A tethering mechanism for length control in a processive carbohydrate polymerization

John F. May[a], Rebecca A. Splain[b], Christine Brotschi[b], and Laura L. Kiessling[a,b,1]

[a]Department of Biochemistry, University of Wisconsin, 433 Babcock Drive, Madison, WI 53706; and [b]Department of Chemistry, University of Wisconsin, 1101 University Avenue, Madison, WI 53706

Edited by Jerrold Meinwald, Cornell University, Ithaca, NY, and approved May 18, 2009 (received for review February 7, 2009)

Carbohydrate polymers are the most abundant organic substances on earth. Their degrees of polymerization range from tens to thousands of units, yet polymerases generate the relevant lengths without the aid of a template. To gain insight into template-independent length control, we investigated how the mycobacterial galactofuranosyltransferase GlfT2 mediates formation of the galactan, a polymer of galactofuranose residues that is an integral part of the cell wall. We show that isolated recombinant GlfT2 can catalyze the synthesis of polymers with degrees of polymerization that are commensurate with values observed in mycobacteria, indicating that length control by GlfT2 is intrinsic. Investigations using synthetic substrates reveal that GlfT2 is processive. The data indicate that GlfT2 controls length by using a substrate tether, which is distal from the site of elongation. The strength of interaction of that tether with the polymerase influences the length of the resultant polymer. Thus, our data identify a mechanism for length control by a template-independent polymerase.

galactofuranose | mycobacteria | polymerase | polysaccharide | processivity

Carbohydrate polymers have myriad and fundamental physiological roles that depend on their length. Polysaccharides that form macromolecular structures require physical and chemical properties that arise from high degrees of polymerization. Examples of critical polymers that contain hundreds to thousands of monomeric units include cellulose (1–3), hyaluronic acid (HA) and related materials found in the extracellular matrix of eukaryotes (4–6), and polysaccharides that form protective capsules around bacteria (7). An alternative role for high-molecular-weight polysaccharides is as reservoirs of stored energy, and polymers in this class, such as starch, form semicrystalline granules that allow energy to be densely deposited (8–10). Contrastingly, polysaccharides that function in signaling tend to be shorter, presumably because signaling depends on the recognition of distinct epitopes. Indeed, short HA polymers of only 100 monomeric units induce inflammatory responses and activate parts of the immune system, whereas full-length HA strikingly mediates the opposite effects (4–6). A role for polysaccharide length in function also is observed for polysaccharide virulence determinants, such as the O-antigens. Their ability to protect Gram-negative bacteria from environmental stress and to facilitate immune evasion depends on their length, which can range from tens to hundreds of monomeric residues (7, 11–14). Last, oligosaccharides that function to connect 2 larger macromolecules are usually short. For example, oligosaccharides that link the mycolic acids to the peptidoglycan in the mycobacterial cell wall contain tens of monomeric residues (15, 16). These examples underscore that control of polysaccharide length is critical for proper biological function.

Little is known about how biosynthetic enzymes control polysaccharide length. Carbohydrate polymers are synthesized by enzyme-catalyzed chain-growth polymerization reactions in which monomer units add successively to the growing end of an acceptor. For biosynthetic polymerizations, the mechanism of length control depends on the mechanism of elongation, and elongation can occur by either a distributive or a processive mechanism. A distributive polymerase releases the elongating polymer into solution after each catalytic addition of monomer unit, whereas a processive polymerase retains the elongating polymer through multiple catalytic rounds of monomer addition (17). In a distributive polymerization, product lengths occur statistically in a Poisson distribution (18). In a processive polymerization, however, product lengths are determined by when the enzyme releases the growing polymer into solution. For template-dependent processive polymerizations, such as transcription and translation, polymer length is encoded by the template. In contrast, template-independent processive polymerizations, such as those that mediate polysaccharide synthesis, lack a template-encoded termination signal. Thus, processive carbohydrate polymerizations require other mechanisms for length control.

Our investigations into the synthesis of an essential polysaccharide in mycobacteria have yielded insight into length control for a template-independent polymerization. The mycobacterial cell wall contains a polymer of galactofuranose (Gal*f*) residues, termed the galactan, which serves as a covalent connector between the peptidoglycan and the mycolic acid–arabinan layer. The galactan is a linear polymer composed of 20–40 Gal*f* residues (19). Its formation is mediated by the essential galactofuranosyltransferase GlfT2 (the product of the *Mycobacterium tuberculosis* gene *glfT2*, also known as *Rv3808c*) (20, 21). It was postulated that GlfT2 serves as a polymerase that catalyzes multiple transfers of Gal*f* residues from the donor UDP-Gal*f* to the nonreducing end of a lipid-linked initiator oligosaccharide acceptor (Fig. 1, compound **1**) (22–26). GlfT2 had been shown to add up to 4 Gal*f* residues to synthetic acceptors (26). Still, the reported products of GlfT2-catalyzed reactions with synthetic acceptor analogs did not approach the length of the galactan. Thus, it was unknown how galactan length is regulated and whether GlfT2 alone exerts control.

We used synthetic chemistry to elucidate factors that influence galactan length. We focused on incorporating key features of the natural substrate into synthetic acceptors. The presence of a lipid in the putative endogenous acceptor suggested to us that this moiety plays a role in GlfT2 binding. To examine this issue, we synthesized compounds **2**–**6**, which contain disaccharides of Gal*f* attached to lipids of various lengths (Fig. 1 and Fig. S1). Investigations with these compounds reveal that GlfT2 can control polymer length through a tethering mechanism.

Results

Length Is Controlled by GlfT2. We first determined whether GlfT2 alone could regulate galactan length. One explanation for the inability of GlfT2 to convert known synthetic acceptors into polymeric products is that additional proteins are required (27). Alternatively, known synthetic acceptors could lack key features required for polymerization. We favored the latter explanation and postulated that the substrate lipid is important for GlfT2 activity. To test

Author contributions: J.F.M., R.A.S., C.B., and L.L.K. designed research; J.F.M. performed research; J.F.M., R.A.S., and C.B. contributed new reagents/analytic tools; J.F.M., R.A.S., C.B., and L.L.K. analyzed data; and J.F.M. and L.L.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: kiessling@chem.wisc.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0901407106/DCSupplemental.

