Enzyme Catalysis

Deutsche Ausgabe: DOI: 10.1002/ange.201506505
Internationale Ausgabe: DOI: 10.1002/anie.201506505

Probing the Catalytic Promiscuity of a Regio- and Stereospecific C-Glycosyltransferase from *Mangifera indica*

*Dawei Chen, Ridao Chen, Ruishan Wang, Jianhua Li, Kebo Xie, Chuancai Bian, Lili Sun, Xiaolin Zhang, Jimei Liu, Lin Yang, Fei Ye, Xiaoming Yu, and Jungui Dai**

Abstract: *The catalytic promiscuity of the novel benzophenone C-glycosyltransferase, MiCGT, which is involved in the biosynthesis of mangiferin from Mangifera indica, was explored. MiCGT exhibited a robust capability to regio- and stereospecific C-glycosylation of 35 structurally diverse drug-like scaffolds and simple phenolics with UDP-glucose, and also formed O- and N-glycosides. Moreover, MiCGT was able to generate C-xylosides with UDP-xylose. The OGT-reversibility of MiCGT was also exploited to generate C-glucosides with simple sugar donor. Three aryl-C-glycosides exhibited potent SGLT2 inhibitory activities with IC_{50} values of 2.6×, 7.6×, and 7.6×10^{-7} M, respectively. These findings demonstrate for the first time the significant potential of an enzymatic approach to diversification through C-glycosidation of bioactive natural and unnatural products in drug discovery.*

Sugar moieties are often essential for the biological activities and pharmacological properties of many natural products.[1] Most of these products are O-glycosylated. However, N-, C-, and S-glycosides are also found in nature.[2] Compared with other glycosides, C-glycosides are of particular pharmaceutical interest because of their pronounced stability against hydrolysis in vivo.[3] Chemical glycosylation remains restricted by such disadvantages as poor regio- and stereoselectivities, and the protection and deprotection of functional groups, but enzymatic syntheses catalyzed by specific glycosyltransferases (GTs; EC 2.4) can alleviate these disadvantages and are thus powerful tools.[4] Significant progress has been achieved recently in the O-glycodiversification of natural and unnatural products by OGTs.[5] However, GTs which are naturally responsible for forming C-glycosidic bonds (CGTs) appear to be comparatively sparse and exhibit stringent specificities which limit their availability and scope for synthesis.[6]

In the past few years, studies on microbial CGTs have attracted considerable interest and achieved great progress.[6a,7] Structurally diverse C-glycosides have also been isolated from plants, and implies the existence of corresponding CGTs.[3a] However, the identification of CGTs involved in plant secondary metabolism was rather limited until 2009, and these CGTs have priority over C-glycosylating flavonoids.[8] Furthermore, the majority of known CGTs exhibit relatively narrow substrate spectra, and few CGTs have been applied as enzymatic tools for the C-glycosylation of various natural and unnatural compounds. Therefore, mining CGTs with catalytic promiscuity, but regio- and stereospecificity, from plants will certainly facilitate the production of a broad range of valuable products.

Mangifera indica is a traditional Chinese medical herb and a wide variety of bioactive natural glycosides, including xanthone and benzophenone C-glycosides, have been isolated from its leaves and bark.[9] However, little is known about the enzymes that are responsible for the biosyntheses of these compounds, particularly in terms of the key step of C-glycosylation, which inspired us to search for CGTs with specific catalytic properties. Herein, we report for the first time, the identification of the novel benzophenone C-glycosyltransferase, MiCGT, which is involved in the biosynthesis of mangiferin from *M. indi*ca. Most importantly, this communication highlights the catalytic promiscuity and regio- and stereospecificity of MiCGT.

