



CrossMark  
 click for updates

Cite this: *RSC Adv.*, 2017, 7, 16217

## Biosynthesis of natural and non-natural genistein glycosides†

Prakash Parajuli,<sup>a</sup> Ramesh Prasad Pandey,<sup>ab</sup> Trang Thi Huyen Nguyen,<sup>a</sup> Biplav Shrestha,<sup>a</sup> Tokutaro Yamaguchi<sup>ab</sup> and Jae Kyung Sohng<sup>\*ab</sup>

Genistein is the principal isoflavone constituent of soybean. It has attracted more attention than other plant polyphenols because of its significant affinity with medical interests. Herein, we biosynthesized and structurally characterized ten different natural and non-natural analogues of genistein glycopyranosides using versatile glycosyltransferases (GTs) and sugar-*O*-methyltransferases (SOMTs). Two GTs, AtUGT89C1 from the *Arabidopsis thaliana* plant and YjIC from *Bacillus licheniformis* DSM-13 bacteria, catalyzed the glycosylation reaction. Two SOMTs transferred the methyl group to two specific hydroxy positions in the sugar moiety of genistein rhamnopyranoside, thus diversifying the sugar attached to genistein. Among the biosynthesized compounds, four of the 2-deoxy-*D*-glycopyranosides, 7-*O*-(2-deoxy- $\beta$ -*D*-*arabino*-hexopyranosyl)genistein, 4'-*O*-(2-deoxy- $\beta$ -*D*-*arabino*-hexopyranosyl)genistein, 4',7-bis-*O*-(2-deoxy- $\beta$ -*D*-*arabino*-hexopyranosyl)genistein, and 4',5,7-tri-*O*-(2-deoxy- $\beta$ -*D*-*arabino*-hexopyranosyl)genistein, and two of the *O*-methyl rhamnopyranosides, 7-*O*-(3-*O*-methyl- $\alpha$ -*L*-rhamnopyranosyl)genistein and 7-*O*-(4-*O*-methyl- $\alpha$ -*L*-rhamnopyranosyl)genistein, are novel compounds that have not been previously reported. The other glycopyranosides are 7-*O*-( $\alpha$ -*L*-rhamnopyranosyl)genistein, 7-*O*-( $\beta$ -*D*-glucopyranosyl)genistein, 4'-*O*-( $\beta$ -*D*-glucopyranosyl)genistein, and 4',7-*O*-( $\beta$ -*D*-glucopyranosyl)genistein. Microbial production of these novel compounds and other glucopyranosides is appreciable in shake flask culture. This study highlights the application of versatile enzymes in the production of diverse glycosides of medicinally important genistein, which can have positive impacts on a variety of molecular targets in future studies, as shown by genistein with remedied drawbacks.

Received 13th December 2016  
 Accepted 22nd February 2017

DOI: 10.1039/c6ra28145a

[rsc.li/rsc-advances](http://rsc.li/rsc-advances)

### 1. Introduction

Genistein is an isoflavone and a sub-class of large phenolic constituents of plant secondary metabolites, which shares the typical carbon skeleton of C6–C3–C6, but the phenyl substituent is linked to position 3 instead of position 2, as occurs in the flavan and flavone skeleton. These phenolic compounds are predominantly present in berries, wine, grains, nuts, soybeans, and other legumes including *Pueraria lobata* (Kudzu root), *Apios americana* (Peanuts), and *Cicer arietinum* (Chickpeas).<sup>1</sup> Soybean (*Glycine max*) is the largest contributor to dietary isoflavone, which mainly represents the core structures daidzein and genistein. Although legumes and beans are consumed worldwide, their consumption is higher in Asian countries. Moreover, consumption of beans and legumes is directly associated with lower incidences of cancers and oxidation-linked diseases in

older people in this region compared to those in the West.<sup>2–4</sup> More importantly, genistein is the predominant compound present in beans and legumes that structurally resembles the 17- $\beta$ -estradiol estrogen hormone and has various agonistic and antagonistic interactions with estrogens receptors. Thus, this molecule is also referred to as phytoestrogen.<sup>5–7</sup>

Among daidzein, glycitein, and their precursors formononetin and biochanin A, genistein has attracted more attention than other plant polyphenols because of its significant alignment with medical interests.<sup>5,8</sup> It has been investigated as a potential chemopreventive agent and a drug supplement for relieving the side effects of estrogen deficiency in menopausal women.<sup>9</sup> Various experiments have also shown its interference in biochemical pathways of living cells including biological targets such as protein tyrosine kinases (induce G2/M arrest and apoptosis),<sup>10</sup> topoisomerase II,<sup>11,12</sup> enzymes involved in phosphatidylinositol turnover,<sup>11</sup> estrogen receptors,<sup>12</sup> and ABC transporters.<sup>13</sup> Although genistein itself exhibits biological activity with potential medicinal applications, many derivatives of genistein have been biosynthesized to explore its application in diverse fields. Several drawbacks have been reported for genistein including poor water solubility and bioavailability.<sup>14,15</sup> To overcome these major drawbacks, derivatives of genistein have been produced by enzymatic

<sup>a</sup>Department of Life Science and Biochemical Engineering, Sun Moon University, 70 Sunmoon-ro 221, Tangjeong-myeon Asan-Si, Chungnam 31460, Republic of Korea. E-mail: sohng@sunmoon.ac.kr

<sup>b</sup>Department of BT-Convergent Pharmaceutical Engineering, Sun Moon University, 70 Sunmoon-ro 221, Tangjeong-myeon Asan-Si, Chungnam 31460, Republic of Korea

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6ra28145a



post modifications, such as glycosylation, methylation, hydroxylation, prenylation, and acylation, or by chemical approaches involving the fabrication of biodegradable nanoparticles to enhance its therapeutic effects.<sup>15–17</sup> Moreover, some glycosylated derivatives such as talosins (talose sugar-conjugated genistein) (talosin A and B) were found to have additional effects such as antifungal activity<sup>18</sup> and potent antidiabetic activity was shown by the 8-C-( $\beta$ -D-glucopyranosyl)genistein treatment on a diabetic animal.<sup>19</sup> In plants, genistein is present in the acylated and malonylated glycoside forms. Besides the core skeleton, other tailoring groups such as acetyl, malonyl, and glucosyl contribute to the bioavailability of genistein for maximum utilization in plants or following soy food consumption in humans.<sup>5</sup> Other examples include genistein glucopyranoside, which was found to have enhanced bioavailability compared to its aglycon.<sup>20</sup>

Although there are few reports about the biosynthesis of deoxy sugar-conjugated genistein, we are familiar with the antifungal activity of talosins where genistein is decorated with  $\alpha$ -6-deoxy L-talose sugar (talose).<sup>18</sup> Talose is an analogue of dTDP- $\alpha$ -L-rhamnose, in which the 4-hydroxyl group is epimerized. Many bacterial natural products contain highly modified and specified deoxy sugar residues, which are often crucial for biological activities but are not common in plant secondary metabolites.<sup>21,22</sup> Thus, the incorporation of sugar moieties in natural products could have different effects on their core biological activities.

In this experiment, we biosynthesized and structurally characterized ten different genistein glycopyranosides (Fig. 1) using versatile glycosyltransferases from plants and bacteria. The previously reported glycosyltransferase (GT) (AtUGT89C1)

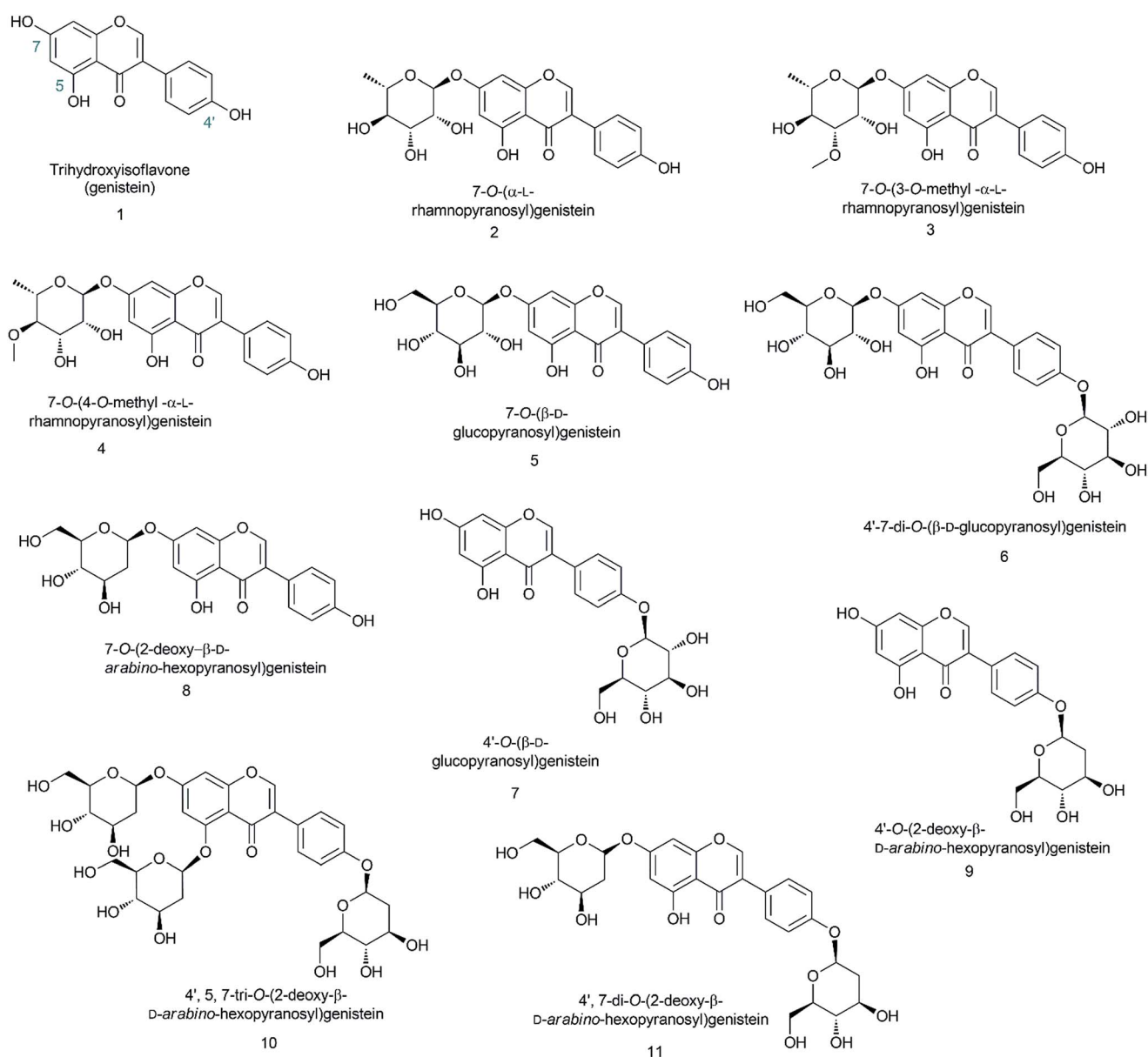


Fig. 1 Structures of all the biosynthesized genistein glycopyranosides in this study.



from *Arabidopsis thaliana*, which is capable of accepting diverse classes of flavonoids to conjugate microbial indigenous dTDP- $\alpha$ -L-rhamnose to the 7-hydroxy position,<sup>23</sup> was used along with YjiC from *Bacillus licheniformis* DSM-13.<sup>24-26</sup> Besides these two GTs, two sugar-O-methyltransferases (SOMTs) from the spinosyn producer *Saccharopolyspora spinosa* NRRL 18395 were used. Both *in vitro* and *in vivo* approaches were used to produce diverse genistein glycosides. Herein, genistein glucopyranosides were biosynthesized using the biotransformation approach, whereas genistein 2-deoxyglucopyranosides were enzymatically biosynthesized. All the compounds were structurally characterized using various analytical approaches. Their anti-fungal and anti-cancer activities were assessed and the results were compared with those of genistein and its derivatives.

## 2. Materials and methods

### 2.1. Chemicals and reagents

Authentic genistein was purchased from Sigma-Aldrich (St. Louis, MO, USA). dTDP- $\alpha$ -2-Deoxy-D-glucose was previously synthesized *via* a one-pot enzymatic approach.<sup>27</sup> All the other chemicals were of high-grade quality and were purchased from commercially available sources.

