ChemComm
Chemical Communications

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: N. Wong, S. Zhong, W. Li, F. Zhou, Z. Deng and Y. Zhou, *Chem. Commun.*, 2020, DOI: 10.1039/D0CC04948D.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemcomm

View Article Online
DOI: 10.1039/D0CC04948D

COMMUNICATION

Selective profiling of steviol-catalyzing UDP-glycosyltransferases with a metabolically synthesized probe†

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Nai-Kei Wong,‡[ac] Suyun Zhong,‡[a] Weichao Li,[d] Fugui Zhou,[ab] Zhangshuang Deng,*[ba] and Yiqing Zhou*[ad]

Selective profiling of steviol-catalyzing UDP-glycosyltransferases in plants was accomplished with a probe metabolically synthesized from two substrate-derived components comprising an alkynylated sugar receptor (steviol) module and a diazirine-modified sugar donor (UDP-glucose) module, thereby illustrating a facile approach to harnessing biosynthetic enzymes of natural glycosides in plants for synthetic biology.

Natural products are ubiquitous secondary metabolites found in animals, plants, and microbes. Their extraordinary structural diversity and scope of bioactivities have inspired numerous discoveries of lead compounds and therapeutics in the development pipeline.[1] Synthetic biology has rapidly emerged as a promising approach to producing natural products on an industrial scale, allowing high precision and control over quality without drawbacks inherent in natural extraction or total synthesis.[2] In the case of plant natural products, functional elucidation of biosynthetic genes typically represents a first and rate-limiting step toward practical utilization of biosynthetic pathways.[3] Nevertheless, technical complexities vary in mining biosynthetic pathways across different classes of organisms. In prokaryotes such as bacteria, due to the presence of gene clusters of secondary metabolites, conventional "genome mining" strategies can be readily applied to discover and annotate biosynthetic genes.[4] In contrast, in eukaryotes such as plants, biosynthetic genes are not as compactly clustered as in prokaryotes, making it extra time-consuming to identify the functional enzymes in plants by classical genomic or transcriptomic approaches contingent on biochemically focused characterization.[5] Thus in any plants, the task of determining the biosynthetically pivotal biological parts remains technically daunting.[6] Recent progress in chemical

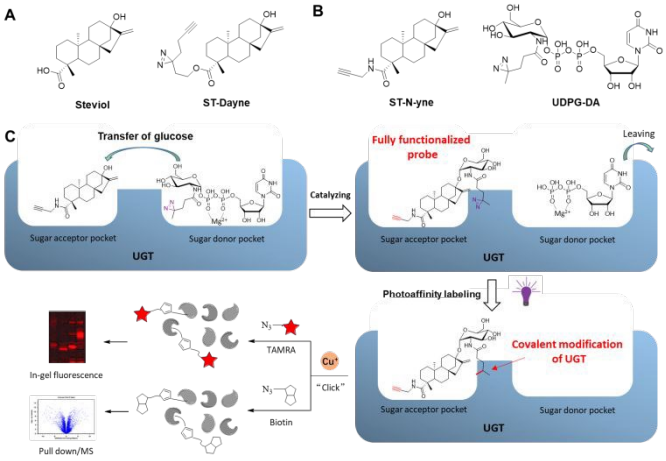
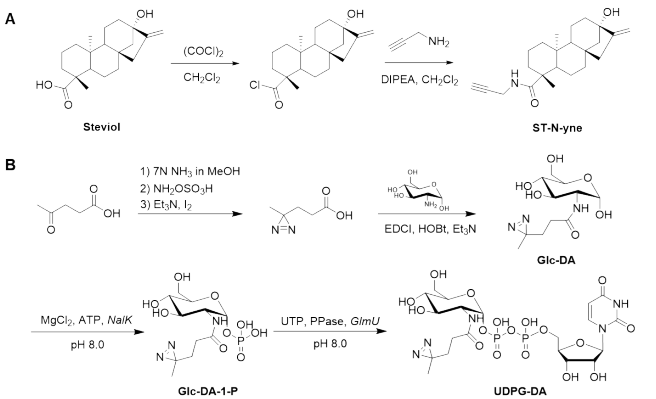


Fig. 1 (A) Structures of steviol and **ST-Dayne**, a previously reported steviol-derived fully functionalized photoaffinity probe used in profiling of the steviol glycosides biosynthesis pathway; (B) Structures of bi-substrate probe modules **ST-N-yne** and **UDPG-DA** used in current study; (C) A bi-substrate probe strategy for identifying biological parts with UGT activities for steviol.



