Enzyme Mechanisms

DOI: 10.1002/anie.201206141

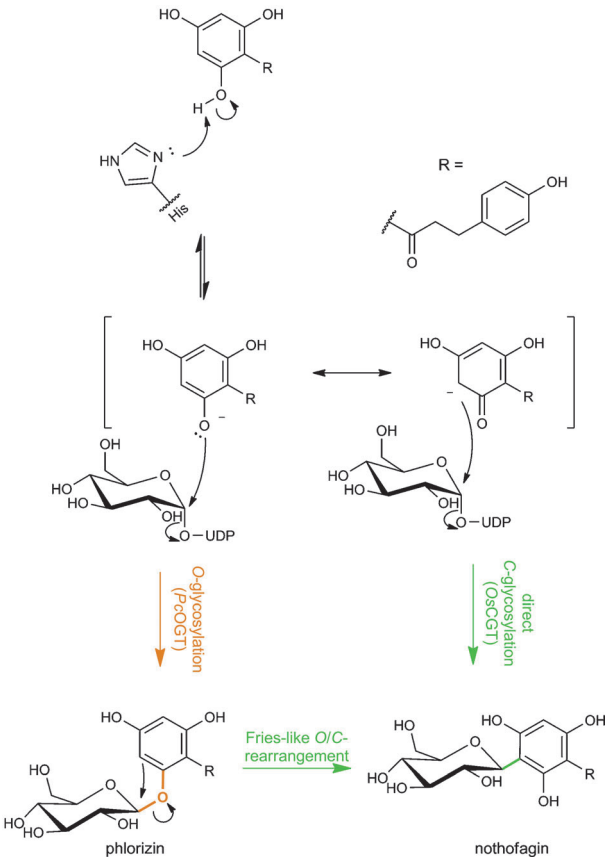
Switching between *O*- and *C*-Glycosyltransferase through Exchange of Active-Site Motifs**

*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 specificity 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



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.

[*] 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

[**] Financial support from the Austrian Science Fund (DK Molecular Enzymology W901-B05) is gratefully acknowledged. Prof. R. Edwards and Prof. K. Stich kindly provided the *Os*CGT and *Pc*OGT gene, respectively.

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

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

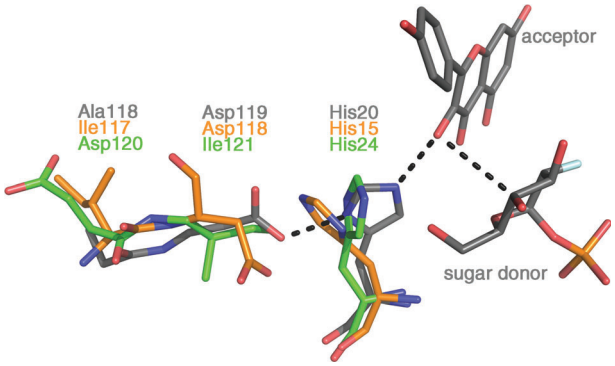


Figure 1. Overlay of modeled structures of *Os*CGT (green) and *Pc*OGT (orange) with the experimental structure of the *Vv*OGT ternary complex (gray: PDB code 2C1Z).[13] *Os*CGT shows disruption of the His-Asp dyad present in the two OGTs due to positional shift of the Asp residue.

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*-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).

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.

