Letter

Development of Photoaffinity Probe for the Discovery of Steviol Glycosides Biosynthesis Pathway in Stevia rebuadiana and Rapid Substrate Screening

Weichao Li, Yiqing Zhou, Wenjing You, Mengquan Yang, Yanrong Ma, Mingli Wang, Yong Wang, Shuguang Yuan, and Youli Xiao

ACS Chem. Biol., **Just Accepted Manuscript** • DOI: 10.1021/acschembio.8b00285 • Publication Date (Web): 04 Jun 2018

Just Accepted

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Development of Photoaffinity Probe for the Discovery of Steviol Glycosides Biosynthesis Pathway in *Stevia rebuadiana* and Rapid Substrate Screening

Weichao Li[1,†], Yiqing Zhou[1,†], Wenjing You[1,2,†], Mengquan Yang[1,2], Yanrong Ma[1], Mingli Wang[1], Yong Wang[1], Shuguang Yuan[3], Youli Xiao[1,2,*]

[1]CAS Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China

[2]University of Chinese Academy of Sciences, Beijing 100039, China

[3]Laboratory of Physical Chemistry of Polymers and Membranes, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH B3 495 (Bâtiment CH) Station 6, CH-1015 Lausanne, Switzerland

[†]These authors contributed equally to this work.

* Corresponding author

Prof. Youli Xiao

CAS Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China

Phone: 86-21-54924226

ylxiao@sibs.ac.cn

ABSTRACT

Functional discovery and characterization of the target enzymes responsible for the biosynthesis pathway coded for the genes is on-going and the unknown functional diversity of this class of enzymes has been revealed by genome sequencing. Commonly, it is feasible in annotating of biosynthetic genes of prokaryote due to the existence of gene clusters of secondary metabolites. However, in eukaryote the biosynthetic genes are not compactly clustered as the way of prokaryote. Hence, it remains challenging to identify the biosynthetic pathways of newly-discovered natural products in plants. Steviol glycosides are one class of natural sweeteners found in high abundances in the herb *Stevia rebaudiana*. Here, we applied the chemoproteomic strategy for the proteomic profiling of the biosynthetic enzymes of steviol glycosides in *Stevia rebaudiana*. We not only identified a steviol-catalyzing UDP-glycosyltransferase (UGT) UGT73E1 involved in steviol glycosides biosynthesis, but also built up a probe-based platform for the screening of potential substrates of functional uncharacterized UGT rapidly. This approach would be a complementary tool in mining novel synthetic parts for assembling of synthetic biological systems for the biosynthesis of other complex natural products.

■ INTRODUCTION

Steviol glycosides (SGs) are novel natural sweeteners isolated from the leaves of the Paraguayan perennial herb *Stevia rebaudiana* (*S. rebaudiana*) Bertoni, which have been used as one kind of sugar-free and noncariogenic food additive worldwide.[1] In the past decades, most genes responsible for the biosynthesis of SGs have been cloned and well characterized.[2] The biosynthesis of steviol glycosides starts from steviol. Two main products, stevioside and rebaudioside A, were formed by a series of glycosylation reactions catalyzed by four UGTs, including UGT85C2, UGT74G1, UGT91D2, and UGT76G1 (Supplementary Figure S1).[3] In plants, by utilizing UDP-glucose (UDPG) as the sugar donor, UGTs transfer the glucose moiety to a number of secondary metabolites.[4] Actually, in the later stages of biosynthesis of numerous natural products in plant, glycosylation is frequently happened for the modification.[5] In line with the complexity and diversity of natural products in plants, a lot of UGTs have been found to be responsible for the glycosylation of these metabolites.[6] Although recent development of next-generation sequencing technique accelerated the discovery of UGTs involved in the biosynthesis of natural products, the actual

biochemical activities of proteins are still unknown due to limited functional studies, especially for non-model plants.[7-9] Therefore, it is a challenging task to discover and assign biosynthesis genes in plant via conventional functional genomics methods.

Recently, affinity-based protein profiling (A*f*BPP) has been widely used in the identification of physiological and therapeutical targets of numerous bioactive small molecules in human proteome.[10-15] Based on the specific interaction of enzyme (glucosyltransferases) and substrate (glucose receptor) during bio-catalytic process, it is also rational to fish out the binding proteins, including proteins involved in biosynthetic pathways of the SGs in their native proteomes through A*f*BPP. Here, we describe the chemoproteomic profiling of biosynthetic UGTs of steviol glycosides in *S. rebaudiana*. A steviol-derived photoaffinity probe containing both a terminal alkyne and an alkyl diazirine groups, **ST-Dayne** (Figure 1a) was designed for the specific labeling and proteome-wide profiling of UGTs accounting for steviol glycosides biosynthesis in *S. rebuadiana*. As well, this probe can be further utilized in gel-based screening of additional substrates of specific binding UGTs, which provided a rapid screening tool for the characterization of novel biological parts.

■ RESULTS AND DISCUSSION

First, we designed an affinity-based probe bearing conventional biotin tag which was modified at the carboxylic position of steviol (Figure 1a). Unfortunately, no binding sign was detected by pull-down experiment with recombinant UGT85C2 (Supplementary Figure S2), a known UGT involved in SG biosynthesis.[3] This result implied that the binding affinity of UGT with the sugar acceptor, steviol in our case, was not tight enough for affinity-based enrichment in a simple pull-down experiments. In the past decades, photoaffinity labeling has been developed as a powerful technique in covalent capturing of reversible binding partners of small molecules.[16] To this end, **ST-Dayne** was designed by coupling a minimal-sized clickable, photoreactive linker to steviol to the same carboxyl group (Figure 1a).[17,18] We speculated that related UGTs recognizing the steviol scaffold could be covalently captured by the activation of diazirine moiety under UV light, followed by click conjugation to fluorophore or biotin tags for gel- and mass spectrometry-based analysis, respectively (Figure 1b).

