Organic & Biomolecular Chemistry

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: B. Fan, W. Dong, T. Chen, J. Chu and B. He, *Org. Biomol. Chem.*, 2018, DOI: 10.1039/C8OB00376A.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/obc

View Article Online
DOI: 10.1039/C8OB00376A

Journal Name

ARTICLE

Switching glycosyltransferase $UGT_{BL}1$ regioselectivity toward polydatin synthesis using semi-rational design

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

DOI: 10.1039/x0xx00000x

www.rsc.org/

Bo Fan[a], Wenxin Dong[a], Tianyi Chen[a], Jianlin Chu*[b,c] and Bingfang He*[a,b,c]

The 62nd residue of glycosyltransferase $UGT_{BL}1$ was identified as a "hot spot" for glycosylation at 3-OH of resveratrol. Via semi-rational design including structure-guided alanine scanning and saturation mutations, the mutation I62G significantly switched the regioselectivity from 4'-OH to 3-OH of resveratrol and mainly produced polydatin (87.7%), a therapeutic natural product.

Introduction

Resveratrol (*trans*-3,5,4′-trihydroxystilbene) is a natural polyphenol with a wide range of physiological and pharmacological activities[1], such as antioxidative[2], anti-obesity[3, 4], anti-inflammatory[5], and antitumor[6]. Despite the importance of resveratrol, its application is limited because of its poor solubility and bioavailability, which is dramatically improved through the glycosylation of resveratrol[7-9]. Polydatin (3-O-β-glucoside of resveratrol or piceid), with about 18-fold water solubility of resveratrol, possesses improved bioavailability and pharmaceutical properties[8, 10, 11]. Polydatin presents various kinds of biological activities, such as anti-aging[10] and antiproliferative effects[11]. Recently, polydatin was used to relieve the toxic and side effects of cisplatin [12] and treat acute severe hemorrhagic shock [13]. A phase II clinical study of polydatin injection (HW6) is in progress (clinicaltrials.gov: NCT01780129). A growing appreciation for the important biological roles and the therapeutic potential of polydatin has resulted in an increasing interest in its synthesis. Selectivity is key for glycosylation of polyphenols. The 4′-OH group of free resveratrol was verified to be important for various biological activities [8, 9].

In the glycosylation of a compound with several possible positions, the chemists should develop effective protective group technology to achieve the synthesis of the target glycoside[14], requiring sequential manipulation of protecting/deprotecting and resulting in relatively low "atom economy". For example, four *tert*-butyldimethylsilyl (TBS) protected resveratrols were prepared in one pot and separated using column chromatography. Then the 4′-OH-protected resveratrol was further chosen for the synthesis of polydatin[15](with a total yield of about 20%). Regioselectivity is one of the typical problems that must be considered in synthetic organic chemistry, especially in the synthesis of glycosides[16, 17]. For stereo- and regioselective glycosylation, the advantages of biocatalysis over chemical processes are obvious[16].

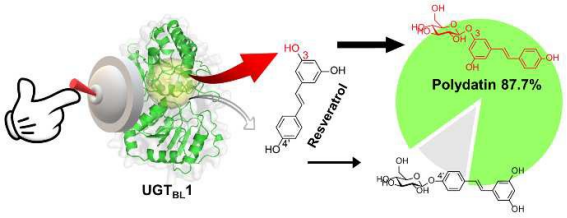

Fig. 1 Switch glycosyltransferase $UGT_{BL}1$ regioselectivity towards polydatin synthesis.

