Organic & Biomolecular Chemistry

Cite this: *Org. Biomol. Chem.*, 2012, **10**, 4074

www.rsc.org/obc

PAPER

Synthetic UDP-galactofuranose analogs reveal critical enzyme–substrate interactions in GlfT2-catalyzed mycobacterial galactan assembly†

Myles B. Poulin,‡ Ruokun Zhou‡ and Todd L. Lowary*

Received 20th January 2012, Accepted 2nd March 2012
DOI: 10.1039/c2ob25159k

Mycobacterial cell wall galactan, composed of alternating β-(1→5) and β-(1→6) galactofuranosyl residues, is assembled by the action of two bifunctional galactofuranosyltransferases, GlfT1 and GlfT2, which use UDP-galactofuranose (UDP-Gal*f*) as the donor substrate. Kinetic analysis of synthetic UDP-Gal*f* analogs identified critical interactions involved in donor substrate recognition by GlfT2, a processive polymerizing glycosyltransferase. Testing of methylated UDP-Gal*f* analogs showed the donor substrate-binding pocket is sterically crowded. Evaluation of deoxy UDP-Gal*f* analogs revealed that the C-6 hydroxyl group is not essential for substrate activity, and that interactions with the UDP-Gal*f* C-3 hydroxyl group orient the substrate for turnover but appears to play no role in substrate recognition, making the 3-deoxy-analog a moderate competitive inhibitor of the enzyme. Moreover, the addition of a Gal*f* residue deoxygenated at C-5 or C-6, or an L-arabinofuranose residue, to the growing galactan chain resulted in "dead end" reaction products, which no longer act as an acceptor for the enzyme. This finding shows dual recognition of both the terminal C-5 and C-6 hydroxyl groups of the acceptor substrate are required for GlfT2 activity, which is consistent with a recent model developed based upon a crystal structure of the enzyme. These observations provide insight into specific protein–carbohydrate interactions in the GlfT2 active site and may facilitate the design of future inhibitors.

Introduction

Mammalian glycoconjugates contain galactose residues exclusively in the thermodynamically favored six-membered pyranose ring form (Gal*p*). However, galactose in the five-membered furanose ring form (Gal*f*) is found in many microorganisms.[1,2] Among these are mycobacteria, which continue to have a significant impact on world health. *Mycobacterium tuberculosis*, the causative agent of tuberculosis (TB), infects approximately one-third of the world's population and results in nearly three million deaths annually.[3–5] Recently, concern over TB has increased due to the emergence of multi-drug resistant and extensively-drug resistant strains of the organism.[6] The need for new anti-mycobacterial therapeutics, and the absence of Gal*f* in mammalian tissues, has lead to an interest in Gal*f* metabolism as a potential target for drug action.[7]

In mycobacteria, Gal*f* residues are found in the complex and glycan-rich cell wall, specifically, the mycolyl–arabinogalactan (mAG) complex, a lipidated polysaccharide composed almost entirely of furanose carbohydrates.[8,9] This glycan is the largest structural component of the mycobacterial cell wall and is covalently attached to cell wall peptidoglycan through an α-L-Rha*p*-(1→3)-α-D-Glc*p*NAc-phosphate disaccharide.[9] The core of the mAG is a galactan composed of 30–35 D-Gal*f* residues connected through alternating β-(1→5) and β-(1→6) linkages. Three arabinan domains composed of D-arabinofuranosyl (Ara*f*) residues are attached at the C-5 hydroxyl group of the eighth, tenth and twelfth Gal*f* residue of the galactan. These arabinan domains are further esterified with mycolic acids, large C_{70}–C_{90} branched lipids that impart significant hydrophobicity to the mycobacterial cell wall.

Many of the pathophysiological features of mycobacterial infections are attributed to the cell wall.[10] For example, this structure contributes to difficulties in treating mycobacterial infections, by acting as a permeability barrier to antibiotics.[11] As a result, the enzymes involved in mAG biosynthesis are attractive targets for new anti-mycobacterial theraputics.[12] Two of the standard drugs currently used to treat TB, isoniazid and ethambutol, target the biosynthesis of the mycolic acid and arabinan components of the mAG, respectively.[13,14] However, no currently used TB drugs are known to target the assembly of the mAG galactan.

The biosynthesis of mAG galactan involves two bifunctional galactofuranosyl-transferase enzymes, GlfT1 and GlfT2 (Fig. 1).[15]

Alberta Glycomics Centre and Department of Chemistry, Gunning–Lemieux Chemistry Centre, University of Alberta, Edmonton, AB T6G 2G2, Canada. E-mail: tlowary@ualberta.ca; Fax: +1 780-492-7705; Tel: +1 780-492-1861

†Electronic supplementary information (ESI) available. See DOI: 10.1039/c2ob25159k

‡These authors contributed equally to this work.

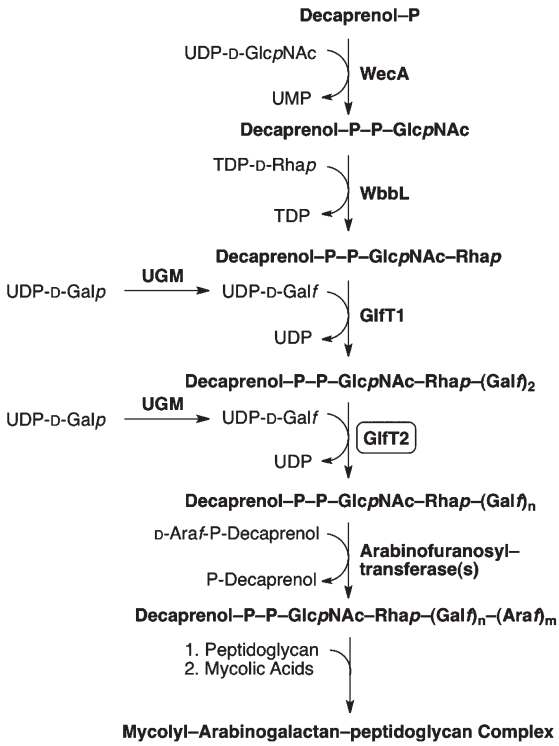


Fig. 1 Proposed pathway for mAG biosynthesis. GlfT2, the galactofuranosyltransferase of interest in this study, is highlighted. $n \approx 30–35$; $m \approx 93$.

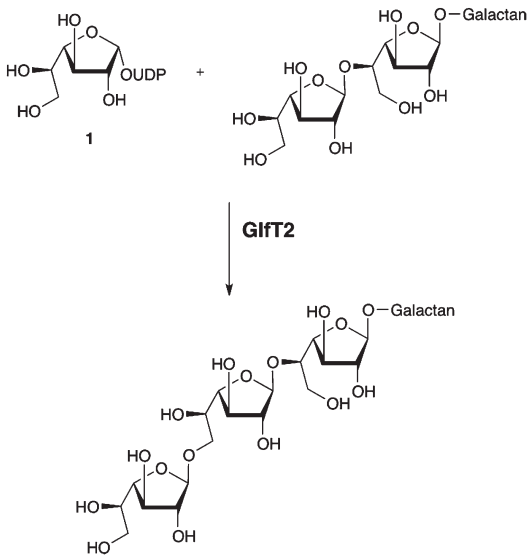


Fig. 2 A representative reaction catalyzed by GlfT2.

As the source of Gal*f*, both enzymes use the sugar nucleotide UDP-Gal*f* (**1**), which is biosynthesized from UDP-galactopyranose (Gal*p*) by the action of UDP-galactopyranose mutase (UGM).[16] GlfT1 is responsible for adding the first and second Gal*f* residues to an α-L-Rha*p*-(1→3)-α-D-Glc*p*NAc-decaprenyl phosphate acceptor.[15,17] The final product of the GlfT1-catalyzed reaction is the initial GlfT2 acceptor substrate.

The gene encoding GlfT2 was first identified in 2000,[18] and recombinant GlfT2 has since been expressed and purified.[17,19] Subsequent studies on the enzyme have demonstrated that GlfT2 is bifunctional, and synthesizes both β-Gal*f*-(1→5)-Gal*f* and β-Gal*f*-(1→6)-Gal*f* linkages using a single active site (Fig. 2).[20,21] More recent crystallographic investigations have confirmed the presence of only one catalytic site.[22] Using synthetic galactan fragments (*e.g.*, **2** and **3**, Fig. 3) we, and others, have demonstrated that GlfT2 requires a β-(1→5)- or β-(1→6)-linked Gal*f* disaccharide as the minimum acceptor substrate.[19,23] GlfT2 acts as a processive polymerase[24] adding the third and subsequent Gal*f* residues to a growing galactan chain without release of the acceptor substrate after each glycosyl transfer reaction. Other studies have suggested that galactan length is controlled by tethering of the polyprenol aglycone to the protein.[25] However, despite these investigations, there is still relatively little known about the specific protein–carbohydrate interactions that GlfT2 uses to recognize its substrates.

The progress in establishing the mechanism of GlfT2 has not been matched by the identification of potent and selective inhibitors of this enzyme. Methylated disaccharide acceptor analogs showed toxicity at micromolar concentrations to live mycobacteria; however, this effect appeared to be due to a nonspecific surfactant effect, as these compounds failed to inhibit GlfT2 activity in a cell free assay.[23] More recent studies on carbasugar disaccharide acceptor analogs showed greater then 50% inhibition of the enzyme, but only at millimolar concentrations.[26] Amino sugar analogs of Gal*f* have also been designed as mimics of the GlfT2 transition state, but again they afforded only weak ($IC_{50} \sim 4.8$ mM) inhibition activity.[27,28] A later transition state analog, incorporating UDP, gave 80% inhibition at 1 mM concentration,[29] but no further work was done to establish its specific inhibitory activity. The most potent GlfT2 inhibitor reported to date is a sugar–amino acid–nucleoside analog of UDP-Gal*p*, with an IC_{50} value of 332 μM.[30] However, the galactose in this analog is in the pyranose ring form and its selectivity for GlfT2, as opposed to other galactosyltrasferases, is unknown.

