Comparing Galactan Biosynthesis in *Mycobacterium tuberculosis* and *Corynebacterium diphtheriae*

Received for publication, September 18, 2016; and in revised form, December 28, 2016. Published, JBC Papers in Press, December 30, 2016, DOI 10.1074/jbc.M116.759340

Darryl A. Wesener, Matthew R. Levengood, and Laura L. Kiessling

Edited by Gerald W. Hart

The suborder Corynebacterineae encompasses species like *Corynebacterium glutamicum*, which has been harnessed for industrial production of amino acids, as well as *Corynebacterium diphtheriae* and *Mycobacterium tuberculosis*, which cause devastating human diseases. A distinctive component of the Corynebacterineae cell envelope is the mycolyl-arabinogalactan (mAG) complex. The mAG is composed of lipid mycolic acids, and arabinofuranose (Ara) and galactofuranose (Gal) carbohydrate residues. Elucidating microbe-specific differences in mAG composition could advance biotechnological applications and lead to new antimicrobial targets. To this end, we compare and contrast galactan biosynthesis in *C. diphtheriae* and *M. tuberculosis*. In each species, the galactan is constructed from uridine 5'-diphosphate-α-D-galactofuranose (UDP-Gal), which is generated by the enzyme UDP-galactopyranose mutase (UGM or Glf). UGM and the galactan are essential in *M. tuberculosis*, but their importance in Corynebacterineae species was not known. We show that small molecule inhibitors of UGM impede *C. glutamicum* growth, suggesting that the galactan is critical in corynebacteria. Previous cell wall analysis data suggest the galactan polymer is longer in mycobacterial species than corynebacterial species. To explore the source of galactan length variation, a *C. diphtheriae* ortholog of the *M. tuberculosis* carbohydrate polymerase responsible for the bulk of galactan polymerization, GlfT2, was produced, and its catalytic activity was evaluated. The *C. diphtheriae* GlfT2 gave rise to shorter polysaccharides than those obtained with the *M. tuberculosis* GlfT2. These data suggest that GlfT2 alone can influence galactan length. Our results provide tools, both small molecule and genetic, for probing and perturbing the assembly of the Corynebacterineae cell envelope.

*Mycobacterium tuberculosis* and *Corynebacterium diphtheriae*, the etiological agents of tuberculosis and diphtheria, respectively, are notorious members of the bacterial suborder Corynebacterineae. This taxon includes other pathogenic bacterial species, including the causative agents of leprosy, nocardiosis, and buruli ulcers. One notable feature of Corynebacterineae is the cell envelope, which has a unique composition. The Corynebacterineae cell envelope contains mycolic acids appended to branched polymers of Ara (arabinan), which are linked to peptidoglycan through a linear polymer of Gal (galactan) (1, 2). The macromolecular structure that extends beyond the peptidoglycan is referred to as the mAG complex. Neither the mAG nor its individual components are present in host mammals. Moreover, in *M. tuberculosis* the mAG complex serves as a barrier to antitubercular drugs and can modulate the human immune response in favor of bacterial immune evasion (3, 4). A complete understanding of mAG assembly and variation could yield novel strategies for therapeutic intervention.

Structural features of the mycobacterial and corynebacterial cell envelopes have been characterized (Fig. 1) (5–11). The galactan is anchored to the peptidoglycan at the C-6 position of peptidoglycan muramic acid. The linker includes a phosphodiester-linked 1-α-D-N-acetylglucosamine residue with an α-substituted 1-rhamnopyranosyl residue at the 3 position (Fig. 2). The 4-position of rhamnose is substituted with the galactan, a polymer of alternating β(1–5)- and β(1–6)-linked β-D-Gal residues. The enzymes that mediate galactan biosynthesis in mycobacteria have been identified. Two galactofuranosyltransferases, GlfT1 and GlfT2, generate the galactan using UDP-Gal, the activated Gal sugar donor afforded by the action of UGM (12, 13). GlfT1 catalyzes the addition of two to three Gal residues to the C-4 hydroxyl of 1-rhamnose (14, 15). The galactofuranosyl polymerase GlfT2 then promotes the sequential addition of the alternating β(1–5) and β(1–6) Gal linkages (16–19). After galactan polymerization, 3 α-Araf residues are added to the C-5 hydroxyl groups of the 8th, 10th, and 12th Gal residues of the galactan (20). Arabinofuranosyltransferases elaborate these residues to append a branched Ara/poly saccharide. Subsequently, multiple terminal Ara residues are acylated to form long chain mycolic acid esters. Enzymes involved in mycolic acid biosynthesis or arabinan production are the targets of clinically used antitubercular drugs (21).

---

3 Supported by a National Institutes of Health Postdoctoral Fellowship (F32GM089219).
4 To whom correspondence should be addressed: Dept. of Chemistry, University of Wisconsin-Madison, 1101 University Ave., Madison, WI, 53706. Tel.: 608-262-0541; Fax: 608-265-0764; E-mail: kiesling@chem.wisc.edu.

© 2017 by The American Society for Biochemistry and Molecular Biology, Inc. Published in the U.S.A.
Although mAG biosynthesis and structure is often studied in *M. tuberculosis*, *Corynebacterium* species have recently been used as models to understand mAG assembly. Their advantages include a decreased doubling time, reduced biosafety designation, and the availability of tools for genetic manipulation (22–25). Although most mAG biosynthetic genes are essential in *M. tuberculosis* (26), their deletion in *C. glutamicum* often yields slow growing but viable mutants. For example, *C. glutamicum* mutants lacking AftA were recently used to determine that this arabinofuranosyltransferase appends three α(1–5)–Araf residues to the galactan to initiate arabinan biosynthesis (27). Another *C. glutamicum* mutant revealed Pks13 catalyzes the final step of mycolic acid biosynthesis in corynebacteria and mycobacteria (24). Unexpectedly, a *C. glutamicum* mutant lacking the first enzyme required for activated Araf donor sugar biosynthesis, *ubiA*, is viable (20). This strain is devoid of Araf, and that it could be isolated and cultured was surprising. These examples highlight how experiments using
Corynebacterial Galactan Biosynthesis

Corynebacterium species can provide important insight into cell envelope biosynthesis within the Corynebacterineae suborder.