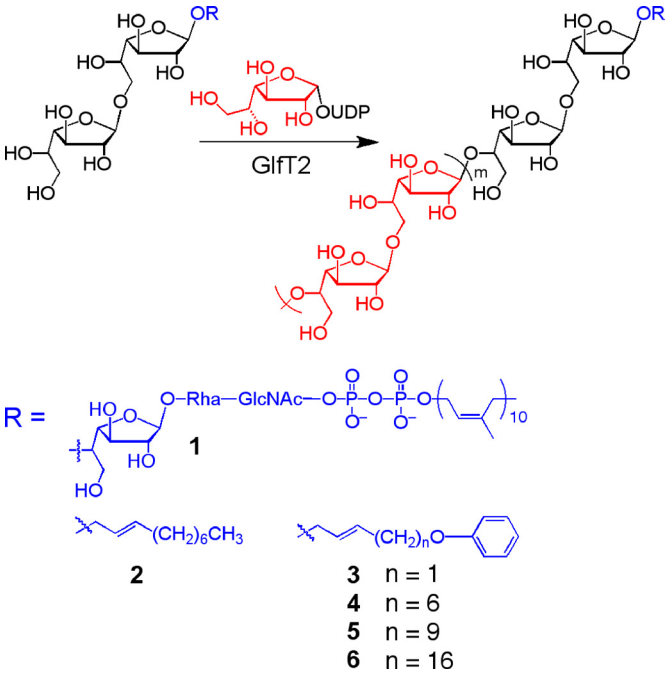

Fig. 1. Proposed role for GlfT2 in mycobacterial galactan biosynthesis. GlfT2 is hypothesized to mediate synthesis of the galactan polymer, which contains between 20 and 40 Gal*f* residues (m = 10–20) connected by alternating β-Gal*f*-(1→5)-Gal*f* and β-Gal*f*-(1→6)-Gal*f* linkages. Elongation occurs by addition of Gal*f* residues from UDP-Gal*f* to the nonreducing end of a lipid-linked acceptor (compound **1**). Synthetic acceptors **2**-**6** share features with the natural acceptor (compound **1**). Rha, rhamnose.

this hypothesis, we assessed the ability of compound **5** (Fig. 1) to serve as an acceptor substrate for GlfT2. When crude enzymatic reaction mixtures were analyzed by MALDI-TOF MS, the products possessed as many Gal*f* residues as are found in the natural galactan (19). Specifically, GlfT2 added up to 27 Gal*f* residues to compound **5** (Fig. 2 and Table S1). These results indicate that GlfT2 alone can catalyze the formation of Gal*f* polymers of the appropriate length for cell wall biosynthesis.

GlfT2 Uses a Processive Mechanism for Galactan Polymerization. We next considered how GlfT2 controls the length of polymers generated. An intriguing feature of the aforementioned reactions

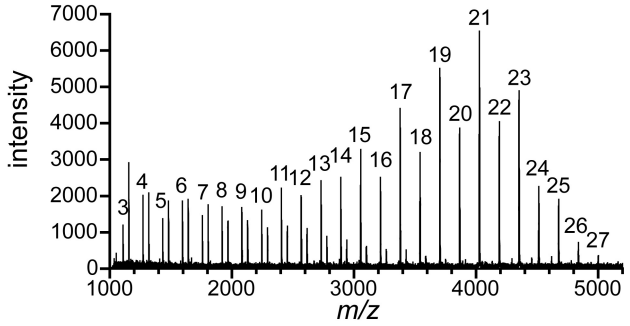

Fig. 2. GlfT2 alone can determine galactan length. The mass spectrum is shown from MALDI-TOF MS analysis of a 20-h incubation of a reaction mixture that contained His_6-GlfT2, UDP-Gal*f*, and compound **5**. Peaks that correspond to *m/z* values of $[M + Na]^+$, in which M equals the mass of compound **5** plus *n* Gal*f* residues, are labeled with the value of *n*. Unlabeled peaks have *m/z* values that correspond to the loss of the lipid moiety from elongation products, presumably because of fragmentation. Reaction products are observed from n = 3 to n = 27.

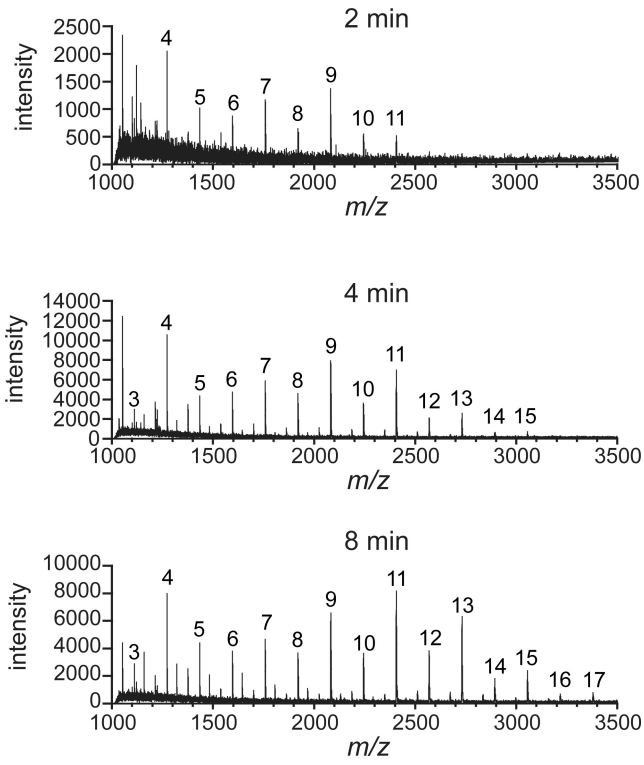

Fig. 3. GlfT2 uses a processive mechanism for polymerization. Mass spectra are shown from MALDI-TOF MS analysis of time points of a reaction that contained His_6-GlfT2, UDP-Gal*f*, and compound **5**. Peaks that correspond to *m/z* values of $[M + Na]^+$, in which M equals the mass of compound **5** plus *n* Gal*f* residues, are labeled with *n*. At 2 min (≈60 turnovers), products are observed from n = 4–11. At 4 min (≈175 turnovers), products are observed from n = 3–15. At 8 min (≈440 turnovers), products are observed from n = 3–17. For all time points, we observed a peak of high intensity corresponding to the mass of unelongated acceptor (n = 0).

suggested a possible mechanism. Specifically, the product distribution was skewed toward the larger polymers (Fig. 2), yet a significant peak corresponding to unelongated compound **5** was observed (Fig. S2*A*). These observations together suggest that, rather than releasing product after the addition of 1 monomer unit, GlfT2 retains the acceptor through multiple cycles of monomer addition. If this analysis is correct, length control by GlfT2 should depend on a processive, not distributive, elongation mechanism.

