Deciphering the sugar biosynthetic pathway and tailoring steps of nucleoside antibiotic A201A unveils a GDP-L-galactose mutase

Qinghua Zhu[a,1], Qi Chen[a,1], Yongxiang Song[a], Hongbo Huang[a], Jun Li[a], Junying Ma[a], Qinglian Li[a], and Jianhua Ju[a,b,2]

[a]Chinese Academy of Sciences (CAS) Key Laboratory of Tropical Marine Bioresources and Ecology, Guangdong Key Laboratory of Marine Materia Medica, Research Network for Applied Microbiology Center for Marine Microbiology, South China Sea Institute of Oceanology, CAS, Guangzhou 510301, China; and [b]College of Earth Sciences, University of Chinese Academy of Sciences, Beijing 10049, China

Edited by Arnold L. Demain, Drew University, Madison, NJ, and approved March 31, 2017 (received for review December 9, 2016)

Galactose, a monosaccharide capable of assuming two possible configurational isomers (D-/L-), can exist as a six-membered ring, galactopyranose (Gal*p*), or as a five-membered ring, galactofuranose (Gal*f*). UDP-galactopyranose mutase (UGM) mediates the conversion of pyranose to furanose thereby providing a precursor for D-Gal*f*. Moreover, UGM is critical to the virulence of numerous eukaryotic and prokaryotic human pathogens and thus represents an excellent antimicrobial drug target. However, the biosynthetic mechanism and relevant enzymes that drive L-Gal*f* production have not yet been characterized. Herein we report that efforts to decipher the sugar biosynthetic pathway and tailoring steps en route to nucleoside antibiotic A201A led to the discovery of a GDP-L-galactose mutase, MtdL. Systematic inactivation of 18 of the 33 biosynthetic genes in the A201A cluster and elucidation of 10 congeners, coupled with feeding and in vitro biochemical experiments, enabled us to: (*i*) decipher the unique enzyme, GDP-L-galactose mutase associated with production of two unique D-mannose-derived sugars, and (*ii*) assign two glycosyltransferases, four methyltransferases, and one desaturase that regiospecifically tailor the A201A scaffold and display relaxed substrate specificities. Taken together, these data provide important insight into the origin of L-Gal*f*-containing natural product biosynthetic pathways with likely ramifications in other organisms and possible antimicrobial drug targeting strategies.

nucleoside antibiotic A201A | biosynthesis | GDP-L-galactose mutase | methyltransferase | desaturase

Galactose (Gal), which can assume two possible stereochemical configurations (D-/L-), is able to exist in the thermodynamically favored six-membered ring form as galactopyranose (Gal*p*), or in the less favored five-membered ring form as galactofuranose (Gal*f*) (1, 2). D-gal in mammals is found only in the pyranose form as D-Gal*p*, whereas in bacteria, fungi, and protozoa it exists as D-Gal*f*; this is the form found also in many pathogenic organisms (3, 4). Indeed, in many pathogenic prokaryotes and eukaryotes, UDP-Gal*f*, the precursor of D-Gal*f*, is usually enzymatically generated from UDP-Gal*p* by UDP-D-galactopyranose mutase (UGM) (3, 5, 6). Since the first identification and cloning of UGM from *Escherichia coli* (7), dozens of UGMs have been discovered and functionally characterized from pathogenic bacteria, fungi, parasites, and nematodes. These findings have highlighted the critical role of UGMs in Gal*f* biosynthesis, especially as relates to virulence, cellular adhesion, parasite viability, and proper developmental processes (8–11). It has been proposed that during the formation of UDP-Gal*f*, a flavin adenine dinucleotide (FAD)-iminium ion intermediate is formed through nucleophilic addition that facilitates galactose ring opening and contraction (5, 12–16). Three dimensional structures for members of the UGM family have been determined by X-ray crystallography in free form and in complex with various substrates, as well as in association with both oxidized and reduced forms of FAD (12, 13, 17–19). Notably, a covalent FAD-Gal*p* intermediate involved in the UGM reaction has recently yielded to elucidation by X-ray crystallography shedding significant insight into UGM catalysis (20). Consequently, UGM is a potential antimicrobial drug target (3, 16).

Compared with the high abundance of D-Gal, L-Gal is rare in both terrestrial and marine environments (21, 22). It is worth noting that L-Gal is a key intermediate in the conversion of D-glucose to ascorbic acid in plants (23). In contrast, L-Gal*f*, the furanose form of L-galactose, is exceptionally rare in the environment. It has been detected only in animal cells (24), plant tissues (24), early weaned piglets (25), botanical species of the Cactaceae family (26), propolis of stingless bees (27), and as a building block for the antimicrobial epipyrones from *Epicoccum purpurascens* (28), as well as in the oligosaccharide fraction isolated from the mycelium of Lingzhi or Reishi medicinal mushrooms *Ganoderma lucidum* (29). In addition, L-Gal*f* has served as a critical intermediate in assorted carbohydrate chemical syntheses (30, 31). Despite these findings, the biosynthetic mechanisms and relevant enzymes that drive the synthesis of L-Gal*f* have yet to be characterized.

The aminonucleoside antibiotic A201A (1) was first isolated from *Streptomyces capreolus* NRRL 3817 (32) and then rediscovered in our laboratory from deep sea-derived *Marinactinospora thermotolerans* SCSIO 00652 (33) (Fig. 1). A201A is composed of a 3′-amino-3′-deoxy adenosine moiety, a *p*-hydroxy-α-methylcinnamic acid moiety, a hexofuranose moiety, an α-D-rhamnose moiety, and is further decorated with unusual sugar unsaturation and *N*- or *O*-methylations (Fig. 1*A*). The compound displays strong antibacterial activities against Gram-positive and anaerobic Gram-negative bacteria by

Significance

L-galactofuranose (L-Gal*f*), the furanose form of L-galactose, is rare in nature yet plays a central role in secondary metabolite structures with medicinal potential. Importantly, the biosynthetic enzymes and mechanisms driving L-Gal*f* biosynthesis have eluded characterization. We report here a previously unknown mutase, MtdL, that generates GDP-L-Gal*f* from its galactopyranose analog GDP-L-Gal*p* using unique chemistry. MtdL mediates the pyranose–furanose conversion solely by the presence of bivalent cation. The ubiquitous nature of this Gal*p* → Gal*f* chemistry is suggested by numerous mutase homologs that are phylogenetically widespread in various phyla of Bacteria, Eukaryota, and Archaea. Thus, a genetic marker for L-Gal*f*-containing natural products and their producers has been discovered. Additionally, mutase homologs may constitute targets for antimicrobial drugs.

