Methylglucosylation of aromatic amino and phenolic moieties of drug-like biosynthons by combinatorial biosynthesis

Linan Xie[a,1], Liwen Zhang[a,1], Chen Wang[a,b,c,d,1], Xiaojing Wang[a,b], Ya-ming Xu[b], Hefen Yu[e], Ping Wu[c,d], Shenglan Li[e], Lida Han[a], A. A. Leslie Gunatilaka[b], Xiaoyi Wei[c,d], Min Lin[a,2], István Molnár[b,2], and Yuquan Xu[a,2]

[a]Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, 100081 Beijing, People's Republic of China; [b]Natural Products Center, University of Arizona, Tucson, AZ 85706; [c]Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, 510650 Guangzhou, People's Republic of China; [d]Guangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Chinese Academy of Sciences, 510650 Guangzhou, People's Republic of China; and [e]Department of Biochemistry and Molecular Biology, School of Basic Medicine, Capital Medical University, 100069 Beijing, People's Republic of China

Edited by Jerrold Meinwald, Cornell University, Ithaca, NY, and approved April 23, 2018 (received for review September 13, 2017)

Glycosylation is a prominent strategy to optimize the pharmacokinetic and pharmacodynamic properties of drug-like small-molecule scaffolds by modulating their solubility, stability, bioavailability, and bioactivity. Glycosyltransferases applicable for "sugarcoating" various small-molecule acceptors have been isolated and characterized from plants and bacteria, but remained cryptic from filamentous fungi until recently, despite the frequent use of some fungi for whole-cell biocatalytic glycosylations. Here, we use bioinformatic and genomic tools combined with heterologous expression to identify a glycosyltransferase–methyltransferase (GT–MT) gene pair that encodes a methylglucosylation functional module in the ascomycetous fungus *Beauveria bassiana*. The GT is the founding member of a family nonorthologous to characterized fungal enzymes. Using combinatorial biosynthetic and biocatalytic platforms, we reveal that this GT is a promiscuous enzyme that efficiently modifies a broad range of drug-like substrates, including polyketides, anthraquinones, flavonoids, and naphthalenes. It yields both *O*- and *N*-glucosides with remarkable regio- and stereospecificity, a spectrum not demonstrated for other characterized fungal enzymes. These glucosides are faithfully processed by the dedicated MT to afford 4-*O*-methylglucosides. The resulting "unnatural products" show increased solubility, while representative polyketide methylglucosides also display increased stability against glycoside hydrolysis. Upon methylglucosidation, specific polyketides were found to attain cancer cell line-specific antiproliferative or matrix attachment inhibitory activities. These findings will guide genome mining for fungal GTs with novel substrate and product specificities, and empower the efficient combinatorial biosynthesis of a broad range of natural and unnatural glycosides in total biosynthetic or biocatalytic formats.

glycosyltransferase | *O*-methyltransferase | combinatorial biosynthesis | fungi | polyketide

Discovery of the next generation of human and veterinary medications and agrochemicals requires access to structurally diverse chemical matter occupying a drug-like chemical space. Combinatorial biosynthesis ventures to supplement combinatorial chemistry, structure-based drug design, and medicinal chemistry to produce such complex chemical matter by exploiting the unparalleled specificity and efficiency of enzymes, and the scalability and mild conditions of microbial fermentations.

Glycosylation is a critical determinant for the biological activities of clinically important natural products, including antibiotics such as vancomycin and erythromycin, or antineoplastic agents such as doxorubicin (1–4). Glycosylation also increases the water solubility and frequently improves the stability and bioavailability of drug-like small molecules (5, 6), inspiring extensive efforts to synthesize varied *O*- and *C*-glycosides in a process termed glycodiversification (7, 8). Chemical glycosylation of complex scaffolds suffers from limited regioselectivity and low stereospecificity, and may be impeded by the unavailability of modified sugar donors. Biological methods, such as in vivo total biosynthesis, whole-cell biocatalysis, and in vitro enzymatic synthesis with isolated glycosyltransferases (GTs) promise a one-pot, one-step, scalable process to realize regio- and stereospecific glycosylations, but suffer from relatively low conversion rates and a limited substrate range. Thus, identification and characterization of efficient GTs with the right balance of substrate promiscuity and product specificity is necessary to obtain useful catalysts (3). Until very recently, bioprospecting to clone such promiscuous GTs that are able to modify drug-like small molecules considered only plants and bacteria as source organisms. However, this choice was highly anomalous because fungi have long been one of the most widely used whole-cell biocatalysts for the glycosylation of various substances. Nevertheless, only

Significance

Glycosylation imparts improved pharmacokinetic and pharmacodynamic properties to many drug candidates. Here we identify the founding member of a new glycosyltransferase (GT) family from *Beauveria bassiana* that is not orthologous to GTs isolated from other fungi. This GT is clustered with a methyltransferase (MT) from a family hitherto characterized only from bacteria. This GT–MT biosynthetic module shows extensive promiscuity in conjugating methylglucose to structurally varied substrates, but yields products with substantial regio- and stereoselectivity. We demonstrate an efficient combinatorial biosynthetic platform to produce glycosylated polyketides unprecedented in nature, some with increased stability and bioactivity. We also use a biocatalytic platform to synthesize methylglucosides of flavonoids, anthraquinones, and naphthalenes, some with an *N*-glucosidic linkage not previously demonstrated with characterized fungal enzymes.

Author contributions: L.Z., A.A.L.G., M.L., I.M., and Y.X. designed research; L.X., L.Z., C.W., X. Wang, Y.-m.X., H.Y., P.W., S.L., A.A.L.G., and X. Wei performed research; H.Y., P.W., L.H., X. Wei, and M.L. contributed new reagents/analytic tools; L.X., L.Z., C.W., Y.-m.X., A.A.L.G., X. Wei, I.M., and Y.X. analyzed data; and L.Z., C.W., I.M., and Y.X. wrote the paper.

Conflict of interest statement: I.M. has a disclosed financial interest in TEVA Pharmaceuticals Hungary, which is unrelated to the subject of the research presented here. All other authors declare no competing financial interests.

This article is a PNAS Direct Submission.

[1]L.X., L.Z., and C.W. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: linmin57@vip.163.com, imolnar@email.arizona.edu, or xuyuquan@caas.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1716046115/-/DCSupplemental.

Published online May 14, 2018.

narrow-spectrum sterol 3β-glucosyltransferases were identified from various yeast strains until recently. While this work was in progress, the discovery of two disparate, nonorthologous phenolic GT groups from three basidiomycete fungi were reported (MhGT1 from *Mucor hiemalis*, and UGT58A1/UGT59A1 from *Rhizopus japonicus* and *Absidia coerulea*, respectively) (6, 9). These discoveries indicated that despite their very low similarity to known enzymes from other organisms (6, 9, 10), promiscuous yet regio- and stereospecific GTs from the kingdom Fungi can indeed be identified and harnessed for synthetic biology (6, 9). Nevertheless, fungal genome sequence data clearly indicate that a much larger variety of diverse GTs remain to be discovered, especially from the filamentous ascomycetes that dedicate 1–2% of their genomes to as yet uncharacterized GTs (6, 9).

The entomopathogenic ascomycetous fungus *Beauveria bassiana* is a useful biocatalyst due to its ability to catalyze various biotransformations, including methylation, hydroxylation, and oxidation (11–19). It also readily glycosylates a large variety of phenolic compounds, such as flavonoids, anthraquinones, and other polyketides (15–17, 20, 21), including benzenediol lactones (BDLs), such as curvularin (21) (**14** in *SI Appendix*, Fig. S1) and desmethyl-lasiodiplodin (DLD) (**1** in Fig. 1 and *SI Appendix*, Fig. S1). BDLs are drug-like fungal secondary metabolites with an astonishing range of bioactivities (22–26), structurally defined by a 1,3-benzenediol moiety fused to a macrocyclic lactone ring. Based on the carbons involved in this fusion, BDLs can be subdivided into resorcylic acid lactones (RALs, C2–C7 as in **1**) and dihydroxyphenylacetic acid lactones (DALs, C3–C8 as in **14**). RALs may incorporate a 14-membered macrocycle (RAL_{14}) such as in radicicol (**6** in *SI Appendix*, Fig. S1), which displays cancer cell antiproliferative and heat-shock modulatory activities. RALs may also contain a 12-membered ring (RAL_{12}), such as in **1**, which shows mineralocorticoid receptor antagonist and prostaglandin biosynthesis inhibitory activities in animals. Natural product DALs most often feature 12-membered macrocycles (DAL_{12}), including 10,11-dehydrocurvularin (**15** in *SI Appendix*, Fig. S1), which modulates the heat-shock response and the immune system. We have been developing combinatorial biosynthetic methods in an engineered *Saccharomyces cerevisiae* production chassis to increase the chemical space accessible by BDL biosynthesis. Thus, we have revealed structural clues that allow morphing RALs and DALs into each other by polyketide synthase (PKS) active site engineering (25); improved the production of unnatural BDL congeners by creating hybrid PKSs (27); and developed PKS subunit shuffling as a practical method to biosynthesize unnatural BDL scaffolds (23) and BDL/azaphilone hybrids (22). Because no glycosylated BDLs are known from fungal producers (26), identifying the *B. bassiana* GT, and co-opting this enzyme for the total biosynthesis of BDL glycosides in a synthetic biology platform with engineered PKSs, would open up an orthogonal chemical dimension for BDLs.

In host tissues, small-molecule glycosides are frequently broken down to their constituent aglycones by glycoside hydrolases of the host or its associated microbiota. Some aglycones show reduced bioactivities (28), but conversely, glycoside hydrolysis may also be exploited in a prodrug strategy to release a highly active aglycone in situ. Modulation of the sensitivity of glycosides to host glycoside hydrolases is thus important to optimize the pharmacokinetic and pharmacodynamic (PK/PD) properties of drugs. Bioactive bacterial glycosides typically contain modified sugar biosynthons (8, 29), while *O*- or *C*-methylation of some flavonoids dramatically increases their metabolic stability (30). In contrast to other fungal biocatalysts, *B. bassiana* typically produces glycosides in which OH-4 is methylated (2, 15, 16). Such a methylation may alter the resistance of these glycosides to hydrolysis. Considering that BDLs failed to live up to their remarkable in vitro potential up to now due to their metabolic instability in vivo (26), modulation of their solubility, metabolic stability, and bioavailability by (methyl)glucosylation may increase their attractiveness for drug development.

