J|A|C|S
ARTICLES
Published on Web 05/03/2006

Expression, Purification, and Characterization of a Galactofuranosyltransferase Involved in *Mycobacterium tuberculosis* Arabinogalactan Biosynthesis

Natisha L. Rose,[†] Gladys C. Completo,[†] Shuang-Jun Lin,[†] Michael McNeil,[‡] Monica M. Palcic,*[,§] and Todd L. Lowary*[,†]

Contributions from the Department of Chemistry and Alberta Ingenuity Centre for Carbohydrate Science, Gunning-Lemieux Chemistry Centre, University of Alberta, Edmonton, AB T6G 2G2, Canada, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Valby, Denmark, Department of Microbiology, Colorado State University, Fort Collins, Colorado 80523-1677

Received December 5, 2005; E-mail: monica@crc.dk; tlowary@ualberta.ca

Abstract: The major structural component of the cell wall of *Mycobacterium tuberculosis* is a lipidated polysaccharide, the mycoyl-arabinogalactan-peptidoglycan (mAGP) complex. This glycoconjugate plays a key role in the survival of the organism, and thus, enzymes involved in its biosynthesis have attracted attention as sites for drug action. At the core of the mAGP is a galactan composed of D-galactofuranose residues attached via alternating β-(1→5) and β-(1→6) linkages. A single enzyme, *glfT*, has been shown to synthesize both glycosidic linkages. We report here the first high-level expression and purification of *glfT* by expression of the *Rv3808c* gene in *Escherichia coli* C41(DE3). Following a three-step purification procedure, 3−7 mg of protein of >95% purity was isolated from each liter of culture. We subsequently probed the substrate specificity of *glfT* by evaluating a panel of potential mono- and oligosaccharide substrates and demonstrated, for the first time, that trisaccharides are better substrates than disaccharides and that one disaccharide, in which the terminal D-galactofuranose residue is replaced with an L-arabinofuranose moiety, is a weak substrate. Kinetic characterization of the enzyme using four of the oligosaccharide acceptors gave K_m values ranging from 204 μM to 1.7 mM. Through the use of NMR spectroscopy and mass spectrometry, we demonstrated that this recombinant enzyme, like the wild-type protein, is bifunctional and can synthesize both β-(1→6) and β-(1→5)-linkages in an alternating fashion. Access to purified *glfT* is expected to facilitate the development of high-throughput assays for the identification of inhibitors of the enzyme, which are potential antituberculosis agents.

Introduction

Tuberculosis (TB) is the world's most lethal bacterial disease, killing nearly 3 million people each year.[1−3] The impact of this disease on world health has been the source of increased recent interest due to the emergence of multi-drug resistant strains of *Mycobacterium tuberculosis*, the organism that causes the disease.[4,5] Of additional concern is the susceptibility of immunocompromised patients to mycobacterial infections, including *M. tuberculosis* and *M. avium*.[6] Mycobacterial diseases are difficult to treat, which is due in part to the particularly impermeable nature of the mycobacterial cell wall. Successful treatment of these infections typically requires drug regimens consisting of multiple antibiotics lasting six months or longer.[7]

The major structural component of the mycobacterial cell wall, and the feature that gives the organism its impermeability, is a lipidated polysaccharide, the mycoyl-arabinogalactan-peptidoglycan (mAGP) complex, a structure of fascinating complexity (Figure 1).[8,9] Connected to cell wall peptidoglycan via an α-L-Rha*p*-(1→3)-α-D-Glc*p*NAc disaccharide, the polysaccharide portion of the mAGP is built upon a galactan core that contains approximately 30 D-galactofuranose (Gal*f*) residues attached via alternating β-(1→5) and β-(1→6) linkages. Along this galactan are branch points to which are attached arabinan chains made up of roughly 70 D-arabinofuranose residues, linked α-(1→5), α-(1→3), or β-(1→2). Recent work on the structure of the mAGP from *Corynebacterium glutamicum* has clearly shown that three arabinan chains are linked to the Gal*f*-containing backbone.[10] It is therefore likely that a similar number of arabinan domains decorate mycobacterial galactan. The nonreducing termini of these arabinan chains are esterified with mycolic acids, $C_{70}-C_{90}$ branched lipids. Through substantial

[†] University of Alberta.
[‡] Colorado State University.
[§] Carlsberg Laboratory.

(1) Paolo, W. F., Jr.; Nosanchuk, J. D. *Lancet Infect. Dis.* **2004**, *4*, 287−293.
(2) Davies, P. D. O. *Ann. Med.* **2003**, *35*, 235−243.
(3) Coker, R. J. *Trop. Med. Int. Health* **2004**, *9*, 25−40.
(4) Nachega, J. B.; Chaisson, R. E. *Clin. Infect. Dis.* **2003**, *36*, S24−S30.
(5) Wade, M. M.; Zhang, Y. *Front. Biosci.* **2004**, *9*, 975−994.
(6) De Jong, B. C.; Israelski, D. M.; Corbett, E. L.; Small, P. M. *Annu. Rev. Med.* **2004**, *55*, 283−301.
(7) Bass, J. B., Jr.; Farer, L. S.; Hopewell, P. C.; Obrien, R.; Jacobs, R. F.; Ruben, F.; Snider, D. E.; Thornton, G. *Am. J. Respir. Crit. Care Med.* **1994**, *149*, 1359−1374.
(8) Brennan, P. J. *Tuberculosis* **2003**, *83*, 91−97.
(9) Lowary, T. L. In *Glycoscience: Chemistry and Chemical Biology*; Springer-Verlag: Berlin, 2001; pp 2005−2080.
(10) Alderwick, L. J.; Radmacher, E.; Seidel, M.; Gande, R.; Hitchen, P. G.; Morris, H. R.; Dell, A.; Sahm, H.; Eggeling, L.; Besra. G. S. *J. Biol. Chem.* **2005**, *280*, 32362−32371.

10.1021/ja058254d CCC: $33.50

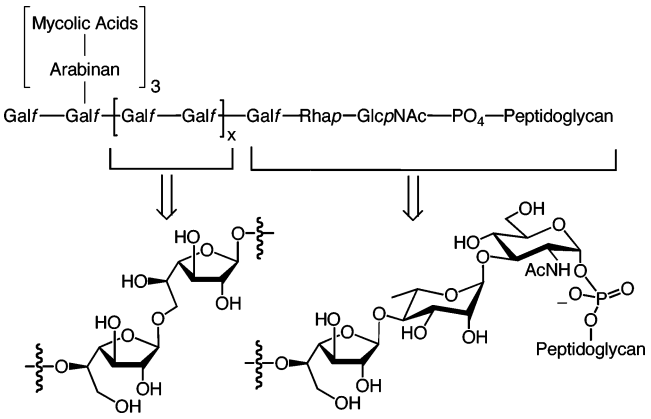

Figure 1. Schematic depiction of the mycobacterial mAGP complex highlighting the structure of the galactan and linker region. $x \approx 13-15$.

van der Waals interactions between these tightly packed lipids, the bacteria presents to its surroundings a facade of extremely low permeability that is sometimes nearly crystalline[11] and that provides the organism with great protection from its environment.

Given the importance of the mAGP to the survival of the organism, it is not surprising that enzymes involved in its biosynthesis have attracted significant attention as sites for drug action.[12−14] Indeed, two of the standard antibiotics used in the treatment of TB target mAGP assembly. Isoniazid blocks mycolic acid biosynthesis,[15] whereas ethambutol inhibits arabinosyltransferases involved in the biosynthesis of the arabinan portions of the polysaccharide.[16] Although none of the antituberculosis agents in use are known to target galactan assembly, its production is essential for the viability of the organism.[17]

A general pathway for mAGP biosynthesis has been proposed and supported by subsequent investigation (Figure 2).[18−20] The process is believed to proceed via the sequential addition of sugar residues to a polyprenol-bound intermediate that is transferred to peptidoglycan before the introduction of the mycolic acids. Although the general features of this process are understood, relatively little is known about the glycosyltransferases that are involved. For example, although the genes thought to encode for one or more of the arabinosyltransferases have been identified,[16,21] the expression of these enzymes in soluble form has, to date, not been reported.

