DOI: 10.1002/cbic.201300474

Biocatalytic Synthesis of Quercetin 3-*O*-Glucoside-7-*O*-Rhamnoside by Metabolic Engineering of *Escherichia coli*

Jonathon Roepke and Gale G. Bozzo*[a]

Flavonol glycosides, like quercetin 3-*O*-glucoside (**1**) and the bisglycoside quercetin-3-*O*-glucoside-7-*O*-rhamnoside (**2**) are plant natural products exhibiting numerous biological activities.[1,2] Compound **2** is rare, but has been described for *Capsicum* species and *Arabidopsis thaliana*, and is often found in complex mixtures with other flavonols.[1,3–5] Hence, purification from plants is not practical. Here, we report the regioselective synthesis and purification of **2** from an *Escherichia coli* expression strain harbouring a rhamnose synthase and a flavonol 7-*O*-rhamnosyltransferase from *Arabidopsis thaliana*.

Approximately 350 quercetin derivatives exist, they all come from plants.[6] In nature, quercetin occurs as glycosides with pentoses or hexoses (e.g., rhamnose) conjugated to any of five available hydroxy groups.[7,8] Glycosylation of flavonols increases their aqueous solubility relative to their aglycones,[9] hence enhancing the absorption of quercetin from the small intestine in humans.[10] In *A. thaliana*, 11 quercetin glycosides are known, including **2** and others distinguished by the regiospecific attachment of one or two sugar moieties at the 3-*O* and/or 7-*O* positions.[8] Flavonol bisglycoside production is catalysed by uridine diphosphate (UDP)-dependent glycosyltransferase (UGT). In *Arabidopsis*, **2** is derived from **1** by a UDP–rhamnose-dependent flavonol 7-*O*-rhamnosyltransferase (AtUGT89C1).[7,11] UDP–rhamnose is derived from UDP–glucose by the activity of rhamnose synthase (RHM). Ablation of *AtRHM1* culminates in a marked reduction in flavonol 3-*O*-glucoside-7-*O*-rhamnoside levels in *Arabidopsis* leaves and flowers,[7] thus pointing to a sequence of steps in the regioselective synthesis of **2** in plants.

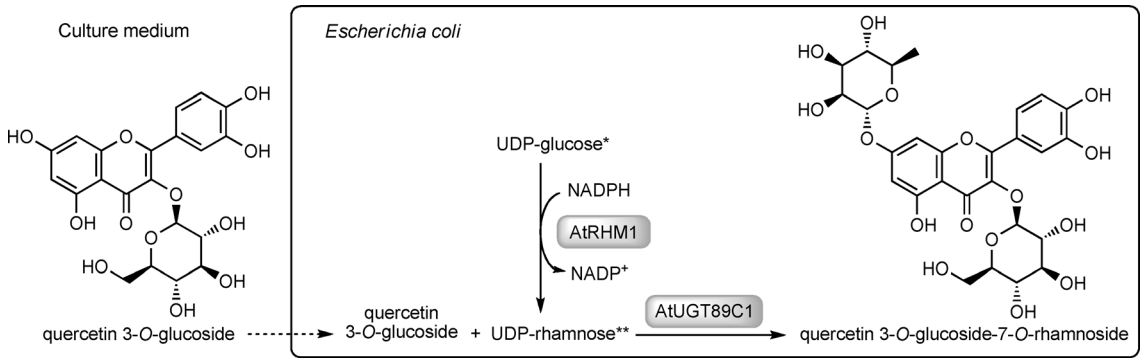
Bacteria are rich in nucleotide sugars and UGTs,[12] but lack flavonols. Moreover, bacteria contain UDP–glucose and thymine diphosphate rhamnose (TDP–rhamnose);[13] however, UDP–rhamnose is absent. Quercetin fed to *E. coli* expressing AtUGT89C1 together with a flavonol 3-*O*-glucosyltransferase (AtUGT78D2) yields a complex mixture consisting of compounds **1** and **2** together with quercetin 3-*O*-glucoside-7-*O*-glucoside, quercetin 3-*O*-(*N*-acetyl)glucosamine and quercetin 3-*O*-(*N*-acetyl)glucosamine-7-*O*-rhamnoside.[14] In the absence of UDP–rhamnose, the aforementioned transformants use U(T)DP–glucose and UDP–*N*-acetylglucosamine in addition to TDP–rhamnose to modify **1** leading to by-products.[14] Alternatively, dual expression of plant UGT and RHM genes in *E. coli* generates rhamnosylated quercetin,[13–15] but this has not been attempted for the rare bisglycoside **2**.

