Alexander Gutmann

From mechanistic insight into flavonoid O- and C-glycosyltransferases to their synthetic application

PhD thesis

Graz University of Technology

Institute of Biotechnology and Biochemical Engineering

Supervisor: Univ.-Prof. Dipl.-Ing. Dr.techn. Bernd Nidetzky

Graz, May 2014

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Dissertation identisch.

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/ resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present doctoral dissertation.

_____   _____
         Datum / Date   Unterschrift / Signature

Danksagung

Die vergangenen Jahre, die schlussendlich zum Verfassen dieser Arbeit führten, waren zweifelsfrei ein intensiver und prägender Lebensabschnitt. Gute Zeiten, voller Fortschritt und Erfolg, wechselten sich mit schwierigen Phasen, geprägt von Stillstand und Rückschlägen, ab. Rückblickend führte wohl primär die notwendige Überwindung von anspruchsvollen Hürden zu einer persönlichen sowie fachlichen Reifung. Unerlässlich war hierbei der moralische und fachliche Beistand zahlreicher Personen, die mich während dieses Weges begleiteten und denen ich hiermit danken möchte. Zuallererst danke ich Bernd Nidetzky für die Möglichkeit, an einem vielfältigen und fordernden Thema zu forschen, wodurch ich an der Entstehung und spannenden Entwicklung eines neuen Projekts von Beginn an teilhaben konnte. Besonders hervorheben möchte ich die Bedeutung des guten Arbeitsklimas für einen facettenreichen Arbeitsalltag und die dadurch bedingte beinahe stete Freude und Motivation an der Forschung. Stellvertretend für alle Institutskollegen, Kollaborationspartner und Studienkollegen danke ich hierfür Linda, Corinna, Michele, Ruth, Alexander, Sandra, Elli und Kim als Projektpartner und Wegbegleiter sowie Thomas, Manuel und Patricia als angenehme und inspirierende Bürokollegen. Des Weiteren möchte ich dem DK Molecular Enzymology für die Ermöglichung meines Auslandsaufenthalts in der Gruppe von Ben Davis an der University of Oxford danken. In diesem Zusammenhang ist die Bedeutung der herzlichen Aufnahme und Hilfsbereitschaft meiner neuen Kollegen für eine schnelle Akklimatisierung hervor zu streichen. Neben dem freundschaftlichen Verhältnis zu zahlreichen Personen aus dem Arbeitsumfeld erwies sich der „fachfremde" Freundeskreis als essentieller Rückhalt in schwierigen Zeiten und willkommene Möglichkeit um dem Arbeitsalltag zu entkommen. Zu guter Letzt möchte ich meiner Familie besonderen Dank für die bedingungslose Unterstützung auf meinem Lebensweg aussprechen.

Abstract

Flavonoids are a large and structurally diverse group of polyphenolic plant secondary metabolites with prospective applications as functional food and pharmaceutical ingredients. To circumvent limited bioavailability due to poor water solubility flavonoids are often modified by attachment of sugars. Glycosylations also play key roles in tailoring pharmacological effectiveness, antioxidative properties, taste and color. While stereo- and regioselective glycosylation by chemical synthesis poses major challenges, Leloir glycosyltransferases (GTs) proved to be highly selective in synthesis of numerous natural product glycosides. Besides ubiquitous *O*-glycosides several phenolic *C*-glycosides are also known. These are outstanding due to their largely increased resistance towards hydrolysis which potentially enhances *in vivo* lifetime of pharmaceuticals. Mechanistic insight in enzymatic *C*-glycosylation was obtained by studying glucosylation of the dihydrochalcone acceptor phloretin from uridine 5'-diphosphate (UDP)-glucose by a homologous pair of plant C- and OGTs. *Os*CGT and *Pc*OGT selectively formed the 3'-*C*- (nothofagin) and 2'-*O*-glucoside (phlorizin), respectively. Distinct active site motifs were identified and their exchange allowed partial or complete reversible switch between C- and OGT. This finding could trigger engineering of new CGTs from abundant OGTs to counteract natural scarcity of CGTs. Furthermore, mechanistic evidence for *C*-glycosylation by single nucleophilic displacement at the glucosyl anomeric carbon was provided. Overcoming common obstacles in GT-catalyzed natural product glycosylations like poor acceptor solubility and supply of costly sugar donors was crucial for synthetic applications. *In situ* regeneration of UDP-glucose was accomplished by coupling GTs with sucrose synthase (SuSy) which catalyzes the reversible transfer of glucose from sucrose to UDP. Besides cost efficient sugar donor supply, yield and rate enhancements were accomplished in one pot GT-SuSy cascade reactions. By applying DMSO as co-solvent in combination with periodic acceptor feed the limitation of poor phloretin solubility was successfully overcome. Thorough biochemical characterizations and reaction engineering were crucial to achieve selective and high yielding synthesis of important glucosides of phloretin and its 4'-deoxy analog davidigenin. The phloretin *C*-glucoside nothofagin is a prominent antioxidant from rooibos tea and phlorizin is the major phenolic compound in apple trees. 2'-*O*- (davidioside) and 4'-*O*-glucosides (confusoside) of davidigenin were synthesized for the first time. This was achieved by exploiting complementary regiospecificities of *Pc*OGT and an *Os*CGT variant. Moreover, irreversibility of phloretin *C*- in contrast to *O*-glycosylation was used for quantitative *O*- to *C*-glucoside rearrangement from phlorizin to nothofagin. Either *Pc*OGT and *Os*CGT or a dual specific *Os*CGT variant were applied in one pot conversions. The two step rearrangement involved intermediary formation of phloretin and UDP-glucose in presence of catalytic amounts of UDP.

Zusammenfassung

Flavonoide sind eine große und vielfältige Gruppe sekundärer phenolischer Pflanzenstoffe mit großer Bedeutung als funktionelle Lebensmittel und Inhaltstoffe von Arzneimitteln. Glykosylierung von Flavonoiden kann verminderter biologischer Verfügbarkeit durch Wasserunlöslichkeit entgegenwirken sowie pharmazeutische und antioxidative Wirkung, Farbe und Geschmack verändern. Während stereo- und regioselektive chemische Glykosylierungen eine große Herausforderung darstellen, erwiesen sich Leloir-Glykosyltransferasen als höchst selektiv in der Glykosylierung zahlreicher Naturstoffe. Neben allgegenwärtigen *O*-Glykosiden existieren auch phenolische *C*-Glykoside, die durch außergewöhnliche Hydrolysebeständigkeit hervorstechen. Dadurch könnten die *in vivo* Halbwertszeiten von Arzneimitteln erhöht werden. Der Glukosyltransfer von Uridin-5'-Diphosphat (UDP)-Glukose auf das Dihydrochalkon Phloretin durch ein homologes Paar an pflanzlichen C- (*Os*CGT) und OGTs (*Pc*OGT) resultierte in der selektiven Bildung des 3'-*C*- (Nothofagin) beziehungsweise des 2'-*O*-Glukosids (Phlorizin) von Phloretin. Unterschiedliche Motive konnten in den aktiven Zentren identifiziert werden und deren Austausch führte zu teilweisem oder vollständigem reversiblen Spezifitätswechsel zwischen *C*- und *O*-Glykosylierung. Dies könnte die Entwicklung neuer CGTs aus zahlreichen OGTs stimulieren, um den Mangel an natürlichen CGTs zu kompensieren. Zusätzlich wurden Beweise für direkte nukleophile Substitution am anomeren Kohlenstoff des Zuckers als *C*-Glykosylierungs-Mechanismus gefunden. Die Hauptprobleme in der Biosynthese von glykosylierten Naturstoffen durch GTs sind die geringe Wasserlöslichkeit der Akzeptoren sowie hohe Kosten der aktivierten Zuckerdonoren. *In situ* Regenerierung von UDP-Glukose wurde durch Kopplung der GTs mit Sucrose Synthase (SuSy) erreicht. Hierbei überträgt SuSy in einer reversiblen Reaktion Glukose von Sucrose auf von GTs erzeugtes UDP. Durch GT-SuSy-Kaskadenreaktionen wurden neben kostengünstiger Zufuhr des Zuckerdonors auch Reaktionsgeschwindigkeiten und Ausbeuten verbessert. Durch Beimengung von DMSO als Kosolvent in Kombination mit periodischer Akzeptorzufuhr wurden Limitierungen durch die geringe Löslichkeit von Phloretin beseitigt. Sorgfältige biochemische Analysen und Reaktionsoptimierungen waren ausschlaggebend, um selektiv bedeutende Glukoside von Phloretin und dessen 4'-Deoxy-Analog Davidigenin in präparativem Maßstab herzustellen. Das Phloretin *C*-Glukosid Nothofagin ist ein bedeutendes Antioxidans im Rooibos-Tee und Phlorizin ist die vorherrschende phenolische Verbindung in Apfelbäumen. Erstmals konnten die 2'-*O*- (Davidiosid) und 4'-*O*-Glukoside (Confusosid) von Davidigenin synthetisiert werden. Dies gelang durch Ausnützen der unterschiedlichen Regiospezifitäten von *Pc*OGT und einer *Os*CGT-Variante. Außerdem wurde die Tatsache, dass die *C*- im Gegensatz zur *O*-Glykosylierung von Phloretin irreversibel ist, für eine quantitative *O*-*C*-Glukosid-Umlagerung von Phlorizin zu Nothofagin ausgenutzt. Die Umsetzungen wurden entweder von *Os*CGT und *Pc*OGT oder einer doppelt spezifischen *Os*CGT-Mutante katalysiert. Die zweistufige Umlagerung geht mit einer zwischenzeitlichen Bildung von Phloretin und UDP-Glukose aus katalytischen Mengen an UDP einher.

Contents

Switching between *O*- and *C*-Glycosyltransferase through Exchange of Active-Site Motifs

Angewandte
International Edition Chemie

Enzyme Mechanisms

Switching between *O*- and *C*-Glycosyltransferase through Exchange of Active-Site Motifs**

*Alexander Gutmann and Bernd Nidetzky**

Many natural products derive their biological activity from the sugars attached to their structure.[1] Their glycosylation therefore often determines efficacy in drug use.[2] Engineering the glycosylation pattern of natural products constitutes a promising way of developing new bioactive molecules with tailored pharmacological properties.[3] Diversification of the glycosylation potentially involves exchange of sugar molecules, but also alterations in type and position of the glycosidic bond(s).[4] Most glycosylated natural products are *O*-glycosylated; however, *N*-, *C*-, and *S*-glycosides are also known.[5] Chemically, *C*-glycosides are outstanding in this class because of their pronounced stability to spontaneous and enzyme-catalyzed hydrolysis.[6] *C*-glycosides have therefore aroused particular interest for medicinal applications where use as isofunctional analogues of the corresponding *O*-glycosides potentially offers the important advantage of enhanced in vivo half-life.[7]

Glycosylations in the biosyntheses of natural products are catalyzed by glycosyltransferases (GTs; EC 2.4). These enzymes use an activated donor substrate, typically a nucleoside diphosphate sugar (e.g. UDP-glucose in Scheme 1), for transferring a glycosyl residue onto the reactive group of an acceptor substrate.[8] GTs show exquisite substrate selectivity and are generally recognized as powerful glycosylation tools for both in vitro and in vivo use.[3] However, GTs naturally capable of forming *C*-glycosidic bonds (CGTs) appear to be sparse, limiting their availability and scope for synthesis.[9] Therefore, development of useful *C*-glycosylation catalysts may have to start from existing *O*-glycosyltransferases (OGTs) using protein engineering. Unfortunately, in contrast to enzymatic *O*-glycosyl transfer which has been studied in great detail,[8] the mechanistic principles underlying the corresponding *C*-glycosyl transfer are not well understood, and this presents severe restriction to enzyme design approaches.[10] GT structural features critical for differentiating between *C*- and *O*-glycosyl transfer are not known. A significant paper from Bechthold and colleagues recently demonstrated the swapping of glycosidic bond-type specific-
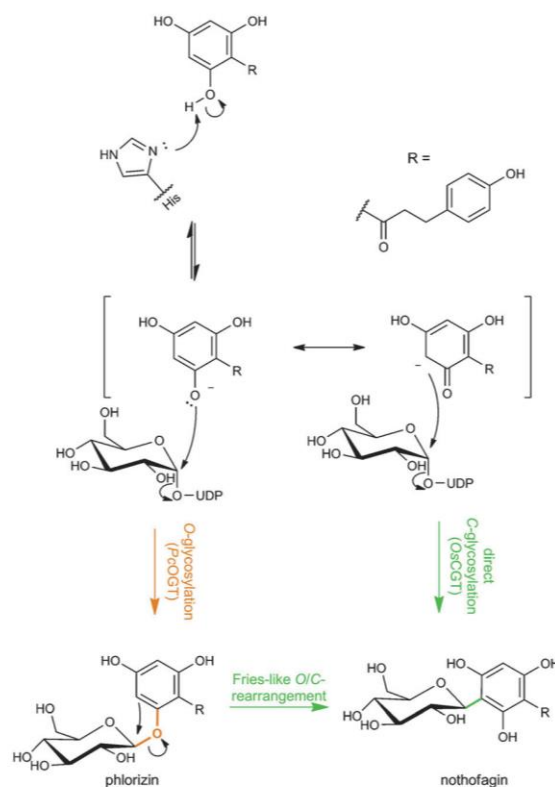
Scheme 1. Formally proposed mechanisms of enzymatic *C*-glycosylation[7b] adapted to the herein examined glucosylation of phloretin by CGT (in green). *C*-glycosylation by means of *O*/*C* rearrangement involves *O*-glycosylation, as catalyzed by OGT (orange), in the first step. Activation of the aryl acceptor substrate by a conserved His is essential in both CGT and OGT.

ity from a bacterial aryl-CGT to a structurally and functionally homologous OGT.[11] Through extensive generation of OGT chimeras that harbored distinct sequence elements from the native CGT,[12] a complete OGT-to-CGT switch was eventually obtained. From protein modeling studies, relevant substitutions in engineered CGT were located within active-site loops that were proposed to adopt highly flexible conformations. However, because of the relatively large number of residue substitutions (≥ 10) required in the native OGT, the molecular interpretation of the specificity change was difficult and a replicable design principle for OGT-to-CGT conversion remained elusive.

We report here the implementation of a reciprocal switch in glycosidic bond-type specificity within a homologous pair

[*] A. Gutmann, Prof. Dr. B. Nidetzky
Institute of Biotechnology and Biochemical Engineering
Graz University of Technology
Petersgasse 12/1, 8010 Graz (Austria)
E-mail: bernd.nidetzky@tugraz.at

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201206141.

WILEY
ONLINE LIBRARY **12879**

2

of plant aryl-OGT and CGT. Both enzymes catalyze glucosyl transfer from UDP-glucose to the dihydrochalcone phloretin, yielding phlorizin (OGT) and nothofagin (CGT) as the product (Scheme 1). Key features of CGT distinguishing it from OGT could therefore be analyzed for two highly similar and conveniently tractable chemical transformations. To achieve specificity switch in each GT, we employed minimal active-site remodeling based on both structural and mechanistic considerations. An Asp residue, which in plant OGT serves to position the catalytic His for function as a Brønsted base[1b] (Scheme 1; Figure 1), is shown to play a key role in
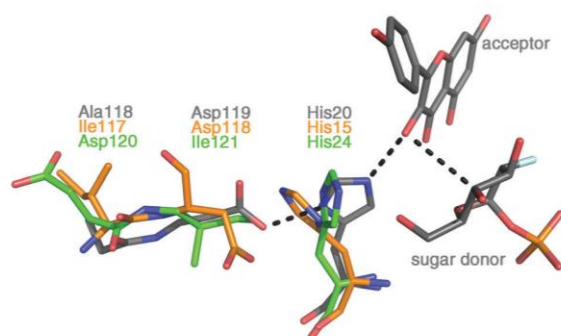
glycosyl transfer in these enzymes (see the Supporting Information; Figure S1). Interestingly, therefore, *Os*CGT features disruption of this "Asp-to-His" arrangement of active-site residues (Figure 1), probably resulting from Ile-Asp to Asp-Ile residue exchange in the sequence of *Os*CGT (Asp120, Ile121) as compared to *Pc*OGT (Ile117, Asp118). The implied relationship between GT structure and catalytic function as OGT or CGT was examined using combinatorial mutagenesis of the Ile-Asp dyad of residues in both enzymes (see Table 1).

Figure 1. Overlay of modeled structures of *Os*CGT (green) and *Pc*OGT (orange) with the experimental structure of the *Vv*OGT ternary complex (gray: PDB code 2C1Z).[13] *Os*CGT shows disruption of the His-Asp dyad present in the two OGTs due to positional shift of the Asp residue.

Table 1: Specific activities of *Os*CGT and *Pc*OGT in wild-type and mutated form, and product distribution from their reactions.

Enzyme	Mutation	Motif	Spec. act.[a] [mU mg^{-1}]	Glycosides [%][a]		
				3'-*C*	2'-*O*	4'-*O*
*Pc*OGT	–	ID	4300	–	100	–
	I117D[b]	DD	>3	<30	>70	–
	D118I[b]	II	>0.3	>90	<10	–
	I117D_D118I[b]	DI	>0.5	>95	–	–
*Os*CGT	–	DI	3300	100	–	–
	I121D	DD	3.4	49	43	8
	D120I	II	0.56	83	7	9
	D120I_I121D	ID	0.078	80	9	11

[a] Specific activity and product distribution were determined by HPLC (100 μM phloretin, 600 μM UDP-glucose, pH 7.0). mU = mUnit = nmol min^{-1}. [b] Owing to low protein expression in *E. coli*, the amount of enzyme usable in the assays was limited and restricted accuracy of the determination to the range given.

OGT–CGT interconversion. Aside from practical ramifications for GT engineering, the evidence presented has also important mechanistic implications, in that it strongly supports aryl-*C*-glycosylation by means of a direct (Friedel–Crafts-like) reaction, while it is not consistent with an alternative two-step reaction consisting of initial *O*-glycosylation followed by a stereochemically controlled Fries-like rearrangement (Scheme 1).[7b] Precise positioning of the acceptor substrate seems to determine the type of glycosidic bond formed, and the conserved His is suggested to play a major role.

As complementary enzymes we chose CGT from *Oryza sativa* (rice; *Os*CGT)[14] and OGT from *Pyrus communis* (pear; *Pc*OGT),[15] which are related to each other by an overall amino acid sequence identity of 30% and common membership to the glycosyltransferase family GT-1. We modeled *Os*CGT and *Pc*OGT against their closest neighbors in plant GTs of the GT-1 family. An active-site structural overlay was made with the closest OGT (*Vv*OGT from *Vitis vinifera*; grape vine) that had had its crystal structure determined in complex with UDP-glucose analogue and aryl acceptor (Figure 1).[13] We found that the position of the catalytic His in the *Os*CGT and *Pc*OGT models is highly similar to that in the experimental structure of *Vv*OGT. Moreover, Asp119 of *Pc*OGT superimposes with the homologous Asp118 of *Vv*OGT which is strongly hydrogen-bonded to His. Both Asp and His are highly conserved in plant OGT sequences,[13,16] suggesting a catalytic consensus motif of *O*-

We obtained *Os*CGT and *Pc*OGT in native and mutated form from *E. coli* cultures that expressed target protein containing N-terminal Strep-tag II for purification (Figure S2). A sensitive HPLC assay for the determination and quantification of the glycosylation products was developed, and product identity was verified using NMR analysis (see the Supporting Information). *Os*CGT formed the 3'-*C*-glycoside nothofagin exclusively, whereas the 2'-*O*-glycoside phlorizin was the sole product of the *Pc*OGT reaction (Scheme 1; Table 1). *Os*CGT mutants to be described later formed a distinct product, which we identified as the 4'-*O*-glycoside (Table 1). The assay was applied for time course analysis of transformations catalyzed by native and mutated enzymes (Figure S3). Table 1 shows the specific activity of each enzyme used and describes the product pattern associated with their reactions. Some GT enzymes promote hydrolysis of their glycosyl donor substrate in a weak side reaction.[17] Mutation potentially enhances this "error hydrolysis". We therefore examined each enzyme in Table 1 for UDP release from UDP-glucose under conditions where phloretin acceptor was absent or present. None showed a detectable hydrolase activity, confirming catalytic behavior of a high-fidelity transferase. Table 1 reveals that site-directed substitutions within Ile-Asp motif of *Pc*OGT went along with substantial loss in specific activity as compared to native enzyme. The effect was even more pronounced in *Os*CGT where modification of the corresponding Asp-Ile motif resulted in a decrease of specific activity by three or more orders of magnitude. However, their low level of activity

3

notwithstanding, mutants of *Pc*OGT and *Os*CGT featuring partial or complete exchange of the active-site motifs showed marked change in glycosidic bond-type specificity, in almost perfect agreement with the hypothesis derived from Figure 1. In *Pc*OGT, OGT-to-CGT specificity switch increased strongly from the Ile117→Asp (I117D) mutant to the Asp118→Ile (D118I) mutant, consistent with the idea that interaction between Asp and His is an essential element of OGT catalytic function and its disruption in the D118I mutant should confer CGT activity. The specificity of the I117D mutant, which showed minor CGT next to main OGT activity, is tentatively explained by weakened His-to-Asp bonding resulting perturbation from the introduced Asp117. The I117D_D118I double mutant of *Pc*OGT, which features complete swap of the OGT by the CGT active-site motif, behaved as a perfect CGT, producing nothofagin as the sole product of glucosyl transfer from UDP-glucose to phloretin.

Proposed structure–function relationships for OGT and CGT were confirmed by reverse engineering of *Os*CGT. Substitution of native Ile121 by Asp potentially reinstalled an OGT-like Asp-to-His active-site group. The resulting I121D mutant is a promiscuous *O/C*-glycosyltransferase that displayed similar levels of OGT and CGT activity. The OGT activity was mainly directed towards the 2′-OH of phloretin, but also included the 4′-OH as alternative glucosylation site. The CGT-to-OGT specificity change due to D120I mutation in *Os*CGT was not the same, but comparable in overall trend to the corresponding OGT-to-CGT specificity change resulting from D118I mutation in *Pc*OGT. The D120I mutant of *Os*CGT produces an Ile-Ile repeat in the active site, just like substitution of Asp118 by Ile does in *Pc*OGT. The *O/C*-glucoside product pattern of the two analogous OGT and CGT mutants was almost identical. Both enzymes showed marked preference for *C*-glycosyl transfer, forming *O*-glucoside in low amounts. These specificities are fully consistent with expectations for enzyme variants that are lacking Asp in a position suitable for catalytic function as OGT. Interestingly, D120I-*Os*CGT differed from D118I-*Pc*OGT in that *Os*CGT mutant was indiscriminate with respect to glucosylation of 2′-OH and 4′-OH of acceptor substrate, while *Pc*OGT mutant was completely specific for reaction at 2′-OH (Table 1). The case of D120I_I121D double mutant of *Os*CGT, however, reveals limitations in the ability of our simple model to predict the magnitude of the specificity change arising from residue substitution within the Ile/Asp active-site motif. While expected to function mainly as OGT, the double mutant rather behaved as an "error-prone" CGT that displayed minor OGT side activity directed towards both 2′-OH and 4′-OH of phloretin. It is known that active-site mutations can have unforeseen consequences on enzyme function, because structural changes caused by them are often not strictly local and affect the active site as a whole.[18] We believe that properties of D120I_I121D double mutant probably reflect such proximally disruptive effects of residue replacements within the *Os*CGT active site. The very low specific activity of this mutant is likely also a manifestation of such secondary effects. Despite these not unusual difficulties, we have clearly identified distinct active-site motifs in plant GTs that define catalytic function as aryl-OGT or aryl-CGT.

The evidence presented immediately suggests a design strategy for engineering glycosidic bond type specificity in plant GTs. Conversion of native OGT into engineered CGT activity would be of special interest for the development of new *C*-glycosylation enzyme catalysts.

We recognized that a dual specific OGT/CGT such as the I121D mutant of *Os*CGT would be instrumental to the mechanistic investigation of enzymatic *C*-glycosyl transfer. If the *C*-glycosylation mechanism involved the rearrangement of initially formed *O*-glycoside (Scheme 1), one would expect that in a reaction catalyzed by a promiscuous OGT/CGT, the relative proportion of *C*-glycoside in the total transfer products formed would increase over time, reflecting the gradual conversion of intermediary *O*-glycosides. The enzymatic reaction by means of direct *O*- and *C*-glycosyl transfer, in contrast, is expected to give a constant ratio of *O/C*-glycosylation products as long as the initial rate conditions apply. We therefore measured time courses of *C*- and *O*-glucoside formation by the I121D mutant and show in Figure 2a that the concentration of each product rose linearly with time, indicating that the molar ratio of glucosyl transfer products remained constant during the reaction. The same time-course analysis was applied to all dual-specific OGT/CGTs in Table 1. The product pattern was invariant with reaction time and the degree of substrate conversion in each case.

We next determined the pH dependence of the glycosidic product formation by the I121D mutant and show our results in Figure 2b. The pH profile of *C*-glycoside synthesis was clearly distinct from the corresponding pH profiles of *O*-glycoside synthesis. Furthermore, the pH change strongly affected the specificity of the enzyme for the glucosylation of the acceptor hydroxy group. The observed pH dependencies seem consistent with the mechanistic proposal for direct *C*-glycosylation. Change in pH might alter the preferred mode of substrate binding to the enzyme and thus influence the specificity for the type of glycosidic linkage formed. Note that phloretin shows multiple ionizations in the pH range examined, its reported pK_a values being 7.0, 9.4, and 10.5.[19] It is more difficult to reconcile pH-dependence data with *C*-glycosylation through the rearrangement mechanism. Almost complete absence of *O*-glucoside formation at high pH would require that catalytic rearrangement occurred at a rate very much faster than the rate of the release of the enzyme-bound *O*-glycoside. A decrease in pH would then have to change this rate ratio in favor of *O*-glycoside dissociation. However, the observation that *O*-glucosylation by the I121D mutant displayed a pH-dependent regioselectivity still necessitates that the accommodation of phloretin at the enzyme's acceptor binding site somehow responds to pH change, such that at pH 7 or lower the glucosylation occurs almost exclusively at the 2′-OH, while at pH 8 or higher the reactivities of 2′-OH and 4′-OH are almost identical.

In further initialkinetic studies with the I121D mutant we analyzed the dependence of the distribution of glucosyl-transfer products on the acceptor substrate concentration used. Figure 2c shows that at low phloretin concentrations, *C*-glucosylation was the predominant path of the enzymatic reaction. At higher acceptor concentrations, however, *O*-
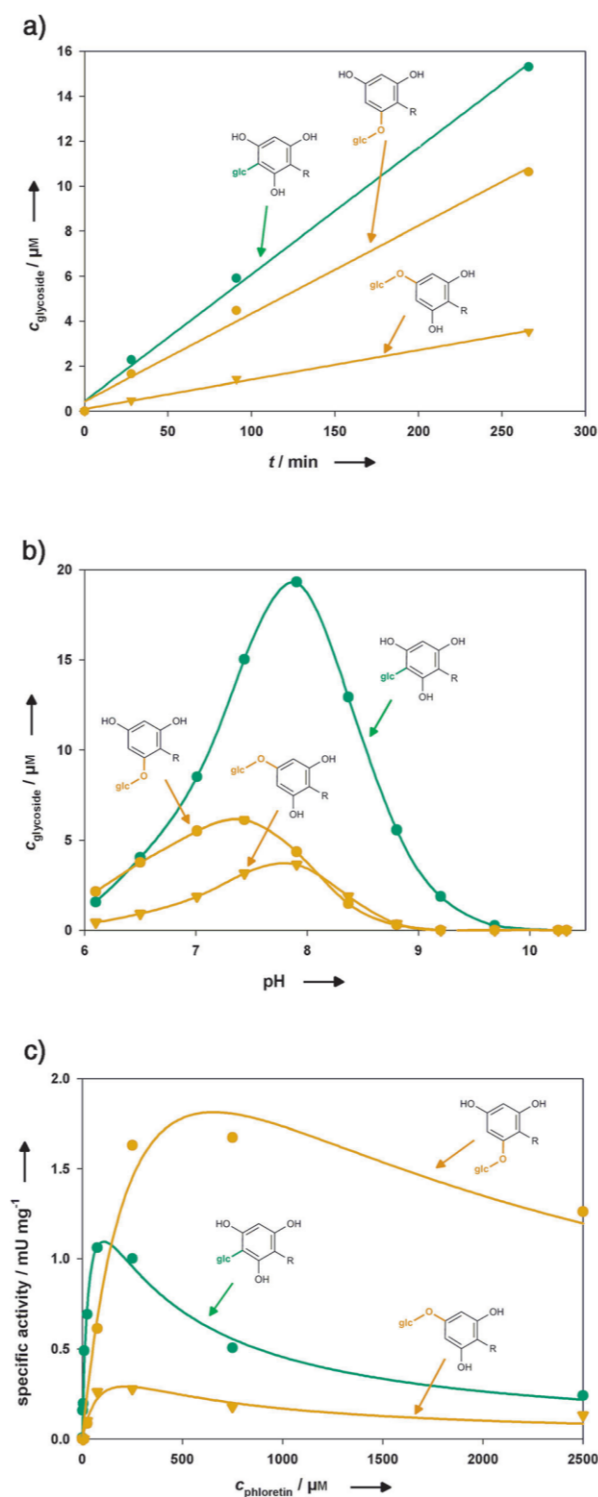
4

Angewandte
Communications

Figure 2. Mechanistic study of *Os*CGT_I121D based on detailed steady-state kinetic analysis of phloretin glucosylation. Time dependence (a) and pH dependence (b) of product formation; c) dependencies of specific rates of product formation on acceptor concentration used.

glucosylation (mainly at 2′-OH, but also at 4′-OH) became increasingly important so that it eventually surpassed *C*-glucosylation, which decreased at high phloretin levels in conjunction with enhanced *O*-glucosyl transfer. Data fitting revealed that the phloretin half-saturation constant for phlorizin formation ($K_M = 324$ μM) was identical within error limit to the phloretin inhibition constant for nothofagin formation (Table S4). According to k_{cat}/K_M, which describes the enzymatic reaction at limiting phloretin concentrations, the I121D mutant prefers nothofagin formation roughly fivefold over phlorizin formation (Table S4). The implied direct competition between *O*- and *C*-glucosyl transfer dependent on the acceptor concentration is unexpected for reaction by means of *O*/*C*-glycoside rearrangement. By contrast, it is in excellent accordance with a direct *O*- and *C*-glycosylation process.

We finally examined the conversion of phlorizin (60 μM) to nothofagin in the presence of I121D (6.7 μM) and UDP (1 mM). There was no reaction above the detection limit (ca. 1%), and phlorizin was stable for 48 h. This result gives strong evidence against a role of catalytic *O*/*C*-rearrangement in the CGT mechanism.

In conclusion, *Pc*OGT and *Os*CGT seem to discriminate between 2′-*O*- and 3′-*C*-glucosylation of phloretin primarily through the relative positioning of the acceptor substrate towards the sugar donor. Partial deprotonation of the phloretin 2′-OH by a conserved His presents a common catalytic feature of both enzymatic reactions. Nucleophilic character is thereby directly generated on O2 and through resonance, on the aromatic C3 in *ortho* position to it (Scheme 1) and *C*-glycosylation is achieved, just like *O*-glycosylation,[8] by a single nucleophilic displacement at the anomeric carbon. Such a reaction can also be described as a Friedel–Crafts-like direct alkylation of the phenolic acceptor in an electrophilic aromatic substitution. The nucleophilic displacement probably involves formation of an oxocarbenium ion-like transition state whereby the positively charged anomeric carbon is stabilized by both the UDP leaving group and the attacking carbanion, as indicated in Scheme 2. Bechthold and co-workers made a similar mechanistic suggestion for a bacterial CGT[11] structurally unrelated to *Os*CGT.

To accommodate the same base-catalytic function from His on an acceptor molecule aligned variably for reaction at O2 or C3, the OGT and CGT active sites appear to have placed their respective His residues in a different microenvironment (Figure 1). Aside from steric effects ("positioning"), other chemical factors of selectivity in competing *O*- and *C*-alkylations of phenoxide ions (e.g. solvent properties, transition-state character) might thus have become optimized in each enzyme.[20] The key notion of this study that a small variation in the active-site structure may result in altered glycosidic bond-type specificity is supported in the literature,[21] where point mutations were used to both graft *N*-glycosylation activity onto OGT and remove it from a dual-specific *O*/*N*-glycosyltransferase. Interestingly, structural change within the His-Asp dyad of OGT (see Figure 1) was thought to be responsible for the unusual *N*-glycoside formation.

5

Angewandte Chemie
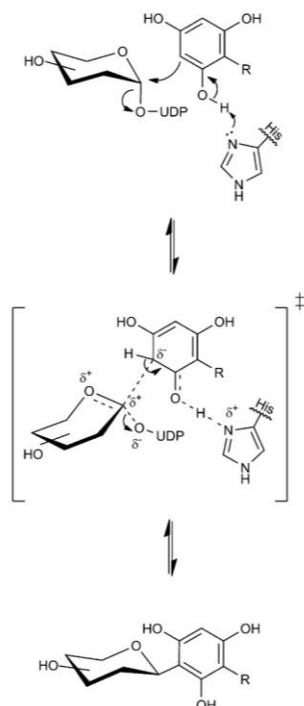
Scheme 2. Proposed mechanism of the aryl-*C*-glycosylation through direct nucleophilic displacement at the anomeric carbon by an aromatic carbanion. The transition state most likely has oxocarbenium ion-like character.

We hope that the advance described herein on the elucidation of structure–function relationships and mechanistic features of plant CGT, in comparison to that of counterpart OGT, will stimulate the generation of novel *C*-glycosylation catalysts through redesign of the active site of the abundant *O*-glycosyltransferases.

[1] a) C. J. Thibodeaux, C. E. Melancon 3rd, H. W. Liu, *Angew. Chem.* **2008**, *120*, 9960–10007; *Angew. Chem. Int. Ed.* **2008**, *47*, 9814–9859; b) X. Q. Wang, *FEBS Lett.* **2009**, *583*, 3303–3309.

[2] B. R. Griffith, J. M. Langenhan, J. S. Thorson, *Curr. Opin. Biotechnol.* **2005**, *16*, 622–630.

[3] G. J. Williams, R. W. Gantt, J. S. Thorson, *Curr. Opin. Chem. Biol.* **2008**, *12*, 556–564.

[4] a) R. W. Gantt, R. D. Goff, G. J. Williams, J. S. Thorson, *Angew. Chem.* **2008**, *120*, 9021–9024; *Angew. Chem. Int. Ed.* **2008**, *47*, 8889–8892; b) G. J. Williams, C. Zhang, J. S. Thorson, *Nat. Chem. Biol.* **2007**, *3*, 657–662.

[5] C. J. Thibodeaux, C. E. Melancon, H. W. Liu, *Nature* **2007**, *446*, 1008–1016.

[6] T. Bililign, C. G. Hyun, J. S. Williams, A. M. Czisny, J. S. Thorson, *Chem. Biol.* **2004**, *11*, 959–969.

[7] a) J. Lee, S. H. Lee, H. J. Seo, E. J. Son, S. H. Lee, M. E. Jung, M. Lee, H. K. Han, J. Kim, J. Kang, J. Lee, *Bioorg. Med. Chem.* **2010**, *18*, 2178–2194; b) C. Dürr, D. Hoffmeister, S. E. Wohlert, K. Ichinose, M. Weber, U. von Mulert, J. S. Thorson, A. Bechthold, *Angew. Chem.* **2004**, *116*, 3022–3025; *Angew. Chem. Int. Ed.* **2004**, *43*, 2962–2965.

[8] L. L. Lairson, B. Henrissat, G. J. Davies, S. G. Withers, *Annu. Rev. Biochem.* **2008**, *77*, 521–555.

[9] M. A. Fischbach, H. N. Lin, D. R. Liu, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 571–576.

[10] T. Bililign, B. R. Griffith, J. S. Thorson, *Nat. Prod. Rep.* **2005**, *22*, 742–760.

[11] J. Härle, S. Gunther, B. Lauinger, M. Weber, B. Kammerer, D. L. Zechel, A. Luzhetskyy, A. Bechthold, *Chem. Biol.* **2011**, *18*, 520–530.

[12] M. Mittler, A. Bechthold, G. E. Schulz, *J. Mol. Biol.* **2007**, *372*, 67–76.

[13] W. Offen, C. Martinez-Fleites, M. Yang, E. Kiat-Lim, B. G. Davis, C. A. Tarling, C. M. Ford, D. J. Bowles, G. J. Davies, *EMBO J.* **2006**, *25*, 1396–1405.

[14] M. Brazier-Hicks, K. M. Evans, M. C. Gershater, H. Puschmann, P. G. Steel, R. Edwards, *J. Biol. Chem.* **2009**, *284*, 17926–17934.

[15] C. Gosch, H. Halbwirth, B. Schneider, D. Holscher, K. Stich, *Plant Sci.* **2010**, *178*, 299–306.

[16] a) L. N. Li, L. V. Modolo, L. L. Escamilia-Trevino, L. Achnine, R. A. Dixon, X. Q. Wang, *J. Mol. Biol.* **2007**, *370*, 951–963; b) L. V. Modolo, L. N. Li, H. Y. Pan, J. W. Blount, R. A. Dixon, X. Q. Wang, *J. Mol. Biol.* **2009**, *392*, 1292–1302; c) H. Shao, X. Z. He, L. Achnine, J. W. Blount, R. A. Dixon, X. Q. Wang, *Plant Cell* **2005**, *17*, 3141–3154.

[17] G. Sugiarto, K. Lau, J. Qu, Y. Li, S. Lim, S. Mu, J. B. Ames, A. J. Fisher, X. Chen, *ACS Chem. Biol.* **2012**, *7*, 1232–1240.

[18] N. Tokuriki, D. S. Tawfik, *Science* **2009**, *324*, 203–207.

[19] C. Valenta, J. Cladera, P. O'Shea, J. Hadgraft, *J. Pharm. Sci.* **2001**, *90*, 485–492.

[20] R. Breslow, K. Groves, M. U. Mayer, *J. Am. Chem. Soc.* **2002**, *124*, 3622–3635, and references therein.

[21] M. Brazier-Hicks, W. A. Offen, M. C. Gershater, T. J. Revett, E. K. Lim, D. J. Bowles, G. J. Davies, R. Edwards, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 20238–20243.

6

Supporting Information

Switching between *O*- and *C*-Glycosyltransferase through Exchange of Active-Site Motifs**

*Alexander Gutmann and Bernd Nidetzky**

anie_201206141_sm_miscellaneous_information.pdf

Table of contents

1.   Methods

1.1   Chemicals and reagents

Unless otherwise mentioned, all chemicals were obtained from Sigma-Aldrich in the highest purity available. DNA modifying enzymes were from Fermentas and PCR primers were purchased from Life Technologies. Phusion® High-Fidelity DNA Polymerase was purchased from New England Biolabs. *Strep*-Tactin® Sepharose® and desthiobiotin were from IBA.

1.2   Strain construction

The *Os*CGT gene (GenBank: FM179712) was received as a kind gift from the group of Prof. Robert Edwards (Centre for Bioactive Chemistry, Durham University, UK). It was provided in a pET-STRP3 vector which is a custom made derivative of pET-24d that enables protein expression with an N-terminally fused *Strep*-tag II.[1] The gene encoding for *Pc*OGT (UGT88F2; GenBank: FJ854496) was kindly provided by the group of Prof. Karl Stich (Institute of Chemical Engineering, Vienna University of Technology, Austria) in a pYES2.1/V5-His-TOPO vector.[2] Before cloning it into the pET-STRP3 vector for expression as N-terminal *Strep*-tag II fusion protein, internal restriction sites for *Nde*I and *Xho*I, respectively were removed by overlap extension (OE)-PCR. Three fragments were amplified by PCR with the primers described in Table S1 and S2. After linking the overlapping fragments by OE-PCR the resulting construct was purified on an agarose gel, digested with *Nde*I and *Xho*I restriction enzymes and cloned into the respective sites of the pET-STRP3 vector. Active site mutants were created by two-stage PCR using the primers described in Table S1 and Table S2. Before transformation of *E. coli* BL21-Gold(DE3) *Dpn*I was added to the amplification product to remove the parental template. The introduction of the desired mutations was verified by sequencing the complete genes.

Table S1: Primers used for amplification of OE-PCR fragments and creation of active site mutants

construct	forward primer	reverse primer
*Pc*OGT OE-PCR 5'-fragment	PcOGT_NdeI_fw	PcOGT_XhoI_rem_rv
*Pc*OGT OE-PCR middle fragment	PcOGT_XhoI_rem_fw	PcOGT_NdeI_rem_rv
*Pc*OGT OE-PCR 3'-fragment	PcOGT_NdeI_rem_fw	PcOGT_XhoI_rv
*Pc*OGT-I117D	PcOGT_I117D_fw	PcOGT_I117D_rv
*Pc*OGT-D118I	PcOGT_D118I_fw	PcOGT_D118I_rv
*Pc*OGT-I117D_D118I	PcOGT_I117D_D118I_fw	PcOGT_I117D_D118I_rv
*Os*CGT-D120I	OsCGT_D120I_fw	OsCGT_D120I_rv
*Os*CGT-I121D	OsCGT_I121D_fw	OsCGT_I121D_rv
*Os*CGT-D120I_I121D	OsCGT_D120I-I121D_fw	OsCGT_D120I-I121D_rv

Table S2: Sequences of PCR primers. Mismatched bases are underlined.

primer	sequence
PcOGT_NdeI_fw	TAACCATATGGGAGACGTCATTGTACTGTACGC
PcOGT_XhoI_rem_rv	GCTAGGTGGCTCCAGCTCTTCGAACGTGTTG
PcOGT_XhoI_rem_fw	CACGTTCGAAGAGCTGGAGCCACCTAGCGTC
PcOGT_NdeI_rem_rv	CATTCCTGTTCATGTGCTGCTCCGCGTAAAGC
PcOGT_NdeI_rem_fw	ACGCGGAGCAGCACATGAACAGGAATGTTC
PcOGT_XhoI_rv	GGTGCTCGAGCTATGTAATGCTACTAACAAAGTTGACCAAG
PcOGT_I117D_fw	GCCTTCATCGATGACCTCTTCTGCACCTCCGCTCTTCC
PcOGT_I117D_rv	GCAGAAGAGGTCATCGATGAAGGCGCGAACGGTGG
PcOGT_D118I_fw	GCCTTCATCATCATTCTCTTCTGCACCTCCGCTCTTCC
PcOGT_D118I_rv	GCAGAAGAGAATGATGATGAAGGCGCGAACGGTGG
PcOGT_I117D_D118I_fw	GCCTTCATCGACATCCTCTTCTGCACCTCCGCTCTTCC
PcOGT_I117D_D118I_rv	GCAGAAGAGGATGTCGATGAAGGCGCGAACGGTGG
OsCGT_I121D_rv	GATGTCAGCGCATCGTCCGTGGCGAGCGCCGAC
OsCGT_I121D_fw	CTCGCCACGGACGATGCGCTGACATCCGTCG
OsCGT_D120I-I121D_rv	GATGTCAGCGCATCGATCGTGGCGAGCGCCGAC
OsCGT_D120I-I121D_fw	CTCGCCACGATCGATGCGCTGACATCCGTCGTC
OsCGT_D120I_rv	GATGTCAGCGCGATAATCGTGGCGAGCGCCGAC
OsCGT_D120I_fw	CTCGCCACGATTATCGCGCTGACATCCGTCGTC

1.3 Enzyme expression and purification

The described *E. coli* strains were cultivated in 1 L baffled shake flasks at 37°C and 120 rpm using 300 ml LB-media containing 50 µg/ml kanamycin until they reached an optical density at 600 nm of 0.8-1.0. At that point the protein expression was induced by adding 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and the temperature was decreased to 25°C. After 18 h the cells were harvested by 30 min centrifugation at 5000 rpm and 4°C. The cells were resuspended in water and stored at -70°C until disruption by repeated passage through a cooled French press at 100 bar. The cell debris was removed by centrifugation for 45 min at 13200 rpm and 4°C. The enzymes were purified by *Strep*-tag affinity chromatography on a 3 ml gravity flow *Strep*-Tactin® Sepharose® column as recommended by the manufacturer IBA. The column was equilibrated with 3 column volumes (CV) of washing buffer W (100 mM Tris/HCl pH 8, 150 mM NaCl, 1 mM EDTA). Cell extract was filtrated through a 1.2 µm cellulose-acetate filter and diluted twofold with buffer W before loading on the column. After washing with 5 CVs of buffer W the proteins were eluted with 3 CVs buffer E (100 mM Tris/HCl pH 8, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin) whereas the first 0.5 CVs were discarded and the rest was pooled. The column was regenerated using 15 CVs of buffer R (100 mM Tris/HCl pH 8, 150 mM NaCl, 1 mM EDTA, 1 mM hydroxy-azophenyl-benzoic acid) and equilibrated with 10 CVs of buffer W. Between the purification of different enzymes or mutants thereof the column was washed with 8 M Guanidin-HCl to eliminate cross-contaminations. The proteins were concentrated and buffer exchanged to 25 mM HEPES pH 7 using centrifugal concentrators with a Molecular Weight Cut Off of 10 kDa. Aliquots of the enzymes were stored at -20°C. Protein concentrations were measured by BCA Protein Assay (Thermo Scientific) on a FLUOstar Omega plate reader from BMG Labtech. The protein concentrations of purified enzymes were furthermore measured on a NanoDrop 2000 system (Thermo Scientific). Therefore molecular weight and molar extinction coefficients were calculated using Peptide Properties Calculator.

1.4 HPLC-based activity assay

Initial rate measurements were done with an HPLC based assay at 30°C. Under standard conditions 600 µM UDP-glc and 100 µM phloretin were used as substrates at pH 7 in a buffer containing 50 mM HEPES, 13 mM MnCl$_2$, 50 mM KCl, 0.13% BSA and 5% ethanol. Reactions were started with the addition of glycosyltransferase and samples were taken by mixing an aliquote of 100 µl with 100 µl acetonitrile to stop the reaction. Precipitated protein was removed by centrifugation for 15 min at 13200 rpm. At least 4 different points in time were used to calculate linear initial rates.

