FEBS Letters 584 (2010) 2258–2264

FEBS Letters

journal homepage: www.FEBSLetters.org

Identification and characterization of glycosyltransferases involved in the biosynthesis of soyasaponin I in *Glycine max*

Masaaki Shibuya [a,b], Kazuya Nishimura [a], Nao Yasuyama [a], Yutaka Ebizuka [a,*]

[a] Graduate School of Pharmaceutical Sciences, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[b] Faculty of Pharmaceutical Sciences, Niigata University of Pharmacy and Applied Life Sciences, Higashijima, Akiha-ku, Niigata 956-8603, Japan

ARTICLE INFO

Article history:
Received 26 January 2010
Revised 7 March 2010
Accepted 23 March 2010
Available online 27 March 2010

Edited by Judit Ovádi

Keywords:
Glycosyltransferase
Soyasaponin I
Glycine max

ABSTRACT

Triterpene saponins are a diverse group of compounds with a structure consisting of a triterpene aglycone and sugars. Identification of the sugar-transferase involved in triterpene saponin biosynthesis is difficult due to the structural complexity of triterpene saponin. Two glycosyltransferases from *Glycine max*, designated as GmSGT2 and GmSGT3, were identified and characterized. In vitro analysis revealed that GmSGT2 transfers a galactosyl group from UDP-galactose to soyasapogenol B monoglucuronide, and that GmSGT3 transfers a rhamnosyl group from UDP-rhamnose to soyasaponin III. These results suggest that soyasaponin I is biosynthesized from soyasapogenol B by successive sugar transfer reactions.

1. Introduction

Triterpene saponins are a diverse group of compounds with a structure consisting of a triterpene aglycone and one or more sugar moieties linked through an ether or ester glycosidic linkage at one or two sites. Triterpene saponins are widely distributed in higher plants [1,2]. Legumes such as soybean (*Glycine max*) [3], licorice (*Glycyrrhiza glabra*) [4], and alfalfa (*Medicago sativa*) [5], as well as some medicinal plants like ginseng (*Panax ginseng*) [6,7] are well known as saponin-producing plants. Triterpene saponins have diverse biological actions including antiinflammatory, antifungal, antimicrobial, antiparasitic, and antitumor activities [2], which attract many researchers from a pharmaceutical point of view.

In addition to diverse biological activities, triterpene saponins have the structural diversity. This diversity is generated by a sequence of three biosynthetic reactions. The first is the cyclization of one common substrate, oxidosqualene, which generates structural diversity of the skeleton and eventually results in more than 100 variations [8,9]. Further structural diversity is generated by oxidation and glycosylation. Among these three reactions, glycosylation contributes far more to the generation of structural diversity than does cyclization or oxidation, as far as various saponins in a particular plant are concerned. For example, although there are only two genuine aglycones of ginsenosides, panaxadiol and panaxatriol in *P. ginseng*, the ensuing glycosylation reactions produce more than 25 ginsenosides with various kinds of sugar moieties [6].

Biosynthetic studies on oxidation and glycosylation in triterpene saponin biosynthesis have lagged behind those on cyclization due to the inherent difficulties of cloning oxidase and glycosyltransferase (GT) caused by the structural diversity of triterpene saponins. Generally, aglycones include multiple hydroxyl (and/or carboxyl) groups, and triterpene saponins include multiple sugars at multiple positions. This suggests that many possible biosynthetic intermediates to triterpene saponins are generated. Preparation of all possible substrates is difficult and makes an assay against them practically impossible, although recent developments of expressed sequence tags (EST) and genome sequence databases have made it easy to select candidates for target hydroxylases and GTs. Thus far, only two triterpene hydroxylases and three GTs have been identified. The former are CYP93E1 from *G. max* with dual (β-amyrin-24-hydroxylase and soforadiol-24-hydroxylase) activities [10], and CYP88D6 from *G. glabra* with β-amyrin-11-oxidase activity [11]. The latter are UGT71G1 and UGT73K1 from *Medicago truncatula* [12], and UGT74M1 from *Saponaria vaccaria* [13].

UGT71G1 showed dual glucosyltransferase activity against hederagenin and some flavonoids, and UGT73K1 showed dual glucosyltransferase activity against hederagenin and soyasapogenol B [12]. As *M. truncatula* is estimated to include more than 27 triterpene saponins with five kinds of aglycones and a variety of sugar moieties [5], more than 27 GTs potentially exist in this plant.

Abbreviations: SMBG, soyasapogenol B monoglucuronide; GT, glycosyltransferase

* Corresponding author. Fax: +81 3 5841 4744.
E-mail address: yebiz@mol.f.u-tokyo.ac.jp (Y. Ebizuka).