[a] *J. Roepke, Dr. G. G. Bozzo*
Department of Plant Agriculture, University of Guelph
50 Stone Road E., Guelph, ON, N1G 2W1 (Canada)
E-mail: gbozzo@uoguelph.ca

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201300474.

To limit the number of by-products produced during biocatalysis the synthesis of **2** may be facilitated by feeding β-glucoside **1** to metabolically engineered *E. coli*. β-Glucosides are actively absorbed by *E. coli*,[16] but are not metabolised by wild-type strains;[17] this eliminates any possibility of their hydrolysis in a biocatalytic reaction. We examined the feasibility of feeding **1** to an *E. coli* expression strain containing two plasmids, one of which carried *AtUGT89C1* and the other *AtRHM1* (Scheme 1). Moreover, we describe a simple chromatography procedure for purifying this bisglycoside.

To determine whether regioselective synthesis of **2** was possible, pET41b and pET32b expression vectors harbouring *AtUGT89C1* and *AtRHM1*, respectively, were co-transformed into *E. coli* BL21(DE3). Compound **1** (1 mg per 50 mL culture) was added to culture media of *E. coli* dually expressing AtUGT89C1 and AtRHM1, *E. coli* harbouring one of these plant genes together with an empty plasmid, or cells transformed with two empty plasmids. At the end of the biocatalysis reaction, quercetin conjugates from the spent culture media and cell lysates of AtRHM1/AtUGT89C1 transformants were extracted and analysed. As much as 40% of **1** was recovered from cell lysates, regardless of whether a product was detected (Figure 1 A); however, it is possible that part of the absorbed compound **1** was refluxed to the medium, as previous studies have shown that 60–80% of flavonol aglycone and monoglycoside are secreted during biocatalysis.[18] HPLC-DAD (DAD = diode array detector) revealed a single product with a retention time (t_R = 6.1 min) distinct from that of **1** (t_R = 8.0 min; Figure 1). The product's absorption maxima (256, 354 nm, in 20% acetonitrile containing 0.1% formic acid; Supporting Information) were comparable to those of compound **2** isolated from whole-plant *Arabidopsis*.[4] Reaction product levels were 1.5 times greater in *E. coli* AtRHM1/AtUGT89C1 transformants than in cells expressing only AtUGT89C1 (Supporting Information). A product in *E. coli* expressing only AtUGT89C1 (Supporting Information) suggests rhamnosylation of compound **1** is supplemented by endogenous TDP–rhamnose.[13] As expected, no product was formed in cells expressing only AtRHM1 or those harbouring two empty plasmids (Supporting Information), thus pointing to a lack of natural flavonol bisglycoside biosynthesis in bacteria. The majority of the reaction product (80%) was present in cells, with the remainder in the spent culture medium (Figure 1 B). Although the proportion of **2** recovered in the spent cell culture medium was lower than in the cellular fraction, this phenomenon is consistent with the partial secretion of quercetin conjugates from *E. coli* flavonol UGT transformants.[18] A previous report found **2** localised primarily to the *E. coli* culture medium;[14] this might be a consequence of a longer incubation period and/or cultivation at 70% higher cell density under



Scheme 1. Biocatalytic synthesis of quercetin 3-*O*-glucoside-7-*O*-rhamnoside (**2**) in *E. coli*. NADPH: reduced nicotinamide adenine dinucleotide phosphate, $NADP^+$: oxidised nicotinamide adenine dinucleotide phosphate, *: endogenous TDP–glucose may substitute for UDP–glucose in the AtRHM1 reaction to yield TDP–rhamnose (**), an alternate nucleotide sugar for AtUGT89C1.[11]

O_2-limiting conditions relative our study. In response to anaerobia, there is an increased incidence of metabolite secretion from *E. coli* cultured at a higher cell density.[19]