10 µl of the supernatant were used for analysis on an Agilent 1200 HPLC equipped with a Chromolith® Performance RP-18e endcapped column (100–4.6 mm) from Merck. The column was thermostatically controlled at 35°C and the separation was monitored by UV detection at 288 nm. Separation of phloretin and its glycosides was achieved by following method using water with 0.1 % TFA as buffer A and acetonitrile with 0.1% TFA as buffer B, respectively. A 7.5 min long linear gradient from 20 to 47.5% B (1 ml/min) was used for product separation. It was followed by 0.05 min of a linear gradient from 47.5 to 100 % B (1 ml/min) and 1.45 min of isocratic flow at 100% B (1.5 ml/min) to wash off hydrophobic compounds. After a 0.05 min linear gradient

2

from 100 to 20 % B (1.5 ml/min) an isocratic flow of 2.45 min at 20% B (1.5 ml/min) was applied to equilibrate the column.

1.5 pH-profile (*Os*CGT_I121D)

The standard activity protocol was modified by using buffer mixtures (25 mM HEPES, 25 mM Tris and 50 mM CAPS) of the respective pH and by replacing $MnCl_2$ with 13 mM $MgCl_2$. Buffers were prepared from pH 6 to 11 in steps of 0.5 pH units. The actual pH of the reaction mixture was determined as the average of pH measurements at the beginning and at the end of the observed time span.

1.6 Variation of acceptor concentration (*Os*CGT_I121D)

For determining the effect of acceptor concentration on product formation by *Os*CGT_I121D the standard activity protocol was modified by varying the concentration of phloretin between 2.5 µM and 2.5 mM (0, 2.5, 5, 10, 25, 75, 250, 750, 2500 µM). The kinetic parameters for half-saturation constant (K_M), turnover number (k_{cat}) and substrate inhibition (K_i) were obtained separately for the formation of 2'-*O*-, 4'-*O*- and 3'-*C*-glycoside by nonlinear regression using equation 1 and 2. Whereby v is the observed reaction rate (mM min^{-1}), v_{max} is the maximal initial rate (mM min^{-1}) and E is the enzyme concentration (mM).

$$v = \frac{v_{max} \cdot c_{phloretin}}{K_M + c_{phloretin} \cdot (1 + \frac{c_{phloretin}}{K_i})} \tag{1}$$

$$k_{cat} = \frac{v_{max}}{E} \tag{2}$$

1.7 UDP-glc hydrolysis

Parallel to the standard activity assay the same amount of enzyme was incubated without the acceptor phloretin under otherwise identical conditions. Aliquots of both reactions were taken over time and stopped by mixing with an equal volume of acetonitrile. The samples were prepared for HPLC-analysis as described but a different HPLC program was used to separate UDP and UDP-glc. 20 µl of sample were injected, the column was thermostatically controlled at 30°C and the separation was monitored by UV detection at 254 nm. Using 20 mM potassium phosphate (pH 6.8) with 2 mM tetrabutylammonium hydrogen sulfate as buffer A and acetonitrile as buffer B following gradient was applied with a constant flow rate of 2 mL/min: 3 min of a linear gradient from 0 to 2% B were followed by a 7 min long linear gradient from 2 to 25% B to separate UDP from UDP-glc. During 2 min of isocratic flow at 25% B hydrophobic compounds were washed of and after a 1 min long gradient from 25 to 0% B, 2 min of isocratic flow at 0% B were applied to equilibrate the column. The last samples of the reactions with acceptor were furthermore measured with the standard HPLC-activity assay to correlate UDP and glycoside formation.

1.8 *O/C*-glycoside rearrangement (*Os*CGT_I121D)

Possible catalysis of *O/C*-glycoside rearrangement by *Os*CGT_I121D was tested by incubation of 60 µM phlorizin and 1 mM UDP with 6.7 µM enzyme (~3.4 mg/ml) for 48 h under conditions otherwise equivalent to the standard activity measurements. The glycosylation pattern was analyzed over time per HPLC as described for the standard activity assay.

1.9 Product identification by NMR

Reactions of *Os*CGT and *Os*CGT_I121D were upscaled and carried out with following alterations of standard reaction conditions. OsCGT: 1 mM phloretin, 2 mM UDP-glc (8.7 mg theoretical yield); OsCGT_I121D: 1 mM phloretin, 1.5 mM UDP-glc, pH 8 (7.4 mg theoretical yield). The dihydrochalcones were isolated by repeated extraction with ethyl acetate. The solvent was removed under reduced pressure and samples were dissolved in MeOD-d$_4$. ^{1}H NMR and ^{13}C NMR spectra were recorded on a Varian Unity Inova 500 MHz spectrometer.

1.10 Sequence alignment and protein structure prediction

The protein sequence of *Os*CGT was aligned with that of several plant OGTs using the multiple sequence alignment program ClustalW2.[3] The tertiary protein structures of *Os*CGT and *Pc*OGT were predicted by the I-TASSER server.[4]

3

2. Results:

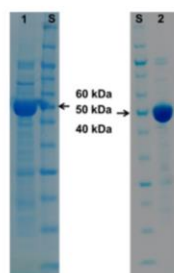### 2.1 Sequence alignment

Figure S1: Sequence alignment of *OsCGT* with plant OGTs (1-7) shows that the catalytic His-Asp dyad (red background) is conserved in plant OGTs. However, like in the structural overlay (Figure 1) a shift of Asp is predicted for *OsCGT* resulting in an Asp/Ile exchange compared to *PcOGT*. Besides the active site motif there are no obvious differences between C- and OGT and generally only few amino acids are conserved (grey background). They are mainly clustered in the plant secondary product glycosyltransferase (PSPG) motif which is responsible for UDP-glc binding (black box). For enzymes 1-4 active site mutants revealed that both Asp and His are essential for *O*-glycoside formation and interaction of these residues was confirmed by crystal structures. Enzymes 5-7 are able to catalyze the *O*-glycosylation of phloretin to phlorizin. Therefore the acceptor binding pockets are expected to be highly similar to that of *OsCGT*. Enzymes: *VvOGT* (P51094) from *Vitis vinifera* (grape vine), UGT71G1 (AAW56092), UGT71A16 (DQ875463) and UGT78G1 (A6XNC6) from *Medicago truncatula* (barrel clover), UGT71A16 (FJ854494), UGT71K2 (FJ854495) and *PcOGT* (FJ854496) from *Pyrus communis* (pear), *OsCGT* (FM179712) from *Oryza sativa* (rice)

4

11

2.2 Protein purification

Figure S2: SDS-PAGE of enzymes purified by *Strep*-tag affinity chromatography: lane 1: *Pc*OGT (55.4 kDa); lane 2: *Os*CGT (51.3 kDa); lane S: Novex® Sharp Unstained Protein Standard (Life Technologies)

2.3 Identification of glycosylation products

Glycosylation products could be clearly distinguished and quantified by reversed phase HPLC (Figure S3). Authentic standards were available for the substrate phloretin and the 2'-*O*-glycoside phlorizin. In agreement with literature the reaction product of *Os*CGT was identified by ^{1}H and ^{13}C NMR as the 3'-*C*-glycoside nothofagin.[5] ^{1}H NMR (499.91 MHz, MeOD-d$_4$) δ: 7.06 (d, J = 8.0 Hz, 2H), 6.70 (d, J = 8.2 Hz, 2H), 5.94 (s, 1H), 4.85 (overlap with H$_2$O signal), 3.97 (t, J = 9.3 Hz, 1H), 3.87 (d, J = 11.7 Hz, 1H), 3.76 (dd, J = 12.2, 5.5 Hz, 1H), 3.25-3.42 (undissolved) 2.87 (t, J = 7.7 Hz, 2H). ^{13}C NMR (125.70 MHz, MeOD-d$_4$) δ: 206.70, 165.71, 164.84, 164.03, 156.44, 133.92, 130.32, 116.09, 105.42, 104.28, 95.87, 85.22, 79.97, 76.05, 73.22, 71.54, 62.55, 47.45, 31.36.
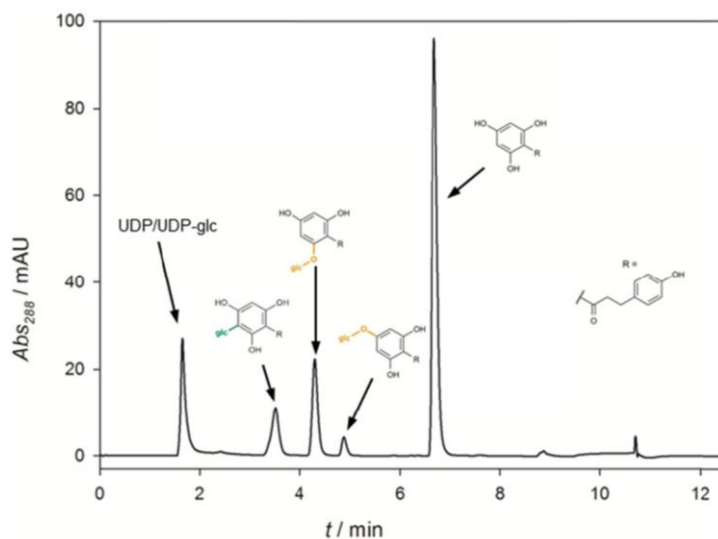
Figure S3: HPLC chromatogram showing base line separation of phloretin (6.7 min) and its glycosides (nothofagin 3.5 min, phlorizin 4.3 min and 4'-*O*-glycoside 4.9 min) after partial conversion of 75 µM phloretin by *Os*CGT_I121D.

HPLC chromatograms of *Os*CGT_I121D (Figure S3) clearly showed three distinct glycosylation products whereby the retention times of two were assigned to nothofagin and phlorizin, respectively. The third glycosylation product was identified as the 4'-*O*-glycoside by analyzing the aromatic ^{1}H signals of the reaction mixture (Table S3 and Figure S4). The shifts of the protons at the B-ring were essentially the same for the aglycon and all of its glycosides indicating exclusive glycosylation at the A ring. Signals of nothofagin, phloretin and phlorizin were assigned by comparison with ^{1}H NMR spectra of the *Os*CGT conversion experiment or of authentic standards, respectively. Only one signal could not be assigned to those compounds: δ 6.10, s, 0.14H. It is well corresponding to the expected signal for two equivalent aromatic protons at positions 3' and 5'.[2]
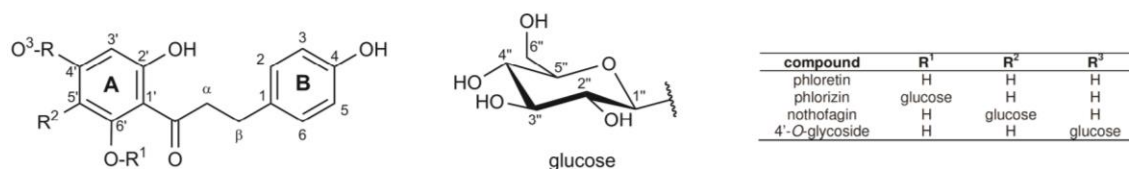
5

Figure S4: Structures and atom numbering of the dihydrochalcone phloretin and the respective glycosides

Table S3: Assignment of aromatic hydrogens to *Os*CGT_I121D conversion products. ^{1}H NMR (499.90 MHz, MeOD-d$_4$)

^{1}H signal	position(s)	compound(s)
7.05/7.07 (d, J = 8.15 Hz, 2.04H)[a]	2, 6	phloretin, nothofagin, phlorizin, 4'-*O*-glycoside
6.69 (d, J = 8.15 Hz, 2H)	3, 5	phloretin, nothofagin, phlorizin, 4'-*O*-glycoside
6.19 (d, J = 1.79 Hz, 0.5H)	3'	phlorizin
6.10 (s, 0.14H)	3', 5'	4'-*O*-glycoside
5.96 (d, J = 1.82 Hz, 0.39H)	5'	phlorizin
5.93 (s, 0.17H)	5'	nothofagin
5.82 (s, 0.20H)	3', 5'	phloretin

[a] overlap of two slightly shifted but otherwise identical peaks

2.4   Kinetic characterization of *Os*CGT_I121D

Table S4: Kinetic constants of *Os*CGT_I121D determined at different acceptor concentrations

product	k_{cat} / s^{-1}	K_M / µM	k_{cat}/K_M / s^{-1}µM^{-1}	K_I / µM
3'-*C*-glycoside (nothofagin)	1.50 ± 0.17	32.5 ± 7.2	$46.3 \pm 15.5 \times 10^{-3}$	352 ± 83
2'-*O*-glycoside (phlorizin)	3.09 ± 0.94	324 ± 162	$9.5 \pm 7.6 \times 10^{-3}$	1325 ± 738
4'-*O*-glycoside	0.50 ± 0.24	111 ± 86	$4.5 \pm 5.7 \times 10^{-3}$	414 ± 338

3.   References

[1]   a) D. P. Dixon, T. Hawkins, P. J. Hussey, R. Edwards, *J. Exp. Bot.* **2009**, *60*, 1207-1218; b) M. Brazier-Hicks, K. M. Evans, M. C. Gershater, H. Puschmann, P. G. Steel, R. Edwards, *J. Biol. Chem.* **2009**, *284*, 17926-17934.

[2]   C. Gosch, H. Halbwirth, B. Schneider, D. Holscher, K. Stich, *Plant Sci.* **2010**, *178*, 299-306.

[3]   a) M. Goujon, H. McWilliam, W. Z. Li, F. Valentin, S. Squizzato, J. Paern, R. Lopez, *Nucleic Acids Res.* **2010**, *38*, W695-W699; b) M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, D. G. Higgins, *Bioinformatics* **2007**, *23*, 2947-2948.

[4]   a) A. Roy, A. Kucukural, Y. Zhang, *Nat. Protoc.* **2010**, *5*, 725-738; b) Y. Zhang, *Proteins: Struct., Funct., Bioinf.* **2007**, *69*, 108-117.

[5]   a) A. Yepremyan, B. Salehani, T. G. Minehan, *Org. Lett.* **2010**, *12*, 1580-1583. b) N. Krafczyk, M. A. Glomb, *J. Agric. Food Chem.* **2008**, *56*, 3368-3376.

Enzymatic *C*-glycosylation: Insights from the study of a complementary pair of plant *O*- and *C*-glucosyltransferases

Enzymatic *C*-glycosylation: Insights from the study of a complementary pair of plant *O*- and *C*-glucosyltransferases*

Alexander Gutmann and Bernd Nidetzky‡

Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, A-8010 Graz, Austria

Abstract: *C*-Glycosylation presents a rare mode of sugar attachment to the core structure of natural products and is catalyzed by a special type of Leloir *C*-glycosyltransferases (C-GTs). Elucidation of mechanistic principles for these glycosyltransferases (GTs) is of fundamental interest, and it could also contribute to the development of new biocatalysts for the synthesis of valuable *C*-glycosides, potentially serving as analogues of the highly hydrolysis-sensitive *O*-glycosides. Enzymatic glucosylation of the natural dihydrochalcone phloretin from UDP-D-glucose was applied as a model reaction in the study of a structurally and functionally homologous pair of plant glucosyltransferases, where the enzyme from rice (*Oryza sativa*) was specific for *C*-glycosylation and the enzyme from pear (*Pyrus communis*) was specific for *O*-glycosylation. We show that distinct active-site motifs are used by the two enzymes to differentiate between *C*- and *O*-glucosylation of the phloretin acceptor. An enzyme design concept is therefore developed where exchange of active-site motifs results in a reversible switch between *C*/*O*-glycosyltransferase (C/O-GT) activity. Mechanistic proposal for enzymatic *C*-glycosylation involves a single nucleophilic displacement at the glucosyl anomeric carbon, proceeding through an oxocarbenium ion-like transition state. Alternatively, the reaction could be described as Friedel–Crafts-like direct alkylation of the phenolic acceptor.

Keywords: active-site motifs; biocatalysis; carbohydrates; catalytic mechanisms; engineering; enzyme mechanisms; *C*-glycosides; glycosyltransferases; natural product synthesis; structure–function.

INTRODUCTION

Many natural products are synthesized as glycosides [1]. Chemically, in the most general sense, glycosylation enhances the water solubility of the often highly hydrophobic core structures of natural products, and thus, it often serves to augment their bioavailability [2]. Furthermore, sugar molecules are important in numerous biological recognition processes [3]. The biological activity and selectivity of natural products are therefore frequently determined by the glycosyl residues attached to them during the biosynthesis [4]. Natural products constitute a vital source of therapeutic agents, and glycosylation can therefore decisively affect their potency as drugs [5]. Glycodiversification, the targeted modification and extension of a compound's glycosylation, is an interesting approach of developing new bio-

active substances with altered or improved properties for medicinal applications [6]. Other promising uses of glycoengineered natural products are as active cosmetic ingredients and functional food additives [7]. Glycodiversification approaches usually involve exchange of sugar molecule(s), but there is also the possibility that the type and position of the glycosidic linkage are modified [8,9].

Glycosylation of natural products usually occurs via *O*-glycosidic linkages, which represent the most prevalent type of glycosidic bond in Nature [10]. However, there are also *N*-glycosidic natural products, and in some quite rare cases, a *C*-glycosidic linkage, typically involving the carbon atom from an electron-rich aromatic aglycon, is formed during natural product glycosylation [11]. *C*-glycosides have previously attracted high interest in medicinal chemistry because it was recognized that they could serve as functional analogues of the corresponding *O*-glycosides [12]. The most striking feature of the *C*-glycosidic linkage compared to its *O*-glycosidic counterpart is the strongly increased resistance to spontaneous and enzyme-catalyzed hydrolysis [13]. Therefore, *C*-glycosides might show biological activity similar to the one of the parent *O*-glycoside, but display at the same time a substantially improved stability against in vivo clearance [14]. The case of dapagliflozin (Scheme 1) serves to illustrate this point. Dapagliflozin is a candidate drug recently recommended by the European Medicines Agency for authorization of treatment of type 2 diabetes. Dapagliflozin is designed to inhibit the sodium-glucose cotransporter 2 (SGLT2) in the kidney, which allows glucose to be reabsorbed into the bloodstream [15]. Due to the compound's action, therefore, glycemic control is improved in diabetes patients without increasing their insulin secretion [16]. The plasma half-life of dapagliflozin in rat is reported as 4.6 h [17]. Phlorizin (Scheme 1) was considered early on as an SGLT2 inhibitor [18]. Phlorizin is widespread in the peelings of pears, apples, cherries, and other fruits [19]. However, its *O*-glycosidic linkage rendered phlorizin susceptible to rapid clearance in vivo and so the compound could not be used clinically [14]. A natural *C*-glycosidic analogue of phlorizin is the antioxidant nothofagin (Scheme 1), which is found in substantial amounts in rooibos herbal tea [20]. Even though a number of *C*-glycosides have been isolated from natural sources, the enzymes responsible for their biosynthesis are known only in very few cases, and biocatalytic approaches of *C*-glycoside production have yet to be established [11]. Therefore, the synthesis of dapagliflozin, just as those of various other *C*-glycosides, has been accomplished purely chemically [11,21,22].
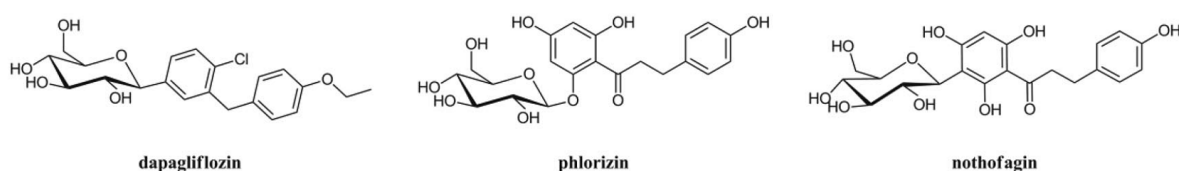
Scheme 1 SGLT2 inhibitors dapagliflozin and phlorizin and the related *C*-glucoside nothofagin.

GLYCOSYLTRANSFERASES: NATURAL CATALYSTS FOR *C*-GLYCOSIDE SYNTHESIS

Glycosylation of natural products is usually performed by classical Leloir glycosyltransferases (GT; EC 2.4). These enzymes catalyze glycosyl transfer from an activated glycosyl donor substrate, most often a nucleoside diphosphate (NDP) sugar, to a specific group on an acceptor molecule [23]. GTs are generally recognized as highly selective glycosylation catalysts, and they are believed to offer huge potential for synthetic use in the applied glycosciences [6]. GTs are distinguished according to whether they retain or invert the α-anomeric configuration of the NDP sugar substrate in the resulting glycosidic product (Scheme 2) [23]. Overall, GT reactions may therefore be viewed as nucleophilic substitutions at the glycosyl anomeric carbon involving axial-to-axial (retaining) or axial-to-equatorial (inverting) stereochemical course.
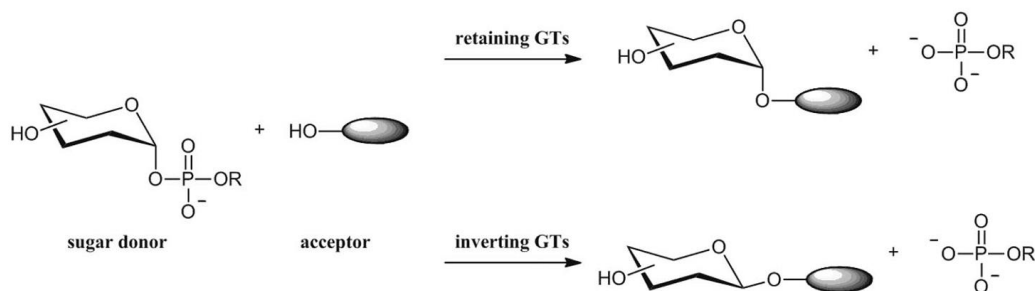
Scheme 2 GTs catalyze the transfer of a glycosyl moiety from an activated donor substrate to an acceptor molecule either under retention or inversion at the anomeric configuration.

Because the few currently characterized *C*-glycosyltransferases (C-GTs) are all inverting enzymes, we shall restrict the discussion later on to this type of enzymatic glycosyl transfer. GTs are categorized according to sequence similarity into currently 94 families [24]. The C-GTs so far characterized biochemically are found in family GT-1, a very large and diverse family of GTs from all three domains of life. The GT families are further aggregated into fold families and clans. The vast majority of GT three-dimensional structures fall into two principal types of GT folds, termed GT-A and GT-B [23]. There is one C-GT, the glycosyltransferase UrdGT2 from *Streptomyces fradiae*, that had its crystal structure determined [25]. The enzyme belongs to fold family GT-B and clan tmpII. Fold GT-B is characterized by the occurrence of two highly similar Rossmann-fold subdomains [26]. The enzyme active site is located in the interdomain cleft. Figure 1 shows the overall structure of UrdGT2 and compares it to the structure of a plant flavonoid *O*-glucosyltransferase [27] that had been a point of departure for this research. A characteristic feature of GTs from fold family GT-B, distinguishing them from enzymes of fold family GT-A, is that their activities are not dependent on divalent metal ions (Mn^{2+}, Mg^{2+}) [23].
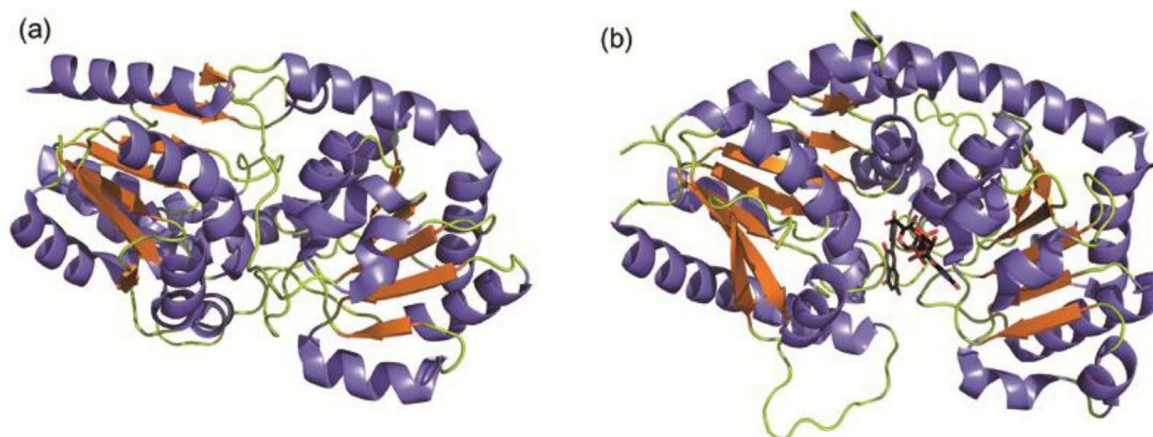

Fig. 1 The bacterial C-GT UrdGT2 (a) and the plant O-GT (from *Vitis vinifera*; grape vine) (b) feature both GT-B fold with two β/α/β Rossmann-like domains. *Vv*_O-GT displays substrates bound in the active site at the interdomain cleft.

Biochemical studies of UrdGT2 revealed a promiscuous GT promoting *C*- as well as *O*-glycosyl transfer from TDP-ᴅ-olivose to hydroxylated anthraquinone acceptors [28]. IroB from *Escherichia coli* catalyzed multiple aryl-*C*-glucosylations of the tricatecholic siderophore enterobactin, using UDP-ᴅ-glucose as donor substrate [29]. The *C*-glucosyl transfer took place to C5 of each of the 2,3-dihydroxy-

Pure Appl. Chem., Vol. 85, No. 9, pp. 1865–1877, 2013

17
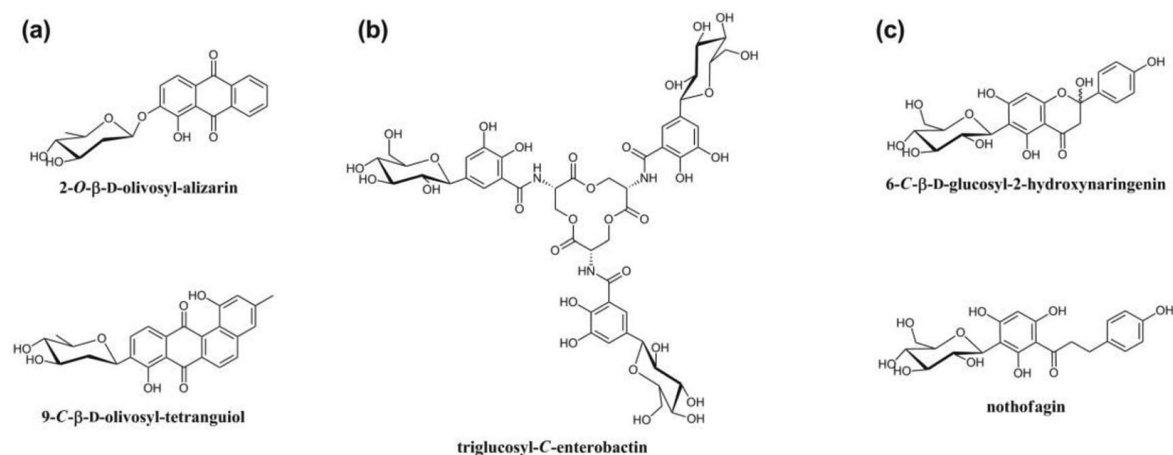
Scheme 3 Conversion products of the *C*-GTs UrdGT2 (a), IroB (b), and *Os*_C-GT (c). Only UrdGT2 is known to form both *O*- and *C*-glycosides.

benzoyl units of the enterobactin acceptor. Another C-GT was identified from rice (*Oryza sativa*; *Os*_C-GT), which preferentially catalyzed glucosyl transfer from UDP-D-glucose to different 2,5,7-tri-hydroxy-substituted flavanone and dihydrochalcone acceptors [30]. Scheme 3 shows typical reaction products of UrdGT2, IroB, and *Os*_C-GT. Generally, enzymatic *C*-glycosylation seems to occur similar to these examples preferentially at positions *ortho* or *para* to phenolic hydroxyls.

The current selection of known C-GTs is by far too small to supply organic synthesis with a sufficient variety of useful *C*-glycosylation catalysts. Identification of generically novel C-GTs from Nature remains a possibility to broaden the scope and enhance the application of enzymatic *C*-glycosylations. However, the relatively rare occurrence of the *C*-glycosidic linkage in natural products seems to imply that natural C-GTs can also not be large in number [1]. An alternative, perhaps more promising possibility would be, therefore, to capitalize on the large diversity of known *O*-glycosyltransferases (O-GTs). Protein engineering might be applied to the existing natural O-GTs for development of new *C*-glycosylation catalysts. Unfortunately, it is not known in terms of protein structure and function, what distinguishes a C-GT from an O-GT. Unlike the catalytic mechanism of (inverting) O-GTs that has been explored in detail [23], the catalytic principles of enzymatic *C*-glycosyl transfer are not well understood. At this point, it is therefore very difficult to apply targeted enzyme design. Evolutionary approaches might still be used, but they would also benefit greatly from an improved knowledge basis that restricts the sequence space to be explored and therefore supports semi-rational strategies.

MECHANISTIC PROPOSALS FOR ENZYMATIC *C*-GLYCOSYL TRANSFER

Scheme 4a shows the current mechanistic proposal for inverting O-GTs. Their reaction is thought to proceed via a direct nucleophilic substitution at the glycosyl anomeric carbon [23]. Departure of the NDP leaving group is assisted by positively charged residues (GT-B fold enzymes) or a bound metal ion (GT-A fold enzymes) in the active site that interact with the pyrophosphate moiety of the NDP. A general catalytic base abstracts a proton from the nucleophile, facilitating its attack on the anomeric carbon. The transition state is widely believed to have oxocarbenium ion-like character. The O-GT mechanism provides a plausible starting mechanistic hypothesis for C-GT that is also shown in Scheme 4b. Deprotonation of the acceptor substrate would induce, through resonance, carbanion character at the reactive carbon, thereby generating a suitable carbon nucleophile for the enzymatic reaction that would therefore take place via a direct *C*-glucosylation process, very much alike *O*-glycosylation. It was shown in Scheme 3 that UrdGT2 and *Os*_C-GT catalyze *C*-glycosylation of aromatic acceptors in posi-
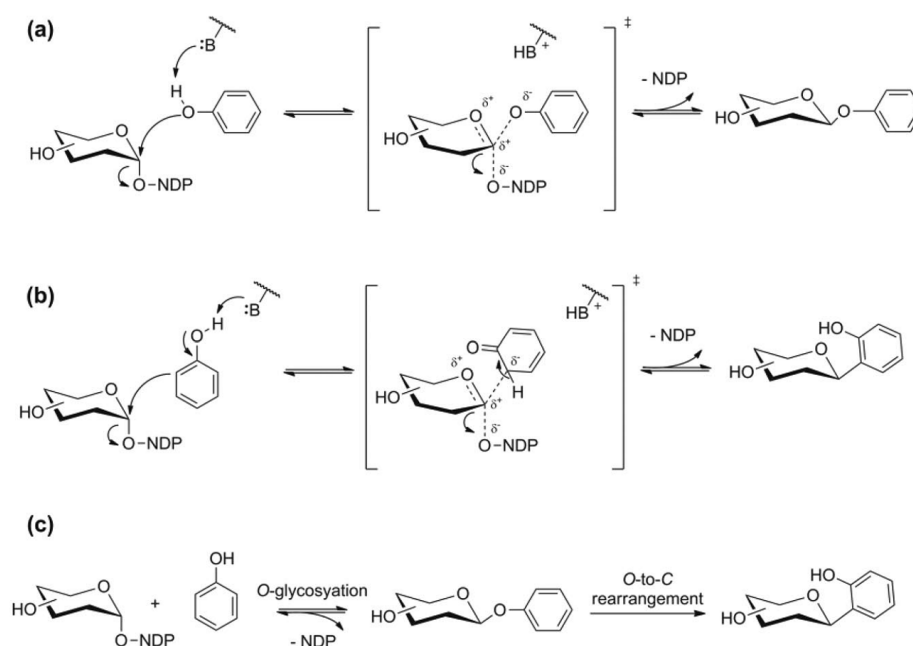
Scheme 4 Mechanistic proposals for enzymatic *O*-glycosylation (a), direct *C*-glycosylation (b), and *C*-glycosylation involving rearrangement of an intermediary *O*-glycoside (c).

tion *ortho* to a phenolic hydroxyl. In previous work on UrdGT2, this regioselectivity of enzymatic *C*-glycosylation has raised the question of an alternative C-GT mechanism where an initial *O*-glycosylation is followed by an *O*-to-*C* rearrangement [28], as depicted in Scheme 4c. While chemically conceivable [31], the proposed rearrangement mechanism is not entirely persuasive, as it requires that the enzyme active site promotes a quite demanding intramolecular multistep glycosyl transfer. Besides the need to stabilize different transition states, the rearrangement mechanism would also involve the major additional difficulty that the relative positions of the glycosyl residue and the acceptor substrate have to change during the reaction. The demand for precise positioning and stereochemical control in the enzymatic *O*-to-*C* rearrangement would therefore be especially high for IroB, where contrary to UrdGT2 and *Os*_C-GT that are selective for reaction at the *ortho* position, the *C*-glycosylation position is *para* to the putative site of *O*-glycosylation in the first catalytic step (Scheme 3) [29]. For intuitive reasons, therefore, mechanistic thinking would clearly favor the comparatively parsimonious reaction coordinate of a direct *C*-glycosylation process catalyzed by C-GT. However, the currently available evidence on enzymatic *C*-glycosyl transfer does not allow for a clear distinction between the two mechanistic possibilities considered. We would like to describe herein new insights into the mechanism of C-GT that were obtained through our recent studies of *Os*_C-GT [32].

SWITCHING BETWEEN O- AND C-GT THROUGH EXCHANGE OF ACTIVE-SITE MOTIFS

Studies of two complementary O- and C-GTs from bacteria

Despite their fundamental disparity, both mechanistic proposals for C-GT agree in that they postulate clear analogies between reactions catalyzed by C-GT and O-GT. This has inspired approaches to change the glycosidic bond-type specificity in O-GT or C-GT using protein engineering. Functional homologues of O-GT and C-GT are therefore of interest. Bechthold and co-workers have examined LanGT2, an O-GT from *Streptomyces cyanogenus* that catalyzes D-olivosyl transfer to tetranguiol, giving 8-*O*-D-olivosyl-11-deoxylandomycinone as product [33]. The C-GT UrdGT2 is active towards the same accep-

© 2013, IUPAC

Pure Appl. Chem., Vol. 85, No. 9, pp. 1865–1877, 2013

19

tor, but it transfers the D-olivosyl residue to C9, yielding 9-*C*-D-olivosyl-tetranguiol (Scheme 3) [33]. Based on the results of structure-based sequence analysis, prediction was made of which substitutions in LanGT2 might be needed to change the original O-GT into C-GT activity [34]. A large variety of LanGT2 chimeras were generated in which distinct sequence elements of the native UrdGT2 were introduced. Using multiple amino acid substitutions in sequence regions of LanGT2, which according to protein modeling belong to flexible loops around the binding pocket for the acceptor substrate, they could generate LanGT2 variants that behaved as low-activity C-GT having UrdGT2 specificity. Moreover, Härle et al. also revealed that the switch from O-GT to C-GT was mainly caused by replacing 10 amino acids in LanGT2 (Val^{52} – Ile^{62}) by the corresponding decapeptide from UrdGT2. However, a reverse swapping experiment where the LanGT2 sequence was introduced to UrdGT2 did not prove fruitful, possibly because the resulting UrdGT2 chimera was not active. Overall, therefore, detailed molecular interpretation of the specificity change in LanGT2 was difficult. Because of the many substitutions required to confer C-GT activity to the native LanGT2, a clear design principle for conversion of O-GT into C-GT remained elusive. Computational docking studies of LanGT2 wild type and mutants were interpreted to support the direct *C*-glycosylation mechanism.

A homologous pair of plant O- and C-GTs

We selected *Os*_C-GT and O-GT from pear (*Pyrus communis*; *Pc*_O-GT) [35] as complementary enzymes. The two GTs are related to each other by an overall amino acid identity of 30 %. Both *Os*_C-GT and *Pc*_O-GT have been classified into GT family GT-1. Each enzyme catalyzes glucosyl transfer from UDP-glucose to phloretin. *Os*_C-GT gives nothofagin as product [30] whereas reaction of *Pc*_O-GT yields phlorizin (Scheme 1) [35]. We reasoned that characteristics of structure and function of *Os*_C-GT distinguishing it from *Pc*_O-GT could thus be analyzed for two very similar and easily tractable chemical transformations. Enzymatic reactions are monitored conveniently by high-performance liquid chromatography (HPLC), as shown in Fig. 2.
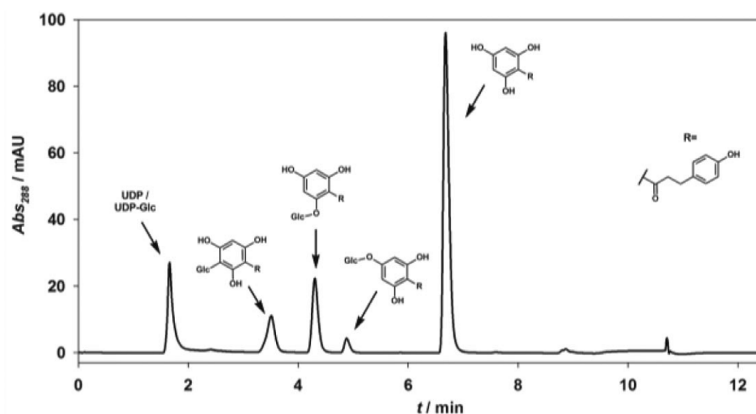

Fig. 2 Using a water to acetonitrile gradient on a reversed phase (C-18) column phloretin and its glycosides could be separated and quantified by HPLC.

No experimental structures are currently available for *Os*_C-GT and *Pc*_O-GT. We therefore obtained structural models of the two enzymes using as templates their closest neighbors of plant GTs in family GT-1. Figure 3a shows the active-site of the crystal structure of an O-GT from grape vine (*Vitis vinifera*; *Vv*_O-GT). It was solved for a ternary complex between the enzyme, UDP-2-deoxy-2-fluoro-D-glucose, an analogue of the native sugar donor that is rendered inactive due to the substitution of the 2-OH by the strongly electron-withdrawing fluorine, and the natural flavonol kaempferol as the
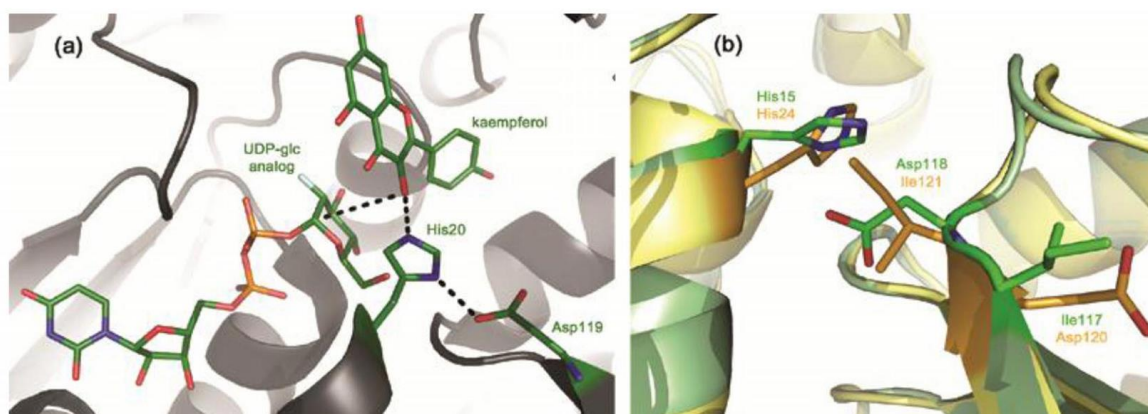
Fig. 3 Crystal structure of ternary *Vv*_O-GT complex (pdb code 2C1Z) (a) and overlay of modeled *Os*_C-GT (yellow) and *Pc*_O-GT (green) structures (b); *Os*_C-GT displays the conserved GT-B fold but disruption in the catalytic His-Asp motif present in O-GTs.

acceptor [27]. The active site of *Vv*_O-GT shows a prominent His residue (His^{20}) that is highly conserved among plant O-GTs of family GT-1 and is also present in *Os*_C-GT (Fig. 3b). The His is the putative catalytic general base that abstracts a proton from the reactive hydroxyl of the acceptor substrate (3-OH of kaempferol) during the reaction. In *Vv*_O-GT, Asp^{119} forms a hydrogen bond with His^{20}. Interaction with Asp^{119} is expected to assist in the proposed catalytic base function of His^{20} in at least two different ways. First, it contributes to the correct orientation of the His side chain in the active site. Second, it will enhance the basicity of His^{20} through pK_a enhancement. Similar dyads of His and Asp have been found in serine proteases amongst other enzymes, and the role of Asp in establishing a proton relay, allowing the His to fulfill its function as general catalytic base, have been well described [36]. Also, the modeled active site of *Pc*_O-GT (Fig. 3b) displays His^{15} and Asp^{118} side chains in close proximity and these two amino acids are generally conserved in the sequences of plant O-GTs of family GT-1, suggesting a catalytic consensus motif of *O*-glycosyl transfer in these enzymes [2].

Figure 3b reveals that the Asp-to-His arrangement thought to be characteristic of plant O-GT function appears to have been disrupted in *Os*_C-GT, probably as result of an Ile-Asp to Asp-Ile residue exchange in the sequence of *Os*_C-GT as compared to *Pc*_O-GT. Conformational and chemical properties of the active-site His will be strongly affected by loss of interaction with the Asp. Combinatorial mutagenesis of the Asp^{120}-Ile^{121} and Ile^{117}-Asp^{118} dyads of *Os*_C-GT and *Pc*_O-GT, respectively, was used to investigate the suggestion that these active-site motifs are crucial for glycosidic bond-type specificity in plant C-GT and O-GT. It is interesting that despite common membership to GT family GT-1, the proposed base-catalytic machinery of the bacterial enzymes (UrdGT2, LanGT2) differs from the one of the plant enzymes in that an active-site His does not seem to have a key role.

Swapping active-site motifs using targeted mutagenesis

The combinatorial approach of site-directed mutagenesis resulted in two single mutants and one double mutant for each enzyme. Thereby, the effect of all possible combinations of Asp and Ile in the putative active-site motif could be tested for both enzymes [32]. The double mutants of *Os*_C-GT and *Pc*_O-GT presented a complete swap of each enzyme's native active-site motif by the active-site motif of the respective other enzyme. The single mutants reflected only a partial motif exchange. Purified preparations of *Os*_C-GT and *Pc*_O-GT in wild-type and mutated form were obtained through recombinant production in *E. coli*. Each enzyme was produced as a chimeric protein that harbored *Strep*-tag II at its N-terminus. Purification was done by affinity chromatography, yielding electrophoretically

A. GUTMANN AND B. NIDETZKY

homogeneous protein in a single step. Protein expression generally occurred at a rather low level in *E. coli*, but except for mutants of *Pc*_O-GT that were produced in very tiny amounts, the procedure was sufficient and also convenient to obtain highly purified protein for characterization. The amount of *Pc*_O-GT mutants obtained after purification was just enough to assay their glycosidic bond-type specificity in the synthetic reaction. Time courses of enzymatic conversions of UDP-glucose and phloretin were recorded using HPLC, and the types of glycosidic products formed were identified from the analysis (Fig. 2). The identity of the product peaks revealed in HPLC had been established using NMR [32]. It needs to be emphasized that mutants of *Os*_C-GT and *Pc*_O-GT were very slow enzymes. Site-directed replacements within the active-site motif went along with a large, between 10^3 to 10^4-fold loss of specific activity in both *Os*_C-GT and *Pc*_O-GT as compared to native enzyme. However, mutants of *Os*_C-GT and *Pc*_O-GT featuring partial or complete exchange of the active-site motif showed a marked change in glycosidic bond-type specificity. The observed specificity change was in very good agreement with the suggestion from Fig. 3 that using a motif swap, a reversible switch between O-GT and C-GT activity might become possible.

Figure 4a shows that in *Pc*_O-GT, there was a gradual alteration in specificity from O-GT to C-GT upon going from the replacement of Ile^{117} by Asp to the replacement of Asp^{118} by Ile. The double mutant of *Pc*_O-GT that harbored a *Os*_C-GT-like active-site motif of Asp^{117} and Ile^{118} behaved as a high-fidelity C-GT, producing nothofagin as sole product of the enzymatic glucosyl transfer to phloretin. No phlorizin was ever detectable in the reaction of the double mutant. The relatively larger effect on change of glycosidic bond type specificity in the Asp$^{118} \rightarrow$ Ile (D118I) mutant as compared to the Ile$^{117} \rightarrow$ Asp (I117D) mutant is consistent with the idea developed from Fig. 3 that a functional active-site dyad of Asp and His is essential for O-GT activity.
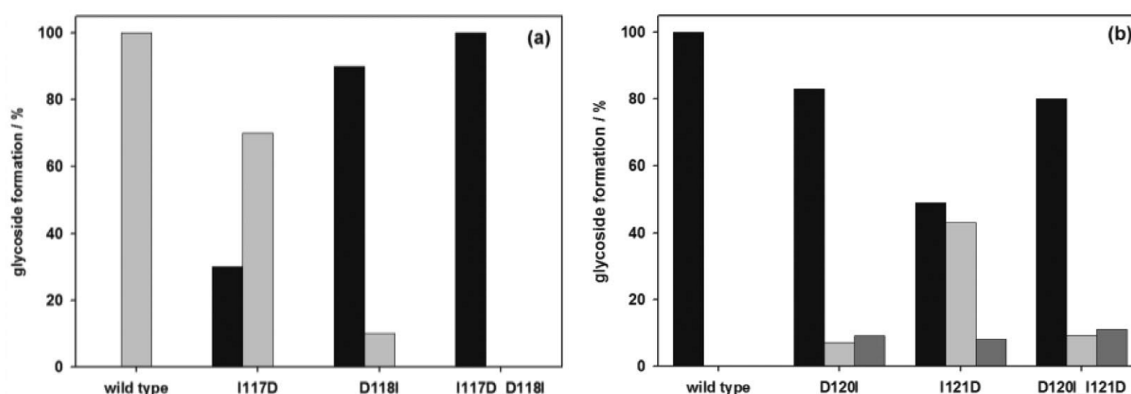

Fig. 4 Distribution of phloretin glucosides produced by wild-type and mutant *Pc*_O-GT (a) and *Os*_C-GT (b); black: 3'-*C*-glucoside (nothofagin); light gray: 2'-*O*-glucoside (phlorizin); dark gray: 4'-*O*-glucoside.