To evaluate the feasibility of this new probe, gel-based labeling experiments on UGT85C2 and UGT76G1, another known UGT responsible for SGs biosynthesis, were carried out. Recombinant UGTs were

incubated with increasing concentrations of **ST-Dayne**, irradiated by UV light (365 nm) for 15 min on ice. The samples were subjected to click reaction with TAMRA-N_3, SDS-PAGE and in-gel fluorescence scanning. As a result, both UGTs were nicely labeled by **ST-Dayne** and pretreatment of either UGT with free steviol (10 μM, 50 μM, and 100 μM) reduced the labeling to variable degrees (Figure 1c). In addition, 10 μM of **ST-Dayne** provided robust fluorescent signals with as low as 2.0 μM (~ 0.1 μg μL^{-1}) of UGTs after 20 min UV irradiation (Supplementary Figure S3). We also assessed the ability of **ST-Dayne** to label overexpressed UGTs in *Escherichia coli* (*E. coli*) cell lysates. Clear fluorescent bands corresponding to UGTs (~ 52 kDa) were labeled with 10 μM of **ST-Dayne** (Figure 1d). Taken together, **ST-Dayne** specifically labels known steviol-catalyzing UGTs *in vitro*.

Having optimized probe concentration and UV-irradiation time using recombinant UGTs (10 μM of probe, 1 hour UV irradiation), we conducted the plant proteome labeling and enrichment studies. Native proteome of *S. rebaudiana* was prepared by grinding the fresh leaves in potassium phosphate (KPi) as previously described.[19] The extracted plant proteome was incubated with **ST-Dayne** (10 μM) either with excess steviol (50 μM, 5 folds) as competitor or alone. After incubation, the proteomes were irradiated with UV light, clicked to TAMRA-N_3, and detected by in-gel fluorescence. Multiple protein bands, including several dominant ~ 50 kDa protein bands were labeled (Figure 2a (lane 2), Supplementary Figure S4); most of the labeled proteins were readily competed by excess steviol (lane 3). As shown in lane 1, UV irradiation is necessary for the labeling of almost all detected proteins, indicating the non-covalent binding property. Next, we performed large scale pull-down experiment to enrich and identify the binding proteins of **ST-Dayne**. After probe incubation and UV crosslinking, the proteome samples were subjected to Cu (I)-catalyzed click reaction in the presence of biotin-azide then enriched by incubation with streptavidin column. Bound proteins in the column were subjected to *in situ* tryptic digestion and the digests were labeled with respective TMT isobaric reagents, and followed by combining for analysis with mass spectrometry. To have functional proteomic information, the public transcriptic data of *S. rebaudiana* from NCBI (SRR1576548)[20] was re-analyzed, generating 62,015 genes which were classified into four main categories by biochemical pathway KEGG database (Figure 2b). In the 13,028 sequences grouped into carbohydrate metabolism, 144 sequences were found to contain the common signature motif Plant secondary product glycosyltransferase (PSPG) box. Among plant UGTs, PSPG box is highly-conserved and has been reported to specifically interact with the UDP moiety of UDPG,[4] indicating the 144 sequences were potential UGTs. We also performed a global proteomic profiling upon *Stevia* leaves with the filter-aided sample preparation

(FASP) protocol.[21] The results showed 68 of the 144 sequences of UGTs were unambiguously identified (Supplementary Figure S5), suggesting the high quality of proteomic database yielded from RNA-seq data. Based on the generated *Stevia* proteomic database, 201 sequences were identified which were ranked in corresponding volcano plots as a ($\log_2$-) ratio of **ST-Dayne** to DMSO treatment against statistical significance (-$\log_{10}$-*p*-value). Proteins enriched by a factor > 1.3 with a *p*-value of 0.05 or below were considered positive hits (Figure 2c). Evaluation of the data revealed seven sequences as positive UGT hits, which finally was corresponding to 5 UGTs (Figure 2d). These UGTs contain two known UGTs involved in SG biosynthesis, namely UGT85C2 (sequence 51051) and UGT76G1 (sequences 65695 and 65696), further validating our approach. Moreover, two previously identified UGTs (UGT79A2, sequence 26485; UGT73E1, sequences 32570 and 32578) with unknown functions as well as an unidentified UGT candidate (UGT76G3, sequence 11768), were also significantly enriched as positive hits (Supplementary Table S2). To validate these UGTs as specific steviol-binding proteins, recombinant UGT73E1 or UGT76G3 was individually treated with **ST-Dayne**, irradiated by UV and followed by click reaction. As predicted, both UGTs were successfully labeled by **ST-Dayne** in UV-dependent manner, and the labeling was competitively abolished by excess steviol, indicating the specific interactions (Figure 2e). To further assess whether both UGTs are targeted by steviol, we performed a thermal-shift assay and compared their thermal stability after steviol- and DMSO-treatment. As detected by antibody of His-tag, both UGTs were obviously stabilized by 10 μM steviol (Supplementary Figure S6). Validation of *Sr*UGT79A2 was not performed due to its insoluble nature during recombinant expression.

With the two soluble UGTs in hands, the glucosyltransferase activity of UGT73E1 and UGT76G3 were evaluated *in vitro*. With steviol or steviol glycosides as substrates (all 1.0 mM), the reactions were catalyzed by both UGTs (20 μM) in the presence of UDPG (1.0 mM) and $MgCl_2$ (3.0 mM) and monitored by LC-MS. UGT73E1 appeared to catalyze the formation of 19-O-*β* glucopyranosol steviol from steviol by addition of one molecule of glucose (Figure 3a). However, further comparison of the product with authentic steviolmonoside, the enzymatic product of steviol catalyzed by UGT85C2, revealed slight difference in retention time. Co-injection of the product with steviolmonoside showed two different peaks were observed (Supplementary Figure S7). Comprehensive analysis of 1D and 2D-NMR data (heteronuclear multiple bond correlation (HMBC) and heteronuclear singular quantum correlation (HSQC)) established that the H-1′ (*δ* 5.41, 1H, d) well correlated with C-19 (*δ* 176.81)

(Supplementary Figure S8), which indicated that the glycosylation reaction occurred in the C-4 carboxyl group of steviol. In addition, UGT73E1 was also found to catalyze the formation of rubusoside from steviolmonoside, in which the same site was glycosylated (Figure 3b, Supplementary Figure S9). Based on these results, we conclude that UGT73E1 is a steviol C-4 glucosyltransferase potentially involved in the glycosylation of steviol in *S. rebaudiana* (Figure 3c, Figure S10).