Many polymerases are processive enzymes: they bind their substrates and perform multiple rounds of catalysis before dissociation (17, 28). To test directly for GlfT2 processivity, we adapted an assay that has been used to evaluate the processivity of nucleotide polymerases (18). Specifically, reactions were conducted under conditions that minimize multiple binding events between the acceptor and the polymerase. In reactions that contain acceptor **5** in 1,000-fold molar excess of GlfT2, a product containing multiple Gal*f* residues would be generated only if GlfT2 retains the precursor for multiple cycles of elongation. Accordingly, a finding that compounds containing multiple Gal*f* residues appear at early time points would indicate that GlfT2 is processive. Conversely, the exclusive appearance of compounds with only 1 or 2 Gal*f* residues would indicate that GlfT2 is distributive. We found that even at short reaction times (2 min), products corresponding to the addition of multiple Gal*f* residues were present (Fig. 3). These data indicate that GlfT2 uses a processive mechanism for polymerization.

A Tethering Model for Length Control. Given that GlfT2 is a processive polymerase, its ability to control polymer length is intriguing. Distributive polymerization reactions afford products with

a statistical distribution of lengths; long lengths are generated only after long reaction times or after many cycles of substrate complexation, monomer unit addition, and product release. In contrast, product lengths that result from processive polymerizations depend on when the polymerase releases the bound polymer. That an enzyme can carry out a processive polymerization reaction yet give rise to length control seems paradoxical: to achieve length control, a processive polymerase must prevent polymer dissociation at shorter lengths yet release the polymer at the appropriate length. For template-dependent polymerases, both requirements for processive length control are met by high-affinity binding of the growing polymer and template to the polymerase. These polymerases, including DNA and RNA polymerases, possess an extended binding site that interacts with repeating units of the growing polymer (17). This subsite-binding model has been generally proposed for processive enzymatic reactions (29), including template-independent polymerizations by members of the glycosyltransferase family (30–33). Although tight binding of repeating saccharide units could prevent polysaccharide dissociation, it is unclear how, in the absence of a template, subsite binding would allow for length control.

To probe this mechanistic question for GlfT2, we tested whether it uses the carbohydrate subsite-binding mechanism for processivity. Reactions of GlfT2 with compound **2** were analyzed. We designed this putative acceptor to have an alkyl lipid similar to those used in known GlfT2 acceptors that, unlike **5**, are not fully elongated (Fig. 1) (25, 26). If GlfT2 used the carbohydrate subsite-binding mechanism for processivity, both compounds **5** and **2** should be polymerized processively. Both acceptors contain identical saccharides, and the addition of Gal*f* residues to either acceptor would provide a growing polysaccharide with identical chemical properties for interaction with carbohydrate-binding subsites of GlfT2 (Fig. 1). In our comparison of acceptors, compound **2** afforded products elongated by, at most, 4 Gal*f* residues (Fig. S2*B*). These results are similar to those obtained by others (26, 27), but they contrast with data obtained by using **5**. The latter findings suggest that the lipid substituent, and not solely the saccharide, is a key determinant of processivity.

To examine further this difference in acceptors, we compared the kinetics of elongation for compound **5** versus **2**. Because GlfT2 produces UDP after addition of each Gal*f* residue to an acceptor, GlfT2 turnover was monitored with a continuous assay in which UDP production was coupled to NADH oxidation (34–36). Both compounds **2** and **5** were similarly competent for catalysis, with steady-state k_{cat} values of 1.04 ± 0.03 s^{-1} (compound **5**) and 0.8 ± 0.1 s^{-1} (compound **2**), but the GlfT2-catalyzed reaction of compound **5** reached half-maximal steady-state velocity at a 27-fold lower concentration of compound **5** ($K_{1/2}$ value of 66 ± 2 μM) than with compound **2** ($K_{1/2}$ value of 1,800 ± 700 μM; Fig. S3). Taken together, the kinetic and mass spectrometry data indicate that GlfT2 can elongate both compounds **5** and **2**, but it functions as a processive polymerase only with compound **5**. The sole structural difference between compounds **5** and **2** is the lipid group distal to the end that undergoes elongation, indicating that the lipid substituent is crucial for processivity.

We further investigated the role of the lipid substituent by examining GlfT2 activity with compounds **3** and **4**, which are analogous to compound **5** yet bear a shorter lipid. We reasoned that this series could provide more insight into whether the length of lipid is important for polymer formation. These studies reveal that although compound **4** gives rise to polymers (Fig. S4), the disaccharide **3**, which possesses the shortest lipid, does not (Fig. S2*C*). These studies indicate that only substrates with lipids of sufficient length can afford polymeric products. One potential explanation for these results is that they are rooted in the abilities of the different substrates to form micelles under the assay conditions. We tested this possibility, and the results indicate that the ability of a given substrate to give rise to polymeric products is not related to its ability to form micelles (Fig. S4). These data reveal the importance of lipid length on processive polymerization.