A galactosyltransferase involved in galactan biosynthesis (*glfT*) is the best characterized of these glycosyltransferases. Analysis of the genome from *M. tuberculosis* H37Rv led to the

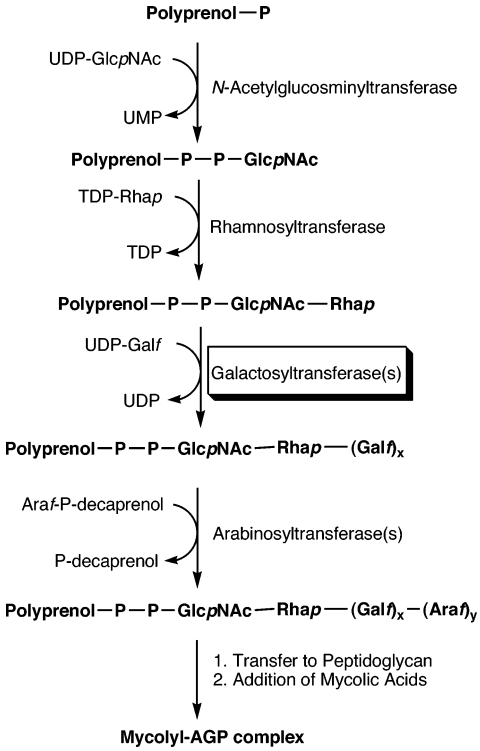

Figure 2. General features of mAGP biosynthesis. The galactosyltransferase of interest to this paper is highlighted. $x \approx 30$; $y \approx 70$.

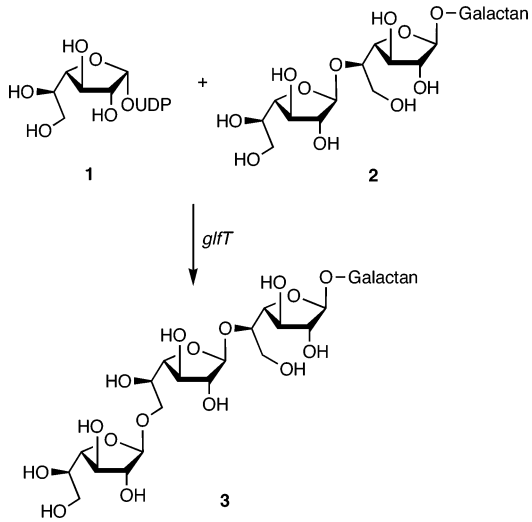

Figure 3. Representative reaction catalyzed by *glfT*.

identification of the gene (*Rv3808c*) encoding for *glfT*,[22] a 68 kD protein that transfers Gal*f* residues from UDP-galactofuranose (UDP-Gal*f*, **1**, Figure 3) to the growing galactan chain (e.g., **2**) yielding an elongated product (e.g., **3**). In the genome, *Rv3808c* overlaps, by four nucleotide bases with *Rv3809c*, the gene encoding for UDP-Gal*p* mutase, the enzyme that converts UDP-Gal*p* into UDP-Gal*f*. Subsequent overexpression of *glfT* in both *M. smegmatis* and *E. coli* led to membrane preparations that have been used to develop an assay for its activity[23] and to screen potential acceptors[24] and inhibitors.[25,26] Interestingly,

(11) Nikaido, H.; Kim, S.-H.; Rosenberg, E. Y. *Mol. Microbiol.* **1993**, *8*, 1025−1030.
(12) Zhang, Y. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 529−564.
(13) Sharma, K.; Chopra, P.; Singh, Y. *Expert Opin. Ther. Targets* **2004**, *8*, 79−93.
(14) Kremer, L.; Besra, G. S. *Expert Opin. Inv. Drugs* **2002**, *11*, 153−157.
(15) Schroeder, E. K.; De Souza, N.; Santos, D. S.; Blanchard, J. S.; Basso, L. A. *Curr. Pharm. Biotech.* **2002**, *3*, 197−225.
(16) Belanger, A. E.; Besra, G. S.; Ford, M. E.; Mikusová, K.; Belisle, J. T.; Brennan, P. J.; Inamine, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 11919−11924.
(17) Pan, F.; Jackson, M.; Ma, Y.; McNeil. M. *J. Bacteriol.* **2001**, *183*, 3991−3998.
(18) Crick, D. C.; Mahaptra, S.; Brennan, P. J. *Glycobiology* **2001**, *11*, 107R-118R.
(19) Yagi, T.; Mahapatra, S.; Mikusová, K.; Crick, D. C.; Brennan, P. J. *J. Biol. Chem.* **2003**, *278*, 26497−26504.
(20) Berg, S.; Starbuck, H.; Torrelles, J. B.; Vissa, V. D.; Crick, D. C.; Chatterjee, D.; Brennan, P. J. *J. Biol. Chem.* **2005**, *280*, 5651−5663.
(21) Escuyer, V. E.; Lety, M.-A.; Torrelles, J. B.; Khoo, K.-H.; Tang, J.-B.; Rithner, C. D.; Frehel, C.; McNeil, M. R.; Brennan, P. J.; Chatterjee, D. *J. Biol. Chem.* **2001**, *276*, 48854−48862.
(22) Mikusová, K.; Yagi, T.; Stern, R.; McNeil, M. R.; Besra, G. S.; Crick, D. C.; Brennan, P. J. *J. Biol. Chem.* **2000**, *275*, 33890−33897.
(23) Kremer, L.; Dover, L. G.; Morehouse, C.; Hitchin, P.; Everett, M.; Morris, H. R.; Dell, A.; Brennan, P. J.; McNeil, M. R.; Flaherty, C.; Duncan, K.; Besra. G. S. *J. Biol. Chem.* **2001**, *276*, 26430−26440.

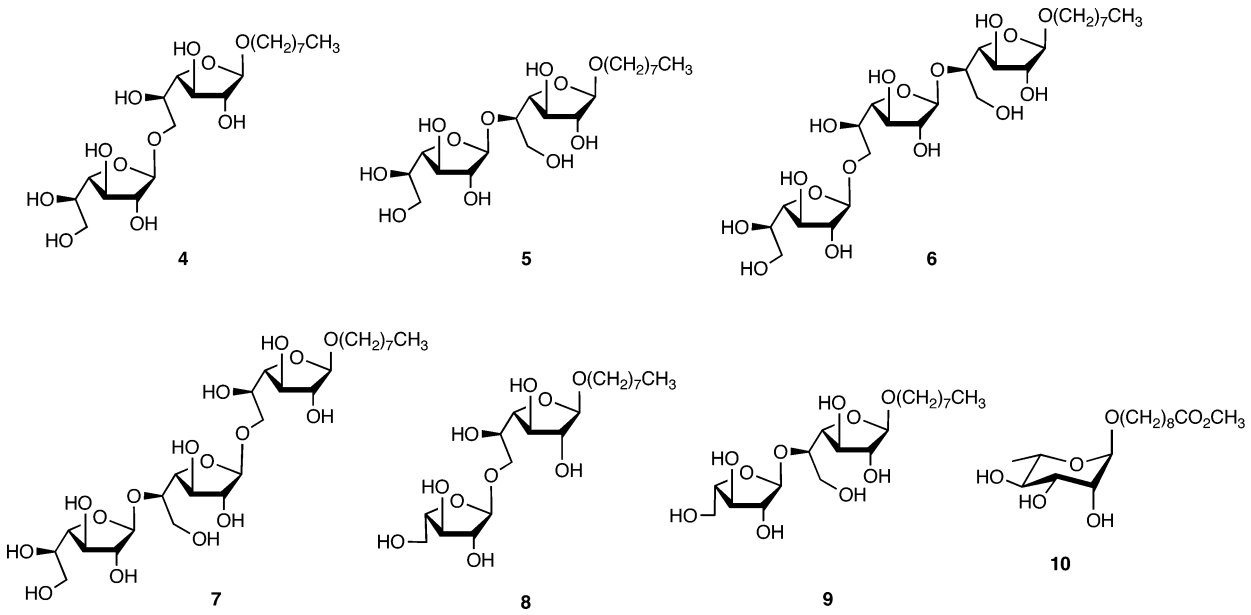

Figure 4. Synthetic compounds screened as potential substrates for *glfT*.

evaluation of two synthetic disaccharide acceptors in this assay demonstrated[23] that the enzyme is bifunctional and can synthesize both β-(1→5) and β-(1→6) Gal*f* linkages. This enzyme, therefore, appears to be one of an increasing number of glycosyltransferases[27−36] that "violate" the "one-enzyme one-linkage" hypothesis that was put forward nearly 40 years ago[37] and that has been a tenet of glycobiology.