Although each Corynebacterineae species possesses an mAG of similar constitution, fine structural features of the mAG can vary. The arabinan of C. diphtheriae lacks the 1,3,5-linked Araf residues that are responsible for branching, suggesting this arabinan is less complex than that of other Corynebacterineae (7). Accordingly, arabinan assembly in mycobacteria is mediated by a larger collection of enzymes. Six or more arabinofuranosyltransferases are involved in mycobacteria, and at least one of these enzymes is inhibited by the first-line antitubercular drug ethambutol (28–30). Of the six mycobacterial arabinofuranosyltransferases, three belong to the Emb family of enzymes. In contrast, corynebacteria encode a single Emb homolog that is most closely related to M. tuberculosis EmbC (20). Thus, the arabinan of C. diphtheriae is generated using fewer enzymes and is simpler than that of M. tuberculosis. The mycolic acids from Mycobacterium and Corynebacterium species also vary. Mycobacterial mycolic acids possess chains of 70–90 carbon atoms, whereas the mycolic acids of C. diphtheriae are shorter, with a chain of 30–36 carbons (7, 31). Additionally, M. tuberculosis mycolic acids commonly contain functionalities such as cis-cyclopropane that are absent from fast-growing Mycobacterium smegmatis or corynebacterial mycolic acids. There also are differences in the galactan. Analyses of the galactan length (20) indicate the corynebacterial galactan is shorter than that of M. tuberculosis (5, 7, 27, 32). Understanding the source of these differences can lend insight into the molecular basis for the cell envelope properties of specific species.

The M. tuberculosis glycosyltransferase GlfT2 (EC 2.4.1.288) is a bifunctional carbohydrate polymerase that generates the bulk of the galactan (17–19, 33). Studies with chain-terminating glycosyl donors indicate that GlfT2 is sequence-selective, and its fidelity for forming a sequence of alternating β(1–5) and β(1–6) linkages is high (34, 35). Within the GlfT2 polypeptide is but a single active site (33, 36); site-directed mutagenesis indicates substitution of key amino acids abrogates the formation of both β(1–5) and β(1–6) linkages (36). Experiments with isotope-labeled acceptors reveal that GlfT2 is processive (37), and its propensity for generating alternating β(1–5) or β(1–6) Galf linkages is a consequence of processive elongation (38). The enzyme not only controls polymer sequence but also polymer length. In a test tube GlfT2 can generate polymers of a similar length to those obtained from cells (19). With synthetic acceptors, the identity of the anomeric lipid was a critical determinant of product polysaccharide length (17–19). Acceptors with short alkyl anomic substituents afforded only short oligosaccharides, whereas those with longer lipids on the acceptor afforded polysaccharide products similar to those in cells. These findings led to the proposal that M. tuberculosis GlfT2 uses the acceptor lipid as a tether and that polymer length is controlled by bivalent substrate binding (19). The DXD motif mediates glycosyltransferase coordination to a divalent cation, and M. tuberculosis GlfT2 variants in this DXD motif afforded truncated oligosaccharide products (36, 39). Thus, GlfT2 is bifunctional with the ability to control the length of polymerized M. tuberculosis galactan. Whether these attributes of the M. tuberculosis GlfT2 are preserved in other Corynebacterineae species is unclear.

To compare and contrast galactan biosynthesis in different species, we examined the galactan biosynthetic enzymes from C. glutamicum and C. diphtheriae NCTC 13129, including GlfT2 and UGM. The putative C. diphtheriae UGM (DIP2203) catalyzes the isomerization of UDP-Galp and UDP-Galf with an efficiency similar to other prokaryotic orthologs. This finding suggests that differences in galactan length are unlikely to arise from differences in cellular concentration of UDP-Galf. To explore the role of the galactan in cells, we employed small molecule UGM inhibitors. These compounds inhibit C. diphtheriae UGM in vitro and prevent the growth of C. glutamicum. These observations support an essential role for the galactan in corynebacteria as well as mycobacteria and suggest that the small molecule UGM inhibitors can be used to probe the cellular roles of Galf. With regard to GlfT2, our results indicate that the DIP2198 gene product is a bifunctional galactofuranosyltransferase similar to the M. tuberculosis GlfT2. The corynebacterial glycosyltransferase can elongate synthetic acceptors to afford polysaccharides commensurate in length to those isolated from corynebacteria. As with the M. tuberculosis GlfT2, the product polysaccharides attained from variants with amino acid changes in the proposed donor binding site are truncated. These findings indicate that GlfT2 orthologs have an intrinsic ability to control polysaccharide length. This length control mechanism differs from that employed in O-antigen biosynthesis in which length is controlled by an auxiliary protein (40–42). Our investigations lay a foundation for dissecting the molecular details and mechanisms of galactan biosynthesis using genetic and chemical genetic tools.

