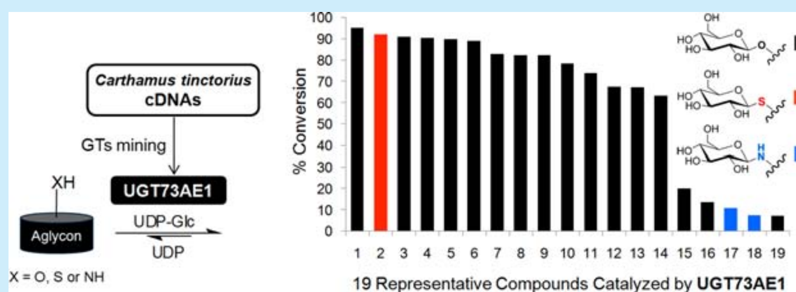
Organic LETTERS
Letter
pubs.acs.org/OrgLett

Exploring the Catalytic Promiscuity of a New Glycosyltransferase from *Carthamus tinctorius*

Kebo Xie, Ridao Chen, Jianhua Li, Ruishan Wang, Dawei Chen, Xiaoxiang Dou, and Jungui Dai*

State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, 1 Xian Nong Tan Street, Beijing 100050, P.R. China

Supporting Information



ABSTRACT: The catalytic promiscuity of a new glycosyltransferase (UGT73AE1) from *Carthamus tinctorius* was explored. UGT73AE1 showed the capability to glucosylate a total of 19 structurally diverse types of acceptors and to generate *O*-, *S*-, and *N*-glycosides, making it the first reported trifunctional plant glycosyltransferase. The catalytic reversibility and regioselectivity were observed and modeled in a one-pot reaction transferring a glucose moiety from icariin to emodin. These findings demonstrate the potential versatility of UGT73AE1 in the glycosylation of bioactive natural products.

Sugar moieties, as part of many bioactive natural products, have important effects on the physiological activity, selectivity, and pharmacological properties of these products.[1] Although enzymatic glycosylation engineering has been applied to attach or alter the sugar moieties of numerous natural products,[2] the process remains primarily restricted by enzyme specificity and the availability of suitable glycosyltransferases (GTs) for the targets of interest.[3] GTs with substrate promiscuity are generally recognized as powerful tools in the glycodiversification of natural products for both *in vitro* and *in vivo* use.[3] Thus, the discovery of GTs with catalytic promiscuity and novel specificity is of necessity in practice, and great progress has been achieved, especially for GTs from microbes.[2a,c,d,3b−d] However, in the past few years, studies on plant GTs have revealed that dramatically varied GTs are involved in plant secondary metabolism;[4] e.g., the model plant *Arabidopsis thaliana* contains >100 GTs dedicated to small-molecule conjugation.[4a] Furthermore, plant GTs have advantages in glycosylating botanical natural products, derivatives of which constitute a large proportion of clinical drugs.[5] However, most of the reported plant GTs show relatively narrow substrate spectra and few plant GTs have been used as enzymatic tools for the glycosylation of natural products with structural diversity. Thus, mining GTs from plants with catalytic promiscuity is important to synthesize bioactive natural product glycosides. Here, we report a new glycosyltransferase (UGT73AE1) from *Carthamus tinctorius*, which can tolerate a number of structurally different acceptors, form *O*-, *S*-, and *N*-glycosidic bonds, and catalyze the reverse reaction, along with its use in a deglucosylation reaction and an aglycon exchange reaction.

C. tinctorius (L.) (Honghua) is a traditional Chinese medicinal herb, and a wide variety of bioactive natural glycosides, including the rare *C*-glycosylated quinochalcones and *O*-glycosylated flavonoids, have been isolated from its florets.[6] The diverse glycosylated secondary metabolites imply the existence of corresponding GTs, which inspired us to seek GTs with interesting substrate specificity. To clone permissive GTs from *C. tinctorius*, a degenerate PCR primer (GT-5′ RACE, Table S2) for 5′ RACE was designed based on the conserved PSPG (Plant Secondary Product Glycosyltransferases) motif of five plant species GTs (Figure S1).[7] Combined with 3′ RACE, nine new *C. tinctorius* GTs (CtGTs) were successfully cloned by RT-PCR amplification using the total RNA from *C. tinctorius* florets as a template and heterologously expressed in *Escherichia coli* as described in the Supporting Information (SI).