To purify as much of the reaction product as possible, spent culture medium was pooled with methanol cell extracts from a 500 mL culture of AtUGT89C1/AtRHM1 transformants. This culture volume is within the range (5 to 2000 mL) of previous experiments.[13–15,18,20] After acidification, the sample was passed through an Amberlite XAD-2 column; this adsorbent purifies quercetin glycosides away from most interfering phenolics and cellular debris.[20] HPLC-DAD revealed that 95% of the biocatalytic product was recovered in the XAD-2 eluate (Table 1). The XAD-2 eluate was dried under vacuum, resuspended in 20% methanol and passed through a Sephadex LH-20 column. This size-exclusion chromatography is ideal for quercetin glycoside purification from complex matrices such as plant-tissue extracts.[21] The chromatographic strategy outlined here is a standard approach for purifying flavonol glycosides to homogeneity, and is cost effective as it does not require preparative HPLC.[20,21] HPLC-DAD analysis revealed that some fractions of the Sephadex LH-20 eluate contained a contaminant-free molecule with a t_R of 6.1 min (Figure 2B), which is similar to that of the biocatalysis product. Preceding fractions from the LH-20 step were dominated by non-flavonol molecules, and fractions 27–31 consisted mostly of compound **1** (Figure 2A).

Table 1. Purification of compound **2** from *E. coli* culture (500 mL) expressing AtUGT89C1 and AtRHM1.

	Biocatalysis [mg]		XAD-2	LH-20
	0 h	24 h	[mg]	[mg]
compound **1**[a]	20	n.d.	n.d.	n.d.
compound **2**[b]	0	10	9.5	3.4
yield [%]		50	47.5	17

[a] Quantity used for biocatalysis. [b] Calculated as mg of **2** on the basis of a molecular mass of 610.2. n.d.: not determined.

Fractions 21–26 of the LH-20 step were pooled and subjected to structural analysis. The molecular mass of the product was 610.2, which is 146 more than that of compound **1**, thus indicating the presence of a rhamnose moiety (Figure 2C). Collision-induced dissociation spectra of the parent ion gave rise to $[M-H]^-$ fragments of 463.2, 446.2 and 301.1, which corresponded to the loss of a rhamnose, a glucose and both sugars respectively (Figure 2C); all indicating that a quercetin bisglycoside was produced from **1**. The structure of the biocatalysis product was investigated by NMR spectroscopy and confirmed to be $C_{27}H_{30}O_{16}$ (Figure 3). HMBC spectra determined a correlation between the C-1‴ anomeric carbon of rhamnose and quercetin H8. Moreover, heteronuclear interactions between the C-7 of quercetin with the H6 and H8 aromatic protons of the quercetin A ring and the anomeric proton (H1‴) of rhamnose were apparent. This is consistent with structural data for **2** from *Arabidopsis* seeds.[22] Together, the structural data confirmed that the biocatalysis product was indeed quercetin bisglycoside **2**.

Previously, it was shown that 20 kg of *Arabidopsis* yields 4.2 mg of pure compound **2**.[4] The disadvantages of this approach include the lengthy plant cultivation period relative to that of bacteria, and the use of large volumes (186 L) of organic solvents for the extraction and purification of the target metabolite from *Arabidopsis*.[4] Our biocatalysis approach produced higher amounts (Table 1) of this commercially unavailable compound within 24 h and generated only 0.4 L of organic waste, thus it is a cost-effective alternative to extraction from *Arabidopsis* tissues. In addition, prior to purification, our regioselective synthesis produced a single product, whereas quercetin feeding yields multiple quercetin glycosides.[13,14,18,20,23] Moreover, 50% of **1** was converted to **2** within 24 h of feeding; this approximates biocatalysis rates described elsewhere.[14,18] As kaempferol 3-*O*-glycosides are rhamnosylated efficiently by AtUGT89C1,[11] our regioselective synthesis approach could be used to produce their bisglycosides. Biocatalytic synthesis of compound **2** could be improved by controlling the pH and O_2 saturation by biofermentation,[24] metabolic engineering of *E. coli* for increased absorption/minimal secretion of **1**, or transforming *E. coli* with a flavonol 7-*O*-rhamnosylransferase that