Results

Inhibitors of Galactan Biosynthesis Block Corynebacterial Growth—Unlike mycobacteria, corynebacteria do not require mycolic acids (24) or arabinan (20). The survival of corynebacteria without integral mAG components raises the issue of whether these species require the galactan. Small molecule UGM inhibitors provide the means to address this issue (Fig. 2) (43). We previously produced the C. diphtheriae UGM (DIP2203) (44). In this study we characterized the enzyme’s catalytic properties and then identified a C. diphtheriae UGM inhibitor that could be used to probe the role of the galactan. The initial velocities of UDP-Galp production were calculated from reaction mixtures containing a range of UDP-Galf concentrations (45, 46). Steady-state kinetic parameters were determined from the Michaelis-Menten equation, which afforded a K_{m} of 47 ± 5 μM and a k_{cat} value of 20 ± 0.6 s^{-1} (Table 1). The enzymatic efficiency, or k_{cat}/K_{m}, for C. diphtheriae UGM is 4.3 ± 0.5 × 10^{3} m^{-1}s^{-1}. These kinetic constants are similar to those obtained with UGM orthologs from other prokaryotes, including those from Klebsiella pneumoniae and Escherichia coli (46, 47).

Our previous experience suggested that inhibitors of M. tuberculosis UGM should act against C. diphtheriae UGM (44). To assess compound affinity for UGM, we validated that a previously described fluorescence polarization (FP) assay (48) could be applied to the C. diphtheriae UGM. The enzyme

2946 JOURNAL OF BIOLOGICAL CHEMISTRY

VOLUME 292 • NUMBER 7 • FEBRUARY 17, 2017

Downloaded from http://www.jbc.org/ at Massachusetts Institute of Technology on January 30, 2018
bound the fluorescein-bearing UDP derivative with an apparent affinity of 24 ± 2 nM. Small molecules could then be tested for their ability to displace the fluorescent probe from the enzyme in a competition assay (49). Only compounds 1 and 2 were effective. Their calculated dissociation constants were similar, 8 ± 1 μM for 1 and 10 ± 2 μM for 2. We next evaluated the UGM ligands as inhibitors. As expected from the FP assay results, UDP and compound 3 did not substantially inhibit C. diphtheriae UGM activity at 15 μM, but compounds 1 and 2 impeded conversion of UDP-Gal to UDP-Galp by 83 ± 2% and 69 ± 3%, respectively (Fig. 3B). At 50 μM UDP-Galp, a concentration near the Km value we report here, the IC50 value of compound 1 was determined to be 5 ± 1 μM.

We tested whether inhibitor 1 could be used ascertain the consequences of inhibiting galactan biosynthesis. C. glutamicum (ATCC 13032) is a non-virulent genetically tractable model organism that had been employed to conclude that corynebacteria can survive without mycolic acids or arabinan (20, 24). We anticipated the UGM inhibitor would be effective because the sequence similarity of the C. diphtheriae and C. glutamicum UGM is 85%. Indeed, exposure of C. glutamicum to compound 1 led to growth inhibition (Fig. 3C). The results of this chemical genetics experiment suggest that the galactan is the outermost polysaccharide required for corynebacterial growth.

**Galactan Polymerization by C. diphtheriae—**The C. diphtheriae GlfT2 (DIP2198 gene) and M. tuberculosis GlfT2 share 57% amino acid identity, but cell wall analysis indicates that the length of the galactan in corynebacteria is shorter than that in M. tuberculosis. This difference prompted us to compare the catalytic properties of the GlfT2 orthologs. If GlfT2 can control galactan length, the C. diphtheriae enzyme should give rise to shorter polymers than those afforded by M. tuberculosis GlfT2. We produced a His-tagged C. diphtheriae GlfT2 and monitored its carbohydrate polymerase activity using matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (19, 37). The reaction conditions employed with acceptor 4a (Fig. 4) were similar to those used to assay M. tuberculosis GlfT2 activity. Polysaccharides were attained demonstrating that the DIP2198 gene product can generate polymers of galactofuranose (Fig. 4B, left). On the basis of previous studies with M. tuberculosis GlfT2 (17, 18, 33, 36), we anticipated that the C. diphtheriae GlfT2 product polysaccharides would be composed of alternating β(1–6) and β(1–5)–Galp linkages. We, therefore, also examined elongation from disaccharide acceptor, compound 5, which has a β(1–5)–Galp linkage and, therefore, is a regiosomer of acceptor 4a (36). The C. diphtheriae GlfT2 also elongated compound 5 to afford polymers of similar length to those attained with 4a (Fig. 4B, right). The ability of both synthetic acceptors to function as substrates of the C. diphtheriae GlfT2 enzymes generate Galp polymers, but mass spectrometry analysis suggested that the polymeric products from the C. diphtheriae ortholog are substantially shorter than those from M. tuberculosis GlfT2 (Fig. 4C). We, therefore, examined the rate of polymerization by C. diphtheriae GlfT2. We employed compound 4a in a coupled continuous enzyme assay that monitors the release of UDP (50). The data indicate that C. diphtheriae GlfT2 is a slower polymerase than is M. tuberculosis GlfT2. Under the same conditions (1.25 mM UDP-Galp and 0.2 mM 4a), the C. diphtheriae GlfT2 was ~17-fold slower than the M. tuberculosis GlfT2 (0.47 ± 0.24 μM/min for C. diphtheriae GlfT2 versus 8.2 μM/min for M. tuberculosis GlfT2). The slower rate of the C. diphtheriae enzyme was not a function of decreased protein stability as the C. diphtheriae GlfT2 was less prone to aggregation and precipitation than M. tuberculosis GlfT2.

**Anomeric Lipid Influences Polysaccharide Length—**We found previously that the products of the M. tuberculosis GlfT2 depend on the attributes of the anomeric lipid of the acceptor. With longer lipid substituents, longer polysaccharide products were obtained (19). We tested whether the C. diphtheriae ortholog affords similar results. We used compounds 4b and 4c as substrates; the former has a shorter lipid group than our benchmark substrate 4a, and the latter has a longer lipid substituent. When 4b was incubated with C. diphtheriae GlfT2, no
Polymeric products were observed (Fig. 5A). In contrast, GlfT2-catalyzed elongation of 4c, a substrate with an extended lipid group, yielded polysaccharide products nearly twice as long, +24 Galf residues (Fig. 5B).