To investigate the glycosylation capability of the CtGTs *in vitro*, UDP-glucose (UDPG) along with phloretin (**9**) and resveratrol (**11**), which are usually used as acceptors in glycosylation by GTs involved in the plant secondary metabolism,[8] were used in the enzymatic assay. The detecting reactions (50 mM Tris-HCl, pH 7.4; 0.5 mM UDPG; 0.25 mM aglycon; 500 μg of crude CtGTs; 30 °C, 6 h) were analyzed by HPLC-UV/MS (high-performance liquid chromatography-UV

Received: August 11, 2014
Published: September 5, 2014

 © 2014 American Chemical Society

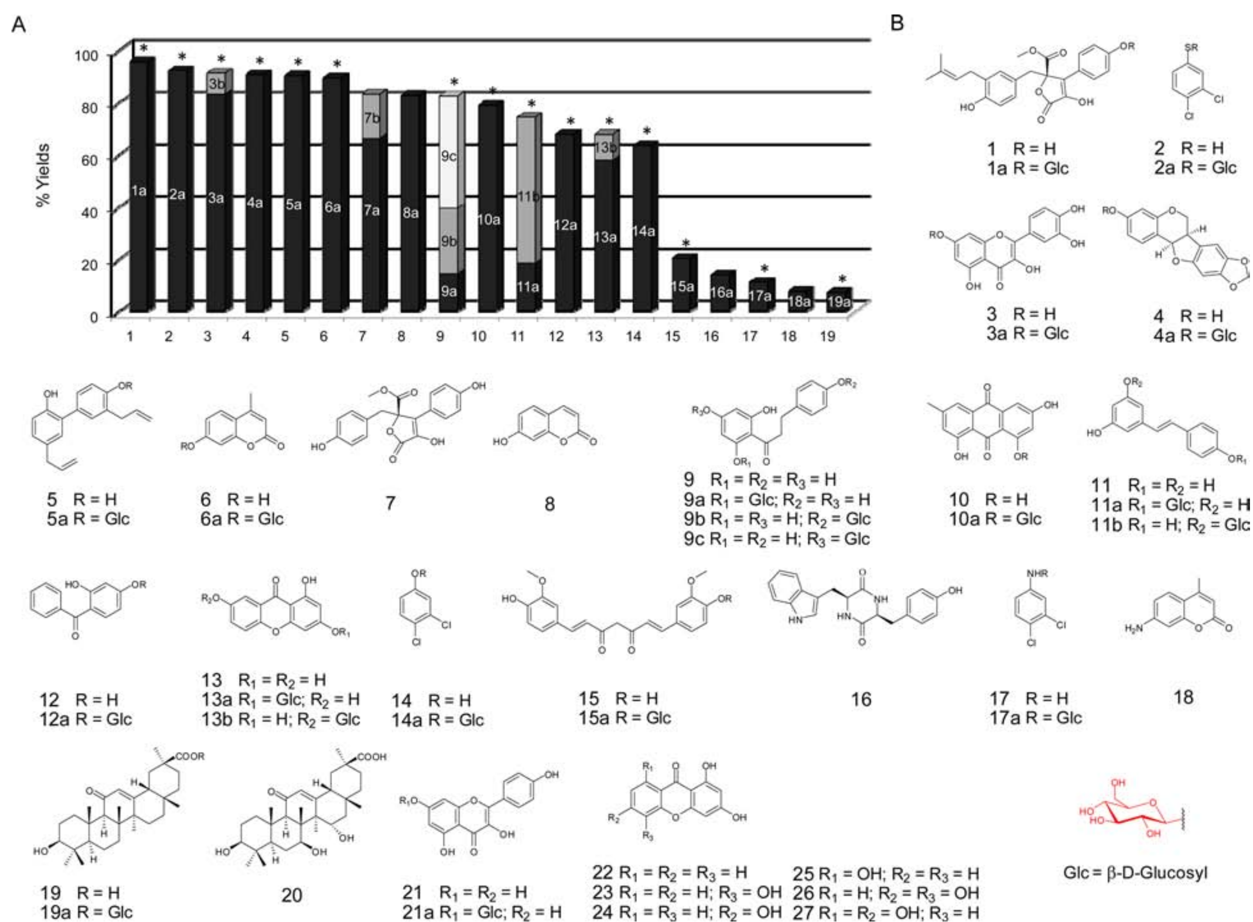


Figure 1. Exploring the catalytic promiscuity of UGT73AE1. (A) Percent yields of glucosylated products catalyzed by UGT73AE1. Members (**1**–**19**) are listed in descending order of yields with numbering corresponding to the structures listed in part B. The black, gray and white columns represent the conversion of each glucosylated product. (B) Structures of the library members and corresponding glucosylated products. The percent yields of compounds **20**–**27** are shown in the SI. "*" represents the glucosylated products (including **21a**) that were isolated and confirmed by MS and ^{1}H, ^{13}C NMR. **3b** was not isolated due to the low yield.

absorption/mass spectrometry). Of the nine recombinant CtGTs, only CtGT6 showed glycosylation activity to all of the hydroxyls of phloretin (**9**) and resveratrol (**11**) with high conversion rates (Figure 1). Control reactions lacking either enzyme or UDPG confirmed that the reactions were dependent upon both the enzyme and UDPG. The glucosylated products were isolated from the preparative-scale reactions, and their structures, including the glucosylated positions and anomeric stereochemistry, were characterized by MS, ^{1}H NMR, and ^{13}C NMR spectroscopic data analyses compared with reported data (SI). The observed large anomeric proton-coupling constants (J = 6.8–7.8 Hz) indicated the formation of the β-anomers and an inverting mechanism for UGT73AE1. The cDNA sequence of *CtGT6* (1476 bp, GenBank accession number KJ956788) contained an ORF encoding 491 amino acids, and this permissive GT was named UGT73AE1 according to the UGT Naming Committee.