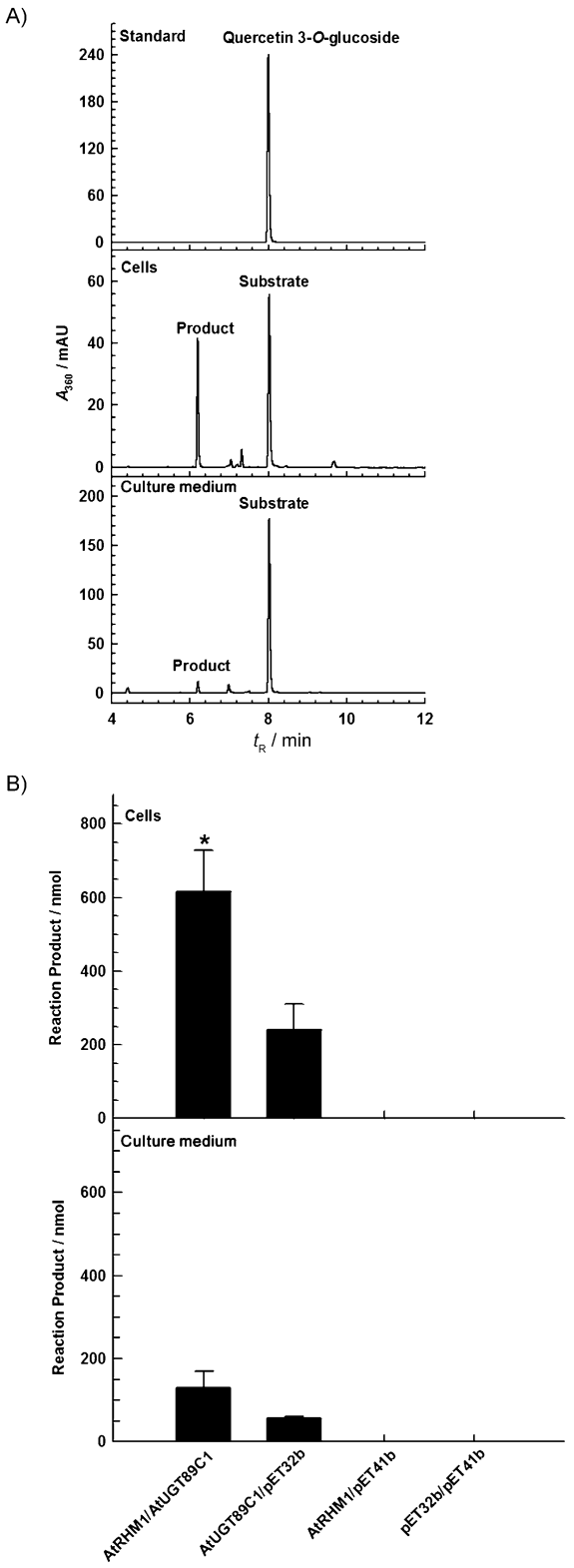


Figure 1. A) HPLC analysis of quercetin glycosides in cells and spent culture media of *Escherichia coli* AtRHM1/AtUGT89C1 transformants. In all cases, controls included *E. coli* transformed with AtRHM1/AtUGT89C1, cells harbouring either AtRHM1/pET41b or AtUGT89C1/pET32b and cells containing two empty plasmids. The control HPLC profiles are provided in the Supporting Information. The standard profile represents authentic compound **1** (1 nmol). B) Total amount of product in cells and spent culture media expressed as equivalents of compound **1**. Data represent the mean ± SE of three biological replicates; $*p < 0.05$ (Student's t-test).