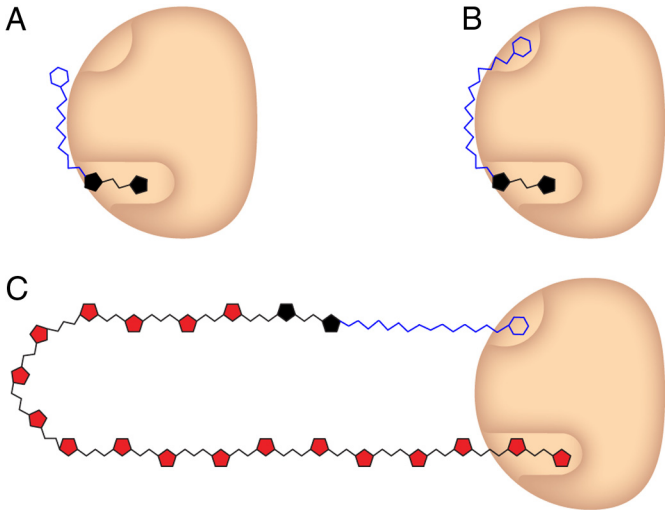

Fig. 4. Proposed tethering model for GlfT2 (tan). (*A*) With acceptors that cannot occupy the lipid-anchoring site (e.g., compounds **2** or **3**), only short oligomers are generated. (*B*) When GlfT2 interacts with acceptors that can bind to both sites (e.g., compound **5**), longer oligomers are obtained. (*C*) When the polymer chain becomes long, polymer dissociation competes with further elongation.

The finding that the substrate lipid substituent is an important determinant of processive elongation led us to postulate that it serves as a tether (Fig. 4). Thus, the acceptor substrate occupies not only the active site but also a lipid-binding secondary site. Because lipid binding would prolong the lifetime of the enzyme–acceptor complex, this bivalent binding mode should facilitate processive polymerization. To test the feasibility of the tethering mechanism, reaction conditions were varied to distinguish between monovalent and multivalent binding. The tethering hypothesis involves multipoint binding, and interactions of this type are kinetically labile in the presence of unbound ligand (37, 38). Specifically, if a tethered ligand dissociates from 1 subsite, free ligand can compete for this unoccupied subsite (39). In contrast, free ligand concentration does not affect the rate of dissociation of a ligand bound at a single binding site. Thus, if GlfT2 binds its substrate through tethering, higher concentrations of unelongated acceptor in solution will increase the rate of complete dissociation, and thereby promote polymer termination; therefore, an increase in the population of short polymers should be observed at higher acceptor concentrations. When reactions were conducted with lower initial concentrations of acceptor **5**, at both early (Fig. S5) and late (20 h; Fig. 5) time points, longer polymers were obtained (Fig. 5). These results indicate that compound **5** engages in multipoint binding.

The tethering model gives rise to an inherent mechanism for length control. As the length of the polymer increases, dissociation will compete more effectively with elongation (39, 40), and the processive polymerization will terminate (Fig. 4). Accordingly, in the presence of a compound that can compete for one of the subsites and thereby disrupt tethering, shorter polymers should be produced. Because the endogenous GlfT2 acceptor contains a polyprenyl lipid, a lipid such as geranylgeranyl pyrophosphate (GGPP) should compete with the substrate for GlfT2 binding and thereby inhibit polymer formation. Indeed, in enzymatic reactions of **5**, the presence of GGPP resulted in shorter products (Fig. S6). Another prediction of our model is that a tether with augmented affinity will afford longer polymers. We therefore synthesized compound **6**, which bears a longer lipid than compound **5** (Fig. 1). We predicted that compound **6** will bind with increased affinity to the secondary site on GlfT2. Thus, if elongation occurs through

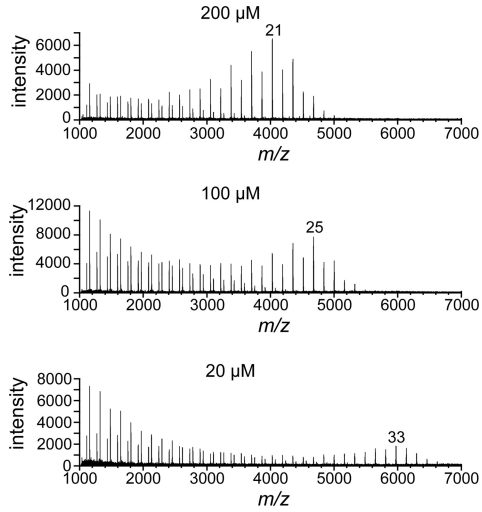

Fig. 5. Termination of polymerization depends on bivalent interactions of the substrate. Mass spectra from MALDI-TOF MS analysis of 20-h reactions with lower initial concentrations of compound **5** showed products with a greater degree of polymerization than those from reactions with higher initial substrate concentrations. In each spectrum, the peak that had the highest intensity of the group of high-molecular-weight peaks is labeled with *n*, which corresponds to $[M + Na]^+$, where M equals the mass of compound **5** plus *n* Gal*f* residues. The value of *n* increased with decreasing initial concentrations of compound **5**.

tethering, the product distribution obtained from compound **6** versus compound **5** should be shifted to higher-molecular weight polymers. Indeed, compound **6** was elongated to give products with a mass distribution that is higher than that obtained for compound **5** (Fig. 6, Fig. S7, and Table S2). These data provide additional evidence for a role for the lipid in tethering and highlight how the tether can influence polymer length.

Discussion

Tethering Facilitates Both Processive Polymerization and Length Control by GlfT2. Our results reveal how the template-independent carbohydrate polymerase GlfT2 controls polysaccharide length. GlfT2 carries out a processive polymerization that yields polysaccharides with degrees of polymerization (DPs) in accord with those of the mycobacterial galactan (Figs. 2 and 3). The data indicate it achieves this control through tethering. Processive polymerization and length control can result when GlfT2 interacts with both acceptor ends (Fig. 4). Processive complexes of GlfT2 and acceptor are labile in the presence of unelongated acceptor (Fig. 5), and stronger tethering by the lipid affords longer polymers (Fig. 6). These results are consistent with bivalent binding of acceptor to GlfT2.