Selective biocatalytic glycosylation avoids using protective group chemistry with exquisite stereoselectivity (α-anomer or β-anomer according to the kind of enzyme). Therefore, it presents a highly attractive route for polydatin synthesis. Polydatin could be produced by plant culture, while the stilbene glucoside derivatives were mixed[18]. In the culture of *Phytolacca americana* cells, the ratio of 3-O-β-glucoside of resveratrol (polydatin) to 4′-O-β-glucoside of resveratrol was about 8:5[19]. The glycosylation of resveratrol was also achieved by glycosyltransferase YjiC from *Bacillus licheniformis* DSM 13, which resulted in a complicated mixture of four glucosides: 3-O-, 4′-O-, 3,5-O-, and 3,5,4′-O-β-glucoside[20]. However, the sucrose phosphorylase TtSPP (R134A) from *Thermoanaerobacterium thermosaccharolyticum* was the most effective for the 3-OH glycosylation of resveratrol[21]. Nevertheless, the anomeric carbon was α-configuration, which

[a] College of Biotechnology and Pharmaceutical Engineering, Nanjing Tech University, 30 Puzhunan Road, Nanjing 211816, China.
[b] School of pharmaceutical sciences, Nanjing Tech University, 30 Puzhunan Road, Nanjing 211816, China.
[c] Jiangsu National Synergetic Innovation Center for Advanced Materials, 30 Puzhunan road, Nanjing 211816, China.
† Correspondence and requests for materials should be addressed to Bingfang He (email: bingfanghe@njtech.edu.cn) and Jianlin Chu (email: cjl2fl@126.com).
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

View Article Online
DOI: 10.1039/C8OB00376A

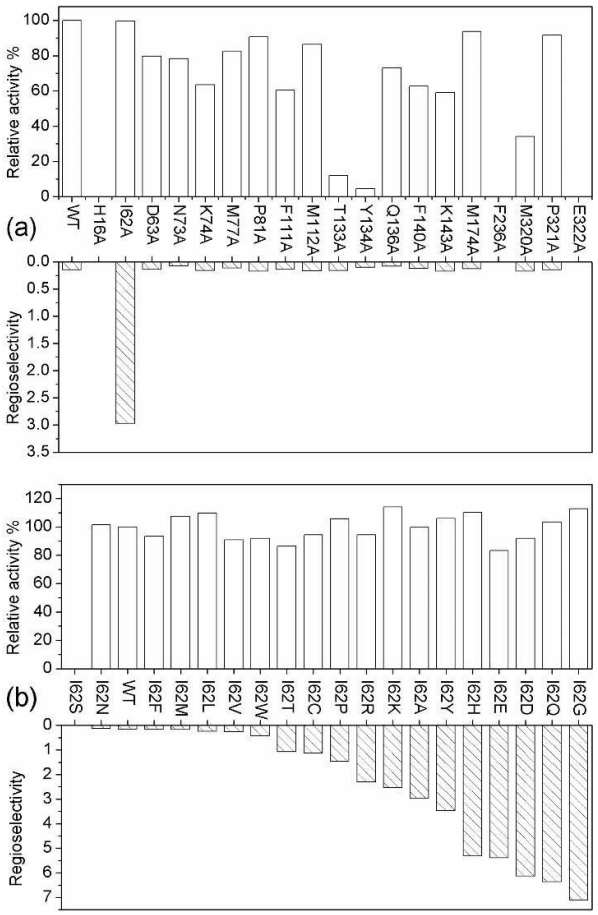

Fig. 2 Activity and regioselectivity of mutants obtained from alanine scanning (a) and Site-saturated mutagenesis of 62nd site (b) of $UGT_{BL}1$. Regioselectivity: ratio of resveratrol-3-O-glucoside (polydatin) to resveratrol-4′-O-glucoside.

was different from the natural polydatin.

The enzymes derived from bacteria show diversity in substrate specificity, but the development of enzyme engineering provides an efficient means for producing designated products[22, 23]. In the engineering of glycosyltransferases, some successes were achieved in modulating the activity and substrate selectivity by structure-guided evolution. For example, the semi-rationally designed glycosyltransferase UGT51 from *Saccharomyces cerevisiae* presented about1800-fold enhanced catalytic efficiency (k_{cat}/K_M) for converting protopanaxadiol to ginsenoside Rh2[24]. The activity of mutant glycosyltransferase OleI (V150A) from *Streptomyces antibioticus* was about 10% higher than that of the wild type[25].