0014-5793/$36.00
doi:10.1016/j.febslet.2010.03.037

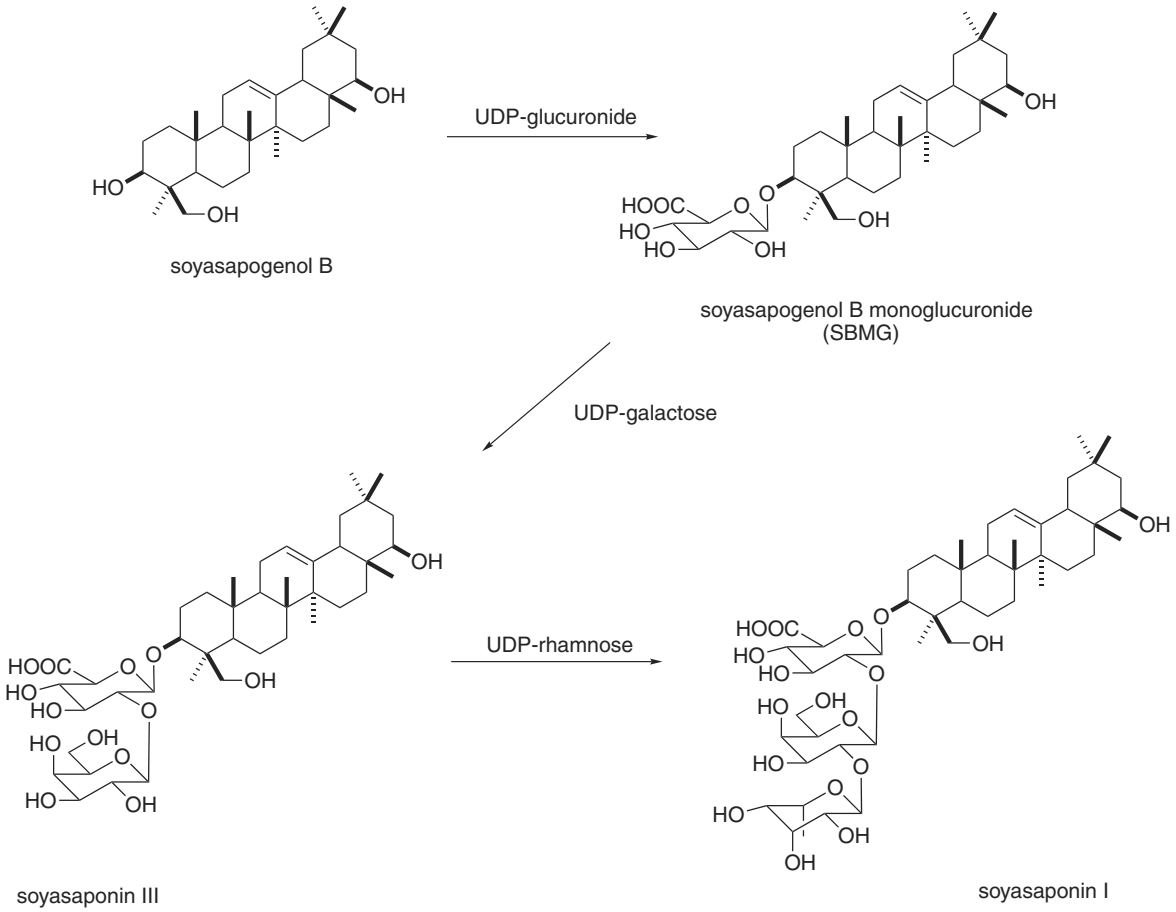

Fig. 1. Formation of soyasaponin I by successive sugar transfers from UDP-sugars to soyasapogenol B.

However, none except for the two glucosyltransferases, UGT71G1 and UGT73K1, were identified since all the other possible substrates necessary for characterizing the function of the remaining candidates were difficult to prepare.

UGT74M1 was mined from the EST library of *S. vaccaria*, which contains vaccarosides [13]. UGT74M1 is an ester-forming GT and is included in a group consisting of other known ester-forming GTs in phylogenetic analysis [13]. The *M. truncatula* triterpene-GTs (UGT71G1 and UGT73K1) were not included in this group. At the present time, phylogenetic analyses cannot reveal whether triterpene-GTs with similar functions share amino acid sequence.

In order to shed light on the nature of triterpene-GT, we have to clone a lot more triterepene-GTs. In this study, we selected *G. max* as a material. *G. max* predominantly produces soyasaponin I [3], which drastically reduces the number of required assays. Soyasaponin I is supposedly biosynthesized by the successive additions of glucuronic acid, galactose, and rhamnose to soyasapogenol B (Fig. 1). Here, using EST information from both *G. max* and *M. truncatula*, we report the successful identification of UDP-galactose: soyasapogenol B monoglucuronide (SBMG)-galactosyltransferase (GmSGT2) and UDP-rhamnose:soyasaponin III-rhamnosyltransferase (GmSGT3) involved in the biosynthesis of soyasaponin I in *G. max*.

2. Materials and methods

2.1. Plant materials and chemicals

Soybean seeds were purchased from a local grocery store. Soyasaponin I was isolated from soybean seeds [3], and SBMG and soyasaponin III were prepared by chemical hydrolysis [14]. Other analytical-grade chemicals were purchased commercially. Synthetic oligonucleotides for PCR primers were obtained from Nihon Bioservice (Saitama, Japan).

2.2. RNA and cDNA preparation

Soybean seeds were germinated in a growth cabinet under dark conditions at 25 °C. After 5 days, the seedlings (4 g) were immediately frozen in liquid nitrogen and homogenized with a mortar and pestle. RNA-extraction was performed with the phenol–chloroform method as reported previously [10] to yield total RNA. A single-stranded cDNA pool was prepared using SuperScript III (Invitrogen, California, USA) with 50 μM of RACE32 (5′-gactcgagtcgacatcgatttttttttttttt-3′) primer according to the manufacturer's protocol.

2.3. Expression of GTXs in Escherichia coli and preparation of cell-free extracts

BL21-CodonPlus(DE3)-RIPL (Novagen) harboring pET21a-GTXs or void vector (pET21a, Novagen) was cultured in LB medium containing 50 μg/ml of ampicillin, by shaking at 220 rpm overnight at 37 °C. The same media (20 ml) was inoculated with a portion (200 μl) of the overnight culture under the same conditions for 2 h. IPTG (final concentration 0.5 mM) was added to the culture, and incubation was continued at 20 °C for 16 h. Cells were collected by centrifugation, suspended in 1 ml of sodium phosphate buffer (50 mM, pH 7.4) containing 300 mM sodium chloride and 20% glycerol, and disrupted three times for 10 s each by sonication.

The supernatant was recovered by centrifugation and used in the screening assay as the cell-free extract.

