A Kinetic Analysis of Regiospecific Glucosylation by Two Glycosyltransferases of Arabidopsis thaliana

Domain Swapping to Introduce New Activities

Adam M. Cartwright1, Eng-Kiat Lim3, Colin Kleanthous5, and Dianna J. Bowles1,2

From the 4Centre for Novel Agricultural Products and the 5Department of Biology, University of York, York YO10 SDD, United Kingdom

Plant Family 1 glycosyltransferases (GTs) recognize a wide range of natural and non-natural scaffolds and have considerable potential as biocatalysts for the synthesis of small molecule glucosides. Regiospecificity of glucosylation is an important property, given that many acceptors have multiple potential glycosylation sites. This study has used a domain-swapping approach to explore the determinants of regiospecific glucosylation of two GTs of Arabidopsis thaliana, UGT74F1 and UGT74F2. The flavonoid quercetin was used as a model acceptor, providing five potential sites for O-glycosylation by the two GTs. As is commonly found for many plant GTs, both of these enzymes produce distinct multiple glycosides of quercetin. A high performance liquid chromatography method has been established to perform detailed steady-state kinetic analyses of these concurrent reactions. These data show the influence of each parameter in determining a GT product formation profile toward quercetin. Interestingly, construction and kinetic analyses of a series of UGT74F1/F2 chimeras have revealed that mutating a single amino acid distal to the active site, Asn-142, can lead to the development of a new GT with a more constrained regiospecificity. This ability to form the 4′-O-glucoside of quercetin is transferable to other flavonoid scaffolds and provides a basis for preparative scale production of flavonoid 4′-O-glucosides through the use of whole-cell biocatalysis.

The glucosylation of small hydrophobic molecules is known to alter their chemical and biological characteristics. The alterations include changes in water solubility, stability, pharmacokinetic properties, and bioactivity (1, 2). Enzymes that catalyze these reactions are Family 1 glycosyltransferases (GTs), with representatives found in a wide range of prokaryotic and eukaryotic organisms.

Extensive study of GT activities has demonstrated that the enzymes display an exquisite regio- (3), enantio- (4), and chemospecificity (5) toward the acceptor molecule. However, as yet the determinants of this specificity are poorly understood. This lack of understanding presents challenges to the interpretation of substrate activity data (6, 7), GT rational redesign (8, 9), and activity prediction (10, 11). Future progress in predicting GT sequence-structure-activity relationships will depend on a greater number of studies that characterize the key determinants of activity and specificity.

Quercetin is the most abundant flavonoid in the human diet (12) and has frequently been used as a model substrate for GT activity (13–15). Quercetin aglycone is found only at low concentrations in the primary dietary source, plants, where glucosylated forms predominate (16). In addition, in mammals, glycosides are the major byproducts of quercetin phase II metabolism (17). Many studies have shown that glycosylation of quercetin significantly affects bioavailability and efficacy with respect to anti-oxidant, anti-proliferative, and anti-cancer properties (18). There are also some data to indicate that the position of glucosylation can affect bioactivity (19–21).

Many plant, animal, and microbial GTs can recognize quercetin as an acceptor when assayed in vitro. Although certain enzymes show some specificity toward individual hydroxyl groups of the aglycone, others can glycosylate multiple hydroxyl groups. For example, the model plant Arabidopsis thaliana contains 107 GT open reading frames that glycosylate a broad range of acceptors in vivo and in vitro (1). The study of Lim et al. (3) revealed the range of A. thaliana GT activities toward quercetin. Some GTs such as UGT78D2 and UGT71C2 glycosylated only the 3-OH or 7-OH positions, respectively, whereas other GTs, such as UGT88A1, were capable of recognizing multiple positions.

To date, analyses of GTs that form multiple monoglycosides of quercetin have typically relied on product formation profiles to give an indication of regiospecificity (22, 23). This provides limited information because the underlying kinetic parameters for each glycosylating reaction are masked by the single reaction condition used to generate the product formation profile. This limitation is equally applicable to studies of GT regiospecificity toward acceptors beyond quercetin.