To clone permissive CGTs from *M. indica*, a degenerate PCR primer for 3′-RACE (see Table S1 in the Supporting Information) was designed based on the conserved plant secondary product glycosyltransferases (PSPG) motif (see Figure S1 in the Supporting Information).[8,10] The 3′-termini of 23 different UGT cDNAs were obtained using the total RNA from *M. indica* leaves as a template. Combined with 5′-RACE, 15 new putative candidates (*MiUGTs*) for CGTs were successfully cloned and heterologously expressed in *Escherichia coli* as described in the Supporting Information. To identify the C-glycosylation capability of the MiUGTs in vitro, UDP-glucose and putative biosynthetic precursors of mangiferin (see Scheme S1 in the Supporting Information), maclurin (**1**), and norathyriol (**25**)[11] were each incubated with the crude recombinant protein, and analyzed with HPLC-UV/MS^2. Of the 15 recombinant enzymes, only MiUGT13 was observed to C-glucosylate **1** to **1a** with the appearance of a parent ion peak at m/z 423 $[M-H]^-$, which was 162 amu more than that of **1**, and the characteristic fragment ions of C-glucoside, $[M-H-90]^-$ and $[M-H-120]^-$ (Figure 1).[8a,12] The product **1a** was prepared in a scale-up reaction, and its structure, including the attached position of the glucosyl moiety and the anomeric stereochemistry, were elucidated by MS, and ^1H and ^{13}C NMR spectroscopic data analyses and compared with those of maclurin 3-*C*-β-D-glucoside.[13] The large coupling constant ($J = 9.8$ Hz) of the anomeric proton (see Figures S59 and S60 in the Supporting

[*] D. Chen, Dr. R. Chen, Dr. R. Wang, Dr. J. Li, K. Xie, C. Bian, L. Sun, X. Zhang, J. Liu, Prof. Dr. L. Yang, Prof. Dr. F. Ye, Prof. Dr. X. Yu, Prof. Dr. J. 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 (China)
E-mail: jgdai@imm.ac.cn

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201506505.

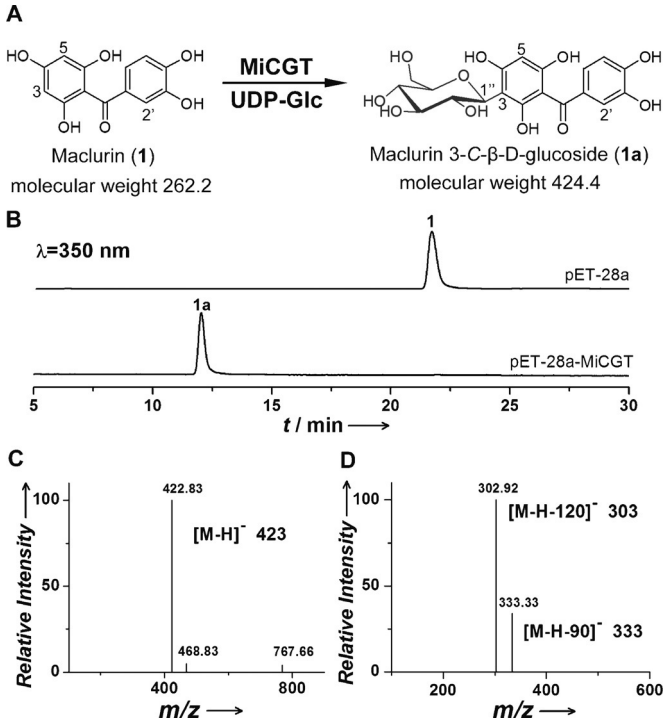

Figure 1. C-glycosylation of **1** catalyzed by recombinant MiCGT. A) The reaction catalyzed by the crude recombinant enzyme. B) HPLC chromatograms of **1** and the enzyme product **1a**. C,D) Typical negative-ion MS and MS^2 spectra for **1a**. The HPLC conditions are provided in Table S2 in the Supporting Information.

Information) indicated the formation of the β-anomer and an inversion mechanism for MiUGT13. The above results unequivocally establish MiUGT13 as a CGT (designated MiCGT, *Mangifera indica* C-glycosyltransferase, GenBank accession No. KT200208). A phylogenetic tree was constructed (see Figure S2 in the Supporting Information) and revealed that MiCGT and other plant CGTs were grouped into a single clade with the highest identity (50%) with UGT708D1 from *Glycine max*.[8e]