The reverse engineering approach of converting *Os*_C-GT into an O-GT was also successful, however, the specificity switch in *Os*_C-GT mutants was not as "clean" as in the corresponding *Pc*_O-GT mutants. Substitution of Ile^{121} by Asp (I121D) is probably most relevant because at the sequence level it implements an O-GT-type His-to-Asp active-site group. We characterized the I121D mutant of *Os*_C-GT and found that it was a promiscuous enzyme, being indiscriminate in forming *C*- and *O*-glycosidic linkages upon glucosyl transfer from UDP-glucose to phloretin, as shown in Fig. 4b. The O-GT activity of the I121D mutant was mainly targeted to the 2'-OH of phloretin, but there was also a significant amount of transfer to the 4'-OH. Replacement of Asp^{120} of *Os*_C-GT by Ile resulted in the same Ile-Ile repeat that was introduced in *Pc*_O-GT through substitution of Asp^{118} by Ile. Interestingly, the glycosidic bond-type specificity in the two mutants, D120I of *Os*_C-GT and D118I of *Pc*_O-GT, was quite similar. Both mutants showed clear preference for the *C*-glycosyl trans-

fer, producing nothofagin in about 5-fold or even larger excess over phlorizin in the enzymatic reaction. The result is important because it shows that complementary mutagenesis experiments in *Os*_C-GT and *Pc*_O-GT according to the proposed principle of C/O-GT structure and function (Fig. 3) gave consistent effects. Moreover, the observed specificities were in excellent agreement with expectations for enzyme variants that do not contain Asp in a position suitable for catalytic function as O-GT. An interesting difference, however, between the D120I mutant of *Os*_C-GT and the D118I mutant of *Pc*_O-GT was the regioselectivity of the *O*-glucosyl transfer. While the D118I mutant was absolutely selective for reaction with the phloretin 2'-OH, the D120I mutant gave unselective glucosylation of 2'-OH and 4'-OH of the acceptor. The double mutant of *Os*_C-GT, featuring complete exchange of C-GT by O-GT active-site motif at the level of primary structure and therefore expected to function mainly as O-GT, gave a pattern of glycosidic products that was similar to the one obtained with the single mutant D120I. The *C*-glucoside nothofagin was formed in 4-fold larger amount than the total of *O*-glucosides including phlorizin. The case of the *Os*_C-GT double mutant clearly emphasizes limitations in the ability of the simple model to precisely predict the glycosidic bond-type specificity of *Os*_C-GT and *Pc*_O-GT variants. However, we think, and the very low specific activity of *Os*_C-GT double mutant supports this notion, that the combination of substitutions of $Asp^{120} \rightarrow Ile$ and $Ile^{121} \rightarrow Asp$ was not well tolerated by *Os*_C-GT and therefore resulted in proximally disruptive effects on the active site as a whole. Such secondary effects are not uncommon in enzyme mutants [37]. However, these peculiarities of the *Os*_C-GT double mutant notwithstanding, evidence from the mutagenesis study is clear in having identified distinct active-site motifs in *Os*_C-GT and *Pc*_O-GT that define catalytic function in *O*- and *C*-glucosyl transfer. The evidence presented immediately suggests a design strategy for engineering glycosidic bond-type specificity in these two and related plant GTs. Conversion of native O-GT into engineered C-GT activity would be of special interest for development of new *C*-glycosylation enzyme catalysts.

Mechanistic implications for *C*-GT

Using the I121D mutant of *Os*_C-GT, which can be described as a dual-specific O-GT and C-GT, we considered experiments that would allow distinction between the two mechanistic possibilities for enzymatic *C*-glycosyl transfer. We figured that involvement of *O*-to-*C* rearrangement in the enzymatic mechanism of wild-type *Os*_C-GT should imply that the I121D mutant catalyzes the conversion of phlorizin to nothofagin in the presence of UDP. Formation of phlorizin in the reaction of the I121D mutant indicates that *O*-glucoside is released from the enzyme-UDP-phlorizin product complex in a post-glucosyl-transfer step. However, the ability to release phlorizin also requires that the I121D mutant is able to bind phlorizin from solution. In the presence of UDP, therefore, it must be possible for this *Os*_C-GT mutant to generate the same ternary complex of enzyme, phlorizin, and UDP without the necessity of a foregoing enzymatic glucosyl transfer. Because the I121D mutant catalyzed *C*-glucosyl transfer just as it catalyzed *O*-glucosyl transfer, postulate of a rearrangement mechanism therefore requires that phlorizin be a substrate for UDP-dependent direct conversion into nothofagin. We tested the implication by incubating phlorizin (60 μM) in the presence of I121D mutant (6.7 μM) and UDP (1 mM) for 48 h [32]. There was no reaction above the detection limit (~1 %), and phlorizin was stable over the whole time-span of the experiment. This result gives strong evidence against a role of catalytic *O*- to-*C*-rearrangement in the C-GT mechanism.

We therefore propose that catalytic reaction of *Os*_*C*-GT proceeds through a direct *C*-glycosylation mechanism. *Pc*_O-GT and *Os*_C_GT seem to discriminate between 2'-*O*- and 3'-*C*-glucosylation of phloretin primarily through relative positioning of the acceptor substrate towards the sugar donor. Partial deprotonation of phloretin 2'-OH by the conserved active-site His presents a common catalytic feature of both enzymatic reactions (Schemes 4a and b). This results in establishment of nucleophilic character on O2 and through resonance, on the aromatic C3 in *ortho* position to it. Just as *O*-glycosylation, *C*-glycosylation would thus be achieved through nucleophilic displacement at the anomeric car-

bon in a single reaction step. Such a reaction could also be described as a Friedel–Crafts-like direct alkylation of the phenolic acceptor in an electrophilic aromatic substitution. The nucleophilic displacement probably involves formation of an oxocarbenium ion-like transition state whereby the positively charged anomeric carbon is stabilized by both, the UDP leaving group and the attacking carbanion, like it is shown in Scheme 4b.

To accommodate essentially the same catalytic function from the active-site His on an acceptor molecule positioned differently for reaction at O2 or C3, it seems that *Pc*_O-GT and *Os*_C_GT have placed their respective His in a different microenvironment (Fig. 3b). In addition to steric effects that will contribute to the relative positioning of reactive and catalytic groups, it is conceivable that other *chemical* factors of selectivity in competing *O* and *C* alkylations of phenoxide ions might thus have become optimized in each enzyme. Solvent dielectric and polarity, S_N1 or S_N2 character of the transition state, and hydrogen bonding have been discussed in literature as factors affecting the position of alkylations of phenoxide ions [38]. Moreover, the hard–soft character of the phenoxide is considered important in determining the selectivity of oxygen vs. carbon alkylation [39]. Reactivity of the acceptor substrate will be partly controlled by the catalytic action of the active-site His, which in turn will be modulated by neighboring residues. An important outcome of this study that small changes in active-site structure can result in marked alteration of glycosidic bond-type specificity is supported in literature where site-directed substitutions were used to install *N*-glycosylation activity in a plant O-GT and remove it from a dual-specific *O*/*N*-GT [40]. Interestingly, structural change within the His-Asp dyad of plant O-GT (see Fig. 3) was thought to be responsible for the unusual *N*-glycoside formation.

OTHER ENZYMATIC *C*-GLYCOSYLATIONS

Three other types of *C*-glycosylation are worth mentioning here. The first is protein α-*C*-mannosylation which occurs at the C2 of the indol ring in the side chain of Trp [41]. A sequon of Trp-*X*-*X*-Trp is often used for recognition, whereby the identity of *X* is variable, even though not random [42]. Dolichyl-phosphate mannose serves as donor substrate, and the reaction seems to be enzyme-catalyzed in the endoplasmatic reticulum [43]. However, a specific GT responsible for *C*-mannosylation has not yet been identified and characterized [44]. Recently, the first *C*-mannosyltransferase (DPY-19 from *Caenorhabditis elegans*) was identified and characterized [44]. The enzyme featured topological and sequence homology to *N*-glycan oligosaccharyltransferases.

Pseudouridine synthase catalyzes isomerization of uridine to pseudouridine in RNA [45]. The enzymatic reaction involves conversion of an *N*- into a *C*-glycosidic linkage, and it proceeds with retention of anomeric configuration [46]. A plausible mechanistic proposal for pseudouridine synthase considers reaction via a configurationally inverted *O*-glycosyl enzyme intermediate (Scheme 5a) [47]. Formation of the acylal adduct implicates a nucleophilic Asp residue in the enzyme's active site which is highly conserved among the known families of pseudouridine synthase [48]. A quite remarkable feature of the proposed mechanism is that it involves three types of glycosidic bonds [47]. Finally, pseudouridine-5'-phosphate glycosidases may catalyze *C*-ribosylation through a ring-opening mechanism including imine formation with a lysine (Scheme 5b) [49].
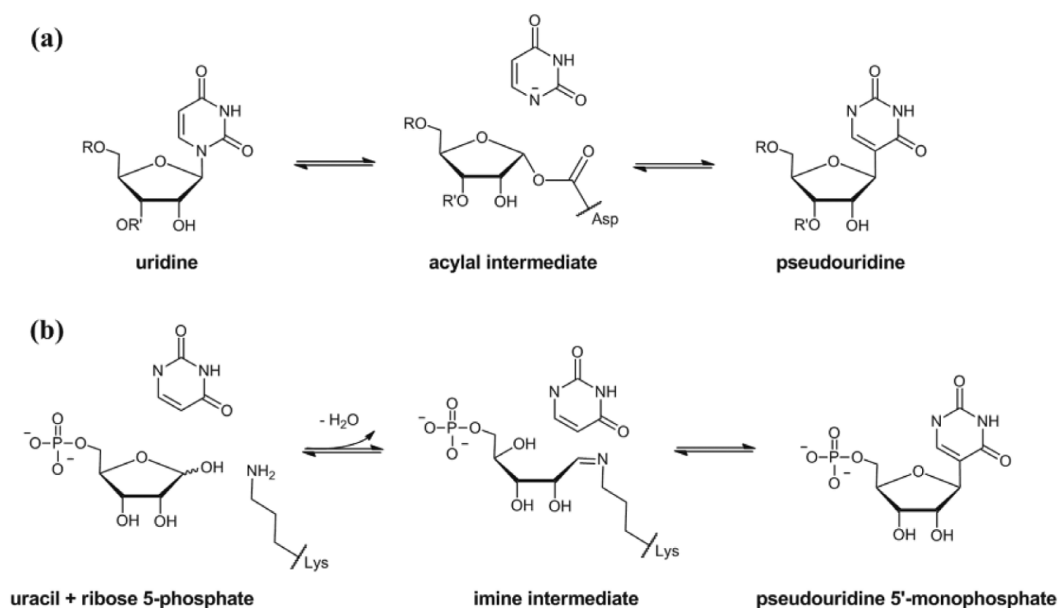
Scheme 5 Isomerization of uridine to pseudouridine by pseudouridine synthase (a) involves most likely an acylal adduct. Pseudouridine glycosidase forms pseudouridine monophosphate through ring opening (b).

CONCLUSIONS

Combinatorial mutagenesis of *Os*_C-GT and *Pc*_O-GT reveals active-site motifs in Leloir-type plant GTs of family GT-1 that are critical for the enzymes to differentiate between *C*- and *O*-glucosylation. This insight might stimulate the development of new *C*-glycosylation catalysts through active-site redesign of the abundant *O*-GTs. Furthermore, analysis of the promiscuous *O*- and *C*-glycosylating activity of the *Os*_C-GT mutant I121D gives strong evidence for an enzymatic *C*-glycosylation mechanism involving single nucleophilic (direct) displacement at the glucosyl anomeric carbon by an acceptor carbanion.

ACKNOWLEDGMENT

REFERENCES

1. C. J. Thibodeaux, C. E. Melancon 3rd, H. W. Liu. *Angew. Chem., Int. Ed.* **47**, 9814 (2008).
2. X. Q. Wang. *FEBS Lett.* **583**, 3303 (2009).
3. J. A. Salas, C. Mendez. *Trends Microbiol.* **15**, 219 (2007).
4. C. Mendez, J. A. Salas. *Trends Biotechnol.* **19**, 449 (2001).
5. S. Singh, G. N. Phillips, J. S. Thorson. *Nat. Prod. Rep.* **29**, 1201 (2012).
6. G. J. Williams, R. W. Gantt, J. S. Thorson. *Curr. Opin. Chem. Biol.* **12**, 556 (2008).
7. C. Luley-Goedl, B. Nidetzky. *Nat. Prod. Rep.* **28**, 875 (2011).
8. R. W. Gantt, R. D. Goff, G. J. Williams, J. S. Thorson. *Angew. Chem., Int. Ed.* **47**, 8889 (2008).
9. B. R. Griffith, J. M. Langenhan, J. S. Thorson. *Curr. Opin. Biotechnol.* **16**, 622 (2005).
10. N. C. Veitch, R. J. Grayer. *Nat. Prod. Rep.* **28**, 1626 (2011).
11. T. Bililign, B. R. Griffith, J. S. Thorson. *Nat. Prod. Rep.* **22**, 742 (2005).

12. D. E. Levy, C. Tang. *The Chemistry of C-glycosides*, pp. 10–22, Pergamon Press, Oxford (1995), and refs. therein.
13. T. Bililign, C. G. Hyun, J. S. Williams, A. M. Czisny, J. S. Thorson. *Chem. Biol.* **11**, 959 (2004).
14. J. Lee, S. H. Lee, H. J. Seo, E. J. Son, S. H. Lee, M. E. Jung, M. Lee, H. K. Han, J. Kim, J. Kang, J. Lee. *Bioorg. Med. Chem.* **18**, 2178 (2010).
15. J. F. List, V. Woo, E. Morales, W. Tang, F. T. Fiedorek. *Diabetes Care* **32**, 650 (2009).
16. E. C. Chao. *Core Evid.* **7**, 21 (2012).
17. W. Meng, B. A. Ellsworth, A. A. Nirschl, P. J. McCann, M. Patel, R. N. Girotra, G. Wu, P. M. Sher, E. P. Morrison, S. A. Biller, R. Zahler, P. P. Deshpande, A. Pullockaran, D. L. Hagan, N. Morgan, J. R. Taylor, M. T. Obermeier, W. G. Humphreys, A. Khanna, L. Discenza, J. G. Robertson, A. Wang, S. Hang, J. R. Wetterau, E. B. Janovitz, O. P. Flint, J. M. Whaley, W. N. Washburn. *J. Med. Chem.* **51**, 1145 (2008).
18. L. Rossetti, D. Smith, G. I. Shulman, D. Papachristou, R. A. Defronzo. *J. Clin. Invest.* **79**, 1510 (1987).
19. C. Gosch, H. Halbwirth, K. Stich. *Phytochemistry* **71**, 838 (2010).
20. D. L. McKay, J. B. Blumberg. *Phytother. Res.* **21**, 1 (2007).
21. M. H. D. Postema. *C-Glycoside Synthesis*, CRC Press, Boca Raton (1995), and refs. therein.
22. S. Lemaire, I. N. Houpis, T. T. Xiao, J. J. Li, E. Digard, C. Gozlan, R. M. Liu, A. Gavryushin, C. Diene, Y. C. Wang, V. Farina, P. Knochel. *Org. Lett.* **14**, 1480 (2012).
23. L. L. Lairson, B. Henrissat, G. J. Davies, S. G. Withers. *Annu. Rev. Biochem.* **77**, 521 (2008).
24. P. M. Coutinho, E. Deleury, G. J. Davies, B. Henrissat. *J. Mol. Biol.* **328**, 307 (2003).
25. M. Mittler, A. Bechthold, G. E. Schulz. *J. Mol. Biol.* **372**, 67 (2007).
26. C. Breton, S. Fournel-Gigleux, M. M. Palcic. *Curr. Opin. Struct. Biol.* **22**, 540 (2012).
27. W. Offen, C. Martinez-Fleites, M. Yang, E. Kiat-Lim, B. G. Davis, C. A. Tarling, C. M. Ford, D. J. Bowles, G. J. Davies. *EMBO J.* **25**, 1396 (2006).
28. C. Durr, D. Hoffmeister, S. E. Wohlert, K. Ichinose, M. Weber, U. von Mulert, J. S. Thorson, A. Bechthold. *Angew. Chem., Int. Ed.* **43**, 2962 (2004).
29. M. A. Fischbach, H. N. Lin, D. R. Liu, C. T. Walsh. *Proc. Natl. Acad. Sci. USA* **102**, 571 (2005).
30. M. Brazier-Hicks, K. M. Evans, M. C. Gershater, H. Puschmann, P. G. Steel, R. Edwards. *J. Biol. Chem.* **284**, 17926 (2009).
31. R. G. dos Santos, A. R. Jesus, J. M. Caio, A. P. Rauter. *Curr. Org. Chem.* **15**, 128 (2011).
32. A. Gutmann, B. Nidetzky. *Angew. Chem., Int. Ed.* **51**, 12879 (2012).
33. A. Luzhetskyy, T. Taguchi, M. Fedoryshyn, C. Durr, S. E. Wohlert, V. Novikov, A. Bechthold. *Chembiochem* **6**, 1406 (2005).
34. J. Harle, S. Gunther, B. Lauinger, M. Weber, B. Kammerer, D. L. Zechel, A. Luzhetskyy, A. Bechthold. *Chem. Biol.* **18**, 520 (2011).
35. C. Gosch, H. Halbwirth, B. Schneider, D. Holscher, K. Stich. *Plant Sci.* **178**, 299 (2010).
36. L. Hedstrom. *Chem. Rev.* **102**, 4501 (2002).
37. N. Tokuriki, D. S. Tawfik. *Science* **324**, 203 (2009).
38. R. Breslow, K. Groves, M. U. Mayer. *J. Am. Chem. Soc.* **124**, 3622 (2002).
39. R. G. Pearson. *J. Am. Chem. Soc.* **85**, 3533 (1963).
40. M. Brazier-Hicks, W. A. Offen, M. C. Gershater, T. J. Revett, E. K. Lim, D. J. Bowles, G. J. Davies, R. Edwards. *Proc. Natl. Acad. Sci. USA* **104**, 20238 (2007).
41. A. Furmanek, J. Hofsteenge. *Acta Biochim. Pol.* **47**, 781 (2000).
42. K. Julenius. *Glycobiology* **17**, 868 (2007).
43. M. A. Doucey, D. Hess, R. Cacan, J. Hofsteenge. *Mol. Biol. Cell* **9**, 291 (1998).
44. F. F. Buettner, A. Ashikov, B. Tiemann, L. Lehle, H. Bakker. *Mol. Cell* **50**, 295 (2013).
45. X. R. Gu, Y. Q. Liu, D. V. Santi. *Proc. Natl. Acad. Sci. USA* **96**, 14270 (1999).
46. C. Hoang, A. R. Ferre-D'Amare. *Cell* **107**, 929 (2001).
47. E. J. Miracco, E. G. Mueller. *J. Am. Chem. Soc.* **133**, 11826 (2011).

48. L. X. Huang, M. Pookanjanatavip, X. G. Gu, D. V. Santi. *Biochemistry* **37**, 344 (1998).
49. S. Huang, N. Mahanta, T. P. Begley, S. E. Ealick. *Biochemistry* **51**, 9245 (2012).

Leloir Glycosyltransferases and Natural Product Glycosylation: Biocatalytic Synthesis of the *C*-Glucoside Nothofagin, a Major Antioxidant of Redbush Herbal Tea

COMMUNICATIONS

Leloir Glycosyltransferases and Natural Product Glycosylation: Biocatalytic Synthesis of the *C*-Glucoside Nothofagin, a Major Antioxidant of Redbush Herbal Tea

Linda Bungaruang,[+a] Alexander Gutmann,[+a] and Bernd Nidetzky[a,*]

[a] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, 8010 Graz, Austria
Fax: (+43)-316-873-8434; phone:(+43)-316-873-8400; e-mail: bernd.nidetzky@tugraz.at

[+] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/adsc.201300251.

Abstract: Nothofagin is a major antioxidant of redbush herbal tea and represents a class of bioactive flavonoid-like *C*-glycosidic natural products. We developed an efficient enzymatic synthesis of nothofagin based on a one-pot coupled glycosyltransferase-catalyzed transformation that involves perfectly selective 3′-*C*-β-D-glucosylation of naturally abundant phloretin and applies sucrose as expedient glucosyl donor. *C*-Glucosyltransferase from *Oryza sativa* (rice) was used for phloretin *C*-glucosylation from uridine 5′-diphosphate (UDP)-glucose, which was supplied continuously *in situ* through conversion of sucrose and UDP catalyzed by sucrose synthase from *Glycine max* (soybean). In an evaluation of thermodynamic, kinetic, and stability parameters of the coupled enzymatic reactions, poor water solubility of the phloretin acceptor substrate was revealed as a major bottleneck of conversion efficiency. Using periodic feed of phloretin controlled by reaction progress, nothofagin concentrations (45 mM; 20 g L^{-1}) were obtained that vastly exceed the phloretin solubility limit (5–10 mM). The intermediate UDP-glucose was produced from catalytic amounts of UDP (1.0 mM) and was thus recycled 45 times in the process. Benchmarked against comparable glycosyltransferase-catalyzed transformations (e.g., on quercetin), the synthesis of nothofagin has achieved intensification in glycosidic product formation by up to three orders of magnitude (μM→mM range). It thus makes a strong case for the application of Leloir glycosyltransferases in biocatalytic syntheses of glycosylated natural products as fine chemicals.

Keywords: carbohydrates; *C*-glycosides; glycosyltransferases; natural products; UDP-glucose recycling

Many bioactive natural products contain sugar molecule(s) as part of their structure.[1] Their physiological activity, selectivity and pharmacological properties are often derived from the sugar component(s).[2] Therefore, glycosylation is often central to a natural product's efficacy in the particular application considered. Aside from therapeutic uses,[3] glycosylated natural products have raised interest as functional food additives and cosmetic ingredients.[4] Glycosylation pattern engineering is regarded as a highly promising way of functional diversification of natural products.[2b,3] This might contribute to the creation of new bioactive substances and drug leads.

In nature, the selective modification of target compounds with sugars is catalyzed by glycosyltransferases (EC 2.4).[5] These enzymes use an activated donor substrate, typically a nucleoside diphosphate (NDP)-sugar, for transfer of the glycosyl residue onto a specific position of an acceptor molecule. Glycosyltransferases display splendid regioselectivity and stereochemical control in the transformations catalyzed,[6] and they are therefore widely recognized as highly valuable glycosylation catalysts.[7] However,
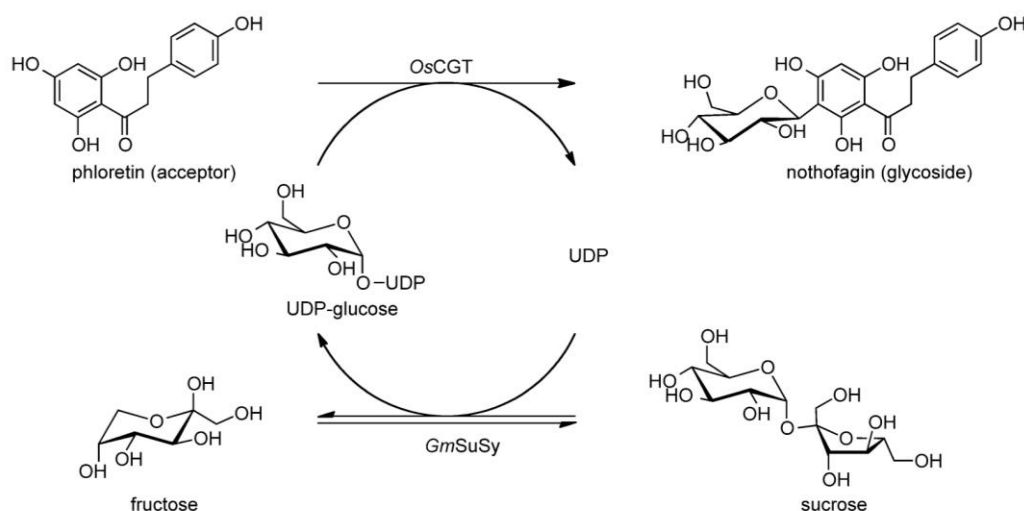
Scheme 1. Synthesis of nothofagin is achieved by enzymatic *C*-glucosylation of phloretin from sucrose *via* UDP-glucose. *Os*CGT and *Gm*SuSy are telescoped in one pot, and the reaction proceeds in the presence of catalytic amounts of UDP. Fructose is the only by-product formed. The overall equilibrium lies far on the *C*-glycoside product side, driven by the reaction of *Os*CGT and the presence of sucrose in excess.

synthetic applications of glycosyltransferases have so far been quite restricted due to complexities of the enzymes (e.g., low specific enzyme activity and stability)[8] and the supply of donor and acceptor substrates for the enzymatic reactions.[9]

The majority of natural product glycosylations involve *O*-glycosidic bonds. Glycosylations at carbon, by contrast, are relatively rare and to date only a small number of natural *C*-glycosyltransferases have been reported.[10] The *C*-glycosidic linkage displays outstanding resistance to chemical or enzyme-catalyzed hydrolysis, surpassing that of the corresponding *O*-glycosidic linkage by a large amount.[11] *C*-glycosides have therefore attracted considerable attention for functional substitution of physiologically active *O*-glycosidic compounds having low *in vivo* lifetimes.[12]

Aryl glucosides derived from flavonoid-like aglycones (Scheme 1) present a very interesting class of plant natural products that may involve either a *C*- or an *O*-glycosidic linkage.[13] They show a highly significant profile of biological activities that typically include strong antioxidant and radical scavenger functions, but also comprise antiviral and cytotoxic effects.[14] Because product isolation directly from the plant is often impractical, compounds must also be prepared by bottom-up synthesis. Despite notable recent advancements,[15] chemical methodologies involve multiple steps and are therefore generally neither atom-efficient nor high-yielding. We demonstrate in this study that single-step glycosyltransferase-catalyzed transformation *in vitro* presents a powerful tool for aryl *C*-glycoside synthesis. We show that when offered the dihydrochalcone phloretin as acceptor, *C*-glycosyltransferase from rice (*Oryza sativa*; *Os*CGT)

reacts with uridine 5′-diphosphate (UDP)-glucose to give the 3′-*C*-aryl β-D-glucoside nothofagin (Scheme 1) as a single transfer product. Nothofagin is a natural substance found in redbush herbal tea and represents a structural class of bioactive aryl *C*-glycosides.[16] For efficient nothofagin synthesis, we coupled the *C*-glucosyltransferase reaction to enzymatic *in situ* supply of the glucosyl donor substrate (Scheme 1): UDP-glucose is produced from sucrose and UDP using recombinant sucrose synthase from soybean (*Glycine max*; *Gm*SuSy). While the applied internal UDP-glucose regeneration is known in principle and has been applied to enzymatic reactions involving different glycosyl donor substrates,[17] a critical test of its performance capability in the synthesis of natural product glucosides such as nothofagin[17e–h] remains outstanding. We show the application of comprehensive step-by-step reaction engineering to overcome complexities inherent to this and similarly coupled glycosyltransferase systems and report an efficient, high-yielding biocatalytic production of nothofagin.

We first examined reactions of *Os*CGT and *Gm*SuSy separately and determined their kinetic and thermodynamic characteristics. Enzymes were obtained from *Escherichia coli* expression cultures and purified to apparent homogeneity by Strep-tag affinity chromatography, as described in the Supporting Information (Methods, Figure S1). Their activities were determined using enzymatic or HPLC-based assays (Supporting Information, Methods). The reaction of *Os*CGT was monitored with an HPLC assay capable of distinguishing between nothofagin and potential alternative products resulting from *O*-glucosyl transfer at the 2′ or 4′ position of the acceptor (Supporting In-
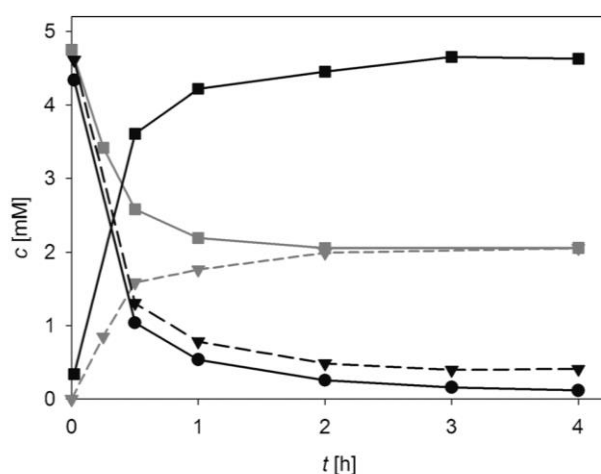
Figure 1. Time course analysis for individual enzymatic reactions catalyzed by *Os*CGT and *Gm*SuSy at pH 7.5 and 30°C. Nothofagin synthesis by *Os*CGT (black symbols): 80 mU mL^{-1}, 5 mM phloretin (triangle, dashed line), 4.75 mM UDP-glucose (circle, solid line), nothofagin (square, solid line). Reactions of *Gm*SuSy (grey symbols): 50 mU mL^{-1}, 5 mM of each substrate, UDP-glucose in sucrose synthesis (squares, solid line) and cleavage (triangles, dashed line).

Table 1. Characterization of glycosyltransferases and their reactions.[a]

Parameter	*Gm*SuSy		*Os*CGT
	Synthesis	Cleavage	
$K_{M\ sucrose}$ [mM]	–	25.5 ± 3.3[b]	–
$K_{M\ UDP}$ [mM]	–	0.13 ± 0.02[b]	–
$K_{M\ fructose}$ [mM]	3.0 ± 0.4[b]	–	
$K_{M\ UDP-glucose}$ [mM]	0.14 ± 0.03[b]	–	0.024 ± 0.004
$K_{M\ phloretin}$ [mM]	–	–	0.009 ± 0.003
k_{cat} [s^{-1}]	7.5 ± 0.4	9.3 ± 0.3	4.4 ± 0.3
Spec. act. [U mg^{-1}]	4.8 ± 0.2	5.9 ± 0.2	5.1 ± 0.3
K_{eq}		0.49 ± 0.01[c,e]	> 400[d,e]
$t_{1/2}$ [h]		18.8 ± 0.9[f]	13.8 ± 1.2[g]

[a] 30°C, 50 mM HEPES pH 7.5, 20% (v/v) DMSO.
[b] 30°C, 20 mM HEPES, pH 7.5.
[c] Cleavage direction (conversion of sucrose and UDP).
[d] Glycosylation direction (conversion of phloretin and UDP-glucose).
[e] Calculated from data in Figure 1.
[f] 30°C, 50 mM HEPES pH 7.5, 20% (v/v) DMSO, 100 mM sucrose.
[g] 30°C, 50 mM HEPES pH 7.5, 20% (v/v) DMSO, 5 mM phloretin.

formation, Figure S2).[18] *Os*CGT displayed absolute selectivity (within an error limit of $\leq 0.5\%$) for 3'-*C*-glycosylation of phloretin. We determined pH-activity dependencies for *Gm*SuSy (sucrose cleavage and synthesis) and *Os*CGT (nothofagin synthesis) at 30°C. The resulting pH profiles revealed suitable overlap of the enzyme activities in the pH range 6.5–8.0 (Supporting Information, Figure S3). The results shown in Figure 1 indicate that *C*-glucosylation of phloretin at pH 7.5 resulted in high conversion of substrates ($\geq 95\%$). Moreover, the enzymatic reverse reaction with nothofagin and UDP was not detectable under these conditions (Supporting Information, Figure S4). We concluded, therefore, that synthesis of nothofagin by *Os*CGT proceeds without critical thermodynamic limitations. The equilibrium for sucrose conversion is pH-dependent, and a low pH of 6 or smaller is known to favor the formation of UDP-glucose.[19] Figure 1 shows that at pH 7.5, the equilibrium constant (K_{eq}) for conversion of sucrose and UDP had a value of 0.49. However, thermodynamic constraints on the supply of UDP-glucose at elevated pH can be eliminated effectively using sucrose in excess. We therefore performed our conversion studies at pH 7.5 and 30°C where both glycosyltransferases showed useful activity and stability (Table 1) and quantitative transformation of sucrose into nothofagin was feasible.

The low water solubility of non-carbohydrate acceptor substrates is an important issue for carrying out natural product glycosylations *in vitro*. In the case of the barely water-soluble phloretin, use of an organic cosolvent was essential to enhance the acceptor substrate availability in *C*-glycosylations catalyzed by *Os*CGT. Whereas both ethanol and DMSO up to 20% by volume caused only minor interference with *Os*CGT activity, *Gm*SuSy displayed a low cosolvent tolerance and its activity was almost completely ($\geq 85\%$) lost in the presence of 15% ethanol. DMSO was less strongly affecting the activity of *Gm*SuSy and around 65% of the specific enzyme activity in purely aqueous buffer were retained in 20% DMSO. Furthermore, *Os*CGT stability and phloretin solubility were superior in DMSO as compared to ethanol. All conversion experiments were therefore performed in 20% DMSO, and the maximum concentration of dissolved phloretin was around 10 mM under these conditions.

Kinetic characterization of *Gm*SuSy (sucrose conversion) and *Os*CGT (nothofagin synthesis) was done at pH 7.5 and results are summarized in Table 1 along with the relevant enzyme stability and reaction thermodynamic parameters under these conditions. Both enzymes showed useful specific activities (≥ 5 units mg^{-1} protein) and were sufficiently stable under the reaction conditions with half-lives of around 19 (*Gm*SuSy) and 14 h (*Os*CGT). The Michaelis–Menten constant (K_M) of *Gm*SuSy for sucrose exceeds the corresponding K_M for UDP by two orders of magnitude. K_M values of *Os*CGT are also much lower than the sucrose K_M. Therefore, this implies that relatively high sucrose concentrations should be used in the coupled reaction to partly saturate and thus make optimum use of the *Gm*SuSy activity pres-

ent. The kinetic requirements of SuSy are therefore in good accordance with the notion of using an excess of sucrose to drive the overall conversion. We noticed that the K_M for UDP was 26-fold higher in our recombinant preparation of *Gm*SuSy as compared to the enzyme isolated from the native source.[20] This large difference in apparent UDP binding affinity might be due to effects of post-translational modification (e.g., covalent phosphorylation) that have been described for sucrose synthases in plants[21] and that may not occur in *E. coli*. The K_M for UDP-glucose was also strongly elevated (12-fold) in recombinant as compared to native *Gm*SuSy while, interestingly, the K_M values for sucrose and fructose were not affected in the recombinant enzyme. Differences between recombinant and native *Gm*SuSy were not further pursued in this study.

We performed synthesis experiments in which the initial concentration of sucrose (5–500 mM; 0.5 mM UDP) or UDP (0.005–1 mM; 100 mM sucrose) was varied while enzymatic activity (10 mU mL^{-1} *Os*CGT/ *Gm*SuSy) and phloretin concentration (5 mM) were constant. The nothofagin production rate (r_P) was measured, and results are depicted in Figure 2. Dependence of r_P on the sucrose concentration was hyperbolic, with a half-saturation constant (27 mM) comparable to the K_M of *Gm*SuSy for sucrose. Therefore, it appears to primarily reflect saturation behavior of *Gm*SuSy, as noted above (Table 1). The dependence of r_P on the UDP concentration was likewise hyperbolic, with a half-saturation level (51 µM) in between the K_M values for UDP and UDP-glucose. Figure 2B) indicates that the applied "nothofagin synthase" activity, derived from the combined activities of *Os*CGT and *Gm*SuSy, was therefore utilized best at a UDP concentration of 0.5 mM or higher. Using a phloretin concentration of 5 mM, this limited the maximum number of UDP-glucose regeneration cycles (RC_{max}) to 10 (=5/0.5). It would certainly be possible to further increase this RC_{max} value by de-
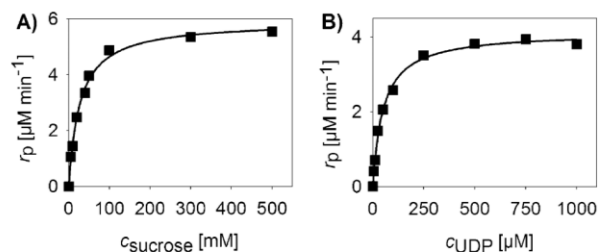
creasing the UDP concentration relative to the phloretin concentration, but this would probably have to occur at the expense of a significant loss in r_P. It is interesting that at the lowest UDP concentration used in Figure 2 (5 µM), the observed r_P was still 10% of its maximum value at saturation with UDP.

Aside from cost-efficient supply of the UDP-glucose donor substrate and favorable thermodynamic effects resulting from the use of high sucrose concentrations, a glucosyltransferase reaction might benefit from its coupling to the SuSy reaction kinetically. Pronounced end-product inhibition by micromolar concentrations of UDP is quite common among flavonoid *O*-glucosyltransferases[17e,g,h] and imposes severe restrictions on the direct synthetic use of these enzymes which could be decreased by continuous removal of the UDP released.[17e,22] We tested the influence of UDP inhibition on nothofagin production by comparing *Os*CGT (50 mU mL^{-1}) conversion of 5 mM phloretin (6 mM UDP-glucose, 100 mM sucrose) in absence and presence of *Gm*SuSy (50 mU mL^{-1}) (Figure 3A). Although in both reactions quantitative conversion (>99.5%) was achieved r_P showed a stronger decrease above ~75% conversion (2 h) without *Gm*SuSy and final conversions (>99%) were only reached after more than 10 h compared to less than 6 h in the presence of *Gm*SuSy. This corresponds to an approximately two-fold gain in space-time yield in nothofagin production (mM product formed/time consumed) resulting from the coupling of *Os*CGT and *Gm*SuSy reactions. UDP-glucose depletion could be excluded as an explanation for the reduction of r_P in the absence of *Gm*SuSy due to excess of UDP-glu-
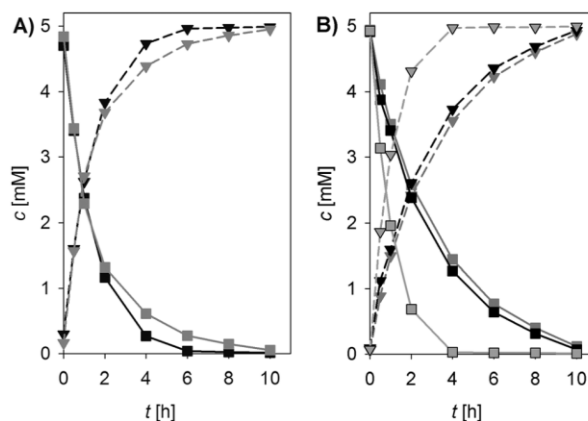

Figure 2. The nothofagin production rate (r_p) in a coupled-enzyme reaction (10 mU mL^{-1} *Os*CGT/*Gm*SuSy, 5 mM phloretin) depends on variation of **A)** the sucrose concentration (0.5 mM UDP) and **B)** the UDP (100 mM sucrose) concentration. Note: because sucrose was not fully saturating in **B)** the achieved r_p at high UDP is slightly lower than in **A)**.

Figure 3. Conversions of 5 mM phloretin (square) to nothofagin (triangle) by *Os*CGT (100 mM sucrose): **A)** Using 50 mU ml^{-1} *Os*CGT and 6 mM UDP-glucose in the absence (grey) and presence (black) of 50 mU mL^{-1} *Gm*SuSy; **B)** variation of *Os*CGT and *Gm*SuSy activity in coupled conversions (0.5 mM UDP): 50 mU mL^{-1} *Os*CGT/*Gm*SuSy (dark grey); 50 mU mL^{-1} *Os*CGT and 250 mU mL^{-1} *Gm*SuSy (black); 250 mU mL^{-1} *Os*CGT and 50 mU mL^{-1} *Gm*SuSy (light grey, black edge).
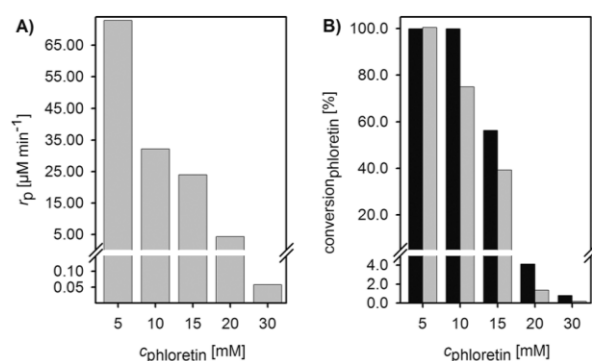
Figure 4. Batch conversions at different phloretin concentrations (100 mM sucrose, 0.5 mM UDP, 190 mU mL^{-1} *Os*CGT, 120 mU mL^{-1} *Gm*SuSy); **A)** Nothofagin production rate decreases at elevated phloretin concentrations; **B)** final conversion of soluble (black) and total applied (grey) phloretin after 24 h is limited at high phloretin concentrations by substrate inhibition and solubility.

cose (1 mM, 40-fold K_M) which was monitored throughout the conversion. Although end product inhibition at low millimolar UDP concentrations was less critical for *Os*CGT than for *O*-glycosyltransferases, *in situ* removal of UDP remains an essential feature for general application of SuSy as UDP-glucose recycling system for high level glycoside production.

Conversion rates were slightly lower when 5 mM UDP-glucose were replaced with 0.5 mM UDP in a coupled glycosyltransferase conversion (50 mU mL^{-1} *Os*CGT/*Gm*SuSy) (Figure 3B; Supporting Information, Figure S5) and complete conversion was only reached after 10 h. Also a 5-fold excess of *Gm*SuSy (250 mU mL^{-1}) over *Os*CGT (50 mU mL^{-1}) did not improve nothofagin production significantly. On the other hand a five-fold excess of *Os*CGT (250 mU mL^{-1}) over *Gm*SuSy (50 mU mL^{-1}) drastically increased the conversion resulting in complete conversion after only 4 h. Thereby *C*-glucosylation was identified as a rate-limiting step at the applied conditions. UDP-glucose levels of roughly 0.1 mM throughout all three conversions coincide with the finding that the UDP-glucose supply was not critical in the coupled conversions. Furthermore, it is worth remarking that the produced nothofagin was stable in all conversions, suggesting that the *C*-glucoside synthesis is conveniently performed under thermodynamic control and in the apparent absence of chemical or enzyme-catalyzed side reactions.

Considering that solubility of phloretin was markedly enhanced ($\geq$5-fold) upon its *C*-glycosylation, we raised the initial concentration of phloretin in various steps to 30 mM, thereby exceeding the solubility limit of the acceptor substrate by at least 3-fold. We figured that insoluble phloretin might still be useful for the continuous *in situ* supply of acceptor substrate when gradual transformation of the dissolved phloretin oc-

curred in the enzymatic reaction (100 mM sucrose, 0.5 mM UDP, 190 mU mL^{-1} *Os*CGT, 120 mU mL^{-1} *Gm*SuSy). Figure 4 shows that nothofagin production could not be upheld under conditions of insoluble acceptor being present ($\geq$10 mM) and also the final product concentration after 24 h was strongly decreased in clear dependence on the phloretin concentration. Using 10 mM phloretin, which was initially dissolved completely, we noticed precipitation of the acceptor substrate over time, limiting the maximum amount of nothofagin obtainable in the reaction under these conditions despite complete conversion of all soluble phloretin (Figure 4B). Furthermore, initial rate studies of *Os*CGT revealed substrate inhibition ($K_i \sim$5 mM) which clearly effected conversion at phloretin concentrations above 1 mM.

To nevertheless increase the end concentration of nothofagin in the enzymatic reaction, we changed the operation mode from batch to fed-batch, adding fresh phloretin to a concentration of 5 or 10 mM once the acceptor substrate had been depleted. Up to 8 rounds of phloretin addition were made, using a highly concentrated stock solution of 500 mM phloretin in pure DMSO to minimize the resulting volume change. Table 2 presents a summary of conditions and results and Figure 5 shows a reaction time course where in each round fresh enzyme was supplied together with 5 mM acceptor to the reaction. Feeding the phloretin acceptor was generally quite effective in enhancing nothofagin production. However, the phloretin conversion rate decreased strongly in dependence on the total amount of acceptor added to the reaction, so that without enzyme feed, the maximum concentration of nothofagin was just around 20 mM (Table 2).

Table 2. Nothofagin synthesis using controlled feed of phloretin and enzyme.[a]

$\Delta c_{phloretin}$[b] [mM]	5	5	10	10
Enzyme feed	no	yes	no	yes
$c_{phloretin}$[c] [mM]	20	45	40	60
Vol. act[c] [mU mL^{-1}]	100	550	100	600
t [h]	27	135	42	120
$c_{nothofagin}$[d] [mM]	14.6	44.1	19.7	46.6
Conversion[e] [%]	88	98	63	90
Precipitation [mM][f]	3.5	<0.1	8.9	8.1
ttn[g] (*Gm*SuSy/*Os*CGT) [1·10^3]	18/16	10/9	25/21	10/8

[a] 300 mM sucrose, 1 mM UDP, 30 °C, 50 mM HEPES pH 7.5, 20% DMSO.
[b] Amount of phloretin added per feeding.
[c] Total amount of phloretin/enzyme activity added.
[d] Final nothofagin concentration in solution.
[e] Based on final nothofagin and phloretin concentrations in solution.
[f] Difference total fed phloretin and final soluble nothofagin and phloretin.
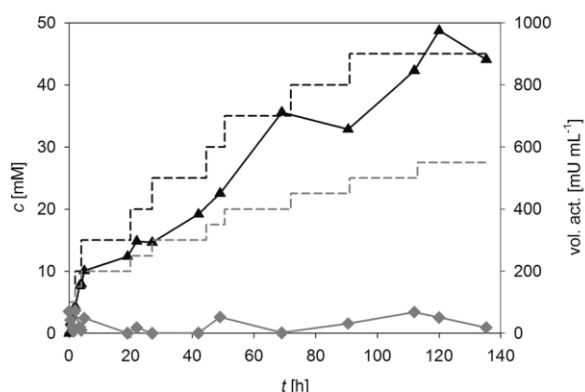[g] Total turnover number (mM nothofagin/mM total enzyme added).