Author contributions: J.J. designed research; Q.Z., Q.C., Y.S., H.H., and J.L. performed research; J.M., Q.L., and J.J. analyzed data; and Q.Z., Q.C., and J.J. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]Q.Z. and Q.C. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: jju@scsio.ac.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1620191114/-/DCSupplemental.

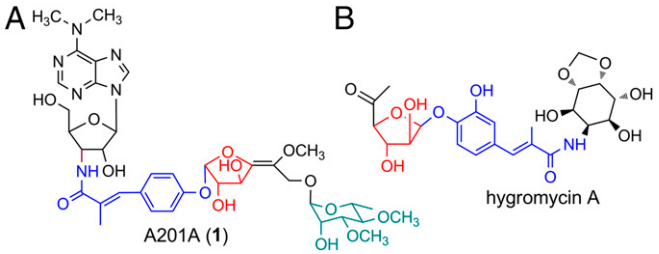

Fig. 1. Structures of A201A (**1**) and hygromycin A.

preventing bacterial ribosomal peptide bond formation or by impairing proper tRNA–ribosome associations during protein synthesis and is only weakly toxic to mammals (32–34). Consequently, A201A possesses great potential for antiinfective drug discovery as in, for instance, the treatment of acne (32, 35, 36). The first total synthesis of A201A was achieved in 47 steps by Yu and coworkers (37). Early investigations into A201A biosynthesis by our group led to the identification of the A201A gene cluster in *M. thermotolerans* SCSIO 00652, demonstrating that MtdA is a negative transcriptional regulator governing A201A biosynthesis (33); efforts elsewhere identified the A201A cluster in *S. capreolus* (38, 39). Notably, *S. capreolus* failed to be genetically amenable to subsequent and more elaborate dissection. Consequently, the exact biosynthetic pathway and machineries driving A201A biosynthesis have eluded characterization over the past 10 y. To further elucidate the biosynthetic apparatus driving A201A production, especially with an eye on L-Gal*f* incorporation, we report herein: (*i*) boundary determinations and heterologous expression of the A201A gene cluster in *Streptomyces lividans* TK64; (*ii*) systematic inactivation of genes coding for the biosynthesis and transfer of the two sugar units, as well as tailoring steps in *M. thermotolerans* enabling us to generate 10 A201A congeners from assorted mutant strains; (*iii*) complete elucidation of metabolite structures; and (*iv*) in vitro biochemical and feeding experiments enabling functional gene assignments related to biosynthesis and transfer of the unique hexofuranose and α-D-rhamnose moieties harboring a GDP-L-galactose mutase MtdL,

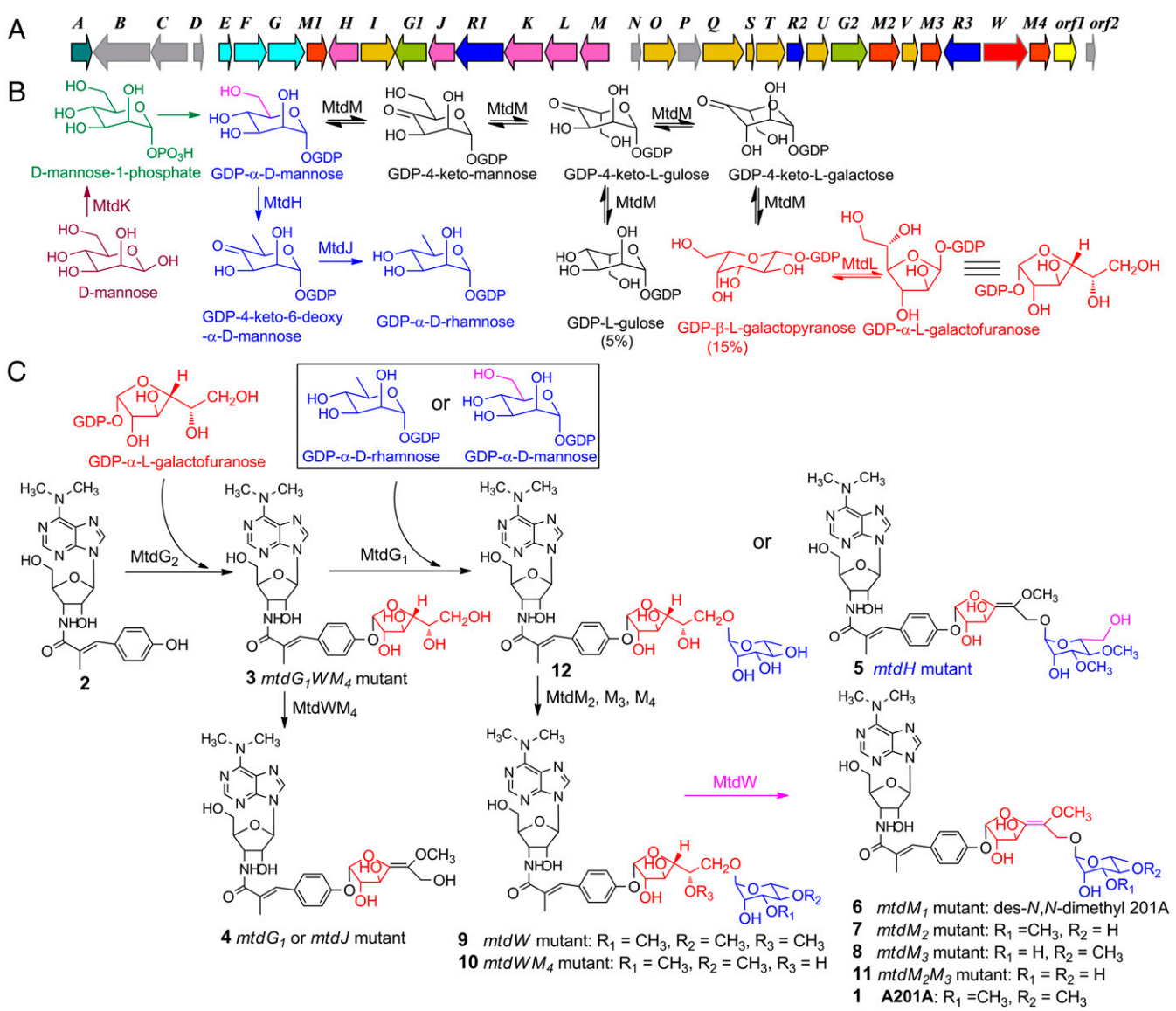

Fig. 2. The biosynthetic gene cluster and proposed biosynthetic pathway of A201A (**1**) in *M. thermotolerans* SCSIO 00652. (*A*) Organization of the A201A biosynthetic genes. (*B*) Biosynthetic pathway of sugar moieties in A201A. (*C*) Tailoring steps involved in A201A assembly.

as well as, four methyltransferases and a regiospecific tailoring desaturase showing relaxed substrate specificity.