To facilitate the design of more potent and selective inhibitors of GlfT2, we have used a panel of singly modified UDP-Gal*f* analogs to probe specific protein–carbohydrate interactions involved in substrate recognition and turnover. Singly methylated and deoxygenated carbohydrate analogs have shown great utility as biological tools to explore protein–carbohydrate binding interactions, and in some cases have led to the identification of specific glycosyltransferase inhibitors.[31–33] Herein, we report the chemo-enzymatic synthesis of a panel of singly modified UDP-Gal*f* (**4–11**) analogs, which were then used to probe the donor specificity of GlfT2, and to explore their effect on galactan polymerization. These studies expand on previous studies focusing on the acceptor substrate specificity of GlfT2.[19,34]

Results and discussion

Preparation of UDP-galactofuranose analogs

To prepare the UDP-Gal*f* analogs, **4–10**, we employed a chemo-enzymatic approach previously used for the preparation of UDP-Gal*f* **1** (Fig. 3).[34–36] This method uses a three enzyme system to convert galactofuranosyl-1-phosphate (Gal*f*-1P) to

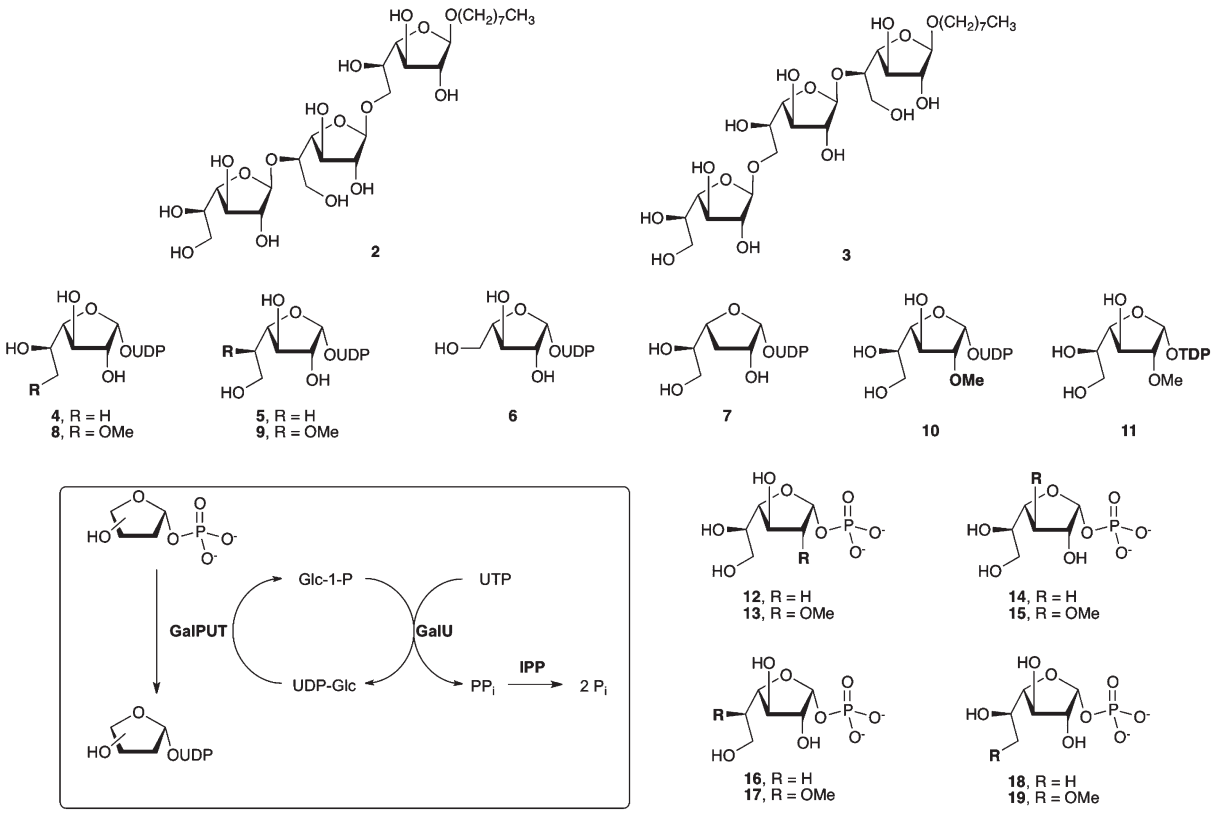
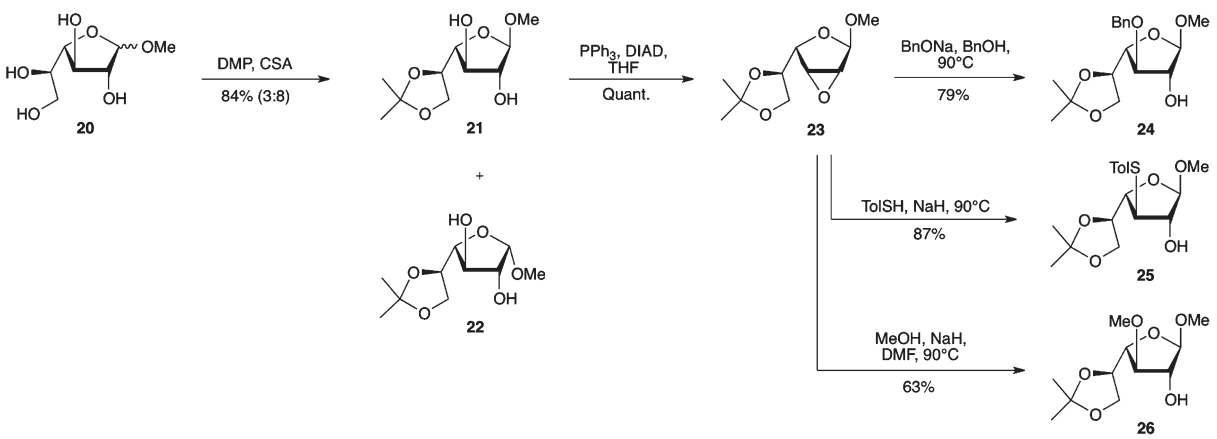


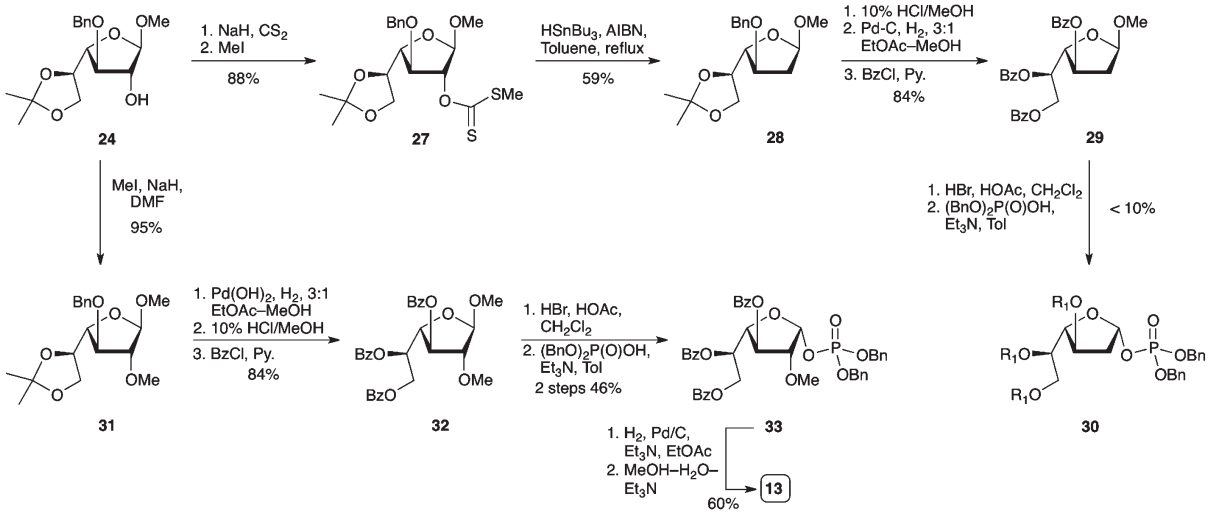
Fig. 3 Synthetic acceptor trisaccharides and UDP-Gal*f* analogs screened to examine GlfT2 specificity in this study. UDP-Gal*f* analogs were prepared using a three enzyme chemo-enzymatic reaction from synthetic Gal*f*-1P precursors **12–19**.



Scheme 1 Synthesis and reactions of key epoxide intermediate **23**.

UDP-Gal*f*. It has been exploited here for the preparation of UDP-Gal*f* analogs **4–10**. To apply this approach, the Gal*f*-1P analogs **12–19** were first prepared. The route to **12–19** involved first the synthesis of the corresponding methyl glycosides and their conversion to the target deprotected Gal*f*-1P analogs as outlined in the following sections.

Gal*f* analogs modified at C-2 and C-3. The synthesis of C-2 and C-3 modified Gal*f* analogs **12–15** started from known methyl galactofuranoside **20**[37,38] as outlined in Scheme 1. The two hydroxyl groups on the acyclic carbon chain were selectively protected as an isopropylidene acetal to afford both **21** and **22** in a ratio of 3 : 8 and in 84% yield. An efficient method to selectively protect both the C-2 and C-3 hydroxyl groups was required and this could be accomplished by first converting these diols into an epoxide by a Mitsunobu reaction, followed by epoxide opening using an appropriate nucleophile.[39] The regioselectivity of the epoxide opening reaction is influenced by the stereochemistry at the anomeric centre, with the α-glycoside expected to give poor regioselectivity and the β-glycoside expected to give exclusively C-3 attack.[40] Therefore, only the minor product, β-glycoside **21**, was carried forward. Treatment of **21** with triphenylphosphine and diisopropyl azodicarboxylate (DIAD) gave epoxide **23** in quantitative yield.



Scheme 2 Synthesis of C-2 modified analogs **12** and **13**.

Epoxide **23** served as a common intermediate for the preparation of all four C-2 and C-3 modified Gal*f* analogs by treating with a variety of nucleophiles. Treatment with the sodium salt of benzyl alcohol at 90 °C gave exclusively the product of C-3 attack, **24**, with a free C-2 hydroxyl group in 79% yield. In the ^{1}H NMR spectrum of **24**, H-1 appeared as a singlet, thus confirming the product stereochemistry.[41] Similarly, treatment of **23** with *p*-toluenethiol and sodium hydride at 90 °C gave the expected C-3 thioether **25** in 87% yield, and heating the epoxide to reflux in the presence of sodium methoxide resulted in the formation of methyl ether **26** in 63% yield. In the latter reaction, the use of DMF as the solvent proved necessary, as no reaction was observed when the reaction was carried out in methanol or THF. Presumably, the polar aprotic character of DMF and the higher reaction temperature help promote opening of the epoxide by the nucleophile.

For the synthesis of the 2-deoxy analog, the free C-2 hydroxyl group of **24** was first removed by means of a Barton–McCombie deoxygenation *via* the requisite xanthate **27** (Scheme 2) to give **28** in 52% yield over the two steps. Deprotection of the isopropylidene ketal under acidic conditions (10% methanolic HCl) followed by removal of the benzyl ether by catalytic hydrogenolysis and then benzoylation of the resulting hydroxyl groups provided the required perbenzoylated 2-deoxy methyl glycoside **29** in an 84% overall yield.

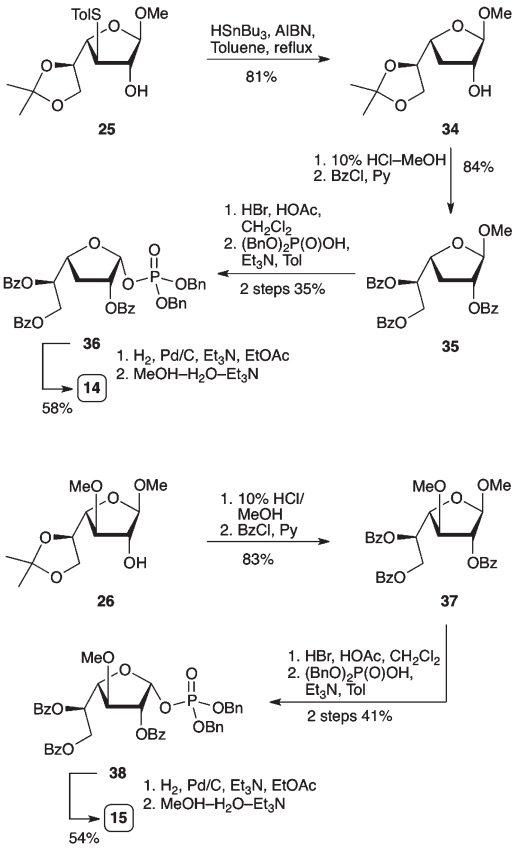
Treatment of methyl galactofuranosides with HBr in acetic acid results in formation of the corresponding glycosyl bromide, which can then be treated with dibenzyl phosphate to provide α-galactofuranosyl phosphates.[42] Unfortunately, treatment of methyl glycoside **29** under the above conditions resulted predominantly in hydrolysis of the methyl glycoside and produced only small (<10%) amounts of the desired phosphate **30**. Attempts to deprotect this intermediate resulted in further hydrolysis giving none of the desired compound **12**. Presumably, removal of the electron withdrawing C-2 oxygen reduces the barrier to oxocarbenium ion formation (and its subsequent trapping with water) from either the glycosyl bromide or dibenzyl phosphate. Given these difficulties, the synthesis of the 2-deoxy Gal*f*-1-phosphate analog **12** was abandoned.