Acceptor Glycan Influences Polysaccharide Length—We previously observed a kinetic lag phase for M. tuberculosis GlfT2-catalyzed elongation of disaccharide acceptors (37). This lag phase was not observed with trisaccharide or tetrasaccharide acceptors (37, 38). To test C. diphtheriae GlfT2 processing of acceptors that vary in the number of saccharide residues, we used a chemoenzymatic approach to generate oligogalactofuranoside acceptors (37). The tetrasaccharide acceptor gave rise to polymeric products up to +32 Galf residues (Fig. 5C). These results are consistent with findings that the catalytic site of GlfT2 can accommodate multiple saccharide residues (37). The tetrasaccharide acceptor gave rise to polymeric products up to +32 Galf residues (Fig. 5C). These results are consistent with findings that the catalytic site of GlfT2 can accommodate multiple saccharide residues (37). Although the tetrasaccharide afforded substantially longer polymeric products, they remain ~10 residues shorter than those obtained using the M. tuberculosis GlfT2 enzyme (38). The rate of C. diphtheriae GlfT2-catalyzed polymerization of the tetrasaccharide acceptor is still significantly slower than that obtained with the M. tuberculosis enzyme (1.8 μmol/min).

Site-directed Mutagenesis for Manipulating Polysaccharide Length—The longer polysaccharide products obtained from the M. tuberculosis GlfT2 versus the C. diphtheriae ortholog suggests that polysaccharide length can be programmed genetically. We previously identified active site residues in the M. tuberculosis GlfT2 important for polymerization, and mutant proteins with substitutions in the putative UDP-Gal binding site led to polysaccharides of attenuated length (36). We tested the generality of the mode of length control by installing point mutations at these sites in the C. diphtheriae GlfT2. From a primary sequence alignment, we identified two DXXD motifs within C. diphtheriae GlfT2 similar to those found in M. tuberculosis GlfT2 (36, 39). By analogy, a DDD motif (residues Asp-273–Asp-275 of C. diphtheriae GlfT2) should mediate cation coordination and binding to UDP-Gal, whereas the DDAE motif (Asp-396–Glu-399) should contain the catalytic base (Fig. 6A). Accordingly, both C. diphtheriae Asp→Ala and Asp→Glu variants were generated at residues Asp-396 and Asp-397. When tested with acceptor 4a, no elongated products were detected. This abrogation of activity is similar to what was found previously for M. tuberculosis GlfT2 (36). The results presented here as well as autoradiographic enzymatic assays and a recent X-ray crystal structure of M. tuberculosis GlfT2 (39) all are consistent with the assignment of Asp-397 as the catalytic base in C. diphtheriae GlfT2.

We also evaluated GlfT2 variants expected to have altered binding to the glycosyl donor UDP-Galf. A series of Asp→Ala...
and Asp→Glu protein substitutions at Asp-273 and Asp-275 were generated. The alanine variants failed to act on compound 4a. These data contrast with our findings with M. tuberculosis GlfT2, as substitution of either aspartic acid residue with an alanine afforded a mutant enzyme that yielded polymeric products (36). Because alanine mutations in the glycosyl donor binding site of C. diphtheriae GlfT2 abrogated polymerase activity, we assayed more conservative Asp→Glu variants. In the reaction mixture with the D273E GlfT2, shorter polymeric products were obtained. The longest oligosaccharide detected in the reaction mixture was \(1/10\) Gal residues, as compared with \(11\) Gal using the wild-type protein (Fig. 6B). The D273E variant similarly processed the tetrasaccharide acceptor to afford shorter oligosaccharide products. With the D275E variant, no products were detected with either compound 4a or tetrasaccharide acceptor. As was observed with the M. tuberculosis GlfT2, C. diphtheriae GlfT2 variants with perturbations in the glycosyl donor binding site afford substantially shorter carbohydrate polymers; the products bore roughly half the number of Gal residues as those produced by wild-type enzyme. These products are unlikely to be substrates for arabinofuranosyltransferases, as arabinan linkage is predicted to occur at Gal residues 8, 10, and 12 within the galactan.

**Discussion**

The taxon Corynebacterineae includes many bacterial species of industrial and medical importance (51–53). Although Corynebacterineae share cell envelope features, different species occupy distinct ecological niches, and the details of their cell wall composition can vary. Survival in specific physiological environments may have been responsible for alterations in the constitution of their cell envelopes and mAG complexes. The few studies characterizing the mAG of Corynebacterineae species have uncovered structural differences (5, 7, 20, 54). We probed galactan biosynthesis in corynebacteria to better understand the mechanisms by which different species assemble this fundamental component of the mAG complex. We also examined the physiological importance of the mAG using a chemical genetic approach.

Whether disruption of galactan biosynthesis would inhibit corynebacterial growth was an open question as previous experiments have revealed that components once thought to be
Corynebacterial Galactan Biosynthesis

essential for Corynebacterineae are dispensable in *C. glutamicum*. Specifically, microbial genetics and chromosomal deletion experiments have shown that mycolic acids (24, 55) and the arabinan are not necessary for viability in this species (20, 27). In contrast, attempts to isolate *M. tuberculosis* mutants with similar genotypes have failed (26). Indeed, no *M. tuberculosis* strains lacking the enzymes that mediate galactan biosynthesis have been isolated, suggesting the galactan is required for *M. tuberculosis* viability.