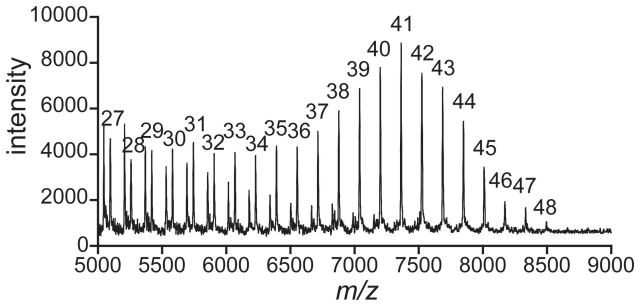

Fig. 6. Improved tethering leads to increased polymer lengths. The mass spectrum, from *m/z* = 5,000 to *m/z* = 9,000, is shown from a MALDI-TOF MS analysis of a 20-h incubation of a reaction mixture that contained His_6-GlfT2, UDP-Gal*f*, and compound **6**. Peaks that correspond to *m/z* values of $[M + Na]^+$, in which M equals the mass of compound **6** plus *n* Gal*f* residues, are labeled with the value of *n*. Unlabeled peaks have *m/z* values that correspond to the loss of the lipid moiety from elongation products, presumably because of fragmentation. Products are observed up to *n* = 48.

By using bivalent binding to enable both processivity and length control, tethering represents a solution to the paradox of how a template-independent polymerase prevents premature termination, yet can release polymers when they are an appropriate length. At initiation, tethering facilitates processive polymerization by prolonging the lifetime of the enzyme–acceptor complex. As polymerization proceeds, tethering influences length by affecting how readily the polymer dissociates. During processive polymerizations, after each addition of monomer unit, polymerization continues because dissociation occurs more slowly than further elongation. When the rate of dissociation exceeds the rate of elongation, however, polymerization terminates (18, 41). For a tethered substrate, the addition of residues to the growing polymer is predicted to increase its conformational entropy, which in turn increases its rate of dissociation.

GlfT2 Exhibits a Degree of Length Control Similar to Other Carbohydrate Polymerases. Although GlfT2 uses a processive mechanism, it does not produce polymers of a single length (Figs. 2 and 3). Rather, it produces a range of polymers that are distributed around an expected length, an observation that is consistent with data indicating galactan polysaccharides isolated from mycobacteria are polydisperse (19, 23). These findings highlight an important difference between template-dependent and template-independent polymerizations. In general, the products of template-dependent polymerizations must be a single specified length to function properly in biological information transfer. Hence, these processes require mechanisms for strict length control. In contrast, the products of template-independent polymerizations, such as those that give rise to polysaccharides, can function over a range of lengths. Therefore, template-independent carbohydrate polymerizations would not require mechanisms for exact control of length to yield functional polysaccharides.

In accord with this prediction, many biologically active polysaccharides are polydisperse. For example, the isolation of bacterial lipopolysaccharide O-antigen carbohydrates affords materials of different sizes (7, 14, 42, 43). The size range is modally distributed around a particular length, and differences in degrees of polymerization correspond to the repeating monomer units of the O-antigen (14, 42). Similarly, a distribution of lengths is observed from analysis of products of peptidoglycan glycosyltransferase reactions (44, 45). Last, large polysaccharides, such as cellulose and HA, are polydisperse (3–5). These examples indicate that the level of length control exhibited by GlfT2 is typical of carbohydrate polymerases.

The polydispersity observed in biological polysaccharides suggests that carbohydrate polymerases generally do not terminate polymerization at a particular threshold length. This product distribution is consistent with that expected from the tethering mechanism. Such reactions generate products with a range of lengths, and the mass distribution is determined by the strength of the interaction of the tether with the polymerase—stronger tethering shifts the range to higher masses. Tethering is a general process that may used by a variety of carbohydrate polymerases that exhibit intrinsic length control. For example, it has been proposed that the lengths of bacterial capsular polysaccharides are dictated by loss of affinity of a glycosyltransferase for a polymer beyond certain lengths (7), especially in ABC-transporter or synthase-dependent pathways (7, 14). It is unclear how other proposed mechanisms for length control (7, 14, 33, 45–47) could lead to the product DPs generated. Polymerases that produce larger polysaccharides, such as HA, may also use tethering for length control. Different isozymes of human hyaluronan synthases inherently produce different size distributions of HA in vitro. It has been hypothesized that increased movement within longer chains could cause polymer release and terminate HA polymerization (5), and tethering could provide a

mechanism for such length-dependent affinity. Thus, tethering may serve as a general strategy for length control in the biosynthesis of a variety of functional polysaccharides.

Polysaccharide-Binding Domains in Other Carbohydrate Polymerases Could Facilitate Processivity Through Tethering. Many biologically important polysaccharides are thought to be synthesized in a processive manner, but it is unknown whether the corresponding glycosyltransferases use tethering. Consistent with the paradigmatic mechanism for processivity, it has been proposed that carbohydrate polymerases bind the repeating saccharide units in the growing polymer by using a glycan-binding domain (30–32). This model resulted from a comparison of sequences of glycosyltransferases within the GT-2 family of inverting β-glycosyltransferases (www.cazy.org). The GT-2 family encompasses nonpolymerizing glycosyltransferases as well as polymerizing glycosyltransferases, such as cellulose synthase, hyaluronan synthase, and GlfT2. Although polymerizing transferases share some conserved sequence features with nonpolymerizing transferases, they also contain other conserved sequence motifs that are unique (30). The latter sequences have been predicted to form a glycan-binding region in the active site to facilitate processivity through tight binding of the nascent glycan chain (48).

Evidence for a glycan-binding domain has been found in a GT-2 family glycosyltransferase that catalyzes the formation of a streptococcal capsular polysaccharide (33). This enzyme requires an octasaccharide primer for processive synthesis, and it was hypothesized that 8 residues are needed for tight association with the enzyme (33). In another example, the formation of polysialic acid by polysialyltransferases requires a domain of basic amino acid residues, and this domain was found only in the sialyltransferases that have polymerase activity (49). These 2 examples support the hypothesis that glycan binding serves as a processivity determinant. Thus, like the lipid in the GlfT2 acceptors, an initiator oligosaccharide sequence could serve as a tether in the biosynthesis of a variety of polysaccharides.