[9] UGT73AE1 showed the highest identity (56%) to 73C2, a predicted UDP-glycosyltransferase from *Vitis vinifera*. Purification of His$_6$-UGT73AE1 was accomplished by His-tag affinity chromatography, and it was analyzed by SDS-PAGE (Figure S2). The biochemical characteristics of purified UGT73AE1 were determined and shown in the SI (Figures S3 and S4). Unlike those in microorganisms, the native substrates of GTs in plants are difficult to identify because of the catalytic promiscuity of GTs *in vitro*, the limitation of the gene knockout technique, the huge number of GTs in a plant, and the long growth cycles of plants. Given the high identities to the reported flavonoid GTs (Figure S1) and the glucosylation activity to quercetin (**3**) and kaempferol (**21**), of which glucosylated derivatives exist in *C. tinctorius*, UGT73AE1 can be tentatively assigned as a flavonoid GT.

GTs usually possess high substrate specificity *in vivo*. For GT-mediated reactions that are performed *in vitro*, however, the enzymes typically tolerate a broader set of substrates and have therefore been useful biocatalysts in the synthesis of various glycosylated derivatives.[10] Thus, to explore the catalytic promiscuity and probe the synthetic utility of UGT73AE1 *in vitro*, a series of representative drug-like compounds with structural diversity were collected. Thus, a compound library including lignans (**1**, **5**, and **7**), flavonoids (**3**, **4**, **9**, and **21**), coumarins (**6** and **8**), anthraquinone (**10**), stilbene (**11**), benzophenone (**12**), xanthones (**13** and **22**–**27**), curcuminoid (**15**), cyclopeptide (**16**), triterpenes (**19** and **20**), and simple aromatics with various nucleophilic groups of –SH, –OH, and –NH$_2$ (**2**, **14**, **17**, and **18**) was employed for enzymatic assays (Figure 1).