On the other hand, we did not detect any catalytic activity of UGT76G3 on all the above tested steviol or steviol glycosides (data not shown) even it had similarity binding affinity comparing to UGT73E1 from the competitive assays (Figure 2e). As non-specific binding is frequently observed in photoaffinity labeling, we then evaluated the binding model through tandem mass analysis using photoactive probe, **ST-Dayne**, combined with molecular modeling. Recombinant UGT76G3 was photo-labeled by **ST-Dayne**, digested and analyzed by LC-MS/MS. The peptide ^{191}SMNDPYAK^{198} was found to exhibit a mass shift of 410.28 (**ST-Dayne**-N_2) and MS2 spectrum of this peptide localized the labeling to Tyr196 (Figure 4a). To further understand the ligand binding mode and specificity of UGT76G3 at atomic level, we constructed its 3D structure using homology modeling. The homology model of UGT76G3 was constructed in Modeller using multi-template method, based on the crystal structures of *Medicago truncatula* UGT85H2 (PDB: 2PQ6) and anthocyanidin 3-O-glucosyltransferase (PDB: 3WC4).[22-24] A UDP molecule was observed in both template crystal structures. Thus, we included a UDPG molecule in an identical conformation in the corresponding position. The overall model of UGT76G3 comprises of twenty α-helices and thirteen β-strands and the UDPG molecule, which is surrounded by several loops, sits in the deep pocket of UGT76G3 (Supplementary Figure S11). From the molecular docking with steviol as ligand, steviol was located between Tyr196 and the sugar ring of UDPG (Figure 4c). However, the carboxylic group of steviol, which was far away from UDPG, rotated almost 90 degrees to form a hydrogen bond with Tyr196. A methyl group connected to carboxylic group formed hydrophobic interactions with UDPG, which prevented the carboxylic group of steviol to be catalyzed. This observed modeling may explain why there was no detectable catalytic activity of UGT76G3 towards steviol.

Compared to other classes of biosynthetic enzymes, plant UGTs have been reported to be more promiscuous and can act on various classes of substrates.[25] Building on the results showing that steviol was also located in the substrate pocket of UGT76G3, we tested whether **ST-Dayne** could be applied in the screening of UGT76G3 substrates. Recombinant UGT76G3 proteins were incubated with a series of

available natural products in our lab, including terpenes, xanthone and flavone, followed by addition of **ST-Dayne** as a "reporter" of substrate pocket occupancy. UV-irradiation, Cu (I)-catalyzed click reaction, then in-gel fluorescence scanning were performed same as above in triplicates and the competition effect was measured by gel densitometry. As shown in Figure 4d, three out of 9 tested compounds, namely isosteviol (lane 2), gibberellic acid 3 (lane 3), and apigenin (lane 6), were found to diminish the probe labeling by more than 60%, as compared to positive control (DMSO, lane 1). The glucosyltransferase activities of UGT76G3 towards these three compounds were further confirmed by HR-ESI-MS (Figure 4e, Supplementary Figure S12), suggesting our probe did can serve as a complementary tool for the rapid screening of UGT substrates. According to the result of molecular docking, isosteviol was also located in an identical place with that of steviol. In contrast, the carboxylic group of isosteviol formed hydrogen bonds with the glucose moiety of UDPG which implied a promising catalytic activity by UGT76G3 (Figure 4b), in agreement with the biochemical results.

Comparing with model species, the public proteome database of a number of non-model plants is not complete. In the present, we generated the protein sequence databases from public RNA-Seq data of *Stevia*, which significantly increased the amount of peptide matching, and hence protein (UGTs in this study) identification. Taking advantage of chemoproteomic strategy, the interactome of steviol in *Stevia* has been profiled, which is much more straightforward and efficient by integrating conventional classical methods which based on genome sequencing and bioinformatics analysis. In this study, the affinity-based probe of steviol was capable of specifically labeling multiple biosynthesis-related UGTs in different biological systems, including recombinant proteins, overexpressed *E. coli* cell lysates, and *Stevia* plant proteome. By this approach we found that UGT73E1, a *Stevia* UGT whose function remains unknown, is capable of catalyze the glycosylation of C4-caboxyl group of steviol. On the other hand, both activity- and affinity-based probes have been widely utilized in the rapid screening and characterization of substrates, inhibitors, cofactors, and other small molecular ligands of enzymes.[15,26,27] Although we cannot characterize the glycosyltransferase of UGT76G3 towards all the available SGs tested, by using of competitive affinity-based labeling strategy, we discovered three substrates of UGT76G3 which also expanded the library of available UGT biological parts. Looking forward, we anticipate that our strategy can be expanded to profile and discovery the biosynthetic pathways of secondary metabolites in other non-model medicinal plant species.

■ METHODS

Experimental methods are detailed in the Supporting Information.

■ ASSOCIATED CONTENT

Supporting information

The Supporting information is available free of charge on the ACS Publications website at DOI: 10.1021/acschembio.xxx.

The synthetic method of biotin-steviol and ST-Dayne, biological experimental procedures, and bioinformatics analysis (PDF).

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 86-21-54924226. E-mail: ylxiao@sibs.ac.cn

ORCID

Weichao Li: 0000-0001-6875-7315

Yiqing Zhou: 0000-0002-6391-3259

Yong Wang: 0000-0001-5541-3978

Shuguang Yuan: 0000-0001-9858-4742

Youli Xiao: 0000-0002-4803-3333

Author Contributions

†These authors contributed equally to this work.

Notes

The authors declare no competing financial interests.