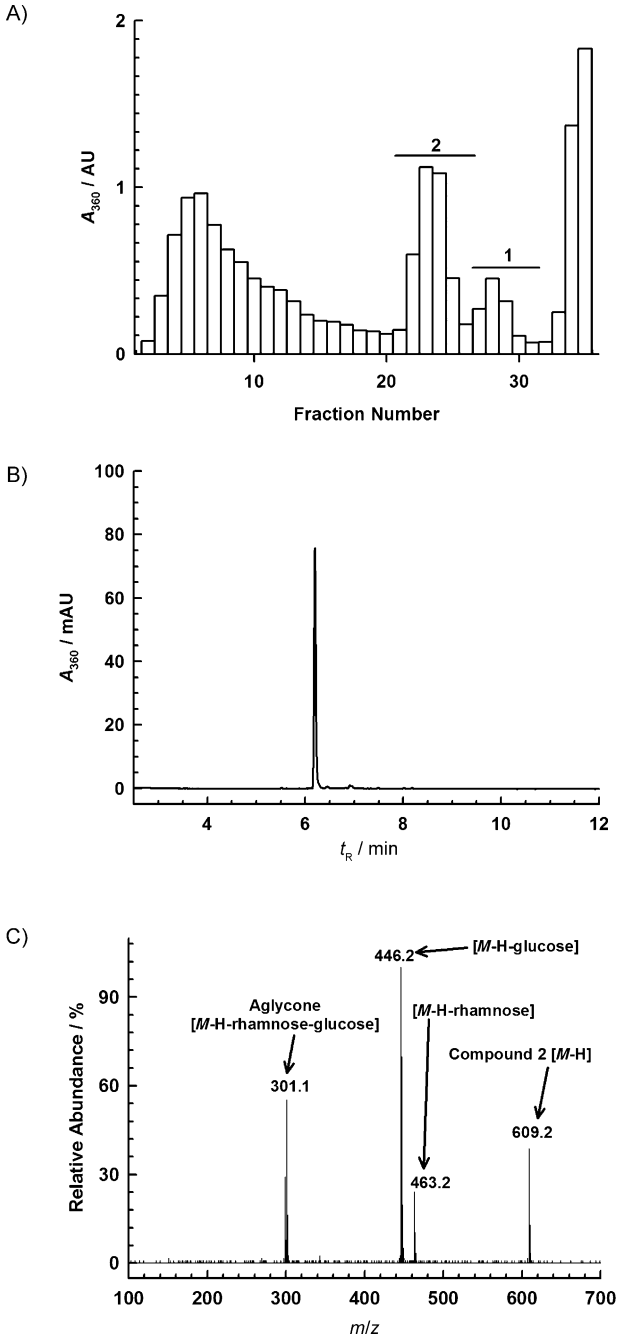


Figure 2. Purification and structural properties of compound **2**. A) Elution profile. After biocatalysis and XAD-2 chromatography, metabolites were applied to a Sephadex LH-20 column. The A_{360} of each elution fraction (10 mL) is shown. B) HPLC-DAD of pooled elution fractions 21–26 (compound **2**). A similar analysis of fractions 27 to 31 revealed a metabolite that co-eluted with an authentic compound **1** standard. HPLC-DAD analysis of all other fractions revealed metabolites with retention times distinct from those of compounds **1** or **2**. C) Quadrupole TOF-MS/MS of the pooled reaction product (LH-20 fractions 21–26) performed in the negative-ion mode implies a molecular mass of 610.2.

has increased specificity for compound **1**. To date, no evidence exists to support these alternative strategies.

Compound **2** displays antioxidant and anti-radical activities.[3] Linden inflorescence extracts containing compound **2** can depress the central nervous system activity and these also act as