Results and Discussion

Determination of the Boundaries of the A201A Gene Cluster in *M. thermotolerans* SCSIO 00652 and Heterologous Expression in *S. lividans* TK64. To narrow down the genes responsible for A201A biosynthesis, we first fully analyzed and annotated the A201A gene cluster (Fig. 2*A* and *SI Appendix*, Table S1). We also inactivated *mtdB* (encoding a peptidase S15), *mtdC* (encoding an unknown protein), *mtdD* (encoding an unknown protein), the *orf1* (encoding an inositol-phosphate phosphatase), and *orf2* genes using our established λ-RED-mediated PCR-targeting mutagenesis strategy in *M. thermotolerans* SCSIO 00652 by replacing them with an apramycin resistance gene cassette (33, 40). Genotypes of the mutants were confirmed by PCR (*SI Appendix*). HPLC analyses of the fermentation extracts of the resultant mutants revealed that all mutants still produced **1** in yields comparable to those of the wild-type strain (*SI Appendix*, Fig. S22), thereby suggesting the dispensability of these genes for A201A biosynthesis.

We subsequently selected cosmid 142H harboring the genes from *mtdC* to *orf3*, integrated elements for conjugation from the pSET152AB vector (41), and transferred them into *S. lividans* TK64 through conjugation. The recombinant *S. lividans* TK64/142H was fermented using the same medium as had been the case for the wild-type producer. HPLC analysis of the fermentation extract revealed that *S. lividans* TK64/142H indeed produced **1** (Fig. 3, trace *ii*), whereas the control did not (Fig. 3, trace *iii*), demonstrating that the narrowed gene cluster spanning *mtdE* to $mtdM_4$ harbors all of the genes necessary for A201A biosynthesis. Production of **1** in *S. lividans* TK64/142H was ∼50% of that seen in the wild-type producer. Within the gene cluster, the gene products $MtdEFGM_1$, homologous to those in the puromycin pathway (42), are proposed to be involved in constructing the *N*, *N*-dimethyl-3′-amino-3′-deoxyadenosine moiety (*SI Appendix*, Fig. S23 and Table S1). The gene products MtdINOSTUV, homologous to those in the hygromycin A (Fig. 1*B*) pathway (43, 44), are proposed to be involved in the biosynthesis of the *p*-hydroxy-α-methylcinnamic acid moiety. MtdQ, a ligase homologous to Hyg12 in the hygromycin A pathway, is proposed to converge the deoxyadenosine and cinnamic acid moieties through an amide bond to form **2** (*SI Appendix*, Fig. S23 and Table S1). A detailed comparison of biosynthetic proteins involved in A201A, hygromycin A, and puromycin is provided in *SI Appendix*, Table S1.

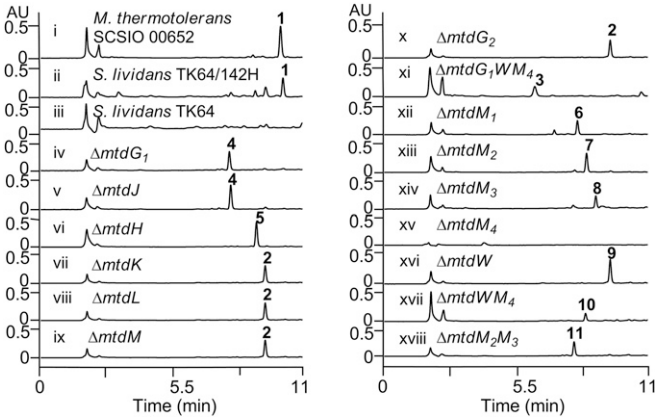

Fig. 3. HPLC metabolite analyses for *S. lividans*, *S. lividans* TK64/142H, *M. thermotolerans* wild type, and mutant strains. See Fig. 2 for structures **1**–**11**. λ = 275 nm.

Biosynthesis and Transfer of the α-D-Rhamnose Moiety Reveals Promiscuous Glycosyltransferase $MtdG_1$ Capable of Processing both GDP-α-D-Mannose and GDP-α-D-Rhamnose. We carried out bioinformatics analysis to identify the genes (*SI Appendix*, Table S1) possibly responsible for the construction and transfer of the α-D-rhamnose moiety in A201A. MtdH and MtdJ belong to members of the extended short-chain dehydrogenase-reductases (SDR) superfamily and share as low as 28% and 15% identity, respectively, with Gmd and Rmd from *Aneurinibacillus thermoaerophilus* L420-91T. Gmd and Rmd proteins are known to be involved in the two-step conversion of GDP-α-D-mannose to GDP-α-D-rhamnose (45). Between MtdH and MtdJ, resides a glycosyltransferase, $MtdG_1$, as a candidate enzyme installing a sugar building block onto the A201A scaffold.

To explore the exact roles of the above-mentioned genes in **1** biosynthesis, we individually inactivated the three genes in *M. thermotolerans* wild type as described before. The metabolite profiles for the $\Delta mtdG_1$ and $\Delta mtdJ$ mutants both revealed that these strains do not produce **1**; instead these strains both led to accumulation of product **4** (Fig. 3, traces *iv* and *v*). The $\Delta mtdH$ mutant also failed to produce **1**, but yielded a product **5** (Fig. 3, trace *vi*). The mutant strains were then fermented on a large scale (8–16 L each) and the structures of **4** and **5** were elucidated (Fig. 2*C*) on the basis of comprehensive high-resolution mass spectrometry (HRMS), 1D and 2D NMR analyses (*SI Appendix*). The $\Delta mtdG_1$ and $\Delta mtdJ$ mutants both yielded **4** devoid of the terminal sugar, demonstrating their involvement in the transfer and synthesis of the terminal α-D-rhamnose moiety. The $\Delta mtdH$ mutant failed to produce **1** but unexpectedly accumulated product **5** in titers comparable to those of **1** in the wild-type producer. That **5** bears a D-mannose unit in place of the D-rhamnose unit found in **1** reveals that MtdH catalyzes 4, 6-dehydration of GDP-α-D-mannose to yield GDP-4-keto-6-deoxy-α-D-mannose (Fig. 2*B*), and that the glycosyltransferase $MtdG_1$ shows significant substrate promiscuity because it can apparently process both GDP-α-D-mannose and GDP-α-D-rhamnose (Fig. 2*C*).