As part of a larger program directed toward the identification of inhibitors of mycobacterial glycosyltransferases,[38−40] we report here the first high-level expression and purification of recombinant mycobacterial *glfT*. In addition, the substrate specificity of this enzyme has also been probed by evaluation of a panel of potential acceptor substrates (**4**−**10**, Figure 4). These studies extend previous investigations on the substrate specificity of this enzyme, which have focused entirely on disaccharides based on β-D-Gal*f*-(1→5)-β-D-Gal*f*, β-D-Gal*f*-(1→6)-β-D-Gal*f*, or β-D-Gal*f*-(1→4)-α-L-Rha*p* scaffolds.[23,24,41]

Experimental Section

Construction of *glfT* Expression Vector. To construct a vector allowing for the expression of a soluble, full-length N-terminal His-tagged *glfT* enzyme, pET-15b/*Rv3808c* was cloned using pViet/*Rv3808c* as a template. The *glfT* gene was PCR-amplified from pViet/*Rv3808c* using the forward primer ONR68 (5′-ATA ATT *CAT ATG* AGT GAA CTC GCC GCG AGC CT-3′) to introduce an *Nde*I site at the 5′ end and the reverse primer ONR69 (5′-ATA T*GG ATC C*TC AGC CAT GCT CGG GCT CTT-3′) to introduce a *Bam*HI site at the 3′ end of the *Rv3808c* gene. The italicized bases represent the *Nde*I and *Bam*HI sites, respectively. The resulting 1933 bp PCR product and the pET-15b vector (Novagen) were digested first by *Nde*I and after by *Bam*HI. The digested fragments were extracted from an agarose gel using a QIAquick Gel Extraction Kit (Qiagen) and ligated to produce pET-15b/*Rv3808c*. The resulting plasmid was transformed into *E. coli* DH5α cells (Invitrogen). Selected transformants were screened for the presence of the recombinant plasmid by restriction mapping and DNA sequence analysis. pET-15b/*Rv3808c* was subsequently transformed into a wide range of compatible host strains, including *E. coli* BL21 (Novagen), BL21-Gold (Stratagene), BL21(DE3) (Novagen), BL21(DE3)pLysS (Novagen), and C41(DE3) (Avidis). *glfT* was expressed from all host strains and screened for the overproduction of soluble protein by SDS−PAGE and Western Blot analysis. SDS−PAGE analysis was performed on NuPage 12% Bis−Tris acrylamide gels (Invitrogen) under conditions recommended by the manufacturer. Blotting to nitrocellulose was performed overnight at 30 V. The membrane was probed with *anti*-His6-peroxidase (Roche), and His-tagged protein was detected by chemiluminescence using SuperSignal West Femto Maximum Sensitivity Substrate (Pierce). The recombinant protein was expressed and purified from *E. coli* C41(DE3), the host strain demonstrating the greatest level of soluble protein expression.

Expression and Purification of *glfT*. *E. coli* C41(DE3) transformed with pET-15b/*Rv3808c* was cultured overnight at 30 °C in 100 mL of TB broth supplemented with 1 × TB salts, 1% glucose, and 100 μg/mL ampicillin. The overnight culture (40−50 mL) was transferred to 1 L of fresh medium and cultured further to an $A_{600\ nm}$ of 0.4−0.8. The culture was then induced for 16−18 h at 25 °C with 1 mM isopropyl β-D-thiogalactopyranoside. Cells were harvested by centrifugation, washed with phosphate buffered saline, and stored at −20 °C until further use. The cell pellet (~15 g wet wt) was resuspended in 50 mM sodium phosphate buffer, pH 8.5, containing 5 mM β-mercaptoethanol, 1 mM EDTA, and 0.1% Triton X-100 supplemented with Complete Protease Inhibitor Cocktail Tablets (Roche) and disrupted by passage

(24) Pathak, A. K.; Pathak, V.; Seitz, L.; Maddry, J. A.; Gurcha, S.; Besra, G. S.; Suling. W. J.; Reynolds, R. C. *Bioorg. Med. Chem.* **2001**, *9*, 3129−3143.
(25) Cren, S.; Gurcha, S. S.; Blake, A. J.; Besra, G. S.; Thomas, N. R. *Org. Biomol. Chem.* **2004**, *2*, 2418−2420.
(26) Cren, S.; Wilson, C.; Thomas, N. R. *Org. Lett.* **2005**, *7*, 3521−3523.
(27) Guan, S.; Clarke, A. J.; Whitfield, C. *J. Bacteriol.* **2001**, *183*, 3318−3327.
(28) Llull, D.; García, E.; López, R. *J. Biol. Chem.* **2001**, *276*, 21053-21061.
(29) Jing, W.; DeAngelis, P. L. *Glycobiology* **2003**, *13*, 661−671.
(30) Jing, W.; DeAngelis, P. L. *Glycobiology* **2000**, *10*, 883−889.
(31) Löbau, S.; Mamat, U.; Brabetz, W.; Brade, H. *Mol. Microbiol.* **1995**, *18*, 391−399.
(32) Toivonen, S.; Aitio, O.; Renkonen, O. *J. Biol. Chem.* **2001**, *276*, 37141−37148.
(33) Wakarchuk, W. W.; Watson, D.; St. Michael, F.; Li, J.; Wu, Y.; Brisson, J.-R.; Young, N. M.; Gilbert, M. *J. Biol. Chem.* **2001**, *276*, 12785−12790.
(34) Van der Wel, H.; Fisher, S. Z.; West, C. M. *J. Biol. Chem.* **2002**, *277*, 46527−46534.
(35) DeAngelis, P. L. *Glycobiology* **2002**, *12*, 9R−16R.
(36) Yada, T.; Gotoh, M.; Sato, T.; Shionyu, M.; Go, M.; Kaseyama, H.; Iwasaki, H.; Kikuchi, N.; Kwon, Y.-D.; Togayachi, A.; Kudo, T.; Watanabe, H.; Narimatsu, H.; Kimata, K. *J. Biol. Chem.* **2003**, *278*, 30235−30247.
(37) Hagopian, A.; Eylar, E. H. *Arch. Biochem. Biophys.* **1968**, *128*, 422−433.
(38) Cociorva, O. M.; Gurcha, S. S.; Besra, G. S.; Lowary, T. L. *Bioorg. Med. Chem.* **2005**, *13*, 1369−1379.
(39) Centrone, C. A.; Lowary, T. L. *J. Org. Chem.* **2002**, *67*, 8862−8870.
(40) Centrone, C. A.; Lowary, T. L. *Bioorg. Med. Chem.* **2004**, *12*, 5495−5503.
(41) Wen, X.; Crick, D. C.; Brennan, P. J.; Hultin, P. G. *Bioorg. Med. Chem.* **2003**, *11*, 3579−3587.