In this study, a domain-swapping strategy of two highly related A. thaliana GTs, UGT74F1 and UGT74F2, has been used to explore the basis of their differing regiospecificity toward quercetin. A high performance liquid chromatography (HPLC)-based kinetic analysis has been established to analyze
Escherichia coli expressed from the pGEX-2T vector as GST fusion proteins in ing cells at 20 °C until late stationary phase ( oligonucleotides is provided in supplemental Table S1. Extension PCR (25) was used to create mutant F22221. A full list of Site-directed Mutagenesis—Site-directed mutagenesis was performed by a whole plasmid method based on that of Hemsley et al. (24). Mutations were confirmed by DNA sequencing. Overlap extension PCR (25) was used to create mutant F22221. A full list of oligonucleotides is provided in supplemental Table S1. Protein Expression and Purification—All proteins were expressed from the pGEX-2T vector as GST fusion proteins in Escherichia coli strain BL21 (DE3). Briefly, this involved culturing cells at 20 °C until late stationary phase (A600 > 2.0) and then inducing with 1 mM isopropyl-1-thio-β-D-galactopyranoside at 20 °C for ~16 h. Cells were lysed by osmotic shock, the lysate clarified, and the GST fusion protein purified using glutathione-Sepharose 4B. Elution from the affinity matrix was achieved by washing twice with an equal bed volume of 100 mM Tris-Cl, pH 8.0, 20 mM reduced glutathione, 120 mM NaCl. The GST fusion partner was cleaved for thermal denaturation studies. Replacing the elution steps, the Sepharose-enzyme pellet was washed with thrombin cleavage buffer (20 mM Tris-Cl, pH 8.4, 150 mM NaCl, 0.25 mM CaCl2) and then incubated with 10 units of thrombin (Sigma) for 20 min at 25 °C. Cleaved GTs were found to be >97% pure by SDS-PAGE analysis (supplemental Fig. S1) and displayed an activity equivalent to the GST fusion (data not shown). Protein concentration was determined using the Bio-Rad Quick Start Bradford Dye Reagent (1×) calibrated using amino acid analysis (Alta Bioscience, Birmingham, UK). Glycosyltransferase Activity Assay—For kinetic analysis a 50-μl volume comprising 100 mM Tris-Cl, pH 8.0, 7 mM uridine diphosphoglucose (UDP), 4–200 μM quercetin, and 0.2–2 μg of enzyme was assayed at 30 °C for 0.5–10 min and then snap-frozen in liquid nitrogen. Prior to HPLC analysis the sample was thawed, and 5 μl of trichloroacetic acid (240 mg/ml) was added and then centrifuged at 13,000 g for 10 min. Reverse-phase HPLC analysis (Waters 2695 separations module with a 486 absorbance detector) was performed on a 5 μ C18 column (Phenomenex) with a 24–59% acetonitrile gradient over 10 min; all solutions contained 0.1% trifluoroacetic acid. Spectra were recorded at 370 nm. GT activity toward additional flavonoid substrates was assayed under the same conditions except that a single time point was analyzed, a single flavonoid concentration (200 μM) was used, and HPLC analysis utilized a 10–75% acetonitrile (0.1% trifluoroacetic acid) gradient over 20 min. To qualify for kinetic analysis by pseudo-single substrate assumptions, the following criteria were defined for all reactions: 1) the other co-substrate in excess, namely UDPG (7 mM, 7–14 × Km wild-type parents) or quercetin (200 μM, 4–50 × Km wild-type parents); 2) the combined amount of product did not exceed 15% of initial substrate concentration; 3) no digluco- side was formed. UGT74F1, UGT74F2, and all mutants can form quercetin diglucosides; their presence indicates product inhibition/competition between the quercetin aglycone and monoglucoside product. Thermal Denaturation Profiles—Recombinant protein, lacking the GST fusion, was dialyzed against 10 mM phosphate buffer, pH 8.0, at 4 °C. Samples were taken at a concentration of 0.24 mg/ml and analyzed on a JASCO J810 CD spectrophotometer with a JASCO PFD-425S Peltier system. Measurements were taken at 220 nm with a data pitch of 2 °C, 1-s response, 1-nm bandwidth, and temperature slope of 2 °C/min. Mean residue weight was calculated according to Kelly et al. (26). Whole-cell Biotransformation and Flavonoid Glucoside Puriﬁcation and Identiﬁcation—Whole-cell biotransformations were performed in 1-liter shake flask cultures of E. coli BL21 (DE3) harboring the appropriate GT in a pGEX-2T vector. Cultures were grown at 28 °C, 150 rpm, in M9 minimal medium containing 0.4% glycerol and 50 μg of ampicillin to A600 0.75 and then induced with 0.1 mM isopropyl-1-thio-β-D-galactopyranoside and incubated for 16 h. Following the induction period, 10 mg of quercetin was added (time point zero). A further 10 mg of quercetin was added at 7, 27, and 31 h. Additional glycerol was added to the cultures, to a final concentration of 0.15% (v/v), at 4, 7, 25, and 31 h. The E. coli culture was centrifuged and vacuum-filtered before application to an Amberlite XAD-2 column (30-ml bed volume). Fisetin glucoside and quercetin glucoside were eluted from the column with 10 and 20% isocratic acetonitrile gradients, respectively. The fractions containing the glucosides were pooled and freeze-dried. The freeze-dried glucosides were applied to a preparative HPLC column (Gemini C18 10 μ 30 mm diameter; Phenomenex) with an isocratic 20% acetonitrile gradient at a flow rate of 10 ml/min. Quercetin glucosides were identified by comparison to known standards (3). The position of fisetin glycosylation was confirmed by proton NMR (supplemental Table S2). HPLC UV chromatogram (370 nm) peaks that did not correspond to a known standard were confirmed as the monoglucosides of the respective aglycone by electrospray ionization mass spectrometry (data not shown). The assignment of glycosylation position to apigenin, kaempferol, and 3’/4’-tri-hydroxyflavone (THF) 4’-O-glucosides and THF 3’-O and 7-O-glucosides was based on comparison of HPLC spectra to known standards and by the comparison of related flavonoid product elution proﬁles.