In this study, a glycosyltransferase $UGT_{BL}1$ (GenBank No. AKA94108) from *B. licheniformis* ZSP01[26, 27] was found to catalyze the glycosylation of resveratrol, yielding resveratrol 4′-O-β-glucoside as the main product and polydatin in small amounts (Fig. S3). From this candidate enzyme, $UGT_{BL}1$ was engineered for 3-OH of resveratrol. However, unveiling the mutants displaying improved properties among the astronomically large sequence space to produce them experimentally is wet-lab intensive and practically limited by the screening capacities. Nevertheless, at this point, a strategy for protein engineering combined with structure-guided alanine scanning and saturation mutations was performed to shift the enzyme's specificity toward polydatin formation (Fig. 1).

Results and discussion

$UGT_{BL}1$ displayed a similar conformation to the OleI (Fig. S1), with about 40% of amino acid sequence identity, and the highly conserved backbone structure. In the glycosyltransferase $UGT_{BL}1$, the substrate was accommodated in the cleft formed between the N- and C-terminal domains[28]. Cumulatively, some successful efforts to engineer glycosyltransferases using structure-guided design are evidence of the importance of amino acid residues close to the substrate-binding pockets in regulating the function of the enzymes. For example, the substrate promiscuity of glycosyltransferase OleD from *S. antibioticus* was expanded by the mutation of the amino acid residues around the binding site [29, 30]. Consequently, residues lining the glycosyl acceptor substrate binding site were considered for potential mutagenesis.

In the first round, alanine scanning mutation on the selected 19 residues nearing the glycosyl acceptor substrate binding site was carried out to find the "hotspot" that could significantly influence regioselectivity of $UGT_{BL}1$. These 19 mutants (the primers, see Table S1) were constructed, and their catalytic activity and regioselectivity to resveratrol were analyzed (Fig. 2a). About one-third of these mutants retained more than 80% activity compared with the wild type, while the mutants H16A, Y134A, F236A, andE322A were almost inactive. Most mutants showed no significant change in regioselectivity. Fortunately, one of the mutants, I62A, significantly switched 4′-O-β-glucoside to 3-O-β-glucoside of resveratrol. The ratio of polydatin to resveratrol-4′-O-glucoside was approximately 3. The key amino acid hot spot 62nd residue, which contributed to the regioselectivity of $UGT_{BL}1$ for polydatin synthesis was identified in the first round. The site 62 was present in a loop region (loop N3) following β-sheet 3 in the N-terminal domain (Fig. S2)[28]. Coincidentally, the importance of residues within the loop N3 with respect to urdamycin (an angucycline-type antibiotic) specificity was shown in studies with the UrdGT1b and UrdGT1c from *S. fradiae*.[31] The amino acid residues located in the loop N3 of UrdGT1b and UrdGT1c were engineered and the engineering glycosyltransferases were used to synthesize the new urdamycin derivatives. In the mutational study of OleD (from *S. antibioticus*), the single mutation (P67T) that contributed the most to substrate specificity was also located within the loop N3[29]. In this study, the present glycosyltransferase mutagenesis highlighted the importance of key residues within the loop N3, which might respond to glycosyltransferase acceptor specificity. It was greatly recommended as a potential focal point for future glycosyltransferase engineering efforts and presented a new scaffold for further (semi-)rational redesign.