Here, we use bioinformatic and genomic tools combined with heterologous expression to identify a glycosyltransferase–methyltransferase (GT–MT) gene pair that encodes a methylglucosylation functional module in *B. bassiana*. The GT is a promiscuous enzyme that efficiently modifies a broad range of natural and unnatural BDLs, as well as drug-like anthraquinone, flavonoid, and naphthalene scaffolds to yield *O*- and *N*-glucosides with remarkable regio- and stereospecificity. These glucosides are faithfully processed by the dedicated MT to afford 4-*O*-methylglucosides. The resulting glycosides show increased solubility, and representative methylglucosides also display higher stability against glycoside hydrolysis. Upon methylglucosylation, some BDL derivatives attain cell line-specific antiproliferative activity or cancer cell–matrix attachment inhibitory activity. These findings will assist further genome mining for fungal GTs with novel substrate- and product-specificities, and empower the efficient combinatorial biosynthesis of a broad range of natural and unnatural glycosides in total biosynthetic or biocatalytic formats.

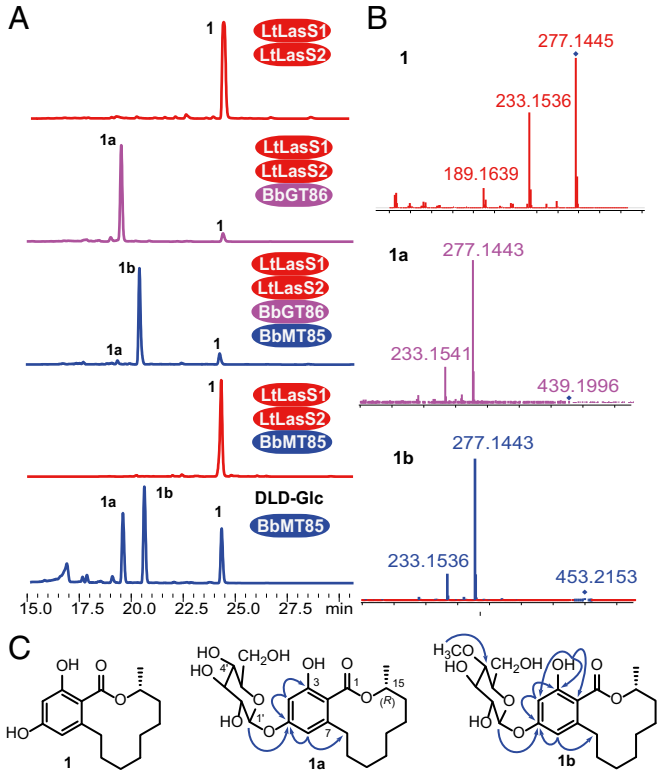

Fig. 1. Identification of the GT–MT pair responsible for the methylglucosylation of desmethyl-lasiodiplodin (**1**). (*A*) Product profiles (reversed-phase HPLC traces recorded at 300 nm) of *S. cerevisiae* BJ5464-NpgA (68) cotransformed with the genes for the indicated *L. theobromae* PKSs (LtLasS1 and LtLasS2) and *B. bassiana* proteins (BbGT86, UDP-dependent GT BBA_08686; BbMT85, FkbM-type MT BBA_08685). DLD-Glc: Purified desmethyl-lasiodiplodin 5-*O*-β-D-glucopyranoside (**1a**) was fed to the culture. (*B*) HRMS/MS spectra of DLD (**1**) and its derivatives. (*C*) Key HMBC correlations (blue arrows) of **1a** and **1b**.

Results

Genes for a 4-*O*-Methylglucose Biosynthon from *B. bassiana*. Previous experiments by us and others indicated that *B. bassiana* is able to methylglucosylate BDL scaffolds, such as DLD (**1** in Fig. 1), curvularin (21) (**14** in *SI Appendix*, Fig. S1), and dihydroresorcylide (31). However, identification of the corresponding enzymes was nontrivial due to the highly divergent primary sequences of the GT

superfamily enzymes, the large number of genes encoding GTs and MTs in fungal genomes, and the complete lack of characterized fungal phenolic GTs at the start of this work. To locate the genes for the enzymes introducing the methylglucose biosynthon into BDLs, we first annotated 57 GT-encoding genes in the genome sequence (32) of *B. bassiana* ARSEF 2860 (*SI Appendix*, Table S1). One of these, BBA_04817, appeared to be severely truncated at its N terminus, and thus was excluded from subsequent analysis. Five indicators were then used to further prioritize the remaining 56 putative GTs for functional screening. First, 12 proteins were shown to belong to GT superfamily-1, many members of which are known to glycosylate small-molecule natural products (1) (*SI Appendix*, Table S2). Next, a collection of 28 voucher GTs was assembled from the genomes of plants, bacteria, and insects that glucosylate substrates similar to those biotransformed by *B. bassiana* (*SI Appendix*, Table S3). A conserved domain search indicated that 22 of these vouchers contain a YjiC-like domain (COG1819) that is present in a large number of uridine 5′-diphosphate (UDP) glucosyltransferases from bacteria, including those that glucosylate flavonoids and macrolides (33, 34). Seven of the 12 superfamily-1 GTs of *B. bassiana* also featured this conserved domain (*SI Appendix*, Table S2). Third, RxnFinder was used to search for enzymes that mediate glucosylation of phenolic substrates and use UDP-glucose as the sugar donor (35, 36) (*SI Appendix*, Table S4). Similarity searches ($e < 0.1$) with representative members of these enzymes returned 6 of the 12 superfamily-1 *B. bassiana* GTs (*SI Appendix*, Table S2). Fourth, we ascertained the expression of *B. bassiana* GTs in a transcriptomic study under culture conditions conducive to methylglucosylation (*SI Appendix*, Table S2). Finally, we considered that *B. bassiana* almost exclusively produces 4-*O*-methylglucopyranosides (15, 16, 20). Promisingly, three of the candidate GT genes were clustered with genes encoding putative MTs, although two of these three GT–MT pairs were only weakly expressed (*SI Appendix*, Table S2).

There are 136 MTs encoded in the *B. bassiana* genome, of which 14 are putative *O*-MTs (*SI Appendix*, Table S5). A multipronged winnowing of these MTs was less informative than the one described above for the GTs. In particular, RxnFinder failed to identify enzymes that methylate *O*-glucosides at OH-4 of the hexose. The most relevant MT was RebM (EC 2.1.1.164) that affords rebeccamycin bearing a *N*-(4-*O*-methyl)glucoside moiety (37). A Blastp search with RebM as the bait returned 19 predicted *B. bassiana* proteins containing MT domains ($e < 0.1$) (*SI Appendix*, Table S6); however, none of these MTs were clustered with GTs.

Considering all of these indicators, the encoding genes of eight putative *B. bassiana* GTs (*SI Appendix*, Table S2) were separately expressed in *S. cerevisiae* BJ5464-NpgA, a host well suited for the expression of fungal enzymes and the reconstitution of fungal polyketide biosynthetic pathways (22–25). DLD (**1**) was selected as the model substrate to represent the BDL polyketide family, based on its efficient conversion by *B. bassiana* ARSEF 2860 to a putative DLD methylglycoside, as verified by HPLC-MS/MS (*SI Appendix*, Fig. S2). To avoid any potential cell permeability issues with an externally supplied substrate, we elected to produce DLD in situ in the recombinant yeast strains by coexpressing the *Lasiodiplodia theobromae* highly reducing PKS-nonreducing PKS pair LtLasS1–LtLasS2 (24) with the target GTs. Strains expressing seven of the eight recombinant GTs yielded only unmodified DLD. However, the strain expressing BBA_08686 (henceforth, BbGT86) produced **1a** as the major product with increased polarity, and only trace amounts of DLD detectable (Fig. 1*A*). The mass-to-charge ratio of the $[M-H]^-$ ion of **1a** was 162 amu higher than that of DLD, indicating a hexose derivative (m/z 439.1996, calculated 439.1974 for the **1a** parent ion in HRMS/MS).

Next, we coexpressed LtLasS1, LtLasS2, and BbGT86 with the MT genes of *B. bassiana* that were found to be clustered with GT-encoding genes (*SI Appendix*, Table S2). Two of these MTs, BBA_03580 and 03582, failed to further modify **1a**. However, the strain expressing BBA_08685 (henceforth, BbMT85), the cognate MT clustered with BbGT86, afforded **1b** as the major product with a decreased polarity compared with **1a**, while producing only trace amounts of DLD **1** or DLD glycoside **1a** (Fig. 1*A*). HRMS/MS analysis of **1b** was consistent with a methylhexose derivative of **1** (m/z 453.2153, calculated 453.2130 for the $[M-H]^-$ parent ion).

The structures of glycosides **1a** and **1b** were elucidated by analyzing the NMR spectroscopic data of the purified compounds. The aglycone of **1a** and **1b** was confirmed as DLD by comparing the ^1H and ^{13}C NMR data with those published (24), and by analyzing the ^1H-^1H COSY, HSQC, and HMBC NMR spectra (*SI Appendix*, Fig. S3 and Table S7). HMBC correlation between the anomeric proton of the hexose (**1a**, δ_H 5.07; **1b**, δ_H 5.04) and C-5 of the aglycone (**1a**, δ_C 162.7; **1b**, δ_C 162.6), together with the presence of a low-field signal for OH-3 (**1a**, δ_H 11.59; **1b**, δ_H 11.62), revealed that the hexose moiety was attached to OH-5. The *O*-β-glycosidic linkage was confirmed by the large coupling constant (**1a**, $J = 7.5$ Hz; **1b**, $J = 7.7$ Hz) of the anomeric proton, and is in agreement with the inversion of the anomeric configuration of the NDP-α-D-sugar substrate upon transfer to the aglycone by superfamily-1 GT enzymes (29). The identity of the hexose fragment as glucose was affirmed by the characteristic ^{13}C NMR resonances (**1a**, signals at δ_C 100.9, 78.00, 77.96, 74.7, 71.29, 62.6), and as D-glucose by acid hydrolysis of **1a** and comparison of the specific rotation of the isolated sugar, $[\alpha]_{25}^{D}$ + 99.6 (c 0.20, H_2O), with published data (38). Finally, HMBC correlation between the methoxyl protons (δ_H 3.56) and C-4 (δ_C 80.0) of the glucose moiety verified that the methylation in **1b** has taken place on OH-4. Thus, **1a** was identified as desmethyl-lasiodiplodin 5-*O*-β-D-glucopyranoside, while **1b** was shown to be desmethyl-lasiodiplodin 5-*O*-β-D-(4-*O*-methyl)glucopyranoside.