In contrast to the 2-deoxy analog, the 2-*O*-methyl analog was easily obtained from **24**, as illustrated in Scheme 2. Methylation of the free C-2 hydroxyl group afforded intermediate **31** in 95% yield. Removal of the 3-*O*-benzyl group by catalytic hydrogenolysis and cleavage of the isopropylidene acetal under acidic conditions, followed by protection of the three resulting hydroxyl groups using benzoyl chloride in pyridine, afforded the desired 2-*O*-methyl methyl glycoside **32** in 84% yield over the three steps. This methyl glycoside was treated with HBr in acetic acid followed by dibenzyl phosphate to give the α-glycosylphosphate **33** as the major product in 46% yield. In this, and all of the successful phosphorylation reactions described below, a small amount (<5%) of the β-galactofuranosyl phosphate was also detected in the ^{1}H NMR spectra of the products. Global deprotection, by first hydrogenolysis of the benzyl groups and then removal of the benzoate esters under weakly basic conditions, afforded a 60% yield of Gal*f*-1P analog **13**.

The synthesis of the 3-deoxy methyl glycoside was achieved from thioether **25** as illustrated in Scheme 3. Reduction of the thioether under radical conditions with AIBN and tributyltin hydride at reflux yielded the 3-deoxy methyl glycoside **34** in 81% yield. To remove the isopropylidene ketal of **34**, 10% methanolic HCl was employed, and the resulting intermediate was benzoylated to produce **35** in 84% yield over the two steps. As described above, treatment with HBr in acetic acid and then dibenzyl phosphate gave **36** as the major product in 35% overall yield. Finally, cleavage of the benzyl ethers and benzoate esters gave, in 58% yield over the two steps, the Gal*f*-1P analog **14**.

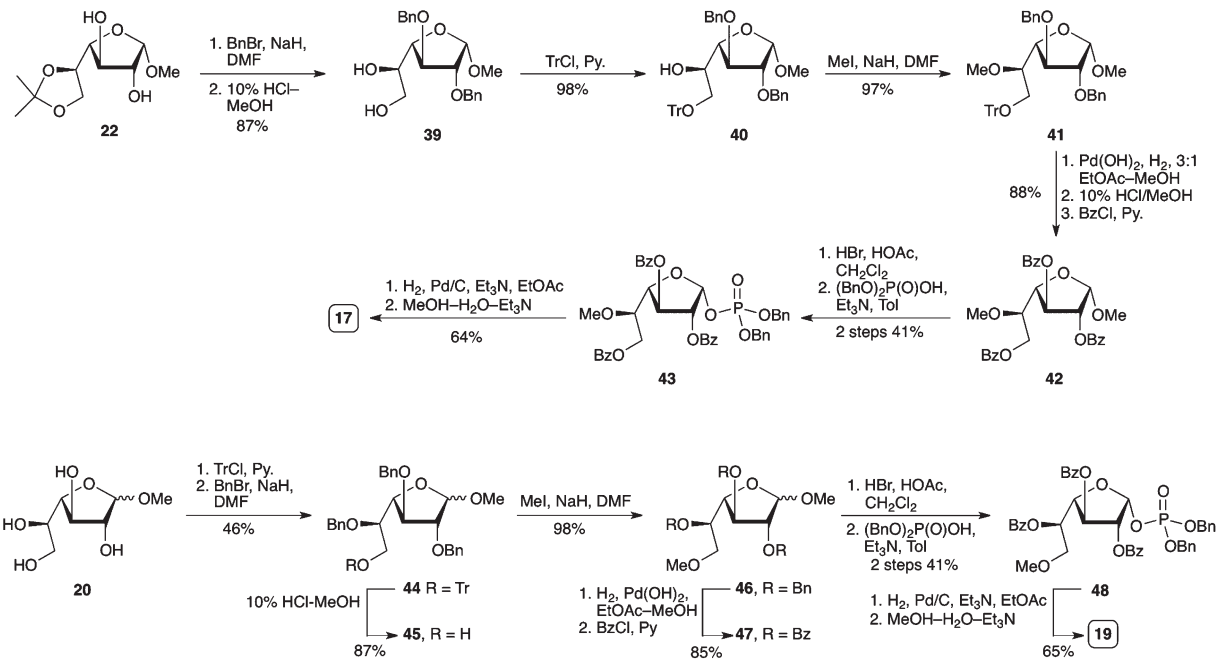
Intermediate **26** was used for the preparation of the 3-*O*-methyl analog **15** (Scheme 3). The isopropylidene acetal was removed by treatment with 10% HCl, and the three resulting hydroxyl groups were protected with benzoyl chloride to give **37** in 83% yield over these two steps. Treating **37** with HBr in acetic acid followed by dibenzyl phosphate provided **38** as the major product in 41% yield after purification. Deprotection of this intermediate, as was done for the preparation of **13** and **14**, provided the 3-*O*-methyl-Gal*f*-1P analog **15** in 54% yield.

Gal*f* analogs modified at C-5″. The 5-deoxy methyl glycoside **16** was prepared as previously described.[43] For the synthesis of 5-*O*-methyl analog **17**, the isopropylidene protected α-glycoside **22** was employed as outlined in Scheme 4. First, the two free hydroxyl groups of **22** were protected as benzyl ethers using standard conditions. Next, removal of the isopropylidene acetal from the resulting intermediate under acidic conditions afforded the C-5/C-6 diol **39** in 87% overall yield. Selective protection of the C-6 hydroxyl group was achieved using trityl chloride in pyridine, giving the C-5 alcohol **40** in excellent yield. Methylation of alcohol **40**, employing methyl iodide and sodium hydride in DMF, afforded the 5-*O*-methyl analog **41**, again in excellent yield. The benzyl and trityl ethers of **41** were removed by catalytic hydrogenation and treatment with 10% methanolic HCl, respectively. Protection of the three resulting hydroxyl groups as benzoate esters afforded methyl glycoside **42** in three steps and 88% overall yield. Finally, treatment with HBr in acetic acid followed by addition of dibenzyl phosphate gave intermediate **43**, which was deprotected as was done for the other analogs to provide a 64% yield of the 5-*O*-methyl Gal*f*-1P analog **17**.



Scheme 3 Synthesis of C-3 modified analogs **14** and **15**.

Gal*f* analogs modified at C-6″. The synthesis of 6-deoxy analog **18** was achieved as previously described.[43] It was originally envisioned that both the 6-deoxy analog **18** and 6-*O*-methyl analog **19** could arise from a common intermediate containing a free hydroxyl group at C-6 and benzoate ester at C-2, C-3 and C-5 (ESI Scheme S1†). However, attempts to methylate this intermediate, even using the weakly basic conditions of methyl iodide and silver oxide, failed to yield the desired 6-*O*-methyl analog **47**, and led instead to significant acyl group migration. Therefore, an alternative approach was used (Scheme 4). The C-6 hydroxyl group of methyl glycoside **20** was selectively protected using trityl chloride and the remaining hydroxyl groups were benzylated to afford **44** in a modest 46% overall yield. Trityl group deprotection with 10% methanolic HCl provided **45** with a free C-6 hydroxyl group. This product was methylated using standard conditions to give the



Scheme 4 Synthesis of C-5 and C-6-*O*-methyl analogs **17** and **19**.

Table 1 Kinetic parameters for GlfT2 with synthetic acceptors **2** and **3** and synthetic donors **1** and **4–11**

Donor substrate	Acceptor substrate	Specific activity[a] (μmol min^{-1} mg^{-1})	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M (μM $min)^{-1}$
1, UDP-Gal*f*	**2**	1.6 (±0.1)	2.5 (±0.4) × 10^2	1.31 (±0.07) × 10^2	5 (±1) × 10^{-1}
4, UDP-6d-Gal*f*	**2**	0.9 (±0.1)	1.0 (±0.2) × 10^3	7.9 (±0.5) × 10^1	7 (±3) × 10^{-2}
5, UDP-5d-Gal*f*	**2**	0.10 (±0.01)	9 (±1.7) × 10^2	10 (±1)	1.2 (±0.6) × 10^{-2}
6, UDP-Ara*f*	**2**	0.14 (±0.01)	9 (±1) × 10^2	14.7 (±0.7)	1.5 (±0.6) × 10^{-2}
7, UDP-3d-Gal*f*	**2**	0.01 (±0.01)	5.0 (±2.8) × 10^2	1.2 (±0.2)	2.2 (±0.8) × 10^{-3}
8, UDP-6OMe-Gal*f*	**2**	7 (±1) × 10^{-3}	n.d.[b]	n.d.	n.d.
9, UDP-5OMe-Gal*f*	**2**	6 (±2) × 10^{-3}	n.d.	n.d.	n.d.
10, UDP-2OMe-Gal*f*	**2**	4 (±3) × 10^{-3}	n.d.	n.d.	n.d.
11, dTDP-Gal*f*	**2**	0.35 (±0.02)	1.9 (±0.3) × 10^3	9.7 (±0.9) × 10^1	2 (±1) × 10^{-2}
1, UDP-Gal*f*	**3**	0.14 (±0.01)	1.0 (±0.3) × 10^3	2 (±1) × 10^1	2 (±3) × 10^{-2}
4, UDP-6d-Gal*f*	**3**	12 (±2) × 10^{-3}	n.d.	n.d.	n.d.
5, UDP-5d-Gal*f*	**3**	0.8 (±0.5) × 10^{-3}	n.d.	n.d.	n.d.
6, UDP-Ara*f*	**3**	10 (±1) × 10^{-3}	n.d.	n.d.	n.d.

[a] Determined at 2.0 mM donor and 2.0 mM acceptor. [b] Kinetic parameters could not be obtained due to low activity.

6-*O*-methyl analog **46** in an excellent yield. Following this, the benzyl groups were exchanged for benzoates to provide the required protected methyl glycoside **47** in 85% yield over two steps. The desired Gal*f*-1P analog **19**, was obtained by treating **47** with HBr in acetic acid followed by the addition of dibenzyl phosphate to give the expected α-Gal*f*-1P derivative **48**, which was deprotected to give **19** in a 64% yield over two steps.

Chemo-enzymatic synthesis of UDP-Gal*f* analogs. We have previously employed a chemo-enzymatic approach for the preparation of UDP-Gal*f* (**1**),[34] as well as UDP-Gal*f* analogs **4–6**.[43] The method employs a promiscuous galactose-1-phosphate uridyltranserase (GalPUT) that was previously shown to convert a variety of hexose-1-phosphate analogs, including hexose sugars in the furanose ring form, into the corresponding UDP-sugars.[35,44,45] This procedure proceeds efficiently and in high yield for a wide range of substrates, offering advantages over the entirely chemical approaches[46–49] that have been used to prepare UDP-Gal*f* and analogs thereof. In the current study, we used this method to produce **4–10** from **12–19**.[36] A limitation of this strategy is the ability of GalPUT to recognize and turn over the modified Gal*f*-1P analogs. The deoxy Gal*f*-1P analogs were well tolerated, yielding 35–78% of the corresponding UDP-Gal*f* derivatives. However, the methylated Gal*f*-1P analogs were very poor substrates providing <5% isolated yields of the product. In the case of 3-*O*-methyl Gal*f*-1P, although product was formed, insufficient material could be isolated to test its activity with GlfT2.