Implications for Mycobacterial Galactan Biosynthesis. The finding that the acceptor lipid is important for GlfT2 processivity provides insight into the site of galactan biosynthesis within mycobacteria. This pathway is thought to be localized to the cytoplasmic face of the membrane because lipid-linked galactan intermediates are found in membrane fractions, and the donor, UDP-Gal*f*, is assumed to be located in the cytoplasm (15). Previous studies showed that GlfT2 activity fractionates with mycobacterial membranes (22, 25), but the role of this association is unclear. Our results that indicate that GlfT2 requires the lipid portion of the acceptor for processive polymerization provide a functional context for the association of GlfT2 with the membrane: interaction of GlfT2 with the membrane would enable GlfT2 to access both ends of the acceptor for polymerization.

Prospects for Inhibiting Mycobacterial Galactan Biosynthesis. The tethering model has ramifications for developing new treatments of diseases caused by mycobacteria, including tuberculosis. Inhibitors of mycobacterial galactan biosynthesis may serve as valuable therapeutic leads for tuberculosis and drug-resistant tuberculosis (21, 50). Because the gene encoding GlfT2 is essential for mycobacterial viability (20, 21), inhibition of GlfT2 activity represents an untapped strategy for preventing mycobacterial growth. The tethering model predicts that inhibitors that target both the lipid-binding and catalytic sites of GlfT2 would be highly potent. Therefore, strategies for fragment-based lead discovery (37) may yield effective inhibitors of GlfT2, and therefore guide the development of new antimycobacterial agents.

Materials and Methods

Synthesis of Acceptor Substrates. Detailed methods and characterization for the synthesis of compounds **2**–**6** can be found in the *SI Methods*.

Engineering of His_6-GlfT2. The gene *glfT2* was amplified from the genomic DNA of *M. tuberculosis* H37Rv (provided through the National Institutes of Health, National Institute of Allergy and Infectious Diseases Contract N01 AI-75320) and cloned into the pET-24a expression vector (Novagen). The His_6 sequence for purification was added to the N terminus of *glfT2* by using QuikChange mutagenesis (Stratagene). Detailed methods, including primer sequences, can be found in the *SI Methods*.

Production of His_6-GlfT2. The construct pET-24a:his_6glfT2 was transformed into Tuner (DE3) *Escherichia coli* cells (Novagen) by electroporation and plated on Luria broth (LB) agar containing 50 μg/mL kanamycin. A starter culture (50 mL of LB with 50 μg/mL kanamycin) was inoculated with a single-colony transformant and incubated overnight at 37 °C. Larger-volume cultures (1 L of LB with 50 μg/mL kanamycin) were inoculated with the starter culture (4 mL) and incubated at 37 °C with shaking until OD at 600 nm exceeded 0.80. The cultures were then cooled in ice water for 1 h, then isopropyl β-thiogalactopyranoside (IPTG) was added to 0.3 mM final concentration, and the cultures were incubated for 18 h at 15 °C. Cultures were harvested by centrifugation (5,000 rpm in a JLA-8.1000 rotor; Beckman–Coulter), and the cell pellets were frozen at −80 °C until further use. Cells were lysed by thawing on ice and resuspending in HisTrap loading buffer (50 mM Hepes, 25 mM imidazole, and 500 mM NaCl, pH 7.4). Protease inhibitor mixture III (Calbiochem), Triton X-100 [0.1% (vol/vol)], and lysozyme were added to the lysis mixture. Cell lysate was then sonicated (Branson Sonifier 450, tip setting no. 7, constant duty cycle, cycles of 10 s on and 2 min 50 s off) until viscosity decreased. Cell debris was removed by centrifugation (22,000 × *g*, 1 h, 4 °C). Soluble lysate was filtered through a 0.44-μm nylon filter (Millipore) and applied at 1.0 mL/min to a 5-mL HisTrap column (GE Healthcare) that had been equilibrated with HisTrap loading buffer on an AKTA FPLC system (GE Healthcare). The column was washed until the UV absorbance at 280 nm reached the baseline level. Then, a linear gradient from 0% to 100% HisTrap elution buffer (50 mM Hepes, 500 mM NaCl, and 500 mM imidazole, pH 7.4) was applied to the column over 20 column volumes. Fractions containing His_6-GlfT2 were identified by SDS/PAGE and were sufficiently pure for enzymatic assays. Glycerol was added to 10% (vol/vol) final concentration to His_6-GlfT2 fractions, and the glycerol stocks were flash-frozen in liquid N_2 and stored at −80 °C until further use.

***Galf* Transferase Activity Assay and Analysis by MALDI-TOF MS.** A sample of His_6-GlfT was dialyzed twice into 2 L of 50 mM Hepes (pH 7.4), 100 mM NaCl, and 5 mM EDTA by using 10,000 molecular weight cut-off dialysis cassettes (Pierce Biotechnology). Protein was assayed by using the BCA assay (Pierce Biotechnology) with BSA as a standard. Reaction mixtures of 120-μL total volume contained final concentrations of 0.2 μM H_6GlfT2, 200 μM acceptor, 1.25 mM UDP-Gal*f* in 50 mM Hepes, 25 mM $MgCl_2$, and 100 mM NaCl, pH 7.0 (buffer was added from 10× concentrated stock solution). Reactions were incubated at room temperature for a specified time, and then 120 μL of 1:1 MeOH:$CHCl_3$ was added to quench the reaction. Quenched reaction mixtures were evaporated to dryness in a SpeedVac SC100 (Varian) under vacuum. Dried reaction mixtures were resuspended in 50 μL of a saturated solution of α-cyano-4-hydroxycinnamic acid in 50% acetonitrile and 0.1% trifluoroacetic acid, and 2 μL of the resuspended reaction solution was spotted onto a stainless steel target (Applied Biosystems) for analysis by MALDI-TOF MS in either linear or reflectron mode (using a Voyager DE Pro; Applied Biosystems). Mass spectra were calibrated by using external standards (angiotensin II and bovine insulin). For the time course assays, larger-volume, single-pot reactions were set up by scaling each component volumetrically. At each time point, 120 μL was withdrawn from the reaction and added to a separate tube that contained the quenching solution. After quenching, the time points were analyzed as described above in this section. Reactions to test the dependence of maximal extent of polymerization on acceptor concentration were set up with various concentrations of acceptor and were allowed to incubate for 20 h at room temperature. These reactions were quenched and analyzed as described above. Compounds **3** and **6** were dissolved as stock solutions in 100% MeOH. Upon dilution, reactions containing compounds **3** or **6** as an acceptor included MeOH [final concentrations were up to 5% (vol/vol)]; these amounts of MeOH did not influence product length in control reactions of GlfT2 with compound **5**. For reaction mixtures containing GGPP (Sigma), GGPP was added to a desired final concentration, and equivalent volumes of MeOH (up to 4 μL) were included in the 120-μL reaction solution. After quenching, the appro-