Taking these data together, we find that the clustered genes for BbGT86–BbMT85 encode a GT–MT pair that effectively biotransforms the BDL model substrate DLD (**1**) to yield the unprecedented methylglucoside **1b**.

Characterization of the *B. bassiana* GT–MT Pair. BbGT86 shares up to 94% amino acid sequence identity with a number of uncharacterized proteins uncovered by genome-sequencing projects of Hypocreales and other ascomycete fungi, many of which are annotated as putative superfamily-1 GTs with a UDP-glucose binding motif. Apart from these uncharacterized fungal proteins, BbGT86 shows the highest similarity (up to 35% sequence identity) to YjiC-like prokaryotic GTs, including biochemically characterized UDP-dependent glycosyltransferases (UGTs), such as those taking part in the glycosylation of macrolides and vancomycin-like nonribosomal peptides. The UDP Glycosyltransferase Nomenclature Committee (39) recognized BbGT86 as the founding member of a new UGT family, and assigned the systematic name UGT61A1 to this enzyme. We have reconstructed the UGT phylogenetic tree based on functionally characterized representatives of 156 named UGT families recognized by the Committee (Fig. 2*A*). The tree topology reveals two large and well-defined clades formed by the animal and virus UGTs on the one hand, and most plant UGTs on the other. Separating these is a small clade consisting of two very recently described basidiomycete phenolic UGT families (9) exemplified by UGT58A1 from *Absidia coerulea* and UGT59A1 from *Rhizopus japonicus*, each showing only marginal sequence similarities to BbGT86 (identities of 13% with UGT58A1, and 15% with UGT59A1). Basal to these clades is a large and taxonomically diverse collection of enzymes that glycosylate small-molecule phenolic, macrolide, sterol, and fatty acyl substrates, including BbGT86 and its yet uncharacterized Hypocrealean orthologs of the UGT61 family. The UGT61 family forms a well-resolved clade, and is sister to a clade of the UGT101 and 102 families of bacterial enzymes, such as OleD (UGT101A1) from *Streptomyces antibioticus*, that is involved in macrolide antibiotic resistance (40), and CrtX (UGT102B1) from *Pantoea ananatis* that catalyzes the glucosylation of zeaxanthin (41). Importantly, the UGT61 clade is very distant from

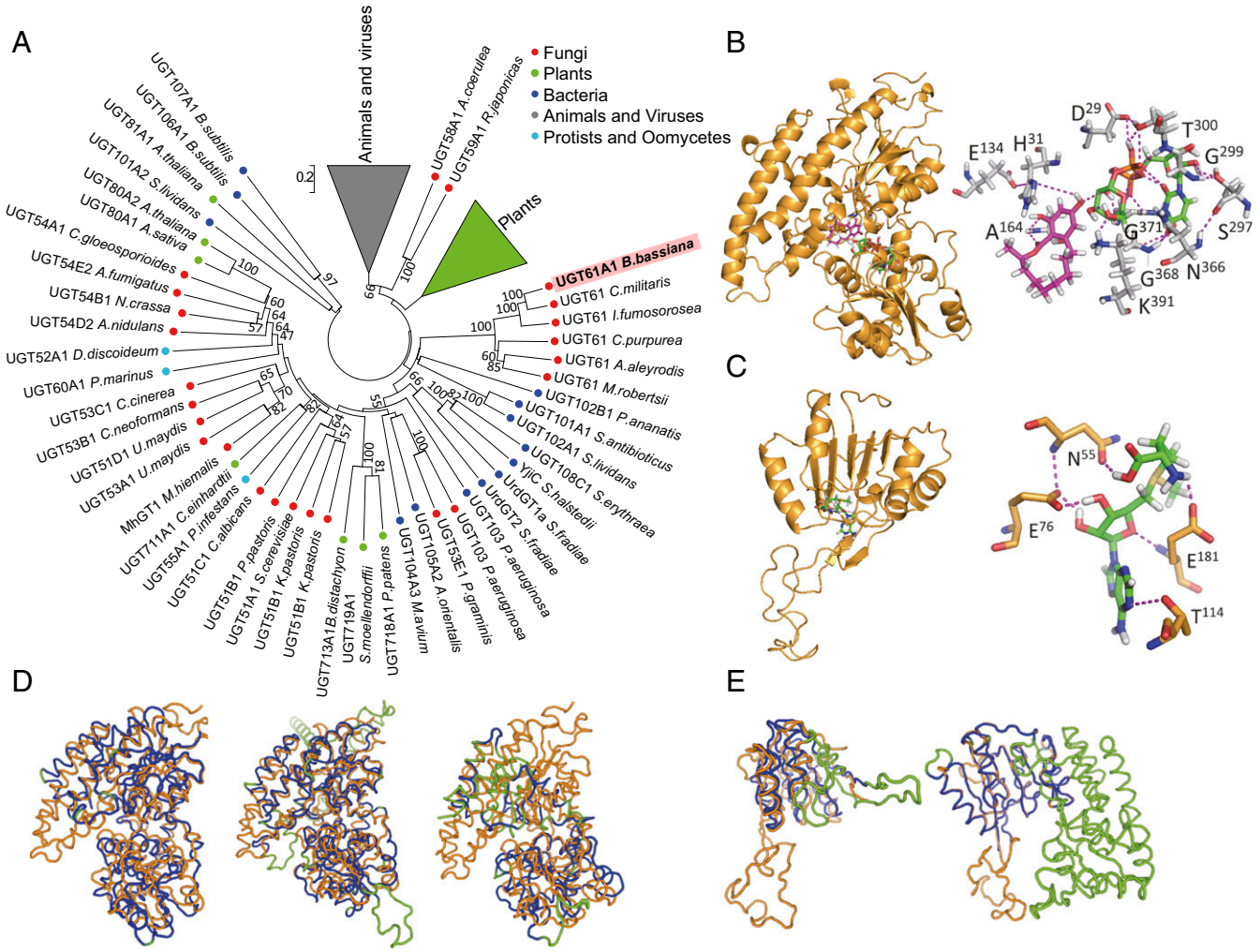

Fig. 2. Phylogenetic and structural analysis of the BbGT86–BbMT85 pair. (*A*) A phylogenetic tree of superfamily-1 UGTs. The clades for UGTs from plants and from animals and viruses were collapsed and represented by the colored triangles. The origins of the enzymes are color-coded as indicated. (*B*) Cartoon representation of a comparative de novo structural model of BbGT86, and its active site canyon with docked substrates UDP-glucose (carbon atoms in green) and DLD (carbons in magenta), both represented as sticks. The deduced active site dyad H^{31} and E^{134} and other residues that may make polar contacts (dashes in magenta) with the substrates are also shown as sticks (carbons in gray) (*SI Appendix, SI Results*). (*C*) Homology structural model of BbMT85 with docked substrate SAM (rendered as sticks with carbon atoms in green). Residues shown as sticks (carbons in gold) may make polar contacts (dashes in magenta) with SAM (*SI Appendix, SI Results*). (*D*) Pairwise structure superimpositions of BbGT86 (in gold) with, from left to right, OleD [Dali server (69) *z*-score 32.6, rmsd of atomic positions 4.1]; UGT58A1 (*z*-score 28.7, rmsd 3.2); and MhGT1 (*z*-score 11.0, rmsd 4.8). (*E*) Pairwise structure superimpositions of BbMT85 (in gold) with, from left to right, 3E05 (*z*-score 12.1, rmsd 2.5); and 2PY6 (*z*-score 21.7, rmsd 0.6). Cartoon representations of the comparator proteins in *D* and *E*: blue, conserved structures; green, divergent structures.

other clades that include characterized fungal UGTs, including the recently described phenolic UGT from the Mucoromycota *Mucor hiemalis* (6) (MhGT1, 15% identity to BbGT86) that in turn clades with basidiomycete sterol UGTs from families 51 and 53.

Structure predictions using Phyre2 (www.sbg.bio.ic.ac.uk/), THMM (www.cbs.dtu.dk/services/TMHMM/), and SignalP (www.cbs.dtu.dk/services/SignalP/) suggested that BbGT86 is an intracellular enzyme with no signal peptide or transmembrane domains. Because of the very low primary amino acid sequence identity of BbGT86 to structurally characterized UGTs, we built a comparative de novo structure model using the Robetta prediction server (42). The resulting model depicts a typical GT-B structural fold (29) that includes an N-terminal acceptor substrate recognition domain and a C-terminal UDP-glucose binding domain, divided by a wide and relatively open canyon for substrate binding (Fig. 2*B*). The volume of this canyon (6,908 $Å^3$), as measured by CASTp (43), exceeds that of the macrolide glucosyltransferase OleD (PDB 4m82.2.A, 5,469 $Å^3$), a representative of the UGT101–102 families that form a sister clade to the UGT61 family. The volume of this canyon also exceeds that of the fungal phenolic glucosyltransferase MhGT1 (5,819 $Å^3$), but is similar to that of UGT58A1 (7,173 $Å^3$), a representative of the third fungal phenolic glucosyltransferase clade. The overall architecture of BbGT86 is in a good agreement with those of OleD and UGT58A1, but only shows reasonable structural similarity to MhGT1 in the C-terminal UDP-glucose binding domain (Fig. 2*D*). Docking of UDP-glucose and DLD to the predicted BbGT86 structure (Fig. 2*B* and *SI Appendix, SI Results*) places these substrates at the deduced conserved active site dyad (29), and suggests that H^{31} acts as a catalytic base that abstracts the proton from OH-5 of DLD, while the transition state may be stabilized by E^{134}. A nucleophilic attack on C-1 of glucose would then break the glycosidic bond with UDP, leading to a direct displacement with inversion of configuration in an overall S_N2-like mechanism.