RESULTS
Regiospecifity of UGT74F1 and UGT74F2 toward Quercetin—A. thaliana GTs UGT74F1 and UGT74F2 share 76% sequence identity and 90% similarity at the amino acid level and are predicted to share an identical secondary structure (Fig. 1). The activities of the two GTs toward the flavonoid quercetin...
Regiospecific Glucosylation by Two *A. thaliana* GTs

When expressed in *E. coli* and assayed as purified recombinant GST fusion proteins (Fig. 2B), UGT74F1 forms monoglucosides on the 7-, 3′-, and 4′-hydroxyls (Fig. 2C), whereas UGT74F2 forms only the 7- and 4′-monoglucosides (Fig. 2D). Variation in amino acid sequence between the enzymes is sufficient to produce different product formation profiles. Thus, to identify crucial amino acids responsible for the differing regiospecificity between the two GTs, segments of sequence were swapped between the two open reading frames.

**Kinetic Characterization of UGT74F1 and UGT74F2 Activities toward Quercetin**—Conditions for the enzyme reactions were established to meet the criteria for pseudo-single substrate kinetics. The data for UGT74F1 and UGT74F2 were analyzed through the use of Hanes plots (Fig. 2, E and F, respectively). The kinetic parameters for the individual glucosides formed were determined when first, quercetin acceptor concentration was varied (Table 1) and second, UDPG donor concentration was varied (Table 2). The *k*ₘ and *Kₘ* values obtained by changing the acceptor concentration explain the observed difference in the UGT74F1 and UGT74F2 product formation profile. It is clear that UGT74F1 and UGT74F2 differ in both enzyme turnover and *Kₘ*. In terms of *k*ₘ, UGT74F1 displays a higher turnover toward the 7-OH and 4′-OH of quercetin than toward the 3′-OH; by comparison, UGT74F2 shows very low turnover toward the 7-OH and 4′-OH positions, with no activity toward the 3′-OH (Table 1). For both the GTs the *Kₘ* for UDPG was near identical, irrespective of the acceptor hydroxyl glycosylated (Table 2).

