nature structural & molecular biology

Structural dissection and high-throughput screening of mannosylglycerate synthase

James Flint[1], Edward Taylor[2], Min Yang[3], David N Bolam[1], Louise E Tailford[1], Carlos Martinez-Fleites[2], Eleanor J Dodson[2], Benjamin G Davis[3], Harry J Gilbert[1] & Gideon J Davies[2]

The enzymatic transfer of activated mannose yields mannosides in glycoconjugates and oligo- and polysaccharides. Yet, despite its biological necessity, the mechanism by which glycosyltransferases recognize mannose and catalyze its transfer to acceptor molecules is poorly understood. Here, we report broad high-throughput screening and kinetic analyses of both natural and synthetic substrates of *Rhodothermus marinus* mannosylglycerate synthase (MGS), which catalyzes the formation of the stress protectant 2-*O*-α-D-mannosyl glycerate. The sequence of MGS indicates that it is at the cusp of inverting and retaining transferases. The structures of apo MGS and complexes with donor and acceptor molecules, including GDP-mannose, combined with mutagenesis of the binding and catalytic sites, unveil the mannosyl transfer center. Nucleotide specificity is as important in GDP-D-mannose recognition as the nature of the donor sugar.

Mannose-containing oligo- and polysaccharides and glycoconjugates are widespread in nature. Gluco- and galactomannans are key components of the plant cell wall, while diverse mannans are also integral to both fungal and bacterial cell walls; consequently, their recognition is central to the innate immune response. In higher organisms, both *O*- and *N*-linked oligosaccharides of glycosylated proteins contain mannose, and therefore defects in mannosyl transfer lead to a number of different human congenital disorders. Because of the key and diverse roles of mannose in complex bioactive molecules, there is considerable interest in the enzyme-catalyzed transfer of the sugar from its activated donor species to a raft of acceptors. Enzymes that catalyze mannose transfer not only represent important therapeutic targets but also are of considerable industrial importance, as mannose-containing polymers have many biotechnological applications.

Mannose transfer is catalyzed by activated sugar–dependent glycosyltransferases (GTs). GTs, although classified into 77 (as of 25 March 2005) sequence-derived GT families[1] comprising >14,000 members, remain a poorly characterized enzyme class, with representative structures for just 16 families. These structures have shown just two protein scaffolds (and variants[2] thereof), termed 'fold families' GT-A and GT-B, following their original observation in the *Bacillus subtilis* SpsA[3] and bacteriophage T4 DNA β-glucosyltransferase structures[4], respectively. In the context of mannosyl transfer, the structure of Mnt1p, which uses GDP-mannose (GDP-Man) as a donor and has the GT-A fold, has been recently reported[5].

MGS catalyzes the formation of 2-*O*-α-D-mannosylglycerate from GDP-Man and D-glycerate. This glycocarboxylic acid is important for protection against temperature and osmotic stress in hyperthermophilic bacteria such as *R. marinus*[6,7], while also mediating protein stabilization *in vitro*[8,9]. Although MGS, at the sequence level, is most similar to family-GT2 inverting GTs (with up to 30% identity in the N-terminal domain), MGS is a retaining GT; the newly formed glycosidic linkage has the same α-configuration as the donor GDP-α-D-Man. MGS is thus a mechanistic oddity whose sequence lies at the cusp of inverting and retaining transferases, and as such it presents an alluring system in which to study mannosyl transfer.

To investigate the mechanism of MGS here, we exploited a high-throughput broad screen for potential donors and acceptors, which showed notable plasticity in substrate donor and acceptor recognition. The three-dimensional structure of MGS in complex with both substrates and products, as well as the biochemical properties of active site variants, reveals the structural basis for GDP-Man and glycerate recognition and subsequent mannosyl transfer. The data emphasize the capacity of GTs to use a variety of sugar donors and acceptors, suggesting their potential exploitation in the synthesis of new glycoconjugates both *in vitro* and *in vivo*.

RESULTS

Donor and acceptor promiscuity of MGS

MGS kinetics, using the natural substrate pair GDP-Man as donor and D-glycerate as acceptor, were studied via mass spectrometry to determine the rate of product formation[10]. NMR analysis of the product formed confirmed its identity as α-D-mannosyl-D-glycerate (**Supplementary Methods** online), demonstrating that catalysis by MGS occurs with retention of anomeric configuration, as implied previously[11]. MGS is inactive in the presence of EDTA and has an

[1]Institute for Cell and Molecular Biosciences, University of Newcastle upon Tyne, The Medical School, Newcastle upon Tyne NE2 4HH, UK. [2]York Structural Biology Laboratory, Department of Chemistry, University of York, York, YO10 5YW, UK. [3]Department of Chemistry, The University of Oxford, Mansfield Road, Oxford OX1 3TA, UK. Correspondence should be addressed to G.J.D. (davies@ysbl.york.ac.uk).

Published online 12 June 2005; doi:10.1038/nsmb950

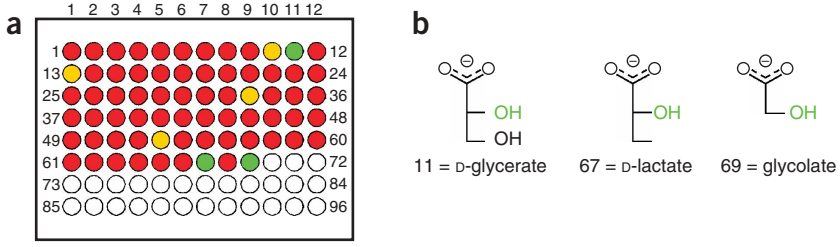

Figure 1 High-throughput screening of GT specificity. (**a**) An example of GAR[10] screening of MGS acceptors (colored by analogy with DNA microarray data, with green indicating a reaction with significant signal-to-noise ratio). In the case of MGS, this allowed the ready identification of five donors (GDP-Man, UDP-Man, GDP-Glc, UDP-Glc and GDP-L-Fuc) and three acceptors, 11, 67 and 69, whose full kinetics were subsequently determined. (**b**) Chemical structures of the MGS acceptors identified by the GAR screen.