■ ACKNOWLEDGMENTS

This work was supported by the “Thousand Talents Program” young investigator award, the Strategic Priority Research Program "Molecular mechanism of Plant Growth and Development" of CAS (Grant XDPB0402), International Partnership Program of CAS (Grant 153D31KYSB20170121 and

153D31KYSB20160074), the NSF-China (Grants 21572243 and 21502206), the Key Projects of Shanghai Committee of Science and Technology (Grant 15JC1400402). The molecular modeling was performed at the Interdisciplinary Centre for Mathematical and Computational Modeling in Warsaw (grant no. GB70-3 and GB71-3). We thank W. Wu for sharing the information of UGT unigenes. We also thank Y. Shan, Y. Liu, and S. Bu in the Core Facility Centre of the Institute of Plant Physiology and Ecology for mass spectrometry and NMR assistance.

■ REFERENCES

1. Geuns, J. Stevioside. (2003) *Phytochemistry 64*, 913-921.
2. Brandle, J. E. and Telmer, P. G. (2007) Steviol glycoside biosynthesis. *Phytochemistry 68*, 1855-1863.
3. Richman, A., Swanson, A., Humphrey, T., Chapman, R., McGarvey, B., Pocs, R., and Brandle, J. (2005) Functional genomics uncovers three glucosyltransferases involved in the synthesis of the major sweet glucosides of *Stevia rebaudiana*. *Plant J. 41*, 56-67.
4. Lairson, L., Henrissat, B., Davies, G., and Withers, S. (2008) Glycosyltransferases: structures, functions, and mechanisms. *Ann. Rev. Biochem. 77*, 521-555.
5. Vogt, T. and Jones, P., (2000) Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends Plant Sci. 5*, 380-386.
6. Bowles, D., Isayenkova, J., Lim, E. K., and Poppenberger, B. (2005) Glycosyltransferases: managers of small molecules. *Curr. Opin. Plant Biol. 8*, 254-263.
7. Lau, W. and Sattely, E. S. (2015) Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science 349*, 1224-1228.
8. Medema, M. H. and Osbourn, A. (2016) Computational genomic identification and functional reconstitution of plant natural product biosynthetic pathways. *Nat. Prod. Rep. 33*, 951-962.
9. Van Der Hoorn, R. A., Colby, T., Nickel, S., Richau, K.H., Schmidt, J., and Kaiser, M. (2011) Mining the active proteome of *Arabidopsis thaliana*. *Front. Plant Sci. 2,* 89-95.
10. Hulce, J. J., Cognetta, A. B., Niphakis, M. J., Tully, S. E., and Cravatt, B. F. (2013) Proteome-wide mapping of cholesterol-interacting proteins in mammalian cells. *Nat. Methods 11*, 259-264.
11. Wright, M. and Sieber, S. (2016) Chemical proteomics approaches for identifying the cellular targets of natural products. *Nat. Prod. Rep. 33*, 681-708.
12. Zhou, Y., Li, W., and Xiao, Y. (2016) Profiling of multiple targets of artemisinin activated by hemin in cancer cell proteome. *ACS Chem. Biol. 11*, 882-888.
13. Shi, H., Zhang, C. J., Chen, G. Y., and Yao, S. Q. (2012) Cell-based proteome profiling of potential

dasatinib targets by use of affinity-based probes. *J. Am. Chem. Soc. 134*, 3001-3014.

14. Rix, U., and Superti-Furga, G. (2009). Target profiling of small molecules by chemical proteomics. *Nat. Chem. Biol. 5*, 616-624.

15. Horning, B. D., Suciu, R. M., Ghadiri, D. A., Ulanovskaya, O. A., Matthews, M. L., Lum, K. M., Backup, K. M., Brown, S. J., Rosen, H, and Cravatt, B. F. (2016). Chemical proteomic profiling of human methyltransferases. *J. Am. Chem. Soc. 138*, 13335-13343.

16. Kawatani, M., and Osada, H. (2014) Affinity-based target identification for bioactive small molecules. *Medchemcomm. 5*, 277-287.

17. Li, Z., Hao, P., Li, L., Tan, C. Y., Cheng, X., Chen, G. Y., Sze, S. K., Shen, H. M., and Yao, S. Q. (2013). Design and synthesis of minimalist terminal alkyne-containing diazirine photo-crosslinkers and their incorporation into kinase inhibitors for cell- and tissue-based proteome profiling. *Angew. Chem. Int. Ed. 52*, 8551-8556.

18. Li, Z., Wang, D., Li, L., Pan, S., Na, Z., Tan, C. Y., and Yao, S. Q. (2014). "Minimalist" cyclopropene-containing photo-cross-linkers suitable for live-cell imaging and affinity-based protein labeling. *J. Am. Chem. Soc. 136*, 9990-9998.

19. Shibata, H., Sonoke, S., Ochiai, H., Nishihashi, H., and Yamada, M. (1990) Glucosylation of steviol and steviol-glucosides in extracts from Stevia rebaudiana bertoni. *Plant Physiol. 95*, 152-156.

20. Chen, J., Hou, K., Qin, P., Liu, H., Yi, B., Yang, W., and Wu, W. (2014) RNA-Seq for gene identification and transcript profiling of three Stevia rebaudiana genotypes. *BMC genomics 15*, 571-578.

21. Wiśniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nat. Methods 6*, 359-362.

22. Osmani, S., Bak, S., and Møller, B. (2009) Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. *Phytochemistry 70*, 325-347.

23. Li, L., Modolo, L. V., Escamilla-Trevino, L. L., Achnine, L., Dixon, R. A., and Wang, X. (2007) Crystal structure of *Medicago truncatula* UGT85H2--insights into the structural basis of a multifunctional (iso)flavonoid glycosyltransferase. *J. Mol. Biol. 370*, 951-963.

24. Hiromoto, T., Honjo, E., Tamada, T., Noda, N., Kazuma, K., Suzuki, M., and Kuroki, R. (2013) Crystal structure of UDP-glucose:anthocyanidin 3-O-glucosyltransferase from *Clitoria ternatea*. *J. Synchrotron Radiat. 20*, 894-898.

25. Blanchard, S. and Thorson, J. S. (2006). Enzymatic tools for engineering natural product glycosylation. *Curr. Opin. Chem. Biol.*, *10*, 263-271.

