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

Linan Xie¹,², Liwen Zhang³,¹, Chen Wang³,⁴, Chen, Xiaoqin Wang³,¹,², Ya-ming Xu³, Hefen Yu⁴, Ping Wu⁴, Shenglan Li⁵, Lida Han⁶, A. A. Leslie Gunatilaka⁷, Xiaoyi Wei⁴, Min Lin⁸,³, István Molnár⁹,², and Yuquan Xu³,²

¹Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, 100081 Beijing, People’s Republic of China; ²Natural Products Center, University of Arizona, Tucson, AZ 85706; ³Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, 510650 Guangzhou, People’s Republic of China; ⁴Guangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Chinese Academy of Sciences, 510650 Guangzhou, People’s Republic of China; and ⁵Department 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-glycosides with remarkable regio- and stereospecificity, a spectrum not demonstrated for other characterized fungal enzymes. These glycosides are faithfully processed by the dedicated MT to afford 4-O-methylglycosides. 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

D

i

covery 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-glycosidic linkage not previously demonstrated with characterized fungal enzymes.


Conflict of interest statement: I.M. has a disclosed financial interest in TEVA Pharmaceuticals Hungry, 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.

Published under the PNAS license.

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

2To whom correspondence may be addressed. Email: linmin57@vip.163.com, imolnar@ email.arizona.edu, or xyuyuquan@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.

www.pnas.org/cgi/doi/10.1073/pnas.1716046115

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 Morchella 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 glucosylates 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). Based on the carbons involved in this fusion, BDLs can be subdivided into resorcylic acid lactones (RALs, C2 to C8) and desmethyl-lasiodiplodin (DLD) (15). The entomopathogenic ascomycetous fungus B. bassiana is a useful biocatalyst due to its ability to catalyze various biotransformations, including methylation, hydroxylation, and oxidation (11–19). It also readily glucosylates 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). Based on the carbons involved in this fusion, BDLs can be subdivided into resorcylic acid lactones (RALs, C2 to C8) and desmethyl-lasiodiplodin (DLD) (15).

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 Biosynthetic 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 dihydroxycoryloside (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 methylglycosyl 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 glycosylate 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) glycosyltransferases from bacteria, including those that glycosylate 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, RfnFinder 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 (c < 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 methylglycosylation (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, RfnFinder 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)glycoside moiety (37). A Blastp search with RebM as the bait returned 19 predicted N-MTs (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). DLL (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 DLL 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 DLL 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 DLL. However, the strain expressing BBA_08686 (henceforth, BbGT86) produced 1b as the major product with increased polarity, and only trace amounts of DLL detectable (Fig. S4). The mass-to-charge ratio of the [M–H]− ion of 1a was 162 amu higher than that of DLL, 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 DLL 1 or DLL glycoside 1a (Fig. L4). 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 DLL by comparing the 1H and 13C 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, δH 5.07; 1b, δH 5.04) and C-5 of the aglycone (1a, δC 162.7; 1b, δC 162.6), together with the presence of a low-field signal for OH-3 (1a, δH 11.59; 1b, δH 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 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 confirmed by the characteristic 13C NMR resonances (1a, signals at δC 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, [α]D25 +99.6 (c 0.20, H2O), with published data (38). Finally, HMBC correlation between the methoxyl protons (δH 3.56) and C-4 (δC 80.0) of the hexose 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 DLL (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 p-xylosyl 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. 2A). 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 (40) (Fig. 2A), and UGT59A1 from Asb altaica and UGT59A1 from Rhizopus japonicus, each showing only marginal sequence similarities to BbGT86 (identities of 13% with UGT59A1, 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
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/THMM/), 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. 2B). 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. 2D).