Figure 5. Controlled feeding of phloretin is useful to enhance the nothofagin concentration in the biocatalytic synthesis. *Reaction conditions:* 100 mU mL^{-1} *Os*CGT/*Gm*SuSy, 5 mM phloretin, 300 mM sucrose, 1 mM UDP. After acceptor substrate depletion, 5 mM phloretin and 50 mU mL^{-1} *Os*CGT/*Gm*SuSy were added. Symbols: phloretin added (black dashed), *Os*CGT/*Gm*SuSy added (grey dashed), nothofagin (black), phloretin (grey)

Co-addition of enzyme alleviated restrictions on the product concentration (substrate conversion), which probably resulted from the combined effect of true enzyme activity loss and product inhibition. *Gm*SuSy is inhibited by D-fructose with a reported K_i of 9 mM.[20] This could also explain why lower total turnover numbers (ttn) were obtained with enzyme feed (~10,000) than without (~20,000). We also noticed the requirement to carefully control the phloretin feed to keep the acceptor concentration well below its solubility limit during the reaction. Figure S6 (Supporting Information) depicts *in situ* precipitation of phloretin under conditions where the acceptor feeding rate was not matched to the enzymatic consumption rate. Phloretin precipitation was clearly reflected in the mass balance in solution (Table 2).

Applying a suitable stepwise feeding of phloretin (5 or 10 mM) that included supplementation with fresh enzyme, it was possible to accumulate nothofagin in a concentration of around 45 mM, equivalent to 20 g L^{-1} (Figure 5, Table 2) which corresponded to RC_{max} values of approximately 45. It has to be noted that only with addition of enzyme and a low acceptor feed (5 mM) (Figure 5) could precipitation be avoided and almost quantitative conversion (~98%) was achieved. Nothofagin was isolated from reaction mixtures in a single step using preparative reversed phase C-18 HPLC with water to methanol gradient elution. Typically more than 80% of the initially applied phloretin (≥25 mg) could be recovered as highly pure nothofagin (Supporting information, Figure S8C). The isolation procedure is simple and not limited in scale.

Glycosyltransferases currently underachieve to a large extent their often-quoted high potential as synthetically usable glycosylation catalysts.[9] Notable

exceptions in the field of complex oligosaccharides notwithstanding,[17a,d,22] glycosyltransferase transformations have been realized almost exclusively at the analytical or minute preparative scale.[23] Glycosylations of similar poorly water-soluble natural product core structures such as the flavonoid quercetin have furnished hardly more than micromolar concentrations of the desired glycosidic compound.[17g,h,24] On comparison with the literature, therefore, the herein described enzymatic process of glucosylation of phloretin from sucrose stands out due to intensification, by up to three orders of magnitude, in terms of the product concentration that it has achieved for a coupled glycosyltransferase-catalyzed conversion. Furthermore, with the notable exception of the recently reported application of *Os*CGT for the production of nothofagin (~200 μM) and 2-hydroxynaringenin *C*-glucoside using engineered *S. cerevisiae* strains for whole cell conversions[23c] this is the first synthetic use of a *C*-glucosyltransferase. The here reported nothofagin process features efficient assembly of isolated *C*-glucosyltransferase in one pot with an adaptable module for UDP-glucose supply from sucrose, which serves as a highly expedient glucosyl donor for the overall conversion. We show that systematic analysis of thermodynamic conditions, kinetic properties of the glycosyltransferases and their stabilities is the key for identifying and thus eliminating critical constraints on the multi-component reaction system. Integration of biochemical optimization with reaction engineering was essential to overcome the restriction of acceptor substrate solubility. The number of UDP-glucose regeneration cycles was brought into a range (around 50) where one begins to truly capitalize on the coupling with the SuSy reaction. Reported RC_{max} values in literature are by far too small[17e 17g,17h] to justify enzymatic recycling of UDP-glucose. However, costs of the donor substrate are expected to prohibit the direct use of UDP-glucose for synthesis. Considering that unprocessed redbush tea contains nothofagin to just about 4.31 g kg^{-1} freeze-dried matter,[25] the high-yielding enzymatic synthesis developed herein is expected to remove compound availability as a critical bottleneck of the various medical and food-related applications of this aryl-*C*-glucoside.[14b,16,26] This study therefore makes a strong and so far missing case for the application of glycosyltransferases in biocatalytic synthesis of glycosylated natural products as fine chemicals.

Experimental Section

Coupled Enzymatic Conversions

Unless otherwise mentioned, standard reaction mixtures contained 5 mM phloretin, 100 mM sucrose, 0.5 mM UDP,

Advanced
Synthesis &
Catalysis

13 mM $MgCl_2$, 50 mM KCl, 0.13% (w/v) BSA, and 20% (v/v) DMSO in 50 mM HEPES buffer pH 7.5. Reactions were started by the addition of the indicated amounts of *Os*CGT and *Gm*SuSy. Enzymatic reactions were performed in 1.5 mL Eppendorf tubes at 30 °C using a thermomixer comfort for temperature control and agitation at 400 rpm. Samples were mixed with an equal volume of acetonitrile to stop the reaction. Precipitated protein was removed by centrifugation (13,200 rpm). The supernatant was analyzed using a reversed phase C-18 HPLC-assay.[18] All compounds of the reaction were analyzed. Reported conversions are confirmed by closed mass balance.

A detailed description of all experimental procedures can be found in the Supporting Information, which comprises: cloning of *Gm*SuSy; expression and purification of *Os*CGT and *Gm*SuSy; enzyme assays; details of the biocatalytic transformations performed; product isolation and identification; and analytical methods used. Any associated references are also given.

Acknowledgements

References

[1] a) C. J. Thibodeaux, C. E. Melancon 3rd, H.-W. Liu, *Angew. Chem.* **2008**, *120*, 9960–10007; *Angew. Chem. Int. Ed.* **2008**, *47*, 9814–9859; b) X. Wang, *FEBS Lett.* **2009**, *583*, 3303–3309.

[2] a) C. Mendez, J. A. Salas, *Trends Biotechnol.* **2001**, *19*, 449–456; b) R. W. Gantt, P. Peltier-Pain, J. S. Thorson, *Nat. Prod. Rep.* **2011**, *28*, 1811–1853.

[3] B. R. Griffith, J. M. Langenhan, J. S. Thorson, *Curr. Opin. Biotechnol.* **2005**, *16*, 622–630.

[4] C. Luley-Goedl, B. Nidetzky, *Nat. Prod. Rep.* **2011**, *28*, 875–896.

[5] L. L. Lairson, B. Henrissat, G. J. Davies, S. G. Withers, *Annu. Rev. Biochem.* **2008**, *77*, 521–555.

[6] M. M. Palcic, *Curr. Opin. Chem. Biol.* **2011**, *15*, 226–233.

[7] G. J. Williams, R. W. Gantt, J. S. Thorson, *Curr. Opin. Chem. Biol.* **2008**, *12*, 556–564.

[8] T. Desmet, W. Soetaert, P. Bojarova, V. Kren, L. Dijkhuizen, V. Eastwick-Field, A. Schiller, *Chem. Eur. J.* **2012**, *18*, 10786–10801.

[9] A. M. Daines, B. A. Maltman, S. L. Flitsch, *Curr. Opin. Chem. Biol.* **2004**, *8*, 106–113.

[10] a) M. Brazier-Hicks, K. M. Evans, M. C. Gershater, H. Puschmann, P. G. Steel, R. Edwards, *J. Biol. Chem.* **2009**, *284*, 17926–17934; b) T. Bililign, B. R. Griffith, J. S. Thorson, *Nat. Prod. Rep.* **2005**, *22*, 742–760.

[11] a) J. Harle, S. Gunther, B. Lauinger, M. Weber, B. Kammerer, D. L. Zechel, A. Luzhetskyy, A. Bechthold, *Chem. Biol.* **2011**, *18*, 520–530; b) T. Bililign, C. G. Hyun, J. S. Williams, A. M. Czisny, J. S. Thorson, *Chem. Biol.* **2004**, *11*, 959–969.

[12] J. Lee, S.-H. Lee, H. J. Seo, E.-J. Son, S. H. Lee, M. E. Jung, M. Lee, H.-K. Han, J. Kim, J. Kang, J. Lee, *Bioorg. Med. Chem.* **2010**, *18*, 2178–2194.

[13] N. C. Veitch, R. J. Grayer, *Nat. Prod. Rep.* **2011**, *28*, 1626–1695.

[14] a) J. B. Harborne, C. A. Williams, *Phytochemistry* **2000**, *55*, 481–504; b) P.-G. Pietta, *J. Nat. Prod.* **2000**, *63*, 1035–1042; c) T. P. Cushnie, A. J. Lamb, *Int. J. Antimicrob. Agents* **2005**, *26*, 343–356; d) W. Ren, Z. Qiao, H. Wang, L. Zhu, L. Zhang, *Med. Res. Rev.* **2003**, *23*, 519–534.

[15] a) R. G. dos Santos, A. R. Jesus, J. M. Caio, A. P. Rauter, *Curr. Org. Chem.* **2011**, *15*, 128–148; b) D. Y. W. Lee, M. He, *Curr. Top. Med. Chem.* **2005**, *5*, 1333–1350; c) K. W. Wellington, S. A. Benner, *Nucleosides Nucleotides Nucleic Acids* **2006**, *25*, 1309–1333.

[16] N. Krafczyk, M. A. Glomb, *J. Agric. Food Chem.* **2008**, *56*, 3368–3376.

[17] a) X. Chen, J. Zhang, P. Kowal, Z. Liu, P. R. Andreana, Y. Lu, P. G. Wang, *J. Am. Chem. Soc.* **2001**, *123*, 8866–8867; b) L. Elling, M. Grothus, M. R. Kula, *Glycobiology* **1993**, *3*, 349–355; c) C. H. Hokke, A. Zervosen, L. Elling, D. H. Joziasse, D. H. Van den Eijnden, *Glycoconjugate J.* **1996**, *13*, 687–692; d) Z. Liu, Y. Lu, J. Zhang, K. Pardee, P. G. Wang, *Appl. Environ. Microbiol.* **2003**, *69*, 2110–2115; e) S. Masada, Y. Kawase, M. Nagatoshi, Y. Oguchi, K. Terasaka, H. Mizukami, *FEBS Lett.* **2007**, *581*, 2562–2566; f) C. Rupprath, M. Kopp, D. Hirtz, R. Mueller, L. Elling, *Adv. Synth. Catal.* **2007**, *349*, 1489–1496; g) M. H. Son, B.-G. Kim, D. H. Kim, M. Jin, K. Kim, J.-H. Ahn, *J. Microbiol. Biotechnol.* **2009**, *19*, 709–712; h) K. Terasaka, Y. Mizutani, A. Nagatsu, H. Mizukami, *FEBS Lett.* **2012**, *586*, 4344–4350; i) A. Zervosen, L. Elling, *J. Am. Chem. Soc.* **1996**, *118*, 1836–1840.

[18] A. Gutmann, B. Nidetzky, *Angew. Chem.* **2012**, *124*, 13051–13056; *Angew. Chem. Int. Ed.* **2012**, *51*, 12879–12883.

[19] L. Elling, M. R. Kula, *Enzyme Microb. Technol.* **1995**, *17*, 929–934.

[20] M. Morell, L. Copeland, *Plant Physiol.* **1985**, *78*, 149–154.

[21] R. Anguenot, S. Yelle, B. Nguyen-Quoc, *Arch. Biochem. Biophys.* **1999**, *365*, 163–169.

[22] C.-H. Hsu, S.-C. Hung, C.-Y. Wu, C.-H. Wong, *Angew. Chem.* **2011**, *123*, 12076–12129; *Angew. Chem. Int. Ed.* **2011**, *50*, 11872–11923.

[23] a) B. G. Kim, H. J. Kim, J. H. Ahn, *J. Agric. Food Chem.* **2012**, *60*, 11143–11148; b) M. Zhou, A. Hamza, C. G. Zhan, J. S. Thorson, *J. Nat. Prod.* **2013**, *76*, 279–286; c) M. Brazier-Hicks, R. Edwards, *Metab. Eng.* **2013**, *16*, 11–20.

[24] E.-K. Lim, D. A. Ashford, B. Hou, R. G. Jackson, D. J. Bowles, *Biotechnol. Bioeng.* **2004**, *87*, 623–631.

[25] E. Joubert, *Food Chem.* **1996**, *55*, 403–411.

[26] P. W. Snijman, E. Joubert, D. Ferreira, X.-C. Li, Y. Ding, I. R. Green, W. C. A. Gelderblom, *J. Agric. Food Chem.* **2009**, *57*, 6678–6684.

Advanced
Synthesis &
Catalysis

Supporting Information

SUPPORTING INFORMATION

Leloir glycosyltransferases and natural product glycosylation: biocatalytic synthesis of the *C*-glucoside nothofagin, a major antioxidant of redbush herbal tea

Linda Bungaruang,[+a] Alexander Gutmann,[+a] and Bernd Nidetzky[a,*]

[a] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, 8010 Graz, Austria
Fax: (+43)-316-873-8434; phone:(+43)-316-873-8400; e-mail: bernd.nidetzky@tugraz.at

[+] These authors contributed equally.

Table of contents:

1 Methods

1.1 Chemicals and reagents

Unless otherwise indicated, all chemicals were purchased from Sigma-Aldrich in the highest purity available. Phloretin (98% purity) was obtained from AK Scientific. Enzymes for DNA manipulation and GeneJET™ Plasmid Miniprep Kit were from Fermentas. *Strep*-Tactin®Sepharose® and desthiobiotin were from IBA. BCA assay kit was purchased from Thermo Scientific.

1.2 Preparation of purified *Os*CGT and *Gm*SuSy

Construction of expression strains

The *Os*CGT gene (GenBank: FM179712) was received as a kind gift from the group of Prof. Robert Edwards (Centre for Bioactive Chemistry, Durham University, UK). It was provided in a pET-STRP3 vector which is a custom made derivative of pET-24d that enables protein expression with N-terminally fused *Strep*-tag II.[1, 2] The codon optimized gene of *Gm*SuSy (GenBank: AF030231) was synthesized with flanking *Nde*I and *Xho*I restriction sites and cloned in the plasmid pUC57 by GenScript. The gene was cut out using *Nde*I and *Xho*I restriction enzymes and inserted into the respective sites of the pET-STRP3 vector for expression as fusion protein with N-terminal *Strep*-Tag II.

Expression strains were created by transformation of electro-competent *E. coli* BL21-Gold (DE3) cells. The correct sequences were verified by sequencing the complete genes.

Protein expression and purification

The described *E. coli* strains were cultivated in 1 L baffled shaking flasks containing 250 mL Luria Bertani (LB) medium with 50 μg mL^{-1} kanamycin on a rotary shaker at 37°C and 120 rpm. Protein expression was induced at an optical density at 600 nm between 0.8 and 1.0 by addition of 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and the expression was carried out overnight at 25 °C. The cells were harvested by 30 min centrifugation at 4°C and 5000 rpm. After resuspension in water they were stored at -70°C until disruption by repeated passage through a cooled French press at 100 bar. Cell debris was removed by centrifugation for 45 min at 4°C and 13200 rpm.

The *Strep*-tag II fusion proteins were purified from the cell extract by affinity chromatography on *Strep*-Tactin®Sepharose® columns according to instructions of the manufacturer IBA. The columns had a volume of 3 mL and were operated by gravity flow. Equilibration of the column was done with 3 column volumes (CVs) of washing buffer W (100 mM Tris/HCl pH 8, 150 mM NaCl, 1 mM EDTA). The cell extract was diluted twofold with buffer W and filtrated through a 1.2 μm cellulose-acetate syringe filter before loading on the column. After washing with 5 CVs of buffer W the proteins were eluted with 3 CVs buffer E (100 mM Tris/HCl pH 8, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin) whereas the first 0.5 CVs were discarded and the rest was pooled. The column was regenerated using 15 CVs of buffer R (100 mM Tris/HCl pH 8, 150 mM NaCl, 1 mM EDTA, 1 mM hydroxy-azophenyl-benzoic acid) and equilibrated with 10 CVs of buffer W. Between the purification of different enzymes the columns were washed with 8 M guanidine-HCl to rule out cross-contaminations. Eluted enzymes were

S2

concentrated and buffer exchanged to 50 mM HEPES buffer pH 7.5 using centrifugal concentrators with a Molecular Weight Cut Off of 10 kDa. Enzymes were aliquoted and small aliquots were thawed for experiments as required and did not undergo multiple freeze-thaw cycles.

Protein concentration and purity

Protein concentrations were determined using the BCA method with bovine serum albumin (BSA) as standard. Purities were estimated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analyses and Coomassie Blue staining.

1.3   Determination of enzymatic activities and conversions

HPLC-based determination of phloretin and nothofagin (*Os*CGT activity)

A HPLC-assay was used for determination of dihydrochalcone concentrations (phloretin and nothofagin) and it was also applied as main assay for measurement of nothofagin production rates in *Os*CGT conversions. Reactions were performed in 1.5 mL reaction tubes at 30°C in a thermomixer at 400 rpm. Concentrations of UDP-glucose and phloretin as well as the exact buffer conditions are listed at the respective experiment. Reactions were started with the addition of *Os*CGT and stopped by mixing an aliquot of 100 µL with 100 µL acetonitrile. At least four samples (typically every 20 min) were taken during the linear initial rate to determine activities. Precipitated protein was removed by centrifugation for 20 min at room temperature and 13200 rpm. Depending on the concentrations 5 or 10 µL of the supernatant were used for analysis on an Agilent 1200 HPLC equipped with a Chromolith® Performance RP-18e endcapped column (100-4.6 mm) from Merck. The column was thermostatically controlled at 35°C and the separation was monitored by UV detection at 288 nm. Separation of phloretin and its glycosides was achieved by following method using water with 0.1 % TFA (trifluoroacetic acid) as solvent A and acetonitrile with 0.1% TFA as solvent B, respectively. A 7.5 min long linear gradient from 20 to 47.5% B (1 mL min^{-1}) was used for product separation. It was followed by 0.05 min of a linear gradient from 47.5 to 100 % B (1 mL min^{-1}) and 1.45 min of isocratic flow at 100% B (1.5 mL min^{-1}) to wash off hydrophobic compounds. After a   0.05 min linear gradient from 100 to 20 % B (1.5 mL min^{-1}) an isocratic flow of 2.45 min at 20% B (1.5 mL min^{-1}) was applied to equilibrate the column.

One Unit of *Os*CGT used in batch or fed-batch conversions for nothofagin production was defined as the amount of enzyme producing 1 µmol nothofagin per minute under following conditions: 0.6 mM UDP-glucose, 5 mM phloretin, 50 mM HEPES, pH 7.5, 13 mM MgCl$_2$, 50 mM KCl, 0.13% BSA and 20% (v/v) DMSO.

HPLC-based determination of UDP and UDP-glucose

Determination of UDP and UDP-glucose concentrations was done by HPLC using the sample preparation protocol with acetonitrile addition described for dihydrochalcone measurements. 10 µL sample were applied on an Agilent 1200 HPLC equipped with a Chromolith® Performance RP-18e endcapped column which was thermostatically controlled at 30°C. Separation was monitored by UV detection at 254 nm. Using 20 mM potassium phosphate buffer pH 6.8 with 2 mM tetrabutylammonium hydrogen sulfate as solvent A and acetonitrile as solvent B following gradient was applied with

S3

constant flow rate of 2 mL min^{-1}: 3 min of a linear gradient from 0 to 2% B were followed by a 7 min long linear gradient from 2 to 25% B to separate UDP-glucose from UDP. During 2 min of isocratic flow at 25% B hydrophobic compounds were washed off and after a 1 min long gradient from 25 to 0%, 2 min of isocratic flow at 0% B were applied to equilibrate the column.

Photometric assay for *Os*CGT and *Gm*SuSy sucrose synthesis activity

Activity of *Os*CGT as well as *Gm*SuSy activity in sucrose synthesis direction were also determined with a photometric assay using a modified spectrophotometric method.[3] It is based on coupling of UDP formation through pyruvate kinase (PK) and lactic dehydrogenase (LD) to equimolar NADH (ε = 6220 M^{-1} cm^{-1}) depletion which can be followed photometrically at 340 nm. The assay was performed in a discontinuous way whereas the reaction conditions for *Os*CGT measurements were as described for the HPLC based assay. Concentrations of UDP-glucose and fructose for *Gm*SuSy activity measurements are listed together with the used buffer in the corresponding sections. Samples of 150 µL were stopped by heating to 95 °C for 5 min. After removing precipitated protein by centrifugation at room temperature and 13200 rpm for 20 min, 100 µL of the supernatant were mixed in a Half Micro Cuvette with 400 µL of measuring solution. It consisted of 0.42 mM phosphoenolpyruvate and 0.18 mM NADH in a 50 mM HEPES buffer pH 7 containing 13 mM MnCl$_2$, 50 mM KCl and 0.13 % (w/v) BSA. The absorbance at 340 nm was determined using a Beckman Coulter DU 800 UV/VIS spectrophotometer before addition of 0.5 µL of a solution containing the coupling enzymes (PK: 682 U mL^{-1}; LD: 990 U mL^{-1} from rabbit muscle, Sigma-Aldrich). After incubation of the sealed cuvette at 30 °C for 45 min the absorbance at 340 nm was measured again and the concentration of UDP was calculated by the difference in absorbance before and after enzyme addition.

One Unit of *Gm*SuSy used in batch or fed-batch conversions for nothofagin production was defined as the amount of enzyme producing 1 µmol NAD$^+$ per minute under following conditions: 0.2 mM UDP, 300 mM sucrose, 50 mM HEPES, pH 7.5, 13 mM MgCl$_2$, 50 mM KCl, 0.13% BSA and 20% (v/v) DMSO.

Photometric assay for *Gm*SuSy sucrose cleavage activity

The measurement of linear initial rates for sucrose cleavage by *Gm*SuSy were performed by coupling the production of UDP-glucose to NADH formation using human UDP-α-D-glucose 6-dehydrogenase (hUGDH).[4] The oxidation of one molecule UDP-glucose to UDP-α-D-glucuronic acid is accompanied with the reduction of two NAD$^+$ molecules to NADH (ε = 6220 M^{-1}cm^{-1}) which is monitored photometrically at 340 nm. Concentrations of UDP and sucrose as well as the exact buffer conditions are listed at the respective experiment. The reaction volume was typically 1 mL and conversions were performed in 1.5 mL reaction tubes at 30 °C in a thermomixer at 400 rpm. The conversion was started by adding *Gm*SuSy. At least 4 samples (typically every 10 min) of 150 µL were withdrawn during the linear initial rate of the reaction. The reactions were stopped by heating the aliquots to 95 °C for 5 min. Precipitated protein was removed by 20 min of centrifugation at room temperature and 13200 rpm. 100 µL of the supernatant were mixed with 400 µL of measuring solution containing 2.5 mM NAD$^+$, 0.05 % Triton^{TM} X-100 and 100 mM HEPES pH 8.0 in a Half Micro Cuvette. The absorbance at 340

nm was determined using a Beckman Coulter DU 800 UV/VIS spectrophotometer before addition of 1.5 mU of hUGDH. After incubation of the sealed cuvette at 30°C for 45 min the absorbance was measured again and the concentration of UDP-glucose was calculated by the difference in absorbance before and after hUGDH addition.

1.4 Characterization of single enzymes

Determination of pH-profiles

The standard protocol of photometric *Os*CGT activity assay was modified by using a buffer mixture of 12.5 mM HEPES, 12.5 mM Tris and 25 mM CAPS of the respective pH. Reaction buffers were prepared from pH 6.0 to 11.0 in steps of 0.5 pH units. The HEPES concentration of the measuring solution was increased from 50 to 100 mM and it contained 16.25 mM $MgCl_2$ instead of 13 mM $MnCl_2$. The actual pH of the reaction mixture was determined as the average of pH measurements at the beginning and at the end of the observed time span.

Measurements of linear initial rates for sucrose cleavage by *Gm*SuSy were performed at 30°C by using the standard protocol of the photometric *Gm*SuSy activity assay. The reactions mixtures contained 100 mM sucrose, 2 mM UDP and 20 mM of buffer. 20 mM MES (pH 4.5 - 7.0), 20 mM HEPES (pH 7.0 - 8.5) and 20 mM CHES (pH 8.5 - 10.5) buffers were prepared in steps of 0.5 pH units. The actual pH of the reaction mixture was determined as the average of pH measurements at the beginning and at the end of the observed time span.

Linear initial rates for sucrose synthesis by *Gm*SuSy were measured with the standard photometric assay at 30°C using 15 mM fructose and 2 mM UDP-glucose in the buffers described for sucrose cleavage. Average pH during conversion was again calculated as average from samples withdrawn at beginning and end of the reaction.

Determination of equilibrium constants (K_{eq})

Equilibrium constant of *Os*CGT was determined by running standard glycosylation reaction (80 mU mL^{-1} *Os*CGT, 5 mM phloretin, 4.75 mM UDP-glucose) until no further conversion was observed. A standard buffer (50 mM HEPES, pH 7.5, 50 mM KCl, 13 mM $MgCl_2$, 0.13% (w/v) BSA, 20% (v/v) DMSO) was used. From all samples phloretin and nothofagin as well as UDP and UDP-glucose concentrations were determined on HPLC using the appropriate protocols. However, quantification of UDP failed and therefore only nothofagin, phloretin and UDP-glucose concentrations could be used for K_{eq} determination.

To test if there is any reverse activity detectable a conversion of 5 mM phloretin and 5 mM UDP-glucose (100 mU mL^{-1} *Os*CGT) was completed as described and subsequently another 5 mM UDP were added to test for deglycosylation of nothofagin. Therefore samples were taken as described over 12 hours and phloretin and nothofagin concentrations were determined per HPLC. Furthermore nothofagin and UDP were used directly as substrate in a reverse reaction using 200 mU mL^{-1} *Os*CGT (1 mM nothofagin, 1 mM UDP, 50 mM HEPES, pH 7.5, 13 mM $MgCl_2$, 50 mM KCl, 0.13% (w/v) BSA, 5% (v/v) ethanol). Nothofagin was extracted from an almost quantitative conversion of phloretin by

S5

*Os*CGT and contained therefore still small amounts of phloretin. Samples were taken over 23 hours and the analyzed by HPLC for phloretin production.

Sucrose cleavage (5 mM sucrose, 5 mM UDP) and synthesis (5 mM fructose and 5 mM UDP-glucose) were catalyzed by 50 mU mL^{-1} *Gm*SuSy using otherwise the same conditions as in *Os*CGT conversions. UDP-glucose concentrations were measured in all samples per HPLC. Comparable to *Os*CGT experiments UDP could not be quantified and concentrations of UDP, sucrose and fructose had to be calculated from starting concentrations and final UDP-glucose level. For K_{eq} calculations data from sucrose cleavage and synthesis was combined.

Determination of half saturation constants

Half saturation and phloretin inhibition constant of *Os*CGT were determined by measurement of linear initial rates at 30 °C in 50 mM HEPES, pH 7.5 containing 50 mM KCl, 13 mM MgCl$_2$, 0.13% (w/v) BSA and 20% (v/v) DMSO using the HPLC-assay. For each substrate 10 different concentrations in a suitable range were used. Phloretin was varied from 0.01 - 10 mM (4 mU mL^{-1} *Os*CGT, 2 mM UDP-glucose) and UDP-glucose was varied from 0.005 - 5 mM (4 mU mL^{-1} *Os*CGT, 1 mM phloretin).

For *Gm*SuSy half saturation constants were determined in sucrose cleavage and synthesis direction. All conversions were made with 2 mU mL^{-1} *Gm*SuSy at 30 °C with 20 mM HEPES, pH 7.5 as buffer. Nine different concentrations of each substrate were used and all activities were measured using the appropriate photometric assays. Sucrose was varied from 5 - 300 mM (2 mM UDP), UDP from 0.0025 - 5 mM (400 mM sucrose), fructose from 0.05 - 30 mM (2 mM UDP-glucose) and UDP-glucose from 0.01 - 5 mM (40 mM fructose).

Influence of cosolvents on *Os*CGT and *Gm*SuSy activity

To test the effect of cosolvents on *Os*CGT and *Gm*SuSy (sucrose cleavage) activity initial rate measurements were done at different concentrations of ethanol or DMSO. Reference measurements for *Gm*SuSy (2 mU mL^{-1}, 0.2 mM UDP, 100 mM sucrose, 50 mM HEPES pH 6.5, 10 mM MgCl$_2$) were made without any organic solvent and those for *Os*CGT (0.2 mU mL^{-1}, 0.6 mM UDP-glucose, 0.1 mM phloretin, 50 mM HEPES pH 7, 13 mM MnCl$_2$, 50 mM KCl, 0.13% (w/v) BSA) were done at 5% of the respective solvent. *Gm*SuSy and *Os*CGT activity were measured using described photometric assays.

Determination of enzyme stability

Enzymes were incubated for certain times (0-24 h) in the final reaction mix at 30 °C and 400 rpm on a thermomixer but without UDP (*Gm*SuSy) or UDP-glucose (*Os*CGT), respectively. Reactions were started by addition of UDP or UDP-glucose and the loss of activity in dependence of incubation time was used to fit first order inactivation kinetics for calculation of half-lifes. In case of *Os*CGT glycosylation activity was measured by HPLC (15 mU mL^{-1}, 5 mM phloretin, 1 mM UDP-glucose) and sucrose cleavage of *Gm*SuSy was determined photometrically (2 mU mL^{-1}, 100 mM sucrose, 1 mM UDP). The buffer for *Gm*SuSy was 50 mM HEPES, pH 7.5 with 13 mM MgCl$_2$, and 20 % (v/v) DMSO and that for *Os*CGT additionally contained 50 mM KCl and 0.13 % (w/v) BSA.

S6

1.5 *Os*CGT conversions coupled with *Gm*SuSy

Batch conversions of *Os*CGT coupled with *Gm*SuSy

Standard reaction mixtures contained 5 mM phloretin, 100 mM sucrose, 0.5 mM UDP, 13 mM MgCl$_2$, 50 mM KCl, 0.13 % (w/v) BSA, 20 % (v/v) DMSO and 50 mM HEPES buffer pH 7.5. Enzymatic conversions were initiated by the addition of the respective amount of *Os*CGT and *Gm*SuSy. Reactions were performed in 1.5 mL reaction tubes at 30 °C in a thermomixer at 400 rpm. Samples were analyzed with the HPLC-based *Os*CGT assay to determine the concentrations of phloretin and nothofagin.

For optimizing the conversion conditions the concentrations of the substrate were varied separately: UDP (0.005 - 1 mM; 10 mU mL^{-1} *Os*CGT, 10 mU mL^{-1} *Gm*SuSy), sucrose (5 - 500 mM; 10 mU mL^{-1} *Os*CGT, 10 mU mL^{-1} *Gm*SuSy) and phloretin (5 - 30 mM; 190 mU mL^{-1} *Os*CGT, 120 mU mL^{-1} *Gm*SuSy). On variation of sucrose and UDP only initial nothofagin production rates were analyzed during the first two hours (3-4 samples) but on variation of phloretin complete conversion curves over 24 hours were recorded.

For comparison of phloretin conversion by *Os*CGT with and without coupling to UDP-glucose recycling with *Gm*SuSy instead of 0.5 mM UDP 6 mM UDP-glucose were used. Both enzymes were applied at 50 mU mL^{-1}.

Fed-batch conversions of *Os*CGT coupled with *Gm*SuSy

Phloretin concentrations of the batch conversion were modified for the fed-batch experiments. The reaction solution contained 300 mM sucrose, 1 mM UDP, 13 mM MgCl$_2$, 50 mM KCl, 0.13 % (w/v) BSA, 20 % (v/v) DMSO in 50 mM HEPES buffer (pH 7.5). Separate reactions with 5 and 10 mM phloretin, respectively were made in a total volume of 2 mL. Reactions were started by enzyme addition (100 mU mL^{-1} *Os*CGT/*Gm*SuSy) and incubated at 30 °C and 400 rpm on a thermomixer. Directly after taking samples phloretin and nothofagin concentrations were measured on HPLC and upon depletion of phloretin, new substrate was added from a stock of 500 mM phloretin in DMSO. The increase of substrate concentration was thereby equivalent to the initial concentrations of 5 and 10 mM, respectively. Separate reactions were made with and without feeding of fresh enzyme (50 mU mL^{-1}) at the same time as the substrate feed.

1.6 Product isolation and identification by HPLC and NMR

The product of phloretin conversions by *Os*CGT has previously been identified as nothofagin using ^{1}H and ^{13}C NMR.[5] Therefore 1 mM phloretin was converted to nothofagin (2 mM UDP-glucose in 50 mM HEPES pH 7, 13 mM MnCl$_2$, 50 mM KCl, 0.13% BSA and 5% (v/v) ethanol) and the resulting product was almost quantitatively extracted from the reaction solution by repeated extraction with ethyl acetate. After removing the solvent under reduced pressure the product was dissolved in MeOD$_4$ for NMR analysis. HPLC retention time of the product of coupled conversion of *Os*CGT with *Gm*SuSy was compared with nothofagin for product verification.

Nothofagin was purified by preparative reversed phase C-18 HPLC. Proteins were removed from phloretin conversions by centrifugal concentrators with a Molecular Weight Cut Off of 10 kDa. The HPLC system from Knauer Technologies (smartline pump 1000 V7603, smartline DAD UV detector

2600, manual 6 port injection valve) was equipped with a C18-reversed phase VP 125/21 Nucleodur 100-5 C18ec column. Water containing 0.1% TFA was used as solvent A and methanol as solvent B. Separation of nothofagin from other compounds was achieved by a 20 min long gradient from 20 to 100% B followed by 6 min of isocratic flow at 100% B. Separation was monitored by UV-detection at 288 nm. Fractions containing nothofagin were pooled, methanol was evaporated under reduced pressure and water was removed by freeze drying.

2 Results
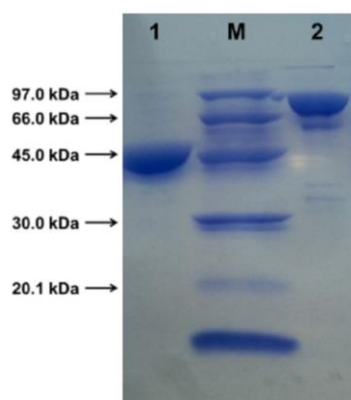
Figure S1: SDS-PAGE of enzymes purified by *Strep-tag* affinity chromatography: Lane 1: *Os*CGT (51.3 kDa); Lane 2: *Gm*SuSy (94.1 kDa); Lane M: Low Molecular Weight standard (GE Healthcare)
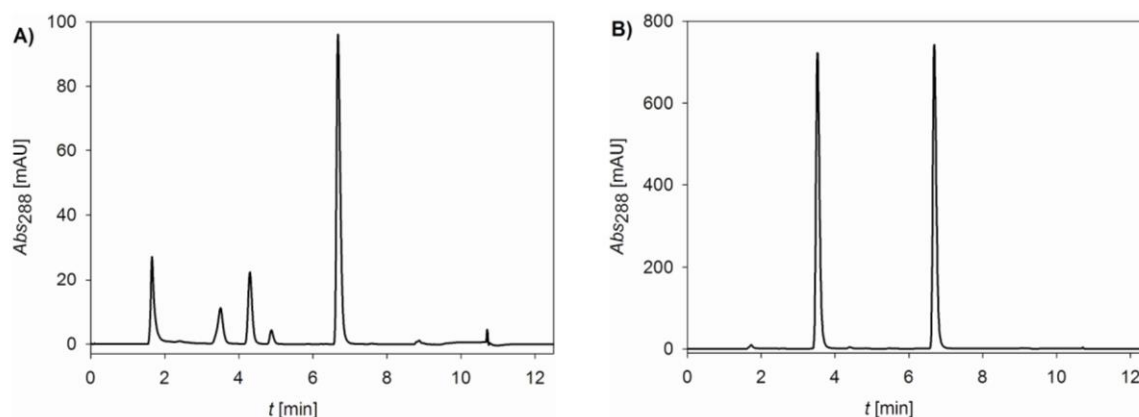
Figure S2: Separation of phloretin from its glycosides using RP-HPLC: A) mixture containing non-binding compounds like UDP(-glucose) (1.7 min), 3'-*C*-glycoside nothofagin (3.5 min), 2'-*O*-glycoside phlorizin (4.3 min), 4'-*O*-glycoside (4.9 min) and phloretin (6.7 min); B) partial phloretin conversion using *Os*CGT and *Gm*SuSy contains only 3'-*C*-glycoside nothofagin (3.5 min), phloretin (6.7 min) and small amounts of UDP(-glucose)

S8
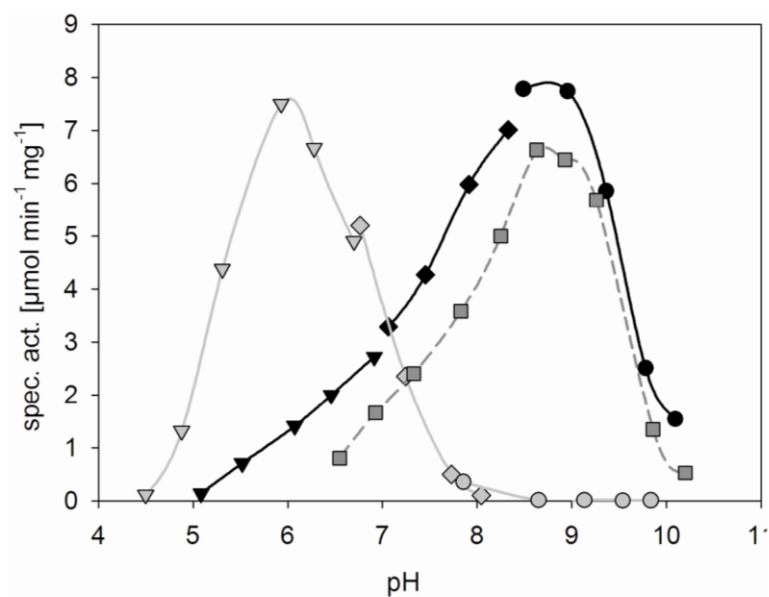
Figure S3. pH profiles for nothofagin synthesis by *Os*CGT (dark grey, dashed line), sucrose cleavage (light grey) and sucrose synthesis (black) by *Gm*SuSy using a mixture of HEPES, Tris and CAPS (square), MES (triangle), HEPES (diamond) or CHES (circle) as buffer.
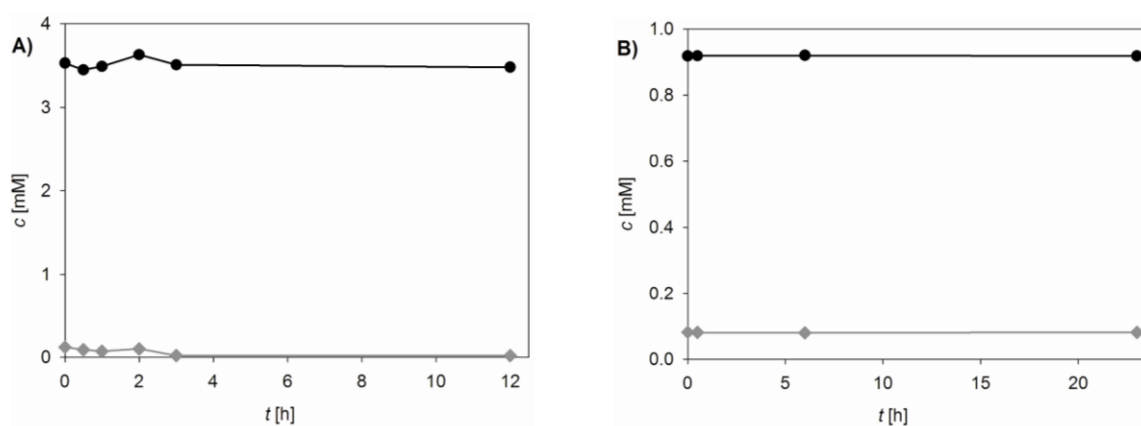
Figure S4: Deglycosylation of nothofagin (black) to phloretin (grey) by *Os*CGT could not be observed: A) 5 mM phloretin and 5 mM UDP-glucose were almost completely converted by 100 mU mL^{-1} *Os*CGT (not shown). When adding further 5 mM UDP (dilution) no deglycosylation of nothofagin was observed; B) Incubation of ~1 mM nothofagin containing some phloretin with 1 mM UDP (200 mU mL^{-1} *Os*CGT) did also not result in any further phloretin production.

S9
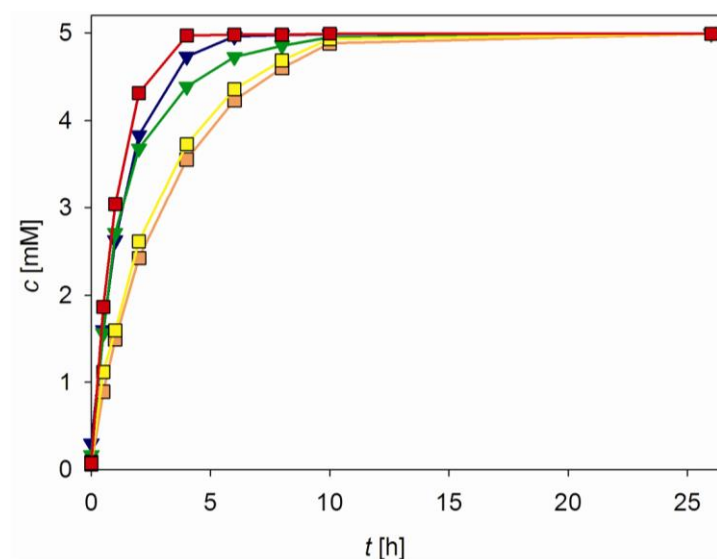
Figure S5. Nothofagin production in *Os*CGT catalyzed conversions of 5 mM phloretin (100 mM sucrose) at various conditions: 50 mU mL^{-1} *Os*CGT, 6 mM UDP-glucose (green); 50 mU mL^{-1} *Os*CGT/*Gm*SuSy, 6 mM UDP-glucose (blue); 50 mU mL^{-1} *Os*CGT/*Gm*SuSy, 0.5 mM UDP (orange); 50 mU mL^{-1} *Os*CGT, 250 mU mL^{-1} *Gm*SuSy, 0.5 mM UDP (yellow); 250 mU mL^{-1} *Os*CGT, 50 mU mL^{-1} *Gm*SuSy, 0.5 mM UDP (red)
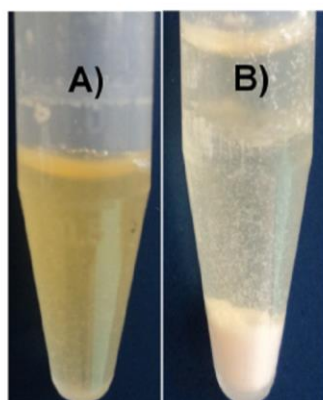
Figure S6: Increased solubility of nothofagin compared to phloretin: A) In a fed-batch conversion of *Os*CGT and *Gm*SuSy with feeding of 5 mM phloretin and enzyme in total 45 mM phloretin were applied and fully converted to nothofagin causing only enzyme to precipitate; B) In an identical solution without enzyme most of the applied 45 mM phloretin was precipitated.

S10
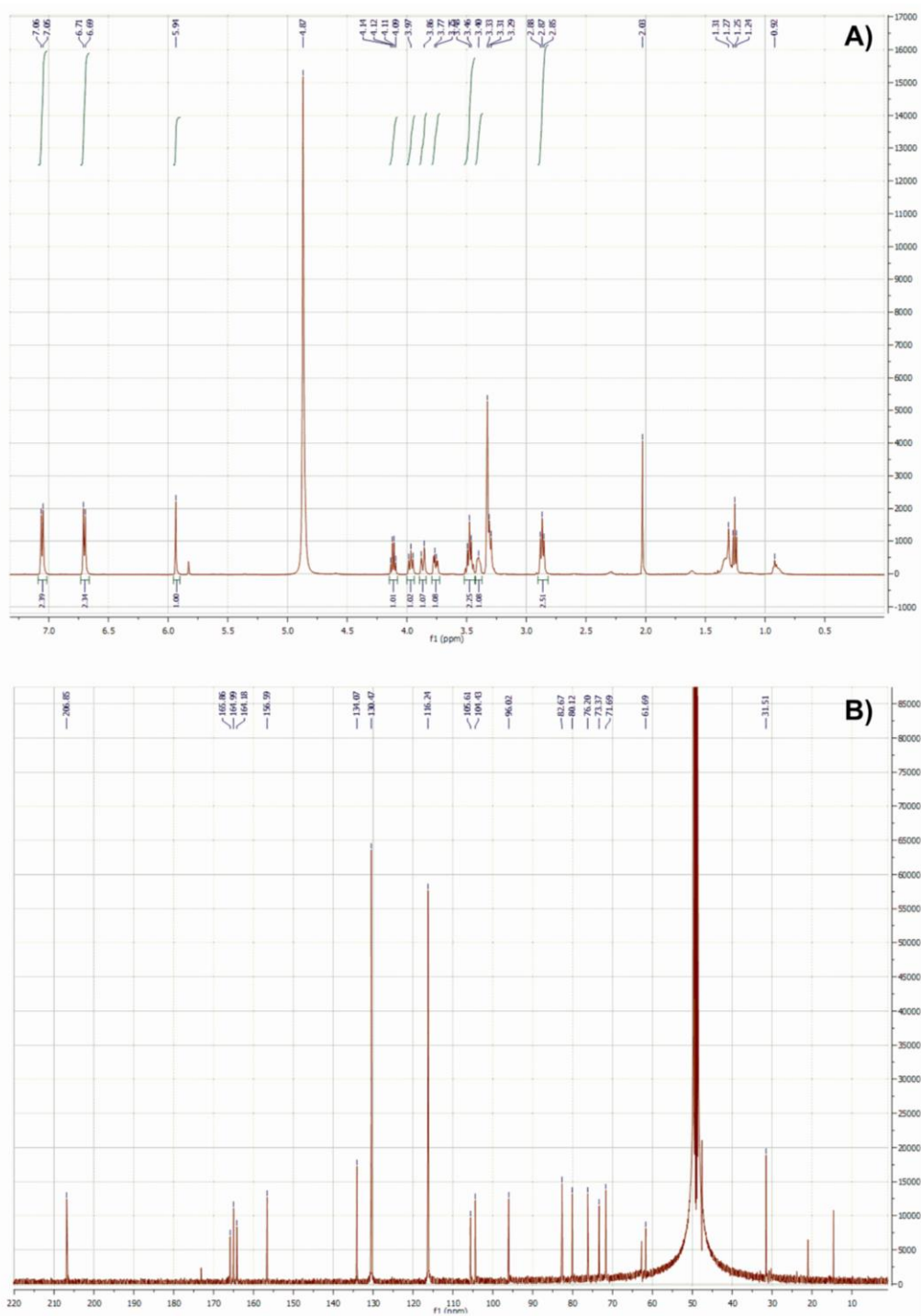
Figure S7: ^{1}H A) and ^{13}C-NMR B) of nothofagin after extraction with ethyl acetate from enzymatic conversion of phloretin by *Os*CGT.