To test whether the galactan is essential for corynebacterial growth, we used a chemical biology strategy. We determined the catalytic activity of the putative *C. diphtheriae* UGM, DIP2203. This enzyme catalyzes the isomerization of UDP-Galp and UDP-Galf with kinetic parameters similar to those of other prokaryotic UGMs (56–58), and the *C. diphtheriae* enzyme can be inhibited by small molecules developed against the *M. tuberculosis* UGM (43, 44). When our most potent small molecule inhibitor of corynebacteria UGM, compound 1, was added to *C. glutamicum*, corynebacterial growth was inhibited. This result suggests the galactan is essential in *C. glutamicum*. This finding implies that the peptidoglycan-galactan conjugate is the minimal cell envelope structure needed for viability of corynebacteria, whereas the polysaccharide emanating from the galactan, the arabinan (20), and the cell surface mycolic acids appear dispensable (24).

Previous cell wall analyses suggest that galactose is present at lower abundance in corynebacteria as compared with mycobacteria. This observation suggests that corynebacteria have a shorter galactan (5, 7, 27, 32). The *in vitro* carbohydrate polymerase activity of *C. diphtheriae* GlfT2 gives rise to shorter polymers, consistent with those observations (19). These results indicate that the GlfT2 glycosyltransferases have an intrinsic ability to control length. Our data further suggest that galactan length can be manipulated through mutation or heterologous expression of polymerases. Thus, polysaccharide length is dictated by the enzyme’s sequence.

Several other microbial cell envelope glycopolymers possess defined length distributions. These include peptidoglycan, lipopolysaccharide (LPS) O-antigens, capsular polysaccharides, and teichoic acids (59–63). Efforts to elucidate how the length of these polysaccharides is controlled have revealed the importance of accessory proteins. For example, in *Wzy*-dependent LPS O-antigen biosynthesis product length is controlled by the polysaccharide co-polymerase protein, termed Wzz (40, 61, 64, 65). An *in vitro* polymerization system was used to demonstrate that Wzy alone is sufficient for polymerization, but that addition of Wzz to the reaction can modulate the product length distribution (41). An LPS O-antigen assembled in an ABC transporter-dependent manner also employs an independent chain-length regulator protein. In *E. coli* serotype O9a O-antigen assembly, an extended coiled-coil in the chain length regulator WbdD operates as a molecular ruler to physically separate the kinase termination activity of WbdD, from the glycosylpolymerase activity of WbDA (42). Unlike these systems, there are no apparent accessory proteins that can modulate the length of polysaccharides generated by GlfT2. The galactan and LPS O-antigen are dissimilar in that the galactan is an essential component of the cell envelope, whereas O-antigen is not required for survival and likely evolved for host interactions (66). Bacteria may have evolved unique accessory proteins to alter O-antigen length as a mechanism for modulating surface antigenicity, but such a selective pressure would not be present for galactan assembly.

The mechanisms underlying length control by GlfT2 are unknown. One possibility for the shortened polysaccharide length obtained with *C. diphtheriae* GlfT2 versus the *M. tuberculosis* ortholog is that the rate of polymerization for the former is 17-fold slower. Still, the kinetic rate alone is unlikely to be solely responsible for the divergent polymer length distributions we observed. Specifically, the tetrasaccharide acceptor gave rise to longer polysaccharides (+32 Galp residues; Fig. 5C) yet the initial velocities observed for *C. diphtheriae* GlfT2 remained slower (1.8 μM/min) than those obtained for *M. tuberculosis* enzyme (8.2 μM/min). Acceptor concentration had little effect on polysaccharide length. When the reaction was conducted at an 8-fold lower acceptor concentration (25 μM compound 4a), the longest polymer was +15 Galp residues, and a nearly identical product profile was attained. The available data indicate that the features of the acceptor and sequence of the glycosyltransferase module produce polysaccharide length (19). For example, the longer polymers obtained from the tetrasaccharide versus those generated from disaccharide 4a may be due to the former binding more efficiently and remaining bound through a larger number of catalytic events. Similarly, the affinity of the glycosyl donor may influence polysaccharide product length control. These two catalytic parameters may contribute to GlfT2 length control. Moreover, the data suggest that length of polysaccharides generated by GlfT2 can be modulated through simple mutagenesis. A similar mutagenesis strategy may be useful for modulating the length of other biologically important polysaccharides. Specifically, both GlfT2 and cellulose synthase belong to the glycosyltransferase family 2 (GT-2) glycosylpolymerases (67), and attenuation of cellulose polymer length may be useful for the industrial production of cellulosic derived commodities such as ethanol (68).

A mechanism for carbohydrate polymerase length control based on sugar donor binding has been suggested in an unrelated process; that is, the synthesis of *Streptococcus pneumoniae* type 3 capsule polysaccharide (69). This polysaccharide is generated by a processive synthase. The enzyme uses two sugar donors, UDP-glucose and UDP-glucuronic acid, for [4→3]-d-GlcUA-β(1→3)-d-GlcUA-β(1→4) polymer formation. Data suggest that the substrate concentrations and the affinity of UDP-glucuronic acid binding to the enzyme influence the product polysaccharide length (69, 70).

The ability of GlfT2 to control polysaccharide length also could rest in an alternative motif. For example, a chain length regulatory proline motif (WPQ) was recently identified in transmembrane glycosyltransferase type 3 arabinofuranosyltransferases (23, 71). Lastly, an alternative hypothesis based on GlfT2 quaternary structure has been put forth (39). The tetrameric structure observed in the *M. tuberculosis* GlfT2 crystals contained a face of mostly hydrophobic and positively charged residues that the authors suggest promotes membrane association. At the center of the tetramer resides a large hollow cavity
that the authors suggest functions as an area for growing galactan to be displaced into during polymerization, with the size of the cavity influencing product polymer length (39).

Several of the aforementioned mechanisms may contribute to regulating galactan length. For example, the cell membrane may function in a similar role as that proposed for the GlfT2 acceptor anomeric lipid. Thus, the distance between the GlfT2 active site and membrane combined with enzyme binding to substrate would influence the length of polymeric products. We anticipate that future experiments will provide additional insight into how galactan length is controlled.