**Construction and Thermal Denaturation Analysis of UGT74F1/UGT74F2 Chimeras**—As a step to identifying the amino acid determinants of regioselective glycosylation of quercetin within the sequences of UGT74F1 and UGT74F2, a series of...
TABLE 1
The kinetic parameters $k_{cat}$ and $K_m$ and the specificity constant ($k_{cat}/K_m$) of UGT74F1, UGT74F2, and UGT74F1 N142Y toward quercetin, with a fixed UDPG concentration

| Quercetin    | UGT74F1 | UGT74F2 | UGT74F1 N142Y |
|--------------|---------|---------|---------------|
| 7-OH         |         |         |               |
| $K_m$ (μM)   | 1074 ± 126 | 554 ± 141 | 1328 ± 188 |
| $k_{cat}$ (s$^{-1}$) | 0.54 ± 0.06 | 0.11 ± 0.004 | 0.026 ± 0.002 |
| $k_{cat}/K_m$ (m$^{-1}$ s$^{-1}$) | 509 ± 123 | 205 ± 64 | 19.6 ± 1.2 |
| 4′-OH        |         |         |               |
| $K_m$ (μM)   | 1202 ± 205 | 507 ± 131 | 1100 ± 435 |
| $k_{cat}$ (s$^{-1}$) | 0.46 ± 0.05 | 0.08 ± 0.004 | 0.020 ± 0.002 |
| $k_{cat}/K_m$ (m$^{-1}$ s$^{-1}$) | 388 ± 27 | 166 ± 38 | 19.4 ± 6.3 |
| 3′-OH        |         |         |               |
| $K_m$ (μM)   | 1185 ± 158 | 0.25 ± 0.03 | 2.12 ± 1 |
| $k_{cat}$ (s$^{-1}$) | 212 ± 1 | 1 | 1 |

TABLE 2
The kinetic parameters $k_{cat}$ and $K_m$ and the specificity constant ($k_{cat}/K_m$) of UGT74F1, UGT74F2, and F22221 toward UDPG, with a fixed quercetin concentration

| Quercetin | UGT74F1 | UGT74F2 | F22221 |
|-----------|---------|---------|--------|
| $K_m$ (μM) | 10.6 ± 0.6 | 55.2 ± 4.8 | 15.5 ± 3.1 |
| $k_{cat}$ (s$^{-1}$) | 0.99 ± 8.4 × 10$^{-2}$ | 0.04 ± 8.8 × 10$^{-3}$ | 0.07 ± 4.1 × 10$^{-3}$ |
| $k_{cat}/K_m$ (m$^{-1}$ s$^{-1}$) | 93,297 ± 3,230 | 1,030 ± 172 | 4,519 ± 1,100 |

Seven chimeras were constructed. This was achieved by chimeric PCR and through the introduction of three unique restriction enzyme sites (see supplemental Fig. S2 for further details). The modified parental open reading frames were designated UGT74F1* and UGT74F2* and contained four “shuffle points” from which chimeras with sequence switched at amino acid positions 67, 112, 153, and 239 were constructed (Fig. 3A). These were purified as recombinant GST fusion proteins (Fig. 3B). Positions 67, 112, and 153 correspond to locations within the predicted N-terminal Rossmann fold-like domain that has been shown to be primarily responsible for acceptor interaction. Position 239 corresponds to a location within the predicted linker region connecting the two Rossmann fold-like domains.

Thermal denaturation profiles of recombinant protein without the GST fusion tag were obtained for UGT74F1, UGT74F2, and chimeras F22221 and F22211, which displayed the greatest modifications. As shown in Fig. 3C the four profiles were near superimposable, implying the thermal stability of the chimeras was not altered.

