

Structural Analysis of *Thermus thermophilus* HB27 Mannosyl-3-phosphoglycerate Synthase Provides Evidence for a Second Catalytic Metal Ion and New Insight into the Retaining Mechanism of Glycosyltransferases^{*[5]}

Received for publication, December 17, 2009, and in revised form, March 17, 2010. Published, JBC Papers in Press, March 31, 2010, DOI 10.1074/jbc.M109.095976

Susana Gonçalves, Nuno Borges, Ana M. Esteves, Bruno L. Victor, Cláudio M. Soares, Helena Santos, and Pedro M. Matias¹

From the Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Apartado 127, 2781-901 Oeiras, Portugal

Mannosyl-3-phosphoglycerate synthase is a glycosyltransferase involved in the two-step synthetic pathway of mannosylglycerate, a compatible solute that accumulates in response to salt and/or heat stresses in many microorganisms thriving in hot environments. The three-dimensional structure of mannosyl-3-phosphoglycerate synthase from *Thermus thermophilus* HB27 in its binary complex form, with GDP- α -D-mannose and Mg^{2+} , shows a second metal binding site, about 6 Å away from the mannose moiety. Kinetic and mutagenesis studies have shown that this metal site plays a role in catalysis. Additionally, Asp¹⁶⁷ in the DXD motif is found within van der Waals contact distance of the C1' atom in the mannopyranose ring, suggesting its action as a catalytic nucleophile, either in the formation of a glycosyl-enzyme intermediate according to the double-displacement S_N2 reaction mechanism or in the stabilization of the oxocarbenium ion-like intermediate according to the $D_N^*A_{Nss}$ (S_{Ni} -like) reaction mechanism. We propose that either mechanism may occur in retaining glycosyltransferases with a GT-A fold, and, based on the gathered structural information, we identified an extended structural signature toward a common scaffold between the inverting and retaining glycosyltransferases.

Mannosylglycerate (α -D-mannopyranosyl-(1,2)-O-D-glycerate) is a compatible solute accumulating in many hyperthermophilic prokaryotes in response to heat and/or osmotic stress (1). Besides the fundamental interest in the synthesis of these compounds in relation with strategies of adaptation to extreme conditions, there is the practical relevance arising from their utilization as stabilizers of a variety of biomaterials (2–4). The role of mannosylglycerate in osmoprotection has been demonstrated in *Thermus thermophilus* (5). The involvement of this glycoside in thermoprotection has not been firmly established,

but its superior ability to stabilize proteins *in vitro* suggests a role in cell protection against heat.

Glycosyltransferases catalyze the synthesis of an immense diversity of oligosaccharides, polysaccharides, and glycoconjugates (6). Mannosyl-3-phosphoglycerate synthase (MpgS)² (EC 2.4.1.217) is a glycosyltransferase (GT) involved in the synthesis of mannosylglycerate (Fig. 1) and is widespread in thermophilic bacteria and hyperthermophilic archaea from marine environments (1, 7). MpgS catalyzes the transfer of the mannosyl moiety from the activated precursor GDP- α -D-mannose (GDP-Man) into the activated nucleophile acceptor D-glycerate-3-phosphate (3-PG; D-3-phosphoglycerate), yielding 2-(α -D-mannosyl)-3-phosphoglycerate (M-3-PG; mannosyl-3-phosphoglycerate) that serves as substrate for mannosyl-3-phosphoglycerate phosphatase (MpgP; EC 3.1.5.70). Biochemical characterization of MpgS (8) has led to its inclusion in the retaining GT55 family (see the CAZY Web site). Alternatively, mannosylglycerate synthesis can proceed via the direct transfer of the mannosyl group in GDP-Man to D-glycerate catalyzed by mannosylglycerate synthase (MgS; EC 2.4.1.-) (7), an enzyme belonging to the retaining GT78 family (9).

Despite the success of the techniques applied to the retaining glycoside hydrolases, to trap the intermediary catalytic state and identify the catalytic nucleophile (10), similar approaches have thus far failed with the retaining GTs. The mechanisms proposed for inverting GTs resemble a single direct displacement S_N2 -like reaction, as established for inverting glycoside hydrolases (11); in contrast, the retaining mechanism for glycosyl transfer is less obvious because there is less stringency for a common scaffold at the β -face of the sugar donor than that required for the double displacement S_N2 -like reaction occurring in retaining glycoside hydrolases (9). Instead, a front-face S_{Ni} -like mechanism was proposed for retaining GTs and later designated $D_N^*A_{Nss}$ according to the IUPAC nomenclature.

^{*} This work was carried out with the support of the Diamond Light Source, funded by Seventh Framework Programme Grant 226716, and Fundação para a Ciência e Tecnologia Grants PTDC/QUI/71142/2006, SFRH/BPD/29708/2006, and SFRH/BD/23222/2005.

[5] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Tables S1 and S2 and Figs. S1–S5.

The atomic coordinates and structure factors (code 2WVL) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

¹ To whom correspondence should be addressed. Tel.: 351-21-4469669; Fax: 351-21-4433644; E-mail: matias@itqb.unl.pt.

² The abbreviations used are: MpgS, mannosyl-3-phosphoglycerate synthase; GpgS, glucosyl-3-phosphoglycerate synthase; GT, glycosyltransferase; GDP-Man, GDP- α -D-mannose; 3-PG, D-glycerate-3-phosphate; M-3-PG, 2-(α -D-mannosyl)-3-phosphoglycerate; MpgP, mannosyl-3-phosphoglycerate phosphatase; MgS, mannosylglycerate synthase; wtMpgS, wild-type MpgS; NCS, non-crystallographic symmetry; NDP-sugar, nucleoside-diphospho-sugar; MAD, multiple-wavelength anomalous dispersion; ESRF, European Synchrotron Radiation Facility; PDB, Protein Data Bank.