absolute requirement for divalent metal ions, including a notable preference for 'harder' metals, with 200 μM of Mg^{2+}, Ca^{2+}, Mn^{2+}, Ni^{2+} and Co^{2+} giving relative rates of 1.0, 0.6, 0.33, 0.16 and 0.15, respectively. Bisubstrate (that is, exploring changes in both donor and acceptor) kinetics at 25 °C in the presence of Ca^{2+} yield a k_{cat} of 1.1 ± 0.3 s^{-1}, with a $K_{m(GDP\text{-}Man)}$ of 81 ± 39 μM^{-1} and a $K_{m(glycerate)}$ of 123 ± 68 μM^{-1}, and with corresponding $K_{I(GDP\text{-}Man)}$ and $K_{I(glycerate)}$ values of 24 ± 3 and 36 ± 3 μM^{-1}, respectively. The low ratio of $K_{m(GDP\text{-}Man)}$ to $K_{m(glycerate)}$ may indicate a random bi-bi mechanism (see discussion in ref. 10).

The high-throughput, broad green-amber-red (GAR)[10] screening approach was used to identify new MGS donors and acceptors. This screen uses liquid chromatography mass spectrometry (LCMS) to screen a large donor and acceptor library (**Supplementary Fig. 1** online) in multiwell format trays. A virtual color of red (no reaction), amber (total ion count (TIC) signal-to-noise > 1) or green (TIC signal-to-noise > 10) is assigned, reflecting the presence of selected ions of mass $n + y$, where y is the mass of the glycosyl moiety transferred. In the case of MGS, the screen (with GDP-Man as the donor) showed flexibility at C3 of the acceptor, with D-glycerate, D-lactate and glycolate all acting as substrates (**Fig. 1**), whereas none of the other 66 molecules assessed could be conjugated to donor sugars by MGS, including glycerol, D-serine and the L-stereoisomers of glycerate and lactate. MGS shows notable donor flexibility and can use α-GDP-D-Man, α-GDP-D-Glc, β-GDP-L-fucose (β-GDP-L-Fuc), α-UDP-D-Man and α-UDP-D-Glc, but not Man-1-P, as substrates. Bisubstrate kinetics at 25 °C with UDP-Glc as the donor yielded a k_{cat} of 0.072 ± 0.009 s^{-1}, with a $K_{A(UDP\text{-}Glc)}$ of 104 ± 14 μM^{-1} and a $K_{B(glycerate)}$ of 80 ± 14 μM^{-1}, and with corresponding $K_{IA(UDP\text{-}Glc)}$ and $K_{IB(glycerate)}$ values of 17 ± 1 and 13 ± 1 μM^{-1}, respectively.

Pseudo single-substrate kinetics were carried out for the new donor and acceptor pairs (**Table 1**). At 25 °C, GDP-linked sugars are transferred about tenfold more efficiently than UDP-linked sugars, and this indicates substantial specificity for the purine nucleotide. With GDP-linked sugars, MGS shows a two- and fivefold preference for D-Man over D-Glc and L-Fuc, respectively. At 65 °C, the k_{cat}/K_m ratio for GDP-Man increases ~6-fold, whereas that for GDP-Glc increases only 1.6-fold, a result suggesting that in the natural milieu the enzyme would be a more specific mannosyltransferase. Full characterization, including NMR analysis of the reaction product generated from GDP-L-Fuc (**Supplementary Methods**), confirms that this donor sugar is transferred to glycerate to form 2-*O*-β-L-fucosyl-D-glycerate. MGS is ~70-fold more active (based on k_{cat}/K_m) when glycerate is the acceptor compared with D-lactate. These data indicate that the carboxylate makes critical interactions with the enzyme, whereas O3 contributes substantially to productive binding. MGS can clearly harness a range of nucleotides, donor sugars and acceptors. These observations add to the emerging picture that GTs are not as specific as at one time assumed and can thus be exploited in the synthesis of new glycoconjugates. The structure of MGS in complex with various ligands was subsequently solved to shed light on the structural basis for substrate recognition and catalysis.

Three-dimensional structure of MGS

MGS crystallizes in several high-solvent-content crystal forms with complex noncrystallographic symmetry and high molecular mobility. The structure was solved through the realization that the threefold symmetry of the native $P3_221$ crystal form had broken down in the selenomethionine (SeMet)-derived space group $C2$ crystals. The initial chain trace showed that the C-terminal 15 residues were disordered in the electron density, and subsequent construction of a C-terminally truncated MGS, MGS-Tr, resulted in isomorphous crystals showing substantially better diffraction qualities. MGS was subsequently solved and refined in four forms (resolutions from 2.9 to 1.95 Å) with citrate; Mn^{2+} and glycerate; Co^{2+} and GDP; and Mn^{2+} and GDP-Man (**Table 2**).

