Pure Appl. Chem., Vol. 85, No. 9, pp. 1865–1877, 2013.
http://dx.doi.org/10.1351/PAC-CON-12-11-24

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.

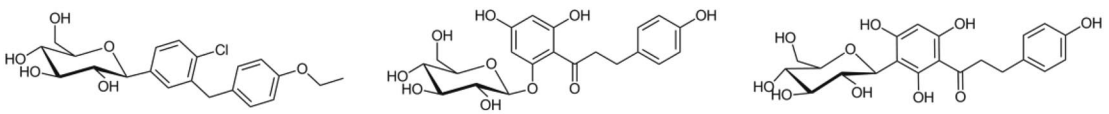
INTRODUCTION

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-

Pure Appl. Chem. **85**, 1759–1900 (2013). A collection of invited papers based on presentations at the 26th International Carbohydrate Symposium (ICS 2012), Madrid, Spain, 22–27 July 2012.
‡Corresponding author

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].

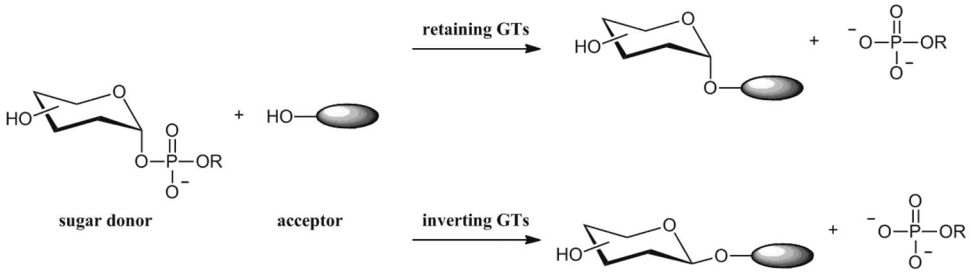
Glycosylation of natural products usually occurs via *O*-glycosidic linkages, which represent the most prevalent type of glycosidic bond in Nature [10]. However, there are also *N*-glycosidic natural products, and in some quite rare cases, a *C*-glycosidic linkage, typically involving the carbon atom 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].

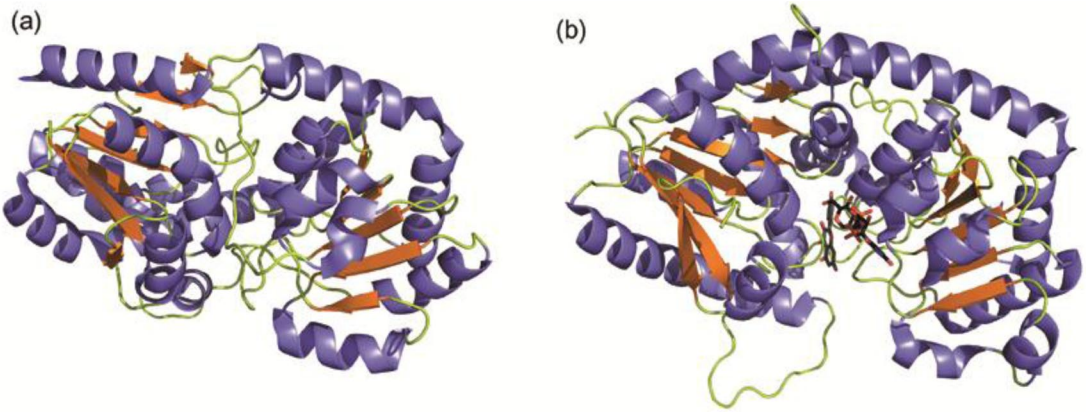
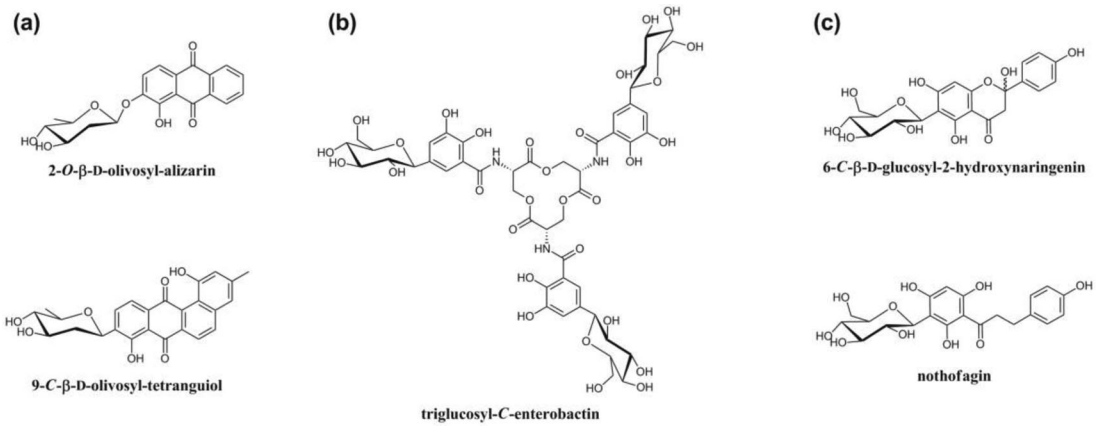