through a French Pressure cell (Thermo Spectronic) at 20 000 psi. The cell lysate was centrifuged at 10 5 000 × *g* for 60 min at 4 °C. The *glfT* enzyme was purified from the resulting supernatant by successive column chromatographies using Q-Sepharose FF (Amersham Biosciences), Ni−NTA Superflow (Qiagen), and Sephacryl S-100-HR (Amersham Biosciences) columns. The buffers used in this purification were as follows: Buffer A, 50 mM sodium phosphate, pH 8.5, containing 5 mM β-mercaptoethanol; Buffer B, 50 mM sodium phosphate, pH 8.0, containing 5 mM β-mercaptoethanol, 500 mM NaCl, and 20 mM imidazole; Buffer C, 50 mM MOPS, pH 7.9, containing 5 mM β-mercaptoethanol and 500 mM NaCl; and Buffer D, 50 mM MOPS, pH 7.9, containing 5 mM β-mercaptoethanol and 100 mM NaCl. The supernatant was applied to a Q-Sepharose FF column that was equilibrated with Buffer A. The column was washed extensively with Buffer A until $A_{280\ nm}$ returned to background levels and eluted with a linear gradient of Buffer A containing 0−1.0 M NaCl. The fractions containing *glfT* were pooled, made up to 20 mM imidazole and 500 mM NaCl, and applied directly to a Ni−NTA Superflow column that had been equilibrated with Buffer B. After washing the column with Buffer B until $A_{280\ nm}$ returned to background levels, the protein was eluted with a linear gradient of Buffer B containing 20−250 mM imidazole. The fractions containing *glfT* were pooled and concentrated to ∼1 mL using an Amicon stirred cell with a YM-30 membrane (Millipore). The concentrated enzyme fraction was applied to a Sephacryl S-100-HR column equilibrated with Buffer C and eluted with the same buffer. Fractions containing *glfT* were pooled and dialyzed extensively against Buffer D before use in activity assays or enzymatic synthesis. Protein concentrations were determined using the Bio-Rad Protein Assay using bovine γ globulin as a standard. Fractions were screened for the presence of *glfT* by SDS−PAGE and Western Blot analysis.

Synthesis of Acceptor Substrates. The compounds screened as potential substrates for the enzyme were synthesized as described elsewhere (**4**−**9**,[42] **10**[43]).

Radiochemical Activity Assay. The *glfT* enzyme activity was determined using a modified assay as previously described by Kremer et al.[23] Synthetic acceptors **4**−**10**, stored as 50 mM methanol stocks, were aliquoted and dried in a vacuum desiccator to remove any residual solvent. The acceptors were resuspended with the remaining constituents of the assay. The reaction mixtures for assessing the incorporation of [^{3}H]Gal consisted of 50 mM MOPS, pH 7.9, 5 mM β-mercaptoethanol, 10 mM $MgCl_2$, 62.5 μM ATP, 10 mM NADH, 1 mM UDP-Gal*p*, 10 μL purified UDP-Gal*p* mutase (3.8 μg/μL), 10 μL purified *glfT* (1.5−2.5 μg/μL), and uridine diphosphate galactose, [galactose-6-^{3}H] (American Radiolabeled Chemicals, Inc., 20 Ci/mmol, 0.5 μL) in a total reaction volume of 40 μL. Purified UDP-Gal*p* mutase was obtained as previously described by Sanders et al.[44] The equilibrium ratio for UDP-Gal*p* mutase is 92:8, thus leading to an approximately 80 μM initial concentration of UDP-Gal*f* in the incubation mixture. For determination of apparent acceptor kinetics, these conditions are assumed to be saturating for the donor. Specific activity was determined for all acceptors at a concentration of 1 mM. The reaction mixtures were incubated at 37 °C for 60 min. Control assays without acceptor or UDP-Gal*p* mutase or with an empty vector were also performed to correct for the presence of endogenous acceptor and the hydrolysis of UDP-[^{3}H]galactose. All assays were normalized to the levels of UDP-Gal*p* mutase activity (∼10% equilibrium). Radiolabeled products were isolated by reverse-phase chromatography on SepPak C_{18} cartridges (Waters). After loading the samples onto the cartridges, they were washed with water (50 mL) to remove unreacted donor and eluted with methanol (3.5 mL). The reaction products in the eluants were quantified by liquid scintillation counting on a Beckman LS6500 Scintillation Counter using 10 mL of Ecolite cocktail.

For the kinetic analysis of *glfT*, SepPak radiochemical C_{18} assays, as described previously,[45] were used with the hydrophobic acceptors **4**−**7**. Kinetic values were determined by varying acceptor concentrations with the donor concentration fixed at 1 mM. Initial rate conditions were linear with no more than 15% of the substrate consumed. The K_m values were calculated within the concentration ranges of 24−3125 μM (**4**), 48−6250 μM (**5**), 4−1000 μM (**6**), and 7−1000 μM (**7**). The kinetic parameters K_m and k_{cat} ($V_{max}/[E]$) were obtained by nonlinear regression analysis of the Michaelis−Menten equation with the GraphPad PRISM 4.00 program (GraphPad Software, San Diego, CA). The catalytic constant or turnover number, k_{cat}, is the maximum number of substrate molecules converted to product per active site per unit time. The k_{cat}/K_m ratio was determined to compare substrate specificity. One milliunit of enzyme activity is defined as the amount of enzyme that catalyzes the conversion of 1 nmol of sugar transferred per minute.

Preparation, Isolation, and Characterization of Reaction Products. To determine if the enzyme was transferring Gal*f* residues to the acceptor substrates via alternating β-(1→5) or β-(1→6) linkages, preparative scale reactions were performed. A reaction containing 50 mM MOPS, pH 7.9, containing 5 mM β-mercaptoethanol, 10 mM $MgCl_2$, 62.5 μM ATP, 10 mM NADH, 100 mM UDP-Gal*p*, 2 units alkaline phosphatase, 80 μL purified UDP-Gal*p* mutase, 80 μL of purified *glfT*, and 5 mM acceptor in a total of 320 μL was incubated at ambient temperature with gentle rotation for 4 days.

The progress of enzymatic reactions was followed by thin-layer chromatography (TLC) on SIL G-25 silica gel plates (Macherey-Nagel) using $CHCl_3/CH_3OH/NH_4OH/H_2O$ (65:25:0.5:3.6) as the eluant.[23] Visualization of compounds was achieved by heating the plates after dipping them in a solution of 3% anisaldehyde in sulfuric acid. At the end of the reaction, separation of the product and unreacted acceptor from other components was achieved by application of the mixture to a Sep-Pak C_{18} reverse phase cartridge. After extensive washing with water to remove buffer, donor, and enzyme, the product was eluted with HPLC-grade methanol (8 mL). The solvent was evaporated to give a residue that was redissolved in water (1 mL) and passed through a 0.22 μm Millex-GV filter. The filtrate was lyophilized, and the resulting residue was characterized by MALDI mass spectrometry on a Voyager Elite time-of-flight spectrometer on samples suspended in 2,5-dihydroxy benzoic acid, using the delayed-extraction mode and positive-ion detection.

For characterization of products by ^{1}H NMR spectroscopy, the reaction products were purified by preparative TLC. The area of the TLC plate corresponding to where the product eluted was scraped from the uncharred plate and dissolved in HPLC-grade methanol. The solvent was evaporated, redissolved in water, and the resulting solution was applied to a Sep-Pak C_{18} reverse phase cartridge. After extensive washing of the cartridge with water, the product was eluted with HPLC-grade methanol (8 mL), and the solvent was evaporated. The residue was dissolved in H_2O (1 mL) and then filtered through a 0.22 μm Millex-GV filter and lyophilized. Subsequent lyophilization of the residue from D_2O (2 × 1 mL) gave a product that was then dissolved in D_2O (∼600 μL). One-dimensional ^{1}H NMR spectra were recorded on a Varian i600 instrument; the HOD signal was suppressed by presaturation.

Results and Discussion

Mycobacterial viability requires galactan biosynthesis,[17] and thus, *glfT* plays an essential role in the life cycle of these organisms. Due to the absence of Gal*f*-containing glyconjugates in mammals, this enzyme is a particularly attractive target for drug development.[46] Previously, using crude membrane and cell

(42) Completo, G. C.; Lowary, T. L. In preparation.
(43) Josephson, S.; Bundle, D. R. *J. Chem. Soc., Perkin Trans.* **1980**, *1* 297−301.
(44) Sanders, D. A. R.; Staines, A. G.; McMahon, S. A.; McNeil, M. R.; Whitfield, C.; Naismith, J. H. *Nat. Struct. Biol.* **2001**, *8*, 858−863.
(45) Palcic, M. M.; Heerze, L. D.; Pierce, M.; Hindsgaul, O. *Glycoconjugate J.* **1988**, *5*, 49−63.