Chemo-enzymatic synthesis of dTDP-Gal*f*. To prepare dTDP-Gal*f* (**11**), we employed a different chemo-enzymatic approach, which uses a bacterial α-D-glucopyranosyl-1-phosphate thymidylyltransferase (Cps2L) to convert sugar-1-phosphates and deoxythymidine 5′-triphosphate (dTTP) into dTDP-sugars.[50] Jakeman and coworkers have previously demonstrated that Cps2L tolerates a broad range of sugar-1-phosphate substrates, including Gal*f*-1P and other furanosyl-1-phosphates.[51] To prepare milligram quantities of dTDP-Gal*f*, we modified the reported procedure by immobilizing Cps2L on Ni-NTA agarose resin, as was done with GalPUT. This both increased protein stability and facilitated product purification. Using this approach, **11** was obtained in 50% overall isolated yield from Gal*f*-1P.

GlfT2 activity and specificity with synthetic donor analogs

Earlier, a continuous spectrophotometric assay was developed to monitor the activity of GlfT2.[34] This assay monitors the formation of UDP liberated from **1** upon GlfT2-mediated transfer of Gal*f* to an acceptor substrate. UDP formation is coupled to the oxidation of NADH *via* pyruvate kinase and lactate dehydrogenase. As the assay is not specific for UDP,[52] it can also be employed to monitor GlfT2 activity with dTDP-Gal*f* through detection of liberated dTDP.

We employed this assay to probe the donor binding site specificity of GlfT2 by measuring the specific transferase activity with **4–11**. In these assays, the trisaccharide β-Gal*f*-(1→5)-β-Gal*f*-(1→6)-β-Gal*f*-octyl (**2**) served as the primary acceptor substrate.[37] Previous studies reported that **2** acts as the preferred GlfT2 acceptor substrate when compared to the isomeric β-Gal*f*-(1→6)-β-Gal*f*-(1→5)-β-Gal*f*-octyl trisaccharide (**3**).[19] Donors **4–6** were previously shown to act as substrates for GlfT2,[43] and were further screened to compare their activity with the other substrates and to determine kinetic parameters. To determine whether GlfT2 donor specificity is influenced by the nature of the acceptor, and glycosidic linkage being formed, the donor analogs were also screened using acceptor **3**.

Effect of deoxy UDP-Gal*f* derivatives on GlfT2 activity. As presented in Table 1 with acceptor **2**, deoxy UDP-Gal*f* donor analogs **4**, **5**, and **7** and UDP-Ara*f* (**6**) all served as GlfT2 substrates with varying degrees of efficiency. Of these, the 6-deoxy analog **4** had the highest relative activity at greater than 55% compared to the natural substrate, **1**. The moderate activity observed for **4** suggests that any hydrogen bonding interactions between GlfT2 and the UDP-Gal*f* C-6 hydroxyl group are not critical for either substrate recognition or turnover. The much lower activity of UDP-Ara*f* analog **6**, which, in contrast to **4**, lacks both the C-6 carbon and hydroxyl group, suggests that hydrophobic interactions with C-6 help facilitate, but are not critical for, donor substrate binding and transferase activity.

Analog **5**, lacking a C-5 hydroxyl group, also demonstrated a decrease in activity similar to **6**. The donor analog **7**, in which the C-3 hydroxyl group has been removed, displayed less then 1% relative activity, which suggests interactions at this position are important in either substrate binding or turnover (see additional discussion below). All four of these donor analogs showed sufficient activity to allow for full kinetic characterization.

Kinetic analysis with deoxygenated UDP-Gal*f* analogs. To better understand the role for each of the interactions described above, the apparent kinetic constants K_M and k_{cat} were determined for **4–7**, using acceptor substrate **2**. As presented in Table 1, the K_M value for **4** was ~4-fold higher compared to the native donor **1**, whereas only a ~1.6-fold decrease in k_{cat} was observed, indicating that removal of the hydroxyl group at C-6 has only a moderate effect on substrate binding without substantially affecting transferase activity. Analogs **5** and **6** demonstrate a similar ~3.6-fold increase in K_M and additionally resulted in a ~12-fold and ~9-fold decrease in k_{cat}, respectively. These results show that interactions with the C-5 hydroxyl or hydroxymethyl group have only a moderate effect on substrate binding, but play a larger role in facilitating substrate turnover. Presumably, these interactions play an important, although not required, role in stabilizing the bound substrate in the optimal conformation for turnover. The low specific activity of 3-deoxy analog **7** complicated kinetic analysis; however, an ~2-fold increase in K_M was observed, indicating hydrogen bonding to the C-3 hydroxyl group is not key to substrate binding. Conversely, this analog showed a greater than 100-fold decrease in k_{cat}, suggesting hydrogen bonding interactions involving the C-3 hydroxyl group play a critical role in orienting the substrate for turnover.

Effect of methyl UDP-Gal*f* derivatives on GlfT2 activity. The three prepared methyl donor analogs **8–10** were also screened for activity with GlfT2, and each demonstrated >230-fold decrease in activity (Table 1). Without results for the 2-deoxy UDP-Gal*f* analog we cannot conclusively show that the decrease in activity observed for **10** is due to additional steric interaction or the disruption of hydrogen bonding interactions. However, the decreased activity of analogs **8** and **9**, when compared to the corresponding deoxy analogs **4** and **5**, imply GlfT2 cannot tolerate additional steric bulk in the donor-binding site and the results for **10** are consistent with this hypothesis. This additional steric bulk either prevents substrate binding or substantially impairs turnover after the substrate is bound. To distinguish between these two possibilities, analogs **8–10** were tested for their ability to inhibit GlfT2 activity. If the analogs were bound by the enzyme but poorly turned over, we expected that they would serve as inhibitors of GlfT2 activity. However, only 26% and 15% inhibition was observed for **8** and **9** respectively, and no inhibitory activity could be detected with **10** (Table 2). It appears therefore, that the addition of the methyl groups disrupt initial substrate binding. This finding is consistent with the model developed as part of a recent X-ray crystallographic study of the enzyme in which a tight donor binding site was proposed.[22]

Effect of acceptor on GlfT2 donor specificity. GlfT2 is a carbohydrate polymerase capable of adding both the β-(1→6)- and β-(1→5)-linked Gal*f* residues to the growing mycobacterial galactan polymer.[10,19] All the measurements discussed above have focused on the β-(1→6)-transferase activity using the synthetic acceptor **2**. We, and others, have previously shown that both transferase activities of GlfT2 originate from the same active site,[20,21] but it is unknown whether the nature of the glycosidic linkage being formed has an influence on the donor substrate binding and specificity of GlfT2. To address this, the same relative activity measurements were performed using the alternative acceptor substrate **3**, which is initially a β-(1→5)-transferase substrate.

Table 2 GlfT2 inhibition with synthetic donors **5–10**

Inhibitor	Inhibition[a]	K_i (μM)
5, UDP-5d-Gal*f*	10%	n.d.
6, UDP-Ara*f*	<1%	n.d.
7, UDP-3d-Gal*f*	68%	1.2 (±0.2) × 10^2
8, UDP-6OMe-Gal*f*	26%	n.d.
9, UDP-5OMe-Gal*f*	15%	n.d.
10, UDP-2OMe-Gal*f*	n.d.[b]	n.d.

[a] Determined at 1.25 mM inhibitor 0.375 mM donor **1** and 2.0 mM acceptor **2**. [b] Inhibitory activity could not be determined.

Compared to **2**, trisaccharide **3** is a much poorer GlfT2 substrate; indeed, the relative activity was only 9% compared to acceptor **2**. When we attempted to measure kinetic parameters we found that the data showed only a modest fit to the Michaelis–Menten equation (ESI Fig. S1†). GlfT2 has been shown to use a processive mechanism, where the growing galactan polymer remains bound between successive transfers of Gal*f* before dissociating, rather then a distributive mechanism, where the galactan dissociates following after each addition.[24,25] Therefore, the observed activity at higher UDP-Gal*f* concentration likely resulted from the addition of multiple Gal*f* residues. When the data was fit to the Michaelis–Menten equation containing a Hill-slope factor, a better fit to the data (ESI Fig. S1-D†) was seen. It appears, therefore, that at a low concentration of donor **1** we detect predominantly β-(1→5)-transferase activity of GlfT2; however, at higher donor concentrations both the β-(1→5)- and β-(1→6)-transferase activities are measured. As the activity with acceptor **3** was lower than with acceptor **2**, only analogs **4–6** were screened as substrates. In this case, the relative activity values varied when compared to the values observed for the same incubations using acceptor **2** (Table 1), but the relative trend is the same. It appears that hydrogen bonding to the C-5 and C-6 hydroxyl groups in the donor play a more important role for the β-(1→5)-transferase activity compared to the β-(1→6)-transferase activity, as evidenced by the lower relative activities observed for analogs **4** and **5** when acceptor **3** was used.

dTDP-Gal*f* activity and kinetics. Recent work studying the substrate binding of GlfT2 using saturation transfer difference (STD)-NMR demonstrated that a greater relative saturation transfer is observed for the protons of the ribose sugar and nucleotide base of **1** relative to the Gal*f* protons, implying the nucleotide is bound more tightly than other parts of the molecule.[53] These observations are consistent with STD-NMR studies on other

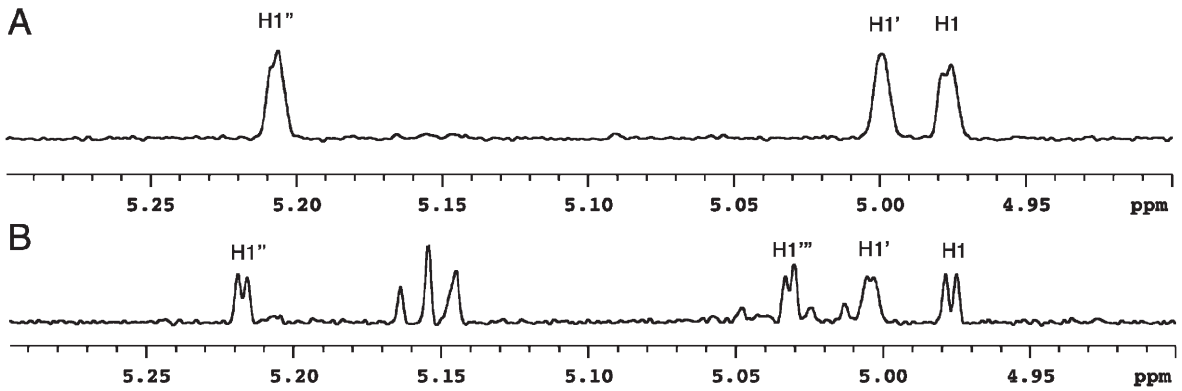


Fig. 4 Partial ^{1}H NMR spectra of trisaccharide acceptor **2** (A) and the product resulting from the incubation of **11** with acceptor **2** and GlfT2 (B). The major signals corresponding to the anomeric hydrogen are labeled.

glycosyltransferases,[54,55] and led to the hypothesis that modifications of the donor nucleotide would result in a large decrease in GlfT2 activity, by disrupting substrate binding. When the donor nucleotide is changed from UDP (**1**) to dTDP (**11**) we observed ~20% relative turnover. Kinetic analysis of **11** revealed an ~8-fold increase and ~1.3-fold decrease in K_M and k_{cat}, respectively, suggesting that the nucleotide portion of the donor is primarily involved in initial substrate binding and recognition, but once bound does not interfere with substrate turnover. The crystal structure of the enzyme with UDP reveals no key hydrogen-bonding interactions with the ribose 2-hydroxyl group are present, and in addition, sufficient room for binding the methyl group of thymidine consistent with our observed results.[22]

The UDP-Gal*f* analog lacking a C-3 hydroxyl group is a moderate GlfT2 inhibitor. As described above, UDP-Gal*f* analogs **5**–**7**, demonstrated low relative activities with GlfT2 while still displaying moderate K_M values. These compounds were therefore screened as inhibitors of the enzyme. Relative inhibition values were obtained by incubating 0.375 mM donor **1** and 2 mM acceptor **2** in the presence of 1.25 mM inhibitor. As presented in Table 2, under these conditions only the 3-deoxy analog (**7**) possessed greater than 30% inhibition. Kinetics measurements revealed that **7** is competitive inhibitor of GlfT2 with a K_i of 120 μM; *cf*. the K_M is 250 μM (Table 1, ESI Fig. S2†). To our knowledge, this represents the most potent GlfT2 inhibitor reported to date.