Scheme 1 Synthesis of bi-substrate probe modules (A) **ST-N-yne** and (B) **UDPG-DA**.

[a] *School of Biotechnology and Food Engineering, Changshu Institute of Technology, Changshu 215500, China. *E-mail: zhyq2012@gmail.com*
[b] *Hubei Key Laboratory of Natural Products Research and Development, College of Biological and Pharmaceutical Sciences, China Three Gorges University, Yichang, 443002, China. *E-mail: d.zhangshuang@gmail.com*
[c] *Department of Pharmacology, Shantou University Medical College, Shantou 515041, China.*
[d] *CAS Key Laboratory of Synthetic Biology, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China.*
† Electronic Supplementary Information (ESI) available. See DOI: 10.1039/x0xx00000x
‡ These authors contributed equally to this work.

View Article Online
DOI: 10.1039/D0CC04948D

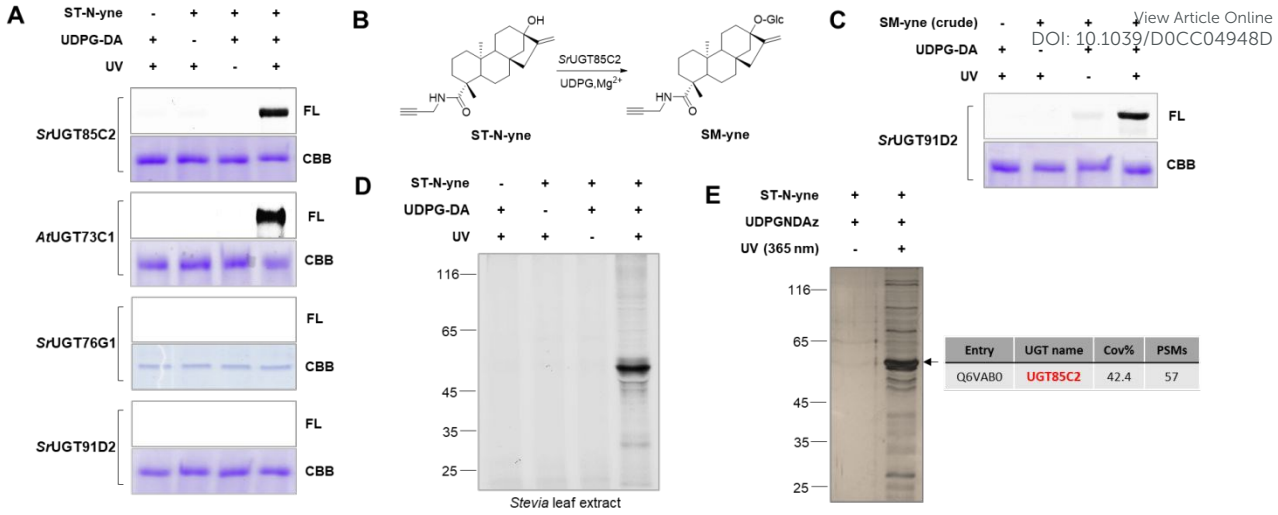


Fig. 2 Proof of concept with recombinant UGTs of known activities and in *Stevia* leaf extracts. (A) Labeling of two steviol-catalyzing UGTs, namely *Sr*UGT85C2 and *At*UGT73C1, by combinatorial use of bi-substrate probe modules. The steviol-noncatalyzing UGTs (*Sr*UGT76G1 and *Sr*UGT91D2) could not be labeled (FL: fluorescence; CBB: Coomassie blue staining); (B) Enzymatic production of the **SM-yne** by recombinant *Sr*UGT85C2; (C) Labeling of *Sr*UGT91D2 by combinatorial use of **SM-yne** (crude) and **UDPG-DA**; (D) Labeling of *Stevia* leaf extracts by combinatorial use of **ST-N-yne** and **UDPG-DA**; (E) Pull-down and identification of the probe labeled band in (D) as *Sr*UGT85C2. See Table S1 in ESI† for full list of identified proteins.