Docking of UDP-glucose and DLD to the predicted BbGT86 structure (Fig. 2E and SI Appendix, SI Results) places these substrates at the deduced conserved active site dyad (29), and suggests that H131 acts as a catalytic base that abstracts the proton from OH-5 of DLD, while the transition state may be stabilized by E134. 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 $\text{S}_\text{N}2$-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 characterized 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 dihydroxy cyclohexane side chain of the macrolactam FKS20 (44) (AAF86398, 34% identity to BbMT85). However, BbMT85 shows no similarity to RebM (Q9KZ34), 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 LDL 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-methylglucosyl biosynthones, 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, ABES011 from Methylobacillus flagellatus (PDB ID code 2PY6). This provided a model with a partial overlap with the template residue covering only the SAM-binding domain (Fig. 2E). This part of BbMT85 (but not the C-terminal part of the protein) also allowed a reasonable structural similarity to the SAM-binding regions of various other MTs, including the precorrin-6y C-5, 15-methyltransferase of Geobacter metallidigens (PDB ID code 3E05) (Fig. 2E). Docking SAM into the BbMT85 model provided a pose with a good overlap with the SAM residing in the experimentally determined 2PY6 structure (Fig. 2C 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 system in Hypocreales to generate any glycosides in 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 methylglucopyranosyl 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. S1A). This included RAL14 (DLD 1, trans-resorcylic 2, and radiplodon 3); RAL14 [monocillin II 4, radicol 6, lasicicol 9, R-zearelanol 10, and trans-14 (15-dehydrozearelenol (DHZ) 12); RAL16 (sediocelic 13); DAL12 [curvularin 14, 10 (11)-dehydrocurvularin 15 and 11-hydroxycurvularin 16]; and DAL14 (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. S1A). Most of these compounds were produced in situ by coexpressing the relevant PKS pair 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 nonmacrocyclic congeners, the presence of an α-pyrene 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 1H and 13C NMR data with those published for their aglycones (22–45), 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 with no precedent among natural fungal BDLs.

Methylglucosyl derivatives invariably dominated the product profiles for all BDL congeners, indicating that BbMT85 efficiently methylates the nonglycosylated 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 regiosomers 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 NMR correlation (Hβ13/α13 or Hβ13/α12) of the anomeric proton (δH 4.96, d, J = 7.4 Hz) and C-7 of the aglycone (δC 157.6). Glycosylation of OH-5 in 17b was affirmed by the HMBC correlation of the anomeric proton (δH 4.90, d, J = 7.4 Hz) with C-5 of the aglycone (δC 160.5) (SI Appendix, Fig. S3 and Table S7), and by detecting a chelated OH (OH-7) at δH 9.83 in its 1H NMR spectrum in DMSO-d6 (SI Appendix, Fig. S15B and Table S7A). 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 pKb 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., pKa 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-zearelanol 10 with that of semisynthetic S-zearelanol 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...
enone double bond is inhibitory toward methylglucosylation in 12-membered macrocycles, but not in 14-membered BDLs. Thus, the enone-DAL$_{12}$ (11)-dehydrocurvularin was not modified, while the expected (21)-O-methylglucopyranoside of curvularin was detected. Similarly, the enone-RAL$_{12}$ trans-resorcylide 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 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. S1B 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

![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.](image)
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 S7, a hydroxynaphthalene shunt product of the Monascus azaphilone pigments) (47). Most aglycones yielded a single prominent methylglucoside, accompanied by minor amounts of glyco-conjugate 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 maringenin 34, the isoflavone genistein 38, and 4-isopropyl-3,5-dihydroxystilbene 42 also yielded relatively large amounts of glucosides 39 and 45. All biotransformation, 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 (secophlavone isoliquiritigenin 40 were not accepted, and preglycosylated flavonoids (baicalin 28, puertarin 39), or the flavonolignan silybin 36 were similarly not transformed. Interestingly, the branched substituent in 4-isopropyl-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 bio-synthesized using compounds 46–50, but 1- or 2-hydroxyanthraquinone 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 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 aminoaanthraquinones (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 which were elucidated (Fig. 3 and SI Appendix, Fig. S3 and Tables S7–S9). 1- or 2-aminoanthraquinones 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 regioselectivity. It is also able to conduct both O- and N-glycosylation, 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 favorable drug properties, bioavailability, pharmacokinetics, and toxicity profiles (50). In particular, flavonoid glucosides 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 de-glycosylation by the intestinal microflora (17, 51).