Kinetic Characterization of UGT74F1*, UGT74F2*, and Chimeric Sequences—The kinetic parameters are provided in Fig. 4 in which $k_{cat}$, $K_m$, and $k_{cat}/K_m$ are illustrated. The data confirm that introduction of the shuffle points into UGT74F1 and UGT74F2 sequences does not significantly alter the enzymic kinetic characteristics; compare data described in Fig. 4 with Table 1.

The chimeric enzymes display a range of alterations in $k_{cat}$ and $K_m$ toward the hydroxyl groups of quercetin. Chimera F22221, in which the entire N-terminal domain of UGT74F2* is attached to the C-terminal domain of UGT74F1, displays UGT74F2-like $k_{cat}$ and does not glucosylate the 3′-OH of quercetin. Similarly, chimera F22211 does not possess quercetin 3′-OH GT activity. However, this chimera shows a marked increase in $k_{cat}$ at the 4′-OH compared with UGT74F2* and F22221. Chimera F22211 does not possess an increased 4′-OH $k_{cat}$ and also forms quercetin 3′-O-glucoside in addition to the 7-O- and 4′-O-glucosides. When chimera F11211 was analyzed it was found that, compared with UGT74F1*, the $k_{cat}$ toward the 7-OH was significantly reduced (10×) whereas the high $k_{cat}$ toward the 4′-OH was retained. However, swapping the remaining individual segments of UGT74F2* sequence into UGT74F1* (F11121, F12111, and F21111) reduced, by varying degrees, the $k_{cat}$ toward all hydroxyl groups glycosylated (Fig. 4A).

In terms of $K_m$, a number of changes were observed in the different chimeras (Fig. 4B). In particular, F11211 illustrates that the $K_m$ in the glucosylation of the 4′-OH of quercetin can be altered (decreased) while $k_{cat}$ remains unchanged, compared with UGT74F1*. Also, for the same chimera to glucosylate the 7-OH of quercetin, the $K_m$ remains unchanged while $k_{cat}$ is reduced significantly.

The $k_{cat}/K_m$ values for UGT74F1*, UGT74F2*, and chimeric enzymes provide a means for direct comparison of GT specificity for the flavonoid quercetin. Fig. 4C shows that the specificity constant for the three hydroxyl groups of quercetin varies significantly between the different enzymes analyzed. The greatest positive shift was observed in the glucosylation of the 4′-OH of quercetin by chimera F11211. Thus, the domain-swapping strategy identified a 40-amino acid region (residues 112–153 in UGT74F1) capable of increased selectivity of catalysis toward the 4′-OH of quercetin. Further interrogation of the amino acids in this region of the protein enabled identification of the specificity determinant.
Analysis of Individual Amino Acid Changes on the Specificity of GT Activity toward Quercetin—Site-directed mutagenesis was used to determine amino acid positions in segment 112–153 that were critical for shifting specificity. In Fig. 5A, the amino acid sequences of UGT74F1 and UGT74F2 within that segment are shown, illustrating the ten amino acids that differ between the two enzymes. When two of those amino acids in UGT74F1, Ser-135 and Asn-142, were mutated to their UGT74F2 equivalents, proline and tyrosine, respectively, the combination mutant UGT74F1 S135P,N142Y possessed a product formation profile similar to F11211. Analysis of the single mutants showed that the N142Y mutation was sufficient to amplify the observed shift in specificity toward the 4′-OH of quercetin (Fig. 5B). Kinetic analysis of UGT74F1 N142Y revealed that the primary kinetic cause of the altered product formation profile was a reduction in turnover at the 7-OH position with a co-concomitant increase in turnover at the 4′-OH of quercetin (Table 1). Further, although the \( K_m \) of UGT74F1 N142Y toward the 7-OH position was not significantly altered with respect to wild type, the \( K_m \) was reduced 3-fold toward the 4′-OH of quercetin. Fig. 5B also shows that mutations at UGT74F1 Asn-142 to a representative positively charged (Arg), negatively charged (Asp), aliphatic (Leu), or hydrogen-bonding
(Ser) amino acid generated a quercetin product formation profile similar to that of F11211.