S11

Synthesis of the *C*-glucoside nothofagin, a major antioxidant of redbush herbal tea
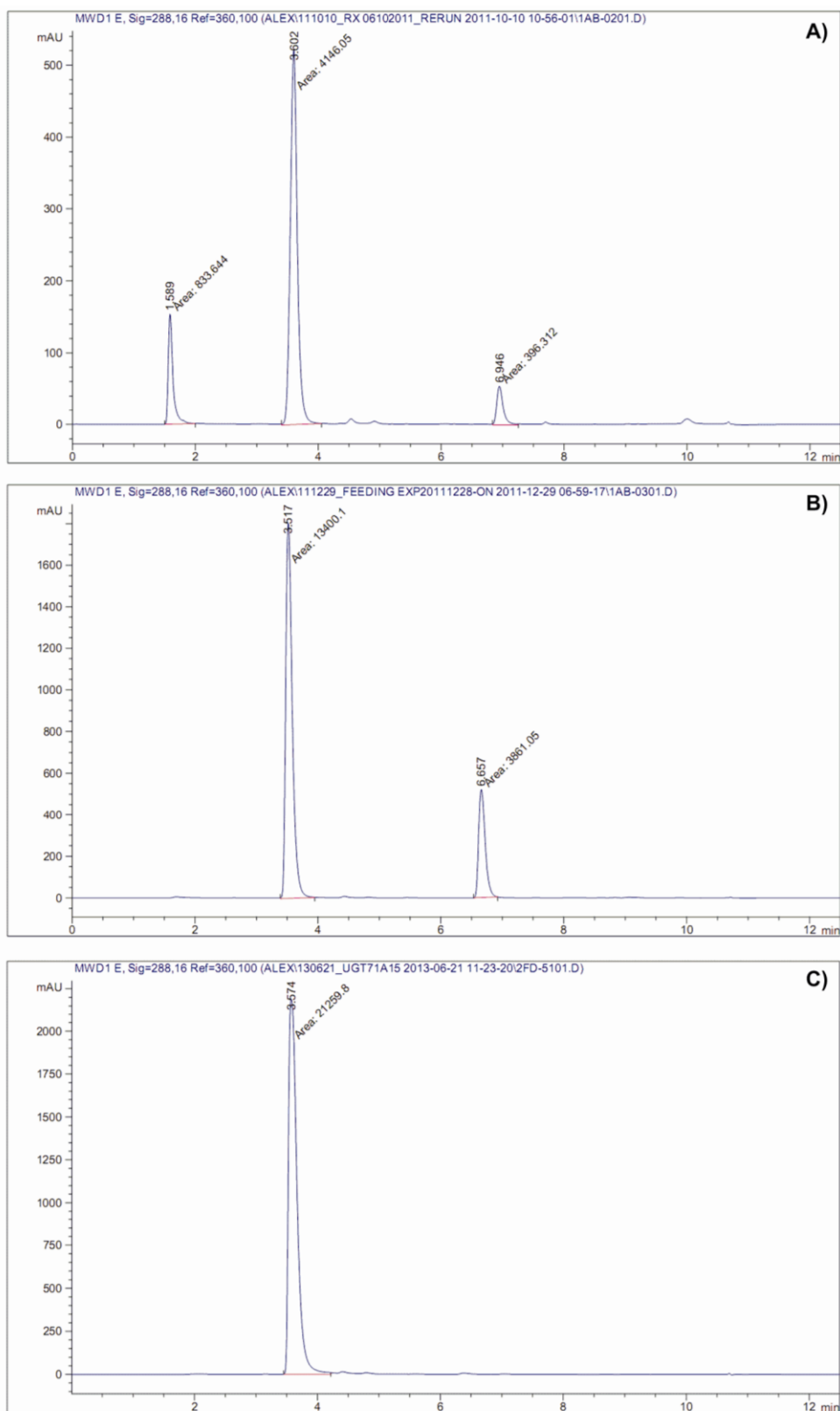
Figure S8: HPLC chromatograms (UV detection at 288 nm) showing nothofagin peak at ~3.5 min either from authentic standard A), after partial phloretin conversion in a coupled reaction of *Os*CGT and *Gm*SuSy B) and after purification by reversed phase HPLC C).

S12

48

3 References

[1] M. Brazier-Hicks, K. M. Evans, M. C. Gershater, H. Puschmann, P. G. Steel, R. Edwards, *J. Biol. Chem.* **2009**, *284*, 17926-17934.
[2] D. P. Dixon, T. Hawkins, P. J. Hussey, R. Edwards, *J. Exp. Bot.* **2009**, *60*, 1207-1218.
[3] S. Gosselin, M. Alhussaini, M. B. Streiff, K. Takabayashi, M. M. Palcic, *Anal. Biochem.* **1994**, *220*, 92-97.
[4] S. Egger, A. Chaikuad, K. L. Kavanagh, U. Oppermann, B. Nidetzky, *Biochem. Soc. Trans.* **2010**, *38*, 1378-1385.
[5] A. Gutmann, B. Nidetzky, *Angew. Chem.* **2012**, 124, 13051–13056; *Angew. Chem. Int. Ed.* **2012**, *51*, 12879-12883.

A two-step *O*- to *C*-glycosidic bond rearrangement using complementary glycosyltransferase activities

ChemComm

COMMUNICATION

A two-step *O*- to *C*-glycosidic bond rearrangement using complementary glycosyltransferase activities†

Alexander Gutmann, Corinna Krump, Linda Bungaruang and Bernd Nidetzky*

An efficient 2′-*O*- to 3′-*C*-β-ᴅ-glucosidic bond rearrangement on the dihydrochalcone phloretin to convert phlorizin into nothofagin was achieved by combining complementary *O*-glycosyltransferase (OGT) and *C*-glycosyltransferase (CGT) activities in a one-pot transformation containing catalytic amounts of uridine 5′-diphosphate (UDP). Two separate enzymes or a single engineered dual-specific O/CGT were applied. Overall (quantitative) conversion occurred in two steps *via* intermediary UDP–glucose and phloretin.
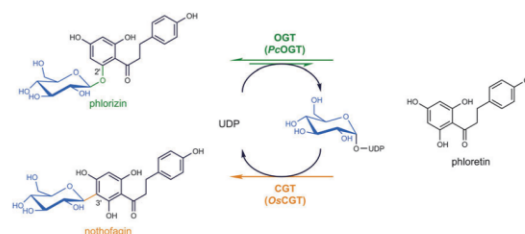
Scheme 1 Conversion of phlorizin into nothofagin using a direct *O*- to *C*-glucoside rearrangement in two catalytic steps.

Leloir glycosyltransferases (GTs) are selective catalysts of synthetically useful glycosylation reactions.[1,2] Naturally they catalyse glycosyl transfer from an activated donor, typically a nucleoside-diphosphate (NDP) sugar, onto metabolic target acceptor(s) (ESI,† Scheme S1a).[3] Recognising the synthetic scope of glycosyltransferase reactions run backwards, researchers introduced two-step exchange reactions to the field (ESI,† Scheme S1b).[4–6] Complementary glycosyltransferase activities are combined in a one-pot conversion where NDP-sugar or an acceptor substrate for the actual synthetic transformation is generated *in situ* from a reactant glycoside and NDP *via* a reverse glycosyltransferase reaction. The overall catalytic conversion is steered to achieve swapping of glycosyl residues between different acceptor substrates (aglycon exchange),[4,6–10] or to result in an alternative glycosylation of a single acceptor compound (sugar exchange).[4,7,11] Exchange processes were exploited in different glycosylations of small molecules and enabled glycoengineering of natural products *in vitro*.[12]

In this communication, we have discovered that phenolic *O*- to aromatic *C*-glycosidic bond rearrangement (Scheme 1; ESI,† Scheme S1c) is achievable by coupled glycosyltransferase reactions. The overall

conversion is formally equivalent to a chemical Fries-type rearrangement;[13] however, the mechanisms are distinct. Glycosyltransferase conversions involve intermediary release of aglycon and NDP-sugar, followed by regio- and stereoselective *C*-glycosylation.

A proof of principle was obtained for conversion of phlorizin, the 2′-*O*-glucoside of the dihydrochalcone phloretin, into the corresponding 3′-*C*-glucoside nothofagin (Scheme 1). The catalytic process involved *O*-glucosyltransferase (OGT) and *C*-glucosyltransferase (CGT) activities, derived from a specificity matched pair of enzymes from pear (*Pyrus communis*; *Pc*OGT)[14] and rice (*Oyrza sativa*; *Os*CGT),[15] or from an engineered "promiscuous" *O*/*C*-glucosyltransferase (O_CGT). Only catalytic amounts of uridine 5′-diphosphate (UDP) were required in the conversion to generate intermediary UDP–glucose and phloretin *via* reverse reaction of the OGT. The overall two-step rearrangement proceeded under thermodynamic control whereby a large driving force for the second step, aromatic *C*-glycosylation, enabled nothofagin formation in quantitative yields. The phlorizin–nothofagin pair stands for a number of homologous *O*- and *C*-glycosidic natural products, in that the *O*-glycoside occurs naturally in relatively high abundance over the quite rare *C*-glycoside.[16,17] Fruit trees (*Rosaceae*) contain large amounts of phlorizin in their bark[18] whereas significant quantities of nothofagin were reported only from redbush herbal tea.[19] Direct *O*- to *C*-glycosidic bond rearrangement makes effective use of both the glycon and the aglycon moiety of the substrate. In principle, therefore, it presents an expedient and particularly atom-efficient transformation of *O*-glycosides

Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12/1, A-8010 Graz, Austria.
E-mail: bernd.nidetzky@tugraz.at

† Electronic supplementary information (ESI) available: Experimental procedures used; the scheme of reverse glycosyltransferase reactions; enzyme purity; deglycosylation of phlorizin by *Pc*OGT; the effect of cosolvents and substrate concentrations on rearrangement; HPLC trace of phloretin (glycosides); potentiometric titrations; optimisation of conversion by *Os*CGT_I121D. See DOI: 10.1039/c4cc00536h

(*e.g.* flavonoid glycosides) readily available from the natural products pool. Due to their high resistance to hydrolysis, the resulting *C*-glycosides are of special interest for development of bioactive substances with enhanced *in vivo* half-lives.[16,20]

Purified preparations of *Pc*OGT and *Os*CGT were obtained from *E. coli* overexpression cultures using His-tag and Strep-tag affinity chromatography, respectively. The Ile[121] → Asp variant (I121D) of *Os*CGT, previously shown to exhibit dual specific O_CGT activity,[21] was produced and isolated in the same way as the wild-type enzyme (ESI,† Fig. S1). All reactions were carried out with 20% (by volume) DMSO to enhance phloretin solubility. We confirmed that the yield and the selectivity of the bi-enzymatic rearrangement were not affected by DMSO (ESI,† Fig. S3). Enzymes were fully active and exhibited suitable stability under these conditions. Phloretin, phlorizin, nothofagin, and other phloretin glucosides were identified and quantified by HPLC (ESI,† Fig. S4b).[21] UDP–glucose and UDP were measured by capillary zone electrophoresis. OGT and CGT activities were determined using HPLC-based assays.

We noticed that rearrangement of phlorizin to nothofagin could become effective only when main requirements concerning enzyme specificity, reaction kinetics, and reaction thermodynamics in each step were adequately met. *Pc*OGT and *Os*CGT were known to glucosylate phloretin at position $2'\text{-}O$[14,21] and $3'\text{-}C$,[21,22] respectively. Interference from secondary *O*-glucoside hydrolase activity in the enzymes used could also be discounted based on earlier evidence.[21] Moreover, *Os*CGT alone was inactive as a "rearrangement enzyme" when phlorizin and UDP were offered as substrates. *C*-Glucosylation of phloretin from UDP–glucose was known to be largely irreversible ($K_{eq} > 400$; 30 °C, pH 7.5).[23] However, kinetic and thermodynamic characteristics of the *Pc*OGT reaction required clarification. Fig. 1a compares the pH-rate profile for deglucosylation of phlorizin to UDP to the pH-rate profile for glucosylation of phloretin from UDP–glucose. *Pc*OGT was more active in the direction of phlorizin synthesis (18.3 U mg^{-1}) than degradation (5.6 U mg^{-1}) at the respective optimum pH. Interestingly, despite the rather uniform pH effects on enzyme activity in each reaction direction (except for

the slight pH range shift), we realised that the reaction equilibrium constant (K_{eq} = [phlorizin][UDP]/[UDP–glucose][phloretin]) increased dramatically in response to pH change from 6.5 to 8.8 (Fig. 1b). Attainment of reaction equilibrium was affirmed rigorously under all conditions, ruling out interference from enzyme activity loss at high or low pH. It was also confirmed that the pH was stable during the conversions. Correlation between $\log K_{eq}$ and pH was linear with a large slope of +1.6 (± 0.1), implying that conversion of phlorizin and UDP involves the uptake of proton(s). Potentiometric titration of each compound present in the reaction (ESI,† Fig. S5) revealed the likely importance of protonation of UDP (pK_a ∼ 5.6). The immediate ramification of results in Fig. 1b is that exploitation of the *Pc*OGT reverse reaction for supplying substrates in adequate steady-state concentrations to the *Os*CGT reaction will only be practical at pH 7.5 or lower. Half-saturation constants of *Os*CGT were determined to be 0.009 mM for phloretin and 0.024 mM for UDP–glucose,[23] defining lower limits to the respective substrate concentration for effective utilization of the *C*-glucosylation activity present.

Rearrangement of phlorizin (5 mM) into nothofagin was examined in a one-pot reaction that contained UDP (2 mM), *Pc*OGT, and *Os*CGT. Because the optimum pH for *C*-glycosylation (pH 8.5)[23] did not match pH conditions applicable to phlorizin conversion by *Pc*OGT, we tested the coupled enzyme reaction at different pH values in the range 5.9–7.8. Nothofagin was produced under all conditions used, demonstrating the system's functionality in principle. Fig. 2a compares phlorizin consumption to the corresponding formation of nothofagin and phloretin after 5 h of reaction. Interestingly, whereas the actual phlorizin conversion was not strongly affected by pH change in the applied range, the resulting product distribution, nothofagin compared to phloretin, showed pronounced pH dependence. Accumulation of phloretin at low pH indicated critical limitations due to insufficient *Os*CGT activity under these conditions. However, *C*-glucosylation was quite effective in the pH range 6.7–7.8 where only small amounts of phloretin were detected next to the main product nothofagin (2.5 mM; 50% substrate conversion).
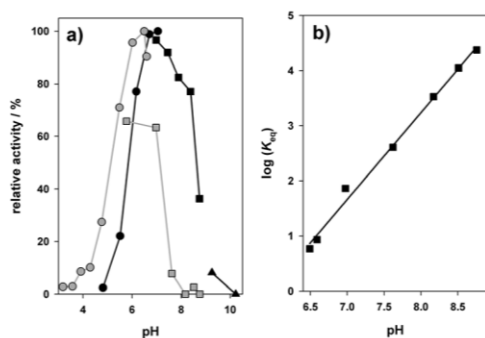

Fig. 1 pH effects on *Pc*OGT activity for phlorizin synthesis and degradation (a), and on the reaction equilibrium constant (b). (a) Relative activities of *Pc*OGT; black symbols (forward reaction: 0.1 mM phloretin, 0.6 mM UDP–glucose, 100% = 18.3 U mg^{-1}); grey symbols (reverse reaction: 1.0 mM phlorizin, 2.0 mM UDP, 100% = 5.6 U mg^{-1}). The buffers used were citrate (circles), tris (squares), and CAPS (triangles). (b) K_{eq} is for synthesis of phlorizin and UDP from phloretin and UDP–glucose.
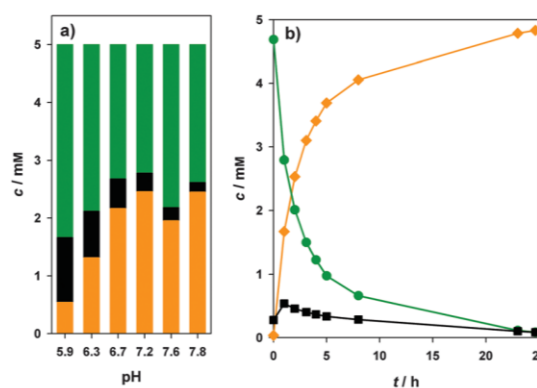

Fig. 2 Conversion of phlorizin (5 mM) *via* *O*- to *C*-glycosidic bond rearrangement in the presence of 2 mM UDP using 100 mU mL^{-1} *Pc*OGT and 50 mU mL^{-1} *Os*CGT. (a) Product distribution after reaction for 5 h at different pH conditions. (b) Reaction time course at pH 7.0. Colours show phlorizin (green), nothofagin (orange), and phloretin (black).
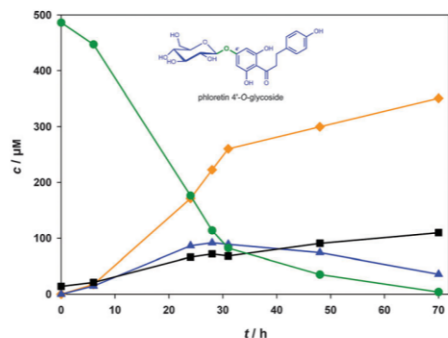
52

A two-step *O*- to *C*-glycosidic bond rearrangement using complementary glycosyltransferase activities

Fig. 3 Conversion of 0.5 mM phlorizin (green) by I121D mutant of *Os*CGT (2.5 mU mL^{-1}) in the presence of 5 mM UDP and 1 mM UDP–glucose. Further symbols: phloretin (black), nothofagin (orange), 4′-*O*-glucoside of phloretin (blue).

Fig. 2b shows a full time course for enzymatic reaction at pH 7.0, demonstrating that quantitative conversion of phlorizin into nothofagin was made possible under these conditions. Isolation of highly pure nothofagin is typically achieved by reversed phase C-18 HPLC in yields of more than 80%.[23] The phloretin concentration attained maximum ($\sim$0.5 mM) early during the reaction, only to drop to a very low level later on. This time dependence was fully consistent with the proposed role of phloretin as an intermediary product in an overall two-step rearrangement process. Evidence that immediately after the reaction started, nothofagin formation was clearly lagging behind the formation of phloretin (see the phloretin "burst" of $\sim$0.3 mM at $t = 0$ h in Fig. 2b) gave additional indication of the reaction in two discrete biocatalytic steps.

Kinetic analysis (ESI,† Fig. S6) showed that the phlorizin consumption rate (r_P) was hyperbolically dependent on UDP and phlorizin concentration, resulting in apparent half-saturation constants of 1.22 ± 0.16 mM (phlorizin) and 0.074 ± 0.010 mM (UDP).

The overall rearrangement rate (r_N; nothofagin formation) paralleled r_P under all conditions, except at high phlorizin concentration ($\geq$3 mM) where r_N dropped off strongly in relation to r_P. The behaviour of r_N is explained by substrate inhibition of *Os*CGT at high phlorizin concentration.[23] However, disparity between r_P and r_N did not diminish the nothofagin yield when 10 mM phlorizin was applied (ESI,† Fig. S4a). High rearrangement rates could be maintained with as little as 0.25 mM UDP (ESI,† Fig. S6b) confirming the requirement for only catalytic amounts of UDP in the overall conversion.

We recognised the interesting possibility of performing the two-step rearrangement using only a single enzyme that exhibits OGT as well as CGT activity. The I121D mutant of *Os*CGT which catalyses glucosyl transfer to O2 and O4, next to C3 of the phloretin acceptor, was considered useful.[21] Kinetic properties of the mutant required adaptation of reaction conditions for synthesis of nothofagin from phlorizin (ESI,† Methods, Fig. S7–S9). In particular, the UDP concentration was increased to 5 mM to promote an otherwise inefficient reverse reaction from phlorizin substrate, which was applied at a lowered concentration of 0.5 mM. Additionally, UDP–glucose was supplied at 1.0 mM to drive the *C*-glucosylation.

Fig. 3 displays the time course of reaction catalysed by I121D mutant under these conditions. Turnover frequency of the mutant ($\sim$0.001 s^{-1})[21] restricted the amount of enzyme activity usable in the reaction, resulting in relatively slow conversion. However, all of the initial phlorizin was converted, and nothofagin was obtained in 70% yield. Unlike in the reaction catalysed by coupled OGT and CGT (Fig. 2b), the phloretin concentration increased steadily during the mutant-catalysed reaction, approaching a maximum value of $\sim$0.1 mM at the time when phlorizin was depleted (Fig. 3). Therefore, this indicates that *C*-glucosylation was rate limiting for the overall rearrangement catalysed by the mutant. Interestingly, small amounts of phloretin 4′-*O*-glucoside were also produced and used up later in the conversion (Fig. 3). Therefore, this implied O2 to O4 positional rearrangement in phloretin *O*-glucosides catalysed by the mutant.

Summarising, the biocatalytic rearrangement discovered in this study is a remarkable chemical transformation that might open up new opportunities for synthesis of aromatic *C*-glycosidic natural products or natural product-like structures from (readily available) phenolic *O*-glycosidic substrates. Unlike glycosyltransferase exchange reactions in which overcoming thermodynamic restrictions presents the main issue,[4,8,24] the example of phlorizin conversion into nothofagin shows that (quantitative) *O*- to *C*-glycosidic bond rearrangement is promoted by a large driving force on the *C*-glycosylation. While the study provides a clear proof of principle, expansion of the synthetic scope of the biocatalytic rearrangement will be important. Development of new pairs of complementary enzymes through discovery work and protein engineering is required whereby identification of new CGT enzymes will be the key.

Notes and references

1 M. M. Palcic, *Curr. Opin. Chem. Biol.*, 2011, **15**, 226–233.
2 C. J. Thibodeaux, C. E. Melançon III and H. W. Liu, *Angew. Chem., Int. Ed.*, 2008, **47**, 9814–9859.
3 L. L. Lairson, B. Henrissat, G. J. Davies and S. G. Withers, *Annu. Rev. Biochem.*, 2008, **77**, 521–555.
4 C. Zhang, B. R. Griffith, Q. Fu, C. Albermann, X. Fu, I. K. Lee, L. Li and J. S. Thorson, *Science*, 2006, **313**, 1291–1294.
5 H. B. Bode and R. Müller, *Angew. Chem., Int. Ed.*, 2007, **46**, 2147–2150.
6 A. Minami, K. Kakinuma and T. Eguchi, *Tetrahedron Lett.*, 2005, **46**, 6187–6190.
7 R. Chen, H. Zhang, G. Zhang, S. Li, G. Zhang, Y. Zhu, J. Liu and C. Zhang, *J. Am. Chem. Soc.*, 2013, **135**, 12152–12155.
8 R. W. Gantt, P. Peltier-Pain, W. J. Cournoyer and J. S. Thorson, *Nat. Chem. Biol.*, 2011, **7**, 685–691.
9 C. Zhang, C. Albermann, X. Fu and J. S. Thorson, *J. Am. Chem. Soc.*, 2006, **128**, 16420–16421.
10 C. Zhang, Q. Fu, C. Albermann, L. Li and J. S. Thorson, *ChemBioChem*, 2007, **8**, 385–390.
11 C. Zhang, E. Bitto, R. D. Goff, S. Singh, C. A. Bingman, B. R. Griffith, C. Albermann, G. N. Phillips and J. S. Thorson, *Chem. Biol.*, 2008, **15**, 842–853.
12 R. W. Gantt, P. Peltier-Pain and J. S. Thorson, *Nat. Prod. Rep.*, 2011, **28**, 1811–1853.
13 R. G. dos Santos, A. R. Jesus, J. M. Caio and A. P. Rauter, *Curr. Org. Chem.*, 2011, **15**, 128–148.
14 C. Gosch, H. Halbwirth, B. Schneider, D. Hölscher and K. Stich, *Plant Sci.*, 2010, **178**, 299–306.

53

15 M. Brazier-Hicks, K. M. Evans, M. C. Gershater, H. Puschmann, P. G. Steel and R. Edwards, *J. Biol. Chem.*, 2009, **284**, 17926–17934.
16 T. Bililign, B. R. Griffith and J. S. Thorson, *Nat. Prod. Rep.*, 2005, **22**, 742–760.
17 J. Härle, S. Günther, B. Lauinger, M. Weber, B. Kammerer, D. L. Zechel, A. Luzhetskyy and A. Bechthold, *Chem. Biol.*, 2011, **18**, 520–530.
18 J. R. L. Ehrenkranz, N. G. Lewis, C. R. Kahn and J. Roth, *Diabetes/Metab. Res. Rev.*, 2005, **21**, 31–38.
19 E. Joubert, *Food Chem.*, 1996, **55**, 403–411.

20 C. Dürr, D. Hoffmeister, S. E. Wohlert, K. Ichinose, M. Weber, U. von Mulert, J. S. Thorson and A. Bechthold, *Angew. Chem., Int. Ed.*, 2004, **43**, 2962–2965.
21 A. Gutmann and B. Nidetzky, *Angew. Chem., Int. Ed.*, 2012, **51**, 12879–12883.
22 M. Brazier-Hicks and R. Edwards, *Metab. Eng.*, 2013, **16**, 11–20.
23 L. Bungaruang, A. Gutmann and B. Nidetzky, *Adv. Synth. Catal.*, 2013, **355**, 2757–2763.
24 P. Peltier-Pain, K. Marchillo, M. Zhou, D. R. Andes and J. S. Thorson, *Org. Lett.*, 2012, **14**, 5086–5089.

54

Two-step *O*- to *C*-glycosidic bond rearrangement using complementary glycosyltransferase activities

Alexander Gutmann, Corinna Krump, Linda Bungaruang and Bernd Nidetzky

Institute of Biotechnology and Biochemical Engineering,
Graz University of Technology, Petersgasse 12/1, A-8010 Graz, Austria

Electronic Supplementary Information

Table of contents

1

1   Methods

1.1   Chemicals and reagents

Unless otherwise noted, all chemicals were obtained from Sigma-Aldrich (Vienna, Austria) in the highest purity available. Phlorizin dihydrate (≥98%) was purchased from Carl Roth (Karlsruhe, Germany). DNA modifying enzymes were from Thermo Scientific (Waltham, MA, US) and PCR primers were obtained from Life Technologies (Carlsbad, CA, US). Phusion® High-Fidelity DNA Polymerase was purchased from New England Biolabs (Ipswich, MA, US). *Strep*-Tactin® Sepharose® and desthiobiotin were obtained from IBA (Goettingen, Germany).

1.2   Strain construction

The *Os*CGT gene (GenBank: FM179712) was kindly provided from the group of Prof. Robert Edwards (Centre for Bioactive Chemistry, Durham University, UK) in a pET-STRP3 vector which is a custom made derivative of pET-24d that enables protein expression with an N-terminally fused *Strep*-tag II.[1-2] Introduction of the mutation I121D was described elsewhere.[3] The *Pc*OGT gene (UGT88F2; GenBank: FJ854496) was a kind gift from the group of Prof. Karl Stich (Institute of Chemical Engineering, Vienna University of Technology, Austria).[4] Flanking restriction sites for *Nde*I and *Xho*I were added and the respective internal restriction sites were removed by overlap extension PCR as described previously.[3] Subsequently the gene was inserted into *Nde*I and *Xho*I sites of a pET-28a vector for expression with N-terminally fused His-tag. Expression strains were created by transformation of electro-competent *E. coli* BL21-Gold (DE3) cells with the described plasmids.

Correct sequences of the complete genes and the regions around both cloning sites of the pET-Strep3 plasmid were verified by sequencing. Vector maps and complete DNA sequences of *Os*CGT and *Pc*OGT as well as the pET-Strep3 plasmid can be found in the Appendix.

1.3   Enzyme expression and purification

Cultivation of *E. coli* cells for protein expression was described elsewhere in detail.[3] Cells were grown at 37°C and 120 rpm in 1 L baffled shake flasks containing 300 mL LB-media (50 µg mL^{-1} kanamycin). At an optical density at 600 nm of 0.8-1.0, protein expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After overnight expression at 25°C cells were harvested by 30 min centrifugation at 5,000 rpm and 4°C, resuspended in water and stored at -70°C until disruption by repeated passage through a cooled French press at 100 bar. Before purification cell debris was removed by centrifugation (45 min, 13,200 rpm, 4°C) and cell extract was filtrated through a 1.2 µm cellulose-acetate filter.

*Os*CGT and *Os*CGT_I121D were purified by *Strep*-tag affinity chromatography on 3 mL gravity flow *Strep*-Tactin® Sepharose® columns at 4°C as previously described.[3] After loading cell extract the column was washed with 5 column volumes (CV) of washing buffer (100 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA). Enzymes were eluted with 3 CVs of elution buffer (100 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin) whereas the first 0.5 CVs were discarded, and the rest was pooled.

*Pc*OGT was purified at 4°C on a 5 mL HiTrap™ Chelating FF column (GE Healthcare) loaded with Ni^{2+}. A flow rate of around 5 mLmin^{-1} was applied manually using a syringe. Cell extract was loaded on the column after equilibration with 5 CVs of buffer W (20 mM Tris pH 7.4, 500 mM NaCl, 20 mM imidazole). After washing with further 5 CVs of buffer W *Pc*OGT was eluted in two steps using 4 CVs buffer E1 (20 mM Tris pH 7.4, 500 mM NaCl, 250 mM imidazole) and 5 CVs of buffer E2 (20 mM Tris pH 7.4, 500 mM NaCl, 500 mM imidazole), respectively.

Fractions containing purified proteins were concentrated and buffer was exchanged to 25 mM HEPES, pH 7.0, using centrifugal concentrators with a Molecular Weight Cut Off of 10,000. Purified enzymes were stored as small aliquots at -20°C to avoid multiple freeze-thaw cycles. Protein concentrations of purified enzymes were determined photometrically on a NanoDrop 2000 system (Thermo Scientific). Molecular weight and molar extinction coefficients were calculated using Peptide Properties Calculator. Enzyme purities were estimated by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) using Coomassie Blue staining.

1.4   Quantification of dihydrochalcones – HPLC based activity assay

Quantification of phloretin and its glycosides was achieved by an HPLC based assay as described elsewhere in detail.[3] It was applied to follow all enzymatic conversions and to determine activities of *Pc*OGT and *Os*CGT (I121D). Briefly, reaction mixtures contained if not mentioned otherwise 0.6 mM UDP-glucose, 1 mM phloretin, 25 mM Tris pH 7, 13 mM MgCl$_2$, 50 mM KCl, 0.13% BSA and 20% DMSO. Conversions were started by glycosyltransferase addition. After certain incubation times at 30°C samples were withdrawn and enzymatic

2

conversions were stopped by mixing with an equal volume of acetonitrile. Precipitated protein was removed by centrifugation for 15 min at 13,200 rpm. Typically 10 µL of supernatant were applied on an Agilent 1200 HPLC system equipped with a Chromolith® Performance RP-18e endcapped column (100–4.6 mm) from Merck. Separation of phloretin and its glycosides was monitored by UV detection at 288 nm under thermostatic control at 35°C by following method. Solvent A: water with 0.1% TFA; Solvent B: acetonitrile with 0.1% TFA; Gradient: 7.5 min: 20-47.5% B (1 mL min^{-1}), 0.05 min: 47.5-100% B (1 mL min^{-1}), 1.45 min: 100% B (1.5 mL min^{-1}), 0.05 min: 100-20% B (1.5 mL min^{-1}), 2.45 min: 20% B (1.5 mL min^{-1})

One unit of *Os*CGT (I121D) or *Pc*OGT activity was defined as the amount of enzyme glucosylating 1 µmol phloretin per minute under the following conditions: 0.6 mM UDP-glucose, 1 mM (*Os*CGT/*Pc*OGT) or 0.1 (OsCGT_I121D) phloretin, 25 mM Tris (*Os*CGT, *Pc*OGT) or BisTris (OsCGT_I121D), pH 7, 13 mM MgCl$_2$, 50 mM KCl, 0.13% BSA and 20% DMSO at 30°C. At least 4 distinct measurements were used to calculate linear initial rates.

1.5 Quantification of UDP and UDP-glucose - capillary zone electrophoresis

Concentrations of UDP and UDP-glucose were determined by capillary zone electrophoresis. Samples were prepared as for HPLC measurements by mixing with an equal volume of acetonitrile to stop enzymatic conversion. Precipitated protein was removed by centrifugation (15 min; 13,200 rpm). Capillary zone electrophoresis was performed using a 3D capillary electrophoresis system (Hewlett Packard) equipped with a fused-silica capillary (56 cm × 50 µm) with an extended light path. The capillary was preconditioned by a 6 min flush with background electrolyte (20 mM sodium borate, pH 9.3). Samples were loaded by pressure injection (50 mbar, 10 s) and compounds were resolved at 22 kV over 22 min at 18°C. UDP and UDP-glucose were monitored on a diode array detector at 262 nm.

1.6 Characterization of *Pc*OGT reaction reversibility

To test if phloretin formation from phlorizin is a result of reverse glycosylation by *Pc*OGT, 1 mM phlorizin was incubated at 30°C in 50 mM BisTris, pH 7.0, 13 mM MgCl$_2$, 50 mM KCl, 0.13% BSA and 20% DMSO. Two mM UDP, 25 mU mL^{-1} *Pc*OGT, or both were added. Conversion of phlorizin was followed by HPLC.

To obtain pH-profiles of *Pc*OGT in glycosylation and deglycosylation direction, the standard protocol of HPLC based activity assay was modified in following manner. A set of reaction buffers was prepared from pH 3 to 10 in steps of 0.5 (25 mM citrate pH 3-7; 25 mM Tris pH 7-9.5, 25 mM CAPS pH 9.5-10). All reactions contained 13 mM MgCl$_2$, 50 mM KCl, 0.13% BSA and 20% DMSO. In glycosylation direction 0.1 mM phloretin and 0.6 mM UDP-glucose were used as substrates and reactions were started by addition of following amounts of *Pc*OGT: 60 ng mL^{-1} (pH 5-5.5); 30 ng mL^{-1} (pH 6-6.5, 9.5-10); 15 ng mL^{-1} (pH 7.0-9.0); Deglycosylation reactions (pH 3-9.5) contained 1 mM phlorizin and 2 mM UDP and were started by addition of 60 ng mL^{-1} *Pc*OGT. All reactions were incubated at 30°C and samples were withdrawn every 20 min for 1h. The actual pH of reaction mixtures was determined as the average of pH measurements at the beginning and at the end of the observed time span.

To reach equilibrium, another 8.3 µg mL^{-1} *Pc*OGT were added to reactions in deglycosylation direction after samples for initial rate measurements were withdrawn. After overnight incubation at 30°C phloretin and phlorizin concentrations were experimentally determined by HPLC. The UDP-glucose concentration was inferred from reaction stoichiometry. Equilibrium constants were calculated in glycosylation direction according to equation 1.

$$K_{eq} = \frac{c_{phlorizin} \cdot c_{UDP}}{c_{phloretin} \cdot c_{UDP-glucose}}$$

(1)

1.7 Potentiometric titration of phlorizin, phloretin, UDP and UDP-glucose

pK_a-values of all compounds involved in phlorizin conversion by *Pc*OGT (phlorizin, phloretin, UDP, UDP-glucose) were determined by potentiometric titration. Of all analytes, solutions of 7.5 mM were made in decarbonated water containing 20% DMSO, 13 mM MgCl$_2$ and 50 mM KCl to reproduce rearrangement conditions. Analyte volumes of 10 (UDP, UDP-glucose) or 20 mL (phlorizin, phloretin) were titrated at 30°C with 150 mM NaOH as titrant using a SenTix® Mic electrode from WTW (Weilheim, Germany) for pH monitoring.

3

1.8   *O-* to *C*-glucoside rearrangement by *Os*CGT and *Pc*OGT

Unless mentioned otherwise reaction mixtures contained 2 mM UDP and 5 mM phlorizin in 50 mM Tris pH 7 containing 13 mM MgCl$_2$, 50 mM KCl, 0.13% BSA and 20% DMSO. Conversions were started by addition of 100 mU mL^{-1} *Pc*OGT and 50 mU mL^{-1} *Os*CGT. All reactions were incubated at 30°C. Aliquots were withdrawn, stopped by acetonitrile addition and analysed by RP-HPLC as described.

To test the influence of cosolvents on *O-* to *C*-glucoside rearrangement conversions were carried out in presence of different concentrations (5, 10, 15 or 20%) of either ethanol or DMSO.

The effect of pH on the rearrangement was studied by replacing the standard buffer with 50 mM HEPES buffers ranging from pH 6 to 8.5 (in steps of 0.5). The actual pH of reaction mixtures was determined as the average of pH measurements at beginning and end of 5 h long conversions at 30°C.

The impact of substrate concentrations on *O-* to *C*-glycosidic bond rearrangement was studied by varying phlorizin (0.05 - 10 mM; 2 mM UDP) and UDP (0.005 - 1.2 mM; 5 mM phlorizin; 5% DMSO). Linear initial phlorizin consumption and nothofagin production rates were determined during initial 25 (phlorizin variation) or 60 min (UDP variation).

1.9   *O-* to *C*-glucoside rearrangement by *Os*CGT_I121D

Unless mentioned otherwise all conversions were made with 1.25 mU mL^{-1} *Os*CGT_I121D in 50 mM BisTris buffer, pH 7.0, containing 13 mM MgCl$_2$, 50 mM KCl, 0.13% BSA and 20% DMSO. During 24-70 h long time courses at least 7 samples were analysed by RP-HPLC as described previously.

For optimisation of substrate concentrations, 0.05, 0.5 and 5 mM phlorizin were tested in combination with 0.5 and 5 mM UDP. Evaluation of pH dependency was done using 50 mM BisTris buffers of pH 6.0, 7.0 and 8.0 for conversion of 0.5 mM phlorizin with 5 mM UDP. The effect of UDP-glucose addition (0, 0.5, 1, 5 mM) was also investigated using 0.5 mM phlorizin and 5 mM UDP. Conversion of 0.5 mM phlorizin under optimized conditions was performed with 2.5 mU mL^{-1} *Os*CGT_I121D in presence of 5 mM UDP and 1 mM UDP-glucose.

4

2. Results

2.1 Synthetic use of reverse glycosyltransferase reactions
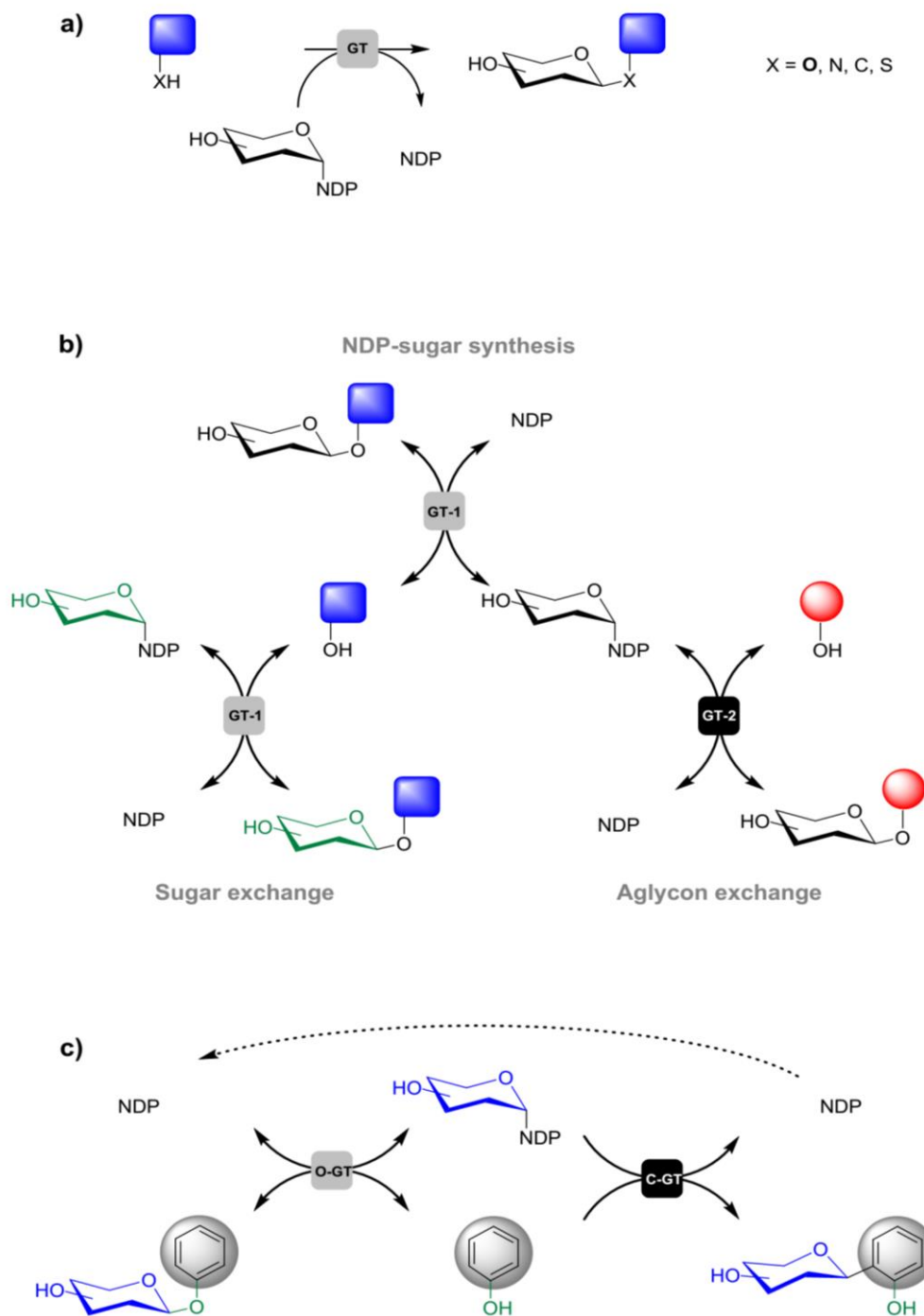
Scheme S1 Natural product glycosylations through a canonical glycosyltransferase reaction (a), two-step exchange processes exploiting a suitable glycosyltransferase reaction in reverse direction (b), and *O*- to *C*-glycosidic bond rearrangement catalysed by complementary glycosyltransferase activities (c)
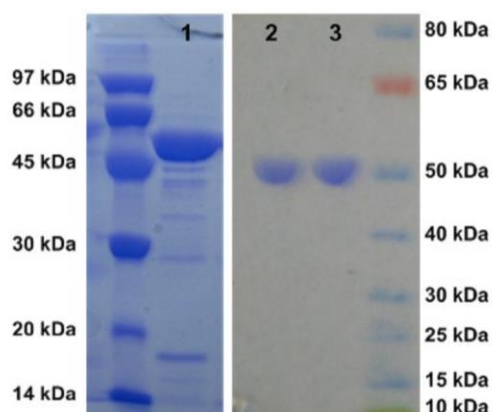
5

2.2 Determination of enzyme purity by SDS-PAGE

Fig. S1 SDS-PAGE of enzymes purified by affinity chromatography: Lane 1: His-tagged *Pc*OGT (55.7 kDa) purified on Ni Sepharose™; Lane 2 and 3: *Strep*-tagged *Os*CGT and *Os*CGT_I121D, respectively purified by *Strep*-tag affinity chromatography (51.3 kDa each)

2.3 Reverse glycosylation of phlorizin by *Pc*OGT
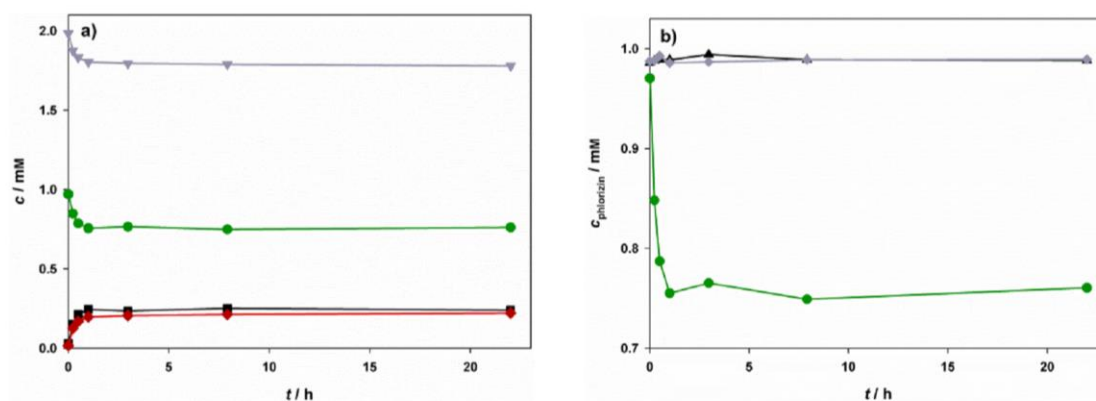
Fig. S2 Deglycosylation of phlorizin by *Pc*OGT at pH 7.0: a) 1 mM phlorizin (green) was partially deglycosylated in presence of 2 mM UDP (grey) and *Pc*OGT (25 mU mL^{-1}). Equal amounts of phloretin (black) and UDP-glucose (red) were formed. b) Phlorizin was only deglycosylated when UDP as well as *Pc*OGT were present (green). In presence of only UDP (grey) or solely *Pc*OGT (black) phlorizin was not converted. Deglycosylation of phlorizin is therefore solely achieved by enzyme-catalysed transfer of glucose to UDP and not by a potential hydrolytic side reaction.