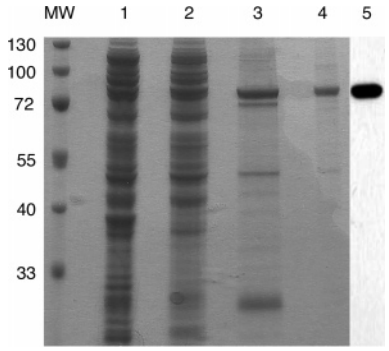

Figure 5. SDS−PAGE and western blot analysis of *glfT*. Fractions of *glfT* after cell lysis (Lane 1), anion-exchange (Lane 2), affinity (Lane 3), and gel filtration (Lane 4) were run on 12% Bis−Tris acrylamide gels and visualized by Coomassie Blue Staining. Purified *glfT* was electrophoresed and immunoblotted with *anti*-His6-peroxidase (Lane 5).

wall fractions as a source of both *glfT* and UDP-Gal*p* mutase, Besra and co-workers demonstrated that *glfT* is a bifunctional enzyme catalyzing the formation of both β-(1→5) and β-(1→6) Gal*f* linkages in mycobacterial galactan.[23,24] Despite these advances, several questions remain unanswered, including potential differences in specificity of the acceptors, differences in the mechanism of transfer, and the potential for *glfT* to possess activity to transfer multiple Gal*f* moieties to a growing oligosaccharide chain. In addition, the identification of inhibitors of the enzyme would be facilitated by the development of a high-throughput assay for activity, ideally using pure enzyme. To begin to address these issues, we expressed and purified a soluble form of the *Rv3808c* gene in *E. coli* and evaluated a range of potential acceptors to measure *glfT* activity using a two-step radiochemical assay.

Expression and Purification of *glfT*. In our hands, initial attempts to use pViet/*Rv3808c* as an expression vector in *E. coli* failed to produce soluble recombinant protein of the desired size. Therefore, to express full length, soluble recombinant *glfT*, the gene sequence encoding *Rv3808c*, was inserted into the expression vector pET-15b between the *Nde*I and *Bam*HI sites and transformed into a wide range of *E. coli* strains. The protein was expressed in each host and further analyzed for the presence of soluble *glfT* protein using SDS−PAGE and Western Blot techniques. Cells were grown at 25 °C to minimize inclusion body formation without significantly decreasing the expression of soluble protein. Compared to the other compatible cell systems, *E. coli* C41(DE3) cells enhanced total expression levels of the enzyme, with an increased secretion of protein but without toxic effects. C41(DE3) mutant cells are increasingly being used for the expression of toxic or difficult globular or membrane proteins.[47]

The enzyme g*lfT* was purified to >95% homogeneity from *E. coli* cell extracts by three chromatographic steps consisting of anion exchange, affinity, and gel filtration columns. SDS−PAGE and Western Blot analysis of the fractions pooled from each successive column illustrated levels of soluble *glfT* expression consistent with the expected size of the N-terminal tagged protein at ~73.5 kDa (Figure 5). After gel filtration chromatography, the final yield of *glfT* ranged from 3 to 7 mg from a 1 L culture. Interestingly, several proteins bound with high affinity to the Ni−NTA Superflow column. Bands at ~28, 50, and 72 kDa may be the result of nonspecific binding of unrelated proteins, degradation of the full length *glfT* protein, or the presence of multiple internal translation sites. If other portions of the DNA are being expressed, it remains uncertain why they exhibit a strong affinity for the Ni−NTA resin, unless they are tightly associated with the full-length, His-tagged recombinant *glfT* protein. Investigations are being performed in our laboratory to determine if the amino acid sequence of these proteins correspond to the sequence of *glfT* and to determine whether they possess galactofuranosyltransferase activity, either alone or in synergy with the full-length protein.

Activity and Specificity of *glfT*. Kremer et al. described a simple two-step radiochemical assay that measured the incorporation of [^{14}C]Gal*f* into β-(1→5) and β-(1→6) linked galactofuranose disaccharides.[23] The assay uses crude mycobacterial membrane and cell wall fractions as sources of *glfT* and UDP-Gal*p* mutase, which converts UDP-Gal*p* to UDP-Gal*f*. Thus, incubation of a potential acceptor and UDP-Gal*p* with this membrane fraction leads to in situ generation of UDP-Gal*f* via the mutase activity and its subsequent transfer to the substrate by *glfT*. Based on this assay system, we developed a radiochemical assay to measure galactosyltransferase activity using UDP-[^{3}H]Gal*p* as the glycosyl donor and purified UDP-Gal*p* mutase[44] and *glfT* as the enzyme sources.

The substrate specificity of *glfT* was probed by comparing the specific transferase activity with a panel of synthetic acceptor substrates (**4**−**10**, Figure 4). Disaccharides **4** and **5** were previously shown[24] to be substrates for *glfT* present in a mycobacterial membrane preparation, and we thus screened them against the purified recombinant protein described here to compare kinetic parameters. In addition, the trisaccharide substrates **6** and **7** were screened to determine the effect of acceptor length on catalysis. Prior to this study, no trisaccharide substrates for this enzyme have been evaluated. The disaccharides containing a terminal L-arabinofuranosyl (Ara*f*) residue, **8** and **9**, were evaluated to determine if the enzyme recognized substrates lacking the terminal hydroxymethyl group. Finally, we tested rhamnoside **10** as a substrate to determine if this bifunctional enzyme can also synthesize the β-D-Gal*f*-(1→4)-L-Rha*p* glycosidic bond that attaches the galactan to the core of the mAGP (Figure 1).

As presented in Table 1, this recombinant *glfT* enzyme, like the ones previously characterized,[23,24] is bifunctional and can synthesize both β-(1→5) and β-(1→6) Gal*f* linkages. The trisaccharides **6** and **7** are the most efficiently galactosylated by *glfT*. All four of these substrates were sufficiently active so that full kinetic characterization could be carried out (see below). The remaining three compounds (**8**−**10**) were substantially poorer *glfT* substrates.

With regard to the Ara*f*-containing disaccharides, the (1→6)-linked isomer **8** is a substrate for *glfT* but exhibits ~50% lower specific activity than either of the Gal*f*-Gal*f* disaccharides **4** or

(46) Pedersen, L. L.; Turco, S. J. *Cell. Mol. Life Sci.* **2003**, *60*, 259−266.
(47) Miroux, B.; Walker, J. E. *J. Mol. Biol.* **1996**, *260*, 289−298.
(48) Cyr, N.; Perlin, A. S. *Can. J. Chem.* **1979**, *57*, 2504−2511.
(49) Gandolfi-Donadío, L.; Gallo-Rodriguez, C.; de Lederkremer, R. M. *J. Org. Chem.* **2003**, *68*, 6928−6934.
(50) Daffe, M.; Brennan, P. J.; McNeil, M. *J. Biol. Chem.* **1990**, *265*, 6734−6743.
(51) McNeil, M.; Daffe, M.; Brennan, P. J. *J. Biol. Chem.* **1990**, *265*, 18200−18206.
(52) Mills, J. A.; Motichka, K.; Jucker, M.; Wu, H. P.; Uhlik, B. C.; Stern, R. J.; Scherman, M. S.; Vissa, V. D.; Pan, F.; Kundu, M.; Ma, Y. F.; McNeil, M. *J. Biol. Chem.* **2004**, *279*, 43540−43546.
(53) Weigel, P. H. *IUBMB Life* **2002**, *54*, 201−211.

Table 1. Acceptor Specificity of Recombinant *glfT*

acceptor substrate	specific activity[a] (mU/mg)	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m ($\mu M \cdot min)^{-1}$
4, (1(6)-di	0.03	635 ± 126	0.0039	6.14e-6
5, (1(5)-di	0.03	1700 ± 385	0.0016	9.41e-7
6, (1(6),(1(5)-tri	0.06	204 ± 34	0.0032	1.57e-5
7, (1(5),(1(6)-tri	0.09	208 ± 50	0.014	6.73e-5
8, (1(6)-Ara-di	0.02	n.d.[b]	n.d.	n.d.
9, (1(5)-Ara-di	0.00	n.d.	n.d.	n.d.
10, Rha-Gr	0.00	n.d.	n.d.	n.d.