26. Niphakis, M. J., and Cravatt, B. F. (2014). Enzyme inhibitor discovery by activity-based protein profiling. *Annu. Rev. Biochem*. *83*, 341-377.

27. Grossman, E. A., Ward, C. C., Spradlin, J. N., Bateman, L. A., Huffman, T. R., Miyamoto, D. K., Kleinman, J. I., and Nomura, D. K. (2017). Covalent ligand discovery against druggable hotspots

targeted by anti-cancer natural products. *Cell Chem. Biol. 24*, 1368-1376.

Figure legends

Figure 1. Strategy and workflow of A*f*BPP utilizing photoaffinity probe for steviol glycosides biosynthesis pathway discovery in *S. rebaudiana*. (a) Structures of steviol, rebaudioside A, **biotin-steviol** and designed photoaffinity probe (**ST-Dayne**). (b) Strategy and workflow of profiling of UGTs with photoaffinity probe. (c) Gel-based competitive labeling of recombinant *Sr*UGT85C2 and *Sr*UGT76G1. (d) Gel-based competitive labeling of overexpressed *Sr*UGT85C2 and *Sr*UGT76G1 in *E. coli* cell lysates.

Figure 2. Chemoproteomic profiling of *S. rebaudiana* proteome using photoaffinity probe. (a) Gel-based competitive labeling of *S. rebaudiana* proteome. (b) Functional annotation and classification of unique putative transcripts from public transcriptic data of *S. rebaudiana* on KEGG categories. (c) Quantitative mass spectrometry-based **ST-Dayne** profiling of *S. rebaudiana* proteome. (d) Veen diagram of enriched proteins. (e) Competitive labeling of recombinant *Sr*UGT73E1 and *Sr*UGT76G3.

Figure 3. Characterization of the glucosyltransferase activity of *Sr*UGT73E1 and redraft of biosynthesis pathway of steviol glycosides. (a) HR-ESI-MS analysis of the product catalyzed by *Sr*UGT73E1 during the *in vitro* assay. (i) Authentic Steviol; (ii) Authentic 19-O-β glucopyranosol steviol; (iii) *Sr*UGT73E1 was used after treatment in 100 °C water for 5 min as control; (iv) *Sr*UGT73E1 reaction product. (b) HR-ESI-MS analysis of the product catalyzed by *Sr*UGT73E1 during the *in vitro* assay. (i) Authentic Steviolmonoside; (ii) Authentic Rubusoside; (iii) *Sr*UGT73E1 was used after treatment in 100 °C water for 5 min as control; (iv) *Sr*UGT73E1 reaction product. (c) Redrafted biosynthesis pathway of steviol glycosides.

Figure 4. Molecular docking and competitive A*f*BPP-based substrate screening of *Sr*UGT76G3. (a) MS based binding site identification of *Sr*UGT76G3. Molecular docking studies of (b) the binding of isosteviol in complex with UGT76G3 (Yellow stick: substrate molecule isosteviol. Green stick: UDPG molecule) and (c) The binding of steviol in complex with UGT76G3 (Cyan stick: inhibitor molecule steviol. Green stick: UDPG molecule). (d) Competitive labeling of *Sr*UGT76G3 by ST-Dayne for substrate screening. (e) Characterization of the glucosyltransferase activity of *Sr*UGT76G3 by HR-ESI-MS analysis.

Figure 1.