From the first-pass analysis with HPLC-UV/MS, enzyme-catalyzed glucosylations of 27 members, including 19 structurally different types, were observed (Figures 1 and

S5–S9). UGT73AE1 provided high conversion (>80%) with 14 (**1**–**9**, **21**, **23**, **24**, **26**, and **27**) out of the 27 substrates. It is important to note that, among the subgroup of library members containing multiple nucleophiles (19 members), nine (**1**, **5**, **10**, **12**, **15**, **19**, **20**, **22**, and **25**) led to a single, chromatographically distinct, monoglucosylated product individually, indicating the regiospecificity of UGT73AE1. Twenty glucosylated products of 16 aglycons were isolated from the preparative-scale reactions, four (**1a**, **2a**, **5a**, and **13b**) of which were novel compounds. The structures were identified by MS and NMR spectroscopic data analysis, and all of the anomers were in the β configuration due to anomeric protons with large coupling constants (J > 6.0 Hz, Table S5). To our surprise, UGT73AE1 could also transfer a glycosyl moiety to the aliphatic carboxyl group of glycyrrhetinic acid (**19**) leading to a glucosyl ester, different from the glucosylating pattern for the other aromatic phenolic substrates. Most of the glucosylated derivatives above are reported to exhibit better druggabilities than the corresponding substrates. For example, the regioselective glucosylation of emodin (**10**) at the 8-OH enhanced the water solubility and even the purgative action.[11]

Interestingly, UGT73AE1 showed *O*-, *S*-, and *N*-glycosylation activities to the simple aromatics simultaneously *in vitro* (Figures 2 and S7). NMR characterizations of the products (**2a**,

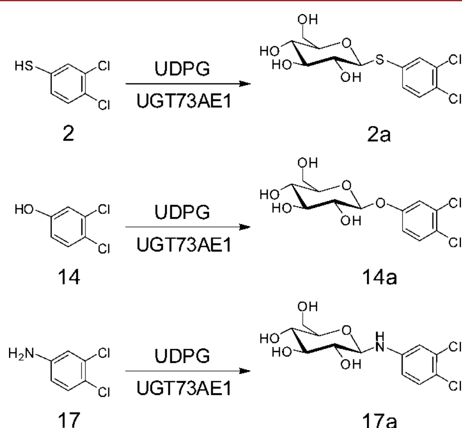


Figure 2. UGT73AE1-catalyzed the *S*-, *O*-, and *N*-glycosidic bond formation. The 3,4-dichlorobenzenethiol (**2**), 3,4-dichlorophenol (**14**), and 3,4-dichloroaniline (**17**) were used as acceptors, and UDPG was used as a donor. The assay conditions and HPLC chromatograms (Figure S10) are available in the SI.

14a, and **17a**) prepared from the preparative-scale reactions were consistent with the β-*S*-, *O*-, and *N*-glucosides (J = 9.6, 6.8, and 8.4 Hz for anomeric protons, respectively). Therefore, UGT73AE1 is a "trifunctional" *O*-, *S*-, and *N*-glycosyltransferase. Although the only example of engineered GTs with *O*-, *S*-, and *N*-glycosylation activity has been obtained from microbes via the directed evolution,[12,3b] to the best of our knowledge, UGT73AE1 is the first reported plant GT capable of catalyzing an *O*-, *S*-, and *N*-glycosidic bond formation.

Generally, GTs are perceived as unidirectional catalysts that drive the formation of glycosidic bonds from NDP-sugar donors and aglycon acceptors.[13] However, since the initial reports,[14,2c] a number of GTs were observed to catalyze reversible, bidirectional reactions, which could be employed for deglycosylation in certain cases or to transfer sugars from glycoside scaffolds to aglycons.[2e,15] Removal of a glucose unit from the 7-*O* position of icariin (**28**) affords the molecule baohuoside I (**28a**) with more permeability and better bioavailability.[16] Thus, to obtain **28a** via regiospecific deglucosylation of **28** as well as explore the reversibility of UGT73AE1, **28** and UDP (molar ratio 1:5) were exemplified as substrates in the enzymatic assay. After being incubated at 30 °C for 12 h, **28a** was detected in 38% yield by analytical HPLC (Figure S11A). The equilibrium constant K_{eq} = 7.5 (30 °C, pH 7.4) was consistent with an equilibrium only moderately favoring the glycoside formation in the UGT73AE1-catalyzed reaction of **28** and UDP.