Characterizing GlfT2 reaction products of synthetic UDP-Gal*f* analogs

Using synthetic UDP-Gal*f* analogs **4**–**11** we uncovered potential hydrogen bonding interactions that are important for GlfT2 activity, but we also wanted to examine the effects of these analogs on both the polymerizing ability of GlfT2, and the regiochemistry of the newly formed glycosidic linkages. To accomplish this, we characterized the products of enzymatic incubations with these analogs by mass spectrometry and, in cases where sufficient product could be produced, ^{1}H NMR spectroscopy.[19]

dTDP-Gal*f* donor has no effect on GlfT2 regioselectivity. The mass spectrum obtained from products isolated from incubations of GlfT2 and **11** with a 4-fold excess of acceptor **2** (ESI Fig. S3†) showed signals at *m*/*z* = 801 and 817, the expected mass for the sodium and potassium adducts of an octyl tetrasaccharide containing four Gal*f* residues. As excess acceptor **2** was used in the incubations, signals at *m*/*z* = 639 and 655 for the sodium and potassium adducts of this trisaccharide were also observed. We further analyzed the purified tetrasaccharide product by ^{1}H NMR spectroscopy (Fig. 4); a new signal at 5.03 ppm, arising from an additional anomeric proton was observed for this product when compared to **2**. This chemical shift of this signal is identical to that observed for the product resulting from incubation of GlfT2 with the natural donor **1**, and is characteristic of a β-(1→6)-linked Gal*f* residue.[19,41] Thus, the replacement of **1** with **11** does not appear to influence the regioselectivity of the reaction.

Products formed using UDP-Gal*f* analogs 4–6. The mass spectrum obtained from enzyme reactions containing 0.5 mM **2** and 2.0 mM of donor analogs **4**–**6** showed signals at *m*/*z* = 785 or *m*/*z* = 771, the expected mass for the sodium adducts of a deoxy-Gal*f* or Ara*f* containing tetrasaccharide product, respectively. In addition, no signal was observed corresponding to the starting material trisaccharide **2**, indicating it was completely consumed (Fig. 5B, ESI Fig. S5†). Similar incubations with 3-deoxy donor analog **7** also showed a signal at *m*/*z* = 785 for the tetrasaccharide product containing deoxy-Gal*f*, but additionally showed a substantial signal for the starting material **2**. This latter observation further demonstrates that **7** is a very poor GlfT2 substrate compared to the other deoxy UDP-Gal*f* analogs examined in this study. ^{1}H NMR analysis of the tetrasaccharide products resulting from each incubation with acceptor **2** revealed a new resonance at ~5.00 ppm, as would be expected for a β-D-Gal*f*-(1→6)- (**4**, **5** and **7**), or α-L-Ara*f*-(1→6)-linkage (**6**), (ESI Fig. S4†).

UDP-Gal*f* analogs result in truncated galactan polymers. Because of the polymerase activity of GlfT2, incubations with an excess of UDP-Gal*f* lead to the formation of longer galactan polymers, with the length depending on the nature of the acceptor substrate and the conditions under which the incubation is

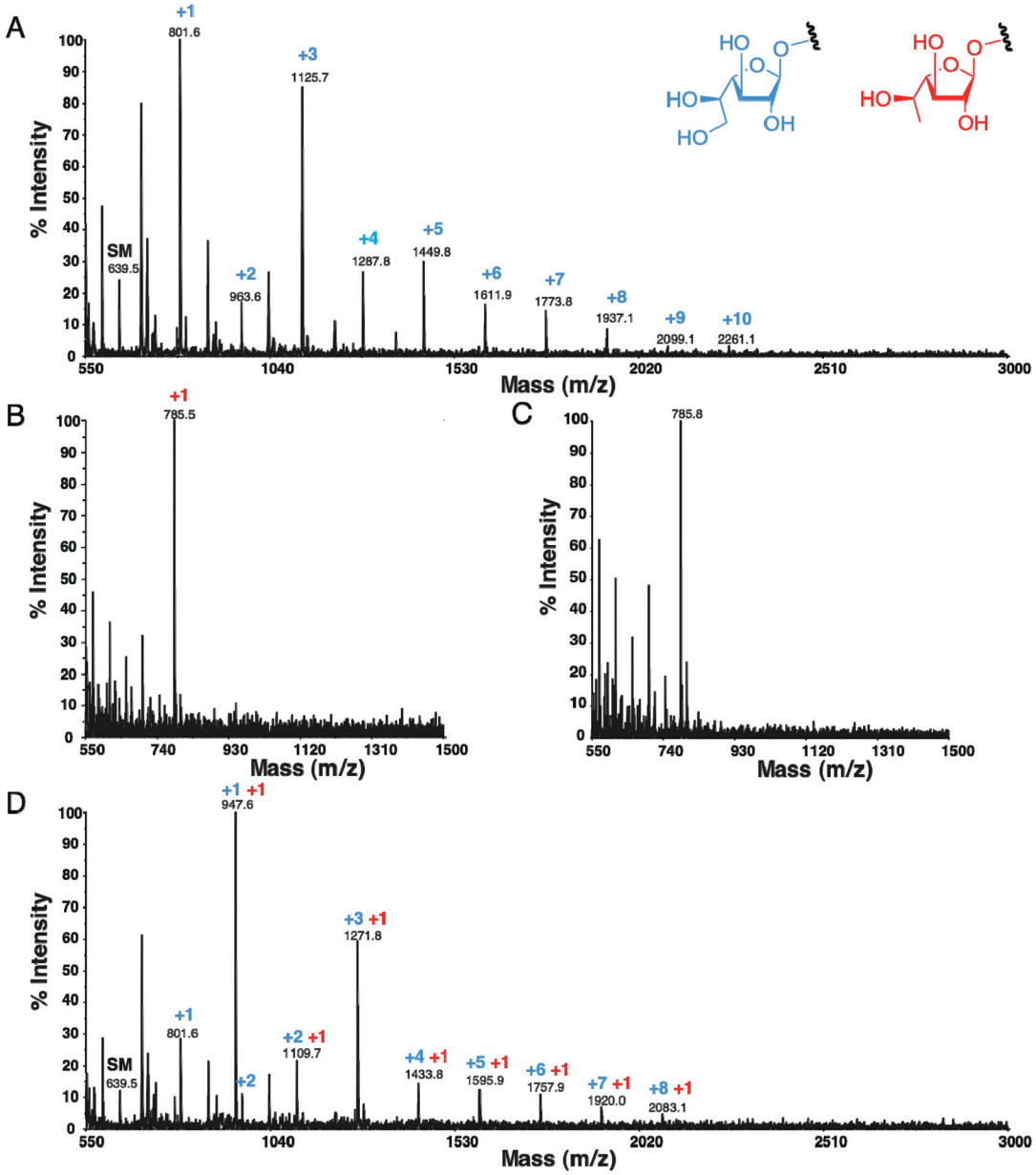


Fig. 5 UDP-Gal*f* analog results in formation of "dead end" products. Spectra are shown for MALDI MS analysis of incubations of acceptor **2** with natural donor **1** (A) or analog **4** (B) and GlfT2. The blue numbers indicate the additional Gal*f* residues added and the red numbers indicate the 6-deoxy-Gal*f* residues added. The peak at m/z = 639 corresponds to the sodium adduct of the starting trisaccharide **2**. When the product isolated from incubation B was further incubated with **1** and GlfT2, no additional products were formed (C). The mass spectrum for an incubation of acceptor **2** with both **1** and **4** and GlfT2 shows products containing 1–8 additional Gal*f* residues and a single 6d-Gal*f* residue (D).

performed.[19,24,25] In particular, the nature of the lipid aglycone used in the acceptor has been shown to influence the product distribution.[25] In our hands, incubations containing acceptors **2** or **3**, which contain an octyl aglycone, with a four-fold excess of UDP-Gal*f* resulted in polymers containing up to an additional 15 Gal*f* residues (Fig. 5A and ESI Fig. S5†), while still showing a signal for the starting trisaccharide acceptor **2** or **3**. In contrast, incubations with the deoxygenated UDP-Gal*f* analogs showed only tetrasaccharide product formation; no longer polymers were observed (Fig. 5B, ESI Fig. S5†). Similarly, incubations containing acceptor **3** and donor **4** or **6** showed only tetrasaccharide products. However, in these cases a substantial amount of starting trisaccharide **3** was observed (ESI Fig. S6†).

GlfT2 normally adds Gal*f* residues through alternating β-(1→6) and β-(1→5)-linkages. From a large scale incubation of GlfT2 with acceptor **2** and 5-deoxy analog **5** we were only able to isolate a tetrasaccharide product and no further polymerized products were observed. This tetrasaccharide product lacks the terminal C-5 hydroxyl group required for the subsequent β-(1→5)-transferase activity, but still possesses a terminal C-6 hydroxyl group. However, we observed no products resulting from subsequent β-(1→6)-transferase activity. Unexpectedly,

incubations of GlfT2 with acceptor **2** and the 6-deoxy or Ara*f* analogs **4** and **6**, whose products would posses terminal C-5 hydroxyl groups, also produced no pentasaccharide or larger products. It is possible that the reduced activity of the modified UDP-Gal*f* donors may have prevented further extension of these tetrasaccharide products. Therefore, we carried out incubations of these isolated tetrasaccharide products with GlfT2 using the natural donor substrate **1**. Again, in all cases, the mass spectrum of the products showed only signals for the starting tetrasaccharides (Fig. 5C, ESI Fig. S4†). This experiment demonstrates that these products containing a terminal Gal*f* residue lacking a C-6 or C-5 hydroxyl group, or a terminal Ara*f* residue are "dead end" products that no longer act as GlfT2 substrates. Similar dead end products were observed when testing **4**–**6** as donor substrates with crude membrane preparations of GlfT1, the first galactofuranosyltransferase involved in mycobacterial galactan biosynthesis.[43]