priate volume of GGPP was added to reactions that had no GGPP to control for any effects of GGPP on the ionization efficiency of the products.

Coupled Enzyme Assay to Measure UDP Production by His_6-GlfT2. In a quartz cuvette, the following were mixed to give a total volume of 120 μL: buffer (50 mM Hepes, 25 mM $MgCl_2$, and 100 mM NaCl (pH 7.0); diluted from a 10× stock): 300 units of pyruvate kinase (Sigma), 20 units of lactate dehydrogenase (Sigma), 250 μM NADH, 500 μM phospho*enol*pyruvate, and 0.2 μM His_6-GlfT2. Absorbance at 340 nm was monitored over time in a Cary 50 Bio UV-Visible Spectrophotometer (Varian) until a steady baseline was reached (usually 2 min), then UDP-Gal*f* was added to 1.25 mM. Absorbance at 340 nm was again monitored, then acceptor substrate was added to the desired concentration. Absorbance at 340 nm was monitored over time. The steady-state rate was calculated from the slope of the linear portion of the decrease in absorbance over time by using $\varepsilon = 6{,}300\ M^{-1}{\cdot}cm^{-1}$ for NADH (34).

ACKNOWLEDGMENTS. We thank H. A. Steinberg for assistance in generating Fig. 4. This research was supported by National Institutes of Health Grant AI063596. J.F.M. was supported by a National Science Foundation Graduate Research Fellowship and by the Molecular Biosciences Training Grant GM007215 at University of Wisconsin, Madison. R.A.S. was supported by an American Chemical Society Division of Medicinal Chemistry Graduate Fellowship. C.B. was supported by the Swiss National Science Foundation Grant PBBE-108553. MALDI-TOF MS data were obtained at the University of Wisconsin, Madison Biophysics Instrumentation Facility, which is supported by the University of Wisconsin, Madison; National Science Foundation Grant BIR-9512577; and National Institutes of Health Grant S10 RR13790. *M. tuberculosis* H37Rv genomic DNA was provided through the National Institutes of Health, National Institute of Allergy and Infectious Diseases Contract N01 AI-75320, titled, ''Tuberculosis Research Materials and Vaccine Testing Contract,'' under the administration of John Belisle, Colorado State University, Fort Collins, CO.