Biosynthesis and Transfer of the Hexofuranose Moiety as Dictated by a GDP-L-Galactose Mutase. Three cotranscribed genes, *mtdK*, *mtdL*, *mtdM* (*mtdKLM*), were identified as candidates possibly involved in the unique hexofuranose moiety sugar unit biosynthesis. MtdK shows as low as 27–32% identity to a group of glucokinases in the GenBank database. MtdL shows the highest homology (61% identity) to Hyg20 in the hygromycin A pathway (*SI Appendix*, Fig. S24 and Table S1) and has been proposed to catalyze the conversion of NDP-L-fucose to NDP-L-fucofuranose (44). MtdM belongs to the extended SDR superfamily with $NAD(P)^+$-binding sites, showing 45–57% identity to a group of GDP-mannose-3′,5′-epimerases from *Arabidopsis thaliana*, and rice (*Oryza sativa*) (*SI Appendix*, Fig. S25) (46, 47). Within the biosynthetic gene cluster of **1**, and downstream of the *mtdKLM*, $mtdG_2$ encoding another glycosyltransferase, is likely responsible for the hexofuranose transfer.

The above four genes (*mtdK*, *mtdL*, *mtdM*, and $mtdG_2$) were similarly individually inactivated in the wild-type strain. LCMS analysis of the fermentation extracts revealed that $\Delta mtdK$, $\Delta mtdL$, $\Delta mtdM$, and $\Delta mtdG_2$ mutants uniformly failed to produce **1** and yielded the same intermediate **2** (Fig. 3, traces *vii*–*x*) devoid of the two sugar moieties as determined by large-scale fermentation, purification, and HRMS and NMR structure elucidation (*SI Appendix*). These data indicate that MtdKLM are all involved in construction of the unique hexofuranose unit and that $MtdG_2$ is responsible for transferring this unit to aglycon **2**.

However, in the structure of **1**, the hexofuranose moiety contains a double bond. Accordingly, the exact origin of this hexofuranose presents a mystery. Based on the bioinformatics data, we first asked whether or not MtdM has the in vitro GDP-mannose-3′,5′-epimerase biochemical activity. We overexpressed MtdM in *E. coli* BL21 (DE3), and purified it to homogeneity as an N-terminus–His_6-tagged

protein. The enzymatic activity of MtdM was then evaluated in the presence of GDP-α-D-mannose (0.2 mM) or GDP-β-L-galactose (0.2 mM), 0.5 mM NAD^+, and 50 mM Tris·HCl buffer (pH = 8.0). An equilibrating mixture of three products, GDP-α-D-mannose, GDP-β-L-galactose, and GDP-β-L-gulose, in a ratio of 80:15:5, was found in the enzymatic reaction (Fig. 4*A*). On the basis of these data it appears that the catalytic function of MtdM is similar to that of the GDP-D-mannose 3′,5′-epimerase (GME) from *A. thaliana* (46) (*SI Appendix*, Fig. S25). This result suggested that MtdL might subsequently catalyze the conversion of GDP-L-galactopyranose into GDP-L-galactofuranose (Fig. 2*B*).

To validate that GDP-L-galactofuranose is incorporated as a precursor into the biosynthetic pathway for **1**, we designed and successfully constructed a triple gene $\Delta mtdG_1WM_4$ mutant strain (in which the *mtdW* and $mtdM_4$ were in-frame deleted, and the $mtdG_1$ was replaced with an apramycin resistance gene cassette) (*SI Appendix*). Fermentations of the $\Delta mtdG_1WM_4$ strain, did indeed, afford important intermediate **3** (Fig. 3, trace *xi*). The structure of **3** (Fig. 2*C*), containing a furanose unit, was fully established upon isolation from a scaled-up fermentation (24 L) and following analyses of HRMS, 1D, and 2D NMR data (*SI Appendix*). Analyses of ^{1}H NMR spectra and comparisons of circular dichroism (CD) data for the methanolyzed and benzoylated hydrolysate of **3** against authentic samples of methanolyzed and benzoylated L-galactose and D-galactose confirmed the presence of the L-galactose moiety in **3** (*SI Appendix* and *SI Appendix*, Figs. S26–S28).

The above results indicated that the two sugar units in **1** might be derived from a uniform precursor, GDP-D-mannose. To validate this hypothesis, we fed 1-^{13}C-labeled D-mannose into fermentations of the wild-type *M. thermotolerans* strain. The corresponding A201A product was purified and subjected to ^{13}C NMR analysis (*SI Appendix*, Figs. S84 and S85). As anticipated, the two anomeric carbons of the two sugar units were clearly enriched, demonstrating that the two sugars use GDP-α-D-mannose as the biosynthetic precursor.

MtdL also shows low homology (30% identity) to a rice UDP-arabinopyranose mutase (UAM) responsible for interconverting UDP-L-arabinofuranose and UDP-L-arabinopyranose in a 10:90 ratio (48, 49). Sequence alignments of MtdL with UAM, Hyg20 in the hygromycin A pathway, which may catalyze a conversion of NDP-L-fucose to NDP-L-fucofuranose, and other homologs in the GenBank with unknown functions, revealed that they all contain the conserved DxD motif usually found in the glycosyltransferase (GT)-A family (50) (*SI Appendix*, Fig. S24); such a motif interacts with the phosphate groups of nucleotide donor through coordination of divalent cations, typically Mn^{2+}, and is responsible for the glycosyl transfer reaction.