Because UDPG was an expensive "waste" in the reaction, the reverse reaction and the emodin (**10**) glycosylation reaction were coupled together as a model, as shown in Figure 3. In this

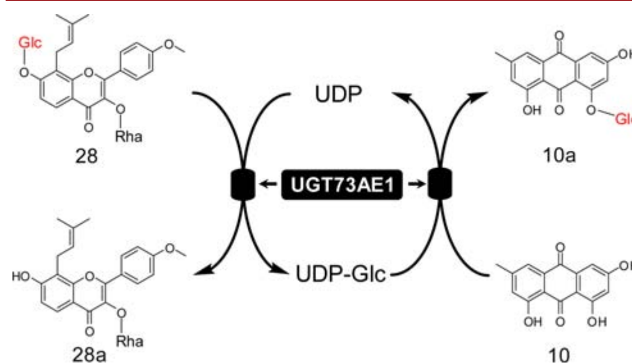


Figure 3. Catalytic reverse reaction and coupled glycosyltransferase-mediated transglucosylation with UGT73AE1. The one-pot reaction combines a selective deglucosylation reaction for nucleotide glucose synthesis with a subsequent regioselective glucosylation to ultimately generate two targeted products, baohuoside I (**28a**) and emodin 8-*O*-glucoside (**10a**), in this case. The assay conditions and HPLC chromatograms (Figure S11) are available in the SI.

"one-pot" reaction, UDP is a double-edged sword, which means that a high concentration of UDP should boost deglucosylation (yielding **28a**), but disfavor glucosylation (affording **10a**) (Figures 3 and S12). The optimized reaction (50 mM Tris-HCl buffer, pH 7.4; 1 mM DTT; 50 μM icariin (**28**); 2.5 μM UDP; 25 μM emodin (**10**); 1 mg UGT73AE1; 30 °C for 12 h) was analyzed by analytical HPLC (Figure S11). Interestingly, the two desired products (**28a** and **10a**) with yields of 24% and 26%, respectively, were generated in the presence of only a very small amount of UDP (1/10 molar of **10**), which meant that UDP was under cyclic utilization throughout the coupled reactions.[17,2c] Above all, this one-pot reaction, providing two desired products via reusing the byproduct without adding the expensive extra UDPG, was economic and environmentally benign.

In summary, the catalytic promiscuity of UGT73AE1, a new glucosyltransferase, cloned from *C. tinctorius* is highlighted. As a plant GT, UGT73AE1 showed robust glucosylation activity to a series of structurally different drug-like compounds and the capability of catalyzing an *O*-, *S*-, and *N*-glycosidic bond formation, which was reported for the first time in a plant. UGT73AE1 with catalytic promiscuity and reversibility may be exploited as a powerful biocatalyst for the enzymatic synthesis of bioactive glycosides or activated sugars, and the present study may facilitate further enzyme engineering to develop novel biocatalysts. Furthermore, future structural studies of UGT73AE1 would provide structural insight into the enzyme catalytic mechanism and specificity for plant GTs,[18,19] particularly to address how the enzyme controls the reaction

dx.doi.org/10.1021/ol502380p | Org. Lett. 2014, 16, 4874–4877

directions.[20] The UGT73AE1 reported here suggests more exciting, novel GTs hiding in the plant kingdom as enzymatic tools for glycosylation in the search for drug leads.

■ ASSOCIATED CONTENT

Supporting Information

The experimental procedures, including CtGTs cloning, expression, reactions analysis and products purification, HPLC/MS, HRESIMS, and NMR characterization data and spectra of products. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: jgdai@imm.ac.cn.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was financially supported by the Fundamental Research Funds for the Central Universities (Grant No. 2012N06) and the National Science & Technology Major Project 'Key New Drug Creation and Manufacturing', China (No. 2012ZX09301002-001-005).