Of the 397 amino acids of the MGS monomer, residues 2–262 adopt the mixed α/β GT-A fold first observed in the structure of the inverting GT2 GT SpsA[3] (**Fig. 2**). An eight-stranded β-sheet core is flanked by three α-helices on either face. An 'Asp-X-Asp' (where X is

Table 1 Donor and acceptor plasticity in MGS from pseudo single-substrate kinetics

Enzyme	Temperature (°C)	Fixed substrate	Varied substrate	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (s^{-1} mM^{-1})
Wild-type MGS	25	D-glycerate[a]	GDP-Man[a]	1.02 ± 0.3	81.2 ± 38.7	12.6
	25	D-glycerate[b]	GDP-Glc[c]	0.63 ± 0.06	124.9 ± 12.6	5.0
	25	GDP-Man[a]	D-glycerate[a]	1.02 ± 0.2	121.9 ± 68.4	8.4
	25	GDP-Man[b]	Glycolic acid[d]	$5.2 \times 10^{-3} \pm 9.5 \times 10^{-4}$	23.7 ± 2.8	0.22
	25	GDP-Man[b]	D-lactate[d]	$2.6 \times 10^{-3} \pm 1.3 \times 10^{-3}$	21.3 ± 4.5	0.12
	65	D-glycerate[b]	GDP-Man[c]	6.1 ± 0.4	89.4 ± 12.0	68.2
	65	GDP-Man[b]	D-glycerate[c]	6.5 ± 0.5	96.5 ± 15.5	67.3
	65	D-glycerate[b]	GDP-Glc[c]	1.1 ± 0.03	138.6 ± 6.8	7.9

[a]Multisubstrate kinetics were carried out as detailed in Methods. All other kinetic parameters were derived from pseudo single-substrate kinetics where the fixed substrate is at a saturating concentration. [b]Fixed substrate is 2 mM. [c]Concentrations of 20, 50, 80, 120 and 150 μM. [d]Concentrations of 5, 15, 25, 50 and 100 μM.

any amino acid) motif implicated in divalent metal-ion binding[3,12] is found at the C terminus of strand β4 and is formed by Asp100 and Asp102 (although metal-ion coordination is actually provided by Asp102 and His217 in an Asp/His motif, described below). In addition, the C-terminal residues 263–381 form an entirely α-helical six-helix domain not previously observed in GTs, which, in the case of MGS, is apparently involved in protein oligomerization. Deletion of this C-terminal domain renders the expressed protein insoluble. MGS forms a tetramer, both in solution, as evidenced by light scattering and gel filtration (**Supplementary Methods**), and in all crystal forms, with the 222 tetramer either presenting noncrystallographically or involving use of a crystallographic twofold axis. One dimer interface is provided through the final C-terminal helix, residues 360–368, and through the final strand of the central domain, residues 248–258, with the second interface provided by regions almost entirely donated from the unusual C-terminal helical bundles interacting with their equivalents from the two-fold related partner. MGS oligomerization does not seem to contribute to the catalytic activity of the enzyme. We see no evidence for cooperativity; the active site of MGS is not located at subunit interfaces, and interactions of the protomers do not alter when in complex with substrate or product. Although the biological importance of the tetramer is unclear, it may contribute to the thermostability shown by the enzyme.

The three-dimensional structure of the 'catalytic domain' of MGS, residues 1–263, shows similarities to that of other GTs that exhibit the GT-A fold, with Dali[13] *Z*-scores from 14.6 to 9.1. The closest structural similarity is with family GT64 α-1,4-*N*-acetylgalactosylaminyl transferase Extl2 (197 Cα residues overlap with an r.m.s. deviation of 3.3 Å)[14]. Of the six most similar structures, four are not GTs but are pyrophosphorylases and nucleotidyltransferases responsible for the synthesis of sugar nucleotides from sugar phosphates and nucleoside triphosphates, respectively; these enzymes catalyze chemistry at the phosphorous centers and not at the anomeric carbon. Despite this, the catalytic centers of all the retaining GTs revealed by this search have conserved structural features, which are discussed in light of several MGS ligand complexes in the next section.

Ligand complexes of MGS

The initial SeMet and 'native' structures of MGS showed density for citrate (present at 1.3 M) in the acceptor-binding site (**Fig. 3a**). Citrate makes many of the interactions of the natural acceptor, D-glycerate (described later), and its binding is reflected in a K_i of 24.6 ± 1.7 μM. The complex with D-glycerate (**Fig. 3b**), at 2.45 Å, shows the interactions of the natural substrate acceptor with MGS. The carboxylate forms hydrogen bonds with the main chain amide hydrogens of Ile138 and Thr139 and with the side chain hydroxyl of the latter. Arg131 sits 'below' the glycerate, interacting with both the carboxylate and C2-OH, whereas C3-OH forms hydrogen bonds with the backbone NH of Met137 and Ala136, respectively.

The structure of the form of MGS with GDP and Co^{2+} (**Fig. 3c**) shows that the nucleotide binds to the N-terminal Rossmann domain of the protein, as expected[3]. The guanidine moiety lies in a hydrophobic pocket, with Gln66 accepting hydrogen bonds from endocyclic N1 and exocyclic N2. Ribose recognition is provided, not by the first aspartate of the Asp-X-Asp motif (where X is Ala101), as expected, but through the interaction of both ribose hydroxyls with Glu11 (whose mutation, notably, increases catalytic activity). Furthermore, ribose O3 accepts a hydrogen bond from the main chain amide of Ala101 of the Asp-X-Asp motif. The divalent cation (Co^{2+}) is coordinated by His217 and the second aspartate of the Asp-X-Asp motif, Asp102. The ion probably has hexagonal coordination with the other two axial ligands, O2 and O1 of the α and β phosphates, but the electron density at 2.95 Å is not sufficiently precise to model the two equatorial water ligands.

The binary complex of MGS with the intact donor Mn^{2+} and GDP-Man (**Fig. 3d**) is highly informative, especially in light of the equivalent donor complex of the retaining UDP-galactosyl transferase, LgtC[15]. In MGS, the mannosyl moiety of GDP-Man binds in a hydrophobic pocket flanked on one side by Met229, with Trp189 (Ile79 and Ile76 in LtgC) providing a hydrophobic platform below the pyranoside ring. The sugar lies in $^4C_1$ chair conformation. Asp192 interacts with both O4 and O6 of the mannose, analogous to Asp188 of LgtC. In LgtC, Asp188 is involved in a hydrogen-bonding network via Arg86 that extends to the first aspartate of the Asp-X-Asp motif. Notably, a similar relay exists in MGS involving Lys76 and Asp100, both of which contribute hydrogen bonds to the O3 position of mannose. The sole direct interaction with the 2-position (which is axial in mannose and equatorial in glucose) of mannose is from the main chain carbonyl of Leu163. This carbonyl is of further interest because it lies in approximately the same position as the side chain carbonyl in LgtC, which is implicated in transition-state stabilization during catalysis by LgtC[15].