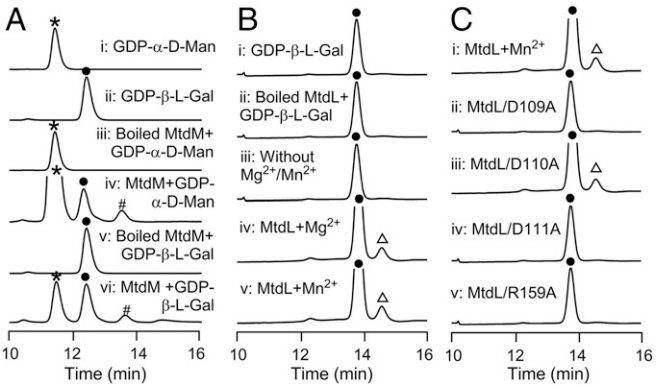

Fig. 4. In vitro characterization of MtdM and MtdL-catalyzed reactions. (*A*) HPLC analysis of in vitro MtdM assays: (*i*) authentic GDP-α-D-mannose (GDP-α-D-Man); (*ii*) authentic GDP-β-L-galactose (GDP-β-L-Gal); (*iii*) control assay with boiled MtdM and GDP-α-D-Man; (*iv*) enzymatic reaction using MtdM and GDP-α-D-Man; (*v*) control assay with boiled MtdM and GDP-β-L-Gal; and (*vi*) enzymatic reaction using MtdM and GDP-β-L-Gal. (*B*) HPLC analysis of in vitro MtdL assays: (*i*) authentic GDP-β-L-Gal; (*ii*) control assay with boiled MtdL and GDP-β-L-Gal; (*iii*) control assay without Mg^{2+}/Mn^{2+}; and (*iv* and *v*) enzymatic reaction with Mg^{2+} or Mn^{2+}, respectively. (*C*) HPLC analysis of the in vitro enzymatic reaction of (*i*) MtdL and its mutant variants; (*ii*) D109A; (*iii*) D110A; (*iv*) D111A; and (*v*) R159A, respectively. *: GDP-α-D-Man; ●: GDP-β-L-Gal; #: GDP-L-gulose; Δ: GDP-α-L-galactofuranose. λ = 254 nm.

To determine whether MtdL could catalyze the transformation of GDP-β-L-galactopyranose to GDP-α-L-galactofuranose in vitro, we overexpressed MtdL in *E. coli* BL21(DE3), and purified it to homogeneity as an N-terminus–His_6-tagged protein. The enzymatic activity of MtdL was tested in a 50-μL reaction containing 0.2 mM GDP-β-L-galactose, in the absence or presence of 5 mM Mg^{2+} or Mn^{2+}, 2 μM MtdL, in 50 mM phosphate buffer (pH = 8.0), at 37 °C for 20 min. HPLC analysis of the enzymatic mixture revealed that MtdL catalyzes the formation of an enzymatic product, which was found to have the identical molecular weight as the substrate by HR-LCMS analysis and the reaction absolutely requires Mg^{2+} or Mn^{2+} (Fig. 4*B*). The enzymatic reaction was performed in situ in an NMR tube and ultimately subjected to ^{1}H NMR analysis following enzyme removal by centrifugal filtration. The ^{1}H NMR spectrum of the enzymatic mixture clearly shows the appearance of the anomeric proton signal characteristic of GDP-α-L-galactofuranose (*SI Appendix*, Figs. S86 and S87). The ratio of the pyranose and furanose form of GDP-L-galactose is 91:9, as judged by the HPLC peak area and integration of the anomeric proton signal in the ^{1}H NMR spectrum (*SI Appendix*, Figs. S86 and S87). It is noted that the coupling constant (J = 6.0 Hz) of the anomeric proton in the GDP-α-L-galactofuranose indicated an α-configuration (Fig. 2*B*) (51, 52), demonstrating that the stereochemistry at C-1 of the GDP-β-L-galactose was reversed during the pyranose–furanose transformation (50).

To investigate which amino acid residues of MtdL are responsible for its enzymatic activity, site-directed mutagenesis experiments were performed based on sequence alignment (*SI Appendix*, Fig. S24). The D109, D110, D111, and R159 residues of MtdL were each converted to alanine and the activities of purified mutant proteins were assessed in enzymatic assays. HPLC analyses of the reaction mixtures revealed that the D109A MtdL mutant (Fig. 4*C*, trace *ii*), together with the D111A (Fig. 4*C*, trace *iv*) and R159A mutants (Fig. 4*C*, trace *v*), was completely devoid of enzymatic activity, whereas the D110A mutation appeared to exert no apparent influence upon MtdL enzymatic activity (Fig. 4*C*, trace *iii*). In the glycosyltransferase A-families, the DxD motif plays a key role in the coordination of Mg^{2+} or Mn^{2+} with the phosphate groups of the nucleotide donor (50); in the UAM protein family, the arginine residue serves as a site of protein autoglycosylation (48, 49). Consequently, D109, D111, and R159 in MtdL are likely to serve as a catalytic triad for the GDP-β-L-galactopyranose and GDP-α-L-galactofuranose transformation.

A plausible chemical mechanism for MtdL-catalyzed pyranose–furanose conversion is envisioned as follows. The anomeric carbon-*O*-GDP bond is broken, and bicyclic intermediate 1, 4-anhydrogalactopyranose, is formed, coincident with ring contraction (*SI Appendix*, Fig. S90*A*). As indicated in *SI Appendix*, Fig. S90*B*, we envisage that Asp109 of MtdL serves as a catalytic base to deprotonate the hydroxyl group of galactose and that Asp111 interacts with the Mg^{2+} or Mn^{2+} ion as Arg159 enters into an electrostatic interaction with the phosphoryl group of GDP. A double displacement mechanism involving discrete nucleophilic catalysis and the formation of a covalently bound glycosyl-enzyme intermediate seems possible for MtdL. Central to this mechanism, in addition to the arrangement of acidic and basic amino acids within the MtdL active site, is that L-galactose is enzymatically amenable to ring contraction as noted in *SI Appendix*, Fig. S90 *A* and *B*. Thus, MtdL uses a flavin-independent mechanism to catalyze the GDP-L-galactose pyranose–furanose transformation. This finding contrasts sharply with the UGMs that use the flavin cofactor as a nucleophile to

BIOCHEMISTRY

carry out the UDP-D-galactose pyranose–furanose exchange (*SI Appendix*, Fig. S90*C*) (11, 20).

Thus, the biosynthetic pathway for sugar moieties in A201A is initiated from the phosphorylated D-mannose catalyzed by MtdK, then converted to GDP-α-D-mannose by a GDP-sugar pyrophosphorylase, and then further processed by MtdM, a GDP-α-D-mannose 3′,5′-epimerase to yield GDP-β-L-galactose (Fig. 2*B*). The GDP-β-L-galactose was then transformed by MtdL, a pyranose–furanose mutase, to yield GDP-α-L-galactofuranose. GDP-α-D-mannose could also be converted to GDP-α-D-rhamnose through the agency of GDP-4-keto-6-deoxy-D-mannose by MtdH and MtdJ successively. The two sugar units were transferred to aglycon **2** by $MtdG_2$ and $MtdG_1$ to yield des-*O*-methyl desaturated A201A **10** (Fig. 2 *B* and *C*).