proteomics has enabled an robust analytical modality for identifying binding proteins of bioactive small molecules including natural products.[7-9] Typically, such binding proteins are classifiable into three categories with respect to the natural products, namely: their functional receptors, transporters, and biosynthetic enzymes, which may be generalized as biological parts.[10] In the past decades, biosynthetic pathways or specific biosynthetic enzymes of multiple classes of natural products including sterols, catechins, and nonribosomal peptides have been successfully profiled by utilizing the affinity-based protein profiling approach.[10-13] More recently, *Morus alba* Diels-Alderase (MaDA), a novel FAD-dependent [4+2] cycloaddition enzyme involved in chalcomoracin biosynthesis, has been discovered with full mechanistic details by using a target identification strategy based on a fully functionalized biosynthetic intermediate probe (BIP).[14]

As one of the most critical modification steps during biosynthesis of diverse natural products, glycosylation enhances their solubility and stability, while boosting their storage and accumulation in cells.[15] In plants, glycosylation is catalyzed by a class of enzymes called UDP-glycosyltransferases (UGTs), which utilize UDP-glucose (UDPG) as donors and transfer the sugar moiety of UDPG to various acceptors.[16] We have previously established a robust chemoproteomics-based platform for rapid identification of steviol-catalyzing UGTs in plants by using a substrate-derived photoaffnity probe, **ST-Dayne** (Fig. 1A).[17,18] Recently, Wang *et al.* reported the chemoproteomic profiling of protein-lipid interactions by using a metabolically synthesized phospholipid probe.[19] Inspired by this seminal work, we advanced herein an approach based on "bi-substrate" probe modules to label, enrich, and identify the biosynthetic UGTs with improved selectivity, sensitivity, and efficiency. Unlike our previously reported fully functionalized probe **ST-Dayne**, an alkyne-tagged sugar receptor module (**ST-N-yne**) and a diazirine-tagged sugar donor module (**UDPG-DA**) were prepared (Fig. 1B). Neither of these modules contains all essential properties of a fully functionalized probe, so that their binding proteins would not be captured in chemical proteomic experiments. However, when catalyzed by a specific UGT, the diazirine-modified sugar moiety in **UDPG-DA** is predictably transferred to the receptor module **ST-N-yne**, generating a more specific product (i.e. steviolmonoside)-derived probe for further affinity-based labeling, capturing, and identification of corresponding UGTs (Fig. 1C).

In previous studies undertaken by our group, we have shown that plant glycosyltransferases can tolerate modifications on the C-13 carboxyl group of steviol.[15] Therefore, the sugar receptor module **ST-N-yne** was synthesized by attaching an alkyne reporter group to this position (Scheme 1A). Meanwhile, a combinatorial synthetic route was designed to prepare a diazirine-containing sugar donor module, **UDPG-DA** by using two reported enzymes for the synthesis of acetylated UDP-glucosamine (**UDPGNAc**).[20] Specifically, diazirine-modified glucosamine (**Glc-DA**) was chemically synthesized by coupling a carboxylated diazirine to glucosamine, and phosphorylated by *N*-acetylhexosamine 1-kinase (NahK) to generate **Glc-DA-1-P**, which was subsequently pyrophosphorylated by *N*-acetylglucosamine uridyltransferase (GlmU) to give **UDPG-DA** (Scheme 1B). With both modules in hand, we then verified by LC-MS whether a fully function probe can be biosynthesized *in vitro*. **ST-N-yne** (200 μM) and **UDPG-DA** (1 mM) were sequentially added to recombinant *Stevia rebaudiana* UGT85C2 (Fig. S1, ESI†), a specific UGT glycosylating steviol on its C-13 hydroxyl group,[21] followed by addition of $MgCl_2$ (10 mM) to initiate the reaction. As shown by LC-MS, a steviolmonoside (**SM**)-based fully functionalized probe **P1** (Fig. S2, ESI†) was generated, suggesting that *Sr*UGT85C2 can tolerate modification in UDPG with the diazirine moiety.