View Article Online
DOI: 10.1039/C8OB00376A

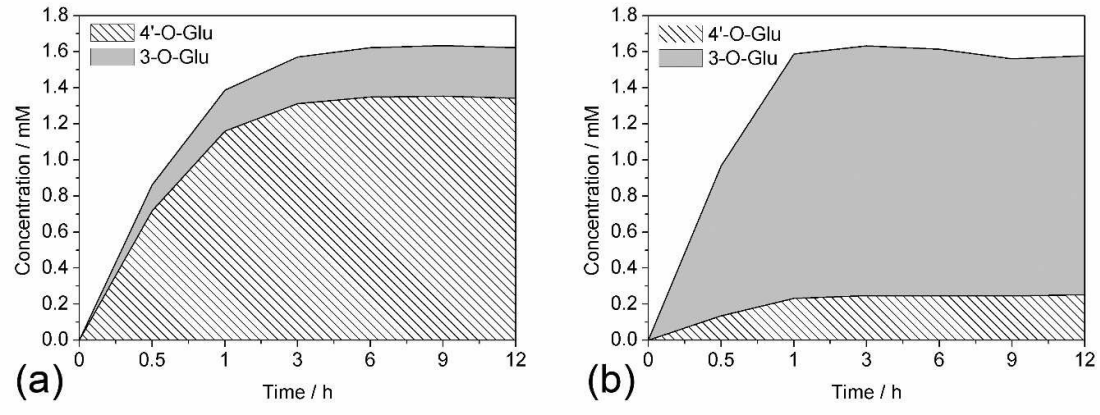

Fig. 3 Time course of the glycosylation of resveratrol by $UGT_{BL}1$ wild type (a) and I62G (b).

Then, the 62nd site of $UGT_{BL}1$ was selected for further saturation mutation to further optimize the selected lead catalyst. All of these mutants retained more than 80% activity compared with the wild type, except for the mutant I62S. From the second screening, the mutants I62G, I62Q, I62D, I62E, and I62H also showed marked preference for the glycosylation of 3-OH. The best mutant, I62G, displayed the largest improvement in regioselectivity compared with I62A, and the ratio of polydatin to resveratrol-4′-O-glucoside was more than 7. Fortunately, the mutant I62G markedly changed the regioselectivity without abolishing but slightly improving the activity (Fig. 2b). In the study of UGT71G1 from *Medicago truncatula*, mutation resulted not only in a change of regioselectivity to flavonol quercetin but also in a prominent decrease in the total activity[32].

The time courses of resveratrol glucosides catalyzed by $UGT_{BL}1$ wild-type and I62G were determined (Fig. 3). A maximal production of glucosides were obtained after a 3-h reaction catalyzed by mutant I62G with the same amount of enzymes, while the reaction catalyzed by wild-type $UGT_{BL}1$ needed more than 6 h to reach the reaction platform. In this work, with an equimolar ratio of UDP-glucose and resveratrol, the conversion of resveratrol reached an equilibrium at about 80%. The conversion rate of resveratrol could be enhanced while increasing the ratio of UDP-glucose to resveratrol (data not shown). The ratio of polydatin/resveratrol-4′-O-β-glucoside was strikingly different for mutant I62G compared with that for the wild type. This can be explained by the apparent kinetic parameters of these two enzymes (Table 1). The wild-type $UGT_{BL}1$ displayed a lower K_M (0.29 mM) toward 4′-OH of resveratrol and a higher K_M (0.37 mM) toward 3-OH of resveratrol, indicating that the wild type $UGT_{BL}1$ had a greater affinity for 4′-OH. The mutant I62G displayed a lower K_M (0.29 mM) for 3-OH and a higher K_M (0.85 mM) for 4′-OH; the enormous increase in K_M for 4′-OH reflected a significant decrease in affinity for 4′-OH. The affinity of mutant I62G for 3-OH was much higher than that for 4′-OH, suggesting that the selectivity of $UGT_{BL}1$ was mainly determined by kinetics. The evidence from enzyme kinetic characterization also supported the idea that in terms of catalytic efficiency (k_{cat}/K_M). The mutant I62G presented about sevenfold higher k_{cat}/K_M (377.2) for 3-OH than for 4′-OH (k_{cat}/K_M: 52.0). According to k_{cat}/K_M, the I62G mutant preferred polydatin formation over 4′-O-glucoside formation by roughly sevenfold, while the result of wild type $UGT_{BL}1$ was just the opposite (Fig. S3). The replacement of Ile to Gly at the 62nd site of $UGT_{BL}1$ could significantly mediate the regioselectivity of glycosylation of resveratrol.