Recombinant His_6-MiCGT was purified with His-tag affinity chromatography and analyzed with SDS-PAGE (see Figure S3 in the Supporting Information) for the biochemical characteristics. MiCGT displayed the maximum activity at pH 9.0 and 45 °C, and was divalent-cation-independent (see Figure S4 in the Supporting Information). MiCGT exhibited K_m values of 47.0 μM and 159.2 μM for **1** and **25** (for structures see Figure 2), respectively, and the corresponding k_{cat} values of 1.6 s^{-1} and 0.8 s^{-1} (see Figure S5 in the Supporting Information). Generally, the C-glucosylation in the mangiferin biosynthesis is hypothesized to occur at the benzophenone (**1**) or the xanthone (**25**) stage (see Scheme S1).[11] In our experiments, we observed that MiCGT could C-glycosylate **1** to **1a**, but only O-glycosylate **25**. MiCGT exhibits a k_{cat}/K_m value of $3.4 \times 10^4\ M^{-1} s^{-1}$ to **1**, which was roughly sevenfold that to that for **25** ($5.2 \times 10^3\ M^{-1} s^{-1}$). Cumulatively, MiCGT can be definitively validated as the first benzophenone CGT. Furthermore, maclurin 3-*C*-β-D-glucoside (**1a**) was enzymatically converted into mangiferin (73% yield for 2 h) when incubated with the crude enzyme from *M. indica* leaves (see Figure S6 in the Supporting Information). Taken together, these results provide the first solid experimental biochemical evidence suggesting that maclurin 3-*C*-β-D-glucoside is an intermediate in the biosynthetic pathway of mangiferin in *M. indica* (see Scheme S1).[11]

To explore the catalytic promiscuity and probe the synthetic utility of this novel CGT in vitro, an acceptor library of representative natural and unnatural compounds with structural diversity was assessed with UDP-glucose (Figure 2). The library members included benzophenones (**1**–**4**, **14**–**16**, **37**–**39**), flavonoids (**11**, **41**–**44**), benzil (**17**), xanthones (**18**–**25**), anthraquinone (**40**), and simple aromatics with various substituents such as OH, NH_2, CH_3, OCH_3, F, or Br (**10**, **26**–**29**, **31**–**36**, **45**–**54**). Additionally, synthetic compounds, including alkaloids (**5**, **8**, **9**), thiophene (**6**), naphthalene (**7**), benzoates (**12**, **13**), and ethyl phloroglucinol (**30**), which closely resemble the natural substrate of **1** featuring a phloroglucinol moiety, were designed and chemically synthesized (see the Supporting Information).

An initial hint about the enzyme's unprecedented broad capability for C-glycosylation was provided by the HPLC-UV/MS^2 analysis (Figure 2; see Figures S7–S27 in the Supporting Information), which revealed that MiCGT was sufficiently flexible to C-glycosylate 35 of the 54 library members (65%), including 18 structurally different types. Such an ability to utilize structurally varied natural and unnatural aglycons appears to be more common among OGTs,[5a,14] but is unusual for CGTs. Moreover, MiCGT resulted in high C-glycosylation conversion rates (>80%) with 13 (**1**–**7**, **10**–**12**, **17**, **26**, **27**) of the 35 substrates. Notably, 34 members led only to a single, chromatographically distinct mono-C-glucoside individually, which indicates the regiospecificity of MiCGT. Weak C-glycosylation activity of MiCGT turned out to be observed with the xanthones (**18**–**24**). Interestingly, MiCGT was able to both C- and O-glycosylate 16 substrates (**14**–**16**, **18**–**24**, **31**–**36**; see Table S3 in the Supporting Information). MiCGT also exhibited N-glycosylation activity toward 3,4-dichloroaniline (**54**; see Figure S28 in the Supporting Information). MiCGT was also observed to be able to form C-xylosides (**1b**, **10b**, **27b**, **28b**) with UDP-xylose (see Figures S29–S32 in the Supporting Information). Based on the above results, MiCGT exhibited unusual substrate (acceptor and donor) promiscuity, which rendered MiCGT a promising enzyme for the construction of C-glycoside libraries that exhibit structural and bioactive diversities.

Given that an understanding of the determinants in the acceptor molecules that control the mode of glycosylation is critical for the generation of novel C-glycosyl derivatives, a detailed structural comparison was performed. It demonstrated that MiCGT appeared to generate only C-glycosides with 2,4,6-tri-hydroxy acceptors at the A ring (**1**–**13**), both C- and O- glycosides with 2,4-di-hydroxyl acceptors at the A ring (**14**–**16**), and only O-glycosides with 2- or 4-mono-hydroxy acceptors at the A ring (**38**, **39**). The same results were observed with simple phenolics (**26**–**36**). Hence, the electron density around the carbon atom which undergoes attack, enhanced by electron-donating hydroxy groups, might play an important role in C-glycosylation, and strongly supports the