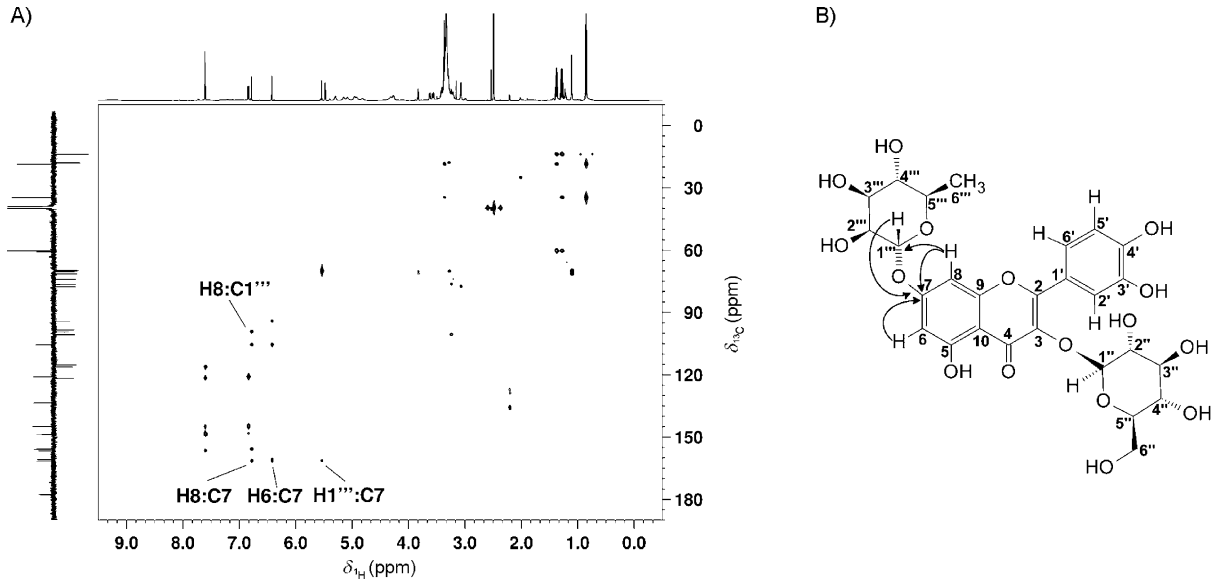


Figure 3. A) $^{1}H,^{13}C$ HMBC spectral map used for demonstrating the interaction between the rhamnose unit and quercetin backbone of compound **2**. B) Structural representation of key HMBC interactions (denoted by arrows).

tranquillisers.[25] The biocatalysis procedure described here has the potential to produce the milligram quantities of **2** required to test its biological/pharmacological activities.

Experimental Section

Cloning: *AtRHM1* cDNA (GenBank accession no. AY081471) was obtained from the *Arabidopsis* Biological Resource Center, amplified with forward (5′-GGATC CAATG GCTTC GTACA CTCCC AAGAA CATT-3′) and reverse (5′-CTCGA GTCAG GTTTT CTTGT TTGGC CCGTA TGCAT A-3′) primers and cloned into pGEMT TA by using standard techniques.[26] The pGEMT TA-*AtRHM1* construct was digested with NcoI/SacI and ligated into the corresponding sites of pET32b (includes N-terminal His_6 tag) and verified by sequencing. *AtUGT89C1* (GenBank accession no. AY093133) cDNA cloned into a pET41b expression vector (includes N-terminal GST and His_6 tags) was provided by the Riken Plant Science Center.[11]

Biocatalysis: *E. coli* BL21(DE3) harbouring both AtRHM1 and AtUGT89C1 were grown in fresh lysogeny broth (50 mL) containing ampicillin and kanamycin (50 mg L^{-1} each) at 37 °C until the A_{600} reached mid-logarithmic growth phase. For controls, *E. coli* harbouring either AtRHM1 or AtUGT89C1 was co-transformed with an empty pET41b or pET32b expression vectors, respectively, and *E. coli* harbouring both empty pET vectors was cultured as described above. Isopropyl-β-D-thiogalactopyranoside (IPTG, 0.1 mM) was added, and the cultures were shaken at 20 °C for 6 h. Quercetin 3-*O*-glucoside (**1**; Extrasynthese, Genay, France) and glucose were added axenically to final concentrations of 20 mg L^{-1} and 4 g L^{-1}, respectively, and the mixture was cultured for 24 h, as described above. Cell pellets were extracted with methanol (10 mL), centrifuged at 5000 *g* for 10 min and dried under vacuum. Spent culture medium was partitioned against an equal volume of water-saturated *n*-butanol (3×),[18] and dried under vacuum. In either case, residues were resuspended in solvent A (10% acetonitrile containing 0.1% formic acid, 200 μL) and passed through a 0.45 μm syringe filter prior to HPLC-DAD analysis.