BbMT85 is an *S*-adenosylmethionine (SAM)-dependent MT of the FkbM family (TIGR 01444). This large and overwhelmingly prokaryotic family features only a few fungal proteins, most from

Hypocreales and other Sordariomycetes, with BbMT85 showing the highest similarities (up to 81% identities) to these uncharacterized fungal FkbM MTs. Among the prokaryotic enzymes, BbMT85 is orthologous to many FkbM MTs from Actinomycetes, including an *O*-MT from *Streptomyces hygroscopicus* ssp. *ascomyceticus* that modifies the dihydroxycyclohexane side chain of the macrolactam FK520 (44) (AAF86398, 34% identity to BbMT85). However, BbMT85 shows no similarity to RebM (Q8KZ94), the rebeccamycin *N*-glucosyl-4′-*O*-methyltransferase from *Lechevalieria aerocolonigenes*. Because FkbM MTs are known to modify polyketide, nonribosomal peptide or flavonoid small-molecule natural products but not sugars, we ascertained that BbMT85 does not methylate DLD or other aglycone small-molecule substrates of BbGT86 (Fig. 1). Instead, the OH-4 functionality of the glucosides created by BbGT86 serves as the substrate: for example, 5-*O*-glucosyl-DLD (**1a**) is readily converted to **1b** by a yeast strain expressing BbMT85 (Fig. 1). This specificity is in accord with bacterial enzymes biosynthesizing polyketide natural products decorated with *O*-methylglycosyl biosynthons, whereby methylation typically does not happen at the NDP-sugar stage, but follows the formation of the glycoside (8). Homology structural modeling of BbMT85 was attempted based on the only available experimentally determined crystal structure of an FkbM-family MT, ABE5011 from *Methylobacillus flagellatus* (PDB ID code 2PY6). This provided a model with a partial overlap with the template structure, covering only the SAM-binding domain (Fig. 2*E*). This part of BbMT85 (but not the rest of the protein) also revealed reasonable structural similarity to the SAM-binding regions of various other MTs, including the precorrin-6y *C*-5, 15-methyltransferase of *Geobacter metallireducens* (PDB ID code 3E05) (Fig. 2*E*). Docking SAM into the BbMT85 model provided a pose with a good overlap with the SAM residing in the experimentally determined 2PY6 structure (Fig. 2*C* and *SI Appendix*, *SI Results*). However, attempts to dock DLD 5-*O*-β-D-glucopyranoside (**1a**) into the BbMT85–SAM model failed to provide any reasonable poses, due to the limitations of the model as evident from the large unstructured loop region (Fig. 2 *C* and *E*).

GT–MT pairs orthologous to BbGT86–BbMT85 are encoded in the genomes of many hypocrealean fungi, but none of these orthologs has been functionally characterized. Interestingly, unlike many bacterial UGTs and FkbM-type MTs, none of the hypocrealean GT–MT pairs is a part of any recognizable small-molecule natural product biosynthetic gene cluster. Even more strikingly, there is no synteni in Hypocreales genomes around the GT–MT pairs (*SI Appendix*, Fig. S4), suggesting that while the GT–MT pair constitutes a coevolving biosynthetic unit, it is functionally independent from other biosynthetic pathways. This and the considerable aglycone promiscuity of BbGT86–BbMT85 (see below) suggests that this GT–MT pair, and by extension its hypocrealean orthologs, act in nature as a phase II/phase III detoxification module against allelochemicals and xenobiotics.

Combinatorial Biosynthesis of BDL Methylglucosides. The utility of the BbGT86–BbMT85 pair for synthetic biology hinges on its ability to conjugate the methylglucopyranoside biosynthon with various drug-like scaffolds in an efficient manner. To investigate the substrate ambiguity of the GT–MT module, we first assembled a collection of model substrates that represent the natural and "unnatural" (combinatorial biosynthetic) structural space of BDLs (*SI Appendix*, Fig. S1*A*). This included RAL_{12} (DLD **1**, *trans*-resorcylide **2**, and radiplodin **3**); RAL_{14} [monocillin II **4**, radicicol **6**, lasicicol **9**, *R*-zearalane **10**, and *trans*-14 (15)-dehydrozearalenol (DHZ) **12**]; RAL_{16} (sedecicol **13**); DAL_{12} [curvularin **14**, 10 (11)-dehydrocurvularin **15** and 11-hydroxycurvularin **16**]; and DAL_{14} (radilarin **17** and lasilarin **18**). We also tested nonmacrocyclic BDL congeners, such as isocoumarins (**19**–**21**); acyl-resorcylic acids (ARA) (**22**–**24**); an acyl-dihydroxyphenylacetic acid (ADA) (**25**); and the benzaldehyde precursor (**26**) of the azaphilone asperfuranone (22) (*SI Appendix*, Fig. S1*A*). Most of these compounds were produced in situ by coexpressing the relevant PKS pairs with the GT–MT pair in the *S. cerevisiae* chassis, although we resorted to feeding some substrates whose de novo production in the GT–MT-expressing yeast would have been too cumbersome. As shown in Fig. 3, the BbGT86–BbMT85 pair successfully methylglucosylated most of these compounds, indicating that the size of the BDL macrocycle, the geometry of the benzenediol lactone (RAL vs. DAL), or even the presence of the macrocyclic ring itself is not crucial for substrate recognition. For the non-macrocyclic congeners, the presence of an α-pyrone ring, the length of the linear chain, or the presence of ester or aldehyde functionalities was not an impediment either. Methylglucosides derived from six RALs, two DALs, three isocoumarins, and three ARAs were isolated, and their structures were elucidated by comparing their ^{1}H and ^{13}C NMR data with those published for their aglycones (22–24, 45), and by analyzing their NMR spectra to ascertain the position of the sugar moiety (*SI Appendix*, Fig. S3 and Table S7). The glycoconjugates of the remaining compounds, including minor glycosylation products, were confirmed by HR-MS/MS (*SI Appendix*, Table S8). All these glycosides represent an orthogonal structural dimension without precedent among natural fungal BDLs.

Methylglucosyl derivatives invariably dominated the product profiles for all BDL congeners, indicating that BbMT85 efficiently methylates the nascent BDL glucoside intermediates (Fig. 3). Notably, BbGT86 maintained a considerable regioselectivity despite the vast structural differences of the acceptor substrates. Di-glycosyl derivatives were not detected. The 5-*O*-β-D-(4-*O*-methyl)glucopyranoside derivatives were the main products for the RALs and their congeners, although 3-*O*-glycosylated regioisomers were also detected at low levels for some compounds (Fig. 3 and *SI Appendix*, Table S8). In contrast, a mixed regioselectivity was observed when DAL-type BDL substrates were offered to BbGT86. Thus, a 7-*O*-β-D-(4-*O*-methyl)glucopyranosyl derivative was exclusively obtained with lasilarin (**18**), a regiospecificity identical to that found earlier during curvularin biotransformation with *B. bassiana* (21). In contrast, an 5-*O*-β-D-(4-*O*-methyl)glucopyranosyl derivative was the only methylglucoside obtained with radilarin (**17**) as the substrate. Attachment of the sugar unit to OH-7 of the aglycone in the lasilarin methylglucoside **18b** was confirmed by detecting an HMBC correlation between the anomeric proton of the hexose (δ_H 4.96, d, J = 7.4 Hz) and C-7 of the aglycone (δ_C 157.6). Glycosylation of OH-5 in **17b** was affirmed by the HMBC correlation of the anomeric proton (δ_H 4.90, d, J = 7.4 Hz) with C-5 of the aglycone (δ_C 160.5) (*SI Appendix*, Fig. S3 and Table S7), and by detecting a chelated OH (OH-7) at δ_H 9.83 in its ^{1}H NMR spectrum in DMSO-d_6 (*SI Appendix*, Fig. S15*B* and Table S7.4). The observed regioselectivity for OH-5 in RALs and radilarin (**17**) is coincident with a stronger acidity of that group [e.g., for DLD (**1**), a pK_a of 8.7 for OH-5 vs. 9.8 for OH-3]. However, the preference of the GT for OH-7 in lasilarin (**18**) is contrary to a weaker acidity of that group (e.g., pK_a of 8.9 for OH-7 vs. 7.8 for OH-5 in **18**). This indicates that the chemical reactivity of the target OH is not the sole determinant of regiospecificity for this GT, and indeed reactivity can be overruled by controlled substrate positioning in the active site pocket of the enzyme.

Nevertheless, not all BDL substrates were accepted, revealing some interesting constraints on the substrate promiscuity of BbGT86. First, BDLs featuring the exocyclic methyl group on a carbon with an *S*-configuration allow only low conversion rates, as seen with curvularin **14** and DHZ **12** (*SI Appendix*, Fig. S1 and Table S9). We validated this constraint by comparing the methylglucosylation efficiencies of *R*-zearalane **10** with that of semisynthetic *S*-zearalane **11**. As expected, **10** turned out to be an excellent substrate, while **11** was only marginally modified (Fig. 3 and *SI Appendix*, Table S9). Second, the presence of an