When one considers representatives of all the retaining GT-A fold GTs (families GT6, 8, 15, 27, 64 and MGS), a pattern of three-dimensional conservation, and one that differs from that of related inverting transferases, begins to emerge (**Fig. 4a**). Retaining GTs from all of these families share several features, such as an aspartate (derived from helix 6) coordinating O6 of the donor sugar, an aspartate from the Asp-X-Asp motif coordinating O3 of the nucleotide ribose and linkage of these carboxylate residues via a positively charged amino acid, such as Lys76 in MGS or Arg86 in LgtC[15] and Mnt1p (ref. 5), for example (**Fig. 4a**). This latter interaction may be especially important in a chemical sense, because the equivalent nitrogen is derived from

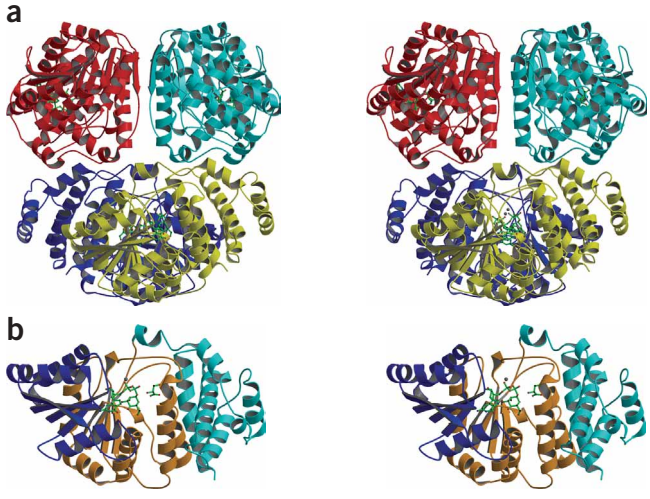

Figure 2 Three-dimensional structure of MGS. (**a**) Structure of an MGS tetramer. The same tetramer is observed in all crystal forms and in solution in both light-scattering and size-exclusion chromatography (not shown). (**b**) Structure of a monomer of MGS colored according to domain, with the N-terminal nucleotide-binding domain dark blue, the central acceptor-binding domain mustard and the unusual C-terminal helical domain pale blue. Mn^{2+} is a shaded sphere with GDP-Man and glycerate (positioned according to its observation in the glycerate complex) in ball-and-stick. This figure, in divergent (wall-eyed) stereo, was drawn with BobScript[26].

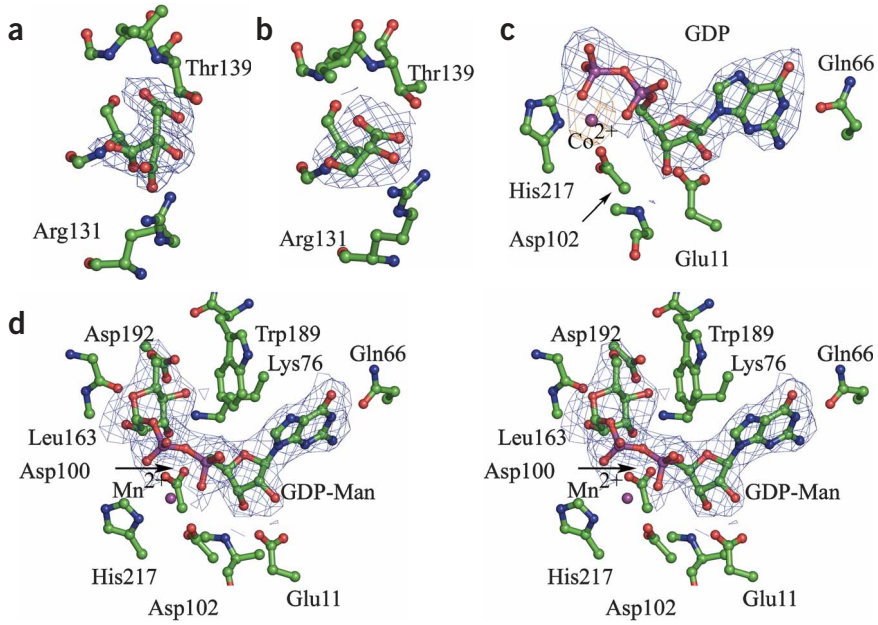

Figure 3 Observed electron density for MGS ligand complexes. (**a**) Citrate. (**b**) D-glycerate. (**c**) Co^{2+} and GDP. The electron density maps shown are $2F_o - F_c$ syntheses (blue) contoured at ~1 σ, and for **c** also the anomalous ($\Delta f''$) difference Fourier in red. (**d**) GDP–Man complex in divergent wall-eyed stereo. This figure and **Figure 4** were drawn with PyMOL (http://pymol.sourceforge.net/).

both different amino acids and different locations in the three-dimensional structure, reflecting a strong convergent evolution. An additional conserved feature is that all known retaining GT-A GTs incorporate at least one histidine side chain into the divalent metal-ion coordination. In GTs from families 6, 8, 15, 27 and 64, this histidine and the position of the metal ion is conserved, but in MGS the histidine (His217) is located two residues farther down the peptide chain, and this creates a slight displacement of the metal ion (**Fig. 4**). His217 lies in the Pro212-Gly222 loop, which shows a large conformational change upon binding of GDP or GDP-Man, with Arg218 and Tyr220 interacting with the α- and β-phosphates, respectively.