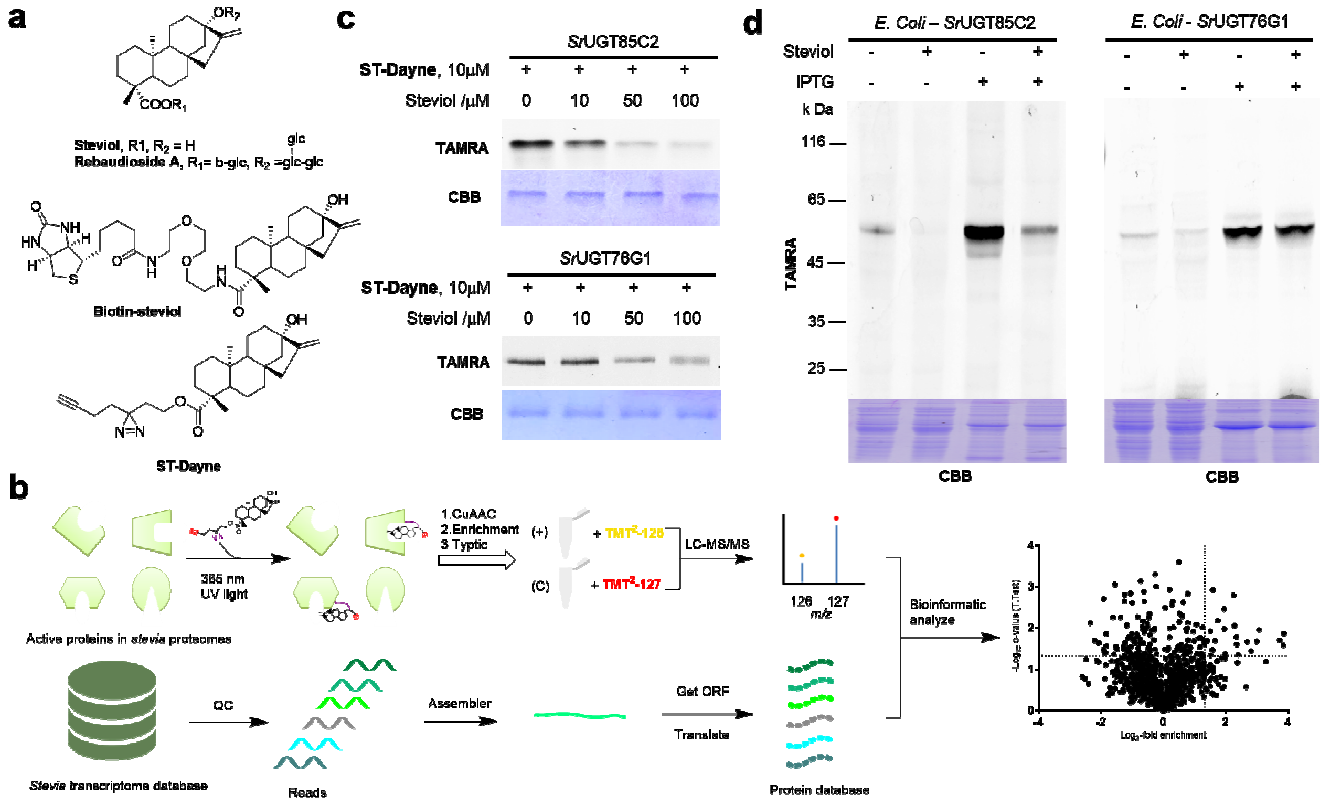

a
Steviol, R1, R2 = H
Rebaudioside A, R1= b-glc, R2 =glc-glc
Biotin-steviol
ST-Dayne
b
Active proteins in stevia proteomes
365 nm
UV light
1.CuAAC
2.Enrichment
3.Typtic
(+)
(C)
+ TMT2-126
+ TMT2-127
LC-MS/MS
126 127
m/z
Bioinformatic analyze
Stevia transcriptome database
QC
Reads
Assembler
Get ORF
Translate
Protein database
c
SrUGT85C2
ST-Dayne, 10μM
Steviol /μM
0 10 50 100
TAMRA
CBB
SrUGT76G1
ST-Dayne, 10μM
Steviol /μM
0 10 50 100
TAMRA
CBB
d
E. Coli – SrUGT85C2
E. Coli - SrUGT76G1
Steviol
IPTG
k Da
116
65
45
35
25
TAMRA
CBB
CBB

Figure 2.

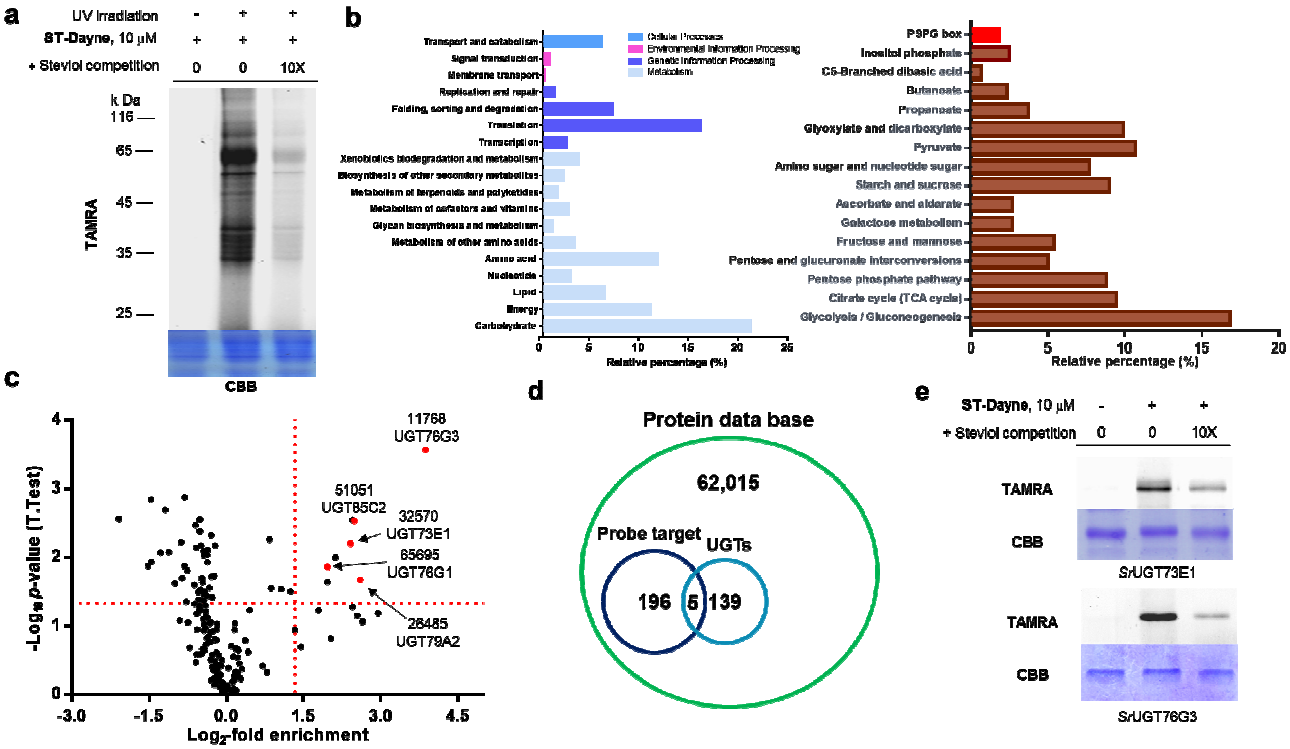

a
UV irradiation - + +
ST-Dayne, 10 μM + + +
+ Steviol competition 0 0 10X
k Da
116
65
45
35
25
TAMRA
CBB
b
Transport and catabolism
Signal transduction
Membrane transport
Replication and repair
Folding, sorting and degradation
Translation
Transcription
Xenobiotics biodegradation and metabolism
Biosynthesis of other secondary metabolites
Metabolism of terpenoids and polyketides
Metabolism of cofactors and vitamins
Glycan biosynthesis and metabolism
Metabolism of other amino acids
Amino acid
Nucleotide
Lipid
Energy
Carbohydrate
Cellular Processes
Environmental Information Processing
Genetic Information Processing
Metabolism
Relative percentage (%)
PSPG box
Inositol phosphate
C5-Branched dibasic acid
Butanoate
Propanoate
Glyoxylate and dicarboxylate
Pyruvate
Amino sugar and nucleotide sugar
Starch and sucrose
Ascorbate and aldarate
Galactose metabolism
Fructose and mannose
Pentose and glucuronate interconversions
Pentose phosphate pathway
Citrate cycle (TCA cycle)
Glycolysis / Gluconeogenesis
Relative percentage (%)
c
11768
UGT76G3
51051
UGT85C2
32570
UGT73E1
85695
UGT76G1
28485
UGT79A2
-Log10 p-value (T.Test)
Log2-fold enrichment
d
Protein data base
62,015
Probe target
UGTs
196
5
139
e
ST-Dayne, 10 μM - + +
+ Steviol competition 0 0 10X
TAMRA
CBB
StUGT73E1
TAMRA
CBB
StUGT76G3

Figure 3.