**Experimental Procedures**

**Synthesis of Substrates, Inhibitors, and Fluorescence Polarization Probe—**UDP-Galαf was prepared as previously reported (72), as were compounds 1, 2, and 3 (43). Acceptor disaccharides 4a-c and 5 were synthesized using published procedures (19). The FP probe UDP-10C-fluorescein was synthesized as previously described (48). A tetrasaccharide acceptor for GlfT2 was prepared as previously reported (43). Acceptor disaccharide, UDP-Gal, was prepared as previously described (37).

**Cloning of Genes Involved in Galactan Biosynthesis in C. diphtheriae—**DIP2198 (glfT2) and DIP2203 (glf) were cloned from *C. diphtheriae* NCTC 13129 genomic DNA (American Type Culture Collection (ATCC)). The DIP2198 gene (glfT2) was amplified via the polymerase chain reaction using the forward primer 5’-GGGAATTC-CATATGCTTACCTGAGGACGCTGCCGAATACGGCTTGCGCG-3’ and the reverse primer 5’-ATAGTTTAACCGGGCGGGTCCTTAGCGCTGAGGACGCTGCCGAATACGGCTTGCGCG-3’. The forward and reverse primers added an NdeI and BamHI restriction site, respectively. The purified PCR product and pET-28a(+) vector (EMD Chemicals) were digested with NdeI and BamHI restriction enzymes (New England Biolabs) and ligated into pET-28a (EMD Chemicals). The DIP2203 gene was amplified via a polymerase chain reaction using the forward primer 5’-GCCACTGTTTTATCAAGTGAGACGTGGAGCCTGAGCATTGGCAGCGG-3’ and the reverse primer 5’-GCCACTGTTTTATCAAGTGAGACGTGGAGCCTGAGCATTGGCAGCGG-3’. The forward and reverse primers added NdeI and NotI restriction sites to the amplicon, respectively. purified PCR product and pET-28a(+) vector (EMD Chemicals) were digested with NdeI and NotI restriction endonucleases (New England Biolabs). The double-digested products were purified using the QIAquick Gel Extraction kit (Qiagen). Digested pET-24a(+) vector and glf insert were ligated with T4 DNA ligase (Fermentas). The vector encoded C-terminal hexahistidine tag, and in-frame insertion of glf was confirmed by DNA sequence analysis.

**C. diphtheriae GlfT2 Mutagenesis—**Sequence analysis was used to identify putative active site residues in *C. diphtheriae* GlfT2. QuikChange mutagenesis (Stratagene) was used to generate point mutations in putative active site residues of the glfT2 gene. The primer and its reverse complement used to generate glfT2 mutants are described in Table 2. The underlined triplet codon denotes the mutated residue. The pET-28a:His6-glfT2 construct generated in this study was used as the template for the polymerase chain reaction. Template DNA was digested with DpnI (Promega) after PCR amplification. The DpnI-digested mixture was transformed into DH5α E. coli by electroporation. Plasmid DNA was isolated using the QiAprep Spin Miniprep kit (Qiagen). Mutation of the desired codon was confirmed by DNA sequence analysis.

**Expression of C. diphtheriae glfT2—**The gene encoding *C. diphtheriae* His6-glfT2 was overexpressed and purified using a protocol similar to that employed for *M. tuberculosis* His6-glfT2 (37). Briefly, a pET-28a plasmid encoding His6-glfT2 was transformed into BL21(DE3) Tuner E. coli cells (Novagen) via electroporation. Cultures from a single colony were grown in Luria-Bertani (LB) medium supplemented with 50 μg/ml kanamycin at 37 °C until A600 = 0.8. At this time cultures were placed on ice for 1 h. Protein overexpression was induced by the addition of isopropyl-β-D-thiogalactopyranoside to 0.3 mM. Cultures were allowed to grow at 18 °C for 18–20 h, at which time cell pellets were harvested and centrifuged (5000 × g), collected, and stored at −80 °C until use. Cell pellets were thawed on ice in a lysis buffer (20 ml/liter of growth) that contained 50 mM potassium phosphate (pH 7.9), 300 mM sodium chloride (NaCl), 20 mM imidazole, 1:200 protease inhibitor mixture III (Calbiochem) and lysozyme. Cells were disrupted by sonication (Branson Sonifer 450, 5 × 10 s cycles at 90% duty cycle). Lysates were cleared by centrifugation for 1 h at 22,000 × g. The soluble lysate was filtered (0.22 μm, Amicon) before loading onto a pre-equilibrated (in lysis buffer) 5-ml HisTrap column (GE Healthcare) at 1.0 ml/min using an AKTA FPLC system (Amerham Biosciences). Protein was eluted using a step gradient of imidazole. Glycerol was added (10% v/v) to 1-ml fractions of >90% purity, with purity assessed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and staining with Coomassie Blue. Samples were vitrified in liquid nitrogen and stored at −80 °C until use. Typical yields were 2 mg/liter culture. When performing enzyme assays, a sample of His6-GlfT2 was removed and dialyzed twice against 2 liters of 50 mM HEPES (pH 7.9) for 16 h before use.