### 2.2. PCR, culture conditions, plasmids, and recombinant strains

*Escherichia coli* XL-1 blue (MRF) (Invitrogen, USA) was used for the cloning and propagation of recombinant plasmids. Recombinant plasmids were confirmed by enzyme restriction digestion and sequencing. *E. coli* BL21 (DE3) was used for protein expression and biotransformation. The pGEM@-T easy vector system (Promega, USA) was used for cloning polymerase chain reaction (PCR) fragments and sequencing. pET41b (+), pET28a (+), and pET32a (+) (Novagen, Germany) were used for

recombinant plasmid construction and protein expression. pET 41b-AtUGT89C1<sup>28</sup> and pET28a-YjiC<sup>25,26</sup> were used from the previously cloned construct. SOMTs spnH (GeneBank Accession no. AAG23269) and spnK (GeneBank Accession no. AAG23272) were PCR amplified from the total DNA of *S. spinosa*. The primers used for PCR amplification were as follows: spnK forward F: 5'-AAGCTTATGTCCACAACGCACGAG-3' and reverse R: 5'-GAATTCTCACTCGTCCTCCGCGCT-3' with restriction sites *EcoRI/HindIII*, and for spnH, forward F: 5'-CATATGATGCCCT CCCAGAACGCG-3' and reverse R: 5'-GATATCTCACCAGCTGCG GCGCCA-3' with restriction sites *EcoRV/NdeI*. *E. coli* strains were grown in Luria-Bertani (LB) broth or on an agar plate supplemented with an appropriate amount of antibiotics (ampicillin, 100  $\mu\text{g mL}^{-1}$  and kanamycin, 50  $\mu\text{g mL}^{-1}$ ). LB broth medium was used for the biotransformation assay. The recombinant plasmids and strains constructed for the experiment are listed in Table 1.

### 2.3. Protein expression

Recombinant strain-5 and strain-6 were used for protein expression. A quantity of 5 mL of these recombinant strains was cultured overnight in 15 mL falcon tubes in LB broth supplemented with appropriate antibiotics at 37 °C. Seed culture of 500  $\mu\text{L}$  was transferred to 100 mL LB medium in an *E. coli* culture flask followed by incubation at 37 °C until the culture was ready for isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) induction. When the culture optical density at 600 nm (OD<sub>600</sub>) reached ~0.6, 0.5 mM final IPTG concentration was added to induce protein expression in each flask. After IPTG induction, the culture was incubated at 20 °C. After 15 h of incubation at 20 °C, the cells were harvested and washed twice with 50 mM Tris-HCl containing 10% glycerol buffer followed by vortexing and centrifugation at 5000  $\times g$ . The cell pellets were then sonicated to harvest the protein. Cell lysates and pellets were

Table 1 Construction of the recombinant strains and plasmids with their descriptions along with the sources

| Recombinant strains                                      | Description  | Strain number | Source/references |
|--|--|---------------|-------------------|
| <i>E. coli</i> BL21 (DE3)                                | ompT hsdT hsdS (rB-mB)- gal (DE3)                                      |               | Novagen           |
| <i>E. coli</i> BL21 (DE3)/pET41b-AtUGT89C1               | BL21(DE3) carrying pET41b-AtUGT89C1                                    | Strain-1      | 23                |
| <i>E. coli</i> BL21 (DE3)/pET28a-YjiC                    | BL21(DE3) carrying pET28a-YjiC   | Strain-2      | 25 and 26         |
| <i>E. coli</i> BL21 (DE3)/pET32a                         | BL21(DE3) carrying pET32a  |               | This study        |
| <i>E. coli</i> BL21 (DE3)/p41b-AtUGT89C1 and pET32a-SpnK | BL21(DE3) carrying p41b-AtUGT89C1 and pET32a-SpnK                      | Strain-3      | This study        |
| <i>E. coli</i> BL21 (DE3)/p41b-AtUGT89C1 and pET32a-SpnH | BL21(DE3) carrying p41b-AtUGT89C1 and pET32a-SpnH                      | Strain-4      | This study        |
| <i>E. coli</i> BL21 (DE3)/pET32a-SpnH                    | BL21(DE3) pET32a-SpnH  | Strain-5      | This study        |
| <i>E. coli</i> BL21 (DE3)/pET32a-SpnK                    | BL21(DE3) pET32a-SpnK  | Strain-6      | This study        |
| Recombinant plasmids                                     | Description  |               | Source/references |
| pGEM@-T easy vector                                      | General cloning vector, T7 and Sp6 promoters, f1 ori, Amp <sup>r</sup> |               | Promega, USA      |
| pET32a-SpnK  | pET32a vector carrying SpnK  |               | This study        |
| pET32a-SpnH  | pET32a vector carrying SpnH  |               | This study        |
| pET41b-AtUGT89C1   | pET41b vector carrying AtUGT89C1                                       |               | 23 and 28         |



separated by high speed centrifugation for 30 min at  $16\,000 \times g$ . The soluble and insoluble fractions of proteins were assayed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for functional expression along with standard sized 100 kDa markers.

#### 2.4. *In vitro* glycosylation reaction for 2-deoxyglucopyranoside analogues

The crude protein of YjiC was used for the *in vitro* glycosylation reaction. The reaction mixture of 25 mL final volume contained 10 mM of genistein dissolved in dimethyl sulfoxide (DMSO), double the substrate concentration of thymidine diphosphate 2-deoxy- $\alpha$ -D-glucose (dTDP- $\alpha$ -2-deoxy-D-glucose) as the sugar donor, 10 mM of  $MgCl_2 \cdot 6H_2O$  containing 100 mM Tris-HCl buffer at pH 7.4, and 5 mL of crude protein. The reactions were incubated for 12 h at 37 °C and quenched using a double volume of chilled methanol. The reaction mixture of 500  $\mu$ L total volume containing all the other components at the same concentration with denatured enzyme was used as the control. Finally, the quenched reaction mixtures were centrifuged ( $16\,000 \times g$ , 20 min) to remove the protein precipitates. This reaction mixture (20  $\mu$ L) was directly injected into high performance liquid chromatography instrument connected to a photodiode array (HPLC-PDA) for chromatographic analyses. The same sample was subjected to liquid chromatography high resolution quadrupole time-of-flight electrospray ionization mass spectrometry (HR-QTOF ESI/MS) for product confirmation. Finally, the large volume reaction mixture was dried and concentrated to purify the genistein 2-deoxy-D-glucopyranosyl derivatives.

#### 2.5. Biotransformation reaction

Recombinant strain-1 and strain-2, *i.e.* *E. coli* BL21 (DE3) harbouring single GT, either AtUGT89C1 or YjiC, respectively, co-expressed with individual SOMTs, spnK and spnH (strain-3 and strain-4, respectively), were used for the biotransformation reaction (Table 1). Seed cultures of these strains were prepared in 100 mL culture volume in 500 mL fed batch flasks until high cell density growth ( $OD_{600\text{ nm}} \sim 1.5$ ) with appropriate antibiotics was achieved. The culture was then aseptically transferred to a 50 mL Falcon tube and centrifuged at  $2000 \times g$  for 10 min. The supernatant culture medium was discarded and viable cells were washed using autoclaved distilled water. Finally, these cells were used as the starter culture for biotransformation in 100 mL freshly prepared LB medium for uniform cell density. When the  $OD_{600\text{ nm}}$  reached  $\sim 0.6$ , 0.5 mM of IPTG was added to induce protein expression, followed by incubation for 12 h at 20 °C. Identical conditions were applied to prepare the culture for biotransformation in all the experiments. Recombinant strain-2 was used for glucopyranoside derivatives production, whereas strain-1 was used for the biosynthesis of the genistein rhamnopyranoside derivative. However, co-transformed recombinant strain-3 and strain-4 were used to biosynthesize the genistein *O*-methyl rhamnopyranosides. For all the biotransformation reactions, 200  $\mu$ M of genistein dissolved in DMSO was exogenously

applied to each flask. After 36 h of incubation at 20 °C, the cultures were mixed with a double volume of ethyl acetate and extracted for 1 h. The organic layer was dried using a rotary evaporator and the extracted compounds were dissolved in 1 mL methanol. These reaction mixtures were then analysed by HPLC-PDA and confirmed by HR-QTOF ESI/MS in the positive ion mode.

For the microbial production of genistein rhamnopyranoside, methyl rhamnopyranosides, and glucopyranosides, recombinant strains strain-1, strain-2, and strain-3, respectively, were used for the biotransformation reaction, in which genistein was exogenously applied (Fig. 1). The seed culture from each strain was individually prepared as abovementioned. Equal optical density seed cultures were transferred to four different flasks each containing 500 mL of LB medium. A total volume of 2-L shake flask cultures was prepared for each recombinant strain to obtain compounds for structural elucidation and product calculation. After proper induction by IPTG, the substrate was exogenously fed at the concentration of  $\sim 118.5 \mu\text{M}$  ( $\sim 32 \text{ mg L}^{-1}$ ) and incubated for 36 h. Compounds were extracted as abovementioned.

#### 2.6. Analytical methods

The culture extracts dissolved in methanol and the *in vitro* reaction mixtures were directly analysed by reverse-phase HPLC-PDA connected to a  $C_{18}$  column (Mightysil RP-18 GP (4.6  $\times$  250 mm, 5  $\mu$ m) using binary conditions of  $H_2O$  (0.05% trifluoroacetic acid buffer) and 100% acetonitrile at the flow rate of 1 mL  $\text{min}^{-1}$  for 30 min. The acetonitrile concentrations were as follows: 10% (0–2 min), 70% (20–24 min), 100% (24–28 min), and 50% (28–30). UV absorbance at 260 nm was used for each compound to monitor their HPLC chromatograms. For quantification of the products, a calibration curve of authentic 7-*O*-( $\beta$ -D-glucopyranosyl)genistein with different concentrations (10, 25, 50, 75, and 100  $\mu\text{g mL}^{-1}$ ) was used. The exact masses of the compounds were analysed by HR-QTOF ESI/MS [ACQUITY (UPLC, Waters, Milford, MA, USA)-SYNAPT G2-S (Waters)] in the positive ion mode. Each product was purified by preparative scale HPLC (prep-HPLC) with a  $C_{18}$  column [YMC-Pack ODS-AQ (250  $\times$  20 mm I.D., 10  $\mu$ m)] connected to a UV detector (260 nm) under the binary condition of  $H_2O$  (0.05% trifluoroacetic acid buffer) and 100% acetonitrile at the flow rate of 10 mL  $\text{min}^{-1}$  for 35 min. The acetonitrile concentrations were 20% (0–5 min), 50% (5–10 min), 70% (10–15 min), 90% (15–25 min), 50% (25–30 min), and 20% (20–35 min). These purified samples were dissolved in hexadeuterio dimethyl sulfoxide ( $DMSO-d_6$ ) and used as the nuclear magnetic resonance (NMR) solvent for further analyses. NMR studies were performed using a Varian Unity INOVA 800 MHz spectrometer (Varian, USA). One-dimensional NMR ( $^1\text{H-NMR}$  and  $^{13}\text{C-NMR}$ ) followed by two-dimensional NMR (HMBC – heteronuclear multiple-quantum correlation spectroscopy) were performed to elucidate the exact structure of the glycosides.

#### 2.7. Relative solubility study

For the determination of the relative water solubility of the glucopyranosides, each strain (strain-1, strain-2, strain-3, strain-4, strain-5, and strain-6) were cultured in 50 mL culture medium



and genistein was exogenously fed at the final concentration of 0.2 mM. This was performed similar to the biotransformation strategy. After 36 h incubation, a double volume of ethyl acetate was added to extract the compounds followed by concentrating the supernatant on a rotary evaporator. The concentrated sample was then mixed in a 1 : 2 ratio of ethyl acetate and water at room temperature (24 °C). The mixture was vortexed well and then subjected to HPLC-PDA of each fraction (aqueous and organic portions). For the 2-deoxyglucosides, an identical laboratory scale *in vitro* reaction was performed at 200  $\mu$ L, as explained in the Materials and method section. The reaction was incubated for 2 h and quenched by methanol. The reaction mixture was dried and mixed with a 1 : 2 ratio of ethyl acetate and water, as abovementioned. Finally, the separate fractions were analysed by HPLC-PDA. The solubility enhancement was compared with the standard genistein in the water fraction.

## 2.8. Biological activity test

The anti-fungal and anti-cancer activities of the synthesized compounds were tested and compared with those of aglycon. For antifungal activity, the fungal strain *Aspergillus nidulans* ATCC 58396 was considered, whereas Nystatin A1 was used as a control antifungal agent. Paper discs of 6 mm diameter were prepared and autoclaved. Different concentrations (from 0.5 mg mL<sup>-1</sup> to maximum 20 mg mL<sup>-1</sup>) of the genistein analogues purified for NMR analyses were used. From the stock, 500  $\mu$ g mL<sup>-1</sup> of each analogue was prepared including genistein dissolved in DMSO to test the anti-fungal inhibition zone and 100  $\mu$ g mL<sup>-1</sup> of Nystatin A1 (high concentration) was used to observe a clear inhibition zone. Potato dextrose agar (PDA) plates were prepared using ready to use media from Sigma. The fungal strain was taken up by an inoculation loop into 100  $\mu$ L of autoclaved distilled water in a tube and mixed well. This mixture was then spread and plated in the PDA plate. Then, each paper disc was dipped into a 500  $\mu$ g mL<sup>-1</sup> solution of genistein analogues, genistein, Nystatin A1, and DMSO (100%) as a control. Finally, they were placed in the fungus-plated PDA plates. These plates were incubated at 28 °C and observed at different time intervals until a clear inhibition zone was seen.