6

2.4   Effect of organic cosolvents on *O*- to *C*-glucoside rearrangement
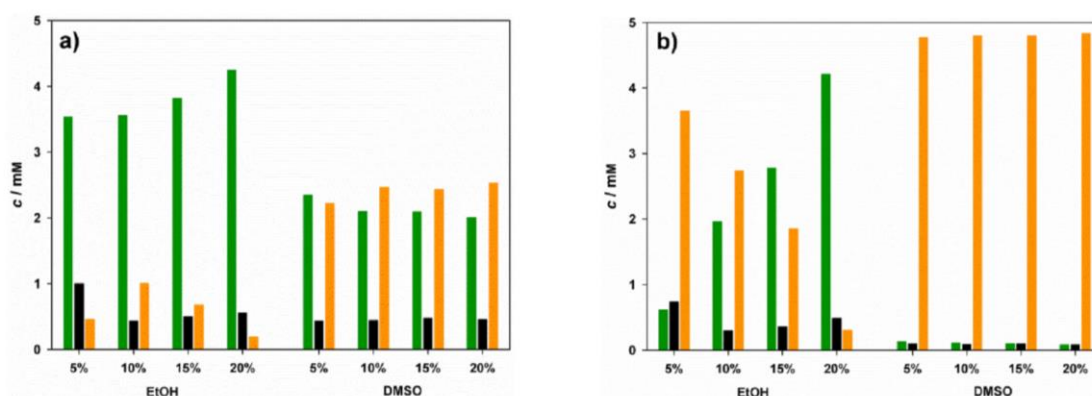

Fig. S3 Yields of *O*- to *C*-glucosidic bond rearrangement at various cosolvent conditions after (a) 2h and (b) 24h of conversion (5 mM phlorizin, 2 mM UDP, 100 mU mL^{-1} *Pc*OGT, 50 mU mL^{-1} *Os*CGT). Phlorizin (green), phloretin (black), nothofagin (orange); Any concentration of ethanol reduced conversions drastically. In contrast even the highest applied DMSO concentration of 20% showed no significant reduction of initial or final conversion.

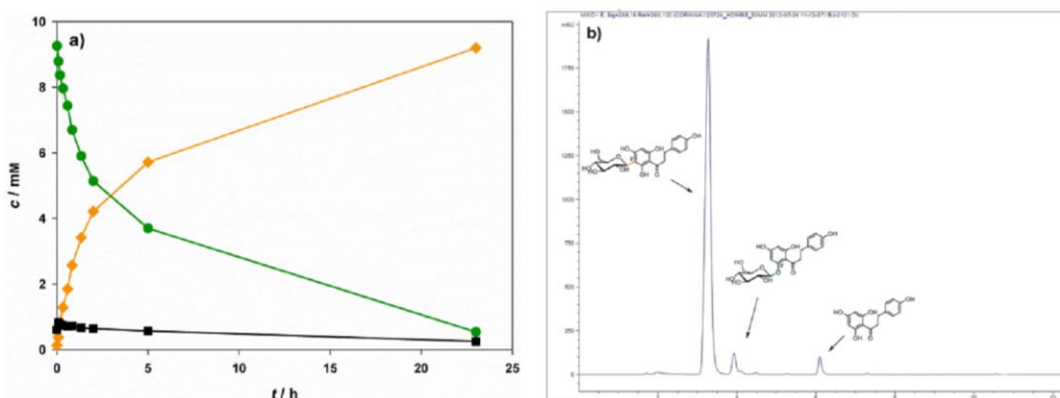**2.5   *O*- to *C*-glucoside rearrangement in the absence of side reactions**

Fig. S4 10 mM phlorizin were almost completely converted to the *C*-glucoside nothofagin within 24 h (10 mM phlorizin, 2 mM UDP, 100 mU mL^{-1} *Pc*OGT, 50 mU mL^{-1} *Os*CGT). a) Time course of rearrangement: Quick initial consumption of phlorizin (green) gradually slows down after around one hour. The aglycon phloretin (black) remains relatively constant over the entire conversion. After 24 h of incubation 92% (9.2 mM) of applied phlorizin are converted to nothofagin (orange). b) HPLC analysis after 24h of conversion: The UV-trace (288 nm) shows a dominant nothofagin signal (3.3 min) with minor peaks for phlorizin (3.9 min) and phloretin (6.1 min). Absence of any further signals indicates that no side products were formed.

7

2.6 Potentiometric titration of phlorizin, phloretin, UDP and UDP-glucose

To investigate the cause for pH dependency of the equilibrium constant in *O*-glycosylation, potentiometric titrations of phloretin and its 4'-*O*-glucoside phlorizin (Fig. S5a) as well as UDP and UDP-glucose (Fig. S5b) were performed. Observed pK_a values were in general agreement with literature data, which allowed their assignment to functional groups (Fig. S5c).[5-6] Slight discrepancies between observed and reported pK_a values are probably caused by differences in experimental conditions (addition of DMSO, temperature, ionic strength).

The titration curve of the aglycon phloretin (Fig. S5a, black) was in good agreement with literature, the reported pK_a values being 7.0, 9.4, and 10.5.[5] Phlorizin shares with phloretin a pK_a value in the neutral pH range. This pK_a supposedly represents the first deprotonation event on one of the three phenolic hydroxyl groups on the aromatic ring involved in glycosylation. The second pK_a value of phloretin of around pH 9.3 is missing in phlorizin (green). This indicates that the pK_a represents a second deprotonation event on the same ring which is missing in phlorizin because of glucose bound to the oxygen at position 2'. A further pronounced plateau at around pH 10.8 is visible in both compounds. It most likely reflects unresolved pK_a values from a final deprotonation on the glycosylated aromatic ring and also deprotonation of the single hydroxyl groups on the second aromatic ring. The observed large increase of the equilibrium constant of *O*-glycosylation with increased pH in the range from 6.5 to 8.8 cannot be explained by deprotonation of phloretin upon glycosylation. Phloretin glycosylation causes no significant proton transfer below pH 8. At higher pH, proton release would rather occur in deglycosylation than in glycosylation direction due to increase of acidity in going from phlorizin to phloretin.

Protonation UDP and UDP-glucose can, however, explain a drastic preference of glycosylation at high pH. UDP (Fig. S5b, grey) displays a pK_a value of around 5.6, which is missing in UDP-glucose. It represents the second deprotonation of the terminal phosphate group, which is glycosylated in the case of UDP-glucose. Fig. S5b reveals that UDP is predominantly deprotonated throughout the relevant pH range (pH 6.5-8.8) where the reaction equilibrium constant exhibited pronounced pH-dependency (Fig. 1b). Sugar donor activation by a substituted phosphate-leaving group is a common feature of glycosyltransferases.[7] Therefore, directional preference for deglycosylation at low pH might be expected as a rather general feature of glycosyltransferase reactions, except in cases where the reactive group of the acceptor substrate matches the low pK_a of the sugar donor's (secondary) phosphate group. The vast majority of hydroxyl acceptor groups have comparably high pK_a and are therefore only deprotonated at high pH. Only a few studies have been performed in which pH effects on glycosyltransferase reaction equilibrium have been examined in detail. However, common trend is that reverse glycosylation is preferred at low pH (equilibrium constants and reaction rates).[8-11] Net proton release during forward glycosyl transfer from NDP-sugar was also exploited for glycosyltransferase assay development based on pH-shift.[12-13]
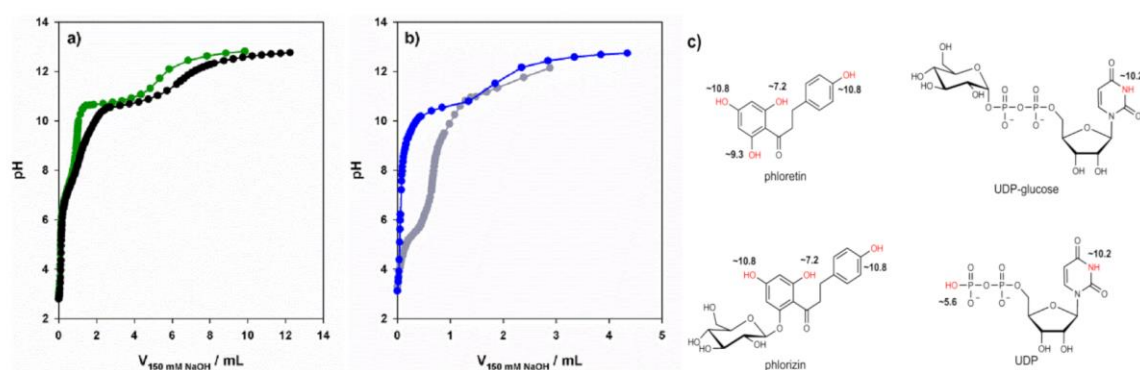

Fig. S5 Potentiometric titrations of substrates and products of *Pc*OGT-catalysed *O*-glycosylation, using 150 mM NaOH at 30°C (7.5 mM analyte in 20% DMSO, 13 mM MgCl$_2$, 50 mM KCl). a) Titration of 20 mL phlorizin (green) and phloretin (black); b) Titration of 10 mL UDP (blue) and UDP-glucose (grey); c) Structures of substrates (phloretin, UDP-glucose) and products (phlorizin, UDP) of the glycosylation reaction catalysed by *Pc*OGT. Approximate pK_a values are indicated at the respective positions.

8

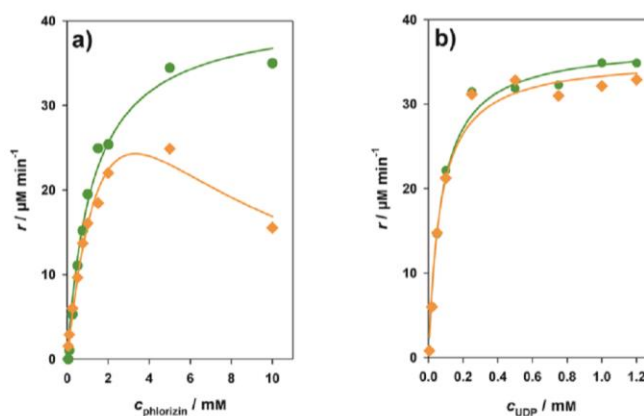## 2.7 Dependency of rearrangement rate on substrate concentrations

Fig. S6 Dependencies of phlorizin consumption (r_P; green) and nothofagin formation (r_N; orange) rates on the substrate concentration used. The constant substrate concentrations were 2 mM UDP (a) and 5 mM phlorizin (b). All reactions contained 100 mU mL^{-1} *Pc*OGT and 50 mU mL^{-1} *Os*CGT. Symbols show measurements and lines are fits to the data.

2.8 Optimisation of *O*- to *C*-glucoside rearrangement by *Os*CGT_I121D

Feasibility of *O*- to *C*-glucoside rearrangement by a single dual-specific *O*- and *C*-glycosyltransferase was evaluated with the promiscuous *Os*CGT mutant I121D. Initially different concentrations of UDP (0.5 and 5 mM) and phlorizin (0.05, 0.5 and 5 mM) were tested (1.25 mU mL^{-1} *Os*CGT_I121D at pH 7). Time courses for 24 h were recorded and a summary of conversions after 6h reflecting the overall trend is depicted in Fig. S7. An increase of UDP concentration from 0.5 to 5 mM as well as rising phlorizin to 5 mM enhanced the formation of nothofagin (Fig. S7a). While presenting a proof of principle, a maximum yield of around 30 µM was by far too low for synthetic use. In all cases more than 90% of the starting material was still present as phlorizin or phloretin (Fig. S7b-c) and only small amounts of 4'-*O*- and 3'-*C*-glycoside (nothofagin) were found. Highest conversion to nothofagin (4.6 %) was obtained when 0.5 mM phlorizin and 5 mM UDP were applied which served as a starting point for optimizations.
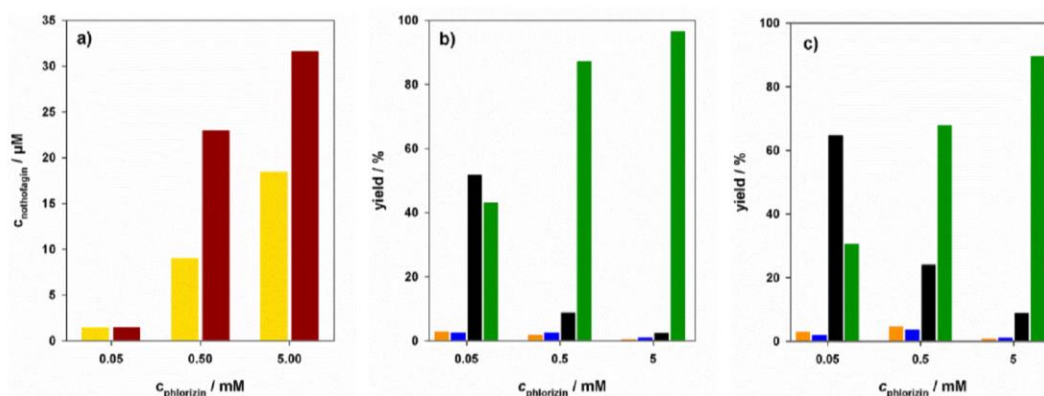

Fig. S7 Yields after 6 h of *O*- to *C*-glycoside rearrangement catalysed by *Os*CGT_I121D (1.25 mU mL^{-1}, pH 7.0) using various substrate concentrations. a) Nothofagin formed from different phlorizin concentrations using 0.5 (yellow) and 5 mM (red) UDP, respectively; Distribution of phloretin and glycosides in presence of 0.5 mM (b) and 5 mM (c) UDP: phlorizin (green), phloretin (black), nothofagin (orange), 4'-*O*-glucoside (blue).

9

Optimisation of the *O*- to *C*-rearrangement with wild type glycosyltransferases revealed that well balanced *O*- and *C*-glycosyltransferase activities and a suitable equilibrium constant of *O*-glycosylation are key for efficient conversions. Besides substrate concentrations, pH is the key parameter to influence the ratio of these enzymatic activities. Therefore, conversion of 0.5 mM phlorizin in presence of 5 mM UDP was evaluated at pH 6.0, 7.0 and 8.0. The effect of pH on yields after 6 h of incubation (Fig. S8) was in general comparable to results with wild type enzymes (Fig. 2a). At low pH formation of the aglycon was favoured but *C*-glycosylation was impaired. Best yield of *C*-glycoside as well as a favourable ratio of 3'-*C*- to 4'-*O*-glycoside were obtained at pH 7.0.
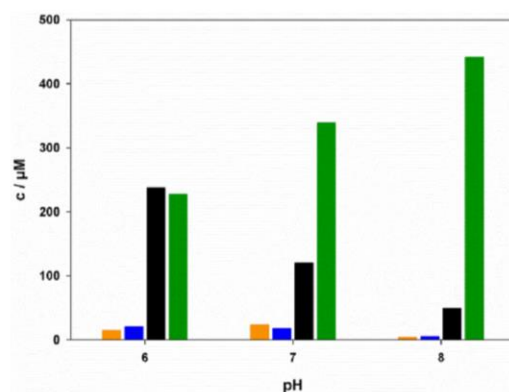

Fig. S8 Distribution of phloretin (glycosides) after 6 h of rearrangement by *Os*CGT_I121D (1.25 mU mL^{-1}) at various pH (50 mM BisTris, 0.5 mM phlorizin, 5 mM UDP): phlorizin (green), phloretin (black), nothofagin (orange), 4'-*O*-glucoside (blue).

Since accumulation of the aglycon phloretin was omnipresent we hypothesized that availability of UDP-glucose could limit *C*-glycosylation. Therefore addition of UDP-glucose (0, 0.5, 1 and 5 mM) was tested. Addition of UDP-glucose should boost final *C*-glycosylation but disfavour initial phloretin production. To analyse potential improvement on *C*-glycoside formation, we focussed on the final yields after 43 h (Fig. S9) rather than on effects on the initial rate. The expected negative effect of UDP-glucose on deglycosylation of phlorizin was detrimental using 5 mM UDP-glucose. At low concentrations (0.5, 1 mM), however, the positive effect of UDP-glucose on *C*-glycosylation prevailed. Although the highest nothofagin yield was obtained in presence of 0.5 mM UDP-glucose we decided to continue working with 1 mM UDP-glucose. We anticipated it to be beneficial for maximizing nothofagin formation since phloretin concentrations (~70 μM) did not seem to be limiting. By doubling the amount of enzyme and increasing the reaction time to 70 h it was finally possible to obtain 70% yield (0.35 mM) in *O*- to *C*-glycoside rearrangement catalysed by *Os*CGT_I121D.
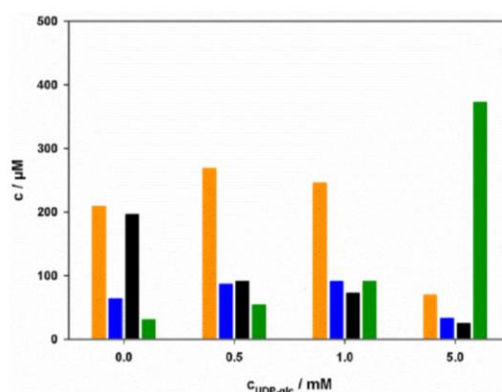

Fig. S9 Distribution of phloretin (glycosides) after 43 h of rearrangement by *Os*CGT_I121D (1.25 mU mL^{-1}) in presence of different UDP-glucose concentrations (0.5 mM phlorizin, 5 mM UDP): phlorizin (green), phloretin (black), nothofagin (orange), 4'-*O*-glucoside (blue).

10

3. References

1. M. Brazier-Hicks, K. M. Evans, M. C. Gershater, H. Puschmann, P. G. Steel, R. Edwards, *J. Biol. Chem.*, 2009, **284**, 17926-17934.
2. D. P. Dixon, A. Lapthorn, P. Madesis, E. A. Mudd, A. Day, R. Edwards, *J. Biol. Chem.*, 2008, **283**, 20268-20276.
3. A. Gutmann, B. Nidetzky, *Angew. Chem. Int. Ed.*, 2012, **51**, 12879-12883.
4. C. Gosch, H. Halbwirth, B. Schneider, D. Holscher, K. Stich, *Plant Sci.*, 2010, **178**, 299-306.
5. C. Valenta, J. Cladera, P. O'Shea, J.Hadgraft, *J. Pharm. Sci.*, 2001, **90**, 485-492
6. R. M. C. Dawson, D. C. Elliot, W. H. Elliot, K. M. Jones, *Data for Biochemical Research*, Clarendon Press, Oxford, 3rd edn., 1989, ch. 5, pp. 103-114 (and references therein).
7. L. L. Lairson, B. Henrissat, G. J. Davies and S. G. Withers, *Annu. Rev. Biochem.*, 2008, **77**, 521-555.
8. L. Elling, M. R. Kula, *Enzyme Microb. Technol.*, 1995, **17**, 929-934.
9. C. Zhang, E. Bitto, R. D. Goff, S. Singh, C. A. Bingman, B. R. Griffith, C. Albermann, G. N. Phillips, J. S. Thorson, *Chem. Biol.*, 2008, **15**, 842-853.
10. K. D. Miller, V. Guyon, J. N. S. Evans, W. A. Shuttleworth, L. P. Taylor, *J. Biol. Chem.*, 1999, **274**, 34011-34019.
11. E. V. Chandrasekaran, J. Xue, J. Xia, R. D. Locke, K. L. Matta, S. Neelamegham, *Biochemistry*, 2008, **47**, 320-330.
12. C. Deng, R. R. Chen, *Anal. Biochem.*, 2004, **330**, 219-226.
13. M. Persson, M. M. Palcic, *Anal. Biochem.*, 2008, **378**, 1-7.

11

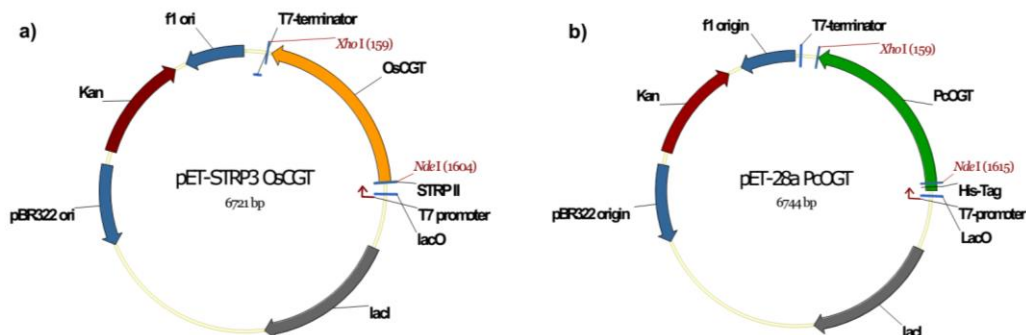## 4. Appendix

4.1 Maps of expression vectors

Fig. S10 Map of vectors for *Os*CGT (pET-STRP3) (a) and *Pc*OGT (pET-28a) (b) expressions. Genes are inserted into *Nde*I and *Xho*I sites for expression with N-terminally fused *Strep*-tag II and His-tag, respectively; f1 ori: f1 phage origin of replication, Kan: kanamycin resistance gene, pBR322 ori: pBR322 origin of replication, lacI: lactose repressor gene, lacO: lactose operator, STRP II: *Strep*-tag II, OsCGT: *Os*CGT gene (GenBank: FM179712), PcOGT: *Pc*OGT gene (GenBank: FJ854496)

4.2 DNA sequence of *Os*CGT in pET-Strep3 expression vector

Bold: confirmed by sequencing
CAPITAL LETTERS: Expressed open reading frame (*Os*CGT + *Strep*-tag II)
Grey: *Strep*-tag II

>pET-STRP3_*Os*CGT
atccggatatagttcctcctttcagcaaaaaaccccctcaagacccgtttagaggccccaagggggttatgctagttattgctcagcggtggcagc
agccaactcagcttcctttcgggctttgttagcagccggatctcagtggtggtggtggtggtgctcgagccaTATCGGTGGCCTGCAGGTCGAG
TCAATTAGTGCGACATGTTCCCCCTTGGCACAGCCGCGCGAACTCGGCCAGACAACGGTGGCTCGATCCACCGCCGGCGACGGCCTTCGCGGCG
GCCTCGGCGAGGCTCGCCGCCTTCATCCGCAACGCCTCGTCCGCCATCGCCGCCTTCACCTTCTCCGATATCTCCTCCGCGCCGATCACCCCGG
CCTCCCCCTCCCAGCTCCAGGTGTCCGCCCACACGCCGAGCCCGGCGCGCGCCACCACGCCGGAGTTCACCCGCTGGTCGCCGAACCTCGGCAG
CGCCAGCACCGGGACGCCGCTCGCCGCCGCCTCCGTCACCGAGTTCCAGCCGCAGTGGCTCACGAACAGCGCCACGGACTCGTGCTTCAGGACC
TCTTCCTGATCCACCCATGCCTTGGTGACGAGGCCTCGCTTCTCCACCCGCTCCAAGAACCCCTCGTCGAGCAGCTCGCCGAGCTCGGCGGCGT
CGTCCCTGTCCACGACGGTGCTCTTCACCACCCACAGGAACCGGTGGCCGCTGCCCTCCAGCCCGGCGGCGAGTTCCCTGAGCTGCTCCCTTGA
GATGGCCTTCCGGCTGCCGAAGCTCACGTACACCACCGACCGCGCCGGCTGCGCGTCGAGCCACTCCATGTAGTTCGCCTGCGGATCTTTCGCC
TGGTTGCTCGCCGGGAGAAGTGGCCCCACGGCGAACACCGGCGGGAAGCCGGAGGCGACCTTGCCCTGCTGCAGGGCCGCGACGGCCTCCGGCT
CCAAGGCGTCGAACGTGTTGACGAGGATGCCGGCGGCGCTCGTGAGGCTCCGGCCGTTGGCGACGAACTGGCGGGTGAAGAGGTGGTTGGGGTC
GTGCAGCGCCTGCGGGATGGAGGCCTTGGGGATGCGGTACACGCCGGGGATGTCGACGTCGCCGACGCCGCCGCCGCCGCCAGCGTTGGCGTCG
AGGTATGTGGGGAAGTAGGCGCAGAGGGAGAGCATCGCGGCGGAGGCGGTGAAGAGGATGTGGCACGGGAGGCCCTGCTCCTTCGCCACGGGTA
TGACGACGGATGTCAGCGCGATGTCCGTGGCGAGCGCCGACGCGCCGGCGCCGGTGAGGAGCGGGCCGAGGAGCGGCGCCGACCGCCGCATGCG
CTCGAACCGGAGGAAGAACGGGTCGGCGCCGGGGAACTCGGACGCGTCGAACGGCGCGAGCTCGAAGTCGAGCCGCCGCACCGCCGGGAACGCG
TCGAACAGCGCGTCGAGGTGCTTCGACTCCGCGGTGGACACGGTGGGGAGCACCGTGACGAGGGAGACGTCGCAGCCGTGGCCGGAGGACAGCG
CCACGGCGAGGCGGCCGAAGGGGACGAGGTGGCCCATGCCGGCGCTCGGGATGAGCACCACATGCGGCCGCCTGCCCGCCGCGTCGCCAGAGCT
CGGCATATGGTTAATTAAGCCTTTCTCGAACTGCGGGTGGCTCCAGCTAGCCATggtatatctccttcttaaagttaaacaaaattatttctag
aggggaattgttatccgctcacaattcccctatagtgagtcgtattaatttcgcgggatcgagatctcgatcctctacgccggacgcatcgtgg
ccggcatcaccggcgccacaggtgcggttgctggcgcctatatcgccgacatcaccgatggggaagatcgggctcgccacttcgggctcatgag
cgcttgtttcggcgtgggtatggtggcaggccccgtggccggggggactgttgggcgccatctccttgcatgcaccattccttgcggcggcggtg
ctcaacggcctcaacctactactgggctgcttcctaatgcaggagtcgcataagggagagcgtcgagatcccggacaccatcgaatggcgcaaa
acctttcgcggtatggcatgatagcgcccggaagagagtcaattcagggtggtgaatgtgaaaccagtaacgttatacgatgtcgcagagtatg
ccggtgtctcttatcagaccgtttcccgcgtggtgaaccaggccagccacgtttctgcgaaaacgcgggaaaaagtggaagcggcgatggcgga
gctgaattacattcccaaccgcgtggcacaacaactggcgggcaaacagtcgttgctgattggcgttgccacctccagtctggccctgcacgcg
ccgtcgcaaattgtcgcggcgattaaatctcgcgccgatcaactgggtgccagcgtggtggtgtcgatggtagaacgaagcggcgtcgaagcct

12

**gtaaagcggcggtgcacaatcttctcgcgcaacgcgtcagtgggctgatcattaactatccgctggatgaccaggatgccattgctgtggaagc
tgcctgcactaatgttcc**ggcgttatttcttgatgtctctgaccagacacccatcaacagtattattttctcccatgaagacggtacgcgactg
ggcgtggagcatctggtcgcattgggtcaccagcaaatcgcgctgttagcgggcccattaagttctgtctcggcgcgtctgcgtctggctggct
ggcataaatatctcactcgcaatcaaattcagccgatagcggaacgggaaggcgactggagtgccatgtccggttttcaacaaaccatgcaaat
gctgaatgagggcatcgttcccactgcgatgctggttgccaacgatcagatggcgctgggcgcaatgcgcgccattaccgagtccgggctgcgc
gttggtgcggatatctcggtagtgccgggatacgacgataccgaagacagctcatgttatatcccgccgttaaccaccatcaaacaggattttcgcc
tgctggggcaaaccagcgtggaccgcttgctgcaactctctcagggccaggcggtgaaggcaatcagctgttgcccgtctcactggtgaaaag
aaaaaccaccctggcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaa
agcgggcagtgagcgcaacgcaattaatgtaagttagctcactcattaggcaccgggatctcgaccgatgcccttgagagccttcaacccagtc
agctccttccggtgggcgcggggcatgactatcgtcgccgcacttatgactgtcttctttatcatgcaactcgtaggacaggtgccggcagcgc
tctgggtcattttcggcgaggaccgctttcgctggagcgcgacgatgatcggcctgtcgcttgcggtattcggaatcttgcacgccctcgctca
agccttcgtcactggtcccgccaccaaacgtttcggcgagaagcaggccattatcgccggcatggcggccccacgggtgcgcatgatcgtgctc
ctgtcgttgaggacccggctaggctggcgggggttgccttactggttagcagaatgaatcaccgatacgcgagcgaacgtgaagcgactgctgct
gcaaaacgtctgcgacctgagcaacaacatgaatggtcttcggtttccgtgtttcgtaaagtctggaaacgcggaagtcagcgccctgcaccat
tatgttccggatctgcatcgcaggatgctgctggctaccctgtggaacacctacatctgtattaacgaagcgctggcattgaccctgagtgatt
tttctctggtcccgccgcatccataccgccagttgtttaccctcacaacgttccagtaaccgggcatgttcatcatcagtaacccgtatcgtga
gcatcctctctcgtttcatcggtatcattacccccatgaacagaaatccccccttacacggaggcatcagtgaccaaacaggaaaaaaccgccct
taacatggcccgctttatcagaagccagacattaacgcttctggagaaactcaacgagctggacgcggatgaacaggcagacatctgtgaatcg
cttcacgaccacgctgatgagctttaccgcagctgcctcgcgcgtttcggtgatgacggtgaaaaacctctgacacatgcagctcccggagacgg
tcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcagggcgcgtcagcgggtgttggcgggtgtcggggcgcagccatgaccca
gtcacgtagcgatagcggagtgtatactggcttaactatcgcagcagcagattgtactgagagtgcaccatatatgcggtgtgaaatacc
gcacagatgcgtaaggagaaaataccgcatcaggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcg
gtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggcca
ggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcg
aaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctg
tccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggct
gtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccact
ggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactaga
aggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggta
gcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatctttctacggggtctgacgctca
gtggaacgaaaactcacgttaagggattttggtcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgttatgagccatat
tcaacgggaaacgtcttgctctaggccgcgattaaattccaacatggatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaa
tcaggtgcgacaatctatcgattgtatgggaagcccgatgcgccagagttgtttctgaaacatggcaaaggtagcgttgccaatgatgttacag
atgagatggtcagactaaactggctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgatgcatggttactcac
cactgcgatccccgggaaaacagcattccaggtattagaagaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgc
cggttgcattcgattcctgtttgtaattgtcctttaacagcgatcgcgtatttcgtctcgctcaggcgcaatcacgaatgaataacggtttgg
ttgatgcgagtgattttgatgacgagcgtaatggctggcctgttgaacaagtctggaaa**gaaatgcataaacttttgccattctcaccggattc
agtcgtcactcatggtgatttctcacttgataaccttatttttgacgaggggaaattaataggttgtattgatgttggacgagtcggaatcgca
gaccgataccaggatcttgccatcctatggaactgcctcggtgagtttttctccttcattacagaaacggcttttttcaaaaatatggtattgata
atcctgatatgaataaattgcagtttcatttgatgctcgatgagtttttctaagaattaattcatgagcggatacatatttgaatgtatttaga
aaaataaacaaataggggggttccgcgcacatttccccgaaaagtgccacctaaattgtaagcgttaatattttgttaaaaattcgcgttaaatttt
tgttaaatcagctcatttttttaaccaataggccgaaatcggcaaaatccctttataaatcaaaagaatagaccgagatagggttgagtgttgttc
cagtttggaacaagagtccactattaaagaacgtggactccaacgtcaaagggcgaaaaaccgtctatcagggcgatggcccactacgtgaacc
atcaccctaatcaagttttttgggggtcgaggtgccgtaaagcactaaatcggaaccctaaagggagccccgatttagagcttgacggggaaag
ccggcgaacgtggcgagaaaggaagggaagaaagcgaaaggagcgggcgctagggcgctggcaagtgtagcggtcacgctgcgcgtaaccacca
cacccgccgcgcttaatgcgccgctacagggcgcgtcccattcgcca**

4.3 DNA sequence of *Pc*OGT

>PcOGT (internal NdeI and XhoI restriction sites removed from GenBank: FJ854496)
atgggagacgtcattgtactgtacgcatctccagggatggggcacatcgtcgccatggtggagctgggcaagttcattgtccaccgctacggcc
cccacaaattctccatcaccattctctacacctgcgggcagcattgtcgacaccgctagcatccccgtctacatccgccgcatctcccactccca
ccctttcatttccttccgccaattccctcgcgtcaccaataatattacccgaaacataagcgtccccgcaatcacgttcgacttcatccgccag
aacgatcctcatgtccgcagtgccctccaagaaatctctaaatccgccaccgttcgcgccttcatcatcgacctcttctgcacctccgctcttc
ccatcgggaaggaattcaacatcccaacatactacttccacacttctggtgccgcagttcttgctgctttttttgtatttgcccaagatcgatga
gcaaaccaaaaccaccgagagtttcaaagacctccgcgacaccgttttcgaattccccgatggaagtctcctctgaaggctacacacatggtc
caactggtgctcgaccggaacgaccctgcttattcggacatgatctattctgctcacatcttcccaaatccaacggaatcatcgtcaacacgt
tcgaagagctcgagccacctagcgtcctccaggccattgctggaggcctgtgtgttcctgatgggccaactccgcccgtgtactacgttggtcc
attgattgaggaagagaaagaattgagtaaggatgcagatcggccgagaaggaggactgcttgtcatggctcgataagcagccaagtcgaagc
gtgctgtttctctgtttcggaagcatgggatcatttccggctgctcaactgaaggagatagcgaacgggttggaggcgagcgggcagaggttcc
tgtgggtggtgaagaagccgccggttgaagagaaatcaaagcaggtccatggagttgacgactttgatttgaagggtgtgttgccagaagggtt
tttggagaggacggacgagcaggggggatggtagtgaagtcatgggcgccgcaggtggtggtgttgaagaaggagtcggttggtgggttcgtgaca
cattgcggatgaactcggtactggaagcagtggttgcggggtgccgatgattgcttggccgctttacgcggagcagcacatgaacaggaatg
ttctagtgacggacatggaaatcgcgatcggggtggagcagagacgaggaaggtgggttcgtgagcggggaagaagtggagaggagagtgag
ggagttgatggagtcggaaggaggaagagcgcttagagagaggtgcaagaaacttgggggagatggcttcggctgctttgggagagaccggttcg
tccaccagaaacttggtcaactttgttagtagcattaca

Towards green synthesis of glycosylated dihydrochalcone natural products using glycosyltransferase-catalysed cascade reactions

Towards green synthesis of glycosylated dihydrochalcone natural products using glycosyltransferase-catalysed cascade reactions

Alexander Gutmann,[a] Linda Bungaruang,[a] Hansjoerg Weber,[b] Mario Leypold,[b] Rolf Breinbauer[b] and Bernd Nidetzky*[a]

Regioselective O-β-D-glucosylation of flavonoid core structures is used in plants to create diverse natural products. Their prospective application as functional food and pharmaceutical ingredients makes flavonoid glucosides interesting targets for chemical synthesis, but selective instalment of a glucosyl group requires elaborate synthetic procedures. We report glycosyltransferase-catalysed cascade reactions for single-step highly efficient O-β-D-glucosylation of two major dihydrochalcones (phloretin, davidigenin) and demonstrate their use for preparation of phlorizin (phloretin 2'-O-β-D-glucoside) and two first-time synthesised natural products, davidioside and confusoside, obtained through selective 2'- and 4'-O-β-D-glucosylation of the dihydroxyphenyl moiety in davidigenin, respectively. Parallel biocatalytic cascades were established by coupling uridine 5'-diphosphate (UDP)-glucose dependent synthetic glucosylations catalysed by herein identified dedicated O-glycosyltransferases (OGT) to UDP dependent conversion of sucrose by sucrose synthase (SuSy; from soybean). The SuSy reaction served not only to regenerate the UDP-glucose donor substrate for OGT (up to 9 times), but also to overcome thermodynamic restrictions on dihydrochalcone β-D-glucoside formation (up to 20% conversion and yield enhancement). Using conditions optimised for overall coupled enzyme activity, target 2'-O- or 4'-O-β-D-glucoside were obtained in ≥ 88% yield from reactions consisting of 5 mM dihydrochalcone acceptor, 100 mM sucrose, and 0.5 mM UDP. Davidioside and confusoside were isolated and their proposed chemical structures confirmed by NMR. OGT-SuSy cascade transformations present a green chemistry approach for efficient glucosylation in natural products synthesis.
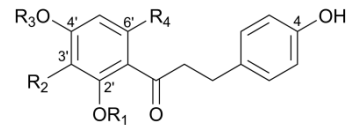
Introduction

Flavonoids are a large and structurally diverse group of natural polyphenols.[1, 2] They are widely distributed in edible plants and therefore constitute an important part of the human diet.[3, 4] Some flavonoids are strong anti-oxidants and thus serve as powerful inhibitors of lipid peroxidation.[5, 6] Dietary intake of flavonoids has been correlated with reduced risk of chronic diseases, coronary heart disease in particular.[7] Dihydrochalcones present a major sub-class of flavonoids and are characterized structurally by two phenolic rings connected through a flexible open-chain three-carbon linker (Table 1).[1] Many dihydrochalcones, such as the common phloretin (**1**), which is abundantly present in the leaves and peel of apples,[8, 9] exhibit a wide spectrum of interesting and pharmacologically relevant bioactivities.[10] Next to a general anti-oxidative property,[11] phloretin shows antithrombotic[12] and hepatoprotective properties,[13] potential suppression of metabolic carcinogen activation[14] and effects intracellular drug accumulation.[15] Davidigenin (**2**), the 4'-deoxy analogue of **1**, is broadly distributed among different plant families.[16] As an antispasmodic compound present in or derived from traditional medicines it is applied to the treatment of intestinal disorders and asthma.[16] **2** exhibits weak antibacterial activity[17] and by inhibiting aldose reductase ALR2 it shows also antidiabetic activity.[18]

Glycosylation is an important biological mechanism for the structural and functional diversification of natural flavonoids.[19, 20] Generally, the attachment of sugar residue(s) to a flavonoid core increases compound bioavailability due to strong water solubility enhancement.[21, 22] Furthermore, physiological and pharmacological properties are often beneficially altered as consequence of glycosylation.[23, 24] Taste properties (e.g. sweetness, bitterness) and colour can also be modified by glycosylation.[25, 26] Interestingly, compared to other flavonoid sub-classes that show broadly diverging and often complex glycosylation patterns, glycosylated dihydrochalcones occur primarily as mono-β-D-glucopyranosides.[1] Sugar attachment to dihydrochalcones often results from a phenolic O-glycosidic bond, although aromatic C-glucosylation is also found in certain compounds.[1, 27] Phloretin is naturally present mainly as the 2'-O-β-D-glucoside, commonly referred to as phlorizin (**3**).[28] Due to inhibition of cellular glucose transport **3** was considered for type II diabetes treatment.[28] The corresponding 4'-O-β-D-glucoside, which is called trilobatin (**4**), is less common.[29] Davidioside (**5**) is the 2'-O-

β-D-glucoside of **2**, while the 4'-O-β-D-glucoside is called confusoside (**6**). Compounds **5** and **6** were previously isolated from plant extracts and characterized structurally.[30, 31] However, with exception of a low yielding total synthesis of **6** preparation of glucosides of **2** was not reported.[32]

Table 1 Major dihydrochalcones and their glucosides

dihydrochalcone	R_1	R_2	R_3	R_4
phloretin (**1**)	H	H	H	OH
davidigenin (**2**)	H	H	H	H
phlorizin (**3**)	glucose	H	H	OH
trilobatin (**4**)	H	H	glucose	OH
davidioside (**5**)	glucose	H	H	H
confusoside (**6**)	H	H	glucose	H
nothofagin (**7**)	H	glucose	H	OH

Due to their various prospective applications as fine chemicals, pharmaceutical ingredients and food additives, dihydrochalcone glucosides **3 – 6** are interesting targets for large-scale preparation.[28, 33, 34] However, their isolation from natural material is complicated by occurrence in complex multi-component mixtures and at relatively low abundance.[30, 31, 35, 36] Selective glucosylation of **1** and **2** would provide convenient access to **3 – 6** through bottom-up synthesis. Process chemistry for this purpose is not well developed, though, and elaborate procedures involving a substantial amount of protecting group chemistry are required for precise instalment of the glucosyl group.[37] Enzymes have earned green credentials for being able to cope with issues of reactivity and selectivity in glycosylation reactions.[38, 39] However, distinction between the different phenolic hydroxyls in dihydrochalcone acceptors presents a notable challenge even for biocatalysis, and practical enzyme catalysts for selective glucosylation of **1** or **2** are to be established. Evidence from different studies of flavonoid glycosylation has shown that only glucosyltransferases (GT), in contrast to alternatively used transglycosidases, exhibit the requisite degree of regioselectivity.[40] GTs utilise an activated donor substrate, typically a nucleoside diphosphate (NDP)-sugar, for transfer of a glycosyl residue onto certain position(s) of the acceptor molecule.[41]

Synthesis of glycosylated dihydrochalcones by glycosyltransferase-cascade reactions

Up to now only one GT from apple (*Malus* x *domestica*) and its ortholog from pear (*Pyrus communis*) (99% sequence identity) were reported to selectively produce analytical amounts of phlorizin (**3**) by glucosylation of **1**.[35, 42] Other recently reported glucosylations of **1** by plant GTs yielded mixtures of monoglucosides.[35, 43] An even more promiscuous bacterial GT glucosylated **1** at 2'-, 4'- and 4-OH, producing a complex mixture of glucosidic compounds including two di- and a triglucoside.[44] Therefore, this emphasises that selectivity-based GT enzyme selection is of key importance for successful development of a biocatalytic synthesis process. Furthermore, dedicated reaction engineering is necessary to solve the problem of cost-effective supply of the NDP-sugar substrate. [45]
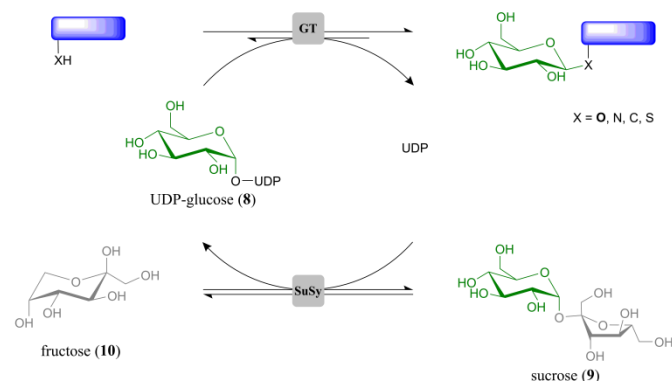
Scheme 1 Formation of natural product glucosides: Coupling of GTs with SuSy in one pot enables cost efficient *in situ* formation of **8** from **9** and catalytic amounts of UDP with concomitant obviation of end-product inhibition by UDP.

Herein we present identification of GT biocatalysts for selective transformation of **1** into **3**, and of **2** into **5** or **6**. A parallel GT cascade reaction (Scheme 1) was established to enable dihydrochalcone glucosylation from sucrose as a highly expedient and inexpensive glucosyl donor substrate. The overall bi-enzymatic transformations were carried out in the presence of catalytic amounts of uridine 5'-diphosphate (UDP) to generate and constantly recycle *in situ* the UDP-glucose (**8**) utilised for glucoside synthesis. Applying reaction conditions optimised for enzyme activity and designed to overcome thermodynamic limitations, target O-β-D-glucoside was obtained as single transfer product in ≥ 88% yield and could be isolated readily. We propose Scheme 1 to be a generally applicable green chemistry approach for efficient glucosylation in natural products synthesis.

Results and discussion

Identification of GT enzymes catalysing regioselective glucosylation of phloretin and davidigenin

A GT from pear (*Pyrus communis; Pc*OGT) was recently reported to be highly selective for transferring the glucosyl moiety of UDP-glucose (**8**) to the 2'-OH of **1**, producing **3**.[35, 46] However, evidences were only from thin layer chromatography[35] or from initial-rate experiments carried out at low conversion (≤ 25%) of both donor and acceptor substrate.[46] Furthermore, the concentrations of **1** used (300 and 100 μM, respectively) were too low to be of synthetic relevance. A number of studies show that GT regioselectivity may change dramatically depending on completeness of conversion or on initial substrate concentration.[46, 47] Therefore, this necessitated rigorous evaluation of *Pc*OGT as catalyst for synthesis of **3** at elevated substrate concentration.

Using recombinant *Pc*OGT (0.1 U mL^{-1}) purified from an *Escherichia coli* overexpression culture (ESI†, Fig. S1),

transformation of **8** and **1** (5 mM each, pH 6.5) into phloretin glucoside(s) was examined in dependence of reaction progress until attainment of apparent equilibrium (~70% conversion of **1**). Product analysis was done using a reversed-phase HPLC method in which the monoglucosidic regio-isomers **3** and **4** were baseline separated (Fig. 1a). Only **3** and no **4** was detected at all times (Fig. 1b). Moreover, close balance between **1** consumed and **3** formed was consistent with enzymatic reaction exclusively at the acceptor's 2'-OH.