Alanine-scanning mutagenesis was used to investigate the functional importance of residues in the active site of MGS. The R131A mutant is inactive (<1,000-fold less active than wild type), consistent with its strong interaction with both the carboxylate and the C2-OH of glycerate. Likewise, H217A and D102A cause activity to drop below detectable levels, which probably reflects the loss of metal-ion coordination and thus interactions between the enzyme and the phosphate component of the nucleotide sugar donor. The divalent metal dependency is exhibited by the complete inactivation of MGS by EDTA. The D100A mutant is also inactive, and thus the proposed interactions between the O3 of the sugar donor and the carboxylate of Asp100 probably play a critical role in substrate recognition and/or catalysis, as has been observed in other retaining GTs[5,14,15]. Mutation of Glu11 (which interacts with ribose) results in a modest potentiation of catalytic activity, with E11A showing a k_{cat} of 2 s^{-1} and a K_m of 105 μM. One interpretation is that Glu11 mutations aid nucleotide departure, which implies that GDP release contributes to the rate-limiting step in the catalytic cycle.

DISCUSSION

Catalytic mechanism in GT-A fold GTs

A feature of the GT sequence classification[1] is that enzymes within a family should catalyze transfer with the same stereochemical outcome. We and others have commented previously that this may be a dangerous assumption[3,16]. MGS performs catalysis with retention of configuration, but it shows closest sequence similarity with family GT2, which contains inverting GTs. On the basis of sequence alone, one could not now predict the retaining stereochemical outcome of the MGS reaction (by inference, several open reading frames of unknown function in GT2 will probably perform catalysis with retention of anomeric configuration). MGS has therefore been classified as the founding member of a new GT family, GT78, into which sequence-related GT2 members may be reclassified should they be shown to possess a retaining catalytic mechanism.

The mechanism of retaining GTs remains a subject of some debate. Although the hydrolysis of glycosides (by retaining glycoside hydrolases) involves a double displacement via the formation of a covalent intermediate, there is little evidence to support such a mechanism for retaining GTs. Many authors favor a 'front-face' departure and attack, related to S_Ni, described further on. Consistent with front-face chemistry, there is, in the case of retaining GTs with the GT-A fold, no conservation of catalytic machinery 'above' (that is, on the β-face) the C1 position, from which the first of any putative double-displacement

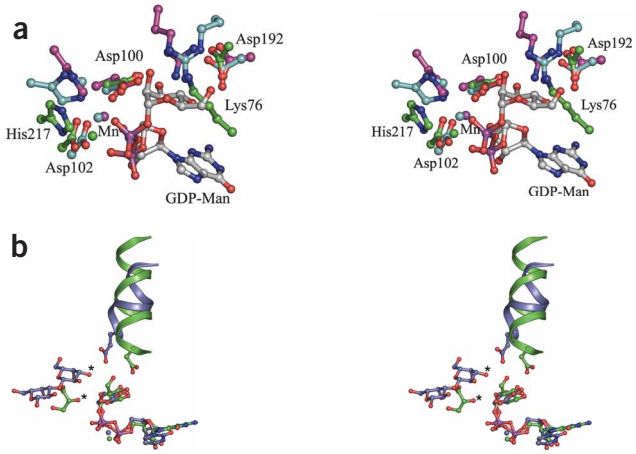

Figure 4 Comparison of GT-A fold retaining and inverting GTs. (**a**) An overlap of the retaining GTs MGS (pale green, this study), LgtC (cyan[15]) and Mnt1p (purple[5]). Of particular note are the chemically invariant interactions around (below) the donor sugar (in this case, GDP-Man from MGS, gray). Residue numbers for MGS are given. (**b**) An overlap of the active centers of the GT43 inverting β-1,3 glucuronyltransferase (slate gray[27]) with acceptor and donor shown and the retaining mannosyltransferase, MGS (pale green). The retaining three-dimensional structures, exemplified by MGS, are characterized by a change in the position of the acceptor from above to slightly below the C1 of the donor and also a change in the angle of helix 6, such that the Brønsted base of the inverting enzymes[27,28] is in a position to interact with the 6-position of the donor sugar of some retaining enzymes. Both structures are a composite of donor and acceptor complexes with the attacking hydroxyl of the acceptor marked with an asterisk.

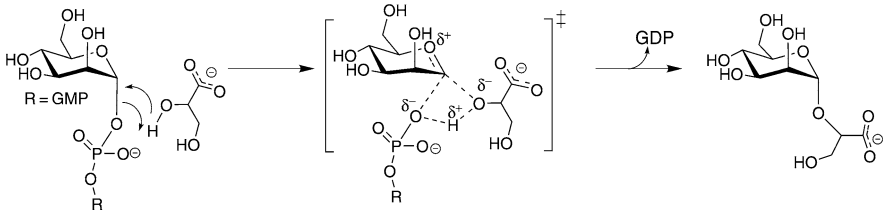

Figure 5 A 'front-face' mechanism for glycosyl transfer with retention of anomeric configuration. Such a mechanism presumably demands asynchronous, but probably concerted, departure of the leaving group and attack by the acceptor through a highly dissociative transition state. Nucleophilic attack, by a poor nucleophile, may involve a close hydrogen bond from one of the oxygens of the leaving-group phosphate oxygen to the hydroxyl of the acceptor, the former acting as Brønsted base for the deprotonation of the latter (in a similar manner to that proposed for the solvolysis of α-glucosyl fluorides by trifluoroethanol[18]).

nucleophilic attack should occur. Recent work[17], however, resulted in the trapping of a covalent glycosyl-enzyme species on a mutant form of LgtC, on residue Asp190 (10 Å away from the active center). If catalytically relevant, such a species would demand a considerable conformational change within the active site, or perhaps the involvement of a dimer partner in catalysis. There is no structurally equivalent residue in MGS to LgtC Asp190, and indeed the environment of this region of retaining GTs is not remotely conserved at the three-dimensional level.