Phylogenetic analyses of MtdL with related proteins in the GenBank database revealed that they belong to a family of proteins with unknown functions represented across three domains: the Bacteria, Archaea, and Eukaryota (*SI Appendix*, Fig. S91). Homologous proteins from the Bacteria domain include proteins from actinomycetes such as *Streptomyces*, *Micromonospora*, *Saccharothrix*, and *Frankia* that have a demonstrated record of producing bioactive natural products. The Bacteria domain homologs also include proteins from *Cyanobacteria*, *Bacteroidetes*, and *Proteobacteria* whose secondary metabolites have only been marginally investigated. MtdL also appears to have a large number of homologous proteins from Archaea, including *Halococcus*, *Halonotius*, and *Salinarchaeum* from which secondary metabolites also have been largely unexplored. Notably, members of the Eukaryota including, but not limited to, *O. sativa*, *A. thaliana*, *Ancylostoma ceylanicum*, and *Necator americanus* also appear to produce MtdL homologs. Thus, it appears completely feasible that the *mtdL* gene represents an effective genetic marker for the mining of L-galactofuranose/L-fucofuranose/L-arabinofuranose–containing natural products in a wide variety of microorganisms.

Tailoring Steps of A201A Involving Four Regiospecific Methyltransferases and a Desaturase MtdW. We next analyzed the genes driving the tailoring steps involved in completing the biosynthesis of **1**. Structurally, A201A contains an *N*, *N*-dimethyl group, three *O*-methyl groups, and an unusual double bond within the hexofuranose. Installation of these moieties is achieved through a number of tailoring steps during A201A biosynthesis. Consistent with these structural considerations and aided by thorough bioinformatics analysis of the A201A gene cluster, we readily identified four methyltransferase genes, $mtdM_1$, $mtdM_2$, $mtdM_3$, and $mtdM_4$, that encode enzymes 234aa, 354aa, 242aa, and 240aa in size, respectively. These four genes were inactivated by replacement with an apramycin gene cassette as described before to yield the respective $\Delta mtdM_1$, $\Delta mtdM_2$, $\Delta mtdM_3$, and $\Delta mtdM_4$ mutants. These mutants were fermented and analyzed by HPLC-UV-MS (Fig. 3, traces *xii*–*xv*). We found that: (*i*) all mutants failed to produce **1**; (*ii*) the $\Delta mtdM_4$ mutant did not yield any "unique" products; (*iii*) $\Delta mtdM_1$ yielded a product **6** that is produced in a 10-fold lower yield than **1** in the wild-type producer; and (*iv*) each of the other mutants, $\Delta mtdM_2$ and $\Delta mtdM_3$, produced analogs **7** and **8**, respectively and, in each case, with titers comparable to those seen with **1** in the wild-type producer. Scaled-up fermentation of these respective mutants led to purification of compounds **6**–**8**, whose structures were unambiguously assigned as shown in Fig. 2*C* upon comprehensive analysis of HRMS, 1D, and 2D NMR spectroscopic data (*SI Appendix*). Judging from the structures of **6**–**8**, we conclude that: (*i*) $MtdM_1$ is responsible for *N*, *N*-dimethylation of the 3′-amino-3′-deoxyadenosine moiety in **1**, because the yield of **6** is significantly lower than that of **1** in the wild-type strain; (*ii*) $MtdM_2$ and $MtdM_3$ are responsible for the *O*-methylation of OH-C-4′ and OH-C-3′ in the α-D-rhamnose moiety of **1**, respectively; and (*iii*) $MtdM_4$ is responsible for the *O*-methylation of OH-C-5′ of the α-L-galactofuranose unit of **1**.

Notably, we found *mtdW* residing at one far end of the A201A gene cluster encoding a 529-aa protein with as low as 37–43% identity to the glucose-methanol-choline (GMC) oxidoreductase family of proteins in GenBank whose functions have not been experimentally verified (53). Sequence alignments (*SI Appendix*, Fig. S29) revealed that these enzymes contain a FAD ribose moiety binding motif $GxGxxG(x)_{18}E$ in the N terminus and an $H(x)_{37}P$ motif in the C terminus presumably involved in substrate oxidation (54). The *mtdW* gene was similarly inactivated and the resultant $\Delta mtdW$ mutant was found to produce compound **9**, as tentatively identified on the basis of HPLC-UV-MS (Fig. 3, trace *xvi*); **9** was generated in titers comparable to **1** produced in wild type. Scaled-up fermentation of the $\Delta mtdW$ mutant, subsequent purification, and structural elucidation on the basis of HRMS, 1D, and 2D NMR spectroscopic studies confirmed the structure of **9** as shown in Fig. 2*C*. Notably, the C-4/C-5 hexofuranose double bond of **1** is saturated in compound **9**. Thus, MtdW represents a desaturase responsible for C-4/C-5 dehydrogenation of the α-L-galactofuranose unit of **1**.

To further diversify the A201A structure, we made a $\Delta mtdWM_4$ dual mutant in which the two genes, *mtdW* and $mtdM_4$, were in-frame deleted in combination, and a $\Delta mtdM_2M_3$ dual mutant in which the $mtdM_2$ was in-frame deleted and the $mtdM_3$ gene was replaced with an apramycin gene cassette (*SI Appendix*). As anticipated, these mutants were fermented and found to produce two analogs **10** and **11,** respectively, upon HPLC-MS analyses (Fig. 3, traces *xvii* and *xviii*). Scaled-up fermentation of these two mutants led to the purification and structure elucidation of **10** and **11** as shown in Fig. 2*C* by analyses of their HRMS, 1D, and 2D NMR spectroscopic data, respectively (*SI Appendix*). The fact that these mutant strains accumulated multiple A201A analogs reveals that the four methyltransferases: $MtdM_1$, $MtdM_2$, $MtdM_3$, and $MtdM_4$, as well as the desaturase MtdW, are highly regiospecific, yet display relaxed substrate specificities (Fig. 2*C*).

Antibacterial Activities of A201A and Its Analogs. Finally, we tested the antibacterial activities of the engineered A201A analogs against a panel of Gram-positive and Gram-negative bacteria (*SI Appendix*, Table S11). We found that **6** displays antibacterial activity on par with that of **1**. A201A analog **6** has potential as an antibacterial lead, in part, by virtue of its improved water solubility relative to **1** {solubility units: **1** as −2.95 [log(mol/L)] vs. **6** as −2.90 [log(mol/L)] at pH 7.3, as calculated by the Water Solubility Module in Percepta software of Advanced Chemistry Development, Inc}. Notably, the removal of any of the sugar-borne methyl groups or their replacement with the hydroxymethyl moiety diminished antibacterial activities by approximately two-fold. Additionally, removal of the terminal sugar or saturation of the double bond within the hexofuranose significantly decreased the bioactivity.