Similarly, for anti-cancer activity, both the cell culture and cell growth assay, AGS (gastric carcinoma cells, ATCC CRL-1739) were maintained in RPMI 1640 medium containing 10% fetal bovine serum (FBS). B16F10 melanoma (skin cancer cell, ATCC CRL-6475), HeLa (cervical carcinoma cells, ATCC CCL-2), and HepG2 hepatocarcinoma cells (liver cancer cells, ATCC 59194) were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% FBS. All the cells were maintained at 37 °C in a humidified 5% CO<sub>2</sub> incubator. For the cell growth assay, various cancer cells were plated at  $2 \times 10^3$  cells per well in 96-well culture plates. The genistein derivatives were added to each well at various concentrations and the cells were incubated for 72 h. Cell growth was measured using a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) colorimetric assay. A 50  $\mu$ L of MTT (2 mg mL<sup>-1</sup> stock solution) was added and the plates were incubated for an additional 4 h. After the removal of the medium, 100  $\mu$ L of dimethyl sulfoxide

(DMSO) was added. The absorbance was measured at 540 nm using a microplate spectrophotometer (Thermo Scientific Multiskan® Spectrum).

## 3. Results

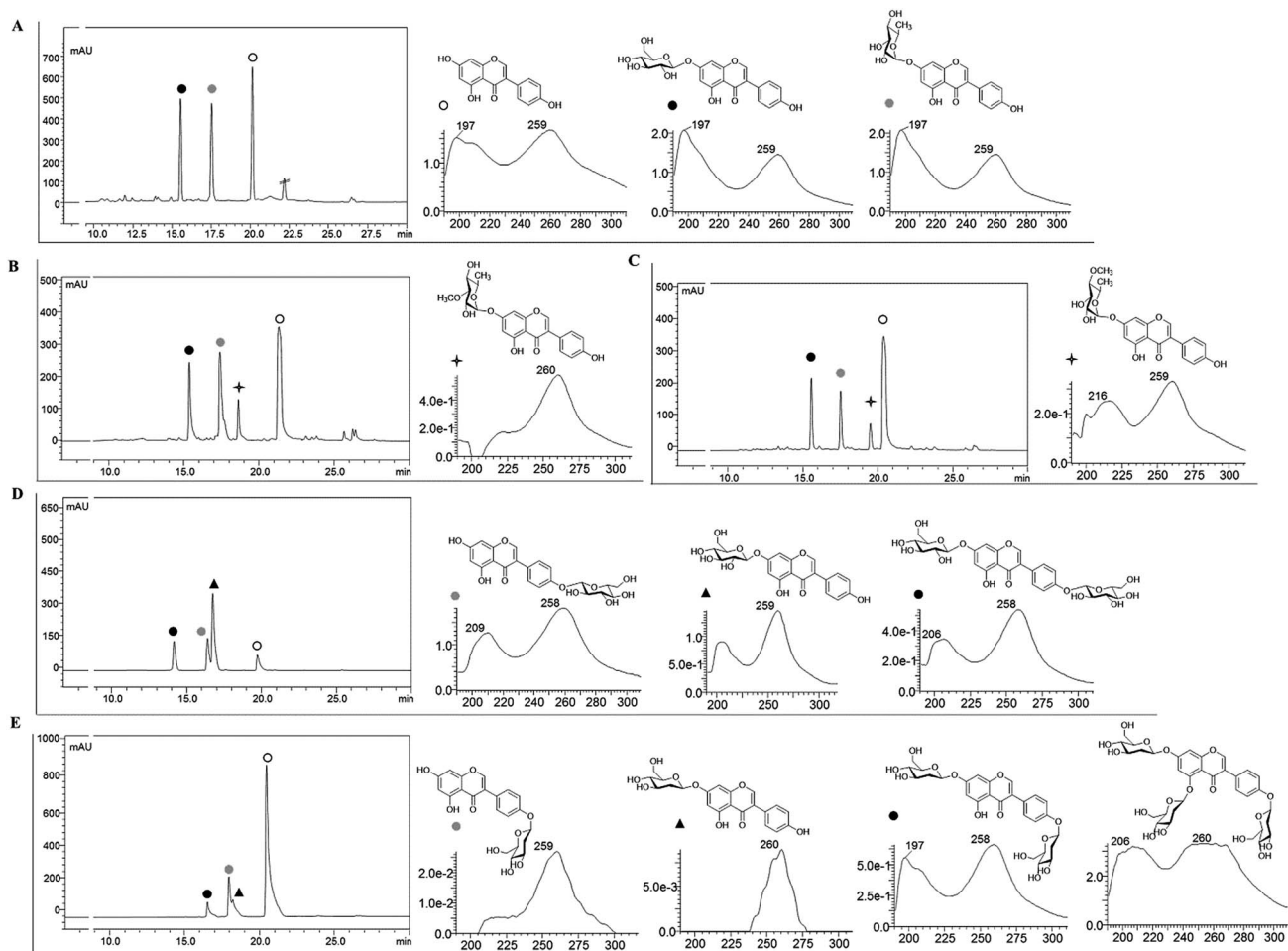
### 3.1. Biosynthesis and identification of genistein rhamnopyranosides

Recombinant strain-1 was used to biosynthesize the rhamnosyl derivative of genistein. The culture was prepared to exogenously feed genistein, as explained in the Materials and methods section. The biotransformation reaction mixture of this strain harbouring only AtUGT89C1 showed two distinct product peaks at the retention times ( $t_R$ ) of 17.6 min and 15.7 min with reference to the standard  $t_R$  at 21 min (Fig. 2A), which were further analysed by HR-QTOF ESI/MS. Mass analysis showed the exact mass  $[M + H]^+ m/z^+$  of  $\sim 417.1192$ , which resembles the genistein rhamnopyranoside calculated mass  $m/z^+$  of 417.1180 for the compound with  $t_R$  17.6 min and  $[M + H]^+ m/z^+$  of  $\sim 433.1129$ , which resembles the calculated mass of the genistein glucopyranoside with the mass  $m/z^+$  of  $\sim 433.1135$  of the compound with  $t_R$  15.7 min (Fig. S3A and B†). These mass spectra results are consistent with the mass of rhamnose ( $t_R$  of 17.6) and glucose ( $t_R$  of 15.7 min)-conjugated genistein. Although AtUGT89C1 previously transferred a single rhamnose sugar to the 7-OH position of flavonoids,<sup>23</sup> herein, the result with genistein was different. Thus, both compounds were purified and further characterized *via* various types of NMR analyses. The <sup>1</sup>H and <sup>13</sup>C-NMR chemical shifts for these compounds are tabulated in Tables 2 and 3.

The <sup>1</sup>H NMR spectra of the compound with a  $t_R$  of 17.6 min (Fig. 2A) display an anomeric proton (1''-H) with a doublet at  $\delta$  5.57 ppm (d,  $J = 1.8$  Hz, 1H) and the other sugar spectra below 5.0 ppm confirm the conjugation of a sugar moiety in the alpha ( $\alpha$ ) conformation (Fig. S4A†). Glycosylation is clearly observed at the 7<sup>th</sup> position since there is absence of the 7-OH proton in the <sup>1</sup>H spectrum, which is lost as water molecules during glycosylation. This result was further supported by the HMBC analysis. In HMBC, a cross peak between an anomeric proton and C-7 is distinctly observed (Fig. S4C†), confirming the position of glycosylation at the 7-OH group. The presence of a doublet peak at  $\delta$  1.12 ppm (d,  $J = 6.0$ ) suggests a methyl group in the sugar moiety, which is further confirmed by the presence of a methyl peak at  $\delta$  17.9 ppm in the <sup>13</sup>C NMR spectra (Fig. S4A and B†). These peaks are characteristic of 6-deoxy sugars. Since AtUGT89C1 was previously characterized as a rhamnosyl-transferase and the presence of these specific peaks in the <sup>1</sup>H and <sup>13</sup>C-NMR spectra and HMBC cross peaks (Fig. S4C†) between the anomeric proton (5.58 ppm) and C7 (162.2 ppm) further provide evidence for the attachment of a rhamnose moiety in genistein at the 7-OH position, thus confirming the molecule to be 7-O-( $\alpha$ -L-rhamnopyranosyl)genistein (2). These NMR spectra were also compared with those of previous reports.<sup>29,30</sup>

The new peak observed at the  $t_R$  of 15.7 min in the biotransformation reaction mixture of the same strain with a molecular mass similar to that of genistein glucopyranoside





**Fig. 2** HPLC-PDA chromatograms analyzed at 260 nm and UV absorbance of the different biotransformation reaction mixtures catalysed by GTs along with SOMTs. (A) Reaction mixture catalyzed by AtUGT89C1, where the black filled circle is 7-*O*-( $\beta$ -D-glucopyranosyl)genistein, gray filled circle is 7-*O*-( $\alpha$ -L-rhamnopyranosyl)genistein, and the open circle is the genistein standard. (B) Reaction mixture catalyzed by AtUGT89C1 along with SpnK where the four-pointed star indicates the specific product 7-*O*-(3-*O*-methyl- $\alpha$ -L-rhamnopyranosyl)genistein. (C) Reaction mixture catalyzed by AtUGT89C1 along with SpnH where the four-pointed star indicates the specific product 7-*O*-(4-*O*-methyl- $\alpha$ -L-rhamnopyranosyl)genistein. (D) Reaction mixture catalyzed by Yjic. Black filled circle is 4',7-di-*O*-( $\beta$ -D-glucopyranosyl)genistein, gray filled circle is 7-*O*-( $\beta$ -D-glucopyranosyl)genistein, and black filled triangle is 4'-*O*-( $\beta$ -D-glucopyranosyl)genistein. (E) *In vitro* reaction mixture catalysed by Yjic. Black filled circle is 4',7-di-*O*-(2-deoxy-*arabino*- $\beta$ -D-hexopyranosyl)genistein, gray filled circle is 7-*O*-(2-deoxy-*arabino*- $\beta$ -D-hexopyranosyl)genistein, and black filled triangle is 4'-*O*-(2-deoxy-*arabino*- $\beta$ -D-hexopyranosyl)genistein.

was also investigated by NMR spectroscopy. The anomeric doublet proton was detected at  $\delta$  5.07 ppm (d,  $J = 7.0$  Hz, 1H) (Fig. S5A<sup>†</sup>), suggesting sugar attachment at the (beta)  $\beta$ -configuration. The other protons of glucose were intact in the region from 3.0 to 4.5 ppm.<sup>30</sup> The <sup>13</sup>C NMR analysis distinctly shows the C-6'' of glucose at  $\delta$  62.6 ppm<sup>31</sup> and an anomeric carbon at  $\delta$  99.6 ppm. This data was compared with that of a previous report.<sup>32</sup> HMBC shows a relationship to an anomeric proton, with the C-7 position of genistein providing evidence of sugar attachment at the 7-hydroxy position of aglycon (Fig. S5C<sup>†</sup>). Further cross peaks of the anomeric protons (5.08 ppm) with C-2'' (76.85 ppm), C-3'' (73.53 ppm), and C-5'' (77.64 ppm) were also observed (Fig. S5D<sup>†</sup>). All these results confirmed the new peak at the  $t_R$  of 15.7 min to be 7-*O*-( $\beta$ -D-glucopyranosyl)genistein (5).

Microbial production of these compounds was performed for product calculation. From 32 mg L<sup>-1</sup> (118.5  $\mu$ M) of substrate,

12 mg L<sup>-1</sup> mg<sup>-1</sup> (28.8  $\mu$ M) of 7-*O*-( $\alpha$ -L-rhamnopyranosyl)genistein and 11 mg L<sup>-1</sup> (25.4  $\mu$ M) of 7-*O*-( $\beta$ -D-glucopyranosyl)genistein yields were recovered in 36 h from the biotransformation strain-1 culture (Fig. 3).

### 3.2. Biosynthesis and identification of genistein-methylated rhamnopyranosides

To synthesize methylated rhamnose-conjugated genistein molecules, two SOMTs, spnK and spnH (sugar 3'-*O*-methyltransferase and sugar 4'-*O*-methyltransferase from the spinosyn producer *S. spinosa* NRRL 18395, respectively), were co-expressed with AtUGT89C1 (strain-3 and strain-4, respectively) (Table 1).