Activity toward Additional Flavonoid Substrates—Given that the UGT74F1 N142Y mutant displayed a significantly altered glucosylation profile of quercetin, a range of additional flavonoids were assayed as acceptors to gain further insights into the interaction between substrate hydroxyls and enzyme. In addition, mutants UGT74F1 N142A/D/R/L/S were assayed against the same substrates to determine the potential role of the amino acid side chain at UGT74F1 position 142. Data for N142Y are illustrated in Fig. 6 for activities toward six additional flavonoids. Data for the remaining mutants are included in supplemental Fig. S3.

Few general conclusions can be drawn from a comparison of the glycosylation profiles of the different flavonoids by UGT74F1 N142Y and the other Asn-142 mutants. However, certain trends did emerge. For example, on all flavonoids assayed, UGT74F1 N142Y showed a decreased level of 7-O-glucoside product. The increased production by UGT74F1 N142Y of a 4-O-glucoside of quercetin was similarly observed for the luteolin, kaempferol, and apigenin acceptors. For the two flavonoids, fisetin and 3’4’7-trihydroxflavone, that do not have a hydroxyl at the 5 position no increase in the level of 4’-O-glucoside was observed. No activity toward morin was found for any enzyme assayed.

Production of Quercetin 4’-O-Glucoside—UGT74F1 N142Y was assessed in a non-optimized E. coli shake flask fermentation for utility in preparative scale synthesis of quercetin 4’-O-glucoside. The GT culture converted quercetin to its 4’-O-glucoside at a linear rate for ~36 h (Fig. 7A). During the fermentation, other glucosides represented <5% of the total glucosylated product (data not shown), indicating that the in vitro specificity of the enzyme is mirrored in vivo. Quercetin and its glucosides were recovered from the culture medium by concentration on an Amberlite XAD-2 column and then separated by HPLC to produce quercetin 4’-O-glucoside. The final amount of purified product was 9.9 ± 0.3 mg/liter culture, a yield of ~17% (Fig. 7).

DISCUSSION

Plants contain a large multigene family of GTs capable of conjugating a diverse range of small lipophilic molecules. These enzymes are classified into Family 1 in the Carbohydrate Active Enzyme (CAZy) data base (27). There is considerable potential for using plant GTs as novel biocatalysts for regiospecific glycosylation and production of high value glycosides. However, in order to realize the full potential of GTs in biocatalysis, it is important to gain a thorough understanding of the kinetic parameters that determine their activity. In this work, we have established an HPLC-based methodology for determining kinetic parameters of the multiple concurrent reactions of a GT toward a single substrate. Using this approach we define the kinetic basis for the differing regiospecificity of two A. thaliana GTs, UGT74F1 and UGT74F2. Further, through a domain-swapping approach we have identified an amino acid distal to the active site that is important for determining the regiospecificity of UGT74F1. Taken together, these data inform both the future design and the analysis of GT engineering experiments.

Several earlier studies have explored the regiospecificity of GTs toward quercetin and other flavonoids (22, 23, 28). However, in each of these, although regiospecificity of products was identified, only a single kinetic parameter was assigned and the contribution of individual reactions could therefore not be disconnected. These studies have involved both individual GTs from different plant species and mutant forms of a single GT. In the latter, He and co-workers identified mutations in a Medi-
Regiospecific Glucosylation by Two A. thaliana GTs

cago truncatula GT, UGT71G1, that significantly shifted the regiospecificity of glycosylation of quercetin (23). Whereas the parental enzyme formed all five potential quercetin glucosides, the F148V mutation was shown to focus activity to the formation of only quercetin 3-O-glucoside. By contrast, all of the other active GT mutants formed at least four quercetin glucosides, each in varying proportion to the parent enzyme. The kinetic approach used in this study would enable assignment of the kinetic parameters $k_{\text{cat}}$ and $K_m$ to each of the glycosylating reactions performed, thereby providing additional quantitative data for interpretation of observed alterations in regiospecificity.