Conclusions

In conclusion, systematic inactivation, individually or in combination, of 18 genes located proximal to the cluster boundaries, purported to be involved in sugar unit construction/linkage, or purported to drive the tailoring steps of A201A biosynthesis in *M. thermotolerans*, enabled us to produce 10 A201A analogs. Gene inactivation, feeding, and in vitro biochemical experiments unveiled that D-mannose is the uniform precursor en route to the two unique sugar units involving MtdL, an enzyme involved in GDP-α-L-galactofuranose generation. Des-*O*-methyl desaturated A201A is decorated by four methyltransferases: $MtdM_1$, $MtdM_2$, $MtdM_3$, and $MtdM_4$, and a desaturase, MtdW. Importantly, the tailoring enzymes all appear to possess exploitable substrate promiscuities. A combinatorial biosynthetic strategy has been credibly applied to the *mtd* pathway leading to a number of structurally diversified natural product analogs.

Materials and Methods

Materials and methods, including heterologous expression, gene inactivations, isolation and structure elucidation of metabolites **2**–**11**, in vitro biochemical assays, large-scale MtdL enzymatic reaction, characterization of the product, etc., are summarized in *SI Appendix, SI Materials and Methods*. Proposed functions of biosynthetic genes, bacteria, plasmids, primers, NMR spectral data, and antibacterial activities of the compounds are summarized in *SI Appendix*, Tables S1–S11. Construction and gel electrophoresis analyses of mutant strains, sequence alignments, NMR spectra of compounds **2**–**11**, plausible chemical mechanisms for MtdL and UGM, and other supporting figures are provided in *SI Appendix*, Figs. S1–S91.

ACKNOWLEDGMENTS. This work was supported in part by the National Science Foundation of China (81425022 and U1501223), the Chinese Academy of Sciences (XDA11030403), and the Guangdong Natural Science Foundation (2016A030312014).