In the HPLC-PDA chromatogram of the reaction mixtures of strain-3 and strain-4, beside genistein glucopyranoside and





**Table 2**  $^1\text{H-NMR}$  spectra of all the biosynthesized genistein glycopyranosides. All the samples were dissolved in  $\text{DMSO}-d_6$  for NMR analyses

| Compounds             | 2                    | 3                    | 4                    | 9                                     | 8                                     | 11                                    | 5                    | 5                    | 6                    | 7                    |
|-----------------------|----------------------|----------------------|----------------------|---------------------------------------|---------------------------------------|---------------------------------------|----------------------|----------------------|----------------------|----------------------|
| Position              | $^1\text{H}$         | $^1\text{H}$         | $^1\text{H}$         | $^1\text{H}$                          | $^1\text{H}$                          | $^1\text{H}$                          | $^1\text{H}$         | $^1\text{H}$         | $^1\text{H}$         | $^1\text{H}$         |
| 5-OH                  | 12.94 (s)            | 12.94 (s)            | 12.94 (s)            | 12.93 (s)                             | 12.89 (s)                             | 12.87                                 | 12.95 (s)            | 12.94 (s)            | 12.82 (s)            | 12.91 (s)            |
| 7-OH                  | —                    | —                    | —                    | —                                     | —                                     | —                                     | —                    | —                    | —                    | —                    |
| 4'-OH                 | 9.64 (s)             | 9.63 (s)             | 9.65 (s)             | 9.68 (s)                              | —                                     | —                                     | —                    | 9.64 (s)             | —                    | —                    |
| H-2                   | 8.4 (s)              | 8.43 (s)             | 8.42 (s)             | 8.42 (s)                              | 8.31 (s)                              | 8.47 (s)                              | 8.42 (s)             | 8.44 (s)             | 8.50 (s)             | 8.38 (s)             |
| H-2'                  | 7.39 (m)             | 7.40 (m)             | 7.39 (m)             | 7.40 (m)                              | 7.48 (m)                              | 7.51 (d)                              | 7.40 (m)             | 7.41 (m)             | 7.53 (m)             | 7.50 (m)             |
| H-6'                  | 7.39 (m)             | 7.40 (m)             | 7.39 (m)             | 7.40 (m)                              | 7.48 (m)                              | 7.51 (d)                              | 7.40 (m)             | 7.41 (m)             | 7.53 (m)             | 7.50 (m)             |
| H-3'                  | 6.83 (m)             | 6.83 (m)             | 6.83 (m)             | 6.83 (m)                              | 7.07 (m)                              | 7.09 (d)                              | 6.84 (m)             | 6.84 (m)             | 7.11 (m)             | 7.10 (m)             |
| H-5'                  | 6.83 (m)             | 6.83 (m)             | 6.83 (m)             | 6.83 (m)                              | 7.07 (m)                              | 7.09 (d)                              | 6.84 (m)             | 6.84 (m)             | 7.11 (m)             | 7.10 (m)             |
| H-8                   | 6.74 (d, $J = 2.1$ ) | 6.76 (d, $J = 2.2$ ) | 6.75 (d, $J = 2.2$ ) | 6.71 (d, $J = 2.2$ )                  | 6.30 (s)                              | 6.73 (d, $J = 2.3$ )                  | 6.72 (d, $J = 2.2$ ) | 6.73 (d, $J = 2.4$ ) | 6.74 (d, $J = 2.4$ ) | 6.40 (d, $J = 2.1$ ) |
| H-6                   | 6.48 (d, $J = 2.1$ ) | 6.51 (d, $J = 2.1$ ) | 6.49 (d, $J = 2.2$ ) | 6.47 (d, $J = 2.2$ )                  | 6.15 (s)                              | 6.48 (d, $J = 2.2$ )                  | 6.48 (d, $J = 2.1$ ) | 6.48 (d, $J = 2.3$ ) | 6.49 (d, $J = 2.2$ ) | 6.23 (d, $J = 2.1$ ) |
| H-1''                 | 5.57 (d, $J = 1.8$ ) | 5.63 (d, $J = 1.9$ ) | 5.57 (d, $J = 1.9$ ) | 5.42 (dd, $J = 9.7, 1.9, 1\text{H}$ ) | 5.29 (dd, $J = 9.7, 1.9, 1\text{H}$ ) | 5.30 (dd, $J = 9.8, 2.1, 1\text{H}$ ) | 5.07 (d, $J = 7.0$ ) | 5.07 (d, $J = 7.4$ ) | 5.08 (d, $J = 7.4$ ) | 4.92 (d, $J = 7.5$ ) |
| H-1'''                | —                    | —                    | —                    | —                                     | —                                     | 5.43 (dd, $J = 9.8, 2.1, 1\text{H}$ ) | —                    | —                    | 4.92 (d, $J = 7.4$ ) | —                    |
| H-2''                 | 3.0–4.0              | 4.09                 | 3.0–4.0              | 3.0–4.0                               | 3.0–4.0                               | 3.0–4.0                               | 3.0–4.0              | 3.0–4.0              | 3.0–4.0              | 3.0–4.0              |
| H-2'''                | —                    | —                    | —                    | —                                     | —                                     | 3.0–4.0                               | 3.0–4.0              | —                    | 3.0–4.0              | —                    |
| H-3''                 | 3.0–4.0              | 3.45                 | 3.0–4.0              | 3.0–4.0                               | 3.0–4.0                               | 3.0–4.0                               | 3.0–4.0              | 3.0–4.0              | 3.0–4.0              | 3.0–4.0              |
| H-3'''                | —                    | —                    | —                    | —                                     | —                                     | 3.0–4.0                               | 3.0–4.0              | —                    | 3.0–4.0              | —                    |
| H-4''                 | 3.0–4.0              | 3.40                 | 3.0–4.0              | 3.0–4.0                               | 3.0–4.0                               | 3.0–4.0                               | 3.0–4.0              | 3.0–4.0              | 3.0–4.0              | 3.0–4.0              |
| H-4'''                | —                    | —                    | —                    | —                                     | —                                     | 3.0–4.0                               | —                    | —                    | 3.0–4.0              | —                    |
| H-5''                 | 3.0–4.0              | 3.40                 | 3.0–4.0              | 3.0–4.0                               | 3.0–4.0                               | 3.0–4.0                               | 3.0–4.0              | 3.0–4.0              | 3.0–4.0              | 3.0–4.0              |
| H-5'''                | —                    | —                    | —                    | —                                     | —                                     | 3.0–4.0                               | —                    | —                    | 3.0–4.0              | —                    |
| 6''-CH <sub>2</sub>   | —                    | —                    | —                    | —                                     | —                                     | 3.0–4.0                               | —                    | —                    | 3.0–4.0              | —                    |
| 6'''-CH <sub>2</sub>  | —                    | —                    | —                    | —                                     | —                                     | 3.0–4.0                               | —                    | —                    | 3.0–4.0              | —                    |
| 6''-CH <sub>3</sub>   | 1.12 (d, $J = 6.0$ ) | 1.12 (d, $J = 6.1$ ) | 1.15 (d, $J = 6.3$ ) | —                                     | —                                     | —                                     | —                    | —                    | —                    | —                    |
| OCH <sub>3</sub> -3'' | —                    | 3.40 (s)             | —                    | —                                     | —                                     | —                                     | —                    | —                    | —                    | —                    |
| OCH <sub>3</sub> -4'' | —                    | —                    | 3.47 (s)             | —                                     | —                                     | —                                     | —                    | —                    | —                    | —                    |
| 2-Deoxy               | —                    | —                    | —                    | 2.19 (ddd, $J = 12, 5, 2.1$ )         | 2.21 (ddd, $J = 12, 5, 1.9$ )         | 2.20 (m)                              | —                    | —                    | —                    | —                    |
| 2-Deoxy               | —                    | —                    | —                    | 1.60 (m)                              | 1.60 (td)                             | 1.60 (m)                              | —                    | —                    | —                    | —                    |

Table 3  $^{13}\text{C}$ -NMR spectra of all the biosynthesized genistein glycopyranosides. All the samples were dissolved in DMSO- $d_6$  for NMR analyses

| Compounds             | 2               | 3               | 4               | 9               | 8               | 11              | 5               | 5               | 6               | 7               |
|-----------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Position              | $^{13}\text{C}$ | $^{13}\text{C}$ | $^{13}\text{C}$ | $^{13}\text{C}$ | $^{13}\text{C}$ | $^{13}\text{C}$ | $^{13}\text{C}$ | $^{13}\text{C}$ | $^{13}\text{C}$ | $^{13}\text{C}$ |
| 5                     | 161.7           | 162.1           | 162.1           | 164.3           | 164.7           | 162.9           | 161.7           | 162.1           | 163.5           | 162.4           |
| 7                     | 161.8           | 162             | 161.4           | 162.8           | 162.3           | 162             | 163             | 163.4           | 162             | 165.2           |
| 4'                    | 157.5           | 157.9           | 157.8           | 157             | 157             | 157.2           | 157.5           | 157.7           | 157.7           | 157.7           |
| 2                     | 154.6           | 155             | 155             | 155             | 154.4           | 155.5           | 154.5           | 155             | 155.5           | 154.8           |
| 2'                    | 130.2           | 130.6           | 130.6           | 130.6           | 130.5           | 130.5           | 130.2           | 130.6           | 130.5           | 130.5           |
| 6'                    | 130.2           | 130.6           | 130.6           | 130.6           | 130.4           | 130.5           | 129.8           | 130.2           | 130.5           | 130.5           |
| 3'                    | 115.1           | 115.5           | 115.5           | 115.5           | 116.3           | 116.3           | 115.1           | 115.5           | 116.5           | 116.5           |
| 5'                    | 115.1           | 115.5           | 115.5           | 115.5           | 116.2           | 116.3           | 115.1           | 115.5           | 116.5           | 116.5           |
| 8                     | 94.6            | 95.1            | 95.1            | 95              | 94              | 95              | 94.5            | 95              | 95              | 94.2            |
| 6                     | 99.7            | 100.1           | 100.1           | 106.5           | 104.7           | 106.5           | 99.6            | 106.5           | 100.7           | 104.7           |
| 1''                   | 98.4            | 98.8            | 98.5            | 98.8            | 97              | 96.9            | 99.9            | 100             | 100.1           | 99.6            |
| 1'''                  | —               | —               | —               | —               | —               | 96.8            | —               | —               | 100.3           | —               |
| 2''                   | 70.3            | 66.3            | 69.1            | 40.2            | 40.2            | 40.3            | 76.4            | 73.5            | 76.8            | 77.1            |
| 2'''                  | —               | —               | —               | —               | —               | 39.7            | —               | —               | 77.1            | —               |
| 3''                   | 69.8            | 57              | 82.4            | 71.6            | 70.6            | 70.6            | 73.1            | 76.8            | 73.5            | 73.7            |
| 3'''                  | —               | —               | —               | —               | —               | 70.8            | —               | —               | 73.7            | —               |
| 4''                   | 71.6            | 70.8            | 60.5            | 70.8            | 71.5            | 71.5            | 69.6            | 77.6            | 77.6            | 77.5            |
| 4'''                  | —               | —               | —               | —               | —               | 71.6            | —               | —               | 77.5            | —               |
| 5''                   | 70.1            | 70.4            | 70.5            | 77.7            | 77.9            | 77.7            | 77.2            | 70              | 70.1            | 70.1            |
| 5'''                  | —               | —               | —               | —               | —               | 77.9            | —               | —               | 70              | —               |
| 6''-CH <sub>2</sub>   | —               | —               | —               | 61.2            | 61.2            | 61.2            | 60.6            | 61.1            | 61.1            | 61.1            |
| 6'''-CH <sub>2</sub>  | —               | —               | —               | —               | —               | 61.2            | —               | —               | 61              | —               |
| 6''-CH <sub>3</sub>   | 17.9            | 18.3            | 18.3            | —               | —               | —               | —               | —               | —               | —               |
| OCH <sub>3</sub> -3'' | —               | 57              | —               | —               | —               | —               | —               | —               | —               | —               |
| OCH <sub>3</sub> -4'' | —               | —               | 60.5            | —               | —               | —               | —               | —               | —               | —               |
| 2-Deoxy               | —               | —               | —               | 40.2            | 40.2            | 40.3            | —               | —               | —               | —               |
| 2-Deoxy               | —               | —               | —               | 40.2            | 40.2            | 39.7            | —               | —               | —               | —               |

rhamnopyranoside peaks, two new product peaks were observed at the  $t_R$  of 18.9 min and 19.6 min (Fig. 2B and C), respectively, which were further analyzed by HR-QTOF ESI/MS. The mass fragments of the new peaks are consistent with the *O*-methyl rhamnose-conjugated genistein. The exact mass observed is approximately the same  $[\text{M} + \text{H}]^+ m/z^+$  of  $\sim 431.1355$  for both peaks ( $t_R$  of 18.9 min and 19.6 min) and are consistent with the calculated mass for the genistein methyl rhamnopyranoside  $m/z^+$  of  $\sim 431.1342$  (S3 C and D). These compounds were purified by preparative HPLC for structural characterization.