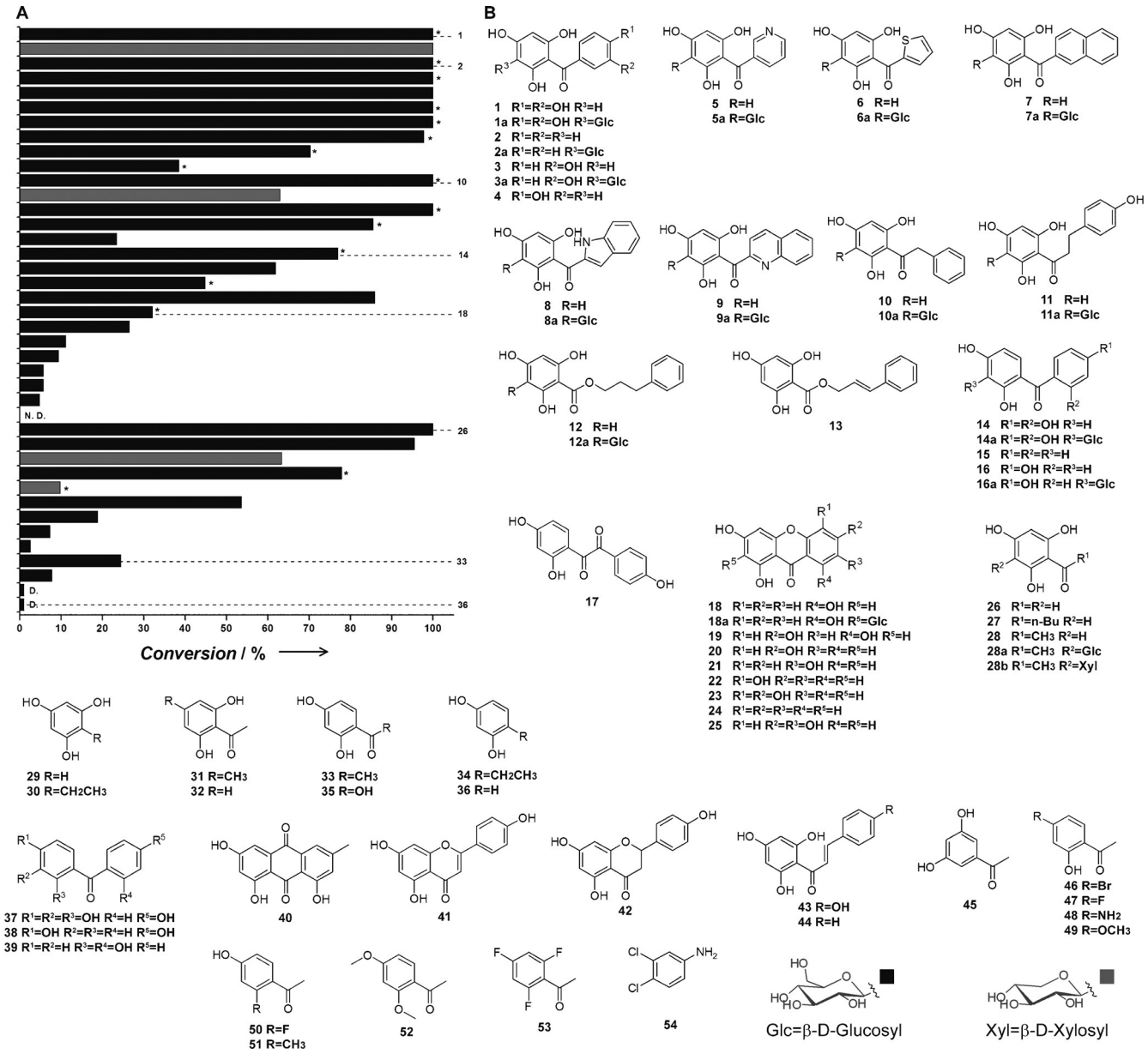

Figure 2. Exploring the catalytic promiscuity of MiCGT. A) Percent conversions of the C-glycosylated products catalyzed by MiCGT. The members are listed based on the structural scaffolds shown in part B. The black and grey columns represent the conversion of C-glucosides and C-xylosides, respectively. B) The structures of the library members and corresponding C-glycosylated products. Compounds **5**, **7**–**9**, **12**, **13**, and **30** are designed novel scaffolds. The O- or N-glucosylated yields of the compounds (**14**–**16**, **18**–**25**, **31**–**54**) are shown in the Supporting Information. * = C-glycosylated products were prepared and confirmed by MS, and ^{1}H and ^{13}C NMR spectroscopy. N.D. = not detected, D. = detected in MS.