**TABLE 2**

<table>
<thead>
<tr>
<th>Primers used to generate mutant glfT2 genes</th>
<th>Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>D273A</td>
<td>5’-GCCCTTCATCCTCGTATTACCGGGGCAGG-3’</td>
</tr>
<tr>
<td>D273E</td>
<td>5’-GCCCTTCATCCTCGTATTACCGGGGCAGG-3’</td>
</tr>
<tr>
<td>D275A</td>
<td>5’-GCCCTTCATCCTCGTATTACCGGGGCAGG-3’</td>
</tr>
<tr>
<td>D275E</td>
<td>5’-GCCCTTCATCCTCGTATTACCGGGGCAGG-3’</td>
</tr>
<tr>
<td>D396A</td>
<td>5’-GCCCTTCATCCTCGTATTACCGGGGCAGG-3’</td>
</tr>
<tr>
<td>D396E</td>
<td>5’-GCCCTTCATCCTCGTATTACCGGGGCAGG-3’</td>
</tr>
<tr>
<td>D397A</td>
<td>5’-GCCCTTCATCCTCGTATTACCGGGGCAGG-3’</td>
</tr>
<tr>
<td>D397E</td>
<td>5’-GCCCTTCATCCTCGTATTACCGGGGCAGG-3’</td>
</tr>
</tbody>
</table>
Corynebacterial Galactan Biosynthesis

7.4), 100 mM NaCl, and 5 mM EDTA in a 10,000 molecular weight cut-off dialysis cassette (Pierce). Protein concentration was determined with the BCA assay (Pierce) using bovine serum albumin as a standard. The point variants of GlfT2 were expressed and purified identically to the wild-type protein.

Production of C. diphtheriae UGM—The pET-24a:glf-His<sub>6</sub> plasmid was transformed into BL21 (DE3) Tuner E. coli (Novagen) by electroporation. Cultures were grown in LB medium supplemented with 50 μg/ml kanamycin at 37 °C until A<sub>600</sub> 0.6. Cells were cooled on ice, and isopropyl-β-d-thiogalactopyranoside was added to 0.1 mM to induce glf-His<sub>6</sub> expression. After 18 h at 20 °C, cells were harvested by centrifugation (5000 × g), and cell pellets were stored at −80 °C until use. Cells were thawed on ice and resuspended in 50 mM potassium phosphate (pH 7.4), 300 mM NaCl, and 25 mM imidazole (20 ml/liter of growth). Cells were disrupted by lysozyme, Triton X-100 (0.1% v/v), and sonication (Branson Sonifier 450), and phenylmethylsulfonlfluoride was added to 1 mM. Cellular debris was cleared by centrifugation (22,000 × g, 1 h, 4 °C). The lysate was filtered (0.22 mm, Amicon) and applied to a 1-ml HisTrap column (GE Healthcare) at 0.75 ml/min on an AKTA FPLC (Amersham Biosciences). Protein was eluted with a linear gradient of 25–400 mM imidazole in 50 mM potassium phosphate (pH 7.4), 300 mM NaCl. Fractions of >80% pure were combined and dialyzed at 4 °C against 2 liters of 50 mM potassium phosphate (pH 7.0), 300 mM NaCl, and then 2 liters of 50 mM potassium phosphate (pH 7.0), 5 mM NaCl. The dialyzed solution was applied to a 1-ml HiTrap Q HP column (GE Healthcare) on an AKTA FPLC. Purified protein was eluted with a gradient using 50 mM potassium phosphate (pH 7.0), 1 mM NaCl. Fractions were analyzed for purity using SDS-PAGE and Coomassie Blue staining. Fractions >95% pure were collected, pooled, and dialyzed against 50 mM potassium phosphate (pH 7.0) and 150 mM NaCl. Typical yields of the two-step purification were 1.5 mg/liter of culture. Protein concentration was determined by absorbance against 50 mM potassium phosphate (pH 7.0) and 150 mM NaCl.

Enzymatic Activity of C. diphtheriae UGM—The enzymatic activity of UGM was determined 1 day after purification using an HPLC-based assay (45, 73). Reactions were performed using 20 nm C. diphtheriae UGM in 50 mM potassium phosphate (pH 7.0) and 20 mM freshly prepared sodium dithionite in a volume of either 50 μl or 80 μl. Enzyme activity was initiated with the rapid addition of chemically synthesized UDP-α-D-Galf at varying concentrations (72). Reactions were quenched by the addition of an equal volume of a 1:1 (v:v) mixture of chloroform:methanol. The aqueous and organic phases were separated by centrifugation, and the aqueous phase was assayed using a CarboPac PA-100 column (Dionex) on a Waters HPLC. The added substrate, UDP-Galf, and major product, UDP-Galp, were separated via isocratic elution using 200 mM ammonium acetate (pH 8.0) and monitored via absorbance of the uridine at 262 nm. Initial velocities were calculated based on the initial concentration of substrate and integration of the HPLC chromatogram. Steady-state kinetic constants were determined by nonlinear regression analysis with Prism 4 (Graphpad). Quantified error represents the S.D. of triplicate measurements.

Inhibition of C. diphtheriae UGM—Inhibition of UGM was assayed using the HPLC-based assay described above. Typical reactions were performed with 20 nm C. diphtheriae UGM, 50 μM UDP-Galf, and 20 mM freshly prepared sodium dithionite in 50 mM potassium phosphate (pH 7.0) and in a final volume of 80 μl. The UDP was dissolved in water and assayed at 15 μM. The 2-aminothiazole-based inhibitors 1–3 were dissolved in DMSO (dimethyl sulfoxide) and were all assayed at 15 μM (total concentration of DMSO equaled 1%). DMSO was used as a vehicle control at 1%. The percent of UDP-Galf converted by the enzyme was determined by integration of the chromatograph, and the data were normalized to the activity of a 50 μM UDP-Galf reaction in the absence of inhibitor or solvent vehicle. Each data point was performed in triplicate, and error bars represent the S.D. of the mean.

The IC<sub>50</sub> of 1 was determined using the HPLC-based assay described above. The inhibitor concentration was adjusted, whereas the other components were at a constant concentration of 20 nm C. diphtheriae UGM, 50 μM UDP-Galf, and 20 mM freshly prepared sodium dithionite. The percentage of DMSO in all reactions was 1%. The fraction of UDP-Galf converted to UDP-Galp was determined by integration of the HPLC chromatograph. Data were plotted using GraphPad Prism 4 and fit using the one site competition model.