2.4. Screening of UDP-glucuronic acid:soyasapogenol B-glucuronosyltransferase from GTXs

Each cell extract (110 μl, see Section 2.3) was incubated with [U-^{14}C]-UDP-glucuronic acid (2.22 kBq, 50 μM, PerkinElmer), soyasapogenol B (50 μM), and magnesium chloride (10 mM) in a total volume of 120 μl at 40 °C for 3 h. The reaction was stopped by the addition of diluted hydrochloric acid (1 M, 10 μl), and the product was extracted with butanol (90 μl). The butanol extract was spotted on a silica-gel TLC plate (Merck Art.11798), which was then developed with the solvent (chloroform:methanol:acetic acid:water = 14:6:1:1). The radio-labeled product was visualized with a BAS-1500 Image Reader V. 1.8 (Fujifilm) and analyzed using Image Gauge (Fujifilm).

2.5. Screening of UDP-galactose:SBMG-galactosyltransferase from GTXs

Each cell extract (110 μl, see Section 2.3) was incubated with [U-^{14}C]-UDP-galactose (2.22 kBq, 50 μM, PerkinElmer) and SBMG (50 μM) in a total volume of 120 μl at 40 °C for 3 h. The procedures for product analysis were the same as described in Section 2.4. Result of *GTX-2* was shown in Fig. 2. DDBJ accession number of *GTX-2* (GmSGT2, UGT73P2) is AB473730.

2.6. Sugar donor specificity of GTX-2

We changed the *E. coli* host from BL21-CodonPlus(DE3)-RIPL to Arctic Express (Stratagene). The transformant harboring pET21a-GT2 was cultured in LB medium containing 50 μg/ml of ampicillin and 20 μg/ml of gentamycin by shaking at 220 rpm overnight at 37 °C. LB media (40 ml) was inoculated with a portion (400 μl) of the overnight culture in the absence of antibiotics and incubated at 30 °C for 3 h (220 rpm). IPTG (final concentration 1 mM) was added to the culture, and incubation was continued by shaking at 220 rpm for 24 h at 12 °C. Cells were collected by centrifugation, suspended in 1 ml of sodium phosphate buffer (50 mM, pH 7.4) containing 300 mM sodium chloride and 20% glycerol, and disrupted three times for 10 s each by sonication. The supernatant was recovered by centrifugation and used as the cell-free extract. This extract (110 μl) and SBMG (100 μM) were incubated with [U-^{14}C]-UDP-glucuronic acid (2.22 kBq, 50 μM, PerkinElmer), [U-^{14}C]-UDP-galactose (2.22 kBq, 50 μM, PerkinElmer), or [U-^{14}C]-UDP-glucose (2.22 kBq, 50 μM, PerkinElmer) in a total volume of 120 μl at 37 °C for 1h. The procedures for product analysis were the same as described in Section 2.4.

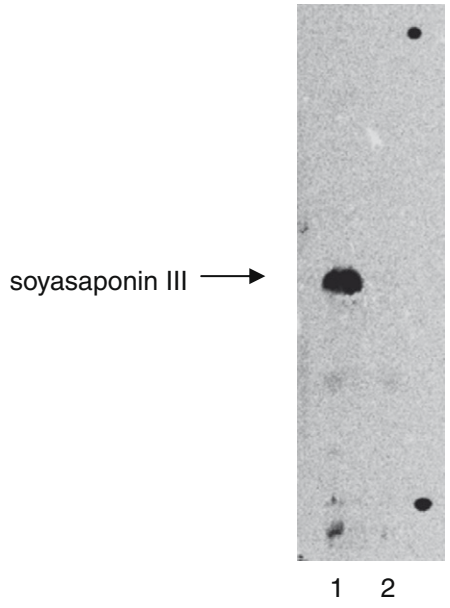

Fig. 2. TLC analysis of the GTX-2 product. Radioactivity was visualized by autoradiography. Lanes 1 and 2, respectively, represent results from *E. coli* with the plasmid harboring GTX-2 cDNA and those from *E. coli* with the plasmid without GTX-2 cDNA.

2.7. Estimation of the K_m of GTX-2

The cell extract (110 μl, see Section 2.6), and [U-^{14}C]-UDP-galactose (3.7 kBq, 200 μM, PerkinElmer), were incubated with SBMG (4.5, 6.25, 7.5, 9.0, 12.0, and 12.5 μM) in a total volume of 100 μl at 37 °C for 3 min. After extraction with butanol and separation by TLC, the radioactivity of the product was measured with a BAS-1500 Image Reader V. 1.8 (Fujifilm) and Image Gauge (Fujifilm). The K_m and V_{max} values were estimated using the Hanes–Woolf plot.

2.8. LC–MS analysis of the GTX-2 product

The cell extract (110 μl, see Section 2.6) was incubated with non-labeled UDP-galactose (100 μM) and SBMG (50 μM) in a total volume of 120 μl at 37 °C for 2 h. The reaction was stopped by the addition of diluted hydrochloric acid (1 M, 10 μl), and the product was extracted with butanol (90 μl). The butanol extract was passed though Cosmospin Filter™ G (0.2 μm, Nacalai tesque). LC–MS was measured using a Micro ToF MS (Bruker Daltonics) equipped with an HP1100 series LC system (Agilent Technologies). HPLC was carried out using a Cosmosil 5C18-MSII (ϕ 2.0 × 150 mm, Nacalai tesque), with a linear gradient of 50–70% acetonitrile containing 0.5% acetic acid during 30 min at a flow rate 0.2 ml/min at ambient temperature. The MS condition for inducing no fragmentation was as follows: ionization and polarity, ESI-negative; capillary exit voltage, 150 V; skimmer 1 voltage, 50 V. The MS condition for inducing fragmentations was as follows: ionization and polarity, ESI-positive; capillary exit voltage, 250 V; skimmer 1 voltage, 50 V.

2.9. Cloning of the UDP-rhamnose synthase RHM1 (At1g78570) from Arabidopsis thaliana

We synthesized PCR primers based on the reported sequence of the UDP-rhamnose synthase RHM1 (At1g78570) from *Arabidopsis thaliana* [15,16]. The primers were RHM-N-*BamH* I (5′-aagcgtggatccatggcttcgtacactccc-3′) and RHM-C-*Xho* I (5′-tcattcctcgagggttttcttgtttggcccgtatgc-3′), respectively. The restriction enzyme sites are underlined. The cDNA pool was prepared from the total RNA of *A. thaliana*, which was obtained from 3- to 4-week-old seedlings [17]. With these primers and the cDNA pool as a template, PCRs were carried out using Phusion High-Fidelity DNA polymerase (New England BioLabs) according to the manufacturer's protocol. The amplified full-length cDNAs were ligated to the *E. coli* expression vector pET21a, and BL21-CodonPlus(DE3)-RIPL (Novagen) was transformed with the resulting plasmid (pET21a-RHM). The cell-free extract was prepared by the same procedure as that described in Section 2.3.

2.10. Screening of UDP-rhamnose:soyasaponin III-rhamnosyltransferase from GTXs

The cell extract (100 μl, see Section 2.9) prepared from BL21-CodonPlus(DE3)-RIPL harboring pET21a-RHM was incubated with [U-^{14}C]-UDP-glucose (2.22 kBq, 50 μM, PerkinElmer) and NADPH

(2 mM) at 37 °C for 2 h. Also the boiled cell extract was incubated with the same mixture to make a negative control which was shown Fig. 5 lane 3. Each of the cell-free extracts (100 μl, see Section 2.3) derived from the GTXs and soyasaponin III (50 μM) were added to this mixture, and incubation was further continued for 2 h. The procedures for product analysis were the same as described in Section 2.4. Result of *GTX-1* was shown in Fig. 5. DDBJ accession number of *GTX-1* (GmSGT3, UGT91H4) is AB473731.

2.11. LC–MS analysis of the GTX-1 product

The cell extract (100 μl, see Section 2.9) prepared from BL21-CodonPlus(DE3)-RIPL harboring pET21a-RHM was incubated with non-labeled UDP-glucose (400 μM) and NADPH (2 mM) at 37 °C for 2 h. The cell-free extracts (100 μl) derived from GTX-1 and soyasaponin III (50 μM) were added to this mixture, and incubation was further continued for 2 h. The product was analyzed by the same procedure as that described in Section 2.8.

3. Results

3.1. Selection of GT candidates, cloning by PCR, and functional expression in E. coli

We found 111 TCs (tentative consensus clones) and 56 singleton ESTs annotated as GT or glucosyltransferase in the soybean EST database. Assuming that target genes are highly expressed and that the number of their ESTs is considerably large, we selected 32 TCs as candidates. Next, assuming that homologues of the target GT must exist in *M. truncatula* given that it also produces soyasaponin I [5], we narrowed down the candidates to 16 TCs. Finally, we obtained 8 full-length clones by PCR, and named them *GTX-1, 2, 3, 4, 5, 6, 7* and *8*. (See Supplementary data, S1.1 and S1.2, for details). Besides these candidates, we selected EV263763 and TC205339, and named these *GTX-9* and *10*, respectively. All 10 candidate clones were expressed in *E. coli* using the *E. coli* expression vector pET21a.

3.2. UDP-galactose:SBMG-galactosyltransferase—the second GT involved in soyasaponin I biosynthesis

The second GT supposedly involved in soyasaponin I biosynthesis is UDP-galactose:SBMG-galactosyltransferase which yields soyasaponin III. Using ^{14}C-labeled UDP-galactose, we examined the enzyme activity in cell-free extracts prepared from *E. coli* harboring each *GTX*. In the extract from *GTX-2*, we detected a radio-labeled spot corresponding to soyasaponin III after TLC separation and autoradiography (Fig. 2), strongly suggesting that *GTX-2* encodes UDP-galactose:SBMG-galactosyltransferase.

In order to rigorously identify the product, we carried out LC–MS analysis using non-labeled substrate. LC gave a peak at 17.9 min (Fig. 3A), with the same retention time as that of authentic soyasaponin III. The MS fragmentation pattern (Fig. 3B) of this peak was identical to that of authentic soyasaponin III (Fig. 3C), with peaks for M + Na^+ (m/z = 819) and M + H^+ (m/z = 797), as well as peaks due to successive elimination of galactose (m/z = 635) and glucuronic acid (m/z = 441) from the molecular ion. These results indicate that the product is soyasaponin III.

Next, we examined the sugar donor specificity of GTX-2. GTX-2 was incubated with ^{14}C-labeled UDP-glucose, UDP-galactose, or UDP-glucuronic acid together with SBMG. The products were analyzed by autoradiography after TLC separation. The radio-labeled product was exclusively detected in the extract with UDP-galactose (Fig. 4), indicating that GTX-2 has a high sugar donor specificity for UDP-galactose.

We also estimated the K_m value for SBMG. The enzyme reaction was achieved with SBMG at various concentrations. The K_m and V_{max} values estimated from Hanes–Woolf plot ($S - S/V$ plot) were

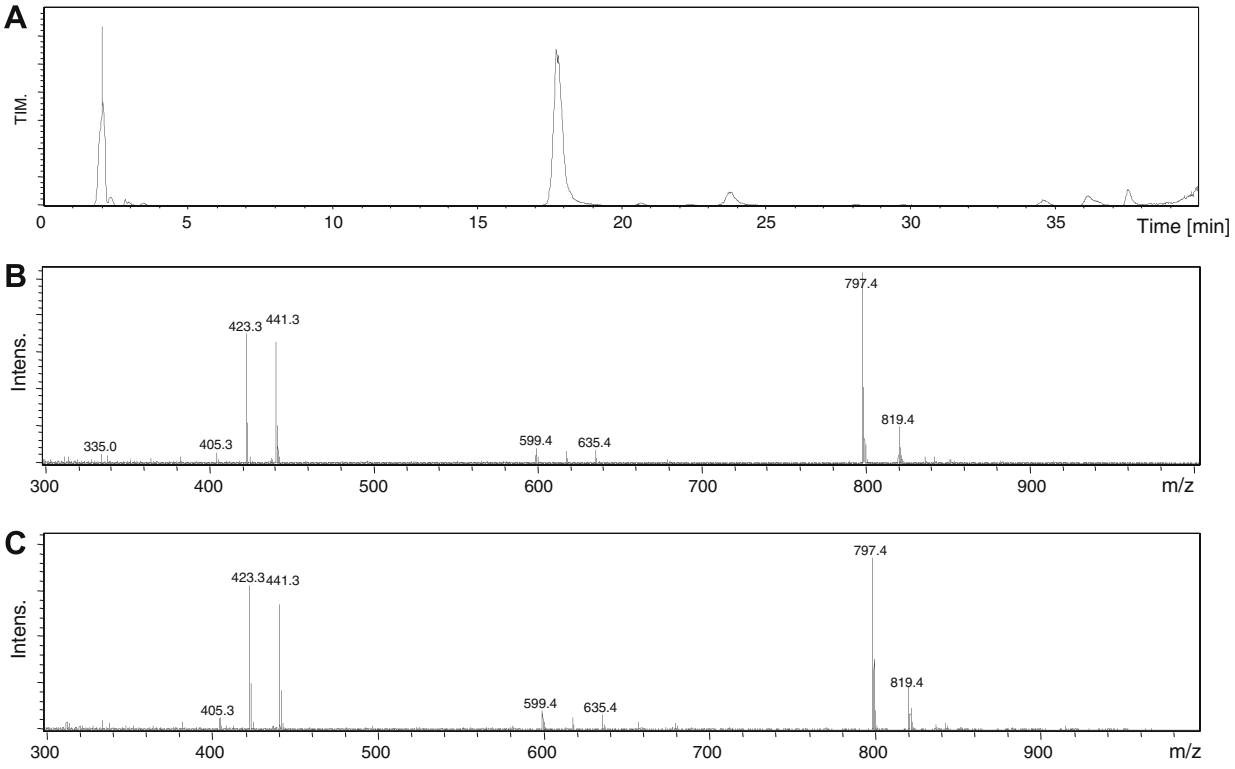

Fig. 3. LC–ESI/MS analysis of product by GTX-2. Elution profile of GTX-2 product was shown in (A). ESI/MS of peak at 17.9 min was shown in (B), and authentic soyasaponin III in (C).

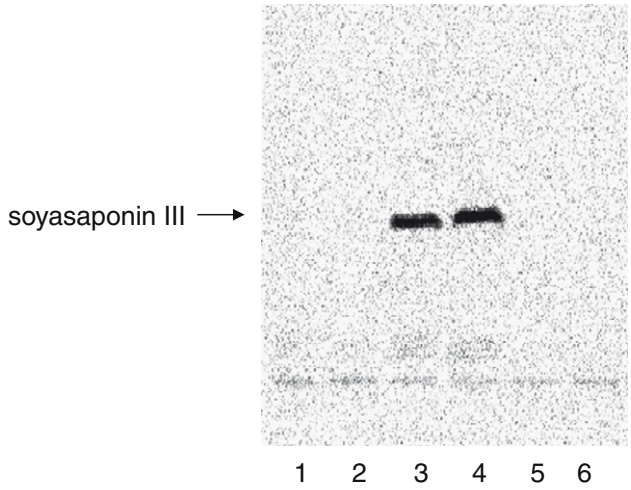

Fig. 4. Sugar donor specificity of GTX-2. Sugar donor specificity was assayed in duplicate, for UDP-glucose (lanes 1 and 2), UDP-galactose (lanes 3 and 4), and UDP-glucuronic acid (lanes 5 and 6). Radioactivity was visualized by radioautography after TLC separation.

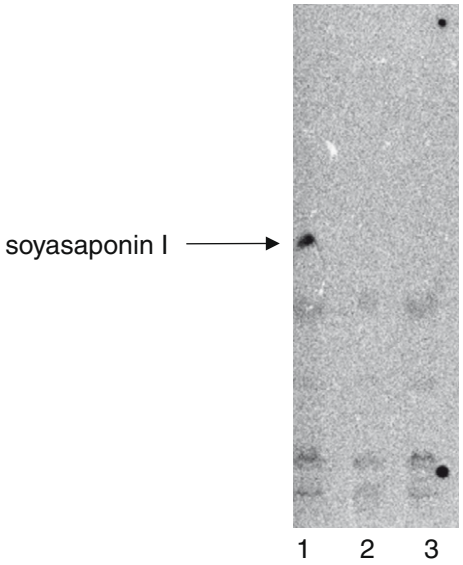

Fig. 5. TLC analysis of GTX-1 product. Lane 1 represents the one with complete assay mixture including soyasaponin III, ^{14}C-labeled UDP-glucose, unpurified RHM1, NADPH and cell-free extract of GTX-1 transformant. Lane 2 represents the one with vector transformant in place of GTX-1 transformant. Lane 3 represents the one with boiled RHM1.

12 μM and 0.02 μM/s, respectively. This K_m value is comparable to those previously reported for known plant GTs-32 μM (BpUGAT, UDP-glucuronic acid:anthocyanin glucuronosyltransferase [18]), 4 μM (GmIF7GT, UDP-glucose:genistein-glucosyltransferase [19]), 15 μM (UFGT, UDP-glucose:quercetin-glucosyltransferase [20]) and 14 μM (UGT85C2, UDP-glucose:steviol-glucosyltransferase [21]), indicating that GTX-2 has high substrate specificity for SBMG.

Based on these results, we conclude that the identity of GTX-2 is UDP-galactose:SBMG-galactosyltransferase, the second GT involved in soyasaponin I biosynthesis in *G. max*. We designated GTX-2 as GmSGT2.

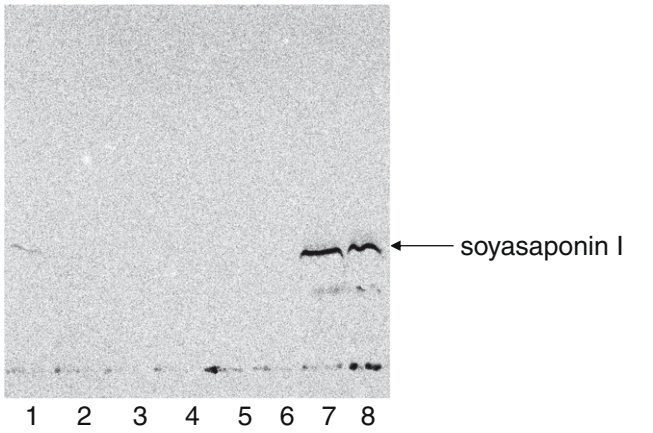

Fig. 6. Sugar donor specificity of GTX-1. Sugar donor specificity was assayed in duplicate, for UDP-glucose (lanes 1 and 2), UDP-galactose (lanes 3 and 4), UDP-glucuronic acid (lanes 5 and 6) and UDP-rhamnose. Radioactivity was visualized by radioautography after TLC separation.

3.3. UDP-rhamnose:soyasaponin III-rhamnosyltransferase—the third GT involved in soyasaponin I biosynthesis

The third GT supposedly involved in soyasaponin I biosynthesis is UDP-rhamnose:soyasaponin III-rhamnosyltransferase. UDP-rhamnose, the sugar donor in this reaction, is not commercially available. Recent studies have identified the UDP-rhamnose synthase RHM1 (At1g78570) from *Arabidopsis thaliana*, which produces UDP-rhamnose from UDP-glucose and NADPH [15,16]. In this study, we expressed RHM1 cDNA in *E. coli* to prepare UDP-rhamnose.

We examined rhamnosyltransferase activity against soyasaponin III in cell-free extracts from each of the GT candidates expressed in *E. coli*. We successfully detected a radio-labeled spot corresponding to soyasaponin I in GTX-1 (Fig. 5, lane 1). The spot was not found in the negative controls represented in lanes 2 (void vector transformant) and 3 (boiled RHM1). These results strongly suggest that *GTX-1* encodes the UDP-rhamnose:soyasaponin III-rhamnosyltransferase.

We next examined the sugar donor specificity of GTX-1. GTX-1 was incubated with ^{14}C-labeled UDP-glucose, UDP-galactose, UDP-glucuronic acid or UDP-rhamnose, together with soyasaponin III. The radio-labeled product was detected exclusively in the extract with UDP-rhamnose (Fig. 6, lanes 7 and 8). These results indicate that GTX-1 has strong sugar donor specificity for UDP-rhamnose.

Subsequently, we determined the product as soyasaponin I by LC–MS analysis. LC gave a peak at 13.2 min (Fig. 7A) with the same retention time as that of authentic soyasaponin I. The MS fragmentation pattern of this peak (Fig. 7B) was identical to that of authentic soyasaponin I (Fig. 7C), with peaks for $M + Na^+$ ($m/z = 965$) and $M + H^+$ ($m/z = 943$) as well as peaks due to successive elimination of rhamnose ($m/z = 797$), galactose ($m/z = 635$) and glucuronic acid ($m/z = 441$) from the molecular ion. These results indicate that product is soyasaponin I, and that *GTX-1* encodes UDP-rhamnose:soyasaponin III-rhamnosyltransferase. We designated GTX-1 as GmSGT3.

4. Discussion

Triterpene saponins are a diverse group of natural compounds and show a variety of biological activities. GTs contribute greatly to generating the structural diversity of triterpene saponins, and are thus important targets in investigating the origins of structural diversity. Despite this importance, few studies have been con-