1. Ross P, Mayer R, Benziman M (1991) Cellulose biosynthesis and function in bacteria. *Microbiol Rev* 55:35–58.
2. Bureau TE, Brown RM, Jr (1987) In vitro synthesis of cellulose II from a cytoplasmic membrane fraction of *Acetobacter xylinum*. *Proc Natl Acad Sci USA* 84:6985–6989.
3. Somerville C (2006) Cellulose synthesis in higher plants. *Annu Rev Cell Dev Biol* 22:53–78.
4. Stern R (2003) Devising a pathway for hyaluronan catabolism: Are we there yet? *Glycobiology* 13:105R–115R.
5. Weigel PH, DeAngelis PL (2007) Hyaluronan synthases: A decade-plus of novel glycosyltransferases. *J Biol Chem* 282:36777–36781.
6. Lee JY, Spicer AP (2000) Hyaluronan: A multifunctional, megaDalton, stealth molecule. *Curr Opin Cell Biol* 12:581–586.
7. Whitfield C (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu Rev Biochem* 75:39–68.
8. Stubbe J, et al. (2005) Nontemplate-dependent polymerization processes: Polyhydroxyalkanoate synthases as a paradigm. *Annu Rev Biochem* 74:433–480.
9. Ball SG, Morell MK (2003) From bacterial glycogen to starch: Understanding the biogenesis of the plant starch granule. *Annu Rev Plant Biol* 54:207–233.
10. Smith AM (2001) The biosynthesis of starch granules. *Biomacromolecules* 2:335–341.
11. Nigou J, et al. (2008) Mannan chain length controls lipoglycans signaling via and binding to TLR2. *J Immunol* 180:6696–6702.
12. Murray GL, Attridge SR, Morona R (2006) Altering the length of the lipopolysaccharide O antigen has an impact on the interaction of *Salmonella enterica* serovar Typhimurium with macrophages and complement. *J Bacteriol* 188:2735–2739.
13. Burns SM, Hull SI (1998) Comparison of loss of serum resistance by defined lipopolysaccharide mutants and an acapsular mutant of uropathogenic *Escherichia coli* O75:K5. *Infect Immun* 66:4244–4253.
14. Raetz CR, Whitfield C (2002) Lipopolysaccharide endotoxins. *Annu Rev Biochem* 71:635–700.
15. Berg S, Kaur D, Jackson M, Brennan PJ (2007) The glycosyltransferases of *Mycobacterium tuberculosis*—roles in the synthesis of arabinogalactan, lipoarabinomannan, and other glycoconjugates. *Glycobiology* 17:35R–56R.
16. Brennan PJ, Nikaido H (1995) The envelope of mycobacteria. *Annu Rev Biochem* 64:29–63.
17. Breyer WA, Matthews BW (2001) A structural basis for processivity. *Protein Sci* 10:1699–1711.
18. McClure WR, Chow Y (1980) The kinetics and processivity of nucleic acid polymerases. *Methods Enzymol* 64:277–297.
19. Besra GS, et al. (1995) A new interpretation of the structure of the mycolyl-arabinogalactan complex of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosylalditol fragments by fast-atom bombardment mass spectrometry and ^{1}H nuclear magnetic resonance spectroscopy. *Biochemistry* 34:4257–4266.
20. Sassetti CM, Boyd DH, Rubin EJ (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* 48:77–84.
21. Pan F, Jackson M, Ma Y, McNeil MR (2001) Cell wall core galactofuran synthesis is essential for growth of mycobacteria. *J Bacteriol* 183:3991–3998.
22. Mikusová K, et al. (2000) Biosynthesis of the galactan component of the mycobacterial cell wall. *J Biol Chem* 275:33890–33897.
23. Belánová M, et al. (2008) Galactosyl transferases in mycobacterial cell wall synthesis. *J Bacteriol* 190:1141–1145.
24. Mikusová K, et al. (2006) Identification of a novel galactosyl transferase involved in biosynthesis of the mycobacterial cell wall. *J Bacteriol* 188:6592–6598.
25. Kremer L, et al. (2001) Galactan biosynthesis in *Mycobacterium tuberculosis*. Identification of a bifunctional UDP-galactofuranosyltransferase. *J Biol Chem* 276:26430–26440.
26. Rose NL, et al. (2006) Expression, purification, and characterization of a galactofuranosyltransferase involved in *Mycobacterium tuberculosis* arabinogalactan biosynthesis. *J Am Chem Soc* 128:6721–6729.
27. Pathak AK, et al. (2001) Studies on (beta,1 -> 5) and (beta,1 -> 6) linked octyl Gal(f) disaccharides as substrates for mycobacterial galactosyltransferase activity. *Bioorg Med Chem* 9:3129–3143.
28. Von Hippel PH, Fairfield FR, Dolejsi MK (1994) On the processivity of polymerases. *Ann N Y Acad Sci* 726:118–131.
29. delCardayré SB, Raines RT (1994) Structural determinants of enzymatic processivity. *Biochemistry* 33:6031–6037.
30. Saxena IM, Brown RM, Fevre M, Geremia RA, Henrissat B (1995) Multidomain architecture of beta-glycosyl transferases: implications for mechanism of action. *J Bacteriol* 177:1419–1424.
31. Keenleyside WJ, Whitfield C (1996) A novel pathway for O-polysaccharide biosynthesis in *Salmonella enterica* Serovar Borreze. *J Biol Chem* 271:28581–28592.
32. Koyama M, Helbert W, Imai T, Sugiyama J, Henrissat B (1997) Parallel-up structure evidences the molecular directionality during biosynthesis of bacterial cellulose. *Proc Natl Acad Sci USA* 94:9091–9095.
33. Forsee WT, Cartee RT, Yother J (2006) Role of the carbohydrate binding site of the *Streptococcus pneumoniae* capsular polysaccharide type 3 synthase in the transition from oligosaccharide to polysaccharide synthesis. *J Biol Chem* 281:6283–6289.
34. Gosselin S, Alhussaini M, Streiff MB, Takabayashi K, Palcic MM (1994) A continuous spectrophotometric assay for glycosyltransferases. *Anal Biochem* 220:92–97.
35. Chen L, et al. (2002) Intrinsic lipid preferences and kinetic mechanism of *Escherichia coli* MurG. *Biochemistry* 41:6824–6833.
36. Rose NL, et al. (2008) Development of a coupled spectrophotometric assay for GlfT2, a bifunctional mycobacterial galactofuranosyltransferase. *Carbohydr Res* 343:2130–2139.
37. Erlanson DA, Wells JA, Braisted AC (2004) Tethering: Fragment-based drug discovery. *Annu Rev Biophys Biomol Struct* 33:199–223.
38. Rao J, Lahiri J, Isaacs L, Weis RM, Whitesides GM (1998) A trivalent system from vancomycin·D-Ala-D-Ala with higher affinity than avidin·biotin. *Science* 280:708–711.
39. Kramer RH, Karpen JW (1998) Spanning binding sites on allosteric proteins with polymer-linked ligand dimers. *Nature* 395:710–713.
40. Flory PJ (1953) *Principles of Polymer Chemistry* (Cornell Univ Press, Ithaca, NY).
41. Fersht A (1985) *Enzyme Structure and Mechanism* (Freeman, New York).
42. Tocilj A, et al. (2008) Bacterial polysaccharide co-polymerases share a common framework for control of polymer length. *Nat Struct Mol Biol* 15:130–138.
43. Goldman RC, Hunt F (1990) Mechanism of O-antigen distribution in lipopolysaccharide. *J Bacteriol* 172:5352–5359.
44. Barrett D, et al. (2007) Analysis of glycan polymers produced by peptidoglycan glycosyltransferases. *J Biol Chem* 282:31964–31971.
45. Wang TS, Manning SA, Walker S, Kahne D (2008) Isolated peptidoglycan glycosyltransferases from different organisms produce different glycan chain lengths. *J Am Chem Soc* 130:14068–14069.
46. Whitfield C, Larue K (2008) Stop and go: Regulation of chain length in the biosynthesis of bacterial polysaccharides. *Nat Struct Mol Biol* 15:121–123.
47. Ventura CL, Cartee RT, Forsee WT, Yother J (2006) Control of capsular polysaccharide chain length by UDP-sugar substrate concentrations in *Streptococcus pneumoniae*. *Mol Microbiol* 61:723–733.
48. Saxena IM, Brown RM, Dandekar T (2001) Structure–function characterization of cellulose synthase: Relationship to other glycosyltransferases. *Phytochemistry* 57:1135–1148.
49. Nakata D, Zhang L, Troy FA (2006) Molecular basis for polysialylation: A novel polybasic polysialyltransferase domain (PSTD) of 32 amino acids unique to the alpha 2,8-polysialyltransferases is essential for polysialylation. *Glycoconj J* 23:423–436.
50. Dykhuizen EC, May JF, Tongpenyai A, Kiessling LL (2008) Inhibitors of UDP-galactopyranose mutase thwart mycobacterial growth. *J Am Chem Soc* 130:6706–6707.