[a] Specific activities measured using 1 mM acceptor. [b] n.d., not determined.

5. The other Ara*f*-containing disaccharide, **9**, is inactive. These results are not unexpected given the processive nature of the enzyme. The parent structure for disaccharide **8** is **4**, and the reduced activity of the Ara*f*-containing analogue implies that the hydroxymethyl group in **8** is important, but not critical, for catalysis. For both **4** and **8**, the sugar residues are attached (1→6), and the Gal*f* moiety added by *glfT* would be attached to the 5 position of the "nonreducing" terminal monosaccharide. As both **4** and **8** have a hydroxyl group at the 5 position, both are substrates, albeit with differing levels of activity. In contrast, the parent structure for **9** is the (1→5)-linked disaccharide **5**, which is a substrate for the β-(1→6)-transferase activity of *glfT*. As the nonreducing residue of disaccharide **9** lacks C-6, and therefore a C-6 hydroxyl group, it is not surprising that this compound is not a substrate.

Screening of rhamnopyranoside **10** against the enzyme indicated that it is also inactive, thus suggesting that another mycobacterial enzyme attaches the first Gal*f* residue to the rhamnopyranoside moiety of the mAGP. However, it is also possible that *glfT* requires a larger substrate to efficiently synthesize the β-D-Gal*f*-(1→4)-α-L-Rha*p* linkage, such as one containing either the D-Glc*p*NAc moiety alone or attached to the polyprenolprenol phosphate scaffold upon which the arabinogalactan is synthesized.

Kinetic analysis of the purified recombinant *glfT* was performed to gain a better understanding of the substrate specificity of the enzyme. The apparent acceptor kinetic constants K_m and k_{cat} were determined using synthetic acceptors **4**−**7** (Table 1). The K_m value for the (1→6)-linked disaccharide **4** (K_m = 635 μM) was ~2.7-fold lower than corresponding (1→5)-linked isomer (K_m = 1.7 mM), demonstrating that **5** is a poorer substrate for *glfT* than **4**. In addition, the k_{cat} value for **4** is ~2.4-fold larger than that for **5**, indicating a faster turnover of the (1→6)-linked disaccharide. Because the k_{cat} and K_m values of **5** were reduced by ~2.4- and 2.7-fold, respectively, compared with **4**, there is only a <1-fold difference in k_{cat}/K_m. That **4** is a better *glfT* substrate than **5** is consistent with earlier work in which bacterial membrane preparations were used as the source of enzyme. In these studies,[23] higher levels of radioactivity incorporation into product were seen when β-D-Gal*f*-(1→6)-β-D-Gal*f*-O$(CH_2)_8CH{=}CH_2$ was used as the substrate as opposed to when β-D-Gal*f*-(1→5)-β-D-Gal*f*-O$(CH_2)_8CH{=}CH_2$ served as the acceptor. Our results are also consistent with previous work in which the Michaelis constants for **4** and **5** were determined using a crude membrane preparation as the source of enzyme.[24] In that study, the K_m values for **4** and **5** were determined to be 2.6 and 3.77 mM, respectively. These values, although higher than those calculated using our recombinant enzyme, demonstrate the same relative efficiency with which *glfT* recognizes these disaccharides. The differences in the absolute kinetic values may be explained by the use of purified enzyme sources in our assay system, as opposed to crude cell membrane preparations, or by differences in the expression system used in each study.

Inhibition of enzyme activity was seen at substrate concentrations greater than 3.13 mM for **4** and 6.25 mM for **5**. A similar trend was previously noted for *glfT* in membrane preparations when high concentrations of neoglycolipid acceptors analogous to **4** and **5** were used as substrates.[23] This inhibition could possibly be due to the detergent-like properties of the acceptors or due to the formation of micelles by the acceptors; however, the critical micellular concentration of these compounds is unknown, and thus, we cannot definitively provide a rationale for the observed inhibition.

In addition to studying the kinetic constants of the disaccharide acceptors, we also evaluated two trisaccharide acceptors to determine how the enzyme recognizes longer acceptors. In an earlier study, Kremer et al.[23] observed an additional band corresponding to a tetrasaccharide product after TLC analysis of the reaction products using a (1→5)- and (1→6)-linked disaccharide acceptors, resulting from further elongation of the initially formed trisaccharide product. Here, we show that the *glfT* enzyme has a greater affinity for longer acceptor substrates. The apparent K_m values of (1→6),(1→5)-trisaccharide **6** (K_m = 204) and the (1→5),(1→6)-trisaccharide **7** (K_m = 208) acceptors were significantly lower than for the disaccharide acceptors. Although the k_{cat} value for **6** is comparable to that of the disaccharide acceptors, the trisaccharide **7** exhibits an approximately 4-fold increase in enzyme turnover. Because the K_m values are similar and the k_{cat} value of **6** was reduced by ~4-fold compared with that of **7**, there is a 4-fold difference in k_{cat}/K_m. We note that the k_{cat} values are low, which may suggest that larger substrates (e.g., those containing the rhamnose and/or the *N*-acetylglucosamine residues) might be the optimal substrates or that a chaperone protein is needed for the enzyme to be fully active. Regardless of the cause, these low activities complicate analysis of the kinetic data.

Inspection of the k_{cat} values shown in Table 1 reveals an apparent flip in enzymatic activity when comparing the disaccharide and trisaccharide substrates. For the disaccharides, the (1→5)-transferase activity is higher (i.e., **4** has a higher k_{cat} than **5**). In contrast, for the trisaccharides, **7** has a higher k_{cat} than **6**, and thus, for these substrates, the (1→6)-transferase activity is more efficient. We are unsure as to the origin of these trends, and the bifunctional nature of the protein makes analysis of this observation difficult. However, it is plausible that these differences result from the processive nature of the enzyme and differences in the k_{cat} value of the trisaccharide substrates. The addition of Gal*f* to disaccharide **4** by *glfT* yields trisaccharide **7**, the best of all of the substrates investigated in this study. It is possible, therefore, that the enzyme may additionally glycosylate the product of the reaction with **4**, yielding not trisaccharide **7** but, instead, a tetrasaccharide (or larger) product. If this were the case, when disaccharide substrates are used, the (1→5)-transferase activity would appear higher than the (1→6)-transferase activity. Support for this hypothesis comes from an earlier study in which the reaction of disaccharide substrates with *glfT* in a crude membrane preparation were studied by thin-layer chromatography.[23] These analyses revealed that when a

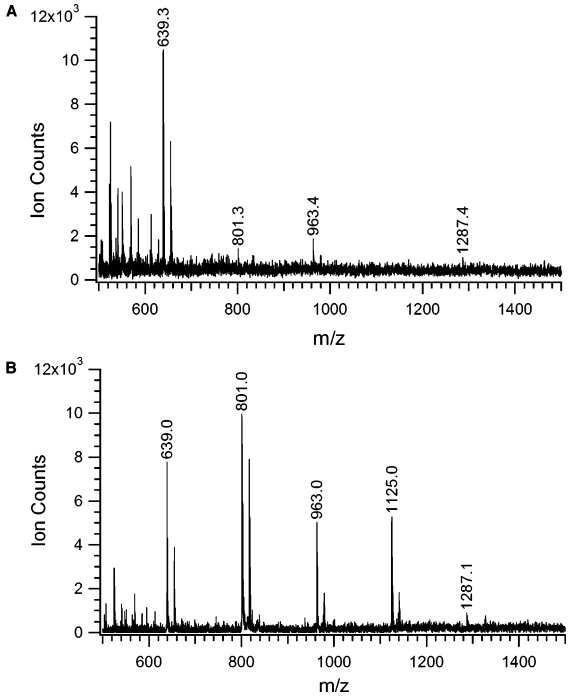

Figure 6. MALDI mass spectra of large-scale incubation mixtures of **6** (A) or **7** (B) with UDP-Gal*f* and *glfT*. The peaks at $m/z = 639$ and 801 correspond to the sodium adduct of the starting trisaccharides and tetrasaccharide products, respectively. Longer oligomers are found at $m/z = 963$, 1125, and 1287. The "shadow" signals at higher m/z are the corresponding potassium adducts.