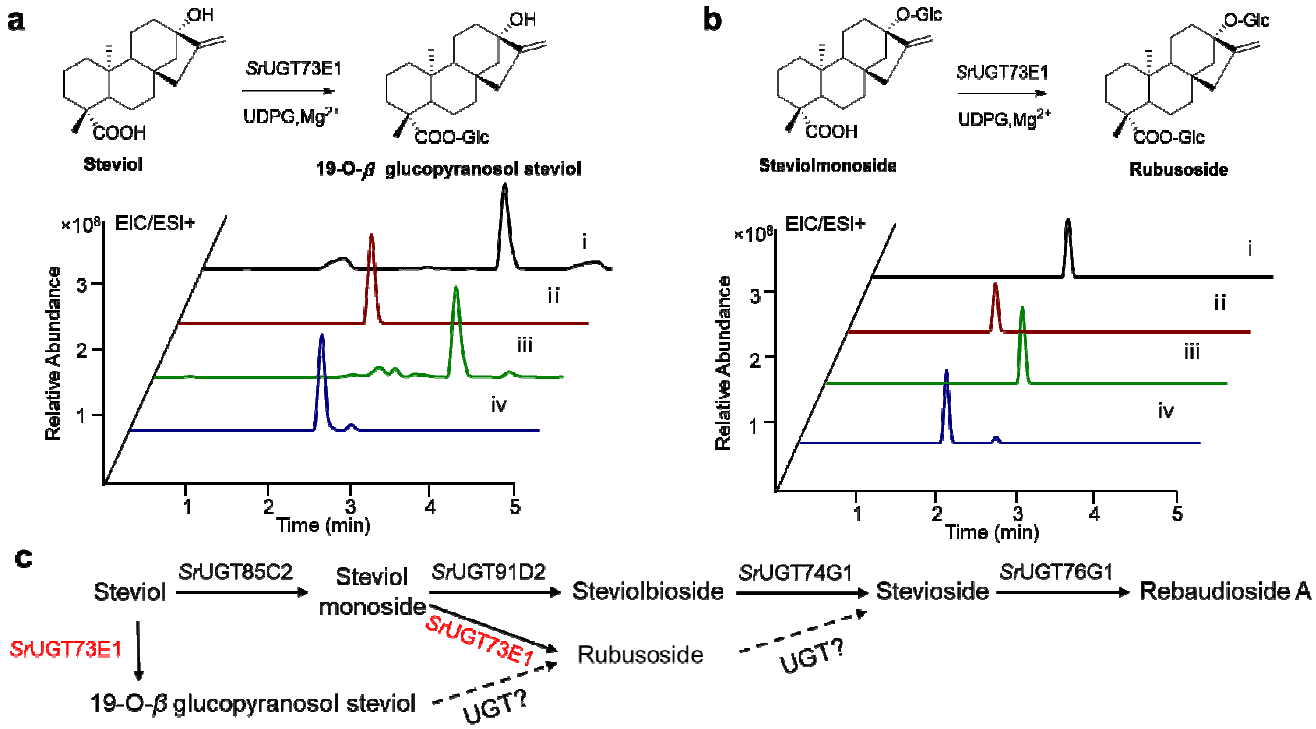
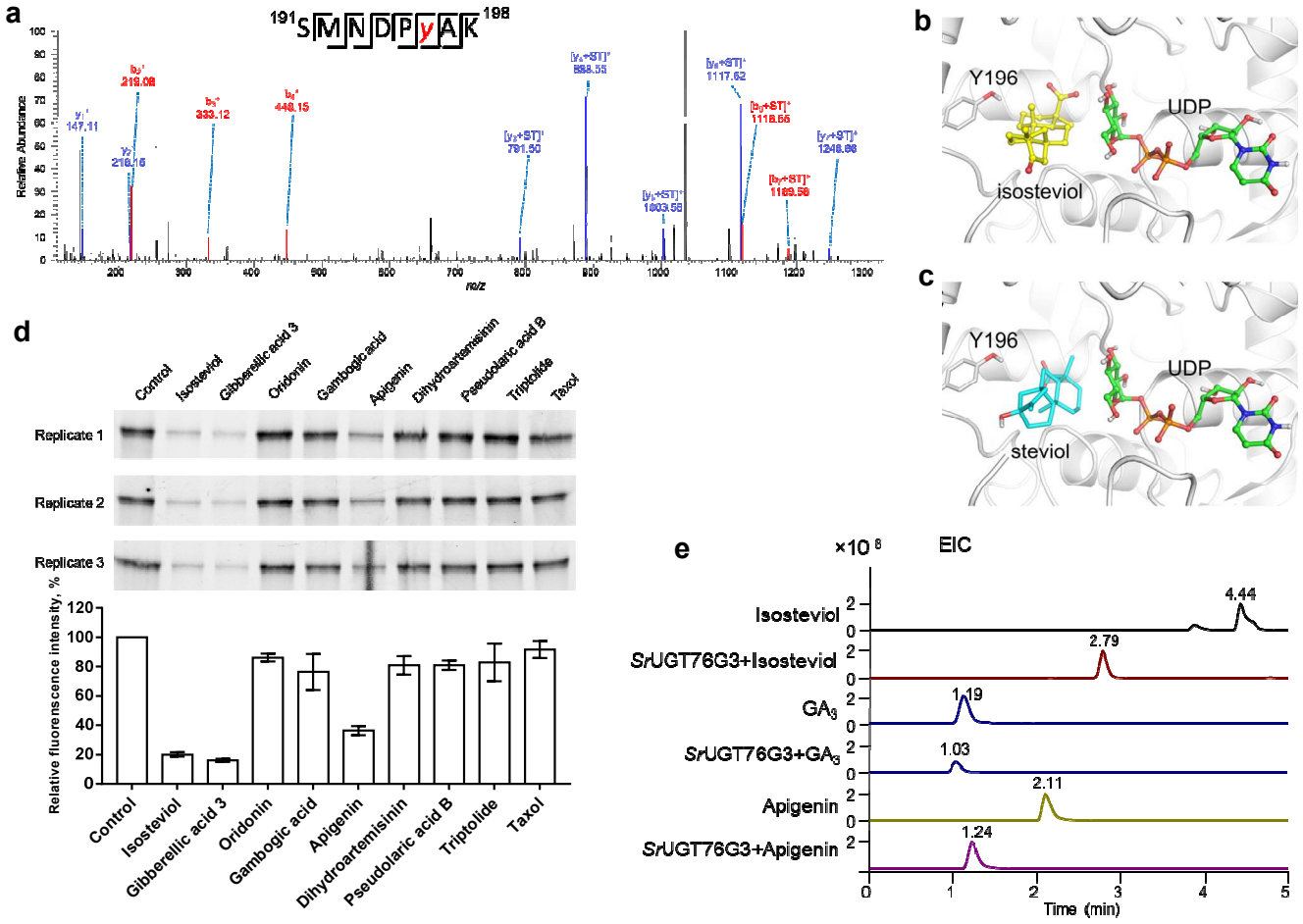
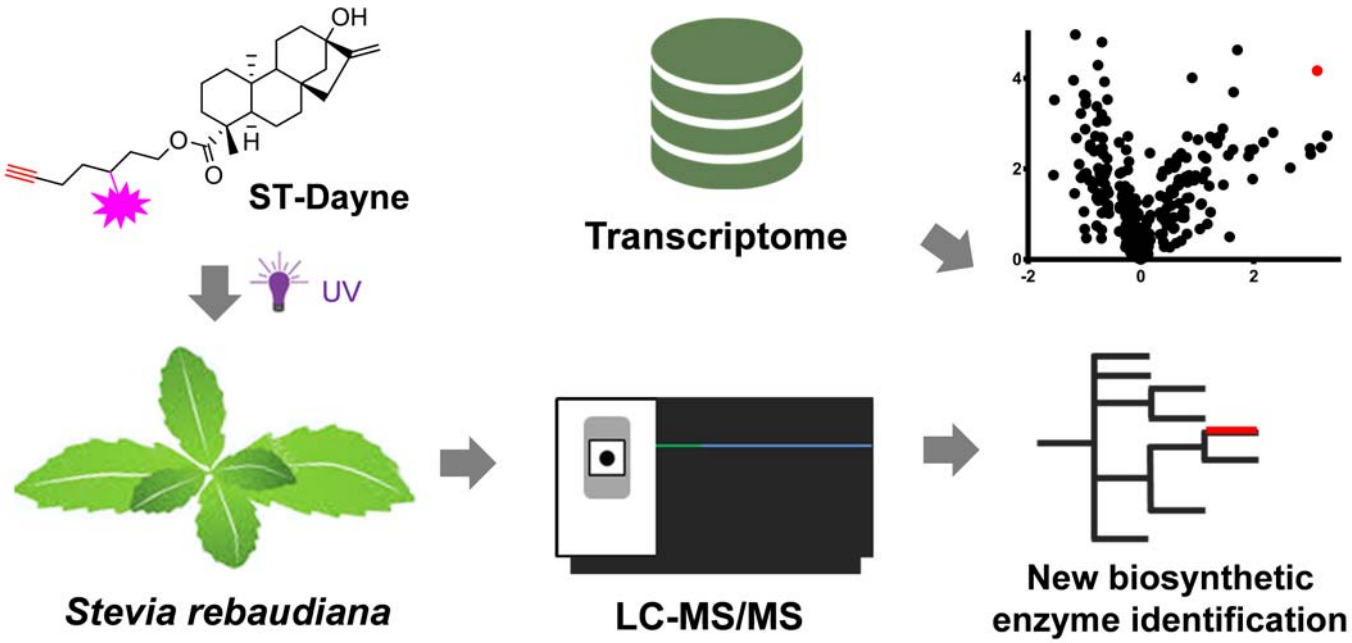