occurrence of aryl-C-glycosylation by a direct Friedel–Crafts-like reaction mechanism.[2a,3b,6,7] Additionally, the facts that neither **45** nor various acetophenone derivatives with either electron-donating groups, such as NH_2, CH_3, and OCH_3 (**48**, **49**, **51**, **52**), or electron-withdrawing groups, such as F and Br (**46**, **47**, **50**, **53**), were accepted indicates that the 2,4-dihydroxy substituents on the aromatic ring are of stringent necessity for MiCGT to C-glycosylate. Thus, the number and position of the electron-donating hydroxy group at an aromatic ring are the critical determinants which mediate the C- or O-glycosylation capacity of MiCGT.

Accordingly, the minimal recognition unit of MiCGT for C-glycosylation was identified as resorcinol (**36**), while that of the previously characterized plant CGTs is 2′,4′,6′-trihydroxyacetophenone (**28**).[8a–c] Most importantly, MiCGT was able to catalyze similar scaffolds (**26**–**36**) to the corresponding C-glycosides, which could serve as common building blocks for the chemical syntheses of structurally varied bioactive C-glycosyl natural products and drugs.

Based on the dual specificity for C- and O-glycosylation of MiCGT and the reversibility of OGTs,[5b,f,14,15] we attempted to synthesize C-glycosides with a simple sugar donor by

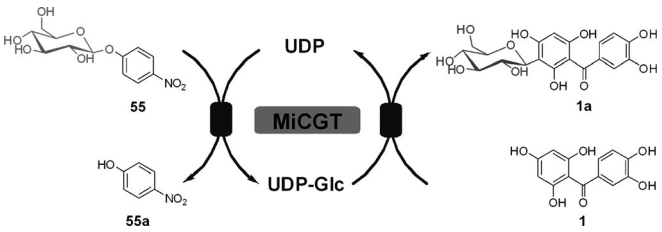

Figure 3. Exploiting the OGT reversibility of MiCGT to generate C-glucoside with a simple sugar donor (the HPLC analysis is shown in Figure S33 in the Supporting Information).

coupled reactions. As shown in Figure 3, the OGT-reversibility of MiCGT was exemplified in the coupling reactions, in which the C-glucosylated product **1a** was obtained in a high yield (77 %) with 4-nitrophenyl-β-D-glucopyranoside (**55**) and only a catalytic amount of UDP (see Figure S33 in the Supporting Information). Without adding relatively costly UDP-glucose, this study provides an economic and environmentally general approach to the C-glycodiversification of bioactive molecules.

To further confirm the catalytic properties of MiCGT and biologically assay the C-glycosides, we obtained 16 C-glycosylated products by preparative-scale reactions, 12 (**3a**, **5a**–**10a**, **12a**, **14a**, **16a**, **18a**, **28b**) of which were novel C-glycosides. Their structures were identified with HRESIMS, and ^{1}H and ^{13}C NMR spectroscopic data. The sugar moieties were observed to be attached at the C3-position of the A ring, a carbon atom in the *ortho*/*para* position to the hydroxy groups, and is similar to the observation of C-glycosylations by all known CGTs.[6–8] The β-glycosidic linkages were deduced from the large coupling constants ($J = 9.7$–9.9 Hz) of the anomeric protons (see Figures S59–S94 in the Supporting Information). Therefore, MiCGT is a regio- and stereospecific C-glycosyltransferase. The resulting 16 C-glycosylated compounds were biologically evaluated for potential effects on type-2 diabetes using human sodium-glucose cotransporter 2 (SGLT2) inhibitory activity in vitro,[16] thus revealing that **10a**–**12a** exhibited potential as lead compounds with the IC_{50} values of $2.6\times$, $7.6\times$, and 7.6×10^{-7} M, respectively.