Given the absence of any conserved conformational signatures above C1 in the reported retaining GTs with the GT-A fold, it is hard to invoke a common mechanism involving a covalent glycosyl-enzyme intermediate. Conversely, the strong structural and chemical conservation below C1 in all of these structures, and the arrangement of donor and acceptor in some complexes, favors a concerted but dissociative and asynchronous (perhaps 'S_Ni-like'[18]) mechanism, with partitioning of nucleotide and acceptor below the sugar (**Fig. 5**). Such a 'front-face' mechanism may have a critical, but unconsidered, feature: the ease of evolving both inverting and retaining GTs upon a common scaffold. The study of primitive archaeal genomes has shown that GTs first appeared as a few GT-A fold inverting transferases (from family GT2) and GT-B fold retaining transferases (from family GT4). More complex archaea seem, however, to have evolved both inverting and retaining transferases from each fold class. A mechanism for retention that does not involve a covalent glycosyl-enzyme intermediate would imply that only small changes in the position of the acceptor, of around 3–4 Å, are required to move it into a position where attack would result in inversion rather than retention (**Fig. 4b**). It is also comparatively simple to introduce a residue able to act as the Brønsted base for the deprotonation of the now inverting acceptor. Indeed, the carboxylate residue that interacts with the donor O6 of retaining transferases (Asp192 of MGS and Asp188 of LgtC, for example) lies on the C-terminal domain, at the N-terminal end of helix 6, in a markedly similar position to the conserved base that activates the acceptor for inversion in fold GT-A inverting transferases. One may speculate that a slight change in the angle of the helix allows this change in function (**Fig. 4b**).

MGS shows marked plasticity in donor sugar specificity, and its interactions with the nucleotide diphosphate make a substantial contribution to defining the substrate specificity of the enzyme, reflecting the minor role O2 plays in sugar-protein recognition. Although it is clear that some GTs have evolved tight specificity for the O2 epimers glucose and mannose, MGS shows no absolute chemical specificity for the configuration at C2; indeed, MGS shows marked plasticity in donor and, to a limited extent, acceptor recognition. Broad substrate specificity renders GTs particularly amenable to exploitation in the chemical synthesis of unusual glycoconjugates. High-throughput screening approaches should reveal whether ligand plasticity is a more general feature of GTs and will contribute to the future dissection and exploitation of these enzymes for the synthesis of new bioactive compounds.

METHODS

Generation of MGS. The forms of MGS used in this study are MGS-His_6, MGS-S and MGS-Tr, encoded by pLT1, pJF1 and pJF2, respectively. MGS-His_6 contains a C-terminal His_6 tag, MGS-S, its original stop codon, and MGS-Tr, a 15-residue C-terminal truncation. These plasmids were generated as described in **Supplementary Methods**. Amino acid substitutions were introduced into MGS by QuikChange site-directed mutagenesis (Stratagene).

To generate MGS-His_6, MGS-S and MGS-Tr, *Escherichia coli* strain Tuner harboring pLT1, pJF1 or pJF2 was cultured as described[19]. Cell-free extracts were heat-treated at 65 °C for 20 min and MGS was purified by immobilized metal-affinity, anion-exchange and size-exclusion chromatography. Purified, electrophoretically homogenous proteins were concentrated to 64 mg ml^{-1} in 10 mM Tris-HCl buffer, pH 8.0. SeMet-labeled MGS-His_6 was produced in *E. coli* B834 (DE3) containing pLT1 with recombinant protein expression induced by 1 mM IPTG and incubation at 16 °C for 18 h. MGS-His_6 was purified as described earlier with 10 mM β-mercaptoethanol in all buffers. Purified protein was exchanged into 10 mM Tris-HCl buffer, pH 8.0, containing 10 mM DTT and concentrated to 64 mg ml^{-1}.

Crystallization of MGS. Crystals of MGS-His_6, MGS-S and MGS-Tr were grown by the hanging-drop method. MGS-His_6 and SeMet MGS-His_6 (16 mg ml^{-1}) were crystallized in 5 mM D-glyceric acid, 0.2 M trisodium citrate, 0.1 M sodium acetate trihydrate, pH 4.6, 30% (v/v) MPD. Crystals of MGS-S in the presence of GDP were grown in 0.01 M $CoCl_2$ hexahydrate, 0.1 M sodium acetate trihydrate, pH 4.6, 1 M 1,6 hexanediol, 10 mM GDP with protein at 16 mg ml^{-1}. Crystals of MGS-Tr in the presence of citrate were grown in 0.3 M trisodium citrate, 0.1 M sodium acetate trihydrate, pH 4.6, 40% (v/v) MPD. Crystals of MGS-Tr in the presence of glycerate and Mn^{2+} were grown from 0.2 M $MnCl_2$, 30% (v/v) MPD 18 mM glycerate with MGS at 18 mg ml^{-1}. Crystals of MGS-Tr (16 mg ml^{-1}) in the presence of GDP-Man were grown in 10 mM NaCl, 0.1 M sodium acetate trihydrate, pH 4.6, 15% (v/v) MPD, 1mM $MnCl_2$, 100 mM GDP-Man, with the concentration of MPD increased to 30% (v/v) for cryocooling. Crystals were flash-frozen in liquid nitrogen.