All Family 1 GTs are predicted to exhibit a GT-B fold comprised of two Rossmann fold-like domains. To date, crystallographic studies of GTs with a low comparative amino acid identity ($\sim 20\% - 30\%$) have shown strong structural conservation of the core $\beta/\alpha/\beta$ Rossmann domains. There is also functional conservation with respect to the interactions of N- and C-terminal domains with the acceptor or sugar donor. Typically, amino acids interacting with the acceptor are located within the N-terminal domain and those interacting with the sugar donor are found in the C-terminal domain. The resulting prediction of a discrete function to a specific domain makes domain swapping an enticing approach for the engineering of GTs. Hoffmeister et al. (29) identified a variable region of amino acid sequence between the urdaminic GTs UrdGT1b and UrdGT1c (91% identity) responsible for differentiating the activity of the two enzymes. By domain swapping between two GTs of 85% identity, Brazier-Hicks et al. (5) found a region of five variable amino acids responsible for enabling the formation of N-glucosides in addition to O-glucosides. Other studies (30, 31) found that at lower levels of sequence identity (75 and 28%, respectively) the generated chimeras were inactive.

The data we present demonstrate the dual benefits of a domain-swapping approach. First, the predicted transfer of a property associated with a whole domain is illustrated clearly by the chimera F22221. In this instance, the whole N-terminal domain of UGT74F2 is fused to the C-terminal domain of UGT74F1 and the chimera displays UGT74F2-like kinetic parameters and regiospecificity toward the quercetin acceptor. As anticipated, the F22221 chimera also displays UGT74F1-like kinetic parameters toward the UDPG donor. Second, as exemplified by another chimera, F11211, an unexpected property was revealed. In this instance, the regiospecificity of the chimera toward the 4'-OH of quercetin was significantly increased relative to the UGT74F1 parent. It will be interesting to explore the limits of GT structural modularity in the transfer of enzymic attributes. This would involve the discovery of enzymes with novel activities such as through the recombination of known activities for a designed outcome. Also, domain-swapping GTs of decreasing amino acid identity at either the Rossmann fold-like domain or structural element level could lead to the discovery of additional unexpected activities.

Many enzyme recombination strategies target active site and linking loops as a means of limiting disruption to the core secondary structure (32). The active site of GT-B fold enzymes is principally comprised of loop regions (33) and is therefore an intuitive starting point for mutagenesis. However, in UGT74F1 the amino acid mutation N142Y that was shown to be responsible for a drastic alteration in $k_{\text{cat}}/K_m$ is predicted to lie within helix No4. A superimposition of the four available plant GT structures clearly showed that the amino acids corresponding to Asn-142 occupy equivalent structural positions in helix No4 (UGT74F1 numbering), but no direct effects on active site residues or on substrate interaction could be directly deduced from their structural position (supplemental Fig. S4). As others have found, amino acid mutations without a direct substrate or catalytic residue(s) interaction can have pronounced effects on GT activity (5, 34). Only the crystallographic three-dimensional structures of UGT74F1 and the N142Y mutant would reveal the structural mechanism by which the regiospecificity of the enzyme is altered.

A recent study, using a high-throughput screen to detect novel donor and acceptor activities, revealed the surprising ability of a Humicola insolens Cell7B glycosynthase mutant (35) to glycosylate flavonoid scaffolds (36). Interestingly, the glycosynthase showed a >95% regiospecificity toward the 4'-OH of quercetin using the activated donor lactosyl fluoride. While glycosynthases have provided powerful tools for biocatalytic synthesis of carbohydrates (37), the study of the Cell7B mutant enzyme demonstrated a potential wider utility. However, in terms of preparative scale synthesis of the 4'-O-glucoside, the GT mutant described in this study provides the added advantage of a whole-cell biocatalysis strategy. This strategy has been
successfully demonstrated for a diverse range of quercetin glycosides and those of other natural product scaffolds (3, 38). Also, in the whole-cell system, a range of sugar donors can be provided by engineered microbial hosts, bypassing the need to supply exogenous cofactor required by in vitro reactions (39–41). When assessed in a non-optimized shake flask fermentation for utility in preparative scale synthesis, E. coli expressing UGT74F1 N142Y converted quercetin aglycone to quercetin 4′-O-glucoside in substantial quantities. Significantly, conjugation of the 4′-position, compared with that of others, has been shown to influence bioactivity of quercetin when assayed in human metabolism studies (42) using in vivo models (20, 21, 43) and in vitro assays (44).