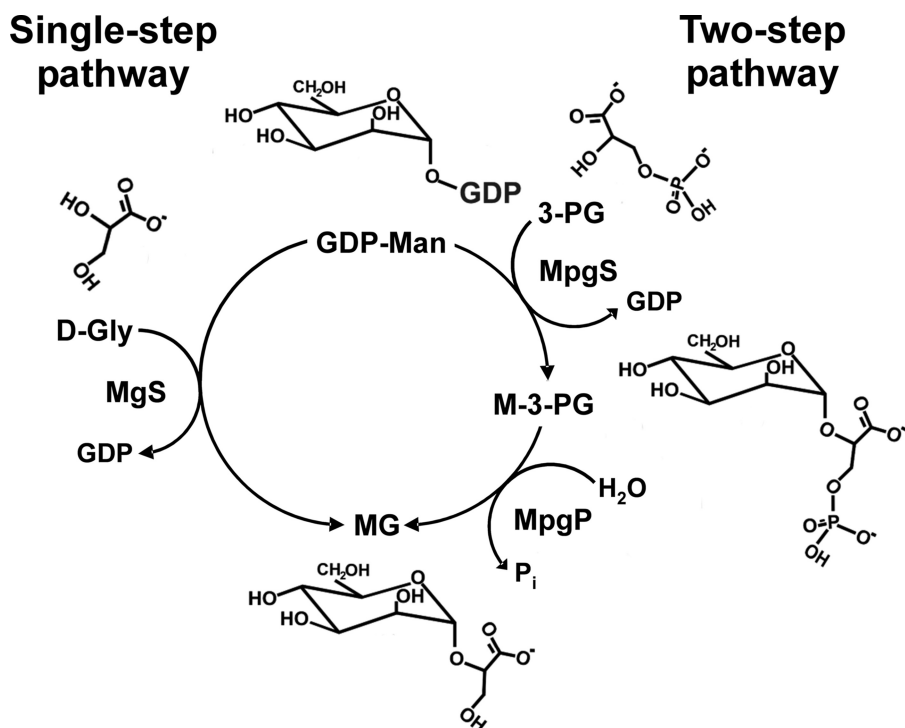


FIGURE 1. **The two pathways for the synthesis of mannosylglycerate.** *Right*, two-step pathway. In the first step, MpgS catalyzes the transfer of the mannosyl moiety from GDP-Man into 3-PG, yielding M-3-PG; in the second step, this intermediate compound is hydrolyzed by MpgP to form the final product α -mannosylglycerate (MG). *Left*, single step pathway. MgS catalyzes the direct glycosyl transfer of GDP-Man into D-glycerate (D-Gly).

This topic has been extensively researched, and a recent comprehensive review is available (12). In the structural studies of the retaining α -1,4-galactosyl transferase in complex with donor and sugar acceptor analogues (13), an argument is provided in favor of the S_Ni -like mechanism by the position of Gln¹⁹⁸ (13). Alternatively, studies on the bovine α -1,3-galactosyltransferase point toward formation of a glycosyl-enzyme intermediate, with Asp³¹⁷ as the likeliest catalytic nucleophile (12, 14). Surprisingly, this enzyme revealed inverting activity when transferring galactose to water or azide (14). Evolutionary studies argue that the retaining GT-A enzymes have evolved from an inverting ancestor (12, 15), and recent work has addressed the elucidation of representative structural scaffolds for both inverting and retaining GTs (9, 12, 16). Currently, the double-displacement S_N2 -like cannot be totally disregarded in favor of the front-face S_Ni -like mechanism. In this work, we describe the three-dimensional structure of *T. thermophilus* HB27 MpgS in three crystal forms: its apo-form and binary complexes with GDP-Man·Mg²⁺ of the wild-type protein and the H309A mutant. The initial structural results prompted additional studies, which revealed the existence of a second catalytic metal center. The protein residues coordinating this metal belong to a sequence motif present in all known MpgS enzymes. In addition, Asp¹⁶⁷ in the DXD motif has been identified as a possible catalytic nucleophile in retaining glycosyltransferases.

EXPERIMENTAL PROCEDURES

Crystallization, Data Collection, and Processing—Wild-type MpgS was produced and crystallized as described previously

(17). The protein was concentrated to 17 mg/ml, and 2- μ l drops (protein/crystallization solution ratio, 1:1) were set up at 20 °C in 24-well crystallization plates using the “hanging drop” vapor diffusion method. The crystallization solution contained 0.2 M magnesium acetate, 0.1 M sodium cacodylate, pH 6.5, 30–35% 2-methyl-2,4-pentanediol and 600 μ M ZnCl₂ as additive.

Crystals of the binary complexes wild-type MpgS (wtMpgS)·GDP-Man·Mg²⁺ and H309A MpgS·GDP-Man·Mg²⁺ were obtained by co-crystallization, using the same conditions as for the apoprotein but supplemented with 2 mM GDP-Man, added to the protein prior to setting up the crystallization drops. Prismatic crystals developed within 5 days, with dimensions varying between 100 and 400 μ m in the longest axis.

A three-wavelength multiple-wavelength anomalous dispersion (MAD) data set to 3.0 Å resolution was measