Available online at www.sciencedirect.com
ScienceDirect
Phytochemistry 68 (2007) 478–486
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Characterization and engineering of glycosyltransferases responsible for steroid saponin biosynthesis in Solanaceous plants

Atsuko Kohara [a,b], Chiharu Nakajima [a], Shigeo Yoshida [a], Toshiya Muranaka [a,*]

[a] *RIKEN Plant Science Center, Tsurumi, Yokohama, Kanagawa 230-0045, Japan*
[b] *RIKEN SPring-8 Center, RIKEN Harima Institute, Kouto, Sayo, Hyogo 679-5148, Japan*

Received 1 June 2006; received in revised form 30 October 2006; accepted 10 November 2006
Available online 3 January 2007

Abstract

Solanaceous plants contain steroid saponins that have diverse biological and pharmacological activities. The structures of their sugar chains play an important role in their activities. A functional glucosyltransferase SaGT4A from *Solanum aculeatissimum* glucosylates both steroidal sapogenins and steroidal alkaloids. A potato (*S. tuberosum*) glycosyltransferase StSGT, which has a high degree of sequence homology with SaGT4A, exhibits the same substrate specificity toward steroidal compounds as SaGT4A. To identify the residues or domain structures responsible for these enzymatic activities, we determined the residues that are essential for SaGT4A activity, compared the specific activities of SaGT4A and StSGT, and constructed several SaGT4A/StSGT chimeric proteins, focusing on the donor-sugar recognition domain. These proteins were heterogeneously expressed in *E. coli* and purified, and their glycosyltransferase activities were evaluated using a coupled assay. His369 and Glu377, located in the consensus motif for plant glycosyltransferases, and Cys121, Cys247, and Cys370 were shown to be important for SaGT4A activity. StSGT exhibited more activity with UDP-galactose as a sugar donor than with UDP-glucose, whereas SaGT4A exhibited glucosyltransferase activity exclusively. The sugar selectivities of SaGT4A and StSGT were not altered by exchanging their domains, and some of the chimeric proteins showed no activity. These results suggest that the differences in the SaGT4A and StSGT amino acid sequences do not simply reflect their distinct sugar-donor specificities. We also successfully converted the non-functional SaGT4A homolog, SaGT4R, into an active glucosyltransferase.

Keywords: Glycosyltransferase; Secondary metabolite; Solanaceae; Steroidal alkaloid; Steroidal sapogenin; *Solanum aculeatissimum*; *Solanum tuberosum*

1. Introduction

Steroidal saponins are a group of plant secondary metabolites that are glycosides. These molecules have long been thought to have pharmacological value, and researchers are becoming increasingly interested in their potential pharmacological activities, especially as anti-cancer agents (Hernandez et al., 2004; Jin et al., 2004; Nartowska et al., 2004; Raju et al., 2004; Takei et al., 2004; Yang et al., 2004; Trouillas et al., 2005). Steroidal saponins have one or two sugar chains attached by glycoside linkages to the aglycone, a non-saccharide portion of the molecule called sapogenin. The sugar chain structure contributes to the biological activities of these molecules (Grüenweller et al., 1990; Osbourn, 1996; Miyakoshi et al., 2000). Steroidal glycoalkaloids in the form of glycosides of steroidal alkaloids, which are nitrogen analogs of steroidal sapogenins, are natural toxins (Schreiber, 1968; Jadhav et al., 1981; Gaffield and Keeler, 1996; Wang et al., 2005). Researchers have found that their toxic properties depend on the type of aglycone and the sugar chain structure (Alzerreca and Hart, 1982; Osman, 1983; Maga, 1994). Glycosylation reactions are thus important to the biosynthesis of both steroidal saponins and steroidal glycoalkaloids. Understanding the diverse bioactivities of steroidal compounds will

* Corresponding author. Tel.: +81 45 503 9651; fax: +81 45 503 9492.
E-mail address: muranaka@riken.jp (T. Muranaka).

0031-9422/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2006.11.020

require the investigation of the glycosyltransferases (GTs) that catalyze the transfer of the sugar molecules.

Previously, we isolated three cDNAs encoding putative glycosyltransferases from *Solanum aculeatissimum* and demonstrated that one of these, *SaGT4A*, encodes a novel functional glucosyltransferase that catalyzes the glucosylation of both spirostane-type steroidal saponins and spirosolane-type steroidal alkaloids. SaGT4R, a homologous protein sharing 86% homology with SaGT4A, did not show any activity (Kohara et al., 2005). We also found that the potato (*S. tuberosum*) glycosyltransferase StSGT, which shows 75% sequence homology with SaGT4A and was originally isolated as a glycosyltransferase that acts on steroidal alkaloids (Moehs et al., 1997), exhibits the same substrate specificity toward steroidal sapogenins as SaGT4A. Furthermore, StSGT prefers UDP-galactose to UDP-glucose as a sugar donor in the heterogeneous protein expression systems of both *E. coli* (Kohara et al., 2005) and yeast (McCue et al., 2005).

In this study, we used *in vitro* mutagenesis to identify the amino acid residues or domains that are essential for the enzymatic activity for SaGT4A, compared the specific activities of the two functional Solanaceous glycosyltransferases (SaGT4A and StSGT), and designed chimeric enzymes. We also engineered a functional enzyme from a non-functional homolog of SaGT4A.

2. Results and discussion

2.1. Residues essential for SaGT4A activity

In this chapter, we focused on the C-terminus of the proteins, which is generally thought to be responsible for donor-sugar recognition in the GT-B superfamily of glycosyltransferases (Vrielink et al., 1994; Bourne and Henrissat, 2001; Hu et al., 2003). Based on sequence data, researchers have proposed a consensus sequence, called the PSPG box, for plant glycosyltransferases that are involved in secondary metabolism (Hughes and Hughes, 1994; Vogt and Jones, 2000). This sequence consists of about 40 amino acids in the C-terminal region of the proteins, and is thought to be the binding site for the nucleotide-diphosphate sugar. Among the highly conserved amino acids within this stretch, the His and Glu residues are the most likely to be involved in a nucleophilic S_N2 reaction mechanism that leads to the inversion of the anomeric sugar configuration (Kapitonov and Yu, 1999).

To verify the importance of the His and Glu residues, we constructed SaGT4A mutant proteins with single amino acid substitutions at His369 or Glu377 (Fig. 1), which correspond to the above residues and are also conserved in other putative Solanaceous glycosyltransferases, such as *S. aculeatissimum* SaGT4R and SaGT6 (Kohara et al., 2005), *S. tuberosum* StSGT, and *S. berthaultii* glucosyltransferase (GenBank accession number AF006081). Table 1 presents the glucosyltransferase activities of these SaGT4A mutants toward diosgenin (**1**) and solasodine (**2**). The substitution of His369 with Ala significantly decreased the enzymatic activities of the protein toward diosgenin (**1**) and solasodine (**2**), to 13.7% and 20.4%, respectively. The substitution of Glu377 with Gln considerably reduced the activity toward diosgenin (**1**), to 8.6%, and resulted in a complete loss of activity toward solasodine (**2**). These results prove that both His369 and Glu377 are essential for the glucosyltransferase activity of SaGT4A. The activity of the E377Q mutant was restored by replacing Gln with Asp, increasing the activities of the protein toward diosgenin (**1**) and solasodine (**2**) to 64.9% and 97.1%, respectively. This indicates that a negative charge in Glu377 is vital for the enzymatic activity. The H369A and E377Q mutants of StSGT also exhibited lowered activity, similar to the SaGT4A mutants (data not shown).

In studies of other plant glycosyltransferases, the residues corresponding to His369 and Glu377 were subjected to site-directed mutagenesis. The arbutin synthase E368D

Fig. 1. Comparison of C-terminal region of SaGT4A and StSGT. Sequence alignment was performed using the ClustalX and BOXSHADE programs. Numbers indicate amino acid positions. Identical residues are boxed in black and similar residues are shaded in gray. The PSPG box is underlined. Highly conserved histidine and glutamic acid residues in the plant secondary product glycosyltransferase sequences are indicated with asterisks. Three regions with low levels of sequence similarity between SaGT4A and StSGT are indicated with dotted lines (V1–V3).

Table 1
Glucosyltransferase activities of SaGT4A mutants

SaGT4A	Specific activity (nkat/mg)	
	Diosgenin (**1**)	Solasodine (**2**)
WT	1.45 (100)	1.37 (100)
H369A	0.20 (14)	0.28 (20)
E377Q	0.13 (9)	ND (0)
E377D	0.95 (66)	1.33 (97)
C121S	0.03 (2)	NT
C247S	0.08 (6)	NT
C276S	1.85 (128)	NT
C334S	1.10 (76)	NT
C370S	0.16 (11)	NT

ND: not detected; NT: not tested. Relative activities are shown in parenthesis.

mutant showed significantly decreased specific activity (Hefner and Stockigt, 2003). In another study of mutated betanidin 5-*O*-glucosyltransferases (Hans et al., 2004), H370A did not show reduced activity but E378A showed a complete loss of activity, consistent with our data. These results indicate that residues that are highly conserved among different GT subfamilies may not contribute to the catalytic mechanism in the same way.

2.2. Substrate specificities of SaGT4A and StSGT

To compare the enzymatic activities of StSGT and SaGT4A toward steroidal sapogenins and steroidal alkaloids, we analyzed the specific activity of StSGT expressed in *E. coli* using the same methodology used for SaGT4A (Kohara et al., 2005). Fig. 2 shows that the proteins share a common substrate specificity toward sugar acceptors with regard to glucosyltransferase activity. However, StSGT exhibited far lower specific activity (represented by shaded bars in Fig. 2) than SaGT4A: 10–60% of the activity of SaGT4A, except toward tigogenin (**3**) (represented by solid bars in Fig. 2). SaGT4A exhibited distinct preferences toward diosgenin (**1**) (1.45 nkat/mg), tomatidine (**4**) (1.41 nkat/mg), and solasodine (**2**) (1.37 nkat/mg), whereas StSGT exhibited no notable preferences.

The galactosyltransferase activity of StSGT (represented by open bars in Fig. 2) was similar to or higher than the glucosyltransferase activity of this protein toward both steroidal sapogenins and steroidal alkaloids.

StSGT exhibited high transferase activity, especially toward solasodine (**2**) (2.09 nkat/mg) and diosgenin (**1**) (1.98 nkat/mg) with UDP-galactose, about three and 15 times that with UDP-glucose, respectively. Sarsasapogenin (**5**) and solanidine (**6**) were galactosylated, but they were not modified with glucose by the same enzyme. A previous study identified StSGT as solanidine UDP-glucose glucosyltransferase (Moehs et al., 1997), and a recent study observed these enzymatic characteristics using recombinant proteins expressed in yeast and in plants as the steroidal alkaloid galactosyltransferase (McCue et al., 2005). Our results indicate that the enzymatic characteristics of this protein should be defined as solasodine UDP-galactose galactosyltransferase with minor glucosyltransferase activity.

2.3. SaGT4A/StSGT chimera mutants

We demonstrated that the sugar-donor specificities of SaGT4A and StSGT differ, even though these enzymes share high levels of sequence similarity and common residues that are essential to their activities. Three regions with low levels of sequence similarity between SaGT4A and StSGT appear in their C-terminal domains (Fig. 1). We constructed and analyzed several SaGT4A/StSGT chimeric enzymes to determine the specific domains or residues that are critical to the sugar recognition of these proteins.

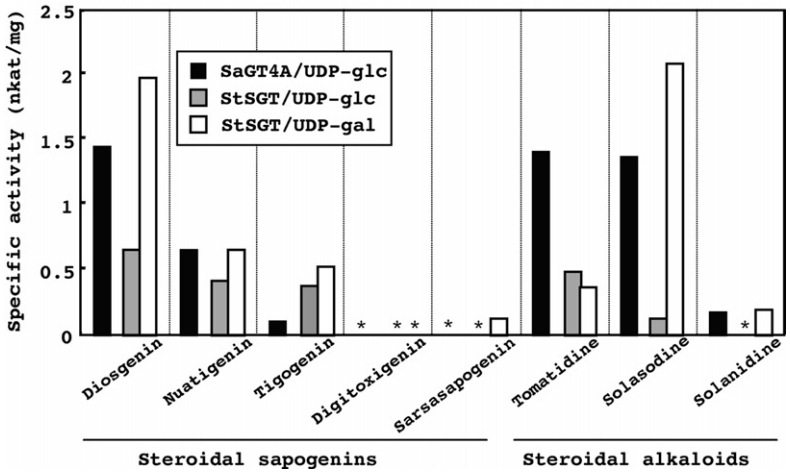


Fig. 2. Glycosyltransferase activities of SaGT4A and StSGT. Specific activity was measured using UDP-glucose or UDP-galactose as sugar donors and several steroidal compounds as sugar acceptors. Diosgenin, nuatigenin, tigogenin, and sarsasapogenin were used as steroidal sapogenins. Tomatidine, solasodine and solanidine were used as steroidal alkaloids. The SaGT4A values are from our previous data (Kohara et al., 2005). Asterisk indicates "Not Detected". All assays were performed in triplicate.

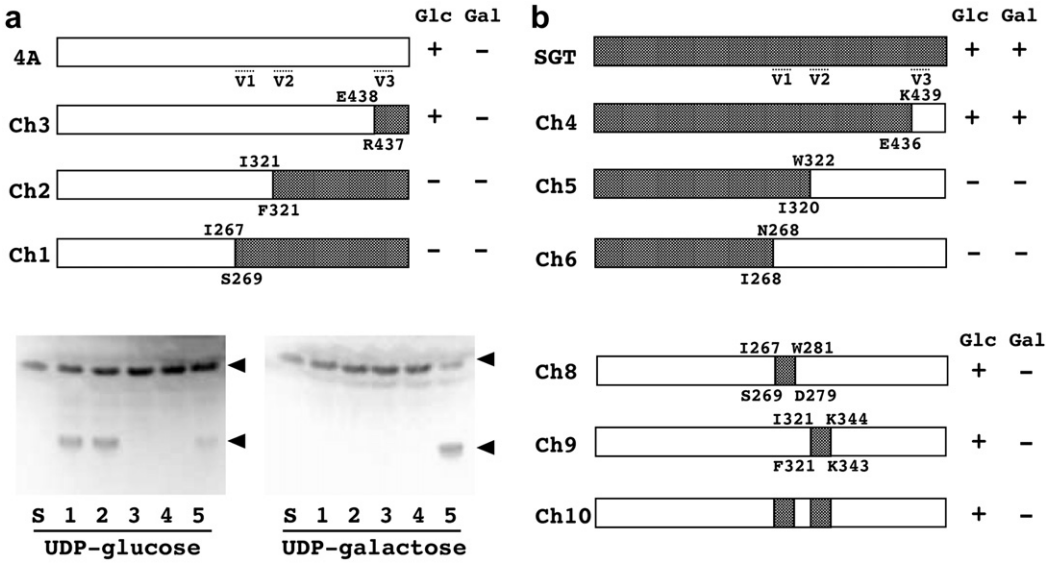


Fig. 3. Glycosyltransferase activities of chimera proteins. (a) (Top) Replacement of the C-terminal portion of SaGT4A with StSGT fragments at V1–V3. Residues at the junction of SaGT4A (top) and StSGT (below) are indicated. (Bottom) TLC analysis of enzymatic activities of chimeric proteins. Diosgenin glycoside (lower arrow) was detected on a Silica Gel 60 plate developed using $CHCl_3$/MeOH (8.5:1.5, v/v). The products were stained with *p*-methoxybenzaldehyde/AcOH/H_2SO_4/EtOH (2.5:1:3.5:104, by vol.). S: 2 μg of diosgenin (upper arrow). The tested proteins include 1, SaGT4A; 2, Ch3; 3, Ch2; 4, Ch1; and 5, StSGT. (b) (Top) Replacement of the C-terminal portion of StSGT with SaGT4A fragments at V1–V3. (Bottom) Replacement of the V1 or/and V2 segment(s) of SaGT4A with the corresponding region(s) of StSGT.

We designed three chimeras (Ch1–Ch3) in which the C-terminal portion of SaGT4A was replaced with a StSGT fragment at V1–V3, respectively (Fig. 3a, top), and examined their enzymatic activities. We first performed qualitative analyses using thin layer chromatography (TLC) to specify the critical region on which to focus subsequent detailed quantitative analyses. Taken together, the results of the TLC and coupled assays indicate that the intensity of the products detected on TLC correlated well with the specific activity of the proteins (data not shown). TLC analysis revealed that Ch3 exhibited only glucosyltransferase activity and no activity toward UDP-galactose (Fig. 3a, bottom). Ch1 and Ch2 showed no activity. Inverse results were obtained with the chimeras Ch4–Ch6 (Fig. 4), in which the C-terminal portion of StSGT was replaced with a SaGT4A fragment (Fig. 3b, top). Ch4 retained both glucosyl- and galactosyltransferase activity, which was also intact in StSGT, but Ch5 and Ch6 showed no activity (TLC data not shown). These results suggest that at least the C-terminal portion of the proteins (including V3) does not influence sugar specificity. It is questionable that the domain exchanges performed in Ch1, Ch2, Ch5, and Ch6 caused a loss of activity, because the amino acid sequences in this region are highly conserved between SaGT4A and StSGT, apart from the areas of V1–V3. It is possible that the proper folding required for the enzymatic activity of these proteins might not have taken place due to the chimeric construction. We observed no problems suggesting misfolding or aggregation of the mutant proteins during protein production and purification. We extended the incubation time of the reactions from 3 h to 16 h, but still no product was detected with Ch1, Ch2, Ch5, or Ch6. This suggests a complete loss of activity in these mutants.

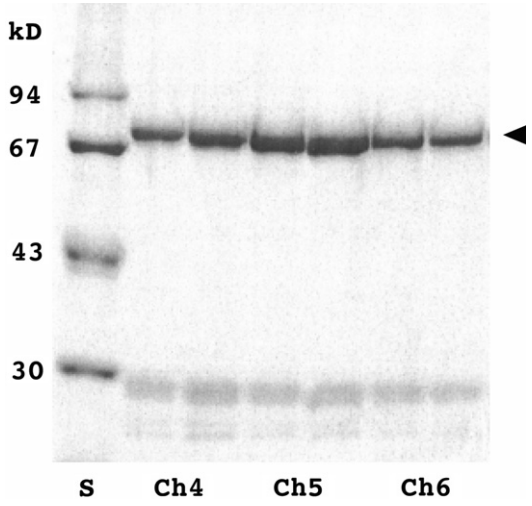


Fig. 4. SDS-PAGE of chimera proteins. Preparation of GST-fusion protein (two clones for each construct). The proteins were purified from *E. coli* using glutathione Sepharose 4B. The protein samples were separated on 10% SDS gel. The gel was Coomassie stained. The positions of the fusion proteins are indicated by arrow. Protein standards (S) contain 1–1.5 μg of each protein.

We then replaced the V1 or/and V2 segment(s) of SaGT4A with the corresponding segment(s) of StSGT, and tested the activity of the partial chimeric proteins Ch8–Ch10 (Fig. 3b, bottom). All three mutants exhibited enzymatic activity, but their activity was limited to glucose. Replacement of these segments did not abolish the activity nor alter the enzymatic characteristics of SaGT4A. These

results indicate that the glucose/galactose discrimination of SaGT4A and StSGT is not defined simply by the highly divergent V1–V3 regions of their C-termini.

Although we were unable to identify the specific residues responsible for sugar selectivity, a previous report analyzed the alteration of the sugar-donor specificity of an anthocyanin 3-*O*-galactosyltransferase from *Aralia cordata*, ACGaT, by a single amino acid substitution (Kubo et al., 2004). When His was replaced by Gln, ACGaT acquired glucosyltransferase activity without losing its original galactosyltransferase activity. The target His is conserved in six galactosyltransferases, and Gln is conserved at the same position in seven glucosyltransferases. The corresponding residue in SaGT4A is Gln394, which is consistent with these results, but the residue at the same position in StSGT is also Gln. The mutant ACGaT appears to have the same enzymatic characteristics as StSGT, which exhibits both glucosyltransferase and galactosyltransferase activity. Recently, a glucosyltransferase SGT2 lacking galactosyltransferase activity was isolated (McCue et al., 2006). SGT2 shows 73% amino acid similarity with StSGT, and the targeted residue in the study of ACGaT is also Gln. Taken together, these observations suggest that, at least in steroidal sapogenin glycosyltransferases, this position is not essential for sugar-donor specificity.

2.4. SaGT4A Cys mutants

The unexpected loss of activity by some of the chimeric enzymes suggests that the mutated proteins had unsuitable conformations. SaGT4A contains five cysteines, at positions 121, 247, 276, 334, and 370, and some of these are thought to participate in disulfide bond formation. To determine the importance of the cysteine residues in SaGT4A, we replaced each Cys with Ser, and analyzed

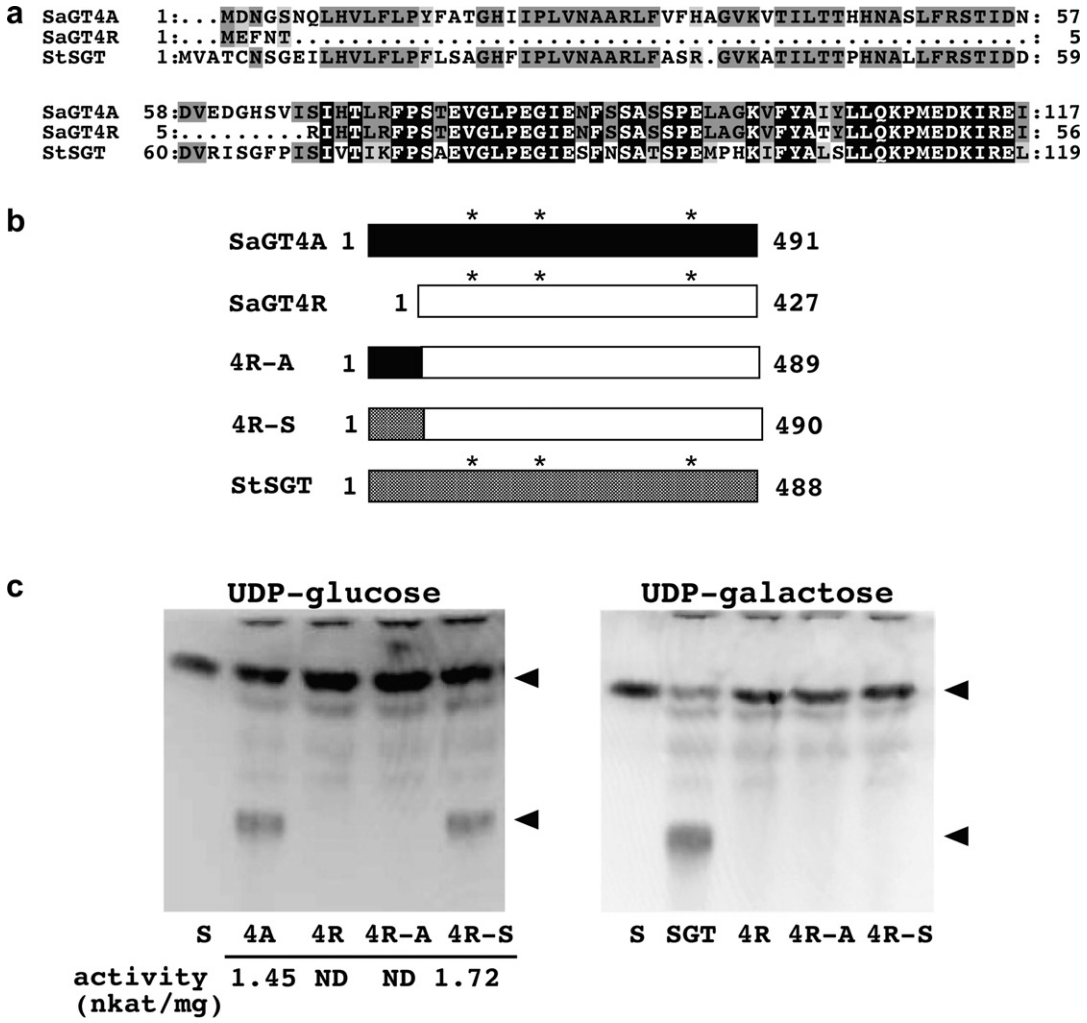


Fig. 5. Putative glycosyltransferase SaGT4R. (a) Sequence alignment of the N-terminal portions of SaGT4A, SaGT4R, and StSGT. Residues identical in all the sequences are boxed in black, residues identical in two sequences are shaded in gray, and similar residues are shaded in pale gray. (b) Construction of SaGT4R mutants. The positions of conserved cysteine residues are indicated with asterisks. (c) TLC analysis and specific activities of SaGT4R mutants. Diosgenin glycoside (lower arrow) was detected on a Silica Gel 60 plate with the same condition as described in Fig. 3. S: 2 μg of diosgenin (upper arrow). ND: not detected. All assays were performed in triplicate.

the activity of the resulting proteins. Drastically decreased activity was observed in C121S, C247S, and C370S, to less than 10% that of the wild-type, whereas C276S and C334S showed normal activity (Table 1). These results indicate that Cys121, Cys247, and Cys370 are important for SaGT4A activity.

Researchers have recently built a structural model of the cyanohydrin glycosyltransferase UGT85B1 from *Sorghum bicolor* (Thorsoe et al., 2005), predicting that Cys369 interacts with the nucleoside component of UDP-Glc. Among the three Cys residues that are indispensable for SaGT4A activity, Cys370 corresponds to the cysteine that is thought to be involved in UDP sugar recognition. Two others, Cys121 and Cys247, may form a disulfide bond. Interestingly, Cys121, Cys247, and Cys370 are also conserved in StSGT at positions 123, 249, and 370, respectively, implying that SaGT4A and StSGT have similar conformations. Nevertheless, the local configuration that contributes to proper protein function would differ between SaGT4A and StSGT, because the chimeras Ch1, Ch2, Ch5, and Ch6 showed no activity even when the Cys residues were conserved after domain replacement. We did not find a significant difference in the expression levels of the recombinant proteins, suggesting suitable folding of these inactive mutants (data not shown).

Previously, we found that other Solanaceous GTs, a non-functional SaGT6 and a *S. berthaultii* protein of unknown function, contain a residue equivalent to Cys370 at positions 362 and 361, respectively, but the other two residues are not conserved (Kohara et al., 2005). In contrast, glycosyltransferases known to act on phenylpropanoids, such as TOGT, DbsGT, and UBGT, show conserved Cys residues corresponding to Cys121 and Cys370, but not Cys247. These observations suggest that Cys247 is a glycosyltransferase residue specific to steroidal sapogenins and steroidal alkaloids. The recently reported crystal structures of flavonoid glycosyltransferases do not indicate the presence of intra-molecular disulfide bonds (Shao et al., 2005; Offen et al., 2006).

2.5. SaGT4R extension mutants

We previously isolated the *SaGT4R* gene, which encodes a putative glycosyltransferase, along with *SaGT4A*, from *S. aculeatissimum*. SaGT4R exhibited no activity despite its considerable sequence similarity to SaGT4A (Kohara et al., 2005). In addition to a PSPG motif, the three Cys residues that are conserved in SaGT4A and StSGT are also found in SaGT4R at positions 60, 186, and 310. However, there is a significant difference between SaGT4R and functional proteins in the N-terminal region. SaGT4R is shorter than SaGT4A and StSGT by about 60 amino acids (Fig. 5a), suggesting that this part of the protein is important for enzyme function.

Based on this speculation, we constructed SaGT4R mutants and examined their activities. In the mutants SaGT4R-A and SaGT4R-S, the N-terminal portion of SaGT4R was replaced with the first 76 residues of SaGT4A and the first 78 residues of StSGT, respectively (Fig. 5b). Although the sequences adopted from SaGT4A and StSGT are very similar (Fig. 5a) and the two mutant proteins exhibited the same properties during preparation (data not shown), only SaGT4R-S, with its extra StSGT N-terminal segment, exhibited glucosyltransferase activity (Fig. 5c). SaGT4R-S exhibited specific activity (1.72 nkat/mg) toward diosgenin. Neither of the mutant proteins showed galactosyltransferase activity (Fig. 5c). These results indicate that the non-functional SaGT4R successfully acquired equivalent or higher levels of SaGT4A enzymatic activity through the addition of an N-terminal segment.

The artificial N-termini of SaGT4R-A and SaGT4-S contain His residues at positions 20 and 23, respectively, the importance of which has been indicated in recent studies (Hans et al., 2004; Thorsoe et al., 2005; Shao et al., 2005; Offen et al., 2006). The crystal structures of flavonoid glycosyltransferases suggest that the highly conserved N-terminal His interacts with acceptor molecules (Shao et al., 2005; Offen et al., 2006). Owing to the difference in chemical structures between flavonoids and steroidal compounds, it is not clear whether the conserved His acts in a similar manner in SaGT4A and StSGT. However, it is possible that subtle differences in the environment surrounding this residue might influence the activities of SaGT4R-A and SaGT4R-S. Further research is required to clarify the compatibility of SaGT4R with the added sequence.

3. Conclusions

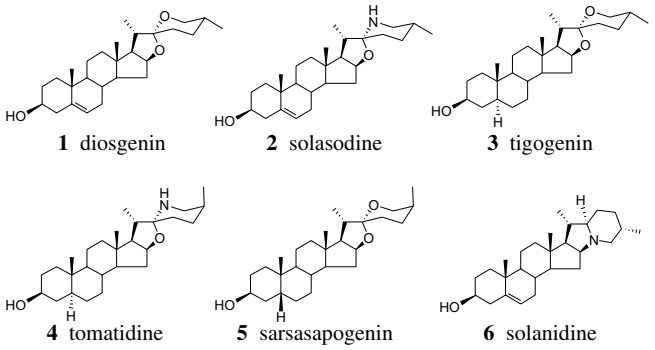
In this study, we showed how the enzymatic characteristics of StSGT differ from those of SaGT4A:StSGT has greater galactosyltransferase activity than glucosyltransferase activity and broad substrate specificity toward steroidal compounds. We found essential amino acid residues for these glycosyltransferase activities. We also converted SaGT4R, a putative glycosyltransferase with no activity, to a functional enzyme by introduction of N-terminal fragment from other enzyme. This is the first step toward the directed evolution of novel glycosyltransferases involved in the modification of bioactive molecules. Our results indicate that, along with the specific amino acids critical for enzymatic activity, the influence of the sequences replaced or introduced in the protein should be considered. Slight variation in the amino acid sequence may lead to distortion of catalytic site of the enzyme resulting in the loss of function.

Our findings provide basic information about the structure–function relationships of glycosyltransferases. Studies combining mutagenesis, molecular modeling, and structural analysis will be necessary to further identify the residues or domain structures that are responsible for the enzymatic activity and substrate specificity.

4. Experimental

4.1. Chemicals

The steroidal compounds diosgenin (**1**), sarsasapogenin (**5**), solanidine (**6**), solasodine (**2**), tigogenin (**3**), and tomatidine (**4**) were purchased from Sigma. UDP-glucose, UDP-galactose, pyruvate kinase, lactate dehydrogenase, phospho(enol)pyruvic acid, and NADH were also obtained from Sigma. *p*-Methoxybenzaldehyde and other organic solvents were obtained from Wako. Nuatigenin was prepared using acid hydrolysis of aculeatiside A as described (Saijo et al., 1983).



4.2. Mutagenesis

Site-directed mutagenesis was performed using PCR with the following primer pairs (substituted amino acids are underlined): sense HA3 (5′-TCATGACAGCCTGCGGTTGGA-3′) and antisense HA4 (5′-CAACCGCAGGCTGTCATGAATC-3′) for *H*369*A*; sense EQ3 (5′-TTCCATACTCCAAGCCATCATCGC-3′) and antisense EQ4 (5′-CGATGATGGCTTGGAGTATGGAATTC-3′) for *E*377*Q*; sense ED3 (5′-TTCCATACTCGATGCCATCATCGC-3′) and antisense ED4 (5′-CGATGATGGCATCGAGTATGGAATTC-3′) for *E*377*D*, sense C1 (5′-CGGATTCTATTTTCTCTGATATGTACCTC-3′) and antisense C2 (5′-GAAAATAGAATCCGGATGGATTTCAC-3′) for *C*121*S*; sense C3 (5′-CTAAATCTTGGCAAATTGGTCCCAT-3′) and antisense C4 (5′-TGCCAAGATTTAGTTTTCTTCACCTTCT-3′) for *C*247*S*; sense CS276s (5′-ACTCATCTGCCATTGTAGAGTGGTTG-3′) and antisense CS276a (5′-AATGGCAGATGAGTTACTTTCATCGA-3′) for *C*276*S*; sense CS334s (5′-CCACGTCTTTGCTGGAGGAGG-3′) and antisense CS334a (5′-AGCAAAGACGTGGTTTCTGCCG-3′) for *C*334*S*; sense C5 (5′-CACACTCCGGTTGGAATTCCATACTC-3′) and antisense C6 (5′-CAACCGGAGTGTGTCATGAATCC-3′) for *C*370*S*.

Two rounds of PCR were carried out using PfuTurbo DNA polymerase (Stratagene). The plasmid pCRII-TOPO (Invitrogen) containing the *SaGT4A* coding region (Kohara et al., 2005) was used as a template. In the first round of PCR, two separate reactions were performed for each mutant, with SA4A-1 (5′-ctggatccctATGGATAACGGCAGCAATCA-3′) and an antisense mutagenic primer in one tube, and with SA4A-2 (5′-gcgtcgacTTAAACGTTCGTCGTCGAAGTAA-3′) and a sense mutagenic primer in another tube. In the second round of PCR, a mixture of the products of the first two PCR reactions was used as a template. The mutated full-length cDNAs were amplified using the primers SA4A-1 and SA4A-2. The resulting products were fused into pCRII-TOPO and sequenced as previously described.

The chimera mutants were constructed using the following primer pairs (nucleotides derived from *SaGT4A* and *StSGT* are presented in upper and lower case, respectively): Sense CH6 (5′-AATAGAGAgactaatgatcagtaatggttctg-3′) and antisense CH5 (5′-tactgatCATTAgTcTCTCTATTGCTTCTTTGA-3′) for *Ch3*; sense CH4 (5′-CCTTTCATTtttgtattgaggcctaatgaa-3′) and antisense CH7 (5′-aggCcTCAaTACAaaAATGAAAGGGATGCTC-3′) for *Ch2*; sense CH2 (5′-AGAGCTGATTtctgagcataacaacaatgag-3′) and antisense CH1 (5′-tatgctCAgaAATCAGCTCTTTTCTACGGA-3′) for *Ch1*; sense CH8 (5′-agaagcaattGAGAaattaatggatgactc-3′) and antisense CH9 (5′-tccattaattTCTCTATTGCTTCTctaatcttg-3′) for *Ch4*; sense CH10 (5′-aatgttCCTTTCATTtgggtagtgaagga-3′) and antisense CH11 (5′-ctacccaAATGAAAGGaacatttgaagcat-3′) for *Ch5*; sense CH12 (5′-aaggaactaATTaatgctgtcgatgaaagtaac-3′) and antisense CH13 (5′-tcgacagcattAATtagttccttactacgg-3′) for *Ch6*; sense CH14 (5′-atagattggttgaacga-gcaggagcataaa-3′) and antisense CH15 (5′-ctgctcgttcaaccaatctataacaatctca-3′) for *Ch8*; sense CH16 (5′-agactaaaaagggtctgattattagagggtg-3′) and antisense CH17 (5′-ataatcagacccttttagtcttgtcctctaa-3′) for *Ch9*.

Two rounds of PCR were carried out as described above. The first set of PCR for *Ch3*, 2 and 1 was performed with SA4A-1 and an antisense mutagenic primer in one tube with the *SaGT4A* coding region as a template, and with GT11 (5′-CGGAATTCCTAATTAAGATTATAATTCTTGATATGTTG-3′) and a sense mutagenic primer in another tube with the *StSGT* coding region as a template. Full-length cDNAs were amplified using the primers SA4A-1 and GT11 in a second PCR reaction containing a mixture of the first reaction products as a template.

In the construction of *Ch4*, 5, and 6, the first reactions were performed with ST-1 (5′-CCTGGATCCCTATGGTAGCAACCTGCAACAGT-3′) and an antisense mutagenic primer in one tube with the *StSGT* coding region as a template, and with SA4A-2 and a sense mutagenic primer in another tube with the *SaGT4A* coding region as a template. Full-length cDNAs were amplified using the primers ST-1 and SA4A-2 in the second PCR.

For the construction of *Ch8* and *Ch9*, we used SA4A-1 with an antisense mutagenic primer in one tube containing the *Ch1* and *Ch2* coding regions, respectively, as templates. SA4A-2 was used with a sense mutagenic primer in another tube with the *SaGT4A* coding region as a template. Full-

length cDNAs were amplified using SA4A-1 and SA4A-2 in the second PCR reaction. *Ch10* was constructed by replacing the 550-bp *Xba*I-*Sal*I fragment of *Ch8* with the corresponding fragment of *Ch9*.

The Sa4R-A and SaGT4R-S mutants were constructed by replacing the 200-bp *Bam*HI-*Nco*I fragment of *SaGT4R* with the 400-bp *Bam*HI-*Nco*I fragment of *SaGT4A* and *Ch7*, respectively. *Ch7* was accidentally obtained during the construction of *Ch4*, resulting from the replacement of the first 76 amino acids of SaGT4A with the first 78 amino acids of StSGT.

4.3. Expression of GT proteins

Recombinant GTs were expressed as fusion proteins with the GST region at the N terminus as described (Kohara et al., 2005). The *E. coli* strain DH5α was transformed with the expression vector and grown under inducible conditions. Protein expression was induced with 0.1 mM isopropyl-1-thio-β-D-galactopyranoside, and the fusion proteins were extracted and purified using glutathione Sepharose 4B (Amersham Pharmacia Biotech). Protein concentrations were determined using densitometry along with SDS-PAGE, using NIH Image (http://rsb.info.nih.gov/nih-image/). Purification steps were performed either at 4 °C or on ice.

4.4. GT activity assay

The standard reaction mixture for TLC assays contained 0.1 M Tris–HCl buffer (pH 8), 0.1 mM UDP-glucose (or UDP-galactose), 0.1 mM steroidal aglycones, and 2 μg of enzyme in 250 μl. After 3 h at 37 °C, the reaction was terminated by heating. Steroidal compounds were then isolated through three extractions with *n*-BuOH 250 μl and separated by TLC (Silica gel 60, Merck). The specific activity of the GT proteins was defined as the release of UDP, which can be measured using a coupled assay containing GT proteins, pyruvate kinase, and lactate dehydrogenase (Jackson et al., 2001). The reaction mechanisms are as follows:

$$\text{Steroidal compound} + \text{UDP-Glc (or Gal)} \Longleftrightarrow \text{steroidal glycoside} + \text{UDP} \quad (1)$$

$$\text{Phospho(enol)pyruvic acid(PEP)} + \text{UDP} \Longleftrightarrow \text{UTP} + \text{pyruvate} \quad (2)$$

$$\text{Pyruvate} + \text{NADH} + \text{H}^+ \Longleftrightarrow \text{lactate} + \text{NAD}^+ \quad (3)$$

The reaction mixture contained 50 mM Tris–HCl (pH 8), 2.5 mM $MgSO_4$, 10 mM KCl, 0.12 mM NADH, 1.6 mM PEP, 500 μg UDP-glucose, 0.6 units of pyruvate kinase, 0.8 units of lactate dehydrogenase, 10 μg steroidal substrate, and 1–1.5 μg enzyme in a total volume of 250 μl. Reaction mixtures lacking enzyme were used as controls. The reactions were incubated at 30 °C for 60 min. Changes in NAD^+ were detected at 340 nm, and the reaction rate was converted to the units of ncat mg^{-1} using the extinction coefficient $6.22 \times 10^3\ M^{-1}\ cm^{-1}$ for NADH. Specific activity was converted to relative activity, defining the value of wild-type SaGT4A toward diosgenin as 100%.

Acknowledgement

This work was supported in part by a Grant-in-Aid for Scientific Research (C) (No. 17510188) to T.M. from the Japan Society for the Promotion of Science, Japan.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.phytochem.2006.11.020.

References

Alzerreca, A., Hart, G., 1982. Molluscicidal steroid glycoalkaloids possessing stereoisomeric spirosolane structures. Toxicol. Lett. 12, 151–155.

Bourne, Y., Henrissat, B., 2001. Glycoside hydrolases and glycosyltransferases: families and functional modules. Curr. Opin. Struct. Biol. 11, 593–600.

Gaffield, W., Keeler, R.F., 1996. Craniofacial malformations induced in hamsters by steroidal alkaloids. J. Nat. Tox. 5, 25–38.

Grünweller, S., Schröder, E., Kesselmeier, J., 1990. Biological activities of furostanol saponins from *Nicotiana tabacum*. Phytochemistry 29, 2485–2490.

Hans, J., Brandt, W., Vogt, T., 2004. Site-directed mutagenesis and protein 3D-homology modelling suggest a catalytic mechanism for UDP-glucose-dependent betanidin 5-*O*-glucosyltransferase from *Dorotheanthus bellidiformis*. Plant J. 39, 319–333.

Hefner, T., Stockigt, J., 2003. Probing suggested catalytic domains of glycosyltransferases by site-directed mutagenesis. Eur. J. Biochem. 270, 533–538.