Purification of quercetin 3-*O*-glucoside-7-*O*-rhamnoside: A culture of *E. coli* (500 mL) dually transformed with AtRHM1 and AtUGT89C1 was cultured as described above. A preliminary experiment displayed no difference in the conversion of **1** at various concentrations (data not shown), hence cultures were supplemented with **1** (40 mg L^{-1}). Methanolic cell extracts were combined with the spent culture media, and the mixture was adjusted to pH 2 and loaded at 6 mL min^{-1} onto a column (2.5×15 cm, i.d.×h) of Amberlite XAD-2 resin. The column was washed with HCl (500 mL, 10 mM, pH 2), followed by HCl (73 mL, 6 μM, pH 5). Quercetin glycosides were eluted (6 mL min^{-1}) with methanol (220 mL), dried under vacuum and re-dissolved in methanol (1 mL, 20% *v/v* in Milli-Q water). The concentrated XAD-2 eluate was passed through a Sephadex LH-20 resin column (2.6×11 cm, i.d.×h; GE Healthcare) connected to an ÄKTA FPLC system that was pre-equilibrated with 20% methanol at a flow rate of 0.85 mL min^{-1}. Substrate and reaction product were eluted initially with a linear gradient of 20–35% methanol over 100 min, followed by a linear gradient of 35–80% methanol over 200 min, collecting 10 mL fractions. Fractions were dried under vacuum, dissolved in solvent A (200 μL) and analysed by HPLC.

HPLC-DAD analysis: Metabolites (5 μL injections) were analysed with an Agilent 1200 HPLC coupled to a diode array detector and a fraction collector (Agilent) and separated on a Kinetex pentafluorophenyl column (100×4.6 mm, 2.6 μm Phenomenex, Torrence, CA). Compound **1** and the reaction product were eluted with a gradient of solvent B (acetonitrile/formic acid 100:0.1) in solvent A of 10–20%, 0–5 min; 20%, 5–10 min, 20–100%, 10–14 min; 100%, 14–17 min at a flow rate of 0.8 mL min^{-1}. HPLC-DAD of quercetin conjugates was performed at 360 nm, and absorption spectra were analysed over the range 230–600 nm. Peak areas (t_R = 6.1 and 8.0 min) were compared to known amounts of an authentic compound **1** standard. No authentic standard for **2** is commercially available.

Structural analyses of the biocatalysis reaction product: The reaction product was analysed by direct infusion on a Waters Q-TOF Ultima Global mass spectrometer (MS; Waters). MS/MS was performed in negative-ion $[M-H]^-$ mode with a cone voltage of 35 V

and a source block temperature of 80 °C. An authentic compound **1** standard fragmented in the range 80–700 *m/z* was used for MS calibration. The product $[M-H]^-$ at 609.2 was subjected to collision-induced dissociation at 30 eV for 15 min. The scan time was 2 s with an interscan delay of 0.1 s. Data were processed with the Masslynx v.4 software provided with the instrument.

To determine the structure of the reaction product, 1D and 2D NMR spectra were taken on a Bruker Avance 600 MHz NMR spectrometer equipped with a 5 mm TCI cryoprobe at 295 K. The reaction product in [D_6]DMSO (800 μL) was transferred to a 5 mm NMR tube. For the ^{1}H NMR experiments, 16 transients were acquired with 2 s relaxation delay with 64 K data points. The 90° pulse duration was 9.0 μs with a spectral width of 8417.5 Hz. ^{13}C NMR experiments were acquired with 755 transients and a 4 s delay for the *J*-modulated spin echo (JMOD) analysis. The 90° pulse duration was 11 s with a spectral width of 32051.3 Hz, and 64 K data points were used. All 2D experiments (^{1}H COSY, ^{13}C HSQC, HMBC) were acquired with 2048 data points for *t*2 increments and 256 for *t*1 increments. NMR spectroscopy data was processed with Topspin v. 2.1.

Acknowledgements

Funding from an NSERC Individual Discovery Grant is gratefully acknowledged. We are also grateful to Richard W. Smith (Mass Spectrometry Facility, University of Waterloo) and Valerie Robertson (NMR Centre, University of Guelph) for technical assistance with HPLC-MS/MS and NMR analyses, respectively.