In summary, the surprising catalytic promiscuity of MiCGT, which is a novel benzophenone CGT involved in the biosynthesis of mangiferin from *M. indica*, was highlighted. MiCGT exhibited robust regio- and stereospecific C-glycosylation activity toward numerous natural and unnatural druglike compounds and simple phenolics with either UDP-glucose or xylose and the abilities to form C-, O-, and N-glycosides. The OGT reversibility of MiCGT was also exploited to economically generate C-glucosides with simple sugar donors without the need for using the expensive UDP-sugar moiety. Three aryl-C-glycosides (**10a**–**12a**) displayed strong inhibitory effects against SGLT2. The present studies have successfully demonstrated, for the first time, the significant potential of an enzymatic approach not only to the chemoenzymatic formation of structurally diverse bioactive aryl-C-glycosides for lead compound discovery, but also for the formation of key and common building blocks for the bioactive compound syntheses. Moreover, the identification of a benzophenone CGT may provide more genetic information for the future verification of additional novel CGTs for the C-glycodiversification of bioactive natural and unnatural products. Given the dual selectivity for both C- and O-glycosylation, MiCGT is probably a good model for future kinetic/mechanistic and structural studies that aim to elucidate the catalytic mechanism and substrate promiscuity, and to address how the enzyme governs the C- and O-glycosylations. The MiCGT reported here hints at more exciting and novel CGTs hiding in plants as enzymatic tools for C-glycosylation in the search for drug leads, and would also facilitate further enzyme engineering for the development of novel biocatalysts in combinatorial biosynthesis.

Acknowledgements

The work was financially supported by the National Natural Science Foundation of China (Grant No. 81273405), and the National Science & Technology Major Project "Key New Drug Creation and Manufacturing", China (Grant No. 2012X09301002-001-005).

Keywords: carbohydrates · enzyme catalysis · glycosylation · regioselectivity · transferases

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 12678–12682
Angew. Chem. **2015**, *127*, 12869–12873

[1] a) A. C. Weymouth-Wilson, *Nat. Prod. Rep.* **1997**, *14*, 99–110; b) A. Chang, S. Singh, G. N. Phillips, J. S. Thorson, *Curr. Opin. Biotechnol.* **2011**, *22*, 800–808.

[2] a) C. J. Thibodeaux, C. E. Melançon, H. w. Liu, *Nature* **2007**, *446*, 1008–1016; b) H. Wang, T. J. Oman, R. Zhang, C. V. Garcia De Gonzalo, Q. Zhang, W. A. van der Donk, *J. Am. Chem. Soc.* **2014**, *136*, 84–87.

[3] a) P. G. Hultin, *Curr. Top. Med. Chem.* **2005**, *5*, 1299–1331; b) T. Bililign, B. R. Griffithb, J. S. Thorson, *Nat. Prod. Rep.* **2005**, *22*, 742–760; c) L. Bungaruang, A. Gutmann, B. Nidetzky, *Adv. Synth. Catal.* **2013**, *355*, 2757–2763.

[4] a) V. Křen, J. Thiem, *Chem. Soc. Rev.* **1997**, *26*, 463–473; b) E. K. Lim, *Chem. Eur. J.* **2005**, *11*, 5486–5494; c) H. Renata, Z. J. Wang, F. H. Arnold, *Angew. Chem. Int. Ed.* **2015**, *54*, 3351–3367; *Angew. Chem.* **2015**, *127*, 3408–3426.

[5] a) S. A. Borisova, C. Zhang, H. Takahashi, H. Zhang, A. W. Wong, J. S. Thorson, H. w. Liu, *Angew. Chem. Int. Ed.* **2006**, *45*, 2748–2753; *Angew. Chem.* **2006**, *118*, 2814–2819; b) C. Zhang, B. R. Griffith, Q. Fu, C. Albermann, X. Fu, I. K. Lee, L. Li, J. S. Thorson, *Science* **2006**, *313*, 1291–1294; c) R. W. Gantt, R. D. Goff, G. J. Williams, J. S. Thorson, *Angew. Chem. Int. Ed.* **2008**, *47*, 8889–8892; *Angew. Chem.* **2008**, *120*, 9021–9024; d) C. J. Thibodeaux, C. E. Melançon III, H. w. Liu, *Angew. Chem. Int. Ed.* **2008**, *47*, 9814–9859; *Angew. Chem.* **2008**, *120*, 9960–10007; e) R. W. Gantt, P. Peltier-Pain, S. Singh, M. Zhou, J. S. Thorson, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 7648–7653; f) K. Xie, R. Chen, J. Li, R. Wang, D. Chen, X. Dou, J. Dai, *Org. Lett.* **2014**, *16*, 4874–4877.