a
OH
COOH
Steviol
SrUGT73E1
UDPG,Mg2+
OH
COO-Glc
19-O-β glucopyranosol steviol
×10⁸ EIC/ESI+
Relative Abundance
Time (min)
i
ii
iii
iv
b
O-Glc
COOH
Steviolmonoside
SrUGT73E1
UDPG,Mg2+
O-Glc
COO-Glc
Rubusoside
×10⁸ EIC/ESI+
Relative Abundance
Time (min)
i
ii
iii
iv
c
Steviol
SrUGT85C2
Steviol monoside
SrUGT91D2
Steviolbioside
SrUGT74G1
Stevioside
SrUGT76G1
Rebaudioside A
SrUGT73E1
19-O-β glucopyranosol steviol
UGT?
SrUGT73E1
Rubusoside
UGT?

Figure 4.


a
191SMNDPyAK198
Relative Abundance
m/z
b
Y196
UDP
isosteviol
c
Y196
UDP
steviol
d
Control
Isosteviol
Gibberellic acid 3
Oridonin
Gambogic acid
Apigenin
Dihydroartemisinin
Pseudolaric acid B
Triptolide
Taxol
Replicate 1
Replicate 2
Replicate 3
Relative fluorescence intensity, %
e
×10 6
EIC
Isosteviol
SrUGT76G3+Isosteviol
GA3
SrUGT76G3+GA3
Apigenin
SrUGT76G3+Apigenin
4.44
2.79
1.19
1.03
2.11
1.24
Time (min)


OH
O
O
H
ST-Dayne
UV
Transcriptome
4
2
0
-2
0
2
Stevia rebaudiana
LC-MS/MS
New biosynthetic
enzyme identification