Hernandez, J.C., Leon, F., Quintana, J., Estevez, F., Bermejo, J., 2004. Icogenin, a new cytotoxic steroidal saponin isolated from *Dracaena draco*. Bioorg. Med. Chem. 12, 4423–4429.

Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., Walker, S., 2003. Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases. Proc. Natl. Acad. Sci. 100, 845–849.

Hughes, J., Hughes, M.A., 1994. Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava. DNA Seq. 5, 41–49.

Jackson, R.G., Lim, E.-K., Kowalczyk, M., Sandberg, G., Hoggett, J., Ashford, D.-A., Bowles, D.-J., 2001. Identification and biochemical characterization of an *Arabidopsis* indole-3-acetic acid glucosyltransferase. J. Biol. Chem. 276, 4350–4356.

Jadhav, S.J., Sharma, R.P., Salunkhe, D.K., 1981. Naturally occurring toxic alkaloids in foods. Crit. Rev. Toxicol. 9, 21–47.

Jin, J.M., Zhang, Y.J., Yang, C.R., 2004. Cytotoxic steroidal saponins from *Polygonatum zanlanscianense*. J. Nat. Prod. 67, 1992–1995.

Kapitonov, D., Yu, R.K., 1999. Conserved domains of glycosyltransferases. Glycobiology 9, 961–978.

Kohara, A., Nakajima, C., Hashimoto, K., Ikenaga, T., Tanaka, H., Shoyama, Y., Yoshida, S., Muranaka, T., 2005. A novel glucosyltransferase involved in steroid saponin biosynthesis in *Solanum aculeatissimum*. Plant Mol. Biol. 57, 225–239.

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

Maga, J.A., 1994. Glycoalkaloids in Solanaceae. Food Rev. Int. 10, 385–418.

McCue, K.F., Shepherd, L.V.T., Allen, P.V., Maccree, M.M., Rockhold, D.R., Corsini, D.L., Davies, H.V., Belknap, W.R., 2005. Metabolic compensation of steroidal glycoalkaloid biosynthesis in transgenic potato tubers: using reverse genetics to confirm the in vivo enzyme function of a steroidal alkaloid alactosyltransferase. Plant Sci. 168, 267–273.

McCue, K.F., Allen, P.V., Shepherd, L.V.T., Blake, A., Whitworth, J., Maccree, M.M., Rockhold, D.R., Stewart, D., Davis, H.V., Belknap, W.R., 2006. The primary in vivo steroidal alkaloid glucosyltransferase from potato. Phytochemistry 67, 1590–1597.

Miyakoshi, M., Tamura, Y., Masuda, H., Mizutani, K., Tanaka, O., Ikeda, T., Ohtani, K., Kasai, R., Yamasaki, K., 2000. Antiyeast steroidal saponins from *Yucca schidigera* (Mohave Yucca), a new anti-food-deteriorating agent. J. Nat. Prod. 63, 332–338.

Moehs, C.P., Allen, P.V., Belknap, W.R., 1997. Cloning and expression of solanidine UDP-glucose glucosyltransferase from potato. Plant J. 11, 227–236.

Nartowska, J., Sommer, E., Pastewka, K., Sommer, S., Skopinska-Pozewska, E., 2004. Anti-angiogenic activity of convallamaroside, the steroidal saponin isolated from the rhizomes and rots of *Convallaria majalis* L. Acta. Pol. Pharm. 61, 279–282.

Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B.G., Tarling, C.A., Ford, C.M., Bowles, D.J., Davies, G.J., 2006. Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. EMBO J. 25, 1396–1405.

Osbourn, A., 1996. Saponins and plant defence: a soap story. Trends Plant Sci. 1, 4–9.

Osman, S.F., 1983. Potato glycoalkaloids. Food Chem. 11, 235–247.

Raju, J., Patlolla, J.M., Swamy, M.V., Rao, C.V., 2004. Diosgenin, a steroid saponin of *Trigonella foenum graecum* (Fenugreek), inhibits azoxymethane-induced aberrant crypt foci formation in F344 rats and induces apoptosis in HT-29 human colon cancer cells. Cancer Epidemiol. Biomarkers Prev. 13, 1392–1398.

Saijo, R., Fuke, C., Murakami, K., Nohara, T., Tomimatsu, T., 1983. Two steroidal glycosides, aculeatiside A and B from *Solanum aculeatissimum*. Phytochemistry 22, 733–736.

Schreiber, K., 1968. Toxic constituents of food plants. Kulturpflanze 16, 255–276.

Shao, H., He, X., Achnine, L., Blount, J.W., Dixon, R.A., Wang, X., 2005. Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. Plant Cell 17, 3141–3154.

Thorsoe, K.S., Bak, S., Olsen, C.E., Imberty, A., Breton, C., Moller, B.L., 2005. Determination of catalytic key amino acids and UDP sugar donor specificity of the cyanohydrin glycosyltransferase UGT85B1 from *Sorghum bicolor*. Molecular modeling substantiated by site-specific mutagenesis and biochemical analyses. Plant Physiol. 139, 664–673.

Trouillas, P., Corbiere, C., Liagre, B., Duroux, J.L., Beneytout, J.L., 2005. Structure–function relationship for saponin effects on cell cycle arrest and apoptosis in the human 1547 osteosarcoma cells: a molecular modelling approach of natural molecules structurally close to diosgenin. Bioorg. Med. Chem. 13, 1141–1149.

Vogt, T., Jones, P., 2000. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. Trends Plant Sci. 5, 380–386.

Vrielink, A., Rüger, W., Driessen, H.P.C., Freemont, P.S., 1994. Crystal structure of the DNA modifying enzyme β-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. EMBO J. 13, 3413–3422.

Wang, S., Panter, K.E., Gaffield, W., Evans, R.C., Bunch, T.D., 2005. Effects of steroidal glycoalkaloids from potatoes (*Solanum tuberosum*) on in vitro bovine embryo development. Anim. Reprod. Sci. 85, 243–250.

Yang, M.F., Li, Y.Y., Gao, X.P., Li, B.G., Zhang, G.L., 2004. Steroidal saponins from *Myriopteron extensum* and their cytotoxic activity. Planta Med. 70, 556–560.