Keywords: biocatalysis · glucosides · metabolic engineering · rhamnose synthase · rhamnosyltransferase

[1] J. B. Harborne, C. A. Williams, *Phytochemistry* **2000**, *55*, 481 – 504.
[2] S. Burda, W. Oleszek, *J. Agric. Food Chem.* **2001**, *49*, 2774 – 2779.
[3] Y. Wahyuni, R. A. Ballester, E. Sudarmonowati, R. J. Bino, A. G. Bovy, *Phytochemistry* **2011**, *72*, 1358 – 1370.
[4] R. Nakabayashi, M. Kusano, M. Kobayashi, T. Tohge, K. Yonekura-Sakakibara, N. Kogure, M. Yamazaki, M. Kitajima, K. Saito, H. Takayama, *Phytochemistry* **2009**, *70*, 1017 – 1029.
[5] J. M. Routaboul, C. Dubos, G. Beck, C. Marquis, P. Bidzinski, O. Loudet, L. Lepiniec, *J. Exp. Bot.* **2012**, *63*, 3749 – 3764.
[6] J. B. Harborne, H. Baxter, *The Handbook of Natural Flavonoids, Vol. 1*, Wiley, Chichester, **1999**, pp. 297 – 549.
[7] K. Yonekura-Sakakibara, T. Toghe, F. Matsuda, R. Nakabayashi, H. Takayama, R. Niida, A. Watanabe-Takahashi, E. Inoue, K. Saito, *Plant Cell* **2008**, *20*, 2160 – 2176.
[8] K. Saito, K. Yonekura-Sakakibara, R. Nakabayashi, Y. Higashi, M. Yamazaki, T. Toghe, A. R. Fernie, *Plant Physiol. Biochem.* **2013**; DOI: 10.1016/j.plaphy.2013.02.001.
[9] E.-K. Lim, D. J. Bowles, *EMBO J.* **2004**, *23*, 2915 – 2922.
[10] P. C. H. Hollman, J. M. P. van Trijp, M. N. C. P. Buysman, M. S. v. d. Gaag, M. J. B. Mengelers, J. H. M. de Vries, M. B. Katan, *FEBS Lett.* **1997**, *418*, 152 – 156.
[11] K. Yonekura-Sakakibara, T. Toghe, R. Niida, K. Saito, *J. Biol. Chem.* **2007**, *282*, 14932 – 14941.
[12] C. J. Thibodeaux, C. E. Melançon, H.-w. Liu, *Nature* **2007**, *446*, 1008 – 1016.
[13] B. G. Kim, H. K. Kim, J. H. Ahn, *J. Agric. Food Chem.* **2012**, *60*, 11143 – 11148.
[14] H. J. Kim, B. G. Kim, J. H. Ahn, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 5275 – 5282.
[15] E.-K. Lim, D. Ashford, D. J. Bowles, *ChemBioChem* **2006**, *7*, 1181 – 1185.
[16] S. Schaefler, *J. Bacteriol.* **1967**, *93*, 254 – 267.
[17] B. G. Hall, S. Yokoyama, D. H. Calhoun, *Mol. Biol. Evol.* **1983**, *1*, 109 – 124.
[18] E.-K. Lim, D. A. Ashford, B. Hou, R. G. Jackson, D. J. Bowles, *Biotechnol. Bioeng.* **2004**, *87*, 623 – 631.
[19] A. Varma, B. W. Boesch, B. O. Palsson, *Appl. Environ. Microbiol.* **1993**, *59*, 2465 – 2473.
[20] A. M. Cartwright, E.-K. Lim, C. Kleanthous, D. J. Bowles, *J. Biol. Chem.* **2008**, *283*, 15724 – 15731.
[21] K. R. Riihinen, T. Gödecke, G. F. Pauli, *J. Chromatogr. A* **2012**, *1244*, 20 – 27.
[22] L. Kerhoas, D. Aouak, A. Cingoz, J.-M. Routaboul, L. Lepiniec, J. Einhorn, N. Birlirakis, *J. Agric. Food Chem.* **2006**, *54*, 6603 – 6612.
[23] J. A. Yoon, B. G. Kim, W. J. Lee, Y. Lim, Y. Chong, J. H. Lim, *Appl. Biochem. Microbiol.* **2012**, *78*, 4256 – 4262.
[24] R. P. Pandey, S. Malla, D. Simkhada, B. G. Kim, J. K. Sohng, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 1889 – 1901.
[25] L. M. Loscalzo, C. Wasowski, M. Marder, *Phytother. Res.* **2009**, *23*, 1453 – 1457.
[26] J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New York, **1989**.

Received: July 21, 2013
Published online on November 6, 2013