FP to Measure Binding Constants to C. diphtheriae UGM—The affinity of the FP probe for C. diphtheriae UGM was determined by varying UGM concentration (17 nm constant probe concentration). Binding was measured in 50 mM potassium phosphate (pH 7.0) using 384-well, black microtiter plates (Corning) with a total volume of 30 μl per well. All points result from assays performed in triplicate, with fluorescence polarization measurements taken on an Infinite M1000 plate reader (Tecan). Data were plotted using GraphPad Prism 4 and fit using the one site competition model. To measure the apparent affinity of the small molecules for C. diphtheriae UGM, an assay similar to that described above was employed. Specifically, 17 nm UDP-10C-fluorescein probe, 20 nm C. diphtheriae UGM, and varying amounts of inhibitor were assayed in 50 mM potassium phosphate (pH 7.0). Each data point is the result of measurements performed in triplicate; error bars represent the S.D. of the mean. Data were plotted using GraphPad Prism 4 and fit with the one site competition model. To determine binding constants, the equation K<sub>app</sub> = K<sub>d</sub>(1 + ([I]/K<sub>i</sub>)) was used, where [I] = 17 nm and K<sub>i</sub> = 25 nm (as determined in this study).

Bioactivity Plate Assay—C. glutamicum ATCC 13032 were obtained from the ATCC. C. glutamicum was grown in LB media at 30 °C until A<sub>600</sub> = 1.5. This culture (400 μl) was used to seed 40 ml of prewarmed LB agar with bacteria. The seeded LB agar was poured into a sterile Omnimart (Nunc), and a sterile 96-well PCR plate was placed on top. After solidification of the agar, the PCR rack was removed to leave a plate with live embedded C. glutamicum and shallow wells for direct comparison of small molecule bioactivity. To each well, 10 μl of a solution of kanamycin, AT91, or ED103 were added. Kanamycin, dissolved in water, was assayed in the following concentrations: 500 ng/ml, 1 μg/ml, 10 μg/ml, 50 μg/ml, 100 μg/ml, 500 μg/ml. Compounds 1 and 3, dissolved in DMSO, were assayed at the following concentrations: 63 μM, 125 μM, 250 μM, 0.5 mM, 1
mm, 2 mm. The plate was incubated at 25 °C for 45 h and imaged using a Fotodyne lightbox and camera. Using this assay, inhibition of growth was observed in the wells where 10 μM of 10 μg/ml (21 μM) kanamycin was added, and 10 μL of 125 μM compound 1 was added. No growth inhibition was observed when using compound 3.

**MALDI Mass Spectrometric Analysis of C. diphtheriae His6-CdipGlfT2 Enzymatic Activity**—Reactions were performed in a total volume of 60 μL and contained 0.2 μM His6-CdipGlfT2, 200 μM acceptor substrate, 1 mM UDP-Galf, 50 mM HEPES (pH 7.0), 25 mM magnesium chloride (MgCl2), and 100 mM NaCl. Reactions were typically allowed to proceed for 20 h at 25 °C until they were quenched with the addition of 60 μL of a 1:1 mixture of chloroform:methanol to precipitate protein. Samples were concentrated by evaporation to dryness under vacuum using a SpeedVac SC1000 (Varian). Samples were resuspended in 50 μL of 50% acetonitrile in water. Samples were spotted at a 1:3 ratio with α-cyano-4-hydroxycinnamic acid matrix. Spectra were collected using a Bruker Ultraflex III mass spectrometer.

**Kinetic Assay of C. diphtheriae His6-CdipGlfT2 Activity**—A coupled enzymatic assay that detects UDP release was employed to measure the kinetic activity of His6-CdipGlfT2 (37, 50). Protein activity was measured in a total volume of 120 μL using a 1-cm path length quartz cuvette containing a solution of 50 mM HEPES (pH 7.0), 25 mM MgCl2, 100 mM NaCl, 300 units of pyruvate kinase (Sigma), 20 units of lactate dehydrogenase (Sigma), 250 μM reduced nicotinamide adenine dinucleotide (NADH), 500 μM phosphoenolpyruvate, and 0.2 μM His6-CdipGlfT2. The oxidation of NADH to yield NAD+ was monitored over time using a Cary 50 Bio UV-visible spectrophotometer (Varian) at 340 nm. Once a steady baseline was observed, UDP-Galf was added to a final concentration of 1.25 mM. Once the baseline again stabilized, acceptor substrate was added to the desired concentration. The rate of UDP-Galf consumption was determined from the negative slope of the linear portion of NADH absorbance using ε₃₄₀ = 6300 M⁻¹ cm⁻¹.

**Author Contributions**—M. R. L. and L. L. K. conceived the study. M. R. L. and D. A. W. performed the studies. D. AW, M. R. L., and L. L. K. analyzed the data and wrote the manuscript.

**Acknowledgments**—MALDI-TOF mass spectrometry data were obtained at the University of Wisconsin (UW)-Madison Chemistry Instrument Center Mass Spectrometry Facility on a Bruker Ultraflex III instrument that was supported by the National Institutes of Health (Center for Research Resources 1S1OR024601). Additional experiments were carried out at the LIW-Madison Biophysics Instrumentation Facility, which is supported by LIW-Madison, National Science Foundation Grant BIR-9512577 and National Institutes of Health Grant S10 RR13790. We acknowledge R. A. Splain for synthesis of the synthetic acceptor substrates used in this study and M. A. Martinez Farias for synthesis of the FP probes. We thank H. L. Hodges and A. M. Justen for helpful discussion and help preparing the manuscript.

**References**


Corynebacterial Galactan Biosynthesis