(1→6)-linked disaccharide analogous to **4** was used as the substrate, a greater amount of longer oligomers was produced compared to when a (1→5)-linked disaccharide was used. In our own hands, we found that TLC and mass spectrometry analysis of a milligram-scale incubation of **4** with the recombinant *glfT* and UDP-Gal*f* revealed only trace amounts of the trisaccharide product, with the major product being a tetrasaccharide (data not shown). A more detailed study of these trends would require either specific inhibitors of each transferase activity or the preparation of mutant proteins that can synthesize only one of the two linkages. These studies are ongoing.

Milligram-Scale Incubations and Establishment of Product Structures. To determine that the recombinant protein was converting these substrates to the expected products, in which there were alternating β-(1→6) and β-(1→5)-linkages, we carried out milligram-scale enzymatic reactions using the trisaccharide substrates and characterized the products by mass spectrometry and ^{1}H NMR spectroscopy. Following the incubation of the substrate with UDP-Gal*f* and the enzyme, the reaction mixture was purified using a C_{18} Sep-Pak cartridge (see Methods) to provide a crude mixture of oligosaccharides, which was analyzed by MALDI mass spectrometry.

The spectrum obtained on the products isolated after incubation of **6** and UDP-Gal*f* with the enzyme (Figure 6A) shows a signal at $m/z = 801$, the mass expected for the sodium adduct of an octyl tetrasaccharide composed of galactofuranose residues. In addition, peaks corresponding to longer oligomers are also seen at $m/z = 963$ and 1287. The mass spectrum of the product obtained from the enzymatic reaction of **7** is shown in Figure 6B. Again, a peak at $m/z = 801$ is observed, clearly indicating the formation of a tetrasaccharide product. The presence of longer oligomers at $m/z = 963$, 1125, and 1287 is also apparent. Assuming that all of the oligosaccharide ionize to similar degrees, the relative abundance of the different oligosaccharides products, as detected in these mass spectra, confirms the greater (1→6)-transferase activity relative to the (1→5)-transferase activity. For example, in the case of **6** (Figure 6A), the first Gal*f* residue is added β-(1→5) and the resulting product is a substrate for the faster (1→6)-transferase activity. The comparatively slow rate of transfer to the 5-position is indicated by the small amount of tetrasaccharide product ($m/z = 801$) relative to the trisaccharide starting material ($m/z = 639$). Once formed, the tetrasaccharide is expected to be converted rapidly to the pentasaccharide ($m/z = 963$), which is present in larger amounts in the mixture of oligosaccharides. The apparent absence of a hexasaccharide ($m/z = 1125$), but presence of a heptasaccharide ($m/z = 1287$), is also consistent with these relative rates. Any hexasaccharide formed via the slow (1→5)-transferase activity would be rapidly converted to the corresponding heptasaccharide by the much faster (1→6)-transferase activity.

Having established by mass spectrometry that treatment of **6** and **7** with UDP-Gal*f* and *glfT* yields a homologous series of oligosaccharides, we then purified these mixtures by preparative TLC and isolated the tetrasaccharide products. Shown in Figure 7 is the anomeric region of the ^{1}H NMR spectra of the purified compounds. One major product is seen in both spectra, contaminated with impurities, which were later shown by mass spectrometry (data not shown) to be longer oligomers, resulting from an incomplete separation of the mixture by TLC. The

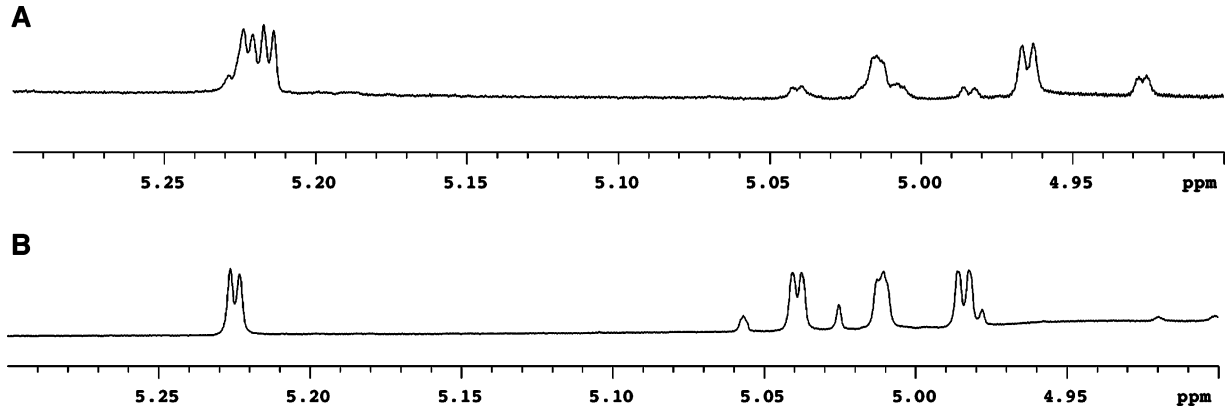

Figure 7. Partial ^{1}H NMR spectra of large-scale incubation mixtures of **6** (A) or **7** (B) with UDP-Gal*f* and *glfT*. The major signals correspond to anomeric hydrogens in tetrasaccharides **11** (A) and **12** (B). The minor resonances arise from longer oligomers that were not effectively separated by preparative TLC.

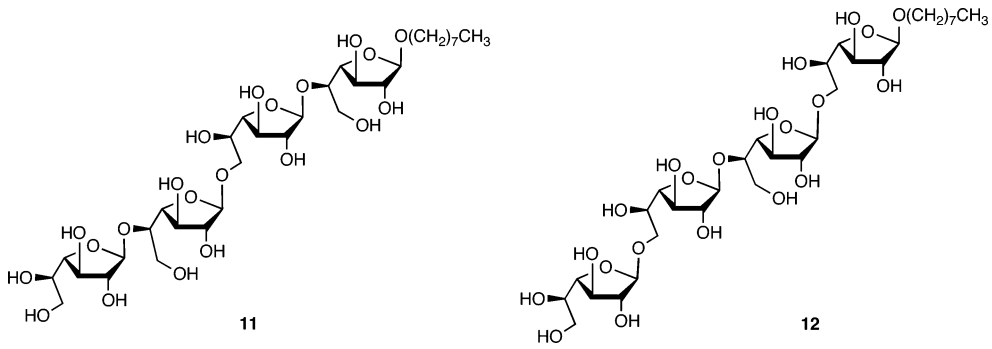

Figure 8. Tetrasaccharides **11** and **12**, the products expected from the *glfT*-catalyzed transfer of a Gal*f* residue to **6** and **7**, respectively.

major product in both spectra has four anomeric anomeric hydrogen signals, thus further confirming these compounds as tetrasaccharides. The β-stereochemistry of all of the Gal*f* residues in both tetrasaccharides was evident by their appearance as a singlet or doublet with $J_{1,2} < 2$ Hz.[48] The alternating β-(1→5) and β-(1→6) sequence of these tetrasaccharides could be established by comparing the chemical shifts of the anomeric hydrogens with data available for **4**−**7** and analogues of **6** and **7**, in which the octyl group is replaced by a decenyl moiety.[49] In all cases, the anomeric hydrogen of a β-(1→5)-linked Gal*f* residue resonates ~0.2 ppm downfield relative to a β-(1→6)-linked Gal*f* residue. For example, in the β-(1→6)-linked disaccharide **4**, the signal for H-1 of the nonreducing end Gal*f* residue resonates at 5.04 ppm, whereas for the β-(1→5)-linked disaccharide **5**, the H-1 chemical shift of the terminal residue is at 5.22 ppm. Further support for this trend can be found in the Supporting Information (Figure S1). Mindful of this information, it was possible to assign the structures of the tetrasaccharide products obtained from **6** and **7** as **11** and **12**, respectively (Figure 8). Thus, the tetrasaccharide obtained from incubation of **6** with the enzyme and UDP-Gal*f* provided an ^{1}H NMR spectrum with H-1 signals at 5.23, 5.21, 5.02, and 4.97 ppm (Figure 7A), thus demonstrating the presence of two β-(1→5)-linked Gal*f* residues. As trisaccharide **6** has only one β-(1→5)-linkage, the residue introduced by the enzyme must be attached to O-5 of the penultimate residue and the structure is **11**. In the ^{1}H NMR spectrum of the product obtained from **7** (Figure 7B), the four anomeric hydrogens were found at 5.23, 5.04, 5.01, and 4.98 ppm, thus revealing only one β-(1→5)-linkage; therefore, the residue added by the enzyme must be β-(1→6)-linked and the tetrasaccharide product is **12**. These NMR experiments provide strong support that this recombinant protein behaves as does the wild-type enzyme in synthesizing alternating β-(1→6) and β-(1→5)-linkages.