With the successful detection of generation of **P1**, we next performed photoaffinity labeling in the same samples to test whether the fully functionalized probe was able to label the corresponding glycosyltransferase, namely, *Sr*UGT85C2.

View Article Online
DOI: 10.1039/D0CC04948D

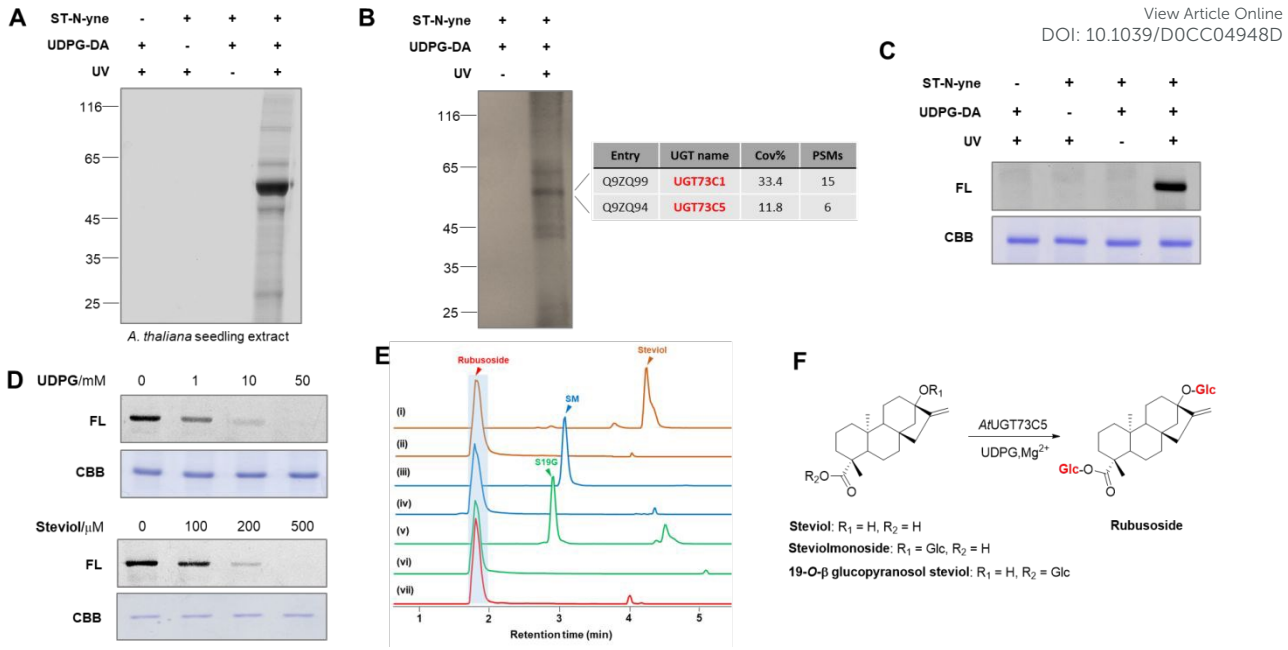


Fig. 3 Discovery of *At*UGT73C5 as a novel steviol-catalyzing UGT in plant model organism *Arabidopsis thaliana*. (A) Labeling of *Arabidopsis* seeding extracts by combinatorial use of **ST-N-yne** and **UDPG-DA**; (B) Pull-down and identification of probe-labeled band in (A) as a combination of *At*UGT73C1 and *At*UGT73C5. See Table S2 in ESI† for full list of identified proteins. (C) Labeling of recombinant *At*UGT73C5 by combinatorial use of bi-substrate probe modules. (D) Competitive labeling of *At*UGT73C5 by either UDPG or steviol. (E) LC-MS analysis of the glycosylated product of steviol, steviolmonoside (**SM**), and 19-*O*-β glucopyranosol steviol (**S19G**) catalyzed by *At*UGT73C5: (i) authentic steviol; (ii) *At*UGT73C5 + steviol; (iii) authentic **SM**; (iv) *At*UGT73C5 + **SM**; (v) authentic **S19G**; (vi) *At*UGT73C5 + **S19G**; or (vii) co-injection of (ii), (iv), and (vi); (F) *At*UGT73C5 catalyses the production of rubusoside from either of steviol, **SM**, and **S19G**.