Conclusion

In summary, two rounds of mutation (alanine scanning and saturation mutation) were carried out for the residues near the substrate binding site based on the semi-rational design strategy. The “hot spot” was found and the exploratory mutational experiments provided a mutant that enhanced the regioselectivity and activity in one case. This study illustrated the ability of structure-guided evolution to significantly alter glycosyltransferase regioselectivity via engineering in the loop N3 and showed that this region might have the potential for the further engineering. It paved the way for a sustainable and scalable biocatalytic process for producing polydatin. Engineered glycosyltransferase exhibiting enhanced regioselectivity could be of interest for diverse glycoside synthesis and drug discovery[22, 23]. The development described in this study might stimulate the generation of novel glycosylation catalysts using (semi-)rational design of the active site pocket of the known glycosyltransferases. Furthermore, the crystallization of $UGT_{BL}1$ and its complex with substrate is ongoing in our laboratory. The high-quality protein crystal structure, in conjunction with kinetic analyses of the wild type enzyme and point mutation variant, might facilitate the exploration of the molecular mechanism underlying glycosyltransferase regioselectivity glycosylation.

Experimental

Materials

Resveratrol was obtained from Nanjing Zelang Medical

View Article Online
DOI: 10.1039/C8OB00376A

	Resveratrol			4'-OH			3-OH		
Enzyme	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M
Wild type	0.31	129.8	418.5	0.29	109.1	376.0	0.37	20.9	56.4
I62G	0.33	145.9	442.0	0.85	44.2	52.0	0.29	109.4	377.2

Table 1 Apparent kinetic parameters for wild type $UGT_{BL}1$ and mutant I62G.

Technology Co.,Ltd, UDP-glucose was obtained from Sigma-Aldrich. Other solvents and reagents used in this study were of analytical grade from commercial sources.

Gene cloning, plasmid construction and mutagenesis

The wild-type $UGT_{BL}1$ gene, *ugt*BL1 (Genbank No., Nucleotide: KP123426; Protein: AKA94108), was introduced into the *Eco*R I and *Xho* I sites of pET28a to generate pET28a-$UGT_{BL}1$ the expression plasmid.

The site-directed mutations were introduced using a modified whole plasmid polymerase chain reaction(PCR) method[33]. The plasmid pET28a-$UGT_{BL}1$ was used as a template. The primers used in the site-directed mutagenesis are listed in the Table S1. The PCR cycles were as follows: 95 °C for 1 min, 98 °C for 10 s, and 68 °C for 7.5 min, with steps 2–3 were run for 30 cycles. The PCR system contained a 2 μL template (5 ng/mL), 1 × 2 μL forward and reverse primers (10 μM), 5 μL primerSTAR buffer, 2.5 μL dNTP mixture (2.5 mM), and 0.25 μL PrimeSTAR HS DNA polymerase (2.5 U/μL, TaKaRa, Dalian, China). The obtained PCR product was digested with *Dpn* I (TaKaRa, Dalian, China) and used for transformation to *Escherichia coli* BL21 (DE3). All of the mutants were confirmed by sequencing (Genewiz, Suzhou, China).