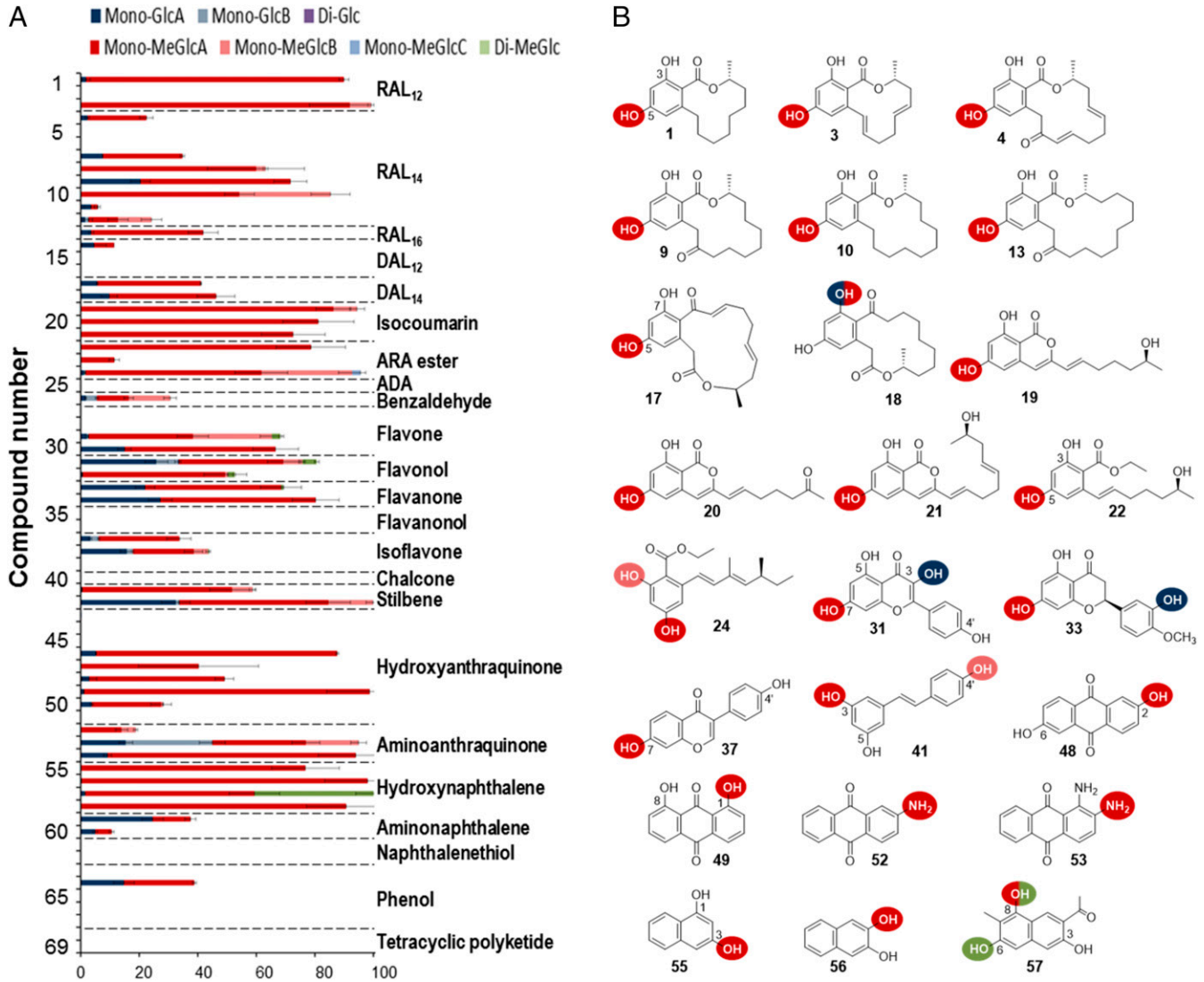

Fig. 3. (Methyl)glucosylation of BDLs and other drug-like compounds. (*A*) Bioconversion of 69 substrates into various glycosides by *S. cerevisiae* BJ5464-NpgA expressing the BbGT86–BbMT85 biosynthetic module. All data represent the means ± SDs in three independent experiments. Mono-GlcA and -B, mono-glucoside regioisomers; Di-Glc, di-glucoside; Mono-MeGlcA, -B, and -C, mono(4-*O*-methyl)glucoside regioisomers; Di-MeGlc, di-(4-*O*-methyl)glucoside. (*B*) Chemical structures of aglycones for which glucosides or methylglucosides were isolated and structurally characterized. Colors indicate the acceptor position(s) of the sugar unit(s) in the glucosides (blue) or (4-*O*-methyl)glucosides (red) or di-(4-*O*-methyl)glucosides (green). See *SI Appendix*, Fig. S1 for additional structures and *SI Appendix*, Tables S8 and S9 for details.

enone double bond is inhibitory toward methylglucosylation in 12-membered macrocycles, but not in 14-membered BDLs. Thus, the enone-DAL_{12} 10 (11)-dehydrocurvularin **15** was not modified, while the expected (21) *O*-methylglucopyranoside of curvularin **14** was detected. Similarly, the enone-RAL_{12} *trans*-resorcylide **2** was not accepted as a substrate, despite the reported ability of *B. bassiana* to methylglucosylate 10,11-dihydroresorcylide (31). At the same time, the enone-RAL_{14} monocillin II (**4**) and the enone-DAL_{14} radilarin **17** were readily modified by the GT-MT pair (Fig. 3 and *SI Appendix*, Table S9). Third, while monocillin II (**4**) was readily accepted as a substrate, its more advanced congener (46), radicicol **6**, was not modified. We disentangled the contributions of the various modifications in **6** toward the inhibition of methylglucosylation. We found that chlorination at C-6 is not an impediment for the GT–MT pair, because pochonin D **8** (also known as 6-chloromonocillin II) was an even better substrate than monocillin II, just as the 6-chloro derivative of dihydroresorcylide was readily methylglucosylated by a *B. bassiana* biocatalyst (31). Consistently, monocillin I (**5**), the deschloro derivative of radicicol, could not be methylglucosylated. The epoxide ring of radicicol was not a barrier for glucosylation either, since pochonin A (**7**) with the 14 (15)-epoxide was still readily modified (Fig. 3 and *SI Appendix*, Table S9). Taken together, this identifies the *cis*-12 (13)-double bond as the structural motif that inhibits methylglucosylation by the BbGT86–08685 pair.

The BbGT86–BbMT85 Pair Acts as a Broad Spectrum Methylglucosylation Module. To further explore the aglycone promiscuity of the BbGT86–BbMT85 pair, we attempted to methylglucosylate a series of compounds in a biocatalytic format with the GT–MT-expressing *S. cerevisiae* strain. The compounds represented diverse drug-like molecules, such as flavonoids, stilbenes, anthraquinones, naphthalenes, and tetracyclic compounds, and a few simple aromatic compounds (Fig. 3 and *SI Appendix*, Fig. S1*B* and Table S9). Of the 43 tested substrates, 25 were successfully glycosylated, as shown by HRMS/MS (Fig. 3 and *SI Appendix*, Tables S8 and S9). Fifteen

representative glucosylated or methylglucosylated compounds, nine of which are new to nature, were then isolated from scaled-up reactions and their structures were elucidated by NMR (*SI Appendix*, Fig. S3 and Table S7). Similar to the case of BDLs and their congeners, di-glycosides were rarely observed (with the notable exception of **57**, a hydroxynaphthalene shunt product of the *Monascus* azaphilone pigments) (47). Most aglycones yielded a single prominent methylglucoside, accompanied by minor amounts of glycoconjugate regioisomers (Fig. 3 and *SI Appendix*, Table S9). In general, fewer glucosylation target sites were found per aglycone than those reported for MhGT1 (6). However, this comparison is made more complicated by the different assay formats (in vitro enzyme assays for MhGT1), the possible variability of the uptake of the aglycones by the cells expressing BbGT86-BbMT85, or the partial degradation of some glycosides in vivo (see the next section) (28).

For the four representative flavonoids where the prominent products were isolated and structurally characterized, OH-7 *para* to the C-4 carbonyl was favored for methylglucosylation. Glucosylation at the pyran OH-3 was also facile for the flavon-3-ol kaempferol **31**, but the resulting glucoside was not accepted by the MT. Notably, other flavonoids, such as apigenin **30** and naringenin **34**, the isoflavone genistein **38**, and 4-*iso*-propyl-3,5-dihydroxystilbene **42** also yielded relatively large amounts of glycosides lacking methylation, although the position of the glycosidic bond was not determined in these cases. Conversely, the phenolic ring of the stilbenoid *trans*-resveratrol **41** was readily methylglucosylated. A flavonoid lacking the C-4 carbonyl (compare catechin **35** and quercetin **32**) (15), a flavone with OH-6 (baicalein **27** vs. apigenin **30**), or the chalcone (*seco*-flavone) isoliquiritigenin **40** were not accepted, and preglucosylated flavonoids (baicalin **28**, puerarin **39**), or the flavonolignan silybin **36** were similarly not transformed. Interestingly, the branched substituent in 4-*iso*-propyl-3,5-dihydroxystilbene **42** was not an impediment for the GT–MT pair for an efficient bioconversion (Fig. 3 and *SI Appendix*, Table S9).

Biotransformation of hydroxyanthraquinones was also readily achieved for five of the nine compounds tested, with two representative major compounds isolated and their structures elucidated by NMR (Fig. 3 and *SI Appendix*, Fig. S3 and Tables S7–S9). OH-1 or OH-2 *O*-β-D-(4-*O*-methyl)glucosides (20) were readily biosynthesized using compounds **46**–**50**, but 1- or 2-hydroxyanthraquinone **43** and **44**, or the substituted anthraquinones rhein **45** and mitoxantrone **51**, were not biotransformed (Fig. 3 and *SI Appendix*, Table S9). 1,3- and 2,3-dihydroxynaphthalene (**55** and **56**) were readily biotransformed to the corresponding methylglucopyranosides (**55b** and **56b**) that were also isolated and characterized by NMR (*SI Appendix*, Fig. S3 and Table S7). Similarly, the substituted trihydroxynaphthalene shunt product **57** of the *Monascus* azaphilone pigments and its spontaneously oxidized benzoquinone derivative **58** (47) were also excellent substrates (Fig. 3 and *SI Appendix*, Table S9). Tetracyclic drugs, such as tetracycline **68** or doxorubicin **69**, were not accepted by the GT–MT pair. We also considered simple phenolic compounds (63–67 in *SI Appendix*, Fig. S1), however only 2,4-dihydroxybenzaldehyde **64** was biotransformed.

Based on our earlier report on *B. bassiana* being able to carry out *N*-(4-*O*-methyl)glucosylation (20), we also investigated the bioconversion of three aminoanthraquinones (**52**–**54** in *SI Appendix*, Fig. S1). All three of these turned out to be excellent substrates yielding the corresponding *N*-(4-*O*-methyl)glucosides, the structures of two of which (**52b** and **53b**) were fully elucidated (Fig. 3 and *SI Appendix*, Fig. S3 and Tables S7–S9). 1- or 2-aminonaphthalenes **59** and **60** were also converted to their *N*-(4-*O*-methyl)glucoside derivatives. However, 1- or 2-naphthalenethiol **61** and **62** (*SI Appendix*, Fig. S1) were not accepted as substrates, consistent with the scarcity of enzymes catalyzing small molecule *S*-glycosylation (8, 48).

Taking these data together, we find that the BbGT86–BbMT85 detoxification system is an extremely broad-spectrum biocatalyst that is able to efficiently decorate a large number of structurally varied substrates with a methylglucose biosynthon (Fig. 3). BbMT85 is able to recognize and efficiently modify most glucosides that it is presented with, indicating that the substrate range of the GT–MT pair is in essence determined by the BbGT86. While this GT is highly promiscuous in terms of its substrates, it displays relatively strict regiospecificity. It is also able to conduct both *O*- and *N*-glucosylation, a property that is very rare indeed for GTs, and has not been reported before for characterized fungal GTs.