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

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*-

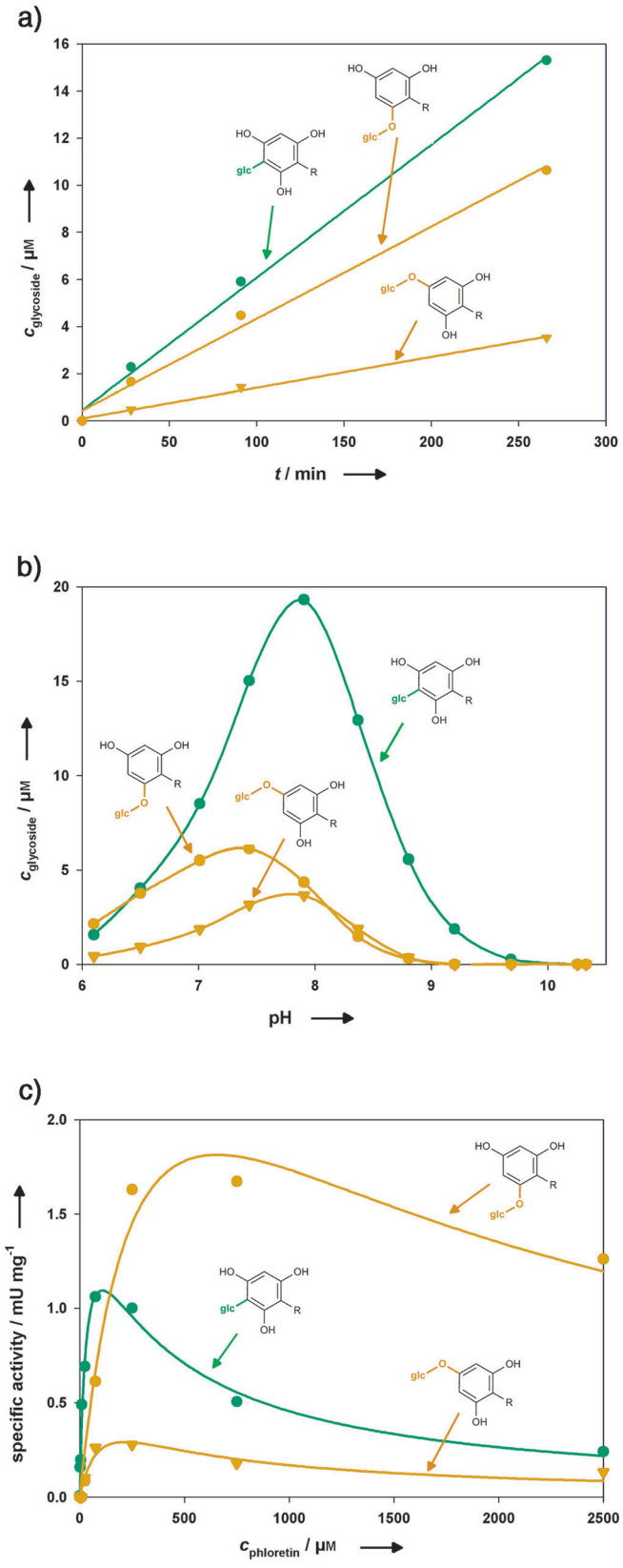


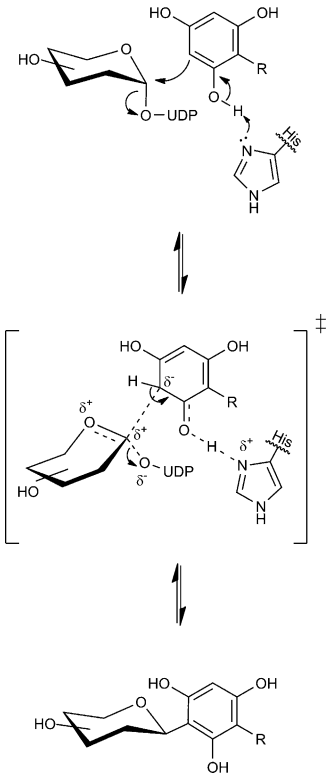
Figure 2. Mechanistic study of *Os*CGT_I121D based on detailed steady-state kinetic analysis of phloretin glucosylation. Time dependence (a) and pH dependence (b) of product formation; c) dependencies of specific rates of product formation on acceptor concentration used.

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.



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.

Received: July 31, 2012
Published online: November 14, 2012

Keywords: *C*-glycosides · enzyme mechanisms · glycosyltransferases · natural products · protein engineering

[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.