Protein expression and purification

The recombinant *E. coli* harboring the appropriate plasmid were cultured in the Lysogeny broth (LB) medium containing 50 μg/mL kanamycin at 37 °C and 200 rpm. When the OD_{600} reached approximately 0.8, the protein expression was induced with 1.0 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) for 6-8 h at 30 °C. The cells were harvested via centrifugation at 12,000 × *g* for 10 min. To purify the target proteins. The cells were resuspended in buffer A (20 mM Tris-HCl, 500 mM NaCl, 20 mM imidazole, and 10% glycerol; pH 7.9) and lysed by sonication to purify the target proteins. The target proteins were purified by nickel-affinity chromatography on Ni-NTA resin (GE Healthcare, Chicago, IL, USA) at AKTA prime plus system (GE Healthcare, Chicago, IL, USA). The clear supernatant of cell lysate was loaded onto the Ni-NTA column. The column was washed with five times volume of buffer B (20 mM Tris-HCl, 500 mM NaCl, 100 mM imidazole, and 10% glycerol; pH 7.9), and then eluted using buffer C (20 mM Tris-HCl, 500 mM NaCl, 150 mM imidazole, and 10% glycerol; pH 7.9). The fractions containing purified protein were concentrated, and the buffer was changed to 50 mM sodium phosphate buffer (pH 7.0) using Amicon Ultra Centrifugal Filter Devices (10 kDa, Millipore, Billerica, MA, USA). Protein concentrations were determined using the bicinchoninic acid (BCA) method with bovine serum albumin (BSA) as the standard. The purified proteins were confirmed by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) (Fig. S4).

Enzyme assays

The reactions were performed in a typical volume of 100 μL containing 2 mM resveratrol and 2 mM UDP-glucose, 50 μg/mL enzyme, and 50 mM sodium phosphate buffer (pH 8.0) to determine the activity and regioselectivity of $UGT_{BL}1$ and its mutant-catalyzed glycosylation. Enzymatic conversion was carried out at 30 °C with an agitation rate of 500 rpm using a thermo-shaker (DITABIS, Pforzheim, Germany). The reaction was stopped by adding 900 μL methanol. The mixture was centrifuged at 12,000 rpm for 10 min, and the supernatant was used for analysis by high-performance liquid chromatography (HPLC).

The apparent kinetic parameters were analyzed by varying resveratrol at a fixed concentration of UDP-glucose.

HPLC and HRMS analysis

The supernatants of sample were filtered through a 0.22 μm nylon filter, and analyzed by HPLC using a C18 column (4.6 mm×250 mm, 5 μm, Kromasil, Bohus, Sweden). The column temperature and flow rate were 30 °C and 1.0 mL/min, respectively. A binary gradient was used. The two solvents were: H_2O (mobile phase A) and methanol (mobile phase B). The method was: 20% (4 min), 20–50% (11 min), 50% (3 min), 50–100% (8 min), 100% (5 min), 100–20% (2 min), 20% (2 min). UV detection was at 310 nm.

The molecular weight of products in the HPLC peaks were analyzed with high resolution mass spectrometry (HRMS) (Fig. S5). The HRMS analysis was performed in positive ion mode on Agilent 6520 Accurate-Mass Q-TOF LC/MS platform (Agilent Technologies, Santa Clara, CA, USA).

Production, purification and identification of the product

Upscaling of the transglycosylation reactions on resveratrol was performed up to 100 mL, contained 2 mM resveratrol, 2mM UDP-glucose, 100 μg/mL enzyme, and 50 mM sodium phosphate buffer (pH 8.0). Threefold volumes of methanol were added and mixed. The mixture was centrifuged at 12000 rpm for 10 min, and the supernatant was concentrated by rotary evaporation. The concentrate was further purified on a semi-preparative C18 reverse-phase column (20.0 mm×250 mm, 10 μm, SinoChrom, Dalian, China). The flow rate was 10 mL/min, and H_2O was used as the mobile phase A and methanol was used as the mobile phase B. The percentage of mobile phase B was manually increased from 10% by increments of 1% per minute until all products were eluted. The collected factions with the same resveratrol glucoside were pooled together and concentrated on a rotary evaporator. The residual water was removed by freeze-drying.

