in this mutant, resulting in abnormal *O*-glucoside formation, as proposed previously [19,32]. Alternatively, enzymatic *C*-glucosylation may proceed via the initial formation of an *O*-glucosyl intermediate, followed by an *O*- to *C*-glucosyl rearrangement reaction [35]. In the latter case, His20 of UGT708D1

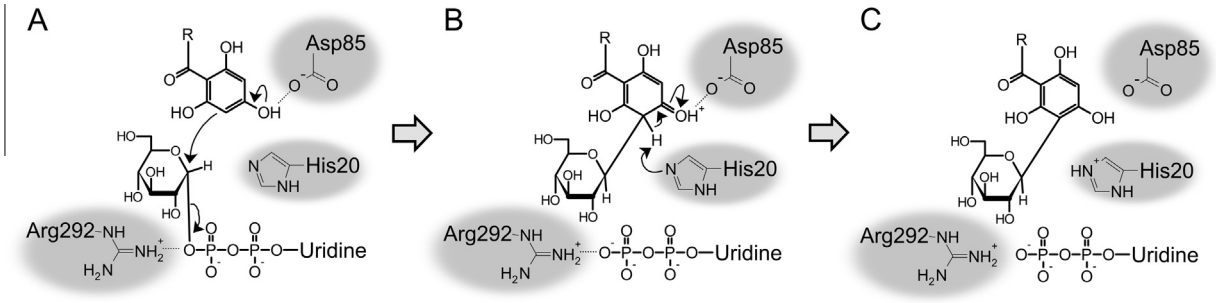

Fig. 7. A possible catalytic model for UGT708D1. (A) The negative charge of the carboxyl group of Asp85 is expected to enhance the nucleophilicity of the aromatic carbon *ortho* to the hydroxyl group of the acceptor substrate through resonance. The positive charge of the guanidino group of Arg292 may enhance the leaving-group ability of the UDP moiety of UDP-glucose. (B and C) After electrophilic attack on the aromatic carbon by the anomeric carbon of UDP-glucose to form the arenium ion (a Wheland intermediate), His20 may act as a catalytic base to deprotonate the cationic non-aromatic intermediate to yield the aryl-*C*-glucoside.

may facilitate the *O*–*C*-glucosyl rearrangement step, given that a mutant CGT enzyme that has lost the *O*–*C* rearrangement activity is expected to display only OGT activity.

Following the assumption that the *C*-glucosylation reaction proceeds through an EAS, we propose an enzymatic reaction model for UGT708D1 (Fig. 7). In the first step, the negatively charged Asp85 activates the aromatic carbon of the acceptor substrate toward EAS (Fig. 7A). After the formation of the arenium ion as an EAS reaction intermediate (Fig. 7B), His20 functions as a base to accept a proton from the cationic non-aromatic intermediate to restore aromaticity (Fig. 7C). Given that Arg292 is located near the UDP-glucose molecule in the structural model (Fig. 3), the positive charge of Arg292 may increase the leaving-group ability of the UDP moiety (Fig. 7A). Thus, in this model, the three conserved residues located at the putative active site of UGT708D1 act cooperatively to facilitate enzymatic *C*-glucosylation.

One of the possible applications of the present study would be the molecular cloning of CGTs using a homology-based method [36,37] from other *C*-glucoside-producing plants whose genomes have not been sequenced, as exemplified by the legume species *Pueraria lobata*, which produces puerarin, an isoflavone 8-*C*-glucoside. *C*-Glucosylation activity, which catalyzes the conversion of isoliquiritigenin (the precursor of puerarin) to puerarin in the presence of UDP-glucose, has been reported in crude extracts of the roots of *P. lobata*, but the gene responsible for the *C*-glucosylation remains unknown [38–40]. Curiously, there have been no reports of *C*-glucosylflavone metabolites in this genus, raising the possibility that a 2-hydroxyflavanone CGT has evolved into an isoflavone CGT. Thus, the unidentified CGT may show high sequence similarity to legume UGT708D1.

In conclusion, the present study not only identified a 2-hydroxyflavanone CGT in soybean but also proposed key residues for CGT activity. This study will provide basic information for understanding the catalytic mechanism of plant CGTs.

Acknowledgments

The authors thank the Center for Medical Research and Education (Osaka University) for HPLC analysis. The authors thank Dr. Motomasa Kobayashi (Osaka University) for supporting the synthesis of 2-hydroxynaringenin, Dr. Antony M. Hooper (Rothamsted Research) for providing the ^{1}H NMR spectrum of 2-hydroxynaringenin, and the anonymous reviewers for their insightful comments. This study was supported by the Graduate School of Frontier Biosciences, Osaka University, and, in part, by JSPS KAKENHI Grant Number 25460127 (to HM).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2015.05.010.