*Pc*OGT was then examined for glucosylation of **2** from **8** (each 5.0 mM, pH 7.5). UV absorbance traces from HPLC analysis of samples taken at different times revealed gradual consumption of **2** with concomitant appearance of a major and a minor signal eluting at positions consistent with glucosidic products (Fig. 1c). According to HPLC peak areas around 80% of **2** were converted to the major glycoside and less than 2% of the by-product were formed when equilibrium was obtained. NMR data recorded from the reaction mixture gave preliminary evidence that the 2'-OH of **2** had been glycosylated, thereby indicating that regioselectivity was maintained upon exchange of acceptor substrate. Specific activities of *Pc*OGT related to glucosyl acceptor substrate were determined from the time courses of consumption of **1** and **2**, and **1** (1.76 U mg^{-1} protein) was an about 500-fold better substrate than **2**.
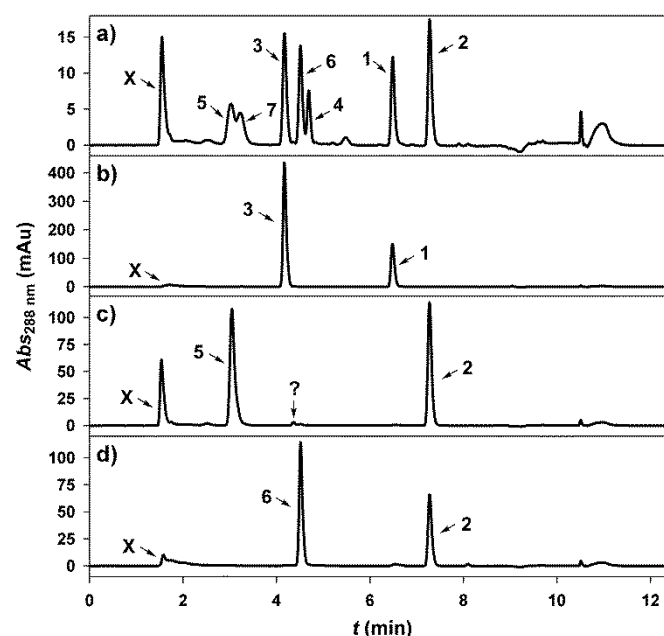

Fig. 1 Reversed-phase C-18 HPLC-analysis of dihydrochalcone (glucosides). (X = unbound compounds including **8** and UDP) (a) Separation of glucosides of **1** (**3**, **4**, **7**) and **2** (**5**, **6**) is achieved.. (b) Conversion of **1** by *Pc*OGT; (c) Conversion of **2** by *Pc*OGT (? = unknown by-product); (d) Conversion of **2** by *Os*CGT I121D

In search of a differently regioselective GT catalysing glucosyl transfer to the 4'-OH of **1** and **2**, we tested a recently described variant of a *C*-glycosyltransferase from rice (*Oryza sativa; Os*CGT) that had its Ile[121] replaced by Asp.[46] Highly purified mutant enzyme was used in all experiments. Wild-type *Os*CGT was shown to catalyse glucosyl transfer from **8** to the aromatic 3'-*C* of **1**, thus producing the *C*-β-D-glucoside nothofagin (**7**).[46, 48] The particular residue substitution in *Os*CGT causes change in reaction selectivity such that *O*-glucosylations of the 2'-OH and 4'-OH occur next to the native *C*-glucosylation of **1**.[46] Using **1** at different concentrations between 2.5 μM and 2.5 mM, it was found that the enzymatic conversion resulted invariably in a mixture of glucosidic products that contained **3** or **7** as the main constituent whereas **4** was present only as a by-product (≤15% of total). It was not possible to enhance the

relative abundance of **4** by running the reaction at different pH in the range 6.2 – 10.5. Ability of the *Os*CGT variant to synthesise substantial amounts of **4** is nevertheless remarkable, and the question of whether isolation of **4** from the enzymatically prepared mixture of phloretin glucosides would still be practical was left for consideration in the future. Notably, though, the *Os*CGT variant was complementarily regioselective to *Pc*OGT in the glucosylation of **2**, as shown in Fig. 1d. The main *Pc*OGT glucosylation product, likely **5**, was completely absent from the reaction sample obtained with the mutated *Os*CGT. However, a single new product peak appeared. Its area increased proportionally to the decreasing peak area of the **2**. Interestingly, the retention time was very similar to that of the trace product formed from **2** by *Pc*OGT, however HPLC analysis of a mixture of both conversions clearly reveals them to be distinct (ESI†, Fig. S2). Based on preliminary NMR data of the reaction mixture it was assumed, and will be confirmed later, that **6** was formed in the reaction of the GT mutant. Specific activity of the *Os*CGT variant for glucosylation of **2** was determined as 22±3 mU mg^{-1}.

In summary, therefore, GT catalysts for selective glucosylation of **1** and **2** have been identified, and they were used further for synthesis of **3**, **5** and **6**.
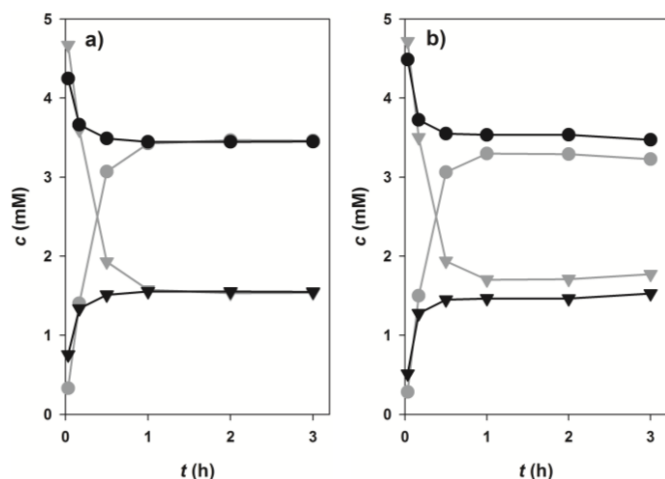

Fig. 2 Phlorizin (**3**) synthesis (grey, 5 mM **1** and **8**) and degradation (black, 5 mM **3** and UDP) by *Pc*OGT (pH 6.5) level out at the same equilibrium concentrations. (a) **1** (triangles), **3** (circles); (b) **8** (triangles), UDP (circles)

Thermodynamic analysis of glucosylation of phloretin (1)

Glucosylation of **1** or **2** from **8** is an equilibrium-controlled process.[49] Thermodynamic constraints on the synthesis of dihydrochalcone β-D-glucosides therefore required clarification. Fig. 2 displays results of time-course analyses for synthesis and degradation of **3**, catalysed by *Pc*OGT at pH 6.5 and 30°C. All compounds present in the reaction were quantified (**1**, **3**, **8**, UDP). Close balance for the proposed overall conversion, **1** + **8** ↔ **3** + UDP, was obtained at each time and in each direction of reaction. Therefore, this indicated absence of any enzyme-catalysed or spontaneous side reactions, such as hydrolysis of **8** for example. Reactions run in forward and reverse direction levelled out at exactly the same end concentrations of **1** and **3** (Fig. 2a), clearly suggesting that the true thermodynamic equilibrium had been attained. Slight differences of final concentrations of UDP and **8** (Fig. 2b) were potentially caused by limited assay accuracy. It was affirmed that addition of fresh enzyme to a reaction mixture at (apparent) equilibrium in Fig. 2 did, as expected, not induce further concentration changes. An equilibrium constant (K_{eq}) of 4.6 (± 0.7) was calculated from the data. Under the conditions

used, therefore, only about 70% of substrate **1** can be converted to product **3**. Development of a glucosyl transfer cascade (Scheme 1) that involved *in situ* generation of the UDP-sugar substrate **8** from sucrose (**9**) and UDP was highly useful to overcome these thermodynamic restrictions.

Glucosyl transfer cascade for efficiency-enhanced glucosylation of phloretin (1)

The general principle of the glucosyl transfer cascade is shown in Scheme 1. Two parallel glucosyl transfer reactions are connected in a one-pot biotransformation via their common UDP-glucose/UDP substrate/product pair. Conversion of (**9**) and UDP is catalysed by sucrose synthase (SuSy) and yields **8** and D-fructose (**10**). The SuSy reaction, which represents the invariant part of the proposed cascade, is flexibly coupled to different GT-catalysed glucosylations of target acceptors under utilisation of **8** as the donor substrate. Benefit of performing synthetic glucosylations through SuSy-GT cascades as compared to single GT reactions is manifested in significant improvement of key parameters of overall transformation efficiency. First of all, cost-effective supply of **8** is accomplished using **9** as a highly expedient and inexpensive donor substrate. Only catalytic amounts of UDP as compared to stoichiometric amounts of **8** are required in the process. By using **9** in suitable excess over the acceptor substrate, thermodynamic restrictions on acceptor glucosylation are overcome effectively (see later). Equilibrium of the SuSy reaction ($K_{eq} \geq 0.5$) favours formation of **8** in a wide pH range (pH ≤ 7.5).[45] Finally, problems of pronounced end-product inhibition by UDP, which have severely restricted direct synthetic use of several flavonoid glycosyltransferases in the past, are brought under control.[50, 51] Continuous removal of UDP due to formation of **8** decreases the steady-state concentration of UDP in the reaction to a value even smaller than the one established from the low (catalytic) amount that was initially added.
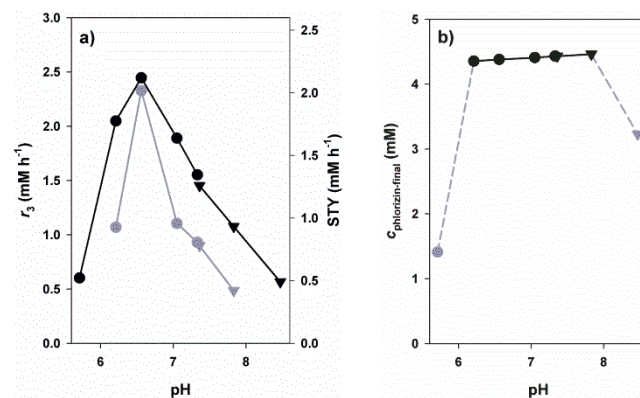

Fig. 3 Influence of pH on efficiency of *Pc*OGT–*Gm*SuSy cascade conversion using BisTris (circles) and TAPS (triangles) buffers (5 mM **1**, 0.5 mM UDP, 100 mM **9**). (a) Initial rate of **3** formation (black), STY for formation of 4 mM **3** (grey). (b) Final equilibrium concentrations of **3** after 24 h (grey: equilibrium was not reached).

To optimise conditions for synthesis of **3**, the coupled reaction of SuSy and *Pc*OGT was studied at different pH values in the range 5.7 − 8.5. Recombinant SuSy from soybean (*Glycine max*; *Gm*SuSy) purified from an *E. coli* expression culture was used. The applied concentrations of **1**, **9** and UDP were 5.0 mM, 100 mM and 0.5 mM, respectively. Despite addition of 20% DMSO as co-solvent, poor aqueous solubility of **1** restricted use of concentrations higher than about 5 mM. The time course of **3**-formation was measured at each pH (ESI†, Fig. S3). The data was used for calculation of initial

production rates of **3** (r_3), final yields at reaction equilibrium (24 h) and space-time yields (STY) for a target product concentration of 4.0 mM (Fig. 3). The optimum pH range for high r_3 was between pH 6.2 and 7.0 (Fig 3a). Except for pH 5.7 and 8.5 where r_3 were small and reaction equilibrium was therefore not attained within the 24-h timespan of the experiment, the final concentrations of **3** were around 4.4 mM, equivalent to a yield of ~90% based on conversion of **1** (Fig 3b). With increasing pH a marginal improvement of equilibrium concentrations of **3** from 4.36 (pH 6.2) to 4.46 (pH 7.8) was observed. However, STY for formation of 4 mM **3** (Fig. 3a) confirmed that pH dependency of initial rates is predominantly determining efficiency of conversions. A clear optimum of STY was found around pH 6.5 with more than 50% loss at pH 6.2 and 7.1, respectively. Overall, compared to synthesis of **3** directly from **1** and **8** (each 5 mM), introduction of the SuSy-GT cascade represented a substantial yield enhancement from ~70 to ~90%. The effect of "thermodynamic push" from **9** on glucoside product formation is noted. It is explained from the net reaction of the glucosyl transfer cascade, which is sucrose (**9**) + phloretin (**1**) ↔ phlorizin (**3**) + D-fructose (**10**).

To further examine the role of **9** in driving the glucosylation of **1**, the cascade reaction was performed at pH 6.5 under conditions where both **9** (100 mM) and **8** (5 mM) were present as donor substrates for glucosylation of **1** (5 mM). Fig. 4 shows the full time course of the conversion along with the corresponding time courses from the coupled reaction without supplementation of **8** (but 0.5 mM UDP) and the direct glucosylation in the absence of *Gm*SuSy. Final conversions, initial rates and space time yields are summarised in Table 2. Product formation in the early reaction phase of cascade reactions benefited somewhat (~25% enhancement of r_3) from the presence of external **8**, probably because the *in situ* produced steady-state concentration of **8** was not sufficient for *Pc*OGT to become fully saturated with glucosyl donor substrate. Use of 5 mM **8** instead of 0.5 mM UDP in SuSy-GT cascade reactions furthermore caused moderate gain in final conversion (~3%) and a more significant improvement of STY for production of 4.0 mM of **3** by about 35%. However, in our opinion a roughly 30% reduced reaction time to obtain a very similar product concentration is not compensating additional costs of replacing UDP with 10-times higher concentrations of the more expensive **8**. Both SuSy-*Pc*OGT cascade reactions clearly outperformed direct glycosylation of **1** by *Pc*OGT. By addition of *Gm*SuSy the final concentration of **3** was increased from 3.61 mM to 4.44 (0.5 mM UDP) and 4.60 mM (5 mM **8**), respectively.
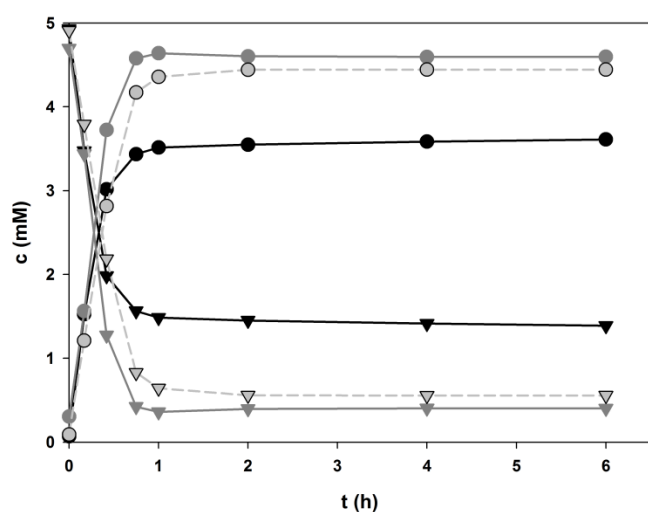

Fig. 4 *Pc*OGT catalysed formation of **3** (circles) from 5 mM **1** (triangles) at pH 6.5: black: without *Gm*SuSy, 5 mM **8**; dark grey: *Gm*SuSy, 5 mM **8**; light grey (dashed): *Gm*SuSy, 0.5 mM UDP

Synthesis of davidoside (5) and confusoside (6) via glucosyl transfer cascade reaction

In order to compensate in some degree the 200-fold decrease in specific activity of *Pc*OGT caused by change of acceptor substrate from **1** to **2**, the pH of the cascade reaction was raised to 7.5 where *Pc*OGT is optimally active and the specific **2** glucosylation rate was therefore enhanced around 2.5-fold as compared to pH 6.5. The Ile[121]→Asp variant of *Os*CGT is also best active at a pH of around 7.5. Loss in specific activity of *Gm*SuSy caused by the rise in pH was negligible in comparison (~35%). The optimum pH of conversion of **9** by *Gm*SuSy was recently determined to be about 6.0.[45]

Table 2 Parameters of direct / GT-cascade catalysed synthesis of **3**, **5** and **6**[a]

		3[b]	**5**[c]	**6**[d]
GT 5 mM **8**	conversion[e] (%)	72.2	79.6	88.3
	$r_{glucoside}$[f] (mM h⁻¹)	6.9	1.3	1.9
	STY[g] (mM h⁻¹)	nd	nd	1.2
GT - *Gm*SuSy 5 mM **8**	conversion[e] (%)	91.9	91.4	95.3
	$r_{glucoside}$[f] (mM h⁻¹)	8.2	1.9	2.2
	STY[g] (mM h⁻¹)	7.6	0.8	1.4
GT - *Gm*SuSy 0.5 mM UDP	conversion[e] (%)	88.9	88.0	91.7
	$r_{glucoside}$[f] (mM h⁻¹)	6.5	1.7	1.3
	STY[g] (mM h⁻¹)	5.6	0.7	1.0

[a] Data extracted from Fig. 4 and 5.
[b] 0.1 U mL⁻¹ *Pc*OGT, (0.1 U mL⁻¹ *Gm*SuSy), pH 6.5, 5 mM **1**
[c] 0.1 U mL⁻¹ *Pc*OGT, (0.1 U mL⁻¹ *Gm*SuSy), pH 7.5, 5 mM **2**
[d] 0.04 U mL⁻¹ *Os*CGT I121D, (0.04 U mL⁻¹ *Gm*SuSy), pH 7.5, 5 mM **2**
[e] At equilibrium after 6 (**3**), 24 (**5**) or 20 (**6**) h conversion
[f] Rate of glucoside formation during initial 25 (**3**) or 60 min (**5**, **6**)
[g] STY for formation of 4 mM glucoside (nd: less than 4 mM product)

Relevant time courses for glucosylation of **2** by *Pc*OGT and *Os*CGT variant are shown in Fig. 5 and summarised in Table 2. The substrate concentrations used were the same as before in the conversions of **1**. The concentration of **2** (5.0 mM) was set according to aqueous solubility in the water-DMSO solvent used. Coupled enzyme reactions carried out in the presence and absence of added **8** are compared to the single enzyme reaction. Confusoside (**6**) was always obtained as single glucosyl transfer product in conversions with *Os*CGT I121D (Fig 5b). However, in *Pc*OGT catalysed conversions besides davidioside (**5**) a second compound, distinct from **6**, accumulated during the first 4 h and diminished again upon prolonged incubation (Fig 5a). We expect it to be the 4-*O*-β-D-glucoside of **2** but failed to confirm this due to very low final concentrations (~0.08 mM, <2 %). Decline of the by-product in favour of **5** formation at longer incubation times could be explained by reversibility of glycosylations. This causes accumulation of the thermodynamically favoured product (here **5**) at equilibrium which we recently exploited for modification of the glucosylation pattern of **1**.[49] More than 50-fold final excess of the main product **5** over the by-product demonstrates that *Pc*OGT can also achieve regiospecific glucosylation of **2**. Comparison of different reaction conditions revealed that initial conversion of **2** proceeded slightly faster and enabled ~3.5% higher conversions in GT-*Gm*SuSy cascade reactions when external **8** was present. Benefit of coupling the synthetic glucosylation to conversion of **9** was again manifested primarily in the final product concentration obtained, which was enhanced by 12% (**5**) and 7% (**6**), respectively in the SuSy-GT cascade reaction as compared to the single GT reference reaction. Product yields in coupled enzyme conversions of **2** were 88% or higher. The STYs of glucosylations of **2** were 5.5 (*Os*CGT variant) and 8 times (*Pc*OGT) lower than those for glucosylation of **1**.
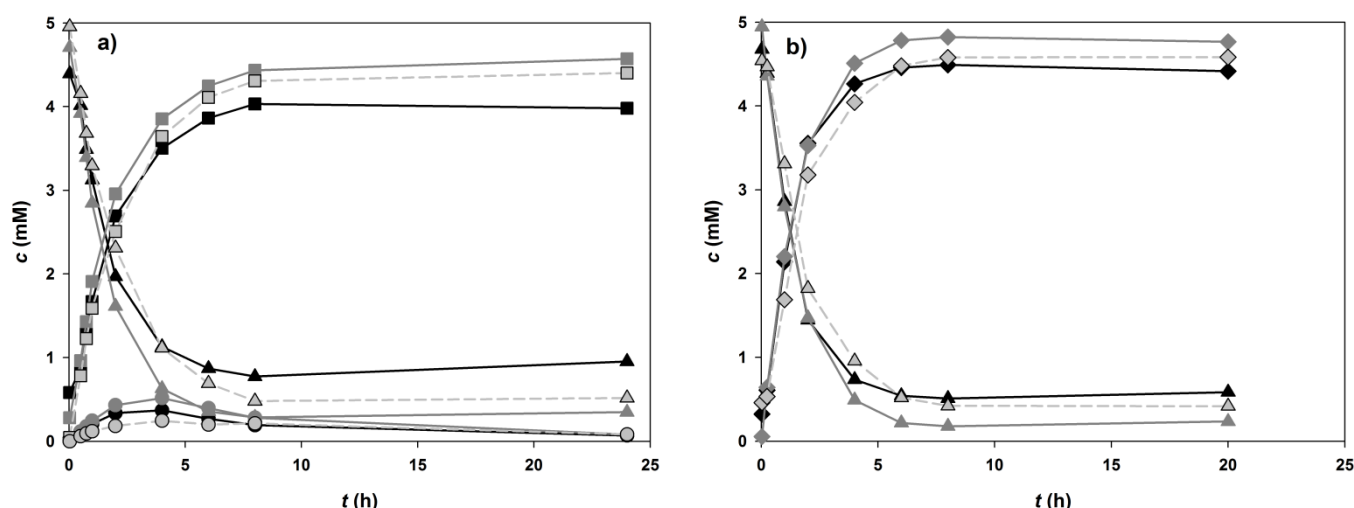
Fig. 5 Glucosylation of 5 mM **2** (triangles) at pH 7.5: black: without *Gm*SuSy, 5 mM **8**; dark grey: *Gm*SuSy, 5 mM **8**; light grey (dashed): *Gm*SuSy, 0.5 mM UDP; (a) *Pc*OGT forms **5** (squares) and an unknown compound (circles). (b) *Os*CGT I121D only produces **6** (diamonds).

Product isolation and NMR spectroscopic characterization

The two products synthesised by glucosylation of 5 mM **2** were isolated from 5 mL reaction mixtures using preparative reversed phase HPLC. **5** and **6** were obtained as white powder after freeze-drying and recovered in ≥ 80% yield. Identities of **5** and **6** were unambiguously assigned from results of a detailed NMR spectroscopic characterization that involved besides ^{1}H and ^{13}C NMR, HMQC (only **5**), COSY and HMBC. Results are summarized in Table S1 and Scheme S1 (ESI†). Although separation was achieved, concentrations of the by-product of **2** glucosylation by *Pc*OGT were too low for NMR analysis. Therefore, davidioside (**5**) and confusoside (**6**) were recovered from glucosylation of **2** by *Pc*OGT and *Os*CGT variant as single compounds at purities >98% based on HPLC peak areas (ESI,† Fig S4).

It follows from the above-mentioned that maximum glucoside product concentration was mainly restricted by aqueous solubility of the comparably hydrophobic dihydrochalcone acceptors. Optimised use of co-solvent,[52] running the reaction in an aqueous-organic two-phase system[53] and feeding the acceptor substrate in accordance to reaction progress[45] are all viable strategies to enhance the effective acceptor concentration in the reaction. However, detailed reaction engineering for targeted "de-bottlenecking" of an enzymatic production was beyond the scope of this study. With the notable exception of highly abundant **3**, no yields above 5 mg g^{-1} dry weight were reported for isolation of **4-6** from plant leaves.[30, 31, 35, 36] Time consuming multi-step procedures and large quantities of organic solvents are required to obtain single flavonoids from the complex mixtures plant materials. Due to simplification of work-up typical yields of the herein described GT-SuSy cascade reactions of around 2 g L^1 (~4.5 mM) seem sufficient to compete with extraction from natural sources. Scarcity in supply of the aglycons **1** and **2** can be eliminated using high yielding single step synthesis by Friedel-Crafts acylation.[54]

Conclusions

Glucosyltransferases catalysing highly regiospecific β-D-glucosyl-transfer from UDP-glucose (**8**) to the 2'-OH of the dihydrochalcones **1** and **2**, and to the 4'-OH of **2** were identified. A two-enzyme one-pot glucosyl transfer cascade for convenient synthesis of the corresponding dihydrochalcone β-D-glucosides (**3**, **5**, **6**) was developed. Coupling of the synthetic

enzymatic glucosyl transfer to conversion of **9** and UDP catalysed by SuSy presents a generally applicable strategy to provide UDP-glucose (**8**) donor substrate in a cost-effective manner and to cope with issues of unfavourable thermodynamics and UDP product inhibition. Green synthesis of flavonoid glucosides is made possible using the approach described herein.

Experimental

Materials

Unless otherwise indicated, all chemicals were from Sigma-Aldrich (Vienna, Austria) in the highest purity available. Phloretin (98%) was from AK Scientific (Union City, CA, US) and phlorizin dihydrate (≥98%), from Carl Roth (Karlsruhe, Germany). *Strep*-Tactin® Sepharose® and desthiobiotin were from IBA (Goettingen, Germany). BCA Protein Assay Kit was from Thermo Scientific (Waltham, MA, US).

Synthesis of davidigenin (2)

Synthesis of **2** was done in a single step from resorcin and phloretic acid using an adopted Friedel-Crafts acylation method.[54] Phloretic acid (6 mmol, 990 mg) and resorcin (6 mmol, 661 mg) were stirred in 4 mL BF$_3$Et$_2$O (32 mmol) at 90°C for 120 min under argon. Complete conversion of resorcin was verified by TLC. The mixture was poured into 200 ml 10% aqueous NaOAc and stirred for 2 h at room temperature. The solution was 3 times extracted with 200 mL EtOAc. Combined EtOAc was washed with 40 mL water and 40 mL brine before drying over MgSO$_4$. The solvent was evaporated under reduced pressure and the residue was chromatographed over silica gel column using cyclohexane-EtOAc mixtures. Product identity was verified by ^{1}H- and ^{13}C-NMR (Bruker AVANCE III 300 spectrometer).

Enzyme production

Escherichia coli BL21-Gold (DE3) expression strains for *Pc*OGT (UGT88F2; GenBank: FJ854496),[46] *Os*CGT (GenBank: FM179712) Ile^{121} to Asp mutant[46] and *Gm*SuSy (GenBank: AF030231)[45] were described elsewhere in detail. Enzyme expression in lysogeny broth

Synthesis of glycosylated dihydrochalcones by glycosyltransferase-cascade reactions

(LB)-medium as N-terminal *Strep*-tag II fusion proteins and enzyme purification by affinity chromatography on *Strep*-Tactin® Sepharose® columns was also reported elsewhere.[45] Expected molecular mass and purity of enzymes were affirmed by SDS polyacrylamide gel electrophoresis (PAGE). Aliquots of final preparations were stored at -70°C and thawed only once prior to their use.

Activity assays

Sucrose cleavage by *Gm*SuSy was measured spectrophotometrically using a discontinuous two-step enzymatic assay. Oxidation of **8**, formed from **9** (100 mM) and UDP (0.5 mM), to UDP-α-D-glucuronic acid is coupled to reduction of two NAD^+ molecules to NADH by human UDP-α-D-glucose 6-dehydrogenase (hUGDH) as described elsewhere in detail.[45] In short, reactions were started by *Gm*SuSy addition and samples of 150 µL were stopped by heating (95°C, 5 min). Precipitated protein was removed by centrifugation (13,200 rpm, 20 min) and 100 µL of the supernatant were mixed with 400 µL of measuring solution (2.5 mM NAD^+, 0.05 % Triton™ X-100 in 100 mM HEPES, pH 8.0) in a Half Micro Cuvette. Absorbance at 340 nm was measured before and after incubation with 1.5 mU of hUGDH and concentration of **8** was calculated by the increase of absorbance ($\varepsilon_{NADH} = 6220$ M^{-1} cm^{-1}). Typically 4 distinct measurements were used to calculate linear initial rates. One Unit of *Gm*SuSy was defined as the amount of enzyme producing 1 µmol **8** (2 µmol NADH) per minute under following conditions: 30°C, pH of cascade reactions (6.5 or 7.5) 0.5 mM UDP, 100 mM **9** in 50 mM BisTris, 13 mM $MgCl_2$, 50 mM KCl, 0.13% BSA and 20% DMSO.

An reversed phase HPLC-based assay for quantification of dihydrochalcones (**1**, **2**) and their glycosides (**3-6**) was used to determine activities of *Pc*OGT and the *Os*CGT variant.[46] In short, conversions of 5 mM dihydrochalcone with 0.5 mM **8** were started by GT addition and conducted under agitation (400 rpm) and temperature control (30°C) using a thermomixer. Typically 4 aliquots of 100 µL were taken within 60 min to determine linear initial rates. Reactions were stopped by mixing with 100 µL acetonitrile. Precipitated protein was removed by centrifugation (13,200 rpm, 20 min) before applying 5-10 µL of supernatant to HPLC analysis. Separation on an Agilent 1200 HPLC equipped with a Chromolith® Performance RP-18e column (100 x 4.6 mm) was performed at 35°C and monitored by UV-detection at 288 nm. Following water (A) to acetonitrile (B) gradient (0.1% trifluoroacetic acid each) was applied: 20-47.5% B (7.5 min, 1 mL min^{-1}), 47.5-100% B (0.05 min, 1 mL min^{-1}), 100% B (1.45 min, 1.5 mL min^{-1}), 100-20% B (0.05 min, 1.5 mL min^{-1}), 20% B (2.45 min, 1.5 mL min^{-1}). One Unit of GT activity was defined as the amount of enzyme glucosylating 1 µmol **1** per min under following conditions: 30°C, 5 mM **1**, 0.5 mM **8** in 50 mM BisTris, pH 6.5 (*Pc*OGT) or 7.5 (*Os*CGT I121D) containing 13 mM $MgCl_2$, 50 mM KCl, 0.13% BSA and 20% DMSO.

Equilibrium of 1 glycosylation (K_{eq}) by *Pc*OGT

K_{eq} of *Pc*OGT was determined by running glucosylation of **1** in forward (5 mM **1**, 5 mM **8**) and reverse direction (5 mM **3**, 5 mM UDP) until no further conversion was observed. Reactions were run at 30°C and started by addition of 100 mU mL^{-1} *Pc*OGT (50 mM BisTris, pH 6.5, 13 mM $MgCl_2$, 50 mM KCl, 0.13% BSA and 20% DMSO). Concentrations of **1** and **3** were determined by the HPLC based GT activity assay.

The same samples (10 µL) were applied for quantification of **8** and UDP using an anion exchange HPLC protocol. An Agilent 1200 HPLC system was used for separation on an Agilent ZORBAX SAX (4.6 x 250 mm) column at 30°C. UDP and 8 were monitored by UV detection at 254 nm. Using 20 and 500 mM potassium phosphate buffer (pH 6.8) as solvent A and B, respectively following gradient was applied at a constant flow rate of 1.5 mL min^{-1}: 0-100% B (7 min), 100% B (2 min), 100-0% B (0.05 min), 0% B (3.95 min)

3, 5 and 6 formation by direct GT and GT-*Gm*SuSy cascade reaction

5 mM **1** and **2** were converted in 50 mM BisTris at pH 6.5 and 7.5, respectively. All reactions contained 13 mM $MgCl_2$, 50 mM KCl, 0.13% BSA and 20% DMSO. As glucose source conversions without *Gm*SuSy contained 5 mM **8** while for those with *Gm*SuSy either 5 mM **8** and 100 mM **9** or 0.5 mM UDP and 100 mM **9** were used. 100 mU mL^{-1} *Pc*OGT or 40 mU mL^{-1} *Os*CGT I121D were applied. In GT-*Gm*SuSy cascade reaction equal activities of both enzymes were added. Conversions were started by GT addition and incubated at 30°C. The GT HPLC protocol was applied for sampling after distinct incubation times and quantification of **1**, **2**, **3**, **5** and **6**.

Coupled glucosylation of 5 mM **1** by *Pc*OGT and *Gm*SuSy (100 mU mL^{-1} each) in presence of 0.5 mM UDP and 100 mM **9** was studied at various pH as described for standard conditions. Buffers were prepared in steps of 0.5 pH units (BisTris pH 5.5-7.5, TAPS pH 7.5-8.5) and added to a final concentration of 50 mM. Actual pH in conversions was calculated as the average of measurements in beginning and at the end of conversions.

Isolation and identification of 5 and 6

5 and **6** from glucosylation of **2** by *Pc*OGT and *Os*CGT I121D mutant were purified by preparative reversed phase C-18 HPLC on an Agilent 1200 system equipped with a SphereClone 5 µm ODS(2) (250 x 10.0 mm) column. Water was used as solvent A and acetonitrile as solvent B (0.1% formic acid each). Separation of **5** and **6** from other compounds was achieved by step gradients from 10% to 100% B at room temperature. **5** eluted at 25% and **6** at 35% B. After removing acetonitrile under reduced pressure water was removed by freeze drying. Purity and identity were confirmed by HPLC and NMR. ^{1}H-NMR, ^{13}C-NMR, COSY, HMBC and HMQC (**5** only) were recorded on a Varian Unity Inova 500 MHz spectrometer.

Acknowledgements

Notes and references

[a] *Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, 8010 Graz, Austria.* E-mail: bernd.nidetzky@tugraz.at; Fax: (+43)-316-873-8434; Tel: (+43)-316-873-8400
[b] *Institute of Organic Chemistry, Graz University of Technology, Stremayrgasse 9, 8010 Graz, Austria*

† Electronic Supplementary Information (ESI) available: protein gel; HPLC traces: distinction of **2** glucosides, purified **5** and **6**; *Pc*OGT-*Gm*SuSy conversions at various pH; NMR data

1. Ø. M. Andersen and K. R. Markham, *Flavonoids: chemistry, biochemistry, and applications*, CRC Press, Boca Raton, FL, 2006.

2. E. Grotewold, *The science of flavonoids*, Springer, New York, 2006.
3. S. Barnes, J. Prasain, T. D'Alessandro, A. Arabshahi, N. Botting, M. A. Lila, G. Jackson, E. M. Janle and C. M. Weaver, *Food Funct.*, 2011, **2**, 235-244.
4. C. Manach, G. Williamson, C. Morand, A. Scalbert and C. Rémésy, *Am. J. Clin. Nutr.*, 2005, **81**, 230S-242S.
5. A. K. Ratty and N. P. Das, *Biochem. Med. Metab. Biol.*, 1988, **39**, 69-79.
6. C. A. Rice-Evans, N. J. Miller and G. Paganga, *Free Radical Biol. Med.*, 1996, **20**, 933-956.
7. M. G. L. Hertog, E. J. M. Feskens, P. C. H. Hollman, M. B. Katan and D. Kromhout, *Lancet*, 1993, **342**, 1007-1011.
8. A. Picinelli, E. Dapena and J. J. Mangas, *J. Agric. Food Chem.*, 1995, **43**, 2273-2278.
9. A. Escarpa and M. C. González, *J. Chromatogr. A. A*, 1998, **823**, 331-337.
10. Y. Nakamura, S. Watanabe, N. Miyake, H. Kohno and T. Osawa, *J. Agric. Food Chem.*, 2003, **51**, 3309-3312.
11. B. M. Rezk, G. R. M. M. Haenen, W. J. F. van der Vijgh and A. Bast, *Biochem. Biophys. Res. Commun.*, 2002, **295**, 9-13.
12. V. Stangl, M. Lorenz, A. Ludwig, N. Grimbo, C. Guether, W. Sanad, S. Ziemer, P. Martus, G. Baumann and K. Stangl, *J. Nutr.*, 2005, **135**, 172-178.
13. R. B. An, E. J. Park, G. S. Jeong, D. H. Sohn and Y. C. Kim, *Arch. Pharmacal Res.*, 2007, **30**, 674-677.
14. C. Pohl, F. Will, H. Dietrich and D. Schrenk, *J. Agric. Food Chem.*, 2006, **54**, 10262-10268.
15. H. Nguyen, S. Zhang and M. E. Morris, *J. Biopharm. Sci.*, 2003, **92**, 250-257.
16. O. Desire, C. Rivière, R. Razafindrazaka, L. Goossens, S. Moreau, J. Guillon, S. Uverg-Ratsimamanga, P. Andriamadio, N. Moore, A. Randriantsoa and A. Raharisololalao, *J. Ethnopharmacol.*, 2010, **130**, 320-328.
17. T. Mencherini, P. Picerno, P. Del Gaudio, M. Festa, A. Capasso and R. Aquino, *J. Nat. Prod.*, 2010, **73**, 247-251.
18. S. Logendra, D. M. Ribnicky, H. Yang, A. Poulev, J. Ma, E. J. Kennelly and I. Raskin, *Phytochemistry*, 2006, **67**, 1539-1546.
19. R. W. Gantt, P. Peltier-Pain and J. S. Thorson, *Nat. Prod. Rep.*, 2011, **28**, 1811-1853.
20. C. J. Thibodeaux, C. E. Melançon III and H. W. Liu, *Angew. Chem., Int. Ed.*, 2008, **47**, 9814-9859.
21. Y. Kaminaga, A. Nagatsu, T. Akiyama, N. Sugimoto, T. Yamazaki, T. Maitani and H. Mizukami, *FEBS Lett.*, 2003, **555**, 311-316.
22. X. Wang, *FEBS Lett.*, 2009, **583**, 3303-3309.
23. C. M. M. Gachon, M. Langlois-Meurinne and P. Saindrenan, *Trends Plant Sci.*, 2005, **10**, 542-549.
24. V. Křen and L. Martínková, *Curr. Med. Chem.*, 2001, **8**, 1303-1328.
25. D. Bowles, J. Isayenkova, E. K. Lim and B. Poppenberger, *Curr. Opin. Plant Biol.*, 2005, **8**, 254-263.
26. M. Montefiori, R. V. Espley, D. Stevenson, J. Cooney, P. M. Datson, A. Saiz, R. G. Atkinson, R. P. Hellens and A. C. Allan, *Plant J.*, 2011, **65**, 106-118.
27. P. W. Snijman, E. Joubert, D. Ferreira, X. C. Li, Y. Ding, I. R. Green and W. C. A. Gelderblom, *J. Agric. Food Chem.*, 2009, **57**, 6678-6684.
28. J. R. L. Ehrenkranz, N. G. Lewis, C. R. Kahn and J. Roth, *Diabetes/Metab. Res. Rev.*, 2005, **21**, 31-38.
29. C. Gosch, H. Halbwirth and K. Stich, *Phytochemistry*, 2010, **71**, 838-843.
30. S. R. Jensen, B. J. Nielsen and V. Norn, *Phytochemistry*, 1977, **16**, 2036-2038.
31. T. Tanaka, K. Kawamura, H. Kohda, K. Yamasaki and O. Tanaka, *Chem. Pharm. Bull.*, 1982, **30**, 2421-2423.
32. G. D. Winget, S. Izawa and N. E. Good, *Biochemistry*, 1969, **8**, 2067-2074.
33. M. Gaucher, T. Dugé de Bernonville, D. Lohou, S. Guyot, T. Guillemette, M. N. Brisset and J. F. Dat, *Phytochemistry*, 2013, **90**, 78-89.
34. H. Q. Dong, M. Li, F. Zhu, F. L. Liu and J. B. Huang, *Food Chem.*, 2012, **130**, 261-266.
35. C. Gosch, H. Halbwirth, B. Schneider, D. Hölscher and K. Stich, *Plant Sci.*, 2010, **178**, 299-306.
36. T. Dugé de Bernonville, S. Guyot, J. P. Paulin, M. Gaucher, L. Loufrani, D. Henrion, S. Derbré, D. Guilet, P. Richomme, J. F. Dat and M. N. Brisset, *Phytochemistry*, 2010, **71**, 443-452.
37. X. Zhu and R. R. Schmidt, *Angew. Chem., Int. Ed.*, 2009, **48**, 1900-1934.
38. B. M. de Roode, M. C. Franssen, A. van der Padt and R. M. Boom, *Biotechnol. Prog.*, 2003, **19**, 1391-1402.
39. T. Desmet, W. Soetaert, P. Bojarová, V. Křen, L. Dijkhuizen, V. Eastwick-Field and A. Schiller, *Chem. - Eur. J.*, 2012, **18**, 10786-10801.
40. A. Bertrand, S. Morel, F. Lefoulon, Y. Rolland, P. Monsan and M. Remaud-Simeon, *Carbohydr. Res.*, 2006, **341**, 855-863.
41. L. L. Lairson, B. Henrissat, G. J. Davies and S. G. Withers, *Annu. Rev. Biochem.*, 2008, **77**, 521-555.
42. H. Jugdé, D. Nguy, I. Moller, J. M. Cooney and R. G. Atkinson, *FEBS J.*, 2008, **275**, 3804-3814.
43. S. R. Werner and J. A. Morgan, *J. Biotechnol.*, 2009, **142**, 233-241.
44. 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.
45. L. Bungaruang, A. Gutmann and B. Nidetzky, *Adv. Synth. Catal.*, 2013, **355**, 2757-2763.
46. A. Gutmann and B. Nidetzky, *Angew. Chem., Int. Ed.*, 2012, **51**, 12879-12883.
47. M. Zhou, A. Hamza, C. G. Zhan and J. S. Thorson, *J. Nat. Prod.*, 2013, **76**, 279-286.
48. M. Brazier-Hicks, K. M. Evans, M. C. Gershater, H. Puschmann, P. G. Steel and R. Edwards, *J. Biol. Chem.*, 2009, **284**, 17926-17934.
49. A. Gutmann, C. Krump, L. Bungaruang and B. Nidetzky, *Chem. Commun.*, 2014, **50**, 5465-5468.
50. K. Terasaka, Y. Mizutani, A. Nagatsu and H. Mizukami, *FEBS Lett.*, 2012, **586**, 4344-4350.
51. D. K. Owens and C. A. McIntosh, *Phytochemistry*, 2009, **70**, 1382-1391.
52. K. De Winter, K. Verlinden, V. Křen, L. Weignerová, W. Soetaert and T. Desmet, *Green Chem.*, 2013, **15**, 1949-1955.
53. K. De Winter, T. Desmet, T. Devlamynck, L. Van Renterghem, T. Verhaeghe, H. Pelantová, V. Křen and W. Soetaert, *Org. Process Res. Dev.*, 2014, DOI: 10.1021/op400302b.
54. V. Siddaiah, C. V. Rao, S. Venkateswarlu and G. V. Subbaraju, *Tetrahedron*, 2006, **62**, 841-846.

Electronic Supplementary Information

76

Towards green synthesis of glycosylated dihydrochalcone natural products using glycosyltransferase-catalysed cascade reactions

Alexander Gutmann,[a] Linda Bungaruang,[a] Hansjoerg Weber,[b] Mario Leypold,[b] Rolf Breinbauer[b] and Bernd Nidetzky*[a]

[a] *Institute of Biotechnology and Biochemical Engineering,*
Graz University of Technology, Petersgasse 12/1, A-8010 Graz, Austria

[b] *Institute of Organic Chemistry,*
Graz University of Technology, Stremayrgasse 9, 8010 Graz, Austria
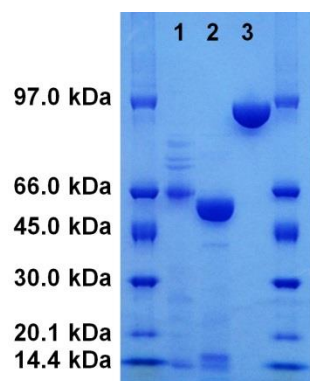
Fig. S1 SDS-PAGE of enzymes from *E. coli* overexpression cultures purified by *Strep*-tag affinity chromatography; lane 1: *Pc*OGT (55.4 kDa), lane 2: *Os*CGT I121D (51.3 kDa), lane 3: *Gm*SuSy (94.1 kDa)
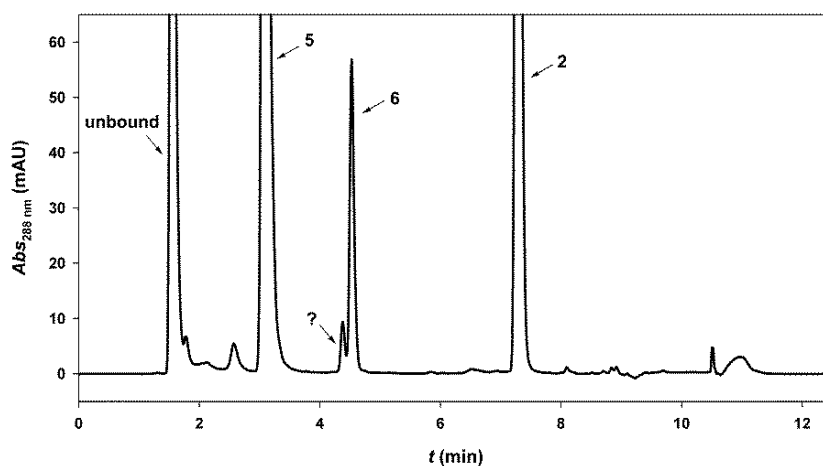
Fig. S2 Reversed-phase C-18 HPLC-analysis of a mixture of glucosylations of **2** by *Pc*OGT and *Os*CGT I121D clearly shows that the minor product of the *Pc*OGT reaction (**?**) is distinct from confusoside (**6**), formed by the *Os*CGT variant.
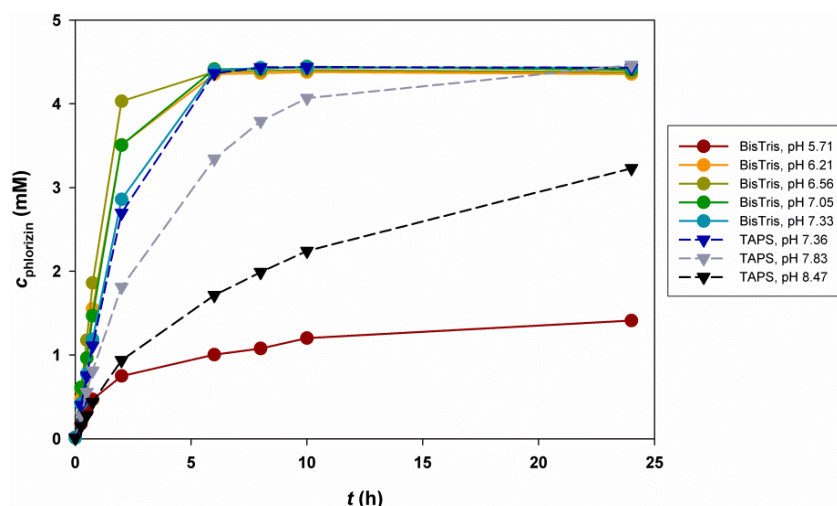
Fig. S3 Time courses of **3** formation through glucosylation of 5 mM **1** by coupled *Pc*OGT-*Gm*SuSy reaction (0.5 mM UDP, 100 mM **9**) using BisTris and TAPS reaction buffers at various pH.
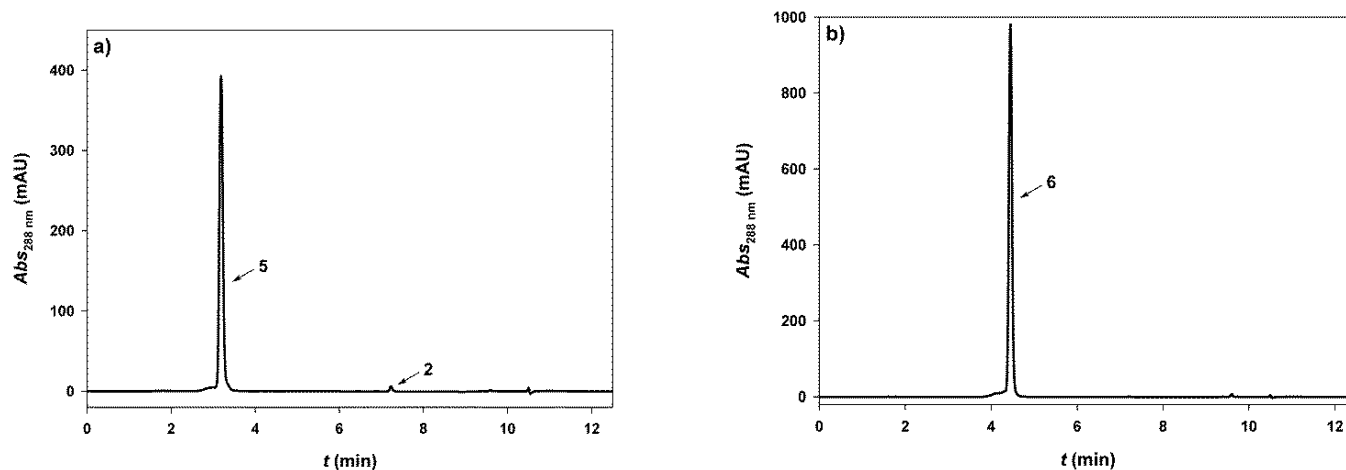
Fig. S4 Reversed-phase C-18 HPLC-analysis of (a) davidioside (**5**) and (b) confusoside (**6**) after purification by preparative HPLC confirms them to be of high purity (>98% based on HPLC peak area).