While analysing the  $^1\text{H}$  NMR spectra of the compound at the  $t_R$  of 18.9 min from strain-3, an anomeric proton with a doublet at  $\delta$  5.63 ppm ( $d, J = 1.9$  Hz, 1H) was obtained, representing the  $\alpha$ -configuration of a sugar attachment, whereas the peak for the 7-OH group was missing. Moreover, there was a characteristic methyl group at 1.12 ppm ( $d, J = 6.1$ ) as in 7-*O*-( $\alpha$ -L-rhamnopyranosyl)genistein (2) (Fig. S6A†). As explained and analysed in previous reports,<sup>29,33,34</sup> a characteristic proton peak for 3''-O-CH<sub>3</sub> was obtained at 3.40 ppm. In the  $^{13}\text{C}$  analysis, a rhamnose methyl peak was obtained at 18.3 ppm, whereas that for 3''-O-CH<sub>3</sub> was obtained at 57 ppm (Fig. S6A and B†). With reference to HMBC, the cross peaks between 1''-H (5.63 ppm) and C-7 (162 ppm) (Fig. S6C†) confirm the conjugation of rhamnose at the 7-hydroxy position of genistein, whereas the cross peaks between 1''-H (5.63 ppm) and C3'' (80.6 ppm) and between 3''-O-CH<sub>3</sub> (3.40 ppm) and C-3'' (80.5 ppm) (Fig. S6D†) further confirm the *O*-methylation at the 3''-OH position of the rhamnose moiety.

Thus, the compound was determined to be 7-*O*-(3-*O*-methyl  $\alpha$ -L-rhamnopyranosyl)genistein (3). The chemical shifts for the  $^1\text{H}$  and  $^{13}\text{C}$  NMR of these compounds are tabulated in Tables 2 and 3. In a previous report, spnK was identified as SOMT catalyzing methyl transfer reactions from SAM to the 3''-OH position of rhamnose moieties.<sup>33</sup> Thus, we were unable to find methylated glucose-conjugated genistein derivatives.

Similarly, the  $^1\text{H}$  NMR spectra of the compound obtained from strain-4 show an anomeric proton with a doublet at  $\delta$  5.57 ppm ( $d, J = 1.9$  Hz, 1H), representing the  $\alpha$ -configuration of a sugar attachment with the peak of a 7-OH group. There was also a characteristic methyl group at  $\delta$  1.15 ppm ( $d, J = 6.3$ ). The characteristic peak for 4''-O-CH<sub>3</sub> was obtained at  $\delta$  3.47 ppm (Fig. S7A†). In the  $^{13}\text{C}$  analysis for this product, the peak for the methyl group of rhamnose was obtained at  $\delta$  18.3 ppm. The 4''-C was shifted to  $\delta$  82.4 and the 4''-O-CH<sub>3</sub> was located at 60.5 ppm (Fig. S7B†), similar to previous reports.<sup>29,33,34</sup> The HMBC also supports these results, showing cross peaks between an anomeric proton 1''-H (5.56 ppm) and C-7 (162 ppm) (Fig. S7C†), which confirm a rhamnose attachment at the 7-OH position of genistein. The HMBC cross peak between 4''-O-CH<sub>3</sub> (3.47 ppm) and C-4'' (60.5 ppm) confirms the methylation of the rhamnose moiety at the 4''-OH position (Fig. S7D†). These results confirm the product to be 7-*O*-(4-*O*-methyl  $\alpha$ -L-rhamnopyranosyl)genistein (4).

Biotransformation resulted in the production of two novel *O*-methyl rhamnose analogues of genistein. To the best of our





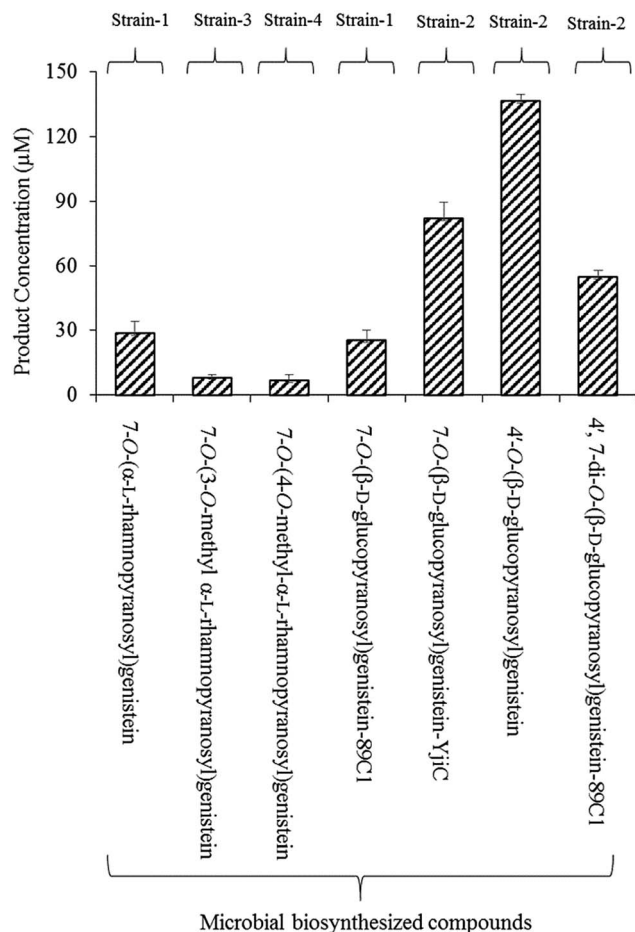


Fig. 3 Production of glycopyranosides in micro molar ( $\mu\text{M}$ ) amounts via the biotransformation reactions catalysed by different recombinant strains (strain-1, strain-2, strain-3, and strain-4). The metabolites were harvested after 36 h incubation at 20 °C. The error bars indicate the standard deviation values for three independent experiments.

knowledge, these sugar *O*-methyltransferases were used for the first time in the generation of methylated rhamnopyranosyl flavonoids, beside glycosylated phenolic compound polycarins.<sup>35</sup> During the microbial production of these compounds using GT and SOMTs co-transformation from 12 mg L<sup>-1</sup> (28.8  $\mu\text{M}$ ) of 7-*O*-( $\alpha$ -L-rhamnopyranosyl)genistein, 3.56 mg L<sup>-1</sup> (8.27  $\mu\text{M}$ ) of 7-*O*-(3-*O*-methyl  $\alpha$ -L-rhamnopyranosyl)genistein and 3.05 mg L<sup>-1</sup> (7  $\mu\text{M}$ ) of 7-*O*-(4-*O*-methyl  $\alpha$ -L-rhamnopyranosyl)genistein yields were recovered (Fig. 3).

### 3.3. Biosynthesis and identification of genistein glycopyranosides

7-*O*-( $\beta$ -D-Glucopyranosyl)genistein (5) was biosynthesized from strain-1 harboring AtUGT89C1, as explained above. However, for other genistein glycopyranosides, recombinant strain-2 harboring YjiC was used, as previously reported.<sup>26</sup> A similar approach was applied in this experiment to biosynthesize particular genistein glycopyranosides and elucidate the structures by NMR. The biotransformation reaction was performed as explained in the Materials and methods section. After 36 h of biotransformation reaction, the HPLC-PDA samples were

prepared to check the production of glucopyranosides. The HPLC-PDA chromatogram showed three distinct product peaks at the  $t_{\text{R}}$  of 14.2 min, 16.5 min, and 16.9 min with reference to the standard peak at 20 min (Fig. 2D). These peaks were subjected to HR-QTOF ESI/MS, resulting in mass fragmentation of genistein di-glucopyranoside  $[\text{M} + \text{H}]^+$   $m/z^+$  of  $\sim 595.1660$ , for which, the calculated mass was  $m/z^+$  of  $\sim 595.1663$  for the  $t_{\text{R}}$  14.2 compound, and  $[\text{M} + \text{H}]^+$   $m/z^+$  of  $\sim 433.1143$  was obtained for the  $t_{\text{R}}$  16.5 min compound, whose calculated mass matches with the genistein glucopyranoside  $m/z^+$  of  $\sim 433.1135$ . Similarly, the obtained mass fragment  $[\text{M} + \text{H}]^+$   $m/z^+$  of  $\sim 433.1148$  resembles the calculated mass for the second genistein glucopyranoside  $m/z^+$  433.1135 for the  $t_{\text{R}}$  16.9 min compound (Fig. S3E–G<sup>†</sup>). Each of the glucopyranosides of genistein produced by YjiC was further structurally elucidated using NMR studies. The <sup>1</sup>H and <sup>13</sup>C NMR chemical shifts for these compounds are also tabulated in Tables 2 and 3.

The compound at the  $t_{\text{R}}$  of 14.2 min showed two doublet anomeric protons at  $\delta$  5.08 ppm (d,  $J = 7.6$  Hz, 1H) and  $\delta$  4.92 ppm (d,  $J = 7.4$  Hz, 1H), with two sugar attachments in the  $\beta$ -configuration (Fig. S8A<sup>†</sup>). Furthermore, this result was supported by the HSQC cross peaks between the anomeric carbon and anomeric protons and HMBC cross peaks, H-1''-H (5.08 ppm) with C-7 (163.5 ppm) and H-1''' (4.93 ppm) with C-4' (157.8 ppm), confirmed the attachment of sugars at the 7-OH and 4'-OH positions of genistein (Fig. S8C<sup>†</sup>). The other peaks for the genistein and glucose moieties were found to be present in their respective positions in the <sup>1</sup>H and <sup>13</sup>C NMR spectra. Thus, this compound was elucidated to be 4',7-di-*O*-( $\beta$ -D-glucopyranosyl)genistein (6). Similarly, for the compound at the  $t_{\text{R}}$  of 16.5 min and  $t_{\text{R}}$  of 16.9 min, doublet anomeric protons at  $\delta$  4.92 ppm (d,  $J = 7.5$  Hz, 1H) and  $\delta$  5.07 ppm (d,  $J = 7.7$  Hz, 1H) were observed, which indicated sugar attachments in the  $\beta$ -confirmation in both molecules (Fig. S9A<sup>†</sup>). Since both of these compounds were genistein glycopyranosides, the HMBC cross peak for the compound at  $t_{\text{R}}$  16.5 min showed a relationship between H-1'' (4.92 ppm) and C-4' (157.7 ppm), confirming the compound to be 4'-*O*-( $\beta$ -D-glucopyranosyl)genistein (7) (Fig. S9C<sup>†</sup>). The cross peak for the compound at  $t_{\text{R}}$  16.9 min showed a relationship between H-1''' (5.07 ppm) and C-7 (163.4 ppm), confirming the compound to be 7-*O*-( $\beta$ -D-glucopyranosyl)genistein (5) (Fig. S10C<sup>†</sup>).

The microbial products of genistein glycopyranosides using strain 2 were as follows: 35.4 mg L<sup>-1</sup> (81.8  $\mu\text{M}$ ) of 7-*O*-( $\beta$ -D-glucopyranosyl)genistein, 59.06 mg L<sup>-1</sup> (136.4  $\mu\text{M}$ ) of 4'-*O*-( $\beta$ -D-glucopyranosyl)genistein, and 32.64 mg L<sup>-1</sup> (54.8  $\mu\text{M}$ ) of 4',7-di-*O*-( $\beta$ -D-glucopyranosyl)genistein, recovered after 36 h culture incubation (Fig. 3).

### 3.4. Biosynthesis and identification of genistein 2-deoxyglucopyranosides

The 2-deoxyglucopyranoside derivatives of genistein were produced from an *in vitro* reaction using YjiC. A laboratory scale glycosylation reaction was performed to check the preliminary glycosylation of genistein and dTDP-2-deoxy- $\alpha$ -D-glucose, along with other reaction components, as explained in the Materials and methods section. HPLC-PDA analysis revealed three distinct product peaks at  $t_{\text{R}}$  16.9 min, 18.1 min, and 18.4 min



with reference to the substrate standard peak at  $t_R$  21 min (Fig. 2E). Subsequently, the mass analysis of each detected peaks resembled the molecular mass for genistein 2-deoxy-D-diglucoopyranoside and two genistein 2-deoxy-D-monoglucoopyranosides. A fourth product was also observed in the mass analysis although it was not observed in HPLC-PDA. The mass for this peak was similar to that of genistein 2-deoxy-D-triglucoopyranoside. In the positive ion mode, the mass for the compound whose peak was invisible in the HPLC-PDA was found to be  $[M + Na]^+ m/z^+ \sim 731.2168$ . The calculated mass for 2-deoxy-D-triglucoopyranoside was  $m/z^+ \sim 709.2344$  (Fig. S3K†). Similarly, the mass of the compound at  $t_R$  16.9 min was  $[M + H]^+ m/z^+ \sim 563.1770$ , and the calculated mass for genistein 2-deoxy-D-diglucoopyranoside was  $m/z^+ \sim 563.1765$  (Fig. S3J†). For the compounds at  $t_R$  18.1 min and  $t_R$  18.4 min, mass analysis showed  $[M + H]^+ m/z^+ \sim 417.1183$  and  $\sim 417.1179$ , respectively. The calculated mass of genistein 2-deoxy-D-monoglucoopyranoside was  $m/z^+ \sim 417.1186$  (Fig. S3I and H†).