View Article Online
DOI: 10.1039/C8OB00376A

The chemical structure of the purified products was identified by the analysis of ^{1}H NMR, ^{13}C NMR and Heteronuclear Multiple-Bond Correlation (HMBC) spectroscopy (polydatin: Fig. S6–S8; resveratrol-4′-O-β-glucoside: Fig. S9–S11, Table S2). All NMR data were recorded in dimethyl sulfoxide–*d*6 on a 400-MHz NMR spectrometer (Bruker, Rheinstetten, Germany).

Homology modeling

Homology modeling was performed with the SWISS-MODEL server[34, 35] using the reported crystal structure (PDB No. 2IYA) of glycosyltransferase OleI[25] from *Streptomyces antibioticus* as a template. The resulting model was superimposed onto the crystal structure of the template OleI in complex with substrate oleandomycin. The residues with 3.5 Å from substrate amount to the resveratrol in homology model $UGT_{BL}1$, were selected for the mutagenesis in the first round of screening. The amino acid residue that significantly affected the regioselectivity of $UGT_{BL}1$ was selected for saturation mutagenesis in the second round of screening.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (21506099 and 81673321). We also acknowledge the fund sponsored by the Program for Innovative Research Team in Universities of Jiangsu Province (2015) and the Jiangsu Synergetic Innovation Center for Advanced Bio-Manufacture.