Structure solution and refinement. MGS crystals appeared in two related crystal forms: $P3_221$ (two molecules in the asymmetric unit with 70% solvent) and its related $C2$ form with six molecules in the asymmetric unit; **Table 2**. Native data were in the trigonal form, whereas SeMet crystals had a breakdown in the threefold axis and were thus reduced in $C2$. Single-wavelength SAD data for SeMet MGS-His_6 extended to 2.9 Å and were collected at a wavelength optimized for the f'' signal of Se on beamline ID14-EH4 of the European Synchrotron Radiation Facility (data not shown). All data were processed and reduced using the HKL suite[20]. Bijvoet differences were used as input to SHELXD[21] and 58 Se sites were identified. Phases were improved with RESOLVE[22], incorporating six-fold noncrystallographic symmetry averaging with the operators automatically derived from the Se positions. A model of MGS was built into the averaged map using QUANTA (Accelrys), and this was used for refinement of a single noncrystallographic symmetry–constrained molecule using CNS[23]. This generated a model with an R_{cryst} of 0.27 and an R_{free} of 0.28, which was used as the starting model for refinement of the truncated and complexed forms.

Crystals for native MGS-Tr in citrate and glycerate-bound forms crystallized in the $C2$ and $P3_221$ forms, respectively. Crystals of MGS in complex with GDP

Table 2 X-ray data and structure-quality statistics

	Native (citrate)	Glycerate	GDP	GDP-Man
Data collection[a]				
Space group	*C*2	$P3_221$	$P2_12_12$	$P2_12_12$
Cell dimensions				
a, *b*, *c* (Å)	261, 151, 153	151, 151, 153	403, 161, 108	403, 161, 108
α, β, γ (°)	90, 90.4, 90	90, 90, 120	90, 90, 90	90, 90, 90
Resolution (Å)	40–1.95 (2.02–1.95)	40–2.45 (2.49–2.40)	40–2.95 (3.11–2.95)	40–2.80 (2.95–2.8)
R_{merge}	0.068 (0.51)	0.067 (0.48)	0.095 (0.51)	0.096 (0.48)
$I / \sigma I$	15 (1.6)	26 (4.0)	13 (2.4)	14.5 (4.0)
Completeness (%)	97 (90)	98 (99)	98 (93)	98 (99)
Redundancy	3.0 (2.1)	6.2 (5.6)	4.3 (3.2)	5.3 (4.8)
Refinement				
Resolution (Å)	20.00–1.95	36.99–2.45	40.00–2.95	40–2.80
No. reflections	392,971	69,404	140,625	162,631
R_{work} / R_{free}	0.18 / 0.19	0.25 / 0.28	0.21 / 0.23	0.20 / 0.22
No. atoms				
Protein	18,828	6,276	31,020	31,380
Ligand / ion	78	14 / 2	280 / 10	390 / 10
Water	1,767	100	127	335
B-factors				
Protein	19	30	63	64
Ligand / ion	37	34 / 49	63 / 67	64 / 64
Water	42	62	43	45
R.m.s. deviations				
Bond lengths (Å)	0.010	0.018	0.015	0.022
Bond angles (°)	1.232	1.367	1.439	1.792

Values in parentheses are for the highest-resolution shell.
[a]All data from ESRF beamlines: citrate and glycerate ID14-EH4, GDP and GDP-Man ID29.

crystallize in space group $P2_12_12$ with ten molecules of MGS in the asymmetric unit (**Table 2**). The structure was solved with MOLREP[24] using the dimer of MGS from the $P3_221$ form as the search model and with the anomalous difference Fourier, showing ten identically bound Co^{2+} ions, providing validation of the difficult solution. Crystals with GDP-Man and Mn^{2+} were also obtained in this $P2_12_12$ form (**Table 2**).

Refinement of all structures used REFMAC[25] with tight noncrystallographic symmetry restraints that were released as judged by the behavior of R_{free}. Manual rebuilding was carried out in weighted $2F_o - F_c$ and $F_o - F_c$ electron density maps, which were averaged according to the noncrystallographic symmetry (2×, 6× or 10×, appropriately). Water molecules were included, automatically, into averaged $F_o - F_c$ maps.

GAR screening and kinetics of MGS. The method has been fully described[10]. Briefly, the screen contains all potential nucleotide-donor sugars plus a large, and expanding, acceptor library containing typical GT ligands including native substrates, as well as antibiotics, flavonoids, coumarins, cinnamic acids, peptides and sugars (**Supplementary Fig. 1**). In the specific case of MGS, enzyme reactions were incubated at 37 °C for 16 h and then analyzed by LCMS to monitor the formation of product ions whose concentration was determined through internal calibration.

For kinetics, the mass spectrometer was calibrated with reaction products and substrates as described[10]. Similar standard curves were constructed for citrate and EDTA (200 and 500 μM, respectively) and in the presence of metal chlorides (200 μM of Mg^{2+}, Mn^{2+}, Ca^{2+}, Ni^{2+} and Co^{2+}). Reactions were carried out in 1 mM Tris-HCl buffer, pH 7.8, containing Ca^{2+} ions (>300 molar excess of enzyme concentration) and 3.51×10^{-8} M MGS at 25 °C unless otherwise stated. Kinetic multisubstrate studies used varied D-glycerate (20 μM, 60 μM and 100 μM) and donor concentrations (20, 40, 60, 80 and 100 μM). For pseudo single-substrate studies, the fixed substrate was at 2 mM, whereas the other substrate varied from 20 to 150 μM. Initial rates were determined by monitoring the reaction every 30 min for up to 480 min. For inhibition studies, citrate was included at concentrations from 0 to 100 μM, the substrate was at 100 μM, and the substrate was varied between 20 and 100 μM. To evaluate how different metal ions influence catalytic activity, enzyme reactions were carried out containing the cognate metal at 200 μM, 40 μM GDP-Man and 40 μM D-glycerate. Comparison of the activity of MGS mutants with wild type used GDP-Man and D-glycerate fixed at 100 μM and an enzyme concentration of 3.51×10^{-8} M.