For the biosynthesis of these 2-deoxyglucoopyranoside derivatives of genistein, a 25 mL reaction volume was prepared. Genistein (68 mg,  $\sim 10$  mM in 25 mL) was used as the acceptor substrate and dTDP- $\alpha$ -2-deoxy-D-glucose (135 mg,  $\sim 10$  mM in 25 mL) was used as the sugar donor. However, the other reaction components were added according to the final volume of the reaction mixture and 5 mL crude protein was added followed by incubation at 37 °C for 12 h. The products were purified and further analyzed by NMR for structural characterization.

Four different compounds were biosynthesized from the *in vitro* reaction using YjiC in the presence of dTDP-2-deoxy- $\alpha$ -D-glucose. The products with the  $t_R$  of 16.9 min showed two anomeric protons at  $\delta$  5.30 ppm (dd,  $J = 9.8, 2.1$  Hz, 1H) and 5.43 ppm (dd,  $J = 9.8, 2.1$  Hz, 1H), where both 7-OH and 4'-OH were absent (Fig. S11A†). The anomeric proton in the conformation of  $\alpha$ -2-deoxy-D-glucose in relation to hydrogen of its 2-position is equatorial-equatorial in orientation. In this case, the coupling constant is  $\sim 3$  Hz. However, in  $\beta$ -2-deoxy-D-glucose, the protons of the sugar 2-deoxy positions are axial-equatorially oriented with respect to the anomeric proton. This provided the coupling values in the range between  $\sim 8$  Hz and  $\sim 2$  Hz. Thus, 2-deoxyglucose can have doublet of doublet spectrum if 1''-H is equatorially oriented. This phenomenon is called W-coupling ( $^4J_{HH}$ ) and it depends on the dihedral angle as well as the coupling constant of the vicinal protons (protons attached to the next carbon in the chain). This is explained by the vicinal Karplus correlation as follows: when there is a long distance from 0 °C to 180 °C, the  $J$  value varies.<sup>36–39</sup> Moreover, the anomeric carbon values were also consistent at  $\delta$  96.9 ppm and  $\delta$  96.8 ppm. The region above 100 ppm, which is known as the  $\alpha$ -linkage region<sup>39</sup> (Fig. S11B†) that directs sugar attachment, was in the  $\beta$ -configuration. The chemical shifts of the anomeric carbons in the  $\alpha$ -linkage range from  $\sim \delta$  97 to  $\delta$  101, whereas the anomeric protons are known to have chemical shifts between 5.1 ppm and 5.5 ppm.<sup>39,40</sup> The characteristic 2-deoxy protons at 2.20 ppm (2H) and 1.60 ppm (2H) and carbon at 40.3 ppm and 39.7 ppm were intact in the  $^1H$ -NMR and  $^{13}C$  spectra, respectively, as reported in previous studies

(Fig. S11A†).<sup>27,41</sup> The cross peaks in HMBC, 5.30 ppm of an anomeric proton with 4'-C, 157.2 ppm and 5.43 ppm of an anomeric proton with C-7, and 163 ppm confirm the position of the sugar attachment in genistein, which is also supported by the HSQC cross peaks (Fig. S11C and D†). These results confirm the product to be 4',7-di-O-(2-deoxy  $\beta$ -D-*arabino*-hexopyranosyl)genistein (**11**). The chemical shifts in the  $^1H$  and  $^{13}C$ -NMR spectra for these compounds are tabulated in Tables 2 and 3.

The exact mass of the two products that appeared at  $t_R$  18.1 min and 18.4 min were consistent with the mass of mono-2-deoxy-D-glucoopyranoside of genistein and were further identified by NMR analyses. The anomeric protons of the compound at  $t_R$  18.1 min showed a doublet of doublet peak at  $\delta$  5.29 ppm (dd,  $J = 9.7, 1.9$  Hz, 1H) and the compound at  $t_R$  18.4 min showed a doublet of doublet peak at  $\delta$  5.42 ppm (dd,  $J = 9.6, 2.1$  Hz, 1H) (Fig. S12A and S13A†). The doublet of doublet effect could be due to the same reason as explained for 4',7-di-O-(2-deoxy- $\beta$ -D-*arabino*-hexopyranosyl)genistein. However, the anomeric carbon values were intact at 97 ppm and 96.8 ppm, suggesting sugar attachment is in the  $\beta$ -configuration (Fig. S12B and S13B†). According to previous reports,<sup>27,35,36</sup> the characteristic protons of 2-deoxy sugars were at 2.19 ppm (1H) and 1.60 ppm (1H), whereas other sugar protons were in the region from 3.0 ppm to 4.5 ppm in the compound with  $t_R$  18.1 min (Fig. S12A†). Similarly, in the compound at  $t_R$  18.4 min, the 2-deoxysugar protons were at 2.21 ppm and 1.60 ppm (Fig. S13A†). The cross peak between 1''-H and 4'-C confirmed the conjugation of 2-deoxy glucose at the 4'-OH position and also confirmed the compound to be 4'-O-(2-deoxy  $\beta$ -D-*arabino*-hexopyranosyl)genistein. The 2-deoxy glucose conjugation was further supported by the cross peaks from HMBC and HSQC (Fig. S12C and D†). Similarly, in the compound with  $t_R$  18.4 min, the cross peaks from HMBC, 1''-H (5.42 ppm) and C-7 (162.9 ppm), and HSQC confirmed the conjugation of 2-deoxy-D-glucose at the 7-OH position. This compound was confirmed to be 7-O-(2-deoxy- $\beta$ -D-*arabino*-hexopyranosyl)genistein (**8**) (Fig. S13C and D†). The  $^1H$  and  $^{13}C$ -NMR chemical shifts of these compounds are tabulated in Tables 2 and 3.

Among the three glucoopyranosides obtained from the 25 mL total reaction volume, 16 mg (28.3  $\mu$ M) of 4',7-di-O-(2-deoxy- $\beta$ -D-*arabino*-hexopyranosyl)genistein (**10**), 30 mg (72.1  $\mu$ M) of 7-O-(2-deoxy- $\beta$ -D-*arabino*-hexopyranosyl)genistein, and 17 mg (40.9  $\mu$ M) of 4'-O-(2-deoxy- $\beta$ -D-*arabino*-hexopyranosyl)genistein were synthesized. The fourth product genistein triglucoopyranoside was produced in trace amounts and was confirmed *via* HR-QTOF ESI/MS analysis. The glycosyltransferases catalysed the  $\beta$ -configuration of the 2-deoxysugar attachment in all the abovementioned products and the availability of a third hydroxyl group at the 5-OH of genistein confirmed this product to be 4',5,7-tri-O-(2-deoxy- $\beta$ -D-*arabino*-hexopyranosyl)genistein (Fig. 2E).

### 3.5. Relative solubility of the glucoopyranoside compounds

For the relative solubility study of glucoopyranosides, the ethyl acetate and water fractions were individually analyzed by HPLC-



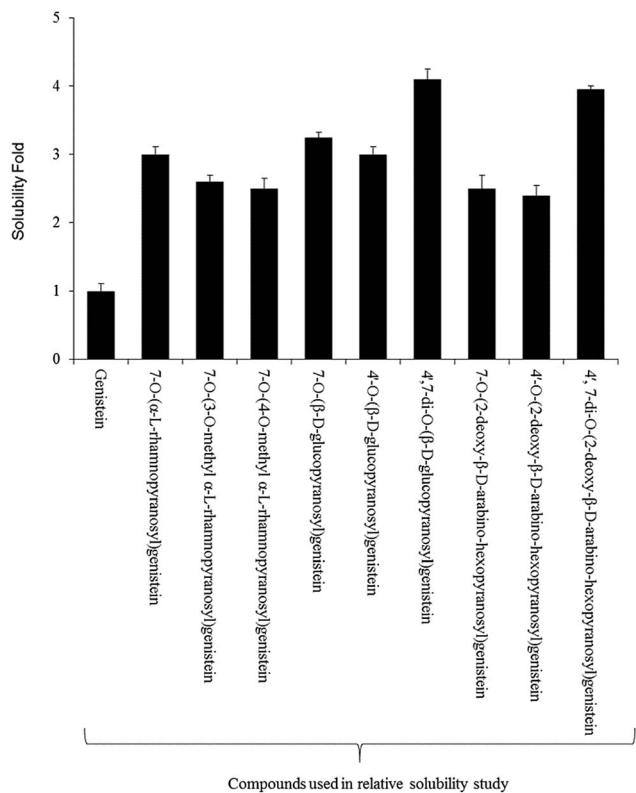


Fig. 4 Relative water solubility of genistein and its glycopyranoside derivatives.

PDA. Although the solubility of genistein varied in different ranges of temperature,<sup>42</sup> we performed the experiment at room temperature ( $\sim 24^\circ\text{C}$ ). The relative solubility of 4',7-di-O-(β-D-glucopyranosyl)genistein (6) and 4',7-di-O-(2-deoxy-β-D-arabino-hexopyranosyl)genistein (11) was  $\sim 3$  fold higher compared to that of genistein aglycon. However, the other glucopyranosides such as 4'-O-(β-D-glucopyranosyl)genistein (7), 7-O-(β-D-glucopyranosyl)genistein (5), 4'-O-(2-deoxy-β-D-arabino-hexopyranosyl)genistein (9), 7-O-(2-deoxy-β-D-arabino-hexopyranosyl)genistein (8), and 7-O-(α-L-rhamnopyranosyl)genistein (2) were  $\sim 2.1$  fold more soluble in water as compared to genistein. Similarly, the solubility of 7-O-(3-O-methyl-α-L-rhamnopyranosyl)genistein (3) and 7-O-(4-O-methyl-α-L-rhamnopyranosyl)genistein (4) was observed to be  $\sim 2.2$  times more as compared to that of genistein standard (Fig. 4).

### 3.6. Anti-fungal and anti-cancer assay

For the anti-fungal test,  $500\ \mu\text{g mL}^{-1}$  of each derivative, including genistein, was used for a disc diffusion experiment against *A. nidulans*. The lowest available concentration of genistein derivative purified for NMR analysis was  $500\ \mu\text{g mL}^{-1}$ ; thus, all the analogues were prepared with the same concentration. However,  $100\ \mu\text{g mL}^{-1}$  Nystatin A1 was used to observe the clear inhibition zone such that it would be easy to compare the inhibition zone with that of the control. The results were observed at 16 h, 30 h, and 50 h incubation time intervals.

However, there was no fungal inhibition zone observed beside Nystatin A1 until  $500\ \mu\text{g mL}^{-1}$  (Fig. S14†).

Since genistein itself shows anti-cancer properties, four different cancer cell lines (AGS, B16F10, HeLa, and HepG2) were assessed against genistein and the genistein derivatives, and the cell growth inhibition was compared. Since some of the genistein derivatives were available in limited concentration, we performed group wise experiments based on their available concentration. In each group, the cell viability inhibited by genistein and genistein derivatives was compared in treatment with different concentrations. In group A, the four cancer cell lines were treated by 7-O-(α-L-rhamnopyranosyl)genistein (2), 7-O-(3-O-methyl α-L-rhamnopyranosyl)genistein (3), and 7-O-(4-O-methyl α-L-rhamnopyranosyl)genistein (4) in the compound concentration range from  $1.56\ \mu\text{M}$  to  $50\ \mu\text{M}$  and the cell viability result was compared with that of genistein (1). These derivatives inhibited the proliferation of each cancer cell line with a different sensitivity to growth inhibition. Among the tested derivatives, 7-O-(3-O-α-L-rhamnopyranosyl)genistein (3) showed an anticancer effect similar to genistein (Fig. 5A). Similarly, in group B, the four cancer cell lines were treated by 4',7-di-O-(2-deoxy-β-D-arabino-hexopyranosyl)genistein (11), 4'-O-(2-deoxy-β-D-arabino-hexopyranosyl)genistein (9), and 7-O-(2-deoxy-β-D-arabino-hexopyranosyl)genistein (8) in the compound concentration range from  $0.31\ \mu\text{M}$  to  $10\ \mu\text{M}$ . These 2-deoxy glucopyranoside derivatives dose dependently inhibited the proliferation of the tested cancer cell lines similar to genistein. Particularly, the inhibitory effect of 4',7-di-O-(2-deoxy-β-D-arabino-hexopyranosyl)genistein was comparable to that of genistein and more effective than those of 4'-O-(2-deoxy-β-D-arabino-hexopyranosyl)genistein and 7-O-(2-deoxy-β-D-arabino-hexopyranosyl)genistein (Fig. 5B). Similarly, in group C, the four cancer cell lines were treated by 4',7-di-O-(β-D-glucopyranosyl)genistein (6), 4'-O-(β-D-glucopyranosyl)genistein (7), and 7-O-(β-D-glucopyranosyl)genistein (5) and the results were compared with that of genistein (1) in the compound concentration range from  $3.12\ \mu\text{M}$  to  $100\ \mu\text{M}$  (Fig. 5C). Although these glucosides showed inhibitory activities on cancer cell growth, they were less effective than genistein. However, these compounds still possess anticancer potential.