[6] a) M. A. Fischbach, H. Lin, D. R. Liu, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 571–576; b) A. Gutmann, B. Nidetzky, *Angew. Chem. Int. Ed.* **2012**, *51*, 12879–12883; *Angew. Chem.* **2012**, *124*, 13051–13056.

[7] a) D. Hoffmeister, G. Dräger, K. Ichinose, J. Rohr, A. Bechthold, *J. Am. Chem. Soc.* **2003**, *125*, 4678–4679; b) J. Härle, S. Günther,

B. Lauinger, M. Weber, B. Kammerer, D. L. Zechel, A. Luzhetskyy, A. Bechthold, *Chem. Biol.* **2011**, *18*, 520 – 530; c) L. Li, P. Wang, Y. Tang, *J. Antibiot.* **2014**, *67*, 65 – 70; d) H. K. Tam, J. Härle, S. Gerhardt, J. Rohr, G. Wang, J. S. Thorson, A. Bigot, M. Lutterbeck, W. Seiche, B. Breit, A. Bechthold, O. Einsle, *Angew. Chem. Int. Ed.* **2015**, *54*, 2811 – 2815; *Angew. Chem.* **2015**, *127*, 2853 – 2857.

[8] a) M. Brazier-Hicks, K. M. Evans, M. C. Gershater, H. Puschmann, P. G. Steel, R. Edwards, *J. Biol. Chem.* **2009**, *284*, 17926 – 17934; b) M. L. Falcone Ferreyra, E. Rodriguez, M. I. Casas, G. Labadie, E. Grotewold, P. Casati, *J. Biol. Chem.* **2013**, *288*, 31678 – 31688; c) Y. Nagatomo, S. Usui, T. Ito, A. Kato, M. Shimosaka, G. Taguchi, *Plant J.* **2014**, *80*, 437 – 448; d) N. Sasaki, Y. Nishizaki, E. Yamada, F. Tatsuzawa, T. Nakatsuka, H. Takahashi, M. Nishihara, *FEBS Lett.* **2015**, *589*, 182 – 187; e) Y. Hirade, N. Kotoku, K. Terasaka, Y. Saijo-Hamano, A. Fukumoto, H. Mizukami, *FEBS Lett.* **2015**, *589*, 1778 – 1786.

[9] J. C. Barreto, M. T. S. Trevisan, W. E. Hull, G. Erben, E. S. de Brito, B. Pfundstein, G. Würtele, B. Spiegelhalder, R. W. Owen, *J. Agric. Food Chem.* **2008**, *56*, 5599 – 5610.

[10] a) J. Hughes, M. A. Hughes, *DNA Sequence* **1994**, *5*, 41 – 49; b) A. Richman, A. Swanson, T. Humphrey, R. Chapman, B. McGarvey, R. Pocs, J. Brandle, *Plant J.* **2005**, *41*, 56 – 67; c) X. Wang, *FEBS Lett.* **2009**, *583*, 3303 – 3309.

[11] M. Fujita, T. Inoue, *Phytochemistry* **1981**, *20*, 2183 – 2185.

[12] F. Gandia-Herrero, A. Lorenz, T. Larson, I. A. Graham, D. J. Bowles, E. L. Rylott, N. C. Bruce, *Plant J.* **2008**, *56*, 963 – 974.

[13] Y. Zhang, L. Han, D. Ge, X. Liu, E. Liu, C. Wu, X. Gao, T. Wang, *J. Agric. Food Chem.* **2013**, *61*, 1884 – 1895.

[14] a) A. Minami, R. Uchida, T. Eguchi, K. Kakinuma, *J. Am. Chem. Soc.* **2005**, *127*, 6148 – 6149; b) A. Minami, T. Eguchi, *J. Am. Chem. Soc.* **2007**, *129*, 5102 – 5107.

[15] A. Gutmann, C. Krump, L. Bungaruang, B. Nidetzky, *Chem. Commun.* **2014**, *50*, 5465 – 5468.

[16] E. C. Chao, R. R. Henry, *Nat. Rev. Drug Discovery* **2010**, *9*, 551 – 559.

Received: July 14, 2015
Published online: September 1, 2015