Glycosylation Modulates the Solubility, Stability, and Bioactivity of BDLs. Compared with their corresponding aglycones, the lipophilicities of the glucosides and methylglucosides were found to be significantly decreased, as evidenced by the large increases of their polarities during reversed-phase chromatography, and by the sizeable decreases of their predicted ClogP values (fragment-based calculation of the logarithm of the partition coefficient between *n*-octanol and water) (49) (*SI Appendix*, Table S10). Although "drug-likeness" for oral drug candidates is a multidimensional property, reduced ClogP values (in the range of 2.5–3.0) have been correlated with higher success rates in market introduction due to more favorable drug potency, bioavailability, pharmacokinetics, and toxicity profiles (50). In particular, flavonoid glycosides are often considered more valuable due to their better solubility and stability, although their bioavailability and absorption from the small intestine is variable, and may be influenced by their differential deglycosylation by the intestinal microflora (17, 51).

In addition to modulating the ClogP of the glycosides, methylation of the glucose moiety also increased the physical and biological stability of selected BDL methylglucosides. This was evident by a much-reduced in-source fragmentation of DLD-methylglucoside **1b** into the aglycone **1** during electrospray ionization-MS, compared with that seen with the DLD-glucoside **1a** (*SI Appendix*, Fig. S5*A*). Intriguingly, little to no conversion of methylglucosides **1b** and **4b** were seen to their respective aglycones when these compounds were incubated with *S. cerevisiae* and *Escherichia coli* cultures that express glucoside hydrolases (28), or when the same compounds were supplemented to MCF-7 (human breast cancer), A549 (human lung cancer), or Vero cells (African green monkey kidney epithelium). In contrast, up to 63.1% and 11.3% of the glucosides **1a** and **4a**, respectively, were converted to their corresponding aglycones by the yeast; 7.0% and 65.1%, respectively, were deglucosylated by the enteric bacterium; and 47.2–57.4% and 32.3–47.3% of the glucosides **1a** and **4a**, respectively, were deglucosylated by the three cell lines (*SI Appendix*, Fig. S5*B* and Table S11).

BDLs show potent cytotoxicity against various cancer cells (26, 52). In particular, DLD (**1**) and monocillin II (**4**) have both been described to suppress the growth of breast cancer cells (53, 54), while **1** was also toxic to lung cancer cells (55). Anticancer activities are also often attributed to various flavonoids and their glycosides (51). To assess the influence of glycosylation on these bioactivities, the cytotoxicities of three sets of BDL and two sets of flavonoid aglycone/glucoside/methylglucoside congeners were compared using untransformed Vero cells and human cancer cell lines MCF-7, A549, HepG2 (human liver hepatocellular carcinoma), and HeLa (human cervix adenocarcinoma) as the targets (Fig. 4*A*). The flavanone hesperetin (**33**), the DAL-type BDL lasilarin (**18**), and their glycosides showed no cytotoxicity against these cell lines. The flavonol kaempferol (**31**) displayed moderate cytotoxicity against HepG2 cells, but its glucoside **31a** was inactive against all cell lines tested. Interestingly, kaempferol methylglucoside **33b** retained weak cytotoxicity against HepG2 cells, and gained weak antiproliferative activity against HeLa cells. While the RAL-type BDL aglycones **1** and **4** showed moderate to strong cytotoxicity against all five cell lines, DLD methylglucoside **1b** and glucosides **1a** and **4a** lost toxicity completely

(IC_{50} > 50 μM) (Fig. 4*A*), similar to the case of flavonoid glycosides that display generally reduced antitumor activities compared with their aglycone counterparts (51). However, monocillin II methylglucoside **4b** still maintained moderate toxicity against MCF-7 cells, and weak toxicity against A549 cells. Importantly, **4b** displayed no toxicity against untransformed Vero cells, revealing an interesting selectivity (IC_{50}^{Vero}/IC_{50}^{MCF-7} > 5.5) toward breast cancer cells (Fig. 4*A*).

Cell–matrix interactions are crucial for many cellular processes and play significant roles in angiogenesis, arteriosclerosis, inflammatory diseases, and cancer metastasis. We have compared the effects of BDLs **1** and **4**, BDL glucosides **1a** and **4a**, and BDL methylglucosides **1b** and **4b** on the attachment of MCF-7 and A549 cells to matrix, using noninvasive electrical cell-substrate impedance sensing (ECIS) technology (56). Monocillin II and its glycosides showed no inhibition of cell attachment. In contrast, DLD methylglucoside **1b** displayed potent inhibition of cell adhesion with both cell lines (Fig. 4 *B* and *C*) in the absence of apparent cytotoxicity (Fig. 4*A*).

Conclusions

Diversity-oriented combinatorial biosynthesis ventures to generate unprecedented chemical matter by appropriating extraneous enzymatic reactions into existing biochemical pathways, or by fusing disparate biosynthons into hybrid molecules. Such unnatural biosynthetic products could then be evaluated for biological activity of pharmaceutical, veterinary, or agricultural interest, or used as value-added reagents in the chemical industries. Using bioactive fungal benzenediol lactones and azaphilones as our model systems, we have previously demonstrated that diversity-oriented polyketide scaffold engineering is feasible with PKS subunit shuffling, and creates novel hybrid polyketide skeletons from defined biosynthons (22, 23). The current work shows that the chemical diversity of such molecules can be further extended toward an orthogonal chemical dimension by decorating these scaffolds with a modified sugar biosynthon, yielding BDL glucosides and methylglucosides not found in nature. The easily scalable total biosynthesis of such molecules is achieved by: (*i*) expressing appropriate PKS subunit pairs to afford natural or unnatural BDL congeners in a yeast chassis; and (*ii*) coexpressing a novel fungal type II/III detoxification module to decorate the polyketide scaffolds with a 4-*O*-methylglucose biosynthon. Although enzymes for the biosynthetic glycosylation of various bioactive aglycones have been characterized from bacterial or plant sources (1, 7), the first fungal

A

	A549	HepG2	HeLa	MCF-7	Vero
1	15.1 ± 1.19	8.7 ± 0.39	19.1 ± 0.32	14.7 ± 0.67	19.1 ± 2.48
1a	>50	18.9 ± 0.91	>50	>50	>50
1b	>50	>50	>50	>50	>50
4	2.8 ± 0.33	5.1 ± 0.80	3.6 ± 0.47	1.9 ± 0.29	3.3 ± 0.015
4a	>50	>50	>50	>50	>50
4b	27.8 ± 3.97	>50	>50	9.1 ± 1.09	>50
18	>50	>50	>50	>50	>50
18a	>50	>50	>50	>50	>50
18b	>50	>50	>50	>50	>50
31	>50	19.9 ± 0.43	>50	>50	>50
31a	>50	>50	>50	>50	>50
31b	>50	28.1 ± 0.71	24.8 ±0.64	>50	26.3 ± 2.37
33	>50	>50	>50	>50	>50
33a	>50	>50	>50	>50	>50
33b	>50	>50	>50	>50	>50
ADM	0.7 ± 0.07	1.2 ± 0.02	0.3 ± 0.02	1.0 ± 0.21	-

B

Cell line	Treatment	Mean ± SD	95% CI	Grouping
A549	DMSO	1.63 ± 0.07	(1.53, 1.73)	A
	1	1.54 ± 0.08	(1.43, 1.65)	AB
	1a	1.51 ± 0.06	(1.42, 1.59)	AB
	1b	1.46 ± 0.04	(1.40, 1.51)	B
MCF-7	DMSO	1.60 ± 0.11	(1.42, 1.78)	A
	1	1.43 ± 0.12	(1.25, 1.61)	AB
	1a	1.42 ± 0.09	(1.24, 1.60)	AB
	1b	1.28 ± 0.02	(1.10, 1.46)	B

C

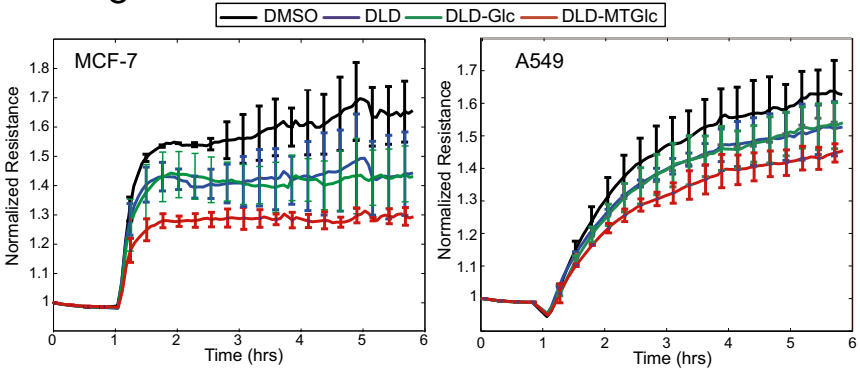

Fig. 4. Evaluation of the cytotoxic and the cell–matrix attachment-inhibitory activities of selected aglycones, glucosides and methylglucosides. (*A*) Cytotoxicities of DLD (**1**), monocillin II (**4**), lasilarin (**18**), kaempferol (**31**), hesperetin (**33**), and their glycosylated derivatives against untransformed African green monkey kidney epithelium (Vero) cells and human cancer cell lines MCF-7, A549, HepG2, and HeLa, with ADM as the positive control. (*B* and *C*) End-point and time-course assays, respectively, for the inhibition of cell–matrix attachment using ECIS technology (56). 95% CI, 95% confidence interval analyzed by one-way ANOVA; Grouping, statistical significance ($P < 0.05$) evaluated by Fisher pairwise comparisons, indicated by different letters. IC_{50} and normalized resistance values represent the means ± SD from three independent experiments.

enzymes with significant aglycone promiscuity were only identified in 2017 (6, 9). Considering the biosynthetic and biodegradative capacities of filamentous fungi, and their proven proficiency in glycosylating various scaffolds (11–16, 20, 41), extending the pool of biocatalysts with novel fungal enzymes, such as the *B. bassiana* GT–MT methylglucosylation module for the glycodiversification (3, 57) of various aglycones, is highly promising.