Notes and references

1 J. A. Baur and D. A. Sinclair, *Nat. Rev. Drug Discov.*, 2006, **5**, 493-506.
2 L. Lu, S. Zhu, H. Zhang, F. Li and S. Zhang, *RSC Adv.*, 2015, **5**, 14114-14122.
3 Y. Qiao, J. Sun, S. Xia, X. Tang, Y. Shi and G. Le, *Food Funct.*, 2014, **5**, 1241-1249.
4 A. Gracia, J. Miranda, A. Fernandez-Quintela, I. Eseberri, M. Garcia-Lacarte, F. I. Milagro, J. A. Martinez, L. Aguirre and M. P. Portillo, *Food Funct.*, 2016, **7**, 1680-1688.
5 J. Walker, K. Schueller, L.-M. Schaefer, M. Pignitter, L. Esefelder and V. Somoza, *Food Funct.*, 2014, **5**, 74-84.
6 D. Vergara, P. Simeone, D. Toraldo, P. Del Boccio, V. Vergaro, S. Leporatti, D. Pieragostino, A. Tinelli, S. De Domenico, S. Alberti, A. Urbani, M. Salzet, A. Santino and M. Maffia, *Mol. Biosyst.*, 2012, **8**, 1078-1087.
7 L. Biasutto, E. Marotta, A. Bradaschia, M. Fallica, A. Mattarei, S. Garbisa, M. Zoratti and C. Paradisi, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 6721-6724.
8 A. Lepak, A. Gutmann, S. T. Kulmer and B. Nidetzky, *ChemBioChem*, 2015, **16**, 1870-1874.
9 I. Medina, D. Alcantara, M. J. Gonzalez, P. Torres, R. Lucas, J. Roque, F. J. Plou and J. C. Morales, *J. Agr. Food Chem.*, 2010, **58**, 9778-9786.
10 H. Wen, X. H. Gao and J. H. Qin, *Integr. Biol.*, 2014, **6**, 35-43.
11 C. E. Storniolo, P. Quifer-Rada, R. M. Lamuela-Raventos and J. J. Moreno, *Food Funct.*, 2014, **5**, 2137-2144.
12 S. Ince, D. A. Acaroz, O. Neuwirth, H. H. Demirel, B. Denk, I. Kucukkurt and R. Turkmen, *Food Chem. Toxicol.*, 2014, **72**, 147-153.
13 X. M. Wang, R. Song, Y. Y. Chen, M. Zhao and K. S. Zhao, *Expert Opin. Inv. Drug*, 2013, **22**, 169-179.
14 B. Yu, J. Sun and X. Yang, *Acc. Chem. Res.*, 2012, **45**, 1227-1236.
15 Z. Zhang, B. Yu and R. R. Schmidt, *Synthesis*, 2006, 1301-1306.
16 X. M. Zhu and R. R. Schmidt, *Angew. Chem. Int. Ed.*, 2009, **48**, 1900-1934.
17 J. Wang, G. Li and M. T. Reetz, *Chem. Commun.*, 2017, **53**, 3916-3928.
18 D. Donnez, P. Jeandet, C. Clément and E. Courot, *Trends Biotechnol.*, 2009, **27**, 706-713.
19 D. Sato, N. Shimizu, Y. Shimizu, M. Akagi, Y. Eshita, S.-i. Ozaki, N. Nakajima, K. Ishihara, N. Masuoka and H. Hamada, *Biosci. Biotechnol. Biochem.*, 2014, **78**, 1123-1128.
20 R. P. Pandey, P. Parajuli, J. Y. Shin, J. Lee, S. Lee, Y.-S. Hong, Y. I. Park, J. S. Kim and J. K. Sohng, *Appl. Environ. Microbiol.*, 2014, **80**, 7235-7243.
21 M. E. Dirks-Hofmeister, T. Verhaeghe, K. De Winter and T. Desmet, *Angew. Chem.*, 2015, **127**, 9421-9424.
22 J. B. McArthur and X. Chen, *Biochem. Soc. T.*, 2016, **44**, 129-142.
23 J. Schmid, D. Heider, N. J. Wendel, N. Sperl and V. Sieber, *Front. Microbiol.*, 2016, **7**, 182.
24 Y. Zhuang, G. Yang, X. Chen, Q. Liu, X. Zhang, Z. Deng and Y. Feng, *Metab. Eng.*, 2017, **42**, 25-32.
25 D. N. Bolam, S. Roberts, M. R. Proctor, J. P. Turkenburg, E. J. Dodson, C. Martinez-Fleites, M. Yang, B. G. Davis, G. J. Davies and H. J. Gilbert, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 5336-5341.
26 B. Fan, T. Chen, S. Zhang, B. Wu and B. He, *Sci. Rep.*, 2017, **7**, 463.
27 S. Zhang, Z. Zhou, Z. Yao and B. He, *Biochem. Eng. J.*, 2013, **71**, 105-110.
28 S. A. Osmani, S. Bak and B. L. Moller, *Phytochemistry*, 2009, **70**, 325-347.
29 G. J. Williams, C. Zhang and J. S. Thorson, *Nat. Chem. Biol.*, 2007, **3**, 657-662.
30 R. W. Gantt, R. D. Goff, G. J. Williams and J. S. Thorson, *Angew. Chem. Int. Ed.*, 2008, **47**, 8889-8892.
31 D. Hoffmeister, B. Wilkinson, G. Foster, P. J. Sidebottom, K. Ichinose and A. Bechthold, *Chem. Biol.*, 2002, **9**, 287-295.
32 X.-Z. He, X. Wang and R. A. Dixon, *J. Biol. Chem.*, 2006, **281**, 34441-34447.
33 J. Sanchis, L. Fernández, J. D. Carballeira, J. Drone, Y. Gumulya, H. Höbenreich, D. Kahakeaw, S. Kille, R. Lohmer, J. J.-P. Peyralans, J. Podtetenieff, S. Prasad, P. Soni, A. Taglieber, S. Wu, F. E. Zilly and M. T. Reetz, *Appl. Microbiol. Biotechnol.*, 2008, **81**, 387-397.
34 K. Arnold, L. Bordoli, J. Kopp and T. Schwede, *Bioinformatics*, 2006, **22**, 195-201.
35 N. Guex, M. C. Peitsch and T. Schwede, *Electrophoresis*, 2009, **30**, S162-S173.