Conclusions

In this paper, we report the first high-level expression and purification of a galactosylfuranosyltransferase (*glfT*) involved in mycobacterial cell wall biosynthesis. This enzyme, which is encoded by the *Rv3808c* gene, was expressed and purified from *E. coli* C41(DE3), and 3−7 mg of pure protein were obtained from a liter of culture. We subsequently evaluated a panel of potential mono- and oligosaccharide substrates (**4**−**10**) and demonstrated, for the first time, that trisaccharides are better substrates than the disaccharides and that one disaccharide, in which the terminal D-Gal*f* residue is replaced with an L-Ara*f* moiety, is a weak substrate. Through the use of NMR spectroscopy and mass spectrometry, we showed that this recombinant enzyme, like the wild-type protein, is bifunctional and can synthesize both β-(1→6) and β-(1→5)-linkages in an alternating fashion.

The kinetic parameters shown in Table 1, along with the earlier structural work, suggest an important hypothesis for the synthesis of mycobacterial galactan in vivo and especially the action of *glfT*. First, the structural work on AG clearly showed that chain elongation from the Gal*f* residue attached to the rhamnosyl moiety occurs from O-5 and not O-6 (Figure 1).[50,51] Second, as shown earlier[23] and confirmed here, *glfT* recognizes Gal*f* residues attached either to O-5 or O-6 of another Gal*f* residue. Given these two facts (and supported by the inactivity of compound **10** as an acceptor), it follows that after the rhamnosyltransferase (WbbL)[52] attaches a rhamnosyl unit to α-D-Glc*p*NAc-P-P-decaprenyl (Figure 2), a galactofuranosyltransferase other than *glfT* must attach the first galactofuranosyl residue. This is reasonable because the acceptor is of an entirely different structure (α-L-Rha*p*-(1→3)-α-D-Glc*p*NAc-P-P-decaprenyl) than the acceptors of *glfT*. Furthermore, it is highly unlikely that *glfT* attaches the second Gal*f* residue because of its specificity to put a Gal*f* residue on position 5 of a 6-linked Gal*f* residue or on position 6 of a 5-linked Gal*f* residue. The acceptor β-D-Gal*f*-(1→4)-α-L-Rha*p*-(1→3)-α-D-Glc*p*NAc-P-P-decaprenyl does not have the structure needed to direct the attachment of the new Gal*f* residue by *glfT*. Our studies here would even suggest that perhaps the third Gal*f* residue is also not attached by *glfT*. This is because the actual substrate, β-D-Gal*f*-(1→5)-β-D-Gal*f*-(1→4)-α-L-Rha*p*-(1→3)-α-D-Glc*p*NAc-P-P-decaprenyl, is analogous to acceptor **5**, which has both a high K_m and low k_{cat} (and, hence, very low k_{cat}/K_m), and it is difficult to imagine *glfT* proceeding in vivo with these parameters. However, once three Gal*f* residues are attached (e.g., structures analogous to acceptors **6** and **7**), the K_m values drop and the k_{cat} values increase. Thus, we hypothesize that three additional enzymes, or at least three additional catalytic activities (regardless of how many proteins on which they are located) are required to build up the lipid bound growing AG molecule to the point where *glfT* begins to act and finish the extension. This hypothesis could be tested by the synthesis of acceptors containing Gal*f* and rhamnosyl residues, which is ongoing. The genes encoding these other putative galactosyltransferases have yet to be identified; a series of initial BLAST searches carried out in our lab led to no clearly similar proteins, and more sophisticated bioinformatics approaches are being investigated.

With access to pure protein, future studies will involve attempts to obtain crystal structures of the enzyme complexed

with substrates or inhibitors. Such studies will provide more detailed information on the bifunctional nature of this enzyme, in particular, whether there are one or two active sites. Based on current literature, the latter seems more likely, as the bifunctional glycosyltransferases for which detailed mechanistic information is available have two active sites each capable of synthesizing a single glycosidic linkage.[29,30,53] The screening of a larger number of potential substrates or inhibitors will also shed light on catalysis by *glfT*. Among important substrates to be tested are larger oligosaccharides and those modified at specific hydroxyl groups, as well as the donor substrate, UDP-Gal*f*. Obtaining kinetic data for the donor will require either the synthesis of radio-labeled UDP-Gal*f* by chemical[54] or chemoenzymatic[55] methods or the development of alternative assays[56−58] in which an unlabeled donor can be used.

Other issues for future research concern the elongation of the alternating Gal*f* polymer by *glfT*. For example, does the enzyme remain bound to the elongating polymer while passing the substrate from one active site to another? The mass spectral evidence tentatively argues against this with the synthetic acceptors because, although multiple additions of Gal*f* are observed, nothing resembling a complete galactan is observed. It should be remembered, however, that the acceptors studied here are soluble and that an acceptor linked to a membrane, which is therefore less diffusible, could give different results. In addition, previous work[59] has shown that in the MALDI analysis of mixtures of oligosaccharides, the ionization response often drops off with increasing molecular weight, and thus, these larger oligomers could be present, but are simply not detected. Questions also remain about how the size of the 25−30 residue galactofuranan is regulated.[60] In addition to these biophysical and mechanistic studies, access to the purified enzyme will facilitate the development of high-throughput assays for the identification of *glfT* inhibitors, which are expected to be lead compounds for new antituberculosis agents.

Acknowledgment. This work was supported by The Alberta Ingenuity for Carbohydrate Science, The Natural Sciences and Engineering Research Council of Canada, The University of Alberta, and NIH Grants AI 33706 and 1P01AI 057836 (MRM). We thank Mr. Blake Zheng, Dr. Randy Whittal, and Dr. Angelina Morales-Izquierdo for technical assistance and Professor David Sanders at the University of Saskatchewan for the gift of the UDP-Gal*p* mutase clone.

Supporting Information Available: Plots of kinetic data and NMR chemical shift information used for determining the structures of the products from the large-scale enzymatic incubations. This material is available free of charge via the Internet at http://www.acs.org.

JA058254D

(54) Mariño, K.; Marino, C.; Lima, C.; Baldoni, L.; de Lederkremer, R. M. *Eur. J. Org. Chem.* **2005**, 2958−2964.
(55) Errey, J. C.; Mukhopadhyay, B.; Kartha, K. P.; Field, R. A. *Chem. Commun.* **2004**, 2706−2707.
(56) Gosselin, S.; Alhussaini, M.; Streiff, M. B.; Takabayashi, K.; Palcic, M. M. *Anal. Biochem.* **1994**, *220*, 92−97.
(57) Zea, C. J.; Pohl, N. L. *Anal. Biochem.* **2004**, *328*, 196−202.
(58) Yang, M.; Brazier, M.; Edwards, R.; Davis, B. G. *ChemBioChem.* **2005**, *6*, 346−357.
(59) Zaia, J. *Mass Spectrom. Rev.* **2004**, *23*, 161−227.
(60) Besra, G. S.; Khoo, K.-H.; McNeil, M.; Dell, A.; Morris, H. R.; Brennan, P. J. *Biochemistry* **1995**, *34*, 4257−4266.