Fig. 1 The bacterial C-GT UrdGT2 (a) and the plant O-GT (from *Vitis vinifera*; grape vine) (b) feature both GT-B fold with two β/α/β Rossmann-like domains. *Vv*_O-GT displays substrates bound in the active site at the interdomain cleft.

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



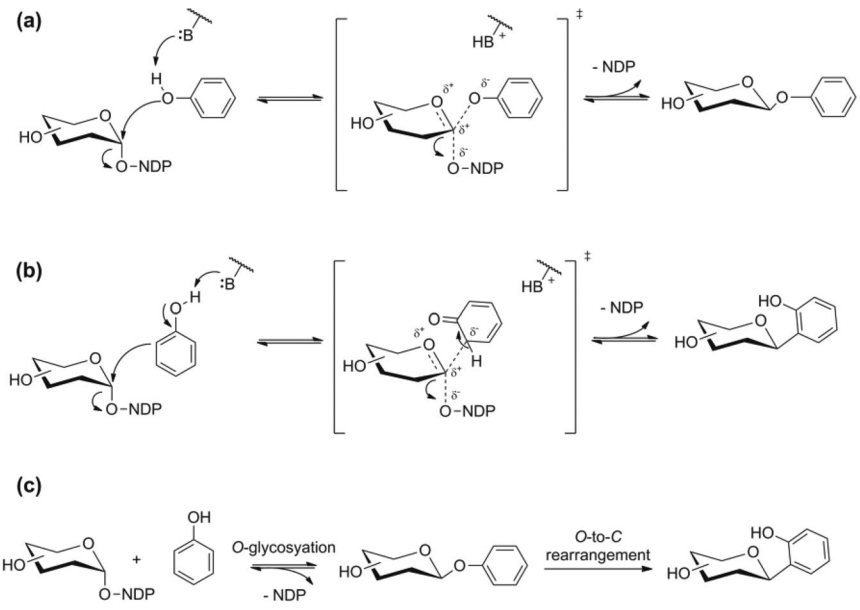
Scheme 3 Conversion products of the *C*-GTs UrdGT2 (a), IroB (b), and *Os*_C-GT (c). Only UrdGT2 is known to form both *O*- and *C*-glycosides.

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-trihydroxy-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-

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.

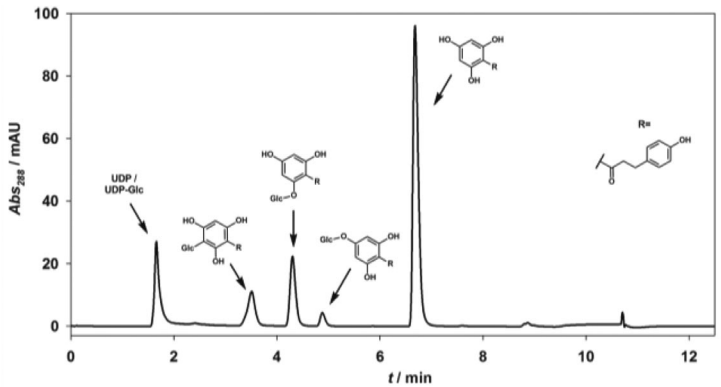


Fig. 2 Using a water to acetonitrile gradient on a reversed phase (C-18) column phloretin and its glycosides could be separated and quantified by HPLC.