Solution of the enzymatic reaction was irradiated by UV (365 nm) for 15 min and concentrated by ultracentrifugation, followed by click conjugation with TAMRA-azide and in-gel fluorescence scanning. As anticipated, recombinant *Sr*UGT85C2 was able to

be labeled only when both modules were present, and UV irradiation also proved indispensable for the labeling (Fig. 2A). To optimize labeling conditions, a series of labeling experiments were performed with various reaction durations and probe modules' concentrations. It was determined that 100 μM **ST-N-yne** and 500 μM **UDPG-DA** for 1 h reaction time are optimal for labeling experiments (Fig. S3, ESI†). Similarly, *Arabidopsis* UGT73C1 (Fig. S1, ESI†), a reported glycosyltransferase with the same activity as *Sr*UGT85C2 towards steviol, was also labeled (Fig. 2A).[15] On the other hand, other UGTs involved in stevioside biosynthesis but with non-steviol substrates, such as *Sr*UGT91D2 and *Sr*UGT76G1 (Fig. S1, ESI†), failed to be labeled (Fig. 2A).[22,23] To further show the superior selectivity of our bi-substrate probe strategy, we prepared crude alkynylated steviolmonoside (**SM-yne**) from **ST-N-yne** with *Sr*UGT85C2 (Fig. 2B and Fig. S2, ESI†) and evaluated its labeling effects on the abovementioned UGTs. Predictably, only *Sr*UGT91D2, the UGT catalyzing the formation of steviobioside from steviolmonoside, was significantly labelled (Fig. 2C and Fig. S4, ESI†). Indeed, *Sr*UGT91D2 could not be labeled by **ST-Dayne** and had not been enriched from *Stevia* proteome in previous studies (Fig. S4, ESI†).[17] Taken together, the bi-substrate probe strategy is notably more selective and sensitive than previous strategies.

Having proved the applicability of the new strategy with recombinantly expressed proteins, we set out to profile UGTs in plant extracts. First, we used **ST-N-yne** and **UDPG-DA** concomitantly to examine the labeling profile of the native proteome extracted from fresh *Stevia* leaves, in which large amounts of steviol glycosides were produced. **ST-N-yne**, in the presence of **UDPG-DA** concomitantly, significantly labeled a band of around 50 kDa upon UV irradiation (Fig. 2D and Fig. S5, ESI†). To identify the fluorescent band, photo-labeled samples were click conjugated to biotin-azide, and biotinylated proteins were enriched by streptavidin sepharose. Following sufficient and rigorous washing, bound proteins were boiled in SDS-PAGE loading buffer for elution, resolved by SDS-PAGE, and visualized via silver staining. Stained bands were excised and subjected to in-gel tryptic digestion, followed by LC-MS/MS analysis. Our results show that *Sr*UGT85C2 was the sole glycosyltransferase significantly enriched (Fig. 2E), which suggests much improved selectivity as compared to results based on the photoaffinity probe **ST-Dayne**.[17]

As mentioned, in previous work we have also identified a heterogeneously glycosyltransferase, UGT73C1, in the leaf extracts of *Arabidopsis thaliana* as a steviol-catalyzing enzyme by using **ST-Dayne**.[18] Here, we employed the bi-substrate probe modules in *Arabidopsis* seedling proteomes to profile other candidates of steviol-catalyzing UGTs. In gel-based labeling experiments, similar to the case of *Stevia* proteome, fluorescent signals around 50 kDa were evident (Fig. 3A and Fig. S5, ESI†). Following pull-down experiment and mass spectrometry, two glycosyltransferases including *At*UGT73C1 and another previously unidentified glycosyltransferase *At*UGT73C5, were significantly enriched and identified (Fig. 3B). *In planta*, UGT73C5 regioselectively glycosylates the C23−OH of brassinosteroids and exhibits acceptor plasticity toward diverse aglycones including flavonoids, phenols, mycotoxins and