■ REFERENCES

(1) (a) Weymouth-Wilson, A. C. *Nat. Prod. Rep.* **1997**, *14*, 99–110. (b) Thorson, J. S.; Hosted, T. J.; Jiang, J.; Biggins, J. B.; Ahlert, J. *Curr. Org. Chem.* **2001**, *5*, 139–167. (c) Osbourn, A. E. *Phytochemistry* **2003**, *62*, 1–4. (d) Ahmed, A.; Peters, N. R.; Fitzgerald, M. K.; Watson, J. A.; Hoffmann, F. M.; Thorson, J. S. *J. Am. Chem. Soc.* **2006**, *128*, 14224–14225.

(2) (a) Hoffmeister, D.; Dräger, G.; Ichinose, K.; Rohr, J.; Bechthold, A. *J. Am. Chem. Soc.* **2003**, *125*, 4678–4679. (b) Williams, G. J.; Goff, R. D.; Zhang, C.; Thorson, J. S. *Chem. Biol.* **2008**, *15*, 393–401. (c) Zhang, C.; Griffith, B. R.; Fu, Q.; Albermann, C.; Fu, X.; Lee, I.; Li, L.; Thorson, J. S. *Science* **2006**, *313*, 1291–1294. (d) Zhou, M.; Hou, Y.; Hamza, A.; Zhan, C.; Bugni, T. S.; Thorson, J. S. *Org. Lett.* **2012**, *14*, 5424–5427. (e) Gutmann, A.; Krump, C.; Bungaruang, L.; Nidetzky, B. *Chem. Commun.* **2014**, *50*, 5465–5468.

(3) (a) Thibodeaux, C. J.; Melançon, C. E.; Liu, H. W. *Angew. Chem.* **2008**, *47*, 9814–9859. (b) Gantt, R. W.; Goff, R. D.; Williams, G. J.; Thorson, J. S. *Angew. Chem.* **2008**, *47*, 8889–8892. (c) Williams, G. J.; Yang, J.; Zhang, C.; Thorson, J. S. *ACS Chem. Biol.* **2010**, *6*, 95–100. (d) Zhou, M.; Thorson, J. S. *Org. Lett.* **2011**, *13*, 2786–2788.

(4) (a) Brazier-Hicks, M.; Offen, W. A.; Gershater, M. C.; Revett, T. J.; Lim, E.; Bowles, D. J.; Davies, G. J.; Edwards, R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20238–20243. (b) Gandia-Herrero, F.; Lorenz, A.; Larson, T.; Graham, I. A.; Bowles, D. J.; Rylott, E. L.; Bruce, N. C. *Plant J.* **2008**, *56*, 963–974. (c) Brazier-Hicks, M.; Evans, K. M.; Gershater, M. C.; Puschmann, H.; Stell, P. G.; Edwards, R. *J. Biol. Chem.* **2009**, *284*, 17926–17934. (d) Ferreyra, M. L. F.; Rodriguez, E.; Casas, M. I.; Labadie, G.; Grotewold, E.; Casati, P. *J. Biol. Chem.* **2013**, *288*, 31678–31688.

(5) (a) Wu, C.; Li, T.; Farh, L.; Lin, L.; Lin, T.; Yu, Y.; Yen, T.; Chiang, C.; Chan, N. *Science* **2011**, *333*, 459–462. (b) Ye, R.; Kong, X.; Yang, Q.; Zhang, Y.; Han, J.; Li, P.; Xiong, L.; Zhao, G. *Neurotherapeutics* **2011**, *8*, 515–525. (c) Xu, L.; Tan, X.; Yun, J.; Shen, S.; Zhang, S.; Tu, C.; Zhao, W.; Tian, B.; Yang, G.; Yao, K. *Ind. Chem. Res.* **2012**, *51*, 11373–11380.

(6) (a) Kazuma, K.; Takahashi, T.; Sato, K.; Takeuchi, H.; Matsumoto, T.; Okuno, T. *Biosci. Biotechnol. Biochem.* **2000**, *64*, 1588–1599. (b) Yoon, H.; Han, H.; Paik, Y. *J. Appl. Biol. Chem.* **2007**, *50*, 175–178. (c) Jiang, J.; He, J.; Feng, Z.; Zhang, P. *Org. Lett.* **2010**, *12*, 1196–1199.