In addition to modulating the ClogP of the glucosides, methylglucosylation of the glucose moiety also increased the physical and biological stability of selected BDLS methylglucosides. This was evident by a much-reduced in-source fragmentation of DLD-methylglucoside 1b into the aglycone 1 in electron spray ionization-MS, compared with that seen with the DLD-glucoside 1a (SI Appendix, Fig. SS4). 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 glucosyl 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 de-glycosylated by the enteric bacterium; and 47.2–57.4% and 32.3–47.3% of the glucosides 1a and 4a, respectively, were de-glycosylated by the three cell lines (SI Appendix, Fig. SSB 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 BDLS 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). The flavonone hesperetin (33), the DAL-type BDL lasilarin (18), and their glucosides 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.
(IC₅₀ > 50 μM) (Fig. 4A), 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 (IC₅₀ HepG2/Vero = 5.1, MCF-7 > 5.5) toward breast cancer cells (Fig. 4A).

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. 4B and C) in the absence of apparent cytotoxicity (Fig. 4A).

**Conclusions**

Diversity-oriented combinatorial biosynthesis ventures to generate unprecedented chemical matter by appropriating extraneous enzymatic reactions into existing biochemical pathways, or by fusing disparate biosynths 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 aza-philones 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 biosynths (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 biosynth, 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 biosynth. Although enzymes for the biosynthetic glycosylation of various bioactive aglycones have been characterized from bacterial or plant sources (1, 7), the first fungal

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Treatment</th>
<th>Mean ± SD</th>
<th>95% CI</th>
<th>Grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td>A549</td>
<td>1</td>
<td>15.1 ± 1.19</td>
<td>19.1 ± 1.32</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>1a</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1b</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2.8 ± 0.33</td>
<td>3.6 ± 0.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4a</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>18a</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>18b</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>31a</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>31b</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>33</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>33a</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>33b</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td></td>
</tr>
</tbody>
</table>

ADM 0.7 ± 0.07 1.2 ± 0.02 0.3 ± 0.02 1.0 ± 0.21 -

**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. IC₅₀ 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 bio-degradative 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 methylglycosylation module for the glycodiversification (3, 57) of various aglycones, is highly promising.

Our work demonstrates that the novel B. bassiana methylglycosylation module is proficient to generate various glycosides of BDL congeners, flavonoids, stilbenes, arthranoquiones, 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, dissected 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 methylglycoside, 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 arthranoquiones (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 ClqP values after glycosylation, and that model BDL methylglucosides display improved physical and biological stability compared with the corresponding aglycones. We have also demonstrated that monocillin II methylglycoside 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-methylglycoside 1b, but not the corresponding glycoside 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 methylglycosylation 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-α (MATα ura3-52 his3-a200 leu2-d1 trp1 pep4::HIS3 prb1 α1.6R 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-α 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 OD600 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. 1H NMR, 13C NMR, 1D-NOESY, and 2D NMR (1H–1H COSY, HSQC, and HMBC) spectra were obtained on a Bruker Avance III 600 spectrometer at 400 MHz for 1H NMR and 100 MHz for 13C 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). Substrate specificity profiling was performed with 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 cancer 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 (IC50) 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 2.0 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 DM50 (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 31500709 (to L.Z.); The China Scholarship Council (X. Wei and C.W.); National Program of China Grant 2015CB755700 (to Y.X. and M.L.); National Natural Science Foundation of China Grants 2013CB736700, 81273096, 81273093, and 81473342 (to M.L.); National Institute of General Medical Sciences Grant R01GM114418-01A1 (to I.M.); and National Institutes of Health-National Institute of General Medical Sciences Grant R01GM114418-01A1 (to I.M.).
16. Wang H, et al. (2016) Engineering Saccharomyces cerevisiae with the deletion of endog-