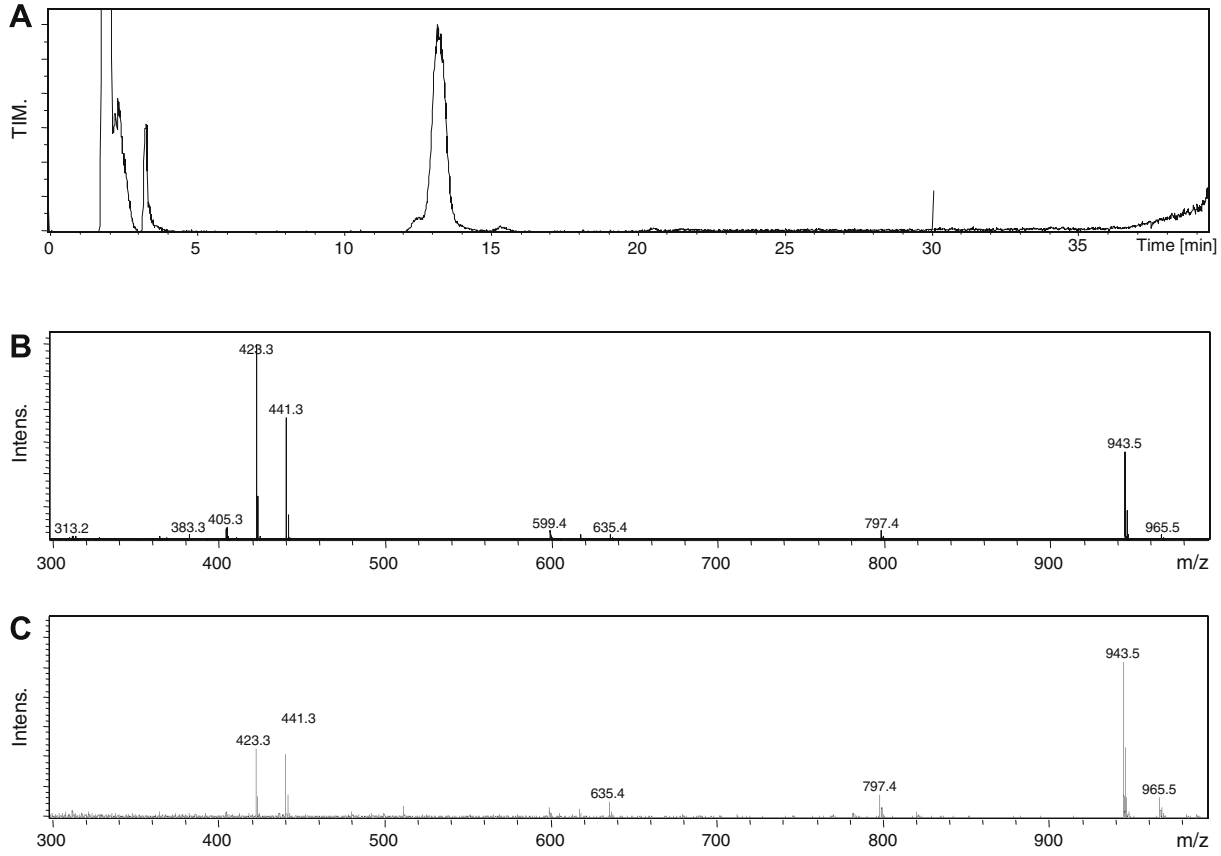

Fig. 7. LC–ESI/MS analysis of product by GTX-1. Elution profile of GTX-1 product was shown in (A). ESI/MS of peak at 13.2 min was shown in (B), and authentic soyasaponin I in (C).

ducted on GTs involved in triterpene saponin biosynthesis. In recent years, UDP-glucose:triterpene-glucosyltransferases have been successfully cloned from *M. truncatula* [12] and *S. vaccaria* [13]. While all of these were GTs which transferred a sugar to a triterpene aglycone to yield a triterpene saponin bearing a monosaccharide, it was unclear how the sugar moiety of triterpene saponins bearing two or more saccharides was constructed. In this study, we successfully identified UDP-galactose:SBMG-galactosyltransferase (GmSGT2) and UDP-rhamnose: soyasaponin III-rhamnosyltransferase (GmSGT3)—GTs involved in soyasaponin biosynthesis which yield triterpene saponins bearing a disaccharide and trisaccharide, respectively. They are the first examples of GTs which transfer the second and third sugar to the biosynthetic intermediate of triterpene saponin. Our results strongly suggest that the sugar chain moiety of triterpene saponin is biosynthesized by successive sugar transfer reactions to aglycones in a manner similar to that observed in anthocyanin biosynthesis [18], rather than a single transfer reaction of a preformed multi-sugar chain to an aglycone.

GmSGT2 and GmSGT3 have been given the names UGT73P2 and UGT91H4, respectively, by the UGT Nomenclature Committee [22,23]. GmSGT2 and GmSGT3 showed high degrees of amino acid sequence identity to full-length EST clones TC108328 and TC100519, respectively, from *M. truncatula*. They might have been included in the candidates of a previous study by Achnine et al. [12], but their enzyme functions were not identified. The cloning and functional analysis of TC108328 and TC100519 are now underway in our laboratory.

GTs' sugar donor specificity also contributes to generating the structural diversity of triterpene saponins. Kubo et al. reported that one amino acid substitution from histidine into glutamine in the UDP-galactose:anthocyanin galactosyltransferase (ACGaT) from *Aralia cordata* altered sugar donor specificity from UDP-galactose into UDP-glucose, although the reverse amino acid substitution from glutamine into histidine in the UDP-glucose:flavonoid glucosyltransferase (UBGT) from *Scutellaria baicalensis* resulted in no reversal of sugar donor specificity [24]. Shao et al., using X-ray crystal structure analysis, found that one role of the corresponding glutamine in UDP-glucose:flavonoid/triterpene glucosyltransferase (UGT71G1) is to form a hydrogen bond between the O2′ atom of UDP-glucose and the OE2 atom of glutamine to fix the position of UDP-glucose, although they failed to uncover the detailed course of changes in donor specificity [25]. However, since histidine is conserved in all cloned galactosyltransferases including GmSGT2 while glutamine is conserved in all glucosyltransferases, and substitution of histidine with glutamine in ACGaT resulted in the alternation of substrate preference, the same substitution in GmSGT2 might result in the same alternation. This could lead SBMG to produce 3-*O*-[glucosyl (1 → 2)-glucuronyl]-soyasapogenol B, a substance which has not been isolated from nature so far.

A clear course of the evolution of triterpene-GTs could not be revealed by phylogenetic analysis (See Supplementary data, S2.2) due to limited number of identified triterpene-GTs. Identification of further clones will enable us to understand the evolutionary course of triterpene-GT and the origins of the structural diversity of triterpene saponins. Unfortunately, we could not identify UDP-glucuronic acid:soyasapogenol B-glucuronosyltransferase (See Supplementary data, S2.1), the first GT. Compared to other GTs, this enzyme is reported to have quite a distinct property such as membrane binding [26], which may severely affect its expression in heterologous host cells. We are re-examining our obtained clones and continuing further cloning in our laboratory.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2010.03.037.

References

[1] Vincken, J.-P., Heng, L., de Groot, A. and Gruppen, H. (2007) Saponins, classification and occurrence in the plant kingdom. Phytochemistry 68, 275–297.

[2] Sparg, S.G., Light, M.E. and van Staden, J. (2004) Biological activities and distribution of plant saponins. J. Ethnopharmacol. 94, 219–243.

[3] Kitagawa, I., Yoshikawa, M. and Yosioka, I. (1976) Saponin and sapogenol. XIII. structures of three soybean saponins: soyasaponin I, soyasaponin II, and soyasaponin III. Chem. Pharm. Bull. 24, 121–129.