References

[1] Jay, M. (1993) *C*-Glycosylflavonoids in: The Flavonoids: Advances in Research Since 1986 (Harborne, J.B., Ed.), pp. 57–93, Chapman & Hall, London.
[2] Jay, M., Viricel, M.R. and Gonnet, J. (2005) *C*-Glycosylflavonoids in: Flavonoids: Chemistry, Biochemistry and Applications (Anderson, Ø.M. and Markham, K.R., Eds.), pp. 857–915, CRC Press, Boca Raton, FL.
[3] Talhi, O. and Silva, A.S. (2012) Advances in *C*-glycosylflavonoid Research. Curr. Org. Chem. 16, 859–896.
[4] Kim, J.H., Lee, B.C., Kim, J.H., Sim, G.S., Lee, D.H., Lee, K.E., Yun, Y.P. and Pyo, H.B. (2005) The isolation and antioxidative effects of vitexin from *Acer palmatum*. Arch. Pharm. Res. 28, 195–202.
[5] Folador, P., Cazarolli, L.H., Gazola, A.C., Reginatto, F.H., Schenkel, E.P. and Silva, F.B. (2010) Potential insulin secretagogue effects of isovitexin and swertisin isolated from *Wilbrandia ebracteata* roots in non-diabetic rats. Fitoterapia 81, 1180–1187.
[6] Nagaprashantha, L.D., Vatsyayan, R., Singhal, J., Fast, S., Roby, R., Awasthi, S. and Singhal, S.S. (2011) Anti-cancer effects of novel flavonoid vicenin-2 as a single agent and in synergistic combination with docetaxel in prostate cancer. Biochem. Pharmacol. 82, 1100–1109.
[7] Brazier-Hicks, M., Evans, K.M., Gershater, M.C., Puschmann, H., Steel, P.G. and Edwards, R. (2009) The *C*-glycosylation of flavonoids in cereals. J. Biol. Chem. 284, 17926–17934.
[8] Falcone Ferreyra, M.L., Rodriguez, E., Casas, M.I., Labadie, G., Grotewold, E. and Casati, P. (2013) Identification of a bifunctional maize *C*- and *O*-glucosyltransferase. J. Biol. Chem. 288, 31678–31688.
[9] Nagatomo, Y., Usui, S., Ito, T., Kato, A., Shimosaka, M. and Taguchi, G. (2014) Purification, molecular cloning, and functional characterization of flavonoid *C*-glucosyltransferases from buckwheat (*Fagopyrum esculentum*) cotyledon. Plant J. 80, 437–448.
[10] Sasaki, N., Nishizaki, Y., Yamada, E., Tatsuzawa, F., Nakatsuka, T., Takahashi, H. and Nishihara, M. (2015) Identification of the glucosyltransferase that mediates direct flavone *C*-glucosylation in *Gentiana triflora*. FEBS Lett. 589, 182–187.
[11] Bowles, D., Isayenkova, J., Lim, E.K. and Poppenberger, B. (2005) Glycosyltransferases: managers of small molecules. Curr. Opin. Plant Biol. 8, 254–263.
[12] Gachon, C.M., Langlois-Meurinne, M. and Saindrenan, P. (2005) Plant secondary metabolism glycosyltransferases: the emerging functional analysis. Trends Plant Sci. 10, 542–549.
[13] Grubb, C.D., Zipp, B.J., Ludwig-Muller, J., Masuno, M.N., Molinski, T.F. and Abel, S. (2004) Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis. Plant J. 40, 893–908.
[14] Brazier-Hicks, M., Offen, W.A., Gershater, M.C., Revett, T.J., Lim, E.K., Bowles, D.J., Davies, G.J. and Edwards, R. (2007) Characterization and engineering of the bifunctional *N*- and *O*-glucosyltransferase involved in xenobiotic metabolism in plants. Proc. Natl. Acad. Sci. U.S.A. 104, 20238–20243.
[15] Shao, H., He, X., Achnine, L., Blount, J.W., Dixon, R.A. and Wang, X. (2005) Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. Plant Cell 17, 3141–3154.
[16] Offen, W., Martinez-Fleites, C., Yang, M., Kiat Lim, E., Davis, B.G., Tarling, C.A., Ford, C.M., Bowles, D.J. and Davies, G.J. (2006) Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. EMBO J. 25, 1396–1405.
[17] Li, L., Modolo, L.V., Escamilla-Trevino, L.L., Achnine, L., Dixon, R.A. and Wang, X. (2007) Crystal structure of *Medicago truncatula* UGT85H2 - insights into the structural basis of a multifunctional (iso)flavonoid glycosyltransferase. J. Mol. Biol. 370, 951–963.
[18] Modolo, L.V., Li, L., Pan, H., Blount, J.W., Dixon, R.A. and Wang, X. (2009) Crystal structures of glycosyltransferase UGT78G1 reveal the molecular basis for glycosylation and deglycosylation of (iso)flavonoids. J. Mol. Biol. 392, 1292–1302.
[19] Gutmann, A. and Nidetzky, B. (2012) Switching between *O*- and *C*-glycosyltransferase through exchange of active-site motifs. Angew. Chem. Int. Ed. Engl. 51, 12879–12883.
[20] Hamilton, M.L., Caulfield, J.C., Pickett, J.A. and Hooper, A.M. (2009) *C*-Glucosylflavonoid biosynthesis from 2-hydroxynaringenin by *Desmodium uncinatum* (Jacq.) (Fabaceae). Tetrahedron Lett. 50, 5656–5659.
[21] Umezawa, T., Sakurai, T., Totoki, Y., Toyoda, A., Seki, M., Ishiwata, A., Akiyama, K., Kurotani, A., Yoshida, T., Mochida, K., Kasuga, M., Todaka, D., Maruyama, K., Nakashima, K., Enju, A., Mizukado, S., Ahmed, S., Yoshiwara, K., Harada, K., Tsubokura, Y., Hayashi, M., Sato, S., Anai, T., Ishimoto, M., Funatsuki, H., Teraishi, M., Osaki, M., Shinano, T., Akashi, R., Sakaki, Y., Yamaguchi-Shinozaki, K. and Shinozaki, K. (2008) Sequencing and analysis of approximately 40,000 soybean cDNA clones from a full-length-enriched cDNA library. DNA Res. 15, 333–346.
[22] Sali, A. and Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779–815.
[23] Smith, T.F. and Waterman, M.S. (1981) Identification of common molecular subsequences. J. Mol. Biol. 147, 195–197.
[24] Pearson, W.R. (1991) Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. Genomics 11, 635–650.
[25] Jay, M., Lameta-D'Arcy, A. and Viricel, M.R. (1984) *C*-Glycosylflavones from roots of *glycine max*. Phytochemistry 23, 1153–1155.
[26] Vogt, T. and Jones, P. (2000) Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. Trends Plant Sci. 5, 380–386.
[27] Li, Y., Baldauf, S., Lim, E.K. and Bowles, D.J. (2001) Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. J. Biol. Chem. 276, 4338–4343.
[28] Yonekura-Sakakibara, K. and Hanada, K. (2011) An evolutionary view of functional diversity in family 1 glycosyltransferases. Plant J. 66, 182–193.
[29] Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C. and Jackson, S.A. (2010) Genome sequence of the palaeopolyploid soybean. Nature 463, 178–183.

[30] Kazuno, S., Yanagida, M., Shindo, N. and Murayama, K. (2005) Mass spectrometric identification and quantification of glycosyl flavonoids, including dihydrochalcones with neutral loss scan mode. Anal. Biochem. 347, 182–192.
[31] Mittler, M., Bechthold, A. and Schulz, G.E. (2007) Structure and action of the C–C bond-forming glycosyltransferase UrdGT2 involved in the biosynthesis of the antibiotic urdamycin. J. Mol. Biol. 372, 67–76.
[32] Harle, J., Gunther, S., Lauinger, B., Weber, M., Kammerer, B., Zechel, D.L., Luzhetskyy, A. and Bechthold, A. (2011) Rational design of an aryl-C-glycoside catalyst from a natural product O-glycosyltransferase. Chem. Biol. 18, 520–530.
[33] Gutmann, A. and Nidetzky, B. (2013) Enzymatic *C*-glycosylation: Insights from the study of a complementary pair of plant *O*- and *C*-glucosyltransferases. Pure Appl. Chem. 85, 1865–1877.
[34] Noguchi, A., Saito, A., Homma, Y., Nakao, M., Sasaki, N., Nishino, T., Takahashi, S. and Nakayama, T. (2007) A UDP-glucose:isoflavone 7-*O*-glucosyltransferase from the roots of soybean (glycine max) seedlings. Purification, gene cloning, phylogenetics, and an implication for an alternative strategy of enzyme catalysis. J. Biol. Chem. 282, 23581–23590.
[35] Bililign, T., Hyun, C., Williams, J.S., Czisny, A.M. and Thorson, J.S. (2004) The hedamycin locus implicates a novel aromatic PKS priming mechanism. Chem. Biol. 11, 959–969.
[36] Nagatoshi, M., Terasaka, K., Nagatsu, A. and Mizukami, H. (2011) Iridoid-specific glucosyltransferase from *Gardenia jasminoides*. J. Biol. Chem. 286, 32866–32874.
[37] Asada, K., Salim, V., Masada-Atsumi, S., Edmunds, E., Nagatoshi, M., Terasaka, K., Mizukami, H. and De Luca, V. (2013) A 7-deoxyloganetic acid glucosyltransferase contributes a key step in secologanin biosynthesis in Madagascar periwinkle. Plant Cell 25, 4123–4134.
[38] Chen, G., Wu, X., Zhou, W. and Li, L. (2010) Preparation and assay of C-glucosyltransferase from roots of *Pueraria lobata*. J. Environ. Biol. 31, 655–660.
[39] Wenling, Z., Yinghua, W., Gang, C., Min, L., Lixia, Y., Xiaoxiang, H., Haihang, L. and Ling, L. (2011) Cloning of a gene encoding glycosyltransferase from *Pueraria lobata* (Wild.) Ohwi and its expression in *Pichia pastoris*. Afr. J. Biotechnol. 10, 85–96.
[40] He, X., Blount, J.W., Ge, S., Tang, Y. and Dixon, R.A. (2011) A genomic approach to isoflavone biosynthesis in kudzu (*Pueraria lobata*). Planta 233, 843–855.
[41] Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.
[42] Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.