The finding that the 6-deoxy-Gal*f*, 5-deoxy-Gal*f* or Ara*f* terminated tetrasaccharide products were not extended, led us to postulate that these synthetic UDP-Gal*f* donor analogs could inhibit GlfT2 catalyzed galactan polymerization by forming prematurely terminated products. We tested this hypothesis by incubating GlfT2 and **2** with both the natural substrate **1** as well as synthetic analog **4** and then monitoring for the formation of truncated products by MALDI MS. As seen in Fig. 5C, the predominant product peaks observed contain a single 6-deoxy-Gal*f* residue containing between one and eight additional Gal*f* residues with the major product being a 6-deoxy-Gal*f* containing pentasaccharide. Other than 6-deoxy-Gal*f* containing products, only a small amount of tetrasaccharide and pentasaccharide were produced and no trisaccharide starting material **2** remained. The same dead end products were observed when analogous reactions were performed with **5** and **6** (data not shown). In addition, incubation of GlfT2 and acceptor **3** containing a mixture of **1** and either **4** or **6** also resulted in the formation of truncated polymers containing a single modified Gal*f* residue (ESI Fig. S5†). In all of these cases no more than a single modified Gal*f* residue was detected in the products. Combined, these results demonstrate that deoxygenated UDP-Gal*f* donors modified at C-6″ and C-5″ can be readily incorporated into a growing galactan chain through the action of GlfT2 resulting in the production of truncated products that prevent further polymerization. This demonstrates dual recognition of the acceptor substrates terminal C-5 and C-6 hydroxyl groups are essential for activity, which is consistent with the interactions proposed to be present in the active site based on X-ray crystallographic investigations of the protein. Notably, the hydroxyl group adjacent to the one undergoing glycosylation is proposed to form (apparently essential) hydrogen bonds with two amino acids on the protein.[22]

Conclusions

In this paper, we have applied two previously reported chemoenzymatic methods to synthesize of a range of deoxygenated and methylated UDP-Gal*f* analogs as well as dTDP-Gal*f* in quantities ranging from hundreds of micrograms to milligrams. With the aid of these compounds we then probed the donor binding site specificity of GlfT2, a bifunctional galactofuranosyltrasferase involved in the biosynthesis of the mycobacterial mAG complex. These results suggest that UDP-Gal*f* binds to GlfT2 in a sterically crowded region of the active site. In addition, there appear to be numerous interactions between the enzyme and the carbohydrate hydroxyl groups of the UDP-Gal*f* donor, as indicated by the reduced activity observed with deoxy donor analogs **4**–**7**. These protein–carbohydrate interactions, which are consistent with a model for donor binding proposed as part of a recent crystallographic investigation of the enzyme,[22] are not critical for initial substrate binding, as only moderate (2- to 4-fold) increases in K_M were observed in all cases. Instead, observations suggest hydrogen bonding to the Gal*f* hydroxyl groups assist to orient the Gal*f* ring for turnover.

Despite using a single active site,[20] our results show, similar to those reported earlier,[19,25] that the β-(1→6)-transferase activity is more efficient that the β-(1→5)-transferase activity of GlfT2, at least with the synthetic trisaccharide acceptors used in this study. Further analysis with acceptor analogs of different lengths would reveal whether the difference between the β-(1→5)- and β-(1→6)-transferase activity is related to the size of the acceptor substrate and these studies are currently under investigation. The difference in β-(1→5)- and β-(1→6)-transferase activity is likely controlled by subtle interactions between the enzyme and the acceptor substrate, as our results show that donor substrate binding is not substantially influenced by the nature of the glycosidic linkage being formed (*i.e.*, the same trends in donor recognition are seen for both acceptors **2** and **3**). This suggests that little if any, reconfiguration of the donor-binding pocket is involved between successive glycosylation events, which is also consistent with the processive nature of the enzyme. However, the details of these interactions are not well understood and available crystallographic data[22] does not provide significant clarity on these differences.

The donor analogs used in this study had no effect on the regioselectivity of the glycosylations catalyzed by GlfT2. The alternating regioselectivity appears to be influenced exclusively by acceptor binding interactions, which, based on these observations, are not influenced by the donor. Nevertheless, the C-6 and C-5 modified UDP-Gal*f* analogs **4**–**6** interfered with normal galactan polymerization. GlfT2 readily incorporates these analogs into a growing galactan; however, the enzyme only adds a single modified Gal*f* residue, in turn producing a product that cannot be further elongated. Similar "dead-end" products have been observed in the study of GlfT1.[43] However, this represents the first observation of these products for GlfT2. These observations demonstrate that hydrogen bonding to the non-reacting C-5 or C-6 hydroxyl group on the terminal residue of the acceptor substrate is essential for GlfT2 activity, consistent with the mechanism for substrate recognition proposed from modeling of the acceptor substrate to the recent crystal structure of GlfT2.[22]

Although the synthetic UDP-Gal*f* analogs prepared in this study have limited potential as chemotherapeutics due to their poor cell permeability, they have proven to be useful probes for studying the process of galactan biosynthesis *in vitro*. These probes have provided valuable information regarding the importance of various carbohydrate–protein interactions occurring in the active site of GlfT2, which can now be explored for the development of novel inhibitors to target galactan biosynthesis. In addition, the chain terminating C-6″ and C-5″ modified

UDP-Gal*f* analogs could serve as useful tools to elucidate the bifunctional activity of GlfT2.

Experimental details

Preparation of acceptor trisaccharides and UDP-Gal*f* donor analogs

Details for the synthesis of acceptor trisaccharides **2** and **3** are reported elsewhere.[37] UDP-Gal*f* (**1**) and UDP-Gal*f* analogs **4–6** were prepared as previously described.[34,36,43] Details for the synthesis of Gal*f*-1-phosphate analogs **13–19** are found in the ESI.†

General procedure for chemo-enzymatic synthesis of UDP-Gal*f* analogs

UDP-glucose pyrophosphorylase (GalU) and resin immobilized-galactose-1-phosphate uridyltransferase (GalPUT) were prepared as previously described.[34,35,44] To a solution of the Gal*f*-1-phosphate analog (10.5 mg, 22 μmol) in 50 mM HEPES buffer pH 8.0 containing 10 mM $MgCl_2$ and 5 mM KCl, was added UTP (12.2 mg, 20 μmol), GalU (10 U), inorganic pyrophosphatase (IPP, 2 U), and immobilized GalPUT (0.6 mL, ~15 U) for a final volume of 1 mL. The reaction was initiated by the addition of UDP-Glc (91 μg, 0.15 μmol) and incubated at ambient temperature under a N_2(g) atmosphere with gentle rotation. After 1–3 days, when analysis of the reaction by HPLC[36] indicated the complete consumption of UTP, the resin bound and soluble proteins were removed by transferring the reaction mixture to a BD column cartridge, washing with Milli-Q water (3–5 mL). The flow through was filtered using a centrifugal filter device with a molecular weight cut off of 10 000 Da. The resulting filtrates were purified by semi-preparative HPLC and gel filtration chromatography as previously described[36] to give the final UDP-Gal*f* analogs as lyophilized white powders.

Uridine 5′-diphospho-3″-deoxy-α-D-*xylo*-hexofuranose (7). (3.7 mg, 31%); ^{1}H NMR (700 MHz, D_2O, δ_H) 7.96 (d, 1 H, J = 8.1 Hz, H-6), 6.00–5.98 (m, 1 H, H-1′), 5.98 (d, 1 H, J = 8.1 Hz, H-5), 5.60 (dd, 1 H, J = 5.3, 4.3 Hz, H-1″), 4.39–4.36 (m, 2 H, H-2′, H-3′), 4.35–4.31 (m, 1 H, H-2″), 4.30–4.28 (m, 1H, H-4′), 4.25 (ddd, 1 H, J = 11.8, 4.5, 2.6 Hz, H-5′a), 4.21 (ddd, 1 H, J = 11.8, 5.7, 2.8 Hz, H-5′b), 4.10 (ddd, 1 H, J = 9.9, 6.5, 6.5 Hz, H-4″), 3.70 (ddd, 1 H, J = 6.8, 6.5, 3.7 Hz, H-5″), 3.66 (dd, 1 H, J = 12.0, 3.7 Hz, H-6″a), 3.56 (dd, 1 H, J = 12.0, 6.8 Hz, H-6″b), 2.31–2.29 (m, 1 H, H-3″a), 1.83 (app. q, 1 H, J = 10.9, H-3″b); ^{13}C NMR (175 MHz, D_2O, δ_C) 167.1 (C-4), 152.7 (C-2), 142.5 (C-5), 103.5 (C-6), 98.5 (d, 1 C, J = 6.0 Hz, C-1″), 89.2 (C-1′), 84.1 (d, 1 C, J = 9.1 Hz, C-4′), 79.7 (C-4″), 75.4 (C-5″), 74.6, 70.5 (C-2′, C-3′), 72.6 (d, 1 C, J = 8.2 Hz, C-2″), 65.7 (d, 1 C, J = 5.3 Hz, C-5′), 63.2 (C-6″), 31.5 (C-3″); MS (ESI) m/z 549 ([M − H]$^-$, 30%), 274 ([M − 2H]$^{2-}$, 100%); HRMS (ESI) m/z Calcd for (M − H)$^-$ $C_{15}H_{23}N_2O_{16}P_2$: 549.0528. Found: 549.0529.

Uridine 5′-diphospho-6″-*O*-methyl-α-D-galactofuranose (8). (0.5 mg, <5%); ^{1}H NMR (700 MHz, D_2O, δ_H) 7.96 (d, 1 H, J = 8.1 Hz, H-6), 6.00–5.98 (m, 1 H, H-1′), 5.98 (d, 1 H, J = 8.1 Hz, H-5), 5.64 (dd, 1 H, J = 5.2, 4.2 Hz, H-1″), 4.38–4.36 (m, 2 H, H-2′, H-3′), 4.30–4.27 (m, 1H, H-4′), 4.24 (ddd, 1 H, J = 11.8, 4.2, 2.6 Hz, H-5′a), 4.22–4.18 (m, 2 H, H-5′b, H-3″), 4.15 (ddd, 1 H, J = 8.3, 4.2, 2.3 Hz, H-2″), 3.90–3.87 (m, 1 H, H-5″), 3.80 (dd, 1 H, J = 7.0, 6.1 Hz, H-4″), 3.60 (dd, 1 H, J = 10.8, 3.9 Hz, H-6″a), 3.53 (dd, 1 H, J = 10.8, 7.2 Hz, H-6″b), 3.39 (s, 3 H, OC*H*$_3$); MS (ESI) m/z 579 ([M − H]$^-$, 5.6%), 289 ([M − 2H]$^{2-}$, 100%); HRMS (ESI) m/z Calcd for (M − 2H)$^{2-}$ $C_{16}H_{24}N_2O_{17}P_2$: 289.0281. Found 289.0281.