1. Richards MR, Lowary TL (2009) Chemistry and biology of galactofuranose-containing polysaccharides. *ChemBioChem* 10:1920–1938.
2. Carlson EE, May JF, Kiessling LL (2006) Chemical probes of UDP-galactopyranose mutase. *Chem Biol* 13:825–837.
3. Oppenheimer M, Valenciano AL, Sobrado P (2011) Biosynthesis of galactofuranose in kinetoplastids: novel therapeutic targets for treating leishmaniasis and chagas' disease. *Enzyme Res* 2011:415976.
4. Tefsen B, Ram AF, van Die I, Routier FH (2012) Galactofuranose in eukaryotes: Aspects of biosynthesis and functional impact. *Glycobiology* 22:456–469.
5. Tanner JJ, Boechi L, Andrew McCammon J, Sobrado P (2014) Structure, mechanism, and dynamics of UDP-galactopyranose mutase. *Arch Biochem Biophys* 544:128–141.
6. Chlubnova I, et al. (2012) Specific and non-specific enzymes for furanosyl-containing conjugates: Biosynthesis, metabolism, and chemo-enzymatic synthesis. *Carbohydr Res* 356:44–61.
7. Nassau PM, et al. (1996) Galactofuranose biosynthesis in *Escherichia coli* K-12: Identification and cloning of UDP-galactopyranose mutase. *J Bacteriol* 178:1047–1052.
8. Beverley SM, et al. (2005) Eukaryotic UDP-galactopyranose mutase (GLF gene) in microbial and metazoal pathogens. *Eukaryot Cell* 4:1147–1154.
9. Poulin MB, et al. (2010) Characterization of a bifunctional pyranose-furanose mutase from *Campylobacter jejuni* 11168. *J Biol Chem* 285:493–501.
10. Oppenheimer M, Poulin MB, Lowary TL, Helm RF, Sobrado P (2010) Characterization of recombinant UDP-galactopyranose mutase from *Aspergillus fumigatus*. *Arch Biochem Biophys* 502:31–38.
11. Wesener DA, May JF, Huffman EM, Kiessling LL (2013) UDP-galactopyranose mutase in nematodes. *Biochemistry* 52:4391–4398.
12. Sanders DA, et al. (2001) UDP-galactopyranose mutase has a novel structure and mechanism. *Nat Struct Biol* 8:858–863.
13. Soltero-Higgin M, Carlson EE, Gruber TD, Kiessling LL (2004) A unique catalytic mechanism for UDP-galactopyranose mutase. *Nat Struct Mol Biol* 11:539–543.
14. Zhang Q, Liu H (2001) Mechanistic investigation of UDP-galactopyranose mutase from *Escherichia coli* using 2- and 3-fluorinated UDP-galactofuranose as probes. *J Am Chem Soc* 123:6756–6766.
15. Sun HG, Ruszczycky MW, Chang WC, Thibodeaux CJ, Liu HW (2012) Nucleophilic participation of reduced flavin coenzyme in mechanism of UDP-galactopyranose mutase. *J Biol Chem* 287:4602–4608.
16. Kizjakina K, Tanner JJ, Sobrado P (2013) Targeting UDP-galactopyranose mutases from eukaryotic human pathogens. *Curr Pharm Des* 19:2561–2573.
17. Gruber TD, Borrok MJ, Westler WM, Forest KT, Kiessling LL (2009) Ligand binding and substrate discrimination by UDP-galactopyranose mutase. *J Mol Biol* 391:327–340.
18. Gruber TD, Westler WM, Kiessling LL, Forest KT (2009) X-ray crystallography reveals a reduced substrate complex of UDP-galactopyranose mutase poised for covalent catalysis by flavin. *Biochemistry* 48:9171–9173.
19. van Straaten KE, et al. (2015) Structural basis of ligand binding to UDP-galactopyranose mutase from *Mycobacterium tuberculosis* using substrate and tetrafluorinated substrate analogues. *J Am Chem Soc* 137:1230–1244.
20. Mehra-Chaudhary R, Dai Y, Sobrado P, Tanner JJ (2016) *In crystallo* capture of a covalent intermediate in the UDP-galactopyranose mutase reaction. *Biochemistry* 55:833–836.
21. Reiter WD (2008) Biochemical genetics of nucleotide sugar interconversion reactions. *Curr Opin Plant Biol* 11:236–243.
22. Gutiérrez M, et al. (2004) L-Galactose as a natural product: Isolation from a marine octocoral of the first α-L-galactosyl saponin. *Tetrahedron Lett* 45:7833–7836.
23. Wheeler GL, Jones MA, Smirnoff N (1998) The biosynthetic pathway of vitamin C in higher plants. *Nature* 393:365–369.
24. Pabst M, et al. (2010) Nucleotide and nucleotide sugar analysis by liquid chromatography-electrospray ionization-mass spectrometry on surface-conditioned porous graphitic carbon. *Anal Chem* 82:9782–9788.
25. Xiao YP, et al. (2012) Response to weaning and dietary L-glutamine supplementation: Metabolomic analysis in piglets by gas chromatography/mass spectrometry. *J Zhejiang Univ Sci B* 13:567–578.
26. Fuentes VM (1996) Compositions containing carbohydrates obtained from plants of the family cactacea. *Eur Pat Appl* EP 0706797.
27. Matienzo AC, Lamorena M (2004) Extraction and initial characterization of propolis from stingless bees (*Trigona biroi* Friese). *Proceeding of the Asian Apicultural Association Conference and 10 BEENET Symposium and Technoflora* (University of the Philippines Los Baños, Laguna, The Philippines), pp 321–329.
28. Van Ginkel R, Selwood AI, Wilkins AL, Ford S (2014) Anti-microbial compositions, US Pat 20140357580.
29. Tello I, et al. (2013) Anticonvulsant and neuroprotective effects of oligosaccharides from Lingzhi or Reishi medicinal mushroom, *Ganoderma lucidum* (Higher Basidiomycetes). *Int J Med Mushrooms* 15:555–568.
30. Hanaya T, Yasuda K, Yamamoto H, Yamamoto H (1993) Stereoselectivity in the preparation of 5,6-dideoxy-5-dimethoxyphosphinyl-D-and-L-hexofuranoses, and an efficient synthesis of 5,6-dideoxy-5-hydroxyphosphinyl-L-galactopyranose (a P-in-the-ring L-Fucose analogue). *Bull Chem Soc Jpn* 66:2315–2322.
31. Binch H, Stangier K, Thiem J (1998) Chemical synthesis of GDP-L-galactose and analogues. *Carbohydr Res* 306:409–419.
32. Kirst HA, et al. (1985) The structure of A201A, a novel nucleoside antibiotic. *J Antibiot (Tokyo)* 38:575–586.
33. Zhu Q, et al. (2012) Discovery and engineered overproduction of antimicrobial nucleoside antibiotic A201A from the deep-sea marine actinomycete *Marinactinospora thermotolerans* SCSIO 00652. *Antimicrob Agents Chemother* 56:110–114.
34. Polikanov YS, et al. (2015) Distinct tRNA accommodation intermediates observed on the ribosome with the antibiotics hygromycin A and A201A. *Mol Cell* 58:832–844.
35. Hamill RL, Hoehn MM (1974) Antibiotics A201A and A201B and process for the production thereof, US Pat 3843784.
36. Ensminger PW, Hamill RL (1979) Method of treating acne with antibiotic A201A, US Pat 4143141.
37. Nie S, Li W, Yu B (2014) Total synthesis of nucleoside antibiotic A201A. *J Am Chem Soc* 136:4157–4160.
38. Saugar I, Sanz E, Rubio MA, Espinosa JC, Jiménez A (2002) Identification of a set of genes involved in the biosynthesis of the aminonucleoside moiety of antibiotic A201A from *Streptomyces capreolus*. *Eur J Biochem* 269:5527–5535.
39. Saugar I, et al. (2017) Characterization of the biosynthetic gene cluster (ata) for the A201A aminonucleoside antibiotic from Saccharothrix mutabilis subsp. capreolus. *J Antibiot (Tokyo)* 70:404–413.
40. Chen Q, et al. (2013) Discovery of McbB, an enzyme catalyzing the β-carboline skeleton construction in the marinacarboline biosynthetic pathway. *Angew Chem Int Ed Engl* 52:9980–9984.
41. Zhang Y, et al. (2013) Identification of the grincamycin gene cluster unveils divergent roles for GcnQ in different hosts, tailoring the L-rhodinose moiety. *Org Lett* 15:3254–3257.
42. Tercero JA, Espinosa JC, Lacalle RA, Jiménez A (1996) The biosynthetic pathway of the aminonucleoside antibiotic puromycin, as deduced from the molecular analysis of the *pur* cluster of *Streptomyces alboniger*. *J Biol Chem* 271:1579–1590.
43. Palaniappan N, Ayers S, Gupta S, Habib S, Reynolds KA (2006) Production of hygromycin A analogs in *Streptomyces hygroscopicus* NRRL 2388 through identification and manipulation of the biosynthetic gene cluster. *Chem Biol* 13:753–764.
44. Palaniappan N, et al. (2009) Biosynthesis of the aminocyclitol subunit of hygromycin A in *Streptomyces hygroscopicus* NRRL 2388. *Chem Biol* 16:1180–1189.
45. Kneidinger B, et al. (2001) Identification of two GDP-6-deoxy-D-*lyxo*-4-hexulose reductases synthesizing GDP-D-rhamnose in *Aneurinibacillus thermoaerophilus* L420-91^T. *J Biol Chem* 276:5577–5583.
46. Major LL, Wolucka BA, Naismith JH (2005) Structure and function of GDP-mannose-3′,5′-epimerase: An enzyme which performs three chemical reactions at the same active site. *J Am Chem Soc* 127:18309–18320.
47. Wolucka BA, Van Montagu M (2003) GDP-mannose 3′,5′-epimerase forms GDP-L-gulose, a putative intermediate for the *de novo* biosynthesis of vitamin C in plants. *J Biol Chem* 278:47483–47490.
48. Konishi T, et al. (2007) A plant mutase that interconverts UDP-arabinofuranose and UDP-arabinopyranose. *Glycobiology* 17:345–354.
49. Rautengarten C, et al. (2011) The interconversion of UDP-arabinopyranose and UDP-arabinofuranose is indispensable for plant development in *Arabidopsis*. *Plant Cell* 23:1373–1390.
50. Breton C, Snajdrová L, Jeanneau C, Koca J, Imberty A (2006) Structures and mechanisms of glycosyltransferases. *Glycobiology* 16:29R–37R.
51. Richards MR, Bai Y, Lowary TL (2013) Comparison between DFT- and NMR-based conformational analysis of methyl galactofuranosides. *Carbohydr Res* 374:103–114.
52. Mangoni A (2012) Strategies for structural assignment of marine natural products through advanced NMR-based techniques. *Handbook of Marine Natural Products*, eds Fattorusso E, Gerwick WH, Taglialatela-Scafati O (Springer, New York), pp 541–543.
53. Cavener DR (1992) GMC oxidoreductases. A newly defined family of homologous proteins with diverse catalytic activities. *J Mol Biol* 223:811–814.
54. Dym O, Eisenberg D (2001) Sequence-structure analysis of FAD-containing proteins. *Protein Sci* 10:1712–1728.