## 4. Discussion

The diphenol structure of genistein resembles that of human endogenous estrogen. Genistein has been extensively studied with numerous molecular targets, from antioxidants to various types of cancer, for the prevention of human immunodeficiency virus (HIV-1) infection of primary macrophages, herpes simplex viral infection,<sup>43–45</sup> and many more. In this study, we biosynthesized a series of genistein glycopyranosides using *in vitro* and *in vivo* glycosylation approaches. Microbial GT YjiC and plant-based GT AtUGT89C1 were flexible enough to utilize the microbial indigenous NDP-sugars as sugar donors for conjugation at various hydroxyl positions of genistein. In addition, we cloned and expressed two SOMTs (*spnK* and *spnH*) from *S. spinosa* (Fig. S1†) and were able to transfer a methyl group to the



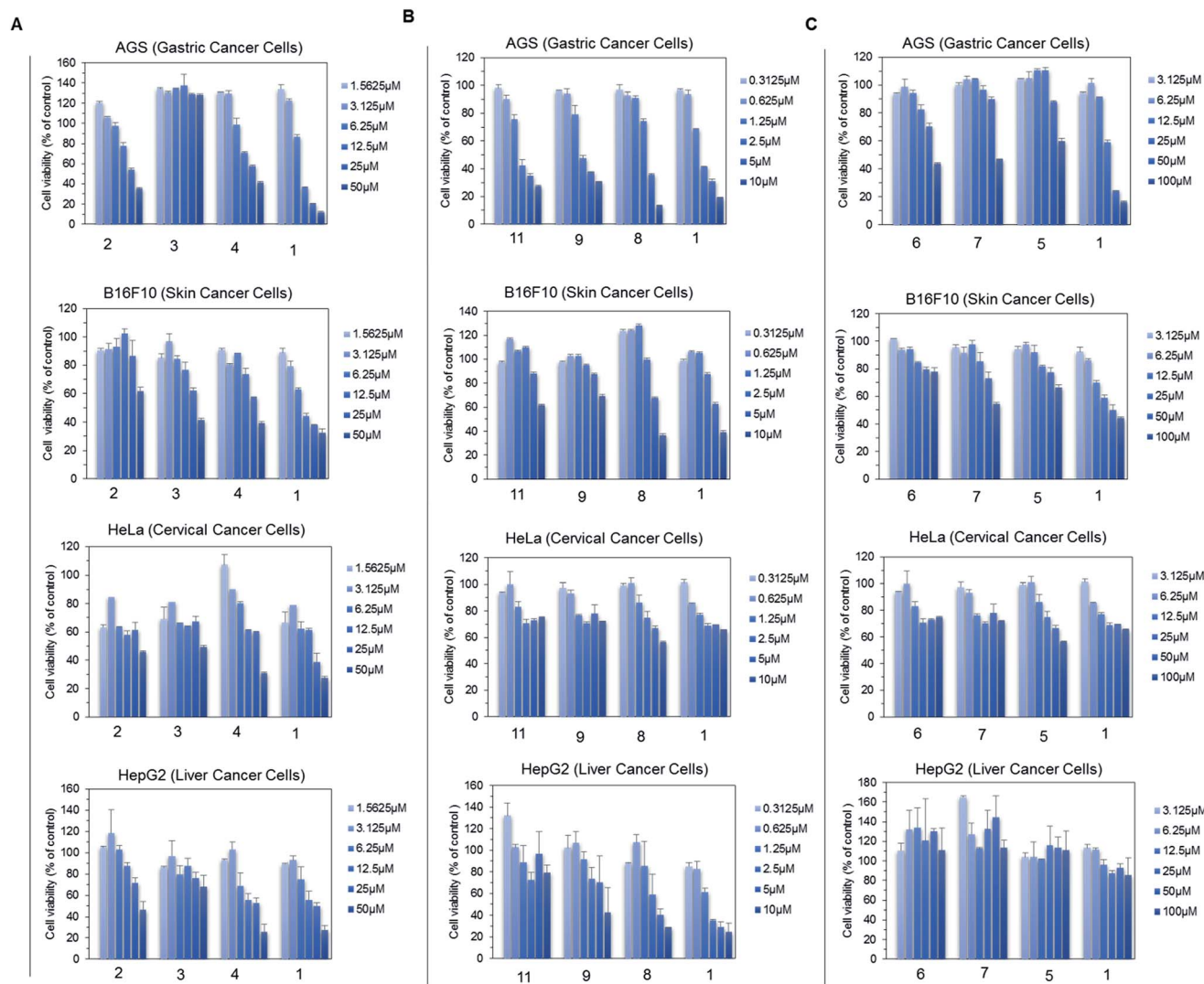


Fig. 5 Anticancer activity (MTT assay) of genistein derivatives against four different cancer cell lines (AGS: gastric cancer cells, B16F10: skin cancer cells, HeLa: cervical cancer cells, and HepG2: liver cancer cells) compared with the cell growth inhibition affected by genistein shown after 72 h incubation time. Three groups of experiments are shown in different compound concentration ranges. (A) Each cancer cell line was treated by different concentrations (range: from 1.56  $\mu\text{M}$  to 50  $\mu\text{M}$ ) of the genistein derivatives 7-*O*-( $\alpha$ -L-rhamnopyranosyl)genistein (2), 7-*O*-(3-*O*- $\alpha$ -L-rhamnopyranosyl)genistein (3), 7-*O*-(4-*O*- $\alpha$ -L-rhamnopyranosyl)genistein (4), and genistein (1) and the results were compared. (B) Each cancer cell line was treated by different concentrations (range: from 0.31  $\mu\text{M}$  to 10  $\mu\text{M}$ ) of the genistein derivatives 4',7-di-*O*-(2-deoxy- $\beta$ -D-*arabino*-hexopyranosyl)genistein (11), 4'-*O*-(2-deoxy- $\beta$ -D-*arabino*-hexopyranosyl)genistein (9), 7-*O*-(2-deoxy- $\beta$ -D-*arabino*-hexopyranosyl)genistein (8), and genistein (1). (C). Each cancer cell line was treated by different concentrations (range: 3.12  $\mu\text{M}$  to 100  $\mu\text{M}$ ) of the genistein derivatives 4',7-di-*O*-( $\beta$ -D-glucopyranosyl)genistein (6), 4'-*O*-( $\beta$ -D-glucopyranosyl)genistein (7), 7-*O*-( $\beta$ -D-glucopyranosyl)genistein (5), and genistein (1). The error bars indicate the standard deviation values for three successive independent experiments.

respective hydroxyl positions of the rhamnose moiety conjugated with genistein.

With reference to our previous study, AtUGT89C1 has shown diverse flexibility in terms of substrates and in using bacterial indigenous dTDP- $\beta$ -L-rhamnose as a sugar donor to transfer the rhamnose moiety at the 7-OH position of various classes of flavonoids.<sup>23</sup> However, in a subsequent study, this GT showed sugar donor flexibility over genistein as the acceptor substrate. We successfully biosynthesized two products from a bioconversion experiment using AtUGT89C1 and further analyses revealed the products to be 7-*O*-( $\beta$ -D-glucopyranosyl)genistein and 7-*O*-( $\alpha$ -L-rhamnopyranosyl)genistein (Fig. 2A). The bioconversion of

genistein into these products was significant compared to that of other flavonoid rhamnopyranosides from the previous report in a shake flask culture.<sup>23</sup> The previously characterized SOMTs were co-transformed with AtUGT89C1 to biosynthesize *O*-methylated rhamnopyranosyl genistein derivatives. These SOMTs were studied by various groups and their functions have been characterized over the native spinosyn biosynthesis pathway. These sugar tailoring enzymes have never been used in the post modifications of phenolic rhamnopyranoside.<sup>33</sup> However, we did not check whether AtUGT89C1 was capable of accepting the *O*-methyl rhamnose moiety for glycosylation. In the biotransformation experiment for AtUGT89C1 along with spnK and spnH



co-transformation, new products were generated and the corresponding peaks were observed before the substrate peak and after the 7-*O*-( $\alpha$ -L-rhamnopyranosyl)genistein peak (Fig. 2B and C). These products were further characterized to be 7-*O*-(3-*O*-methyl  $\alpha$ -L-rhamnopyranosyl)genistein and 7-*O*-(4-*O*-methyl  $\alpha$ -L-rhamnopyranosyl)genistein (Tables 1 and 2, respectively). The microbial products of these compounds were significant in the shake flask culture when 7-*O*-( $\alpha$ -L-rhamnopyranosyl)genistein was used as the substrate without host engineering or supplementation of methyl donors (Fig. 2B and C). Thus, three rhamnosyl genistein derivatives 7-*O*-( $\alpha$ -L-rhamnopyranosyl)genistein, 7-*O*-(3-*O*-methyl  $\alpha$ -L-rhamnopyranosyl)genistein, and 7-*O*-(4-*O*-methyl  $\alpha$ -L-rhamnopyranosyl)genistein are novel products reported for the first time from a microbial source.

For the biosynthesis of other genistein glycopyranosides and 2-deoxy D-glycopyranosides, YjiC from *B. licheniformis* DSM-13 was used. This is a versatile GT that has been extensively studied for the glycosylation of various plants and microbial secondary metabolites, which conjugates diverse NDP-sugars as the sugar donor.<sup>25,46</sup> For example, the glycosylation of geldanamycin, phloretin, flavonols, epothilone A, mupirocin, resveratrol, nargenicin A (1), and anthraquinones.<sup>24,25,47–52</sup> In this experiment, we used enzymatically prepared dTDP-2-deoxy  $\alpha$ -D-glucose as the sugar donor and genistein as the acceptor substrate. Three different 2-deoxy D-glycopyranosides peaks were visible in HPLC-PDA with reference to the substrate peak (Fig. 2E). These products were further confirmed by HR-QTOF ESI/MS and structurally elucidated by NMR spectroscopy to be 7-*O*-(2-deoxy- $\beta$ -D-glycopyranosyl)genistein, 4'-*O*-(2-deoxy- $\beta$ -D-glycopyranosyl)genistein, and 4',7-di-*O*-(2-deoxy- $\beta$ -D-glycopyranosyl)genistein (Fig. 2E and Tables 1 and 2). During mass analyses, a fourth product, which was detected only in trace amounts, was confirmed to be tri-2-deoxy glycopyranoside of genistein, *i.e.* 4',5,7-tri-*O*-(2-deoxy- $\beta$ -D-glycopyranosyl)genistein (Fig. S3K†). Genistein is a tri-hydroxylated isoflavone for which tri-glycopyranoside formation is possible when the reaction is catalyzed by a flexible GT such as YjiC. All these compounds are novel glycopyranoside derivatives.

Similarly, based on our previous study on the biosynthesis of isoflavonoid glycopyranosides,<sup>26</sup> a similar biotransformation approach was used. A single transformant, strain-2, was used to catalyze the production of three different glycopyranosides. Significant consumption of the substrate was achieved in 36 h of incubation to produce three different glycopyranosides: 4',7-di-*O*-(2-deoxy- $\beta$ -D-glycopyranosyl)genistein, 4'-*O*-( $\beta$ -D-glycopyranosyl)genistein, and 7-*O*-( $\beta$ -D-glycopyranosyl)genistein (Fig. 2D and S3E–G†).

*In vivo* and *in vitro* reactions were employed to catalyze the biosynthesis of different glycopyranosyl derivatives of genistein. Further production enhancement is necessary over these bioconversion experiments. As a proof of concept, we experimented with and used various post modifying enzymes in the shake flask culture to biosynthesize natural and non-natural glycopyranosides of genistein in a microbial host. Simply using the same recombinant strains in the fermentation could lead to an enhancement in the production. Moreover, various biotechnological approaches such as microbial host engineering and the

construction of cell factories have significant impact over the production of these valuable metabolites in microbial platforms.<sup>53</sup> Microbial metabolite production and enhancement is an important and economic method for the biosynthesis and post modification of natural products. A number of strategies including system/synthetic biology and metabolic engineering approaches can be applied to design and construct NDP-sugar pathways such that there is a sufficient supply of donor substrates, which is usually a limiting factor in the *in vivo* glycosylation reactions.<sup>54</sup> These strategies could be applied to enhance the production of genistein glycopyranosides in microbial hosts. However, in enzymatic synthesis, the addition of more sugar donors could increase the production of 2-deoxy D-glycopyranosides. However, this would not be an economical approach. Alternatively, a one-pot enzymatic system could be applicable to biosynthesize these glycopyranosides in bulk from cheap sources.<sup>41</sup>

The relative solubilities of the synthesized compounds were experimented and were found to be significant as compared to that of the parent compound (Fig. 4). Thus, the beneficial therapeutic value of the synthesized compounds could be explored in future studies. Genistein itself is considered as a lead compound in anticancer drug design, with several researchers exploring its potential and comparing its activity to those of its glycosidic derivatives since glycosidic bonds are prone to undergo enzymatic hydrolysis by the native glycosidase present in mammals.<sup>55–57</sup> Some deoxy analogues such as talosin A and B possess strong and comparable antifungal activity with amphotericin B although genistein and 7-*O*-( $\beta$ -D-glycopyranosyl)genistein do not show significant antifungal activity, as explained in a previous report.<sup>18</sup> Other deoxy analogues of genistein glycosides could possibly have more potent activity. With these concepts, we assessed the anti-fungal and anti-cancer activity of these synthesized compounds. However, we could not detect anti-fungal activity of any of the compounds tested against *A. nidulans* (Fig. S14†). However, in the MTT assay experiment over different cancer cell lines (AGS, B16F10, HeLa, and HepG2), some derivatives such as 4',7-di-*O*-(2-deoxy- $\beta$ -D-*arabino*-pyranosyl)genistein and 7-*O*-(4-*O*-methyl- $\alpha$ -L-rhamnopyranosyl)genistein showed an anticancer effect similar to genistein. However, the rest of the derivatives dose dependently inhibited the growth of cancer cells treated at various concentrations and their anti-cancer properties were not knock down (Fig. 5 A–C).