Uridine 5′-diphospho-5″-*O*-methyl-α-D-galactofuranose (9). (0.5 mg, <5%); ^{1}H NMR (700 MHz, D_2O, δ_H) 7.96 (d, 1 H, J = 8.1 Hz, H-6), 6.00–5.98 (m, 1 H, H-1′), 5.98 (d, 1 H, J = 8.1 Hz, H-5), 5.67 (dd, 1 H, J = 6.7, 3.8 Hz, H-1″), 4.39–4.37 (m, 2 H, H-2′, H-3′), 4.29–4.28 (m, 1 H, H-4′), 4.24 (ddd, 1 H, J = 11.8, 4.4, 2.6 Hz, H-5′a), 4.20 (ddd, 1 H, J = 11.8, 5.6, 2.8 Hz, H-5′a), 4.16–4.12 (m, 2 H, H-2″, H-3″), 3.88–3.84 (m, 2 H, H-4″, H-6″a), 3.62 (dd, 1 H, J = 12.5, 5.8 Hz, H-6″b), 3.55 (s, 3 H, OC*H*$_3$), 3.55–3.50 (m, 1 H, H-5″); ^{13}C NMR (175 MHz, D_2O, δ_C) 167.1 (C-4), 142.5 (C-5), 103.5 (C-6), 98.7 (d, J = 5.6 Hz, C-1″), 89.2 (C-1′), 84.2 (C-5″), 84.1 (d, J = 8.9 Hz, H-4′), 82.1 (C-4″), 77.6 (d, J = 7.1 Hz, H-2″), 75.1 (C-3″), 74.6, 70.5 (C-2′, C-3′), 65.7 (d, J = 5.9 Hz, H-5′), 60.2 (C-6″), 59.4 (O*C*H$_3$); MS (ESI) m/z 579 ([M − H]$^-$, 18%), 289 ([M − 2H]$^{2-}$, 100%); HRMS (ESI) m/z Calcd for (M − 2H)$^{2-}$ $C_{16}H_{24}N_2O_{17}P_2$: 289.0281. Found: 289.0281.

Uridine 5′-diphospho-2″-*O*-methyl-α-D-galactofuranose (10). (0.03 mg, <5%); ^{1}H NMR (700 MHz, D_2O, δ_H) 7.95 (d, 1 H, J = 8.1 Hz, H-6), 5.98–5.97 (m, 1 H, H-1′), 5.97 (d, 1 H, J = 8.1 Hz, H-5), 5.78 (dd, 1 H, J = 5.5, 4.2 Hz, H-1″), 4.38–4.35 (m, 2 H, H-2′, H-3′), 4.28–4.26 (m, 2 H, H-4′, H-3″), 4.23 (ddd, 1 H, J = 11.8, 4.5, 2.6 Hz, H-5′a), 4.19 (ddd, 1 H, J = 11.8, 5.7, 2.9 Hz, H-5′a), 3.94 (ddd, 1 H, J = 8.5, 4.1, 2.6 Hz, H-2″), 3.82 (dd, 1 H, J = 7.4, 5.2 Hz, H-4″), 3.77–3.73 (m, 1 H, H-5″), 3.69 (dd, 1 H, J = 11.8, 4.4 Hz, H-6″a), 3.62 (dd, 1 H, J = 11.8, 7.2 Hz, H-6″b), 3.49 (s, 3 H, OC*H*$_3$); HRMS (ESI) m/z Calcd for (M − H)$^-$ $C_{16}H_{25}N_2O_{17}P_2$: 579.0634. Found: 579.0634.

Chemo-enzymatic synthesis of dTDP-Gal*f*

Resin immobilized Cps2L protein was prepared from *E. coli* BL21 (DE3) cells containing recombinant pSK001 plasmid.[50] Cells were grown in LB broth (1 L) supplemented with 25 μg mL^{-1} kanamycin. Production of Cps2L was induced by the addition of 375 μM IPTG (isopropyl 1-thio-β-D-galactopyranoside) at a OD_{600} of 0.6 followed by incubation at 30 °C for 4 h. Cells were collected by centrifugation at 11 300 × g_{max} for 15 min and the pellets were then re-suspended in 40 mL of resuspension buffer (20 mM Tris-HCl, pH 8.0, containing 300 mM NaCl and 10 mM imidazole). The cells were lysed using a benchtop cell disruptor (Constant Systems Inc., NC) set to 20 kpsi and the lysate clarified by centrifugation (105 000 × g for 1 h at 4 °C). The lysate was applied to a 5 mL Ni–NTA agarose column and washed with 6 column volumes of wash buffer (20 mM Tris-HCl, pH 8.0, containing 300 mM NaCl and 25 mM imidazole), followed by 6 column volumes of reaction buffer (20 mM Tris-HCl, pH 8.0, containing 300 mM NaCl).

To a solution of Gal*f*-1-phosphate (5 mg, 11 μmol) in reaction buffer containing 2.5 mM $MgCl_2$ was added dTTP (4.8 mg,

10 μmol), IPP (3.75 U), and immobilized Csp2L protein (0.5 mL) for a final volume of 0.6 mL. After incubating for 48 h at ambient temperature under a atmosphere of N_2(g) with gentle rotation, analysis of the reaction by HPLC indicated complete consumption of dTTP. The reaction was incubated for 5 h with alkaline phosphatase (AP, 10 U) to degrade unwanted dTDP and dTMP in the reaction mixture. The resin bound Cps2L was again removed by transferring the reaction mixture to a BD column cartridge and washing with Milli-Q water (3–5 mL). Soluble IPP and AP proteins were removed by filtration of the resulting flow through using a centrifugal filter device with a molecular weight cut off of 10 000 Da. The filtrate was purified by reverse phase semi-preparative HPLC using conditions previously described for the purification of UDP-Gal*f*.[34,36] Purified HPLC fractions were combined, the volume was reduced to 5 mL by evaporation under reduced pressure, and the salts were removed by gel filtration chromatography (Sephadex G-15) eluting with Milli-Q water at a flow rate of 1 mL per minute. Fractions containing the purified product were combined and lyophilized to give dTDP-Gal*f* **11** as a white powder (3.4 mg, 50%); ^{1}H NMR (600 MHz, D_2O, δ_H) 7.72 (br d, 1 H, J = 1.2 Hz, H-6), 6.33 (dd, 1 H, J = 7.5, 6.5 Hz, H-1′), 5.61 (dd, 1 H, J = 5.5, 4.5 Hz, H-1″), 4.60 (app. dt, 1 H, J = 5.9, 3.0 Hz, H-3′), 4.21 (dd, 1 H, J = 8.4, 7.4 Hz, H-3″), 4.18–4.14 (m, 3H, H-4′, H-5′a, H-5′b), 4.12 (ddd, 1 H, J = 8.4, 4.3, 2.4 Hz, H-2″), 3.80 (dd, 1 H, J = 7.4, 5.2 Hz, H-4″), 3.75 (app. dt, 1 H, J = 7.2, 4.8 Hz, H-5″), 3.69 (dd, 1 H, J = 11.8, 4.3 Hz, H-6″a), 3.61 (dd, 1 H, J = 11.8, 7.2 Hz, H-3″), 2.39–2.32 (m, 2 H, H-2′a, H-2′b); MS (ESI) m/z 549 ($[M - H]^-$, 47%), 274 ($[M - 2H]^{2-}$, 100%); HRMS (ESI) m/z Calcd for $(M - H)^-$ $C_{16}H_{25}N_2O_{16}P_2$: 563.0685. Found: 563.0684.

GlfT2 activity, kinetics, and inhibition

The GlfT2 protein was prepared as previously described, and its activity was determined using the coupled spectrophotometric assay reported previously.[34] Assays were performed in 384 well microtiter plates in a volume of 40 μL containing 100 mM MOPS, pH 7.6, 50 mM KCl, 20 mM $MgCl_2$, 1.1 mM NADH, 3.5 mM phosphoenolpyruvate (PEP), 7.5 U pyruvate kinase (PK, EC 2.7.1.40), 16.8 U lactate dehydrogenase (LDH, EC 1.1.1.27), 2 mM acceptor trisaccharide **2** or **3**, and 2 mM donor substrate (**1, 4–11**). The amount of GlfT2 added was controlled to allow for sufficient substrate turnover. Assays were continuously monitored at 37 °C over 20 min and initial velocities were determined from the decrease in NADH absorbance at 340 nm. Specific activities were determined in duplicate for all donor analogs.

Kinetic values were determined by varying the concentration of donor analog (**1, 4–11**) between 0 and 4000 μM while keeping the concentration of acceptor trisaccharide **2** fixed at 2 mM. At this concentration the acceptor trisaccharide is saturating (10 times the K_M reported for **2**),[19] allowing for single substrate kinetics of the donor analogs to be measured. Assays were run in duplicate and initial velocities were determined for each substrate concentration. Kinetic parameters K_M and k_{cat} were obtained by nonlinear regression analysis of the Michaelis–Menten equation using GraphPad PRISM 4 (GraphPad Software, San Diego, CA).

GlfT2 percent inhibition was determined using the coupled spectrophotometric assay with 1250 μM donor analog (**5–10**), and acceptor trisaccharide **2** and UDP-Gal*f* donor **1** concentration fixed at 2000 μM and 375 μM respectively. Donor analogs showing >50% inhibition were further evaluated. Inhibition kinetics were determined by varying the concentration of donor **1** between 0–2500 μM while the concentration of acceptor trisaccharide **2** fixed at 2 mM. Assays were performed at varying concentrations of donor analog (0–1000 μM). The inhibition constant K_i was determined by nonlinear regression analysis using GraphPad PRISM 4 software.

Isolation and characterization of UDP-Gal*f* (1) and dTDP-Gal*f* (11) reaction products

Reactions containing 50 mM MOPS pH 7.6, with 20 mM $MgCl_2$, 500 μM donor **1** or **11**, 3000 μM trisaccharide acceptor **2**, and 100 μg GlfT2 in a total volume of 400 μL were incubated at ambient temperature under a nitrogen gas atmosphere for 4 days with gentle rotation. To ensure only singly glycosylated products were produced, a six-fold excess of trisaccharide **2** was used. Progress of the enzymatic reactions was monitored by thin-layer chromatography (TLC) on SilicaPlate TLC silica gel plates (Silicycle) eluting with $CHCl_3$–CH_3OH–NH_4OH–H_2O (65 : 25 : 0.5 : 3.6) as previously described.[19] Reaction products on TLC were visualized using 3% anisaldehyde in sulfuric acid stain. After 4 days the reactions were diluted to 1 mL with Milli-Q water, filtered through 0.22 μm Millex-GV filters, and 200 μL of the filtrate was lyophilized for MALDI MS analysis using a Voyager Elite time-of-flight spectrometer in positive ion mode. Preparative TLC was used to purify the products from the remaining 800 μL of filtrate. The silica from the area of TLC plate corresponding to the reaction product (R_f 0.31) was scraped from the plate without visualization using the 3% anisaldehyde–sulfuric acid stain. HPLC grade methanol (4 mL) was used to extract the purified reaction product, the silica was filtered, and the methanol was evaporated. The resulting residue was re-suspended in Milli-Q water and passed through a 0.22 μm Millex-GV filter, the filtrate lyophilized, resuspended in D_2O (1 mL) and again lyophilized. Products were dissolved in 700 μL D_2O and one-dimensional ^{1}H NMR spectra were recorded on a Varian i600 instrument with suppression of the HOD signal using a presaturation pulse sequence, irradiating at 4.67 ppm.

Isolation and characterization UDP-Gal*f* analog reaction products

Reactions containing 50 mM MOPS pH 7.6 with 20 mM $MgCl_2$, 5 mM β-mercaptoethanol, 2 mM donor analog **4–7**, 500 μM trisaccharide acceptor **2** or **3**, 50 μg GlfT2, and 2 units of alkaline phosphatase (AP) were incubated under a nitrogen gas atmosphere at ambient temperature for 3 days with gentle rotation. To promote production of polymeric products a four-fold excess of donor was used; also, AP was added to degrade the UDP by-product produced during the reaction, which is known to inhibit GlfT2. Reaction progress was monitored by TLC, again eluting with $CHCl_3$–CH_3OH–NH_4OH–H_2O (65 : 25 : 0.5 : 3.6). After 3 days the reaction products were

purified using a Sep-Pak C_{18} cartridge. After washing with ~10 mL of water to remove the enzyme and unreacted donor, the reaction products were eluted using 4 mL of HPLC grade CH_3OH. The solvent was then evaporated; the products were resuspended in 1 mL water and passed through a 0.22 μm Millex-GV filter. From this solution, 100 μL was lyophilized and re-suspended in 2,5-dihydroxy benzoic acid and characterized by MALDI MS as described above. The remaining 900 μL of the extraction solution was lyophilized and re-suspended in 600 μL D_2O. One-dimensional ^1H NMR spectra were recorded on a Varian i600 instrument with the presaturation of the HOD signal.