Table S1 ^{1}H and ^{13}C-NMR spectral data of davidigenin (**2**), davidioside (**5**) and confusoside (**6**)

nr	davidigenin (2) (R₁, R₂ = H)[a]		davidioside (5) (R₁ = glucose, R₂ = H)[b]		confusoside (6) (R₁ = H, R₂ = glucose)[b]	
	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H
1	133.0		134.0		133.6	
2 / 6	129.2	7.07 (2H, d, J = 8.5 Hz)	130.5	7.03 (2H, d, J = 8.4 Hz)	130.5	7.05 (2H, d, J = 8.4 Hz)
3 / 5	115.0	6.68 (2H, d, J = 8.5 Hz)	116.3	6.67 (2H, d, J = 8.2 Hz)	116.4	6.69 (2H, d, J = 8.4 Hz)
4	155.5	9.17 (1H, s)	156.6		156.9	
C=O	203.9		203.0		206.4	
α	39.4	3.22 (2H, t, J = 7.6 Hz)	46.4	~3.3[c]	41.4	3.23 (2H, t, J = 7.3 Hz)
β	29.1	2.83 (2H, t, J = 7.4 Hz)	31.1	2.86 (2H, t, J = 7.5 Hz)	31.0	2.92 (2H, t, J = 7.3 Hz)
1'	112.5		122.0		116.1	
2'	164.7	10.62 (1H, s)	160.4		165.9	
3'	102.4	6.26 (1H, d, J = 2.2 Hz)	104.1	6.70 (1H, d, J = 2.0 Hz)	105.2	6.58 (1H, d, J = 2.2 Hz)
4'	164.3	12.65 (1H, s)	164.5		165.2	
5'	108.1	6.37 (1H, dd, J = 8.74, 2.3 Hz)	111.0	6.50 (1H, dd, J = 8.5, 1.9 Hz)	109.5	6.62 (1H, dd, J = 8.9, 2.3 Hz)
6'	131.0	7.81 (1H, d, J = 8.8 Hz)	133.4	7.58 (1H, d, J = 8.6 Hz)	133.2	7.81 (1H, d, J = 8.9 Hz)
1"			102.7	4.99 (1H, d, J = 7.1 Hz)	101.5	5.00 (1H, d, J = 7.2 Hz)
2"			75.0		74.9	
3"			78.4	3.33-3.48-(4H, unresolved)	78.1	3.45-3.50 (3H, unresolved)
4"			71.4		71.4	
5"			78.6		78.5	3.41 (1H, m)
6"			62.7	3.91 (1H, dd, J = 12.4, 2.0 Hz) 3.72 (1H, dd, J = 12.1, 5.7 Hz)	62.5	3.89 (1H, dd, J = 12.1, 2.1 Hz) 3.70 (1H, dd, J = 12.3, 5.5 Hz)

[a] ^{1}H: 300.36 MHz, ^{13}C: 75.53 MHz; (DMSO-d_6, δ in ppm)
[b] ^{1}H: 499.89 MHz, ^{13}C: 125.70 MHz; (CD$_3$OD, δ in ppm)
[c] overlap with MeOH signal

Synthesis of glycosylated dihydrochalcones by glycosyltransferase-cascade reactions
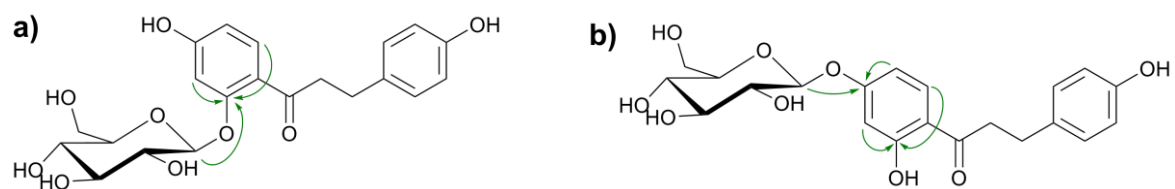
Scheme S1 Key HMBC couplings to identify (a) davidioside (**5**) and (b) confusoside (**6**), respectively
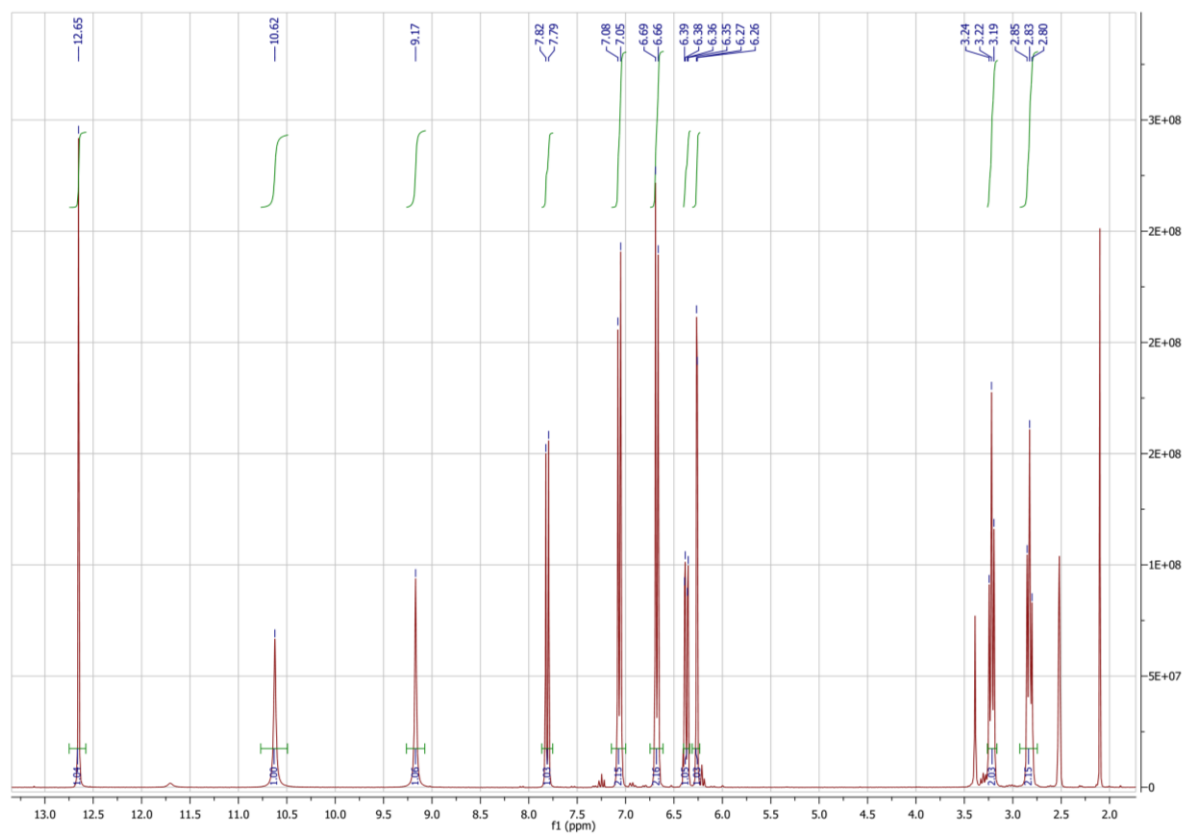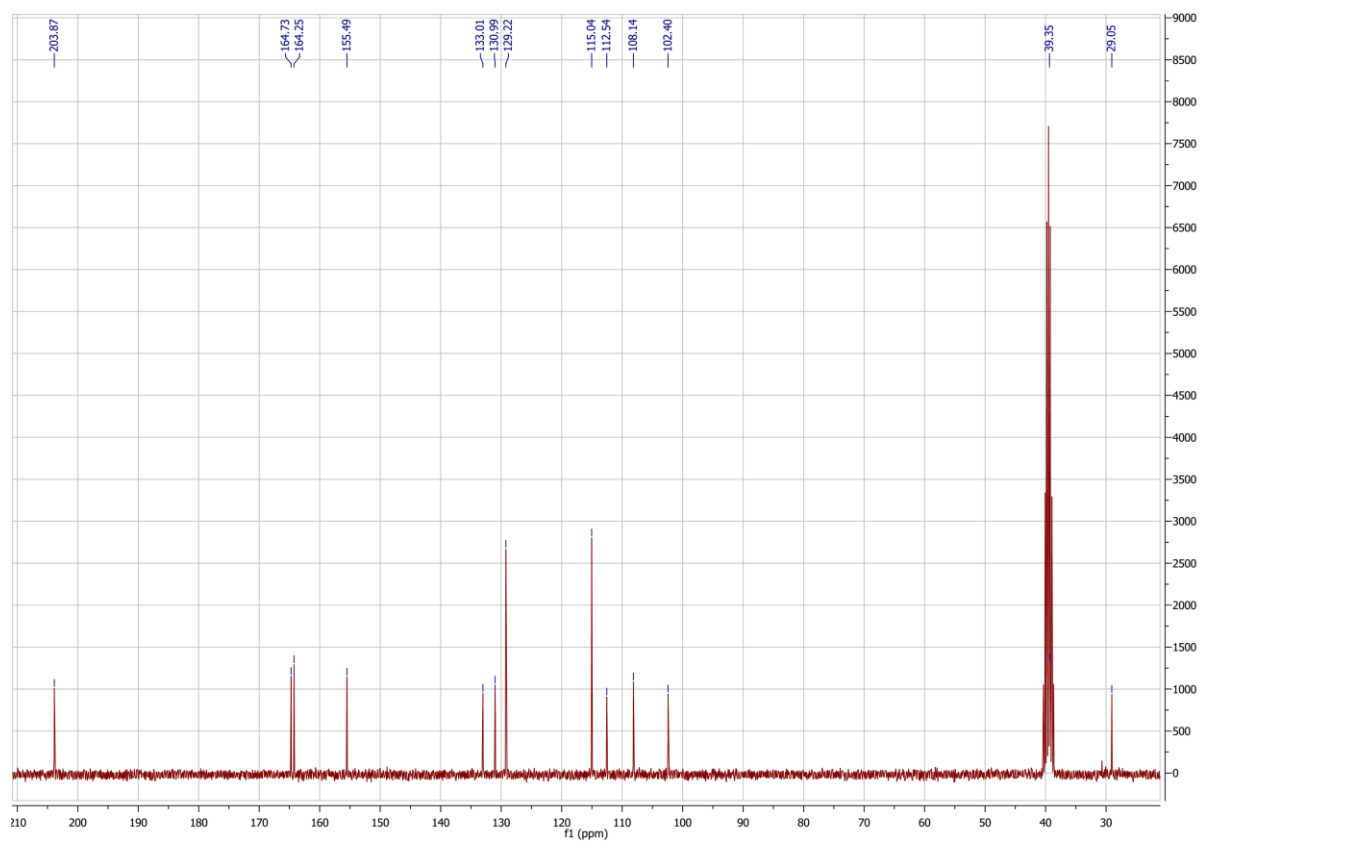

Fig. S5 ^{1}H-NMR of davidigenin (**2**)
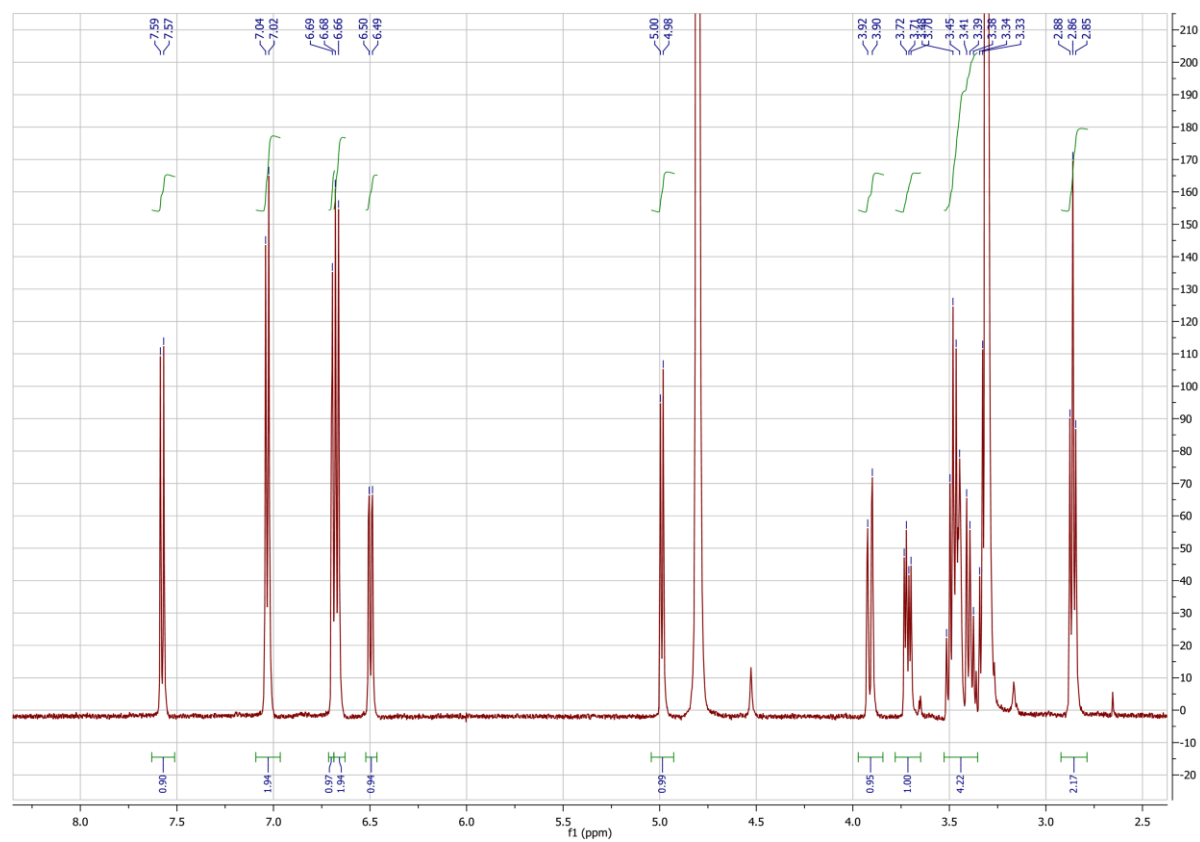
Fig. S6 ^{13}C-NMR of davidigenin (**2**)

Fig. S7 ^{1}H-NMR of HPLC purified davidioside (**5**)

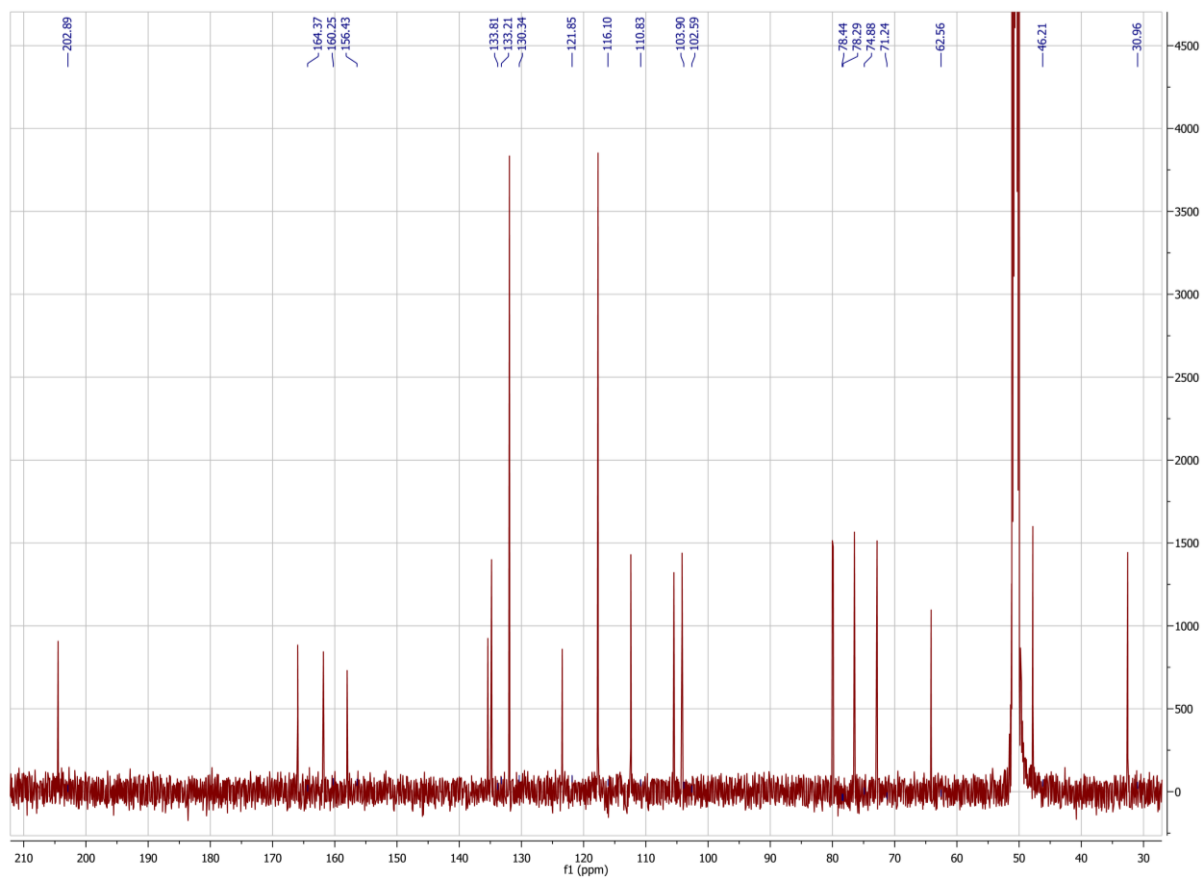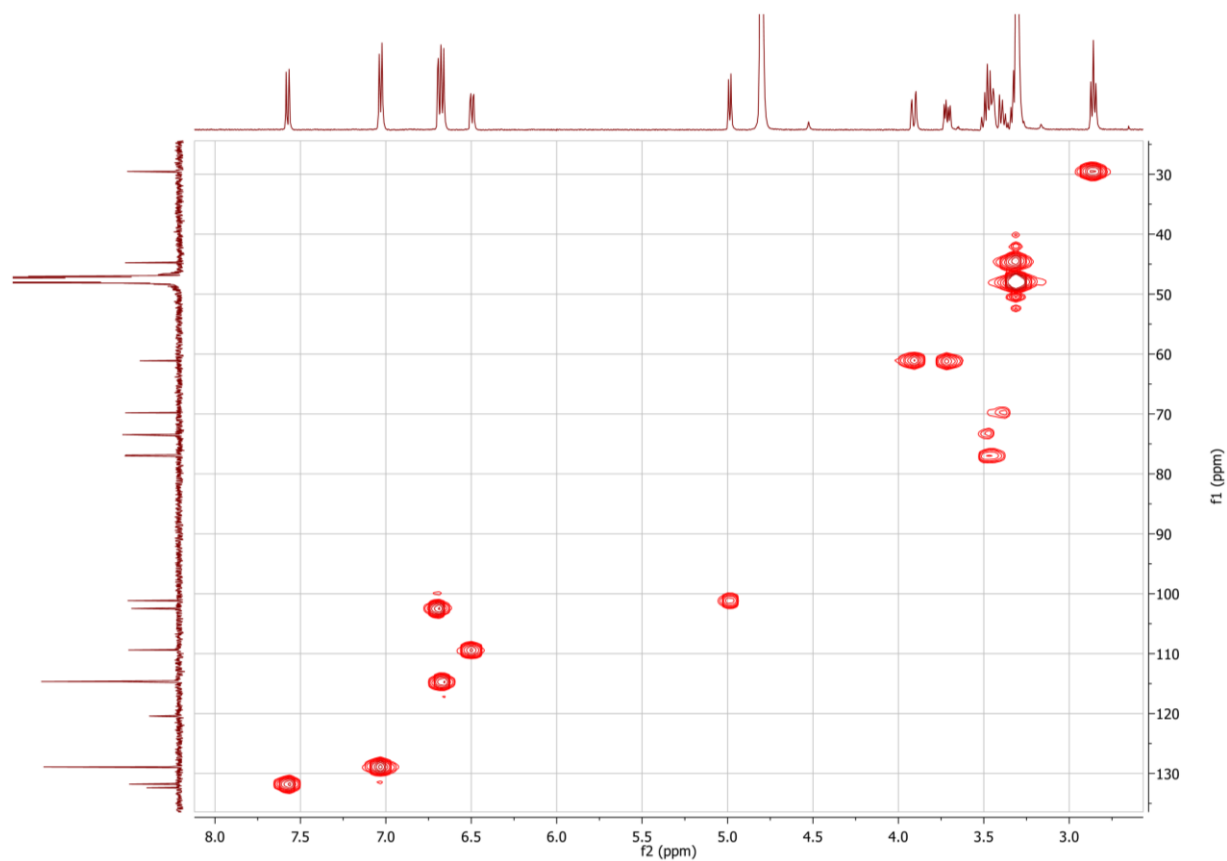**Fig. S8** ^{13}C-NMR of HPLC purified davidioside (**5**)
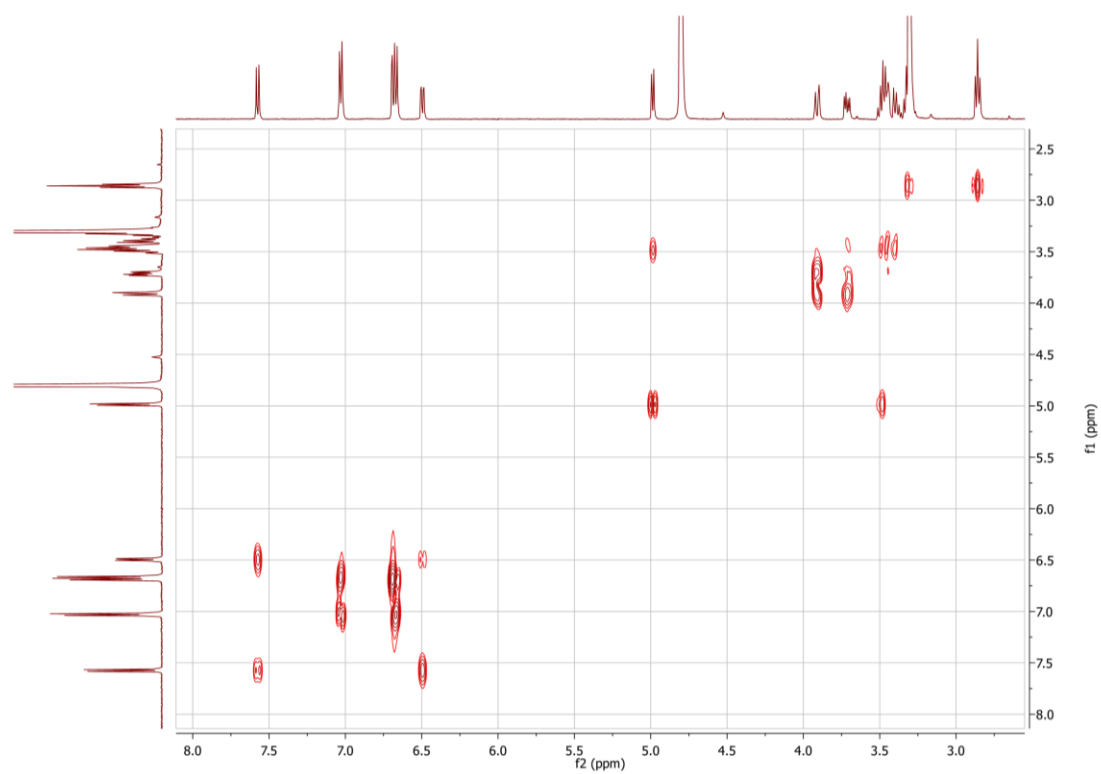

Fig. S9 2D HMQC-NMR of HPLC purified davidioside (**5**)

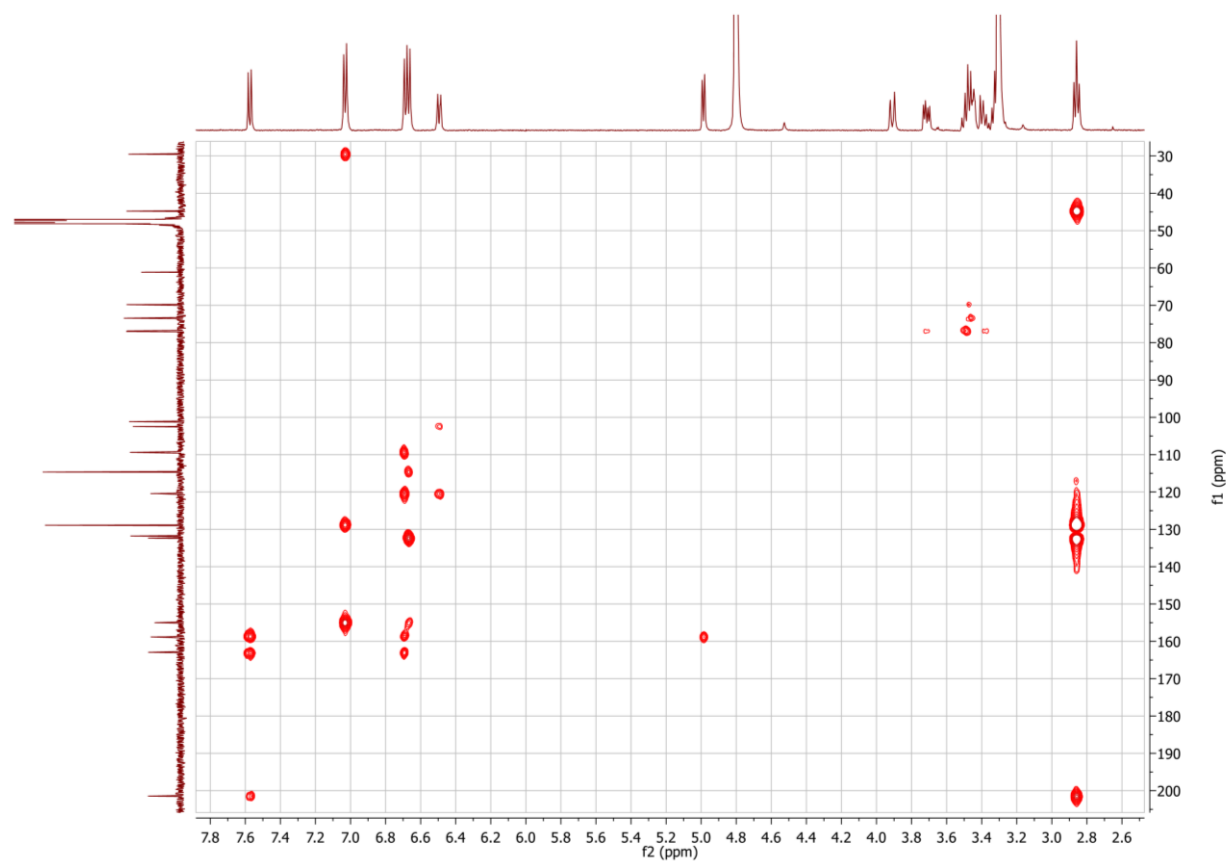**Fig. S10** 2D COSY-NMR of HPLC purified davidioside (**5**)

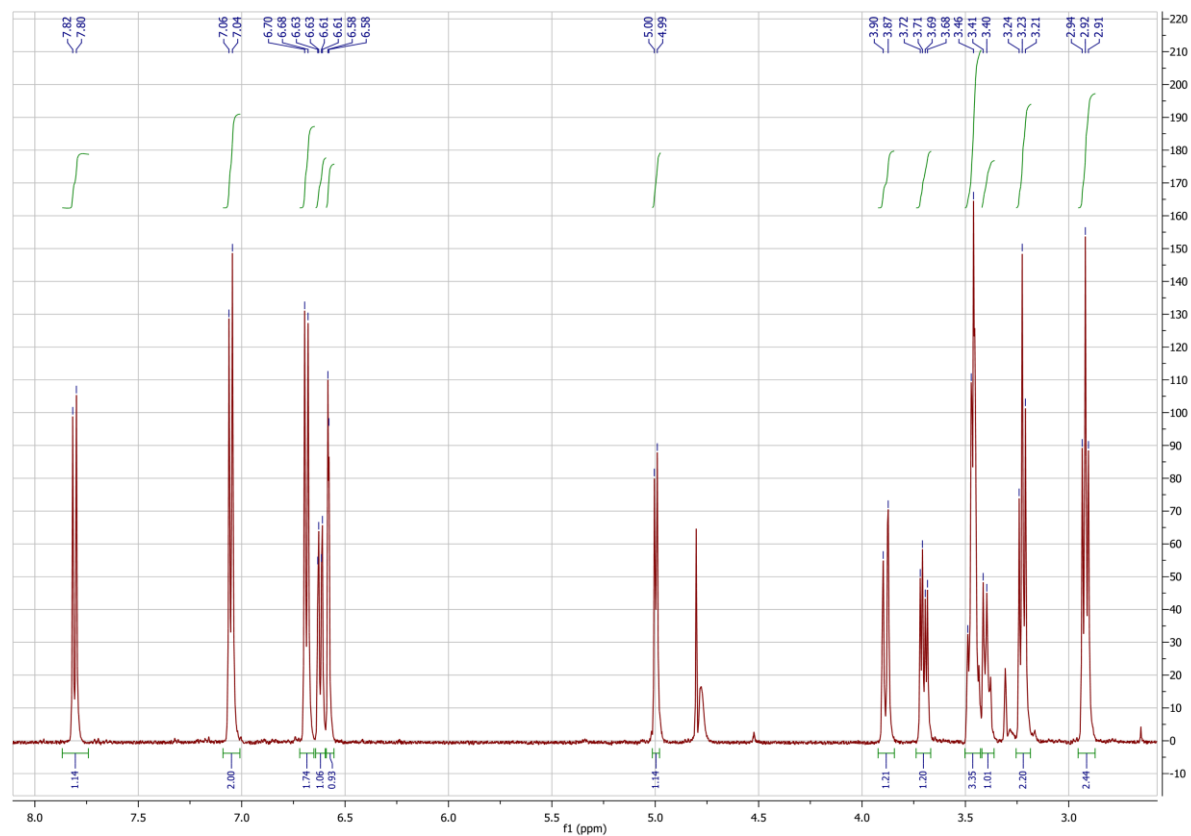**Fig. S11** 2D HMBC-NMR of HPLC purified davidioside (**5**)

Synthesis of glycosylated dihydrochalcones by glycosyltransferase-cascade reactions

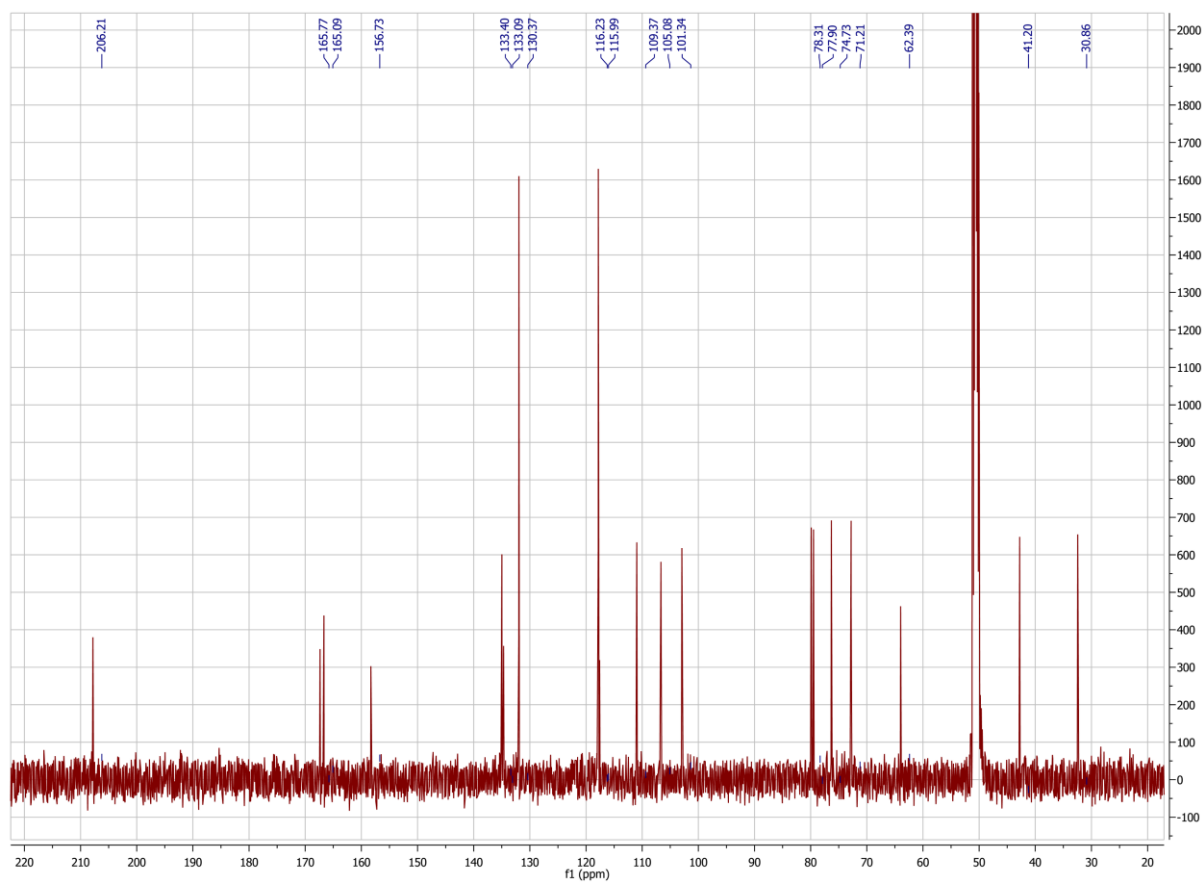**Fig. S12** ^{1}H-NMR of HPLC purified confusoside (**6**)

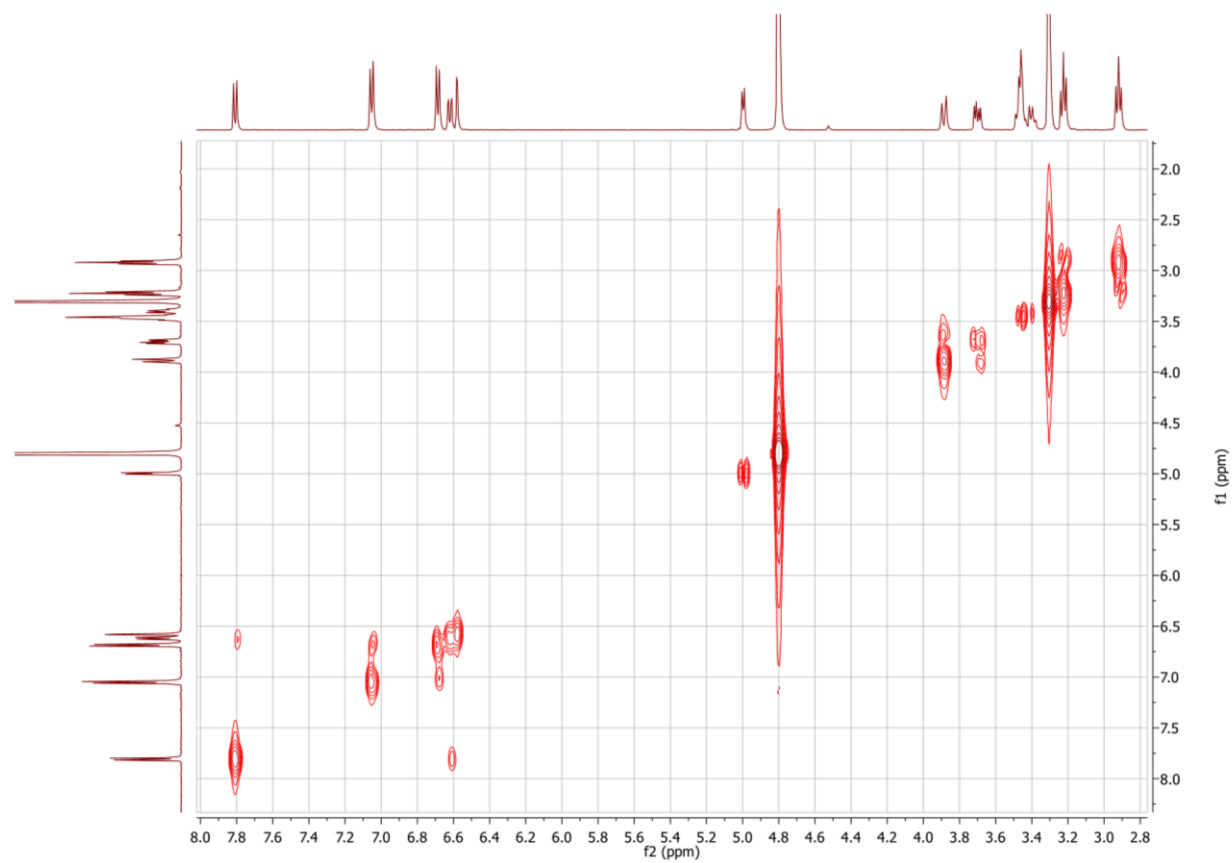**Fig. S13** ^{13}C-NMR of HPLC purified confusoside (**6**)
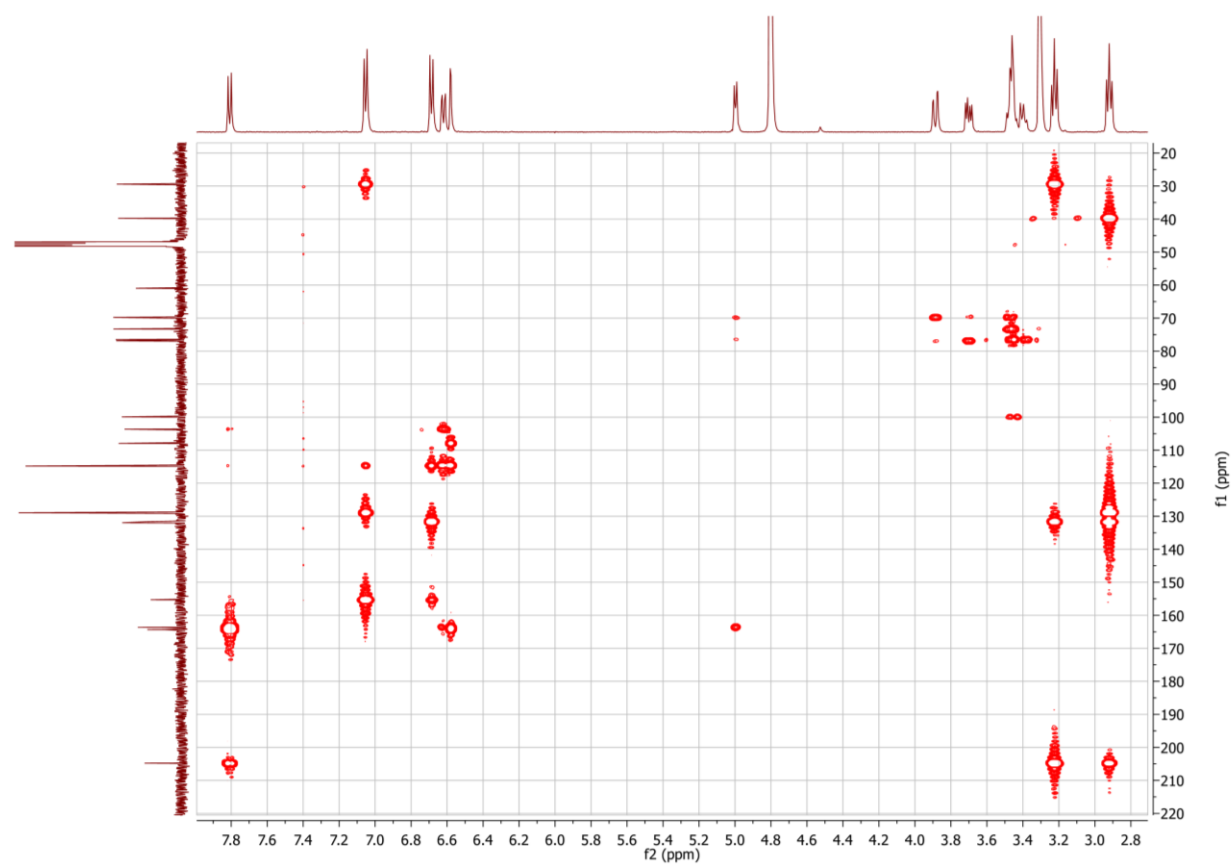
Fig. S14 2D COSY-NMR of HPLC purified confusoside (**6**)

Fig. S15 2D HMBC-NMR of HPLC purified confusoside (**6**)

Scientific record

Publications

Switching between *O*- and *C*-Glycosyltransferase through Exchange of Active-Site Motifs

Alexander Gutmann and Bernd Nidetzky, *Angew. Chem. Int. Ed.* **2012**, *51*, 12879-12883.

Enzymatic *C*-glycosylation: Insights from the study of a complementary pair of plant *O*- and *C*-glucosyltransferases

Alexander Gutmann and Bernd Nidetzky, *Pure Appl. Chem.* **2013**, *85*, 1865-1877.

Leloir Glycosyltransferases and Natural Product Glycosylation: Biocatalytic Synthesis of the *C*-Glucoside Nothofagin, a Major Antioxidant of Redbush Herbal Tea

Linda Bungaruang,[+] Alexander Gutmann,[+] and Bernd Nidetzky, *Adv. Synth. Catal.* **2013**, *355*, 2757-2763.

[+] These authors contributed equally to this work.

A two-step *O*-to *C*-glycosidic bond rearrangement using complementary glycosyltransferase activities

Alexander Gutmann, Corinna Krump, Linda Bungaruang and Bernd Nidetzky, *Chem. Commun.* **2014**, *50*, 5465-5468

Oral presentations

Deciphering the Mechanism of *C*-Glycosidic Bond Formation by a Glycosyltransferase from Rice

Alexander Gutmann and Bernd Nidetzky

15th Austrian Carbohydrate Symposium, *2011*, Graz, Austria

Insight into *C*-glycosidic Bond Formation by Plant Glycosyl Transferases

Alexander Gutmann and Bernd Nidetzky

16th Austrian Carbohydrate Symposium, *2012*, Vienna, Austria

Switching between *O*- and *C*-glycosylation by active-site motif exchange in plant glycosyltransferases

Alexander Gutmann and Bernd Nidetzky

10th Carbohydrate Bioengineering Meeting, *2013*, Prague, Czech Republic

Poster presentations

Elucidation of the role of the key amino acids for the unusual formation of C-C bonds by a glycosyltransferase from *Oryza sativa*

Alexander Gutmann, Michele Tedesco and Bernd Nidetzky

11th European Training Course on Carbohydrates, *2010*, Wageningen, The Netherlands.

Insight into the mechanism of *C*-glycosidic bond formation

Alexander Gutmann and Bernd Nidetzky

9th Carbohydrate Bioengineering Meeting, 2011, Lisboa, Portugal