## 5. Conclusions

The biosynthesis of these compounds is important for new drug discovery. Since genistein itself has been studied with regard to several molecular targets, our synthesized natural and non-natural glycosides (2-deoxyglycopyranosides, *O*-methyl rhamnopyranosides) of genistein could be the lead compounds for future therapeutic applications if investigated over various targets. The application of GTs and SOMTs resulted in diversified flavonoids by decorating diverse sugar moieties and methyl groups, thereby increasing bioactivities. We used SOMTs to conjugate the methyl group at specific hydroxyl positions in 7-*O*- $\alpha$ -L-rhamnopyranosyl genistein, which could add positive



effects to its activity since it was reported to have no antifungal activity at 100  $\mu\text{g mL}^{-1}$  by the Suh group in 2006.<sup>18</sup> Hence, the approach of using versatile post modifying enzymes in this study shows the possibility of biosynthesizing future prospective compounds. It is possible to enhance the biosynthesis of these compounds by the application of further host engineering and synthetic biology tools.

## Acknowledgements

This work was supported by a grant from the Next-Generation BioGreen 21 Program (SSAC, grant#: PJ01111901), Rural Development Administration, Republic of Korea.

## References

- 1 R. Simons, H. Gruppen, T. F. Bovee, M. V. Verbruggen and J. P. Vincken, *Food Funct.*, 2012, **3**, 810–827.
- 2 G. Maskarinec, R. Yamakawa, S. Hebshi and A. A. Franke, *Eur. J. Clin. Nutr.*, 2007, **61**, 255–261.
- 3 P. McCue and K. Shetty, *Crit. Rev. Food Sci. Nutr.*, 2004, **44**, 361–367.
- 4 F. He and J. Chen, *Food Sci. Hum. Wellness*, 2013, **2**, 146–161.
- 5 E. Miadokova, *Interdiscip. Toxicol.*, 2009, **2**, 211–218.
- 6 M. L. Xu, J. Liu, C. Zhu, Y. Gao, S. Zhao, W. Liu and Y. Zhang, *Phytochemistry*, 2015, **14**, 459–467.
- 7 K. Polkowski and A. P. Mazurek, *Acta Pol. Pharm.*, 2000, **57**, 132–155.
- 8 Q. Miao, J. G. Li, S. Miao, N. Hu, J. Zhang, S. Zhang, Y. H. Xie, J. B. Wang and S. W. Wang, *Int. J. Mol. Sci.*, 2011, **13**, 56–70.
- 9 G. R. Yan, C. L. Xiao, G. W. He, X. F. Yin, N. P. Chen, Y. Cao and Q. Y. He, *Proteomics*, 2010, **10**, 976–986.
- 10 Y. Mizushima, K. Shiomi, I. Kuriyama, Y. Yakahashi and H. Yoshida, *Int. J. Clin. Oncol.*, 2013, **43**, 1117–1124.
- 11 F. Schmidt, C. B. Knobbe, B. Frank, H. Wolburg and H. Weller, *Oncol. Rep.*, 2008, **19**, 1061–1066.
- 12 C. Weigt, T. Hertrampf, U. Flenker, F. Hulsemann, P. Kurnaz, K. H. Fritzscheier and P. Diel, *J. Steroid Biochem. Mol. Biol.*, 2015, **154**, 12–22.
- 13 T. A. Boqush, A. V. Konukhocva, A. B. Ravcheeva, T. N. Zabolina, Z. G. Kadagidze, E. A. Bogush, D. V. Komov, B. E. Polotskii, K. K. Laktionov and M. I. Davydov, *Antibiot. Khimioter.*, 2003, **48**, 11–15.
- 14 Z. Yang, K. Kulkarni, W. Zhu and M. Hu, *Adv. Anticancer Agents Med. Chem.*, 2012, **12**, 1264–1280.
- 15 H. Zhang, G. Liu, X. Zeng, Y. Wu, C. Yang, L. Mei, Z. Wang and L. Huang, *Int. J. Nanomed.*, 2015, **10**, 2461–2473.
- 16 A. Rusin, Z. Krawczyk, G. Gryniewicz, A. Goqler, J. Zawisza-Puchalka and W. Szeja, *Acta. Biochem. Pol.*, 2010, **57**, 23–34.
- 17 B. F. Cress, R. J. Linhardt and M. A. G. Koffas, *Isoflavonoid production by genetically engineered microorganisms*, Springer, New York, NY, 2013, DOI: 10.1007/978-3-642-22144-6\_53.
- 18 T. M. Yoon, J. W. Kim, J. G. Kim, W. G. Kim and J. W. Suh, *J. Antibiot.*, 2006, **59**, 633–639.
- 19 A. R. Jesus, C. Dias, A. M. Matos, R. F. de Almeida, A. S. Viana, F. Marcelo, R. T. Reibeiro, M. P. Macedo, C. Airoidi, F. Nicotra, A. Martins, E. J. Cabrita, J. Jimenez-Barbero and A. P. Rauter, *J. Med. Chem.*, 2014, **57**, 9463–9472.
- 20 A. Steensma, M. A. Fassen-Peters, H. P. Noteborn and I. M. Rietjens, *J. Agric. Food Chem.*, 2006, **54**, 8006–8012.
- 21 C. J. Thibodeaux, C. E. Melancon and H. W. Liu, *Nature*, 2007, **446**, 1008–1016.
- 22 J. White-Phillip, C. J. Thibodeaux and H. W. Liu, *Methods Enzymol.*, 2009, **459**, 521–544.
- 23 P. Parajuli, R. P. Pandey, N. T. Trang, T. J. Oh and J. K. Sohng, *Carbohydr. Res.*, 2015, **418**, 13–19.
- 24 P. Parajuli, R. P. Pandey, N. Koirala, Y. J. Yoon, B. G. Kim and J. K. Sohng, *AMB Express*, 2014, **4**, 31.
- 25 R. P. Pandey, P. Parajuli, J. Y. Shin, J. Lee, S. Lee, Y. S. Hong, Y. I. Park, J. S. Kim and J. K. Sohng, *Appl. Environ. Microbiol.*, 2014, **80**, 7235–7243.
- 26 R. P. Pandey, P. Parajuli, N. Koirala, J. H. Lee, Y. I. Park and J. K. Sohng, *Mol. Cells*, 2014, **37**, 172–177.
- 27 Y. H. Yang, Y. B. Kang, D. H. Kim, T. H. Lee, S. H. Park, K. Lee, D. Yoo, D. K. K. Liou, H. C. Lee, J. K. Sohng and B. G. Kim, *J. Mol. Catal. B: Enzym.*, 2010, **62**, 282–287.
- 28 J. Roepke and G. G. Bozzo, *ChemBioChem*, 2013, **14**, 2418–2422.
- 29 W. G. Kim, T. M. Yoon, H. J. Kwon and J. W. Suh, *J. Antibiot.*, 2006, **59**, 640–645.
- 30 M. U. Roslund, P. Tahtinen, M. Niemitz and R. Sjöholm, *Carbohydr. Res.*, 2008, **343**, 101–112.
- 31 P. K. Agrawal, *Phytochemistry*, 1992, **31**, 3307–3330.
- 32 A. P. Rauter, A. Martins, C. Borges, J. Ferreira, J. Jorge, M. R. Bronze, A. V. Coelho, Y. H. Choi and R. Verpoorte, *J. Chromatogr. A*, 2005, **1089**, 59–64.
- 33 H. J. Kim, J. A. White-Phillip, Y. Ogasawara, N. Shin, E. A. Isiorho and H. W. Liu, *J. Am. Chem. Soc.*, 2010, **132**, 2901–2903.
- 34 A. Z. Popper, I. H. Sadler and S. C. Fry, *Biochem. Syst. Ecol.*, 2004, **32**, 279–289.
- 35 J. M. Chen, M. D. Shepherd, J. Horn, M. Leggas and J. Rohr, *ChemBioChem*, 2014, **15**, 2729–2735.
- 36 B. Coxon, *Adv. Carbohydr. Chem. Biochem.*, 2009, **62**, 17.
- 37 G. G. Zapesochnaya, *Chem. Nat. Compd.*, 1982, **18**, 658.
- 38 R. J. Cushley, J. F. Codrington and J. J. Fox, *Can. J. Chem.*, 1967, **46**, 1131.
- 39 N. E. Jacobsen, *NMR Spectroscopy Explained: Simplified Theory, Applications and Examples for Organic Chemistry and Structural Biology*, John Wiley, 2007, DOI: 10.1002/9780470173350.
- 40 W. A. Bubb, *Concepts Magn. Reson., Part A*, 2003, **19**, 1–19.
- 41 T. T. Le, R. P. Pandey, R. B. Gurung, D. Dhakal and J. K. Sohng, *Appl. Microbiol. Biotechnol.*, 2014, **98**, 8527–8538.
- 42 J. G. Wu, J. Ge, Y. P. Zhang, Y. Yu and X. Y. Zhang, *J. Chem. Eng. Data*, 2010, **55**, 5285–5288.
- 43 T. S. Stantchev, I. Markovic, W. G. Telford, K. A. Clouse and C. C. Broder, *Virus Res.*, 2007, **123**, 178–189.
- 44 A. Ghaemi, H. Soleimanjahi, S. Razeghi, A. Gorji, A. Tabaraei, A. Moradi, A. Alizadeh and M. A. Vakili, *Iran. J. Immunol.*, 2012, **9**, 119–127.
- 45 A. A. Date and C. J. Destache, *Drug Discovery Today*, 2016, **21**, 333–341.



- 46 R. P. Pandey, R. B. Gurung, P. Parajuli, N. Koirala, T. T. Le and J. K. Sohng, *Carbohydr. Res.*, 2014, **393**, 26–31.
- 47 C. Z. CZ, J. S. Jang, M. Woo, J. S. Ahn, J. S. Kim and Y. S. Hong, *Appl. Environ. Microbiol.*, 2012, **21**, 7680–7686.
- 48 R. P. Pandey, T. F. Li, E. H. Kim, T. Yamaguchi, Y. I. Park, J. S. Kim and J. K. Sohng, *Appl. Environ. Microbiol.*, 2013, **79**, 3516–3521.
- 49 R. P. Pandey, P. Parajuli, N. Koirala, J. W. Park and J. K. Sohng, *Appl. Environ. Microbiol.*, 2013, **79**, 6833–6838.
- 50 P. Parajuli, R. P. Pandey, A. R. Pokhrel, G. P. Ghimire and J. K. Sohng, *Glycoconjugate J.*, 2014, **31**, 563–572.
- 51 D. Dhakal, T. T. Le, R. P. Pandey, A. K. Jha, R. B. Gurung, P. Parajuli, A. R. Pokhrel, J. C. Yoo and J. K. Sohng, *Appl. Biochem. Biotechnol.*, 2015, **175**, 2934–2949.
- 52 G. P. Ghimire, N. Koirala, R. P. Pandey, H. J. Jung and J. K. Sohng, *World J. Microbiol. Biotechnol.*, 2015, **31**, 611–619.
- 53 P. Parajuli, R. P. Pandey, N. T. Trang, A. K. Chaudhary and J. K. Sohng, *Microb. Cell Fact.*, 2015, **14**, 76.
- 54 R. P. Pandey, P. Parajuli, M. A. Koffas and J. K. Sohng, *Biotechnol. Adv.*, 2016, **S0734-9750**, 30019–30022.
- 55 K. Polkowski, J. Popiolkiewicz, P. Krzeczynski, J. Ramza, W. Pucko, O. Zegrocka-Stendel, J. Boryski, J. S. Skierski, A. P. Mazurek and G. Gryniewicz, *Cancer Lett.*, 2004, **203**, 59–69.
- 56 J. Popiolkiewicz, K. Polkowski, J. S. Skierski and A. P. Mazurek, *Cancer Lett.*, 2005, **229**, 67–75.
- 57 C. O. Junior, S. B. Castro, A. A. Pereira, C. C. Alves, E. E. Oliveira, R. T. Rego, A. P. Ferreira and M. V. de Almeida, *Eur. J. Med. Chem.*, 2014, **85**, 615–620.