Accession codes. Coordinates and observed structure factors have been deposited in the Protein Data Bank (accession codes: 2BO4 (citrate), 2BO6 (glycerate), 2BO7 (GDP) and 2BO8 (GDP-Man)).

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS
The authors thank G. Sheldrick (Göttingen) for assistance with SHELXD. This work was funded by Biotechnology and Biological Sciences Research Council and the Wellcome Trust. G.J.D. is a Royal Society University Research Fellow.

COMPETING INTERESTS STATEMENT
The authors declare that they have no competing financial interests.

Received 31 January; accepted 28 April 2005
Published online at http://www.nature.com/nsmb/

1. Coutinho, P., Deleury, E., Davies, G.J. & Henrissat, B. An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* **328**, 307–317 (2003).
2. Chiu, C.P. *et al.* Structural analysis of the sialyltransferase CstII from *Campylobacter jejuni* in complex with a substrate analog. *Nat. Struct. Mol. Biol.* **11**, 163–170 (2004).
3. Charnock, S.J. & Davies, G.J. Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms. *Biochemistry* **38**, 6380–6385 (1999).

4. Vrielink, A., Rüger, W., Driessen, H.P.C. & Freemont, P.S. Crystal structure of the DNA modifying enzyme β-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *EMBO J.* **13**, 3413–3422 (1994).
5. Lobsanov, Y.D. *et al.* Structure of Kre2p/Mnt1p—a yeast α-1,2-mannosyltransferase involved in mannoprotein biosynthesis. *J. Biol. Chem.* **279**, 17921–17931 (2004).
6. Martins, L.O. *et al.* Biosynthesis of mannosylglycerate in the thermophilic bacterium *Rhodothermus marinus*—Biochemical and genetic characterization of a mannosylglycerate synthase. *J. Biol. Chem.* **274**, 35407–35414 (1999).
7. Borges, N., Marugg, J.D., Empadinhas, N., da Costa, M.S. & Santos, H. Specialized roles of the two pathways for the synthesis of mannosylglycerate in osmoadaptation and thermoadaptation of *Rhodothermus marinus*. *J. Biol. Chem.* **279**, 9892–9898 (2004).
8. Borges, N., Ramos, A., Raven, N.D.H., Sharp, R.J. & Santos, H. Comparative study of the thermostabilizing properties of mannosylglycerate and other compatible solutes on model enzymes. *Extremophiles* **6**, 209–216 (2002).
9. Faria, T.Q., Knapp, S., Ladenstein, R., Macanita, A.L. & Santos, H. Protein stabilisation by compatible solutes: Effect of mannosylglycerate on unfolding thermodynamics and activity of ribonuclease A. *Chembiochem* **4**, 734–741 (2003).
10. Yang, M., Brazier, M., Edwards, R. & Davis, B.H. High-throughput mass spectrometry monitoring for multi-substrate enzymes: determining the kinetic parameters and catalytic activities of glycosyltransferases. *Chembiochem* **6**, 346–357 (2005).
11. Sampaio, M.M., Santos, H. & Boos, W. Synthesis of GDP-mannose and mannosylglycerate from labeled mannose by genetically engineered *Escherichia coli* without loss of specific isotopic enrichment. *Appl. Environ. Microbiol.* **69**, 233–240 (2003).
12. Wiggins, C.A. & Munro, S. Activity of the yeast *MNN1* α-1,3-mannosyltransferase requires a motif conserved in many other glycosyltransferase families. *Proc. Natl. Acad. Sci. USA* **95**, 7945–7950 (1998).
13. Holm, L. & Sander, C. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138 (1993).
14. Pedersen, L.C. *et al.* Crystal structure of an alpha 1,4-*N*- acetylhexosaminyltransferase (EXTL2), a member of the exostosin gene family involved in heparan sulfate biosynthesis. *J. Biol. Chem.* **278**, 14420–14428 (2003).
15. Persson, K. *et al.* Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs. *Nat. Struct. Biol.* **8**, 166–175 (2001).
16. Campbell, J.A., Davies, G.J., Bulone, V. & Henrissat, B. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino-acid similarities. *Biochem. J.* **326**, 929–942 (1997).
17. Lairson, L.L. *et al.* Intermediate trapping on a mutant retaining α- galactosyltransferase identifies an unexpected aspartate residue. *J. Biol. Chem.* **279**, 28339–28344 (2004).
18. Sinnott, M.L. & Jencks, W. Solvolysis of D-glucopyranosyl derivatives in mixtures of ethanol and 2,2,2-trifluoroethano. *J. Am. Chem. Soc.* **102**, 2026–2032 (1980).
19. Proctor, M. *et al.* Tailored catalysts for plant cell-wall degradation: redesigning the exo/endo preference of the *Cellvibrio japonicus* arabinanase 43A. *Proc. Natl. Acad. Sci. USA* **102**, 2697–2702 (2005).
20. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
21. Schneider, T.R. & Sheldrick, G.M. Substructure solution with SHELXD. *Acta Crystallogr. D* **58**, 1772–1779 (2002).
22. Terwilliger, T.C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D* **55**, 849–861 (1999).
23. Brünger, A.T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
24. Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30**, 1022–1025 (1997).
25. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
26. Esnouf, R.M. An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Model.* **15**, 132–134 (1997).
27. Pedersen, L.C. *et al.* Heparan/chondroitin sulfate biosynthesis: structure and mechanism of human glucuronyltransferase I. *J. Biol. Chem.* **275**, 34580–34585 (2000).
28. Tarbouriech, N., Charnock, S.J. & Davies, G.J. Three-dimensional structures of the Mn and Mg dTDP complexes of the family GT-2 glycosyltransferase SpsA: a comparison with related NDP-sugar glycosyltransferases. *J. Mol. Biol.* **314**, 655–661 (2001).