oleanane-type pentacyclic triterpenes.[23] Remarkably, our results are good evidence that *At*UGT73C5 could be a potential steviol-catalyzing enzyme within *Arabidopsis thaliana*. Next, we prepared recombinant *At*UGT73C5 to validate its interactions with both probe modules and assess its enzymatic activity. As anticipated, recombinant *At*UGT73C5 was successfully labeled in the same manner as *Sr*UGT85C2 and *At*UGT73C1 (Fig. 3C). The labeling could be dose-dependently competed by either UDPG (Fig. 3D) or steviol (Fig. 3E), indicating that the interaction is specific. Finally, *in vitro* enzymatic assays were carried out to further confirm the catalytic products. When **ST-N-yne** and **UDPG-DA** coordinately served as substrates, **P1** was predictably detected (Fig. S6, ESI†). Unexpectedly, however, when unmodified substrates (steviol and UDPG) were applied, no steviolmonoside signals were detected despite catalytic actions of *At*UGT73C5. Instead, a more polar product with a molecular weight corresponding to steviol plus two glucose moieties was detected (Fig. S7, ESI†), suggesting that *At*UGT73C5 might additionally catalyze the glycosylation of steviolmonoside. As the C19-carboxyl group represents another potential glycosylation site, we then investigated the catalytic product of *At*UGT73C5 using either steviolmonoside or 19-*O*-β glucopyranosol steviol (**S19G**) as substrate. Interestingly, both substrates were consumed following catalysis, and the resultant products registered the same retention time as that of steviol. Products for the three *At*UGT73C5-catalyzed reactions with distinct substrates were combined and purified by preparative LC, which turned out to be rubusoside in NMR analysis (Fig. S8, ESI†). Rubusoside is a natural sweetener and a solubilizing agent with diverse biological activities, whose *de novo* production in genetically engineered *E. coli.* required at least two UGTs.[25] Our evidence demonstrates that *At*UGT73C5 alone can catalyze the glycosylation of steviol at both C-13 and C-19, which represents an alternative route for producing rubusoside.

In conclusion, we have developed a facile and efficient method for profiling novel UDP-glycosyltransferases potentially applicable to production of natural products via synthetic biological approaches. Compared with previous works, the current strategy based on bi-substrate probe modules permit significant improvements in selectivity and sensitivity toward UDP-glycosyltransferases in native plant proteomes. As an example, *At*UGT73C5 was unequivocally identified for the first time as a steviol-catalyzing enzyme in *Arabidopsis thaliana*. We anticipate that our current strategy is applicable to mining of other multi-substrate biosynthetic enzymes in bioinformatically challenging organisms such as plants. It should be noted that multi-staged synthetic processes based on bi-substrate probe modules could nonetheless reduce the catalytic efficiency of native enzymes. Further optimization of probe structures and module combinations is imperative to help realize the full application potential of this promising method.

This work was supported by the Natural Science Foundation of Jiangsu Province of China (BK20191477), the Natural Science Foundation of the Higher Education Institutions of Jiangsu Province of China (19KJB150001) and the Natural Science Foundation of China (21502206, 21807102).

Conflicts of interest

View Article Online
DOI: 10.1039/D0CC04948D

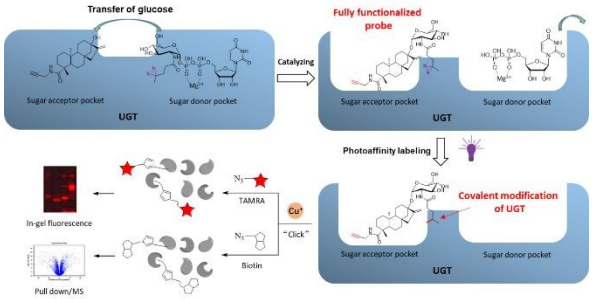
There are no conflicts to declare.