In summary, we have assigned kinetic parameters to the multiple glycosylation reactions of a plant GT toward a single substrate. Further, a domain-swapping strategy has been used to identify an amino acid position in UGT74F1 that is important for determining specificity toward quercetin and other flavonoids. Application of the UGT74F1 N142Y mutant in an E. coli whole-cell fermentation demonstrated the potential utility of this GT for the production of flavonoid 4′-O-glucoside.

Acknowledgments—We thank Dr. Fabian E. Vaistij and Professor Tony Wilkinson for helpful discussions.

REFERENCES
1. Bowles, D., Lim, E. K., Poppenberger, B., and Vaistij, F. E. (2006) Annu. Rev. Plant Biol. 57, 567–597
2. Lim, E. K. and Bowles, D. J. (2004) EMBO J. 23, 2915–2922
3. Lim, E. K., Ashford, D. A., Hou, B. K., Jackson, R. G., and Bowles, D. J. (2004) Biotechnol. Bioeng. 87, 623–631
4. Lim, E. K., Doucet, C. J., Hou, B., Jackson, R. G., Abrams, S. R., and Bowles, D. J. (2005) Tetrahedron Asymmetry 16, 143–147
5. Brazier-Hicks, M., Offen, W. A., Gershater, M. C., Revett, T. J., Lim, E. K., Bowles, D. J., Davies, G. J., and Edwards, R. (2007) Proc. Natl. Acad. Sci. U. S. A. 104, 20238–20241
6. Lim, E. K., Baldauf, S., Li, Y., Elias, L., Worrall, D., Spencer, S. P., Jackson, R. G., Taguchi, G., Ross, J., and Bowles, D. J.(2003) Glycobiology 13, 139–145
7. Hansen, K. S., Kristensen, C., Tattersall, D. B., Jones, P. R., Olsen, C. E., Bak, S., and Moller, B. L. (2003) Phytochemistry 64, 143–151
8. Hofmeister, D., Wilkinson, B., Foster, G., Sidebottom, P. J., Ichinose, K., and Bechthold, A. (2002) Chem. Biol. 9, 287–295
9. Kubo, A., Arai, Y., Nagashima, S., and Yoshikawa, T. (2004) Arch. Biochem. Biophys. 429, 198–203
10. Kama, P., Gokhale, R. S., and Mohanty, D. (2005) Nucleic Acids Res. 33, W220–W225
11. Thorsen, A. E., Kohn, C., Sakhrom, D., Prata, R. T., Willits, M. G., De Luca, V., Steffens, J. C., and Graser, G. (2003) Phytochemistry 64, 1069–1076
12. Modolo, L. V., Blount, J. W., Achrine, L., Naoumikina, M. A., Wang, X. Q., and Dixon, R. A. (2007) Plant Mol. Biol. 64, 499–518
13. Hall, D., and De Luca, V. (2007) Plant J. 49, 579–591
14. Hollman, P. C. H., van Trijffel, J. M. P., Buysman, M. N. C. P., VanderGaag, M. S., Mengelers, M. J. B., de Vries, J. H. M., and Katan, M. B. (1997) FEBS Lett. 418, 152–156
15. van der Woude, H., Boersma, M. G., Vervoort, J., and Rietjens, I. M. C. M. (2004) Chem. Res. Toxicol. 17, 1520–1530
16. Hu, R., Xue, T., and Pei, J. (2003) Biochem. Biophys. Res. Commun. 314, 342–346