Our work demonstrates that the novel *B. bassiana* methylglucosylation module is proficient to generate various glycosides of BDL congeners, flavonoids, stilbenes, anthraquinones, and naphthalenes, and can even be applied to produce biosynthetically rare *N*-glycosides. Using a panel of 69 potential substrates, we detected 117 glycosylated compounds derived from 45 scaffolds. We isolated and elucidated the structures of 32 representative glycosides, 26 of which turned out to be novel compounds. Generation of these glycosides is straightforward in a biotransformation format by feeding a variety of preformed aglycones to an engineered biocatalyst. The presence of orthologous GT–MT modules in nonsyntenic genomic regions of Hypocreales, disjointed from any cognate biosynthetic gene cluster, indicates that these filamentous fungi already utilize these enzymes in an analogous manner. Thus, these enzymes conjugate various external substrates with methylglucose, although probably for self-defense and not in a synthetic capacity. In addition to the biotransformation of preformed aglycones, in situ biosynthesis of diverse aglycones is also feasible as shown by us for the BDL congeners and for the naphthalenes **57** and **58**, and by others for flavonoids (15, 58), chalcones and stilbenes (59), or anthraquinones (20). Just as with the BDL congeners, coupling the biosynthesis of such scaffolds with the expression of the GT–MT module in the same chassis would then yield de novo molecules by "total biosynthesis" (60–62).

"Sugarcoating" drug-like small-molecule scaffolds often provides glycosides that display better solubility and bioavailability, modulating their PK/PD characteristics as drugs or, as is often the case, as prodrugs (51). Macrolides such as erythromycin or tylosin gain antibiotic activities only with glycosylation, while the antineoplastic Adriamycin (ADM, doxorubicin) is inactivated upon deglycosylation in human tissues (4). In this work, we have shown that BDLs achieve more drug-like ClogP values after glycosylation, and that model BDL methylglucosides display improved physical and biological stability compared with the corresponding glucosides. We have also demonstrated that monocillin II methylglucoside **4b** retains significant toxicity against MCF-7 breast cancer cells, and even attains selectivity due to its attenuated toxicity against untransformed control cells. We also show that DLD-methylglucoside **1b**, but not the corresponding glucoside **1a** or aglycone **1**, inhibits the attachment of MCF-7 and A549 (lung cancer) cells to matrix while displaying no direct toxicity to these cells. Such nontoxic cell adhesion inhibitors may find utility in preventing tumor metastasis, and in modulating angiogenesis, arteriosclerosis, and inflammatory diseases. Meanwhile, unnatural BDL glycosides can be investigated for additional, novel, or improved biological activities, and the *B. bassiana* methylglucosylation enzyme module can be utilized for the efficient generation of *O*- and *N*-glycosides of a large variety of drug-like aglycone scaffolds.

Materials and Methods

Strains and Culture Conditions. *E. coli* DH10B and plasmid pJET1.2 (Thermo Fisher) were used for routine cloning and sequencing. *S. cerevisiae* BJ5464-NpgA (*MATα ura3-52 his3-Δ200 leu2-Δ1 trp1 pep4::HIS3 prb1 Δ1.6R can1 GAL*) was used as the host for expression vectors based on plasmids YEpADH2p-URA, YEpADH2p-TRP, and YEpADH2p-LEU (22–25, 27, 45). Cultivation of *B. bassiana* ARSEF 2860, primers used in this study, and details on the construction of expression vectors are described in the *SI Appendix, SI Materials and Methods*. Cultivation of recombinant *S. cerevisiae* BJ5464-NpgA strains for production of polyketides and for biotransformation was carried out as previously described (23–25, 27, 33, 45). Aglycone substrates (10 μg/mL, final concentration) in methanol were supplemented to the culture when it reached an OD_{600} of 0.6, and incubation was continued at 30 °C with shaking at 220 rpm for an additional 2 d. Polyketide production or aglycone biotransformation was analyzed in three to five independent *S. cerevisiae* transformants for each recombinant yeast strain, and fermentations with representative isolates were repeated at least three times.

Chemical Characterization of Glycosides. Extracts were prepared, analyzed by LC-MS, and products were isolated from scaled-up fermentations (1–10 L, depending on yield) as previously described (22–25, 27, 45). HPLC-HRESIMS and MS-MS spectra were acquired on an Agilent 1290 Infinity II HPLC coupled with an Agilent QTOF 6530 instrument. ^{1}H NMR, ^{13}C NMR, 1D-NOESY, and 2D NMR (^{1}H-^{1}H COSY, HSQC, and HMBC) spectra were obtained on a Bruker Avance III 400 spectrometer at 400 MHz for ^{1}H NMR and 100 MHz for ^{13}C NMR. See *SI Appendix* for details.

Protein Structure Modeling. The *B. bassiana* BbMT85 homology protein structure model was built with SWISS-MODEL (63). Comparative de novo protein structure models for GTs were built using the Robetta server (42). Substrates UDP-glucose and DLD for BbGT86, and SAM and DLD-5-*O*-β-D-glucopyranoside for BbMT85 were modeled in Chem3D, and docked with their respective enzymes using Autodock. Protein structures were compared using the DALI server (64), and the volumes of cavities were measured using the CASTp server (65). See *SI Appendix* for details.

Bioactivity Assays. Human lung adenocarcinoma A549, human breast adenocarcinoma MCF-7, human hepatocellular carcinoma HepG2, human cervical carcinoma HeLa, and African green monkey kidney epithelial Vero cell lines were purchased from the Kunming Cell Bank (People's Republic of China). The cytotoxicity of selected compounds and ADM as the positive control was determined with the in vitro tetrazolium-based assay (66) (see *SI Appendix, SI Materials and Methods* for details). The half-maximal inhibitory concentration (IC_{50}) of each test compound was calculated by Probit analysis using SPSS 20 (SPSS). All data represent the means ± SDs of three independent experiments with four replicates each.

Cell adhesion and spreading was determined with an ECIS Zθ system using 8W10E arrays (Applied Biophysics) (67). The array surfaces and electrodes were pretreated with a solution of cysteine (10 mM) and subsequently incubated for 1 h at 37 °C in DMEM Nutrient Mixture F-12 (Gibco) supplemented with 10 mM Hepes. After the addition of A549 cells (1,000,000 per well) or MCF-7 cells (800,000 per well), and the supplementation of the test compound in DMSO (5 μM, final concentration), electrical resistance was continuously monitored at a frequency of 4,000 Hz for 5.5 h at 37 °C. Treatments were in duplicates and experiments were repeated three times. Electrical resistances, normalized to the measured resistance at $t = 0$ min for each treatment, represent the means ± SDs of the three independent experiments. Endpoint resistances after 5.5 h of incubation were analyzed by one-way ANOVA followed by Fisher's pairwise comparison.

ACKNOWLEDGMENTS. This work was supported by National Basic Research Program of China Grant 2015CB755700 (to Y.X. and M.L.); National Natural Science Foundation of China Grants 31570093 (to Y.X.) and 31500079 (to L.Z.); The China Scholarship Council (X. Wei and C.W.); National Program of China for Transgenic Research Grant 2016ZX08009003-002 (to M.L.); National Key Research and Development Program of China Grant 2017YFD0201301-06 (to L.Z.); Joint Genomics Institute of the US Department of Energy WIP ID 1349 (to I.M.); and National Institutes of Health-National Institute of General Medical Sciences Grant R01GM114418-01A1 (to I.M.).

1. Tiwari P, Sangwan RS, Sangwan NS (2016) Plant secondary metabolism linked glycosyltransferases: An update on expanding knowledge and scopes. *Biotechnol Adv* 34:714–739.
2. Losey HC, et al. (2001) Tandem action of glycosyltransferases in the maturation of vancomycin and teicoplanin aglycones: Novel glycopeptides. *Biochemistry* 40:4745–4755.
3. Chen X (2011) Fermenting next generation glycosylated therapeutics. *ACS Chem Biol* 6:14–17.
4. Fujita H, et al. (1986) Pharmacokinetics of doxorubicin, (2″R)-4′-O-tetrahydropyranyl-adriamycin and aclarubicin. *Jpn J Antibiot* 39:1321–1336.
5. Veitch NC, Grayer RJ (2011) Flavonoids and their glycosides, including anthocyanins. *Nat Prod Rep* 28:1626–1695.
6. Feng J, et al. (2017) Regio- and stereospecific *O*-glycosylation of phenolic compounds catalyzed by a fungal glycosyltransferase from *Mucor hiemalis*. *Adv Synth Catal* 359:995–1006.
7. Hofer B (2016) Recent developments in the enzymatic *O*-glycosylation of flavonoids. *Appl Microbiol Biotechnol* 100:4269–4281.
8. Thibodeaux CJ, Melançon CE, 3rd, Liu HW (2008) Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew Chem Int Ed Engl* 47:9814–9859.