[4] Gibson, M.R. (1978) Glycyrrhiza in old and new perspectives. Lloydia 41, 348–354.

[5] Huhman, D.V. and Sumner, L.W. (2002) Metabolic profiling of saponins in *Medicago sativa* and *Medicago truncatula* using HPLC coupled to an electrospray ion-trap mass spectrometer. Phytochemistry 59, 347–360.

[6] Shibata, S. (2001) Chemistry and cancer preventing activities of ginseng saponins and some related triterpenoid compounds. J. Korean Med. Sci. 16 (Suppl), S28–S37.

[7] Park, J.D., Rhee, D.K. and Lee, Y.H. (2005) Biological activities and chemistry of saponins from *Panax ginseng* C.A. Meyer. Phytochem. Rev. 4, 159–175.

[8] Xu, R., Fazio, G.C. and Matsuda, S.P.T. (2004) On the origins of triterpenoid skeletal diversity. Phytochemistry 65, 261–291.

[9] Phillips, D.R., Rasbery, J.M., Bartel, B. and Matsuda, S.P.T. (2006) Biosynthetic diversity in plant triterpene cyclization. Curr. Opin. Plant Biol. 9, 305–314.

[10] Shibuya, M., Hoshino, M., Katsube, Y., Hayashi, H., Kushiro, T. and Ebizuka, Y. (2006) Identification of β-amyrin and sophoradiol 24-hydroxylase by expressed sequence tag mining and functional expression assay. FEBS J. 273, 948–959.

[11] Seki, H., Ohyama, K., Sawai, S., Mizutani, M., Ohnishi, T., Sudo, H., Akashi, T., Aoki, T., Saito, K. and Muranaka, T. (2008) Licorice β-amyrin 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.

[12] Achnine, L., Huhman, D.V., Farag, M.A., Sumner, L.W., Blount, J.W. and Dixon, R.A. (2005) Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. Plant J. 41, 875–887.

[13] Meesapyodsuk, D., Balsevich, J., Reed, D.W. and Covello, P.S. (2007) Saponin biosynthesis in *Saponaria vaccaria*. CDNAs encoding β-amyrin synthase and a triterpene carboxylic acid glucosyltransferase. Plant Physiol. 143, 959–969.

[14] Ikeda, T., Udayama, M., Okawa, M., Arao, T., Kinjo, J. and Nohara, T. (1998) Partial hydrolysis of soyasaponin I and the hepatoprotective effects of the hydrolytic products. Study of the structure–hepatoprotective relationship of soyasapogenol B analogs. Chem. Pharm. Bull. 46, 359–361.

[15] Diet, A., Link, B., Seifert, G.J., Schellenberg, B., Wagner, U., Pauly, M., Reiter, W.-D. and Ringlia, C. (2006) The arabidopsis root hair cell wall formation mutant lrx1 is suppressed by mutations in the RHM1 gene encoding a UDP-l-rhamnose synthase. Plant Cell 18, 1630–1641.

[16] Lim, E.-K., Ashford, D.A. and Bowles, D.J. (2006) The synthesis of small-molecule rhamnosides through the rational design of a whole-cell biocatalysis system. ChemBioChem 7, 1181–1185.

[17] Shibuya, M., Xiang, T., Katsube, Y., Otsuka, M., Zhang, H. and Ebizuka, Y. (2007) Origin of structural diversity in natural triterpenes: Direct synthesis of *seco*-triterpene skeletons by oxidosqualene cyclase. J. Am. Chem. Soc. 129, 1450–1455.

[18] Sawada, S., Suzuki, H., Ichimaida, F., Yamaguchi, M., Iwashita, T., Fukui, Y., Hemmi, H., Nishino, T. and Nakayama, T. (2005) UDP-glucuronic acid:anthocyanin glucuronosyltransferase from red daisy (*Bellis perennis*) flowers. J. Biol. Chem. 280, 899–906.

[19] 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. J. Biol. Chem. 282, 23581–23590.

[20] Ford, C.M., Boss, P.K. and Bordier, P.H. (1998) Cloning and characterization of *Vitis vinifera* UDP-glucose:flavonoid 3-*O*-glucosyltransferase, a homologue of the enzyme encoded by the maize bronze-1 locus that may primarily serve to glucosylate anthocyanidins *in vivo*. J. Biol. Chem. 273, 9224–9233.

[21] Richman, A., Swanson, A., Humphrey, T., Chapman, R., McGarvey, B., Pocs, R. and Brandle, J. (2005) Functional genomics uncovers three glucosyltransferases involved in the synthesis of the major sweet glucosides of *Stevia rebaudiana*. Plant J. 41, 56–67.

[22] Mackenzie, P.I., Owens, I.S., Burchell, B., Bock, K.W., Bairoch, A., Bélanger, A., Fournel-Gigleux, S., Green, M., Hum, D.W., Iyanagi, T., Lancet, D., Louisot, P., Magdalou, J., Chowdhury, J.R., Ritter, J.K., Schachter, H., Tephly, T.R., Tipton, K.F. and Nebert, D.W. (1997) The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. Pharmacogenetics 7, 255–269.

[23] Mackenzie, P.I., Bock, K.W., Burchell, B., Guillemette, C., Ikushiro, S., Iyanagi, T., Minersa, J.O., Owensg, I.S. and Nebert, D.W. (2005) Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily. Pharmacogenet. Genomics 15, 677–685.

[24] Kubo, A., Arai, Y., Nagashima, S. and Yoshikawa, T. (2004) Alteration of sugar donor specificities of plant glycosyltransferases by a single point mutation. Arch. Biochem. Biophys. 429, 198–203.

[25] 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.

[26] Kurosawa, Y., Takahara, H. and Shiraiwa, M. (2002) UDP-glucuronic acid: soyasapogenol glucuronosyltransferase involved in saponin biosynthesis in germinating soybean seeds. Planta 215, 620–629.