(7) (a) Hughes, J.; Hughes, M. A. *DNA Seq.* **1994**, *5*, 41–49. (b) Masada, S.; Terasaka, K.; Mizukami, H. *FEBS Lett.* **2007**, *581*, 2605–2610.

(8) (a) Gosch, C.; Halbwirth, H.; Schneider, B.; Hölscher, D.; Stich, K. *Plant Sci.* **2010**, *178*, 299–306. (b) Terasaka, K.; Mizutani, Y.; Nagatsu, A.; Mizukami, H. *FEBS Lett.* **2012**, *586*, 4344–4350.

(9) (a) Mackenzie, P. I.; Owens, I. S.; Burchell, B.; Bock, K. W.; Bairoch, A.; Bélanger, A.; Fournel-Gigleux, S.; Green, M.; Hum, D. W.; Iyanagi, T.; Lancet, D.; Louisot, P.; Magdalou, J.; Chowdhury, J. R.; Ritter, J. K.; Schachter, H.; Tephly, T. R.; Tipton, K. F.; Nebert, D. W. *Pharmacogenetics* **1997**, *7*, 255–269. (b) Homepage of the UGT Nomenclature Committee, http://www.flinders.edu.au/medicine/sites/clinical-pharmacology/ugt-homepage.cfm.

(10) (a) Koeller, K. M.; Wong, C. H. *Nature* **2001**, *409*, 232–240. (b) Ban, L.; Pettit, N.; Li, L.; Stuparu, A. D.; Cai, L.; Chen, W.; Guan, W.; Han, W.; Wang, P. G.; Mrksich, M. *Nat. Chem. Biol.* **2012**, *8*, 769–773.

(11) Srinivas, G.; Babykutty, S.; Sathiadevan, P. P.; Srinivas, P. *Med. Res. Rev.* **2007**, *27*, 591–608.

(12) Williams, G. J.; Zhang, C.; Thorson, J. S. *Nat. Chem. Biol.* **2007**, *3*, 657–662.

(13) Koeller, K. M.; Wong, C. H. *Chem. Rev.* **2000**, *100*, 4465–4494.

(14) Minami, A.; Kakinuma, K.; Eguchi, T. *Tetrahedron Lett.* **2005**, *46*, 6187–6190.

(15) (a) Gantt, R. W.; Peltier-Pain, P.; Cournoyer, W. J.; Thorson, J. S. *Nat. Chem. Biol.* **2011**, *7*, 685–691. (b) Chen, R.; Zhang, H.; Zhang, G.; Li, S.; Zhang, G.; Zhu, Y.; Liu, J.; Zhang, C. *J. Am. Chem. Soc.* **2013**, *135*, 12152–12155. (c) Zhang, J.; Singh, S.; Hughes, R. R.; Zhou, M.; Sunkara, M.; Morris, A. J.; Thorson, J. S. *ChemBioChem* **2014**, *15*, 647–651.

(16) Chen, Y.; Zhao, Y. H.; Jia, X. B.; Hu, M. *Pharm. Res.* **2008**, *25*, 2109–2199.

(17) (a) Haynie, S. L.; Whitesides, G. M. *Appl. Biochem. Biotechnol.* **1990**, *23*, 155–170. (b) Ichikawa, Y.; Wang, R.; Wong, C. H. *Methods Enzymol.* **1994**, *247*, 107–127.

(18) (a) Shao, H.; He, X.; Achine, L.; Blount, J. W.; Dixon, R. A.; Wang, X. *Plant Cell* **2005**, *17*, 3141–3154. (b) Li, L.; Modolo, L. V.; Escamilla-Trevino, L. L.; Achnine, L.; Dixon, R. A.; Wang, X. *J. Mol. Biol.* **2007**, *370*, 951–963.

(19) Jakeman, D. L. *ChemBioChem* **2011**, *12*, 2540–2542.

(20) Modolo, L. V.; Li, L.; Pan, H.; Blount, J. W.; Dixon, R. A.; Wang, X. *J. Mol. Biol.* **2009**, *392*, 1292–1302.