9. Xie K, et al. (2017) Two novel fungal phenolic UDP glycosyltransferases from *Absidia coerulea* and *Rhizopus japonicus*. *Appl Environ Microbiol* 83:e03103-16.
10. Warnecke D, et al. (1999) Cloning and functional expression of UGT genes encoding sterol glucosyltransferases from *Saccharomyces cerevisiae*, *Candida albicans*, *Pichia pastoris*, and *Dictyostelium discoideum*. *J Biol Chem* 274:13048–13059.
11. de M B Costa EM, Pimenta FC, Luz WC, de Oliveira V (2008) Selection of filamentous fungi of the *Beauveria* genus able to metabolize quercetin like mammalian cells. *Braz J Microbiol* 39:405–408.
12. Zi J, Gladstone SG, Zhan J (2012) Specific 5-hydroxylation of piperlongumine by *Beauveria bassiana* ATCC 7159. *Biosci Biotechnol Biochem* 76:1565–1567.
13. Zymańczyk-Duda E, Brzezińska-Rodak M, Klimek-Ochab M, Lejczak B (2011) Application of the *Beauveria bassiana* strain for the enantioselective oxidation of the diethyl 1-hydroxy-1-phenylmethanephosphonate. *Curr Microbiol* 62:1168–1172.
14. Corrêia Gomes D, Takahashi JA (2016) Sequential fungal fermentation-biotransformation process to produce a red pigment from sclerotiorin. *Food Chem* 210:355–361.
15. Zhan JX, Gunatilaka AAL (2006) Selective 4′-*O*-methylglycosylation of the pentahydroxyflavonoid quercetin by *Beauveria bassiana* ATCC 7159. *Biocatal Biotransform* 24:396–399.
16. Zeng J, Yang N, Li XM, Shami PJ, Zhan J (2010) 4′-*O*-methylglycosylation of curcumin by *Beauveria bassiana*. *Nat Prod Commun* 5:77–80.
17. Sordon S, Popłoński J, Tronina T, Huszcza E (2017) Microbial glycosylation of daidzein, genistein and biochanin A: Two new glucosides of biochanin A. *Molecules* 22:e81.
18. Buchanan GO, Reese PB (2001) Biotransformation of diterpenes and diterpene derivatives by *Beauveria bassiana* ATCC 7159. *Phytochemistry* 56:141–151.
19. Abdel Halim OB, Maatooq GT, Marzouk AM (2007) Metabolism of parthenin by *Beauveria bassiana* ATCC 7159. *Pharmazie* 62:226–230.
20. Zhan J, Gunatilaka AAL (2006) Microbial transformation of amino- and hydroxyanthraquinones by *Beauveria bassiana* ATCC 7159. *J Nat Prod* 69:1525–1527.
21. Zhan J, Gunatilaka AAL (2005) Microbial transformation of curvularin. *J Nat Prod* 68: 1271–1273.
22. Bai J, et al. (2016) Diversity-oriented combinatorial biosynthesis of hybrid polyketide scaffolds from azaphilone and benzenediol lactone biosynthons. *Org Lett* 18: 1262–1265.
23. Xu Y, et al. (2014) Diversity-oriented combinatorial biosynthesis of benzenediol lactone scaffolds by subunit shuffling of fungal polyketide synthases. *Proc Natl Acad Sci USA* 111:12354–12359.
24. Xu Y, et al. (2014) Insights into the biosynthesis of 12-membered resorcylic acid lactones from heterologous production in *Saccharomyces cerevisiae*. *ACS Chem Biol* 9: 1119–1127.
25. Xu Y, et al. (2013) Rational reprogramming of fungal polyketide first-ring cyclization. *Proc Natl Acad Sci USA* 110:5398–5403.
26. Shen W, Mao H, Huang Q, Dong J (2015) Benzenediol lactones: A class of fungal metabolites with diverse structural features and biological activities. *Eur J Med Chem* 97:747–777.
27. Xu Y, et al. (2013) Thioesterase domains of fungal nonreducing polyketide synthases act as decision gates during combinatorial biosynthesis. *J Am Chem Soc* 135: 10783–10791.
28. Wang H, et al. (2016) Engineering *Saccharomyces cerevisiae* with the deletion of endogenous glucosidases for the production of flavonoid glucosides. *Microb Cell Fact* 15:134.
29. Liang D-M, et al. (2015) Glycosyltransferases: Mechanisms and applications in natural product development. *Chem Soc Rev* 44:8350–8374.
30. Koirala N, Thuan NH, Ghimire GP, Thang DV, Sohng JK (2016) Methylation of flavonoids: Chemical structures, bioactivities, progress and perspectives for biotechnological production. *Enzyme Microb Technol* 86:103–116.
31. Zeng J, Valiente J, Zhan J (2011) Generation of two new macrolactones through sequential biotransformation of dihydroresorcylide. *Nat Prod Commun* 6:223–226.
32. Xiao G, et al. (2012) Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. *Sci Rep* 2:483.
33. Davis C, et al. (2011) Single-pot derivatisation strategy for enhanced gliotoxin detection by HPLC and MALDI-ToF mass spectrometry. *Anal Bioanal Chem* 401:2519–2529.
34. Cooper B, Campbell KB, Beard HS, Garrett WM, Islam N (2016) Putative rust fungal effector proteins in infected bean and soybean leaves. *Phytopathology* 106:491–499.
35. Hu Q-N, et al. (2012) Assignment of EC numbers to enzymatic reactions with reaction difference fingerprints. *PLoS One* 7:e52901.
36. Hu Q-N, Deng Z, Hu H, Cao D-S, Liang Y-Z (2011) RxnFinder: Biochemical reaction search engines using molecular structures, molecular fragments and reaction similarity. *Bioinformatics* 27:2465–2467.
37. Sánchez C, et al. (2005) Combinatorial biosynthesis of antitumor indolocarbazole compounds. *Proc Natl Acad Sci USA* 102:461–466.
38. Isaka M, Yangchum A, Auncharoen P, Srichomthong K, Srikitikulchai P (2011) Ring B aromatic norpimarane glucoside from a *Xylaria* sp. *J Nat Prod* 74:300–302.
39. Mackenzie PI, et al. (1997) The UDP glycosyltransferase gene superfamily: Recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics* 7:255–269.
40. Choi SH, Ryu M, Yoon YJ, Kim DM, Lee EY (2012) Glycosylation of various flavonoids by recombinant oleandomycin glycosyltransferase from *Streptomyces antibioticus* in batch and repeated batch modes. *Biotechnol Lett* 34:499–505.
41. Heider SAE, et al. (2014) Production and glucosylation of C50 and C 40 carotenoids by metabolically engineered *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol* 98:1223–1235.
42. Song Y, et al. (2013) High-resolution comparative modeling with RosettaCM. *Structure* 21:1735–1742.
43. Binkowski TA, Naghibzadeh S, Liang J (2003) CASTp: Computed atlas of surface topography of proteins. *Nucleic Acids Res* 31:3352–3355.
44. Wu K, Chung L, Revill WP, Katz L, Reeves CD (2000) The FK520 gene cluster of *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. *Gene* 251:81–90.
45. Xu Y, et al. (2013) Characterization of the biosynthetic genes for 10,11-dehydrocurvularin, a heat shock response-modulating anticancer fungal polyketide from *Aspergillus terreus*. *Appl Environ Microbiol* 79:2038–2047.
46. Wang S, et al. (2008) Functional characterization of the biosynthesis of radicicol, an Hsp90 inhibitor resorcylic acid lactone from *Chaetomium chiversii*. *Chem Biol* 15: 1328–1338.
47. Chen W, et al. (2017) Orange, red, yellow: Biosynthesis of azaphilone pigments in *Monascus* fungi. *Chem Sci (Camb)* 8:4917–4925.
48. Wang H, et al. (2014) The glycosyltransferase involved in thurandacin biosynthesis catalyzes both *O*- and *S*-glycosylation. *J Am Chem Soc* 136:84–87.
49. Leo AJ, Hoekman D (2000) Calculating logP(oct) with no missing fragments; the problem of estimating new interaction parameters. *Perspect Drug Discov Des* 18:19–38.
50. Leeson PD, Springthorpe B (2007) The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat Rev Drug Discov* 6:881–890.
51. Xiao J (2017) Dietary flavonoid aglycones and their glycosides: Which show better biological significance? *Crit Rev Food Sci Nutr* 57:1874–1905.
52. Napolitano C, Murphy PV (2013) Resorcylic acid lactones. *Natural Lactones and Lactams: Synthesis, Occurrence and Biological Activity*, ed Tomasz J (Wiley, Somerset, NJ), pp 273–319.
53. Wei H, et al. (2012) Monocillin II inhibits human breast cancer growth partially by inhibiting MAPK pathways and CDK2 Thr160 phosphorylation. *ChemBioChem* 13:465–475.
54. Hazalin NA, Lim SM, Cole AL, Majeed AB, Ramasamy K (2013) Apoptosis induced by desmethyl-lasiodiplodin is associated with upregulation of apoptotic genes and downregulation of monocyte chemotactic protein-3. *Anticancer Drugs* 24:852–861.
55. Buayairaksa M, et al. (2011) Cytotoxic lasiodiplodin derivatives from the fungus *Syncephalastrum racemosum*. *Arch Pharm Res* 34:2037–2041.
56. Hong J, Kandasamy K, Marimuthu M, Choi CS, Kim S (2011) Electrical cell-substrate impedance sensing as a non-invasive tool for cancer cell study. *Analyst (Lond)* 136:237–245.
57. Chang A, Singh S, Phillips GN, Jr, Thorson JS (2011) Glycosyltransferase structural biology and its role in the design of catalysts for glycosylation. *Curr Opin Biotechnol* 22: 800–808.
58. Herath W, Mikell JR, Hale AL, Ferreira D, Khan IA (2008) Microbial metabolism part 9. Structure and antioxidant significance of the metabolites of 5,7-dihydroxyflavone (chrysin), and 5- and 6-hydroxyflavones. *Chem Pharm Bull (Tokyo)* 56:418–422.
59. Herath W, Khan SI, Khan IA (2013) Microbial metabolism. Part 14. Isolation and bioactivity evaluation of microbial metabolites of resveratrol. *Nat Prod Res* 27: 1437–1444.
60. Fujii R, et al. (2011) Total biosynthesis of diterpene aphidicolin, a specific inhibitor of DNA polymerase α: Heterologous expression of four biosynthetic genes in *Aspergillus oryzae*. *Biosci Biotechnol Biochem* 75:1813–1817.
61. Tagami K, et al. (2014) Rapid reconstitution of biosynthetic machinery for fungal metabolites in *Aspergillus oryzae*: Total biosynthesis of aflatrem. *ChemBioChem* 15: 2076–2080.
62. Minami A, Liu CW, Oikawa H (2016) Total biosynthesis of fungal indole diterpenes using cell factories. *Heterocycles* 92:397–421.
63. Biasini M, et al. (2014) SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* 42:W252–W258.
64. Wheby MS, Jones LG, Crosby WH (1964) Studies on iron absorption. Intestinal regulatory mechanisms. *J Clin Invest* 43:1433–1442.
65. Dundas J, et al. (2006) CASTp: Computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res* 34:W116–W118.
66. Rubinstein LV, et al. (1990) Comparison of in vitro anticancer-drug-screening data generated with a tetrazolium assay versus a protein assay against a diverse panel of human tumor cell lines. *J Natl Cancer Inst* 82:1113–1118.
67. Patani N, Douglas-Jones A, Mansel R, Jiang W, Mokbel K (2010) Tumour suppressor function of MDA-7/IL-24 in human breast cancer. *Cancer Cell Int* 10:29.
68. Ma SM, et al. (2009) Complete reconstitution of a highly reducing iterative polyketide synthase. *Science* 326:589–592.
69. Holm L, Sander C (1995) Dali: A network tool for protein structure comparison. *Trends Biochem Sci* 20:478–480.