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

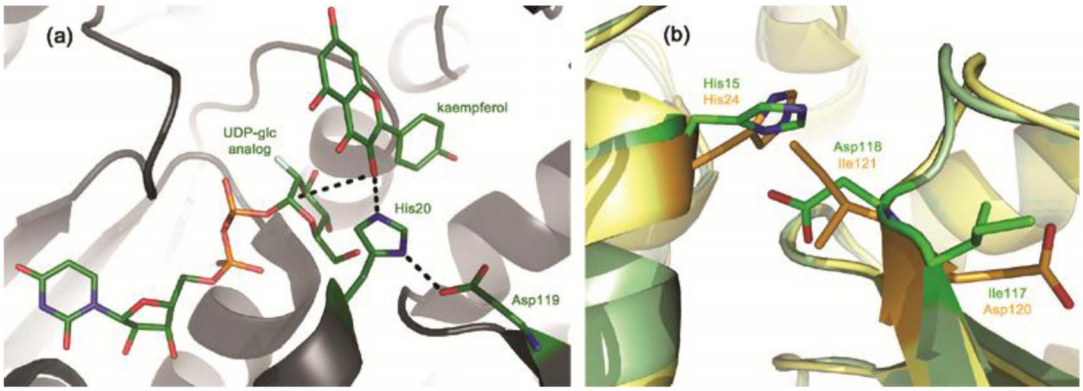


Fig. 3 Crystal structure of ternary *Vv*_O-GT complex (pdb code 2C1Z) (a) and overlay of modeled *Os*_C-GT (yellow) and *Pc*_O-GT (green) structures (b); *Os*_C-GT displays the conserved GT-B fold but disruption in the catalytic His-Asp motif present in O-GTs.

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

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.

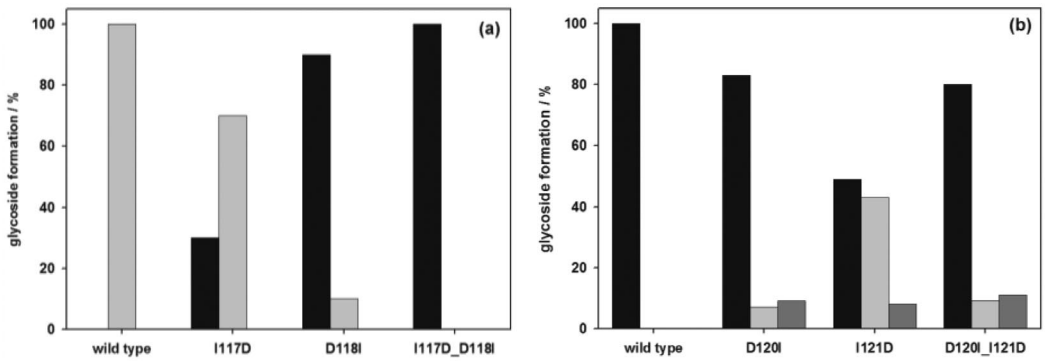


Fig. 4 Distribution of phloretin glucosides produced by wild-type and mutant *Pc*_O-GT (a) and *Os*_C-GT (b); black: 3'-*C*-glucoside (nothofagin); light gray: 2'-*O*-glucoside (phlorizin); dark gray: 4'-*O*-glucoside.

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-

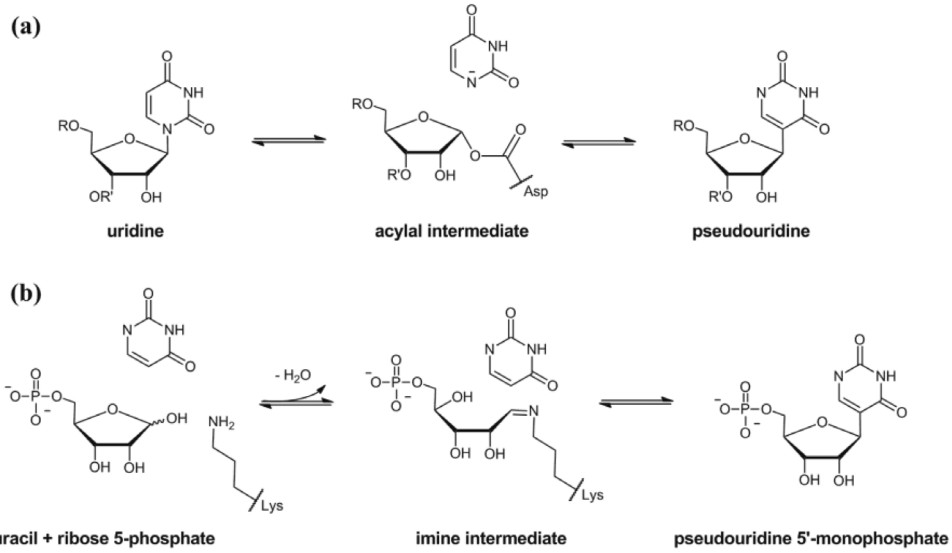
bon in a single reaction step. Such a reaction could also be described as a Friedel–Crafts-like direct alkylation of the phenolic acceptor in an electrophilic aromatic substitution. The nucleophilic displacement probably involves formation of an oxocarbenium ion-like transition state whereby the positively 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

This work was financially supported by the Austrian Science Fund (DK Molecular Enzymology W901-B05).

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).