Notes and references

1 F. E. Koehn and G. T. Carter, *Nat. Rev. Drug Discov.*, 2005, **4**, 206-220.
2 P. Xu, N. Bhan, and M. Koffas, *Curr. Opin. Biotechnol.,* 2013, **24**, 291-299.
3 B. Wilkinson and J. Micklefield, *Nat. Chem. Biol.*, 2007, **3**, 379-386.
4 M. H. Medema, and A. Osbourn, *Nat. Prod. Rep.*, 2016, **33**, 951-962.
5 V. De Luca, V. Salim, S. M. Atsum and F. Yu, *Science*, 2012, **336**, 1658-1661.
6 F. Kellner, J. Kim, B. J. Clavijo, J. P. Hamilton, K. L. Childs, B. Vaillancourt, J. Cepela, M. Habermann, B. Steuernagel, L. Clissold, K. McLay, C. R. Buell and S. E. O'Connor, *Plant J.*, 2015, **82**, 680-692.
7 B. F. Cravatt, A. T. Wright and J. W. Kozarich, *Annu. Rev. Biochem.*, 2008, **77**, 383-414.
8 M. H. Wright and S. A. Sieber, *Nat. Prod. Rep.*, 2016, **33**, 681-708.
9 S. Pan, H. Zhang, C. Wang, S. C. Yao and S. Q. Yao, *Nat. Prod. Rep.*, 2016, **33**, 612-620.
10 S. Zhuang, Q. Li, L. Cai, C. Wang and X. Lei, *ACS Cent. Sci.*, 2017, **3**, 501-509.
11 J. J. Hulce, A. B. Cognetta, M. J. Niphakis, S. E. Tully and B. F. Cravatt, *Nat. Method.* 2013, **10**, 259-266.
12 H. Carrie, D. T. Tran, S. Rousseau, S. Chaignepain, J.-M. Schmitter, D. Deffieux and S. Quideau, *Chem. Commun.*, 2014, **50**, 9387-9390.
13 J. L. Meier and M. D. Burkart, *Chem. Soc. Rev.*, 2009, **38**, 2012-2045.
14 L. Gao, C. Su, X. Du, R. Wang, S. Chen, Y. Zhou, C. Liu, X, Liu, R. Tian, L. Zhang, K. Xie, S. Chen, Q. Guo, L. Guo, Y. Hano, M. Shimazaki, A. Minami, H. Oikawa, N. Huang, K. N. Houk, L. Huang, J. Dai and X. Lei, *Nat. Chem.*, 2020, **12**, 620-628.
15 D. Bowles, J. Isayenkova, E. K. Lim and B. Poppenberger, *Curr. Opin. Plant Biol.*, 2005, **8**, 254-263.
16 L. Lairson, B. Henrissat, G. Davies and S. Withers, *Annu. Rev. Biochem.*, 2008, **77**, 521-555.
17 W. Li, Y. Zhou, W. You, M. Yang, Y. Ma, M. Wang, Y. Wang, S. Yuan and Y. Xiao, *ACS Chem. Biol.*, 2018, **13**, 1944-1949.
18 Y. Zhou, W. Li, W. You, Z. Di, M. Wang, H. Zhou, S. Yuan, N.-K. Wong and Y. Xiao, *Chem. Commun.*, 2018, **54**, 7179-7182.
19 D. Wang, S. Du, A. Cazenave-Gassiot, J. Ge, J.-S. Lee, M. R. Wenk and S. Q. Yao, *Angew. Chem. Int. Ed.* 2017, **56**, 1-6.
20 G. Zhao, W. Guan, L. Cai and P. G. Wang, *Nat. Protoc.*, 2010, **5**, 636-646.
21 A. Richman, A. Swanson, T. Humphrey, R. Chapman, B. McGarvey, R. Pocs and J. Brandle, *Plant J.,* 2005, **41**, 56-67.
22 J. Wang, S. Li, Z. Xiong and Y. Wang, *Cell Res.,* 2016, **25**, 258-261.
23 B. Poppenberger, S. Fujioka, K. Soeno, G. L. George, F. E. Vaistij, S. Hiranuma, H. Seto, S. Takatsuto, G. Adam, S. Yoshida and D. Bowles, *Proc. Natl. Acad. Sci. U.S.A.*, 2005, **102**, 15253-15258.
24 E. K. Lim, D. A. Ashford, B. Hou, R. G. Jackson, D. J. Bowles, *Biotechnol. Bioeng.*, 2004, **87**, 623-631.
25 Y. Sun, Z. Chen, J. Li, J. Li, H. Lv, J. Yang, W. Li, D. Xie, Z. Xiong, P. Zhang and Y. Wang, *Mol. Plant*, 2018, **11**, 1308-1311.

TOC

View Article Online
DOI: 10.1039/D0CC04948D



Selective profiling of steviol-catalyzing UDP-glycosyltransferases via “bi-substrate probe” strategy