Incubation of UDP-Gal*f* analogs (4–6) reaction products with UDP-Gal*f*. The purified and lyophilized reaction products from incubations of GlfT2 with acceptor **2** and UDP-Gal*f* analogs **4–6** were re-suspended in 50 mM MOPS pH 7.6 with 20 mM $MgCl_2$, 5 mM β-mercaptoethanol, 2 mM UDP-Gal*f* **1**, 50 μg GlfT2, and 2 units of alkaline phosphatase (AP) in a final volume of 100 μL. The reaction mixtures were incubated under a nitrogen gas atmosphere at ambient temperature for 1 day. The products were purified and analyzed by MALDI MS as described above.

Inhibition of galactan polymerization by UDP-Gal*f* analogs (4–6). Reactions containing 50 mM MOPS pH 7.6 with 20 mM $MgCl_2$, 5 mM β-mercaptoethanol, 2 mM donor analog **4–6**, 1 mM UDP-Gal*f* **1**, 500 μM trisaccharide acceptor **2** or **3**, 50 μg GlfT2, and 2 units of alkaline phosphatase (AP) in 100 μL final volume, were incubated under a nitrogen gas atmosphere at ambient temperature for 3 days with gentle rotation. The products were purified and analyzed by MALDI MS as described above.

Acknowledgements

This work was supported by the Alberta Glycomics Centre and the Natural Sciences and Engineering Research Council of Canada. MBP is the recipient of a Studentship award from Alberta Innovates–Technology Futures and an Alexander Graham Bell Canada Graduate Scholarship from the Natural Sciences and Engineering Research Council of Canada. The authors would like to acknowledge Ruixiang Blake Zheng for assistance with the coupled spectrophotometric Glf2 assay.

References

1 M. R. Richards and T. L. Lowary, *ChemBioChem*, 2009, **10**, 1920–1938.
2 P. Peltier, R. Euzen, R. Daniellou, C. Nugier-Chauvin and V. Ferrières, *Carbohydr. Res.*, 2008, **343**, 1897–1923.
3 W. F. Paolo and J. D. Nosanchuk, *Lancet Infect. Dis.*, 2004, **4**, 287–293.
4 P. D. O. Davies, *Ann. Med.*, 2003, **35**, 235–243.
5 R. J. Coker, *Trop. Med. Int. Health*, 2004, **9**, 25–40.
6 E. D. Chan and M. D. Iseman, *Curr. Opin. Infect. Dis.*, 2008, **21**, 587–595.
7 L. L. Pedersen and S. J. Turco, *Cell. Mol. Life Sci.*, 2003, **60**, 259–266.
8 P. J. Brennan and H. Nikaido, *Annu. Rev. Biochem.*, 1995, **64**, 29–63.
9 P. J. Brennan, *Tuberculosis*, 2003, **83**, 91–97.
10 D. C. Crick, S. Mahapatra and P. J. Brennan, *Glycobiology*, 2001, **11**, 107R–118R.
11 V. Jarlier and H. Nikaido, *J. Bacteriol.*, 1990, **172**, 1418–1423.
12 F. E. Umesiri, A. K. Sanki, J. Boucau, D. R. Ronning and S. J. Sucheck, *Med. Res. Rev.*, 2010, **30**, 290–326.
13 R. A. Slayden, R. E. Lee and C. E. Barry, *Mol. Microbiol.*, 2000, **38**, 514–525.
14 A. E. Belanger, G. S. Besra, M. E. Ford, K. Mikusova, J. T. Belisle, P. J. Brennan and J. M. Inamine, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 11919–11924.
15 M. Beláňová, P. Dianišková, P. J. Brennan, G. C. Completo, N. L. Rose, T. L. Lowary and K. Mikušová, *J. Bacteriol.*, 2008, **190**, 1141–1145.
16 A. Weston, R. J. Stern, R. E. Lee, P. M. Nassau, D. Monsey, S. L. Martin, M. S. Scherman, G. S. Besra, K. Duncan and M. R. McNeil, *Tubercle Lung Dis.*, 1998, **78**, 123–131.
17 L. J. Alderwick, L. G. Dover, N. Veerapen, S. S. Gurcha, L. Kremer, D. L. Roper, A. K. Pathak, R. C. Reynolds and G. S. Besra, *Protein Expression Purif.*, 2008, **58**, 332–341.
18 L. Kremer, L. G. Dover, C. Morehouse, P. Hitchin, M. Everett, H. R. Morris, A. Dell, P. J. Brennan, M. R. McNeil, C. Flaherty, K. Duncan and G. S. Besra, *J. Biol. Chem.*, 2001, **276**, 26430–26440.
19 N. L. Rose, G. C. Completo, S. J. Lin, M. McNeil, M. M. Palcic and T. L. Lowary, *J. Am. Chem. Soc.*, 2006, **128**, 6721–6729.
20 M. G. Szczepina, R. B. Zheng, G. C. Completo, T. L. Lowary and B. M. Pinto, *ChemBioChem*, 2009, **10**, 2052–2059.
21 J. F. May, M. R. Levengood, R. A. Splain, C. D. Brown and L. L. Kiessling, *Biochemistry*, 2012, **51**, 1148–1159.
22 R. W. Wheatley, R. B. Zheng, T. L. Lowary and K. K. S. Ng, *J. Biol. Chem. (in revision)*.
23 A. K. Pathak, V. Pathak, L. Seitz, J. A. Maddry, S. S. Gurcha, G. S. Besra, W. J. Suling and R. C. Reynolds, *Bioorg. Med. Chem.*, 2001, **9**, 3129–3143.
24 M. R. Levengood, R. A. Splain and L. L. Kiessling, *J. Am. Chem. Soc.*, 2011, **133**, 12758–12766.
25 J. F. May, R. A. Splain, C. Brotschi and L. L. Kiessling, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 11851–11856.
26 J. Frigell, J. A. Pearcey, T. L. Lowary and I. Cumpstey, *Eur. J. Org. Chem.*, 2011, 1367–1375.
27 S. Cren, S. S. Gurcha, A. J. Blake, G. S. Besra and N. R. Thomas, *Org. Biomol. Chem.*, 2004, **2**, 2418–2420.
28 S. Cren, C. Wilson and N. R. Thomas, *Org. Lett.*, 2005, **7**, 3521–3523.
29 A. E. Trunkfield, S. S. Gurcha, G. S. Besra and T. D. H. Bugg, *Bioorg. Med. Chem.*, 2010, **18**, 2651–2663.
30 K. Vembaiyan, J. A. Pearcey, M. Bhasin, T. L. Lowary and W. Zou, *Bioorg. Med. Chem.*, 2011, **19**, 58–66.
31 U. Spohr, O. Hindsgaul and R. U. Lemieux, *Can. J. Chem.*, 1985, **63**, 2644–2652.
32 O. Hindsgaul, K. J. Kaur, G. Srivastava, M. Blaszczykthurin, S. C. Crawley, L. D. Heerze and M. M. Palcic, *J. Biol. Chem.*, 1991, **266**, 17858–17862.
33 S. Laferte, N. W. C. Chan, K. Sujino, T. L. Lowary and M. M. Palcic, *Eur. J. Biochem.*, 2000, **267**, 4840–4849.
34 N. L. Rose, R. B. Zheng, J. Pearcey, R. Zhou, G. C. Completo and T. L. Lowary, *Carbohydr. Res.*, 2008, **343**, 2130–2139.
35 J. C. Errey, B. Mukhopadhyay, K. P. R. Kartha and R. A. Field, *Chem. Commun.*, 2004, 2706–2707.
36 M. B. Poulin and T. L. Lowary, *Methods Enzymol.*, 2010, **478**, 389–411.
37 G. C. Completo and T. L. Lowary, *J. Org. Chem.*, 2008, **73**, 4513–4525.
38 A. K. Pathak, G. S. Besra, D. Crick, J. A. Maddry, C. B. Morehouse, W. J. Suling and R. C. Reynolds, *Bioorg. Med. Chem.*, 1999, **7**, 2407–2413.
39 Y. Bai and T. L. Lowary, *J. Org. Chem.*, 2006, **71**, 9658–9671.
40 N. R. Williams, *Adv. Carbohydr. Chem. Biochem.*, 1970, **25**, 109–179.
41 N. Cyr and A. S. Perlin, *Can. J. Chem.*, 1979, **57**, 2504–2511.
42 R. M. de Lederkremer, V. B. Nahmad and O. Varela, *J. Org. Chem.*, 1994, **59**, 690–692.
43 P. Peltier, M. Beláňová, P. Dianišková, R. Zhou, R. B. Zheng, J. A. Pearcey, M. Joe, P. J. Brennan, C. Nugier-Chauvin, V. Ferrières, T. L. Lowary, R. Daniellou and K. Mikušová, *Chem. Biol.*, 2010, **17**, 1356–1366.
44 Z. Y. Liu, J. B. Zhang, X. Chen and P. G. Wang, *ChemBioChem*, 2002, **3**, 348–355.
45 P. Peltier, J. P. Guegan, R. Daniellou, C. Nugier-Chauvin and V. Ferrières, *Eur. J. Org. Chem.*, 2008, 5988–5994.
46 A. L. Marlow and L. L. Kiessling, *Org. Lett.*, 2001, **3**, 2517–2519.
47 Q. B. Zhang and H. W. Liu, *J. Am. Chem. Soc.*, 2000, **122**, 9065–9070.
48 Y. E. Tsvetkov and A. V. Nikolaev, *J. Chem. Soc., Perkin Trans. 1*, 2000, 889–891.
49 P. Peltier, R. Daniellou, C. Nugier-Chauvin and V. Ferrières, *Org. Lett.*, 2007, **9**, 5227–5230.

50 S. C. Timmons, R. H. Mosher, S. A. Knowles and D. L. Jakeman, *Org. Lett.*, 2007, **9**, 857–860.
51 S. C. Timmons, J. P. M. Hui, J. L. Pearson, P. Peltier, R. Daniellou, C. Nugier-Chauvin, E. C. Soo, R. T. Syvitski, V. Ferrieres and D. L. Jakeman, *Org. Lett.*, 2008, **10**, 161–163.
52 I. Neverova, C. H. Scaman, O. P. Srivastava, R. Szweda, I. K. Vijay and M. M. Palcic, *Anal. Biochem.*, 1994, **222**, 190–195.
53 M. G. Szczepina, R. X. B. Zheng, G. C. Completo, T. L. Lowary and B. M. Pinto, *Bioorg. Med. Chem.*, 2010, **18**, 5123–5128.
54 M. A. Macnaughtan, M. Kamar, G. Alvarez-Manilla, A. Venot, J. Glushka, J. M. Pierce and J. H. Prestegard, *J. Mol. Biol.*, 2007, **366**, 1266–1281.
55 J. Angulo, B. Langpap, A. Blume, T. Biet, B. Meyer, N. R. Krishna, H. Peters, M. M. Palcic and T. Peters, *J. Am. Chem. Soc.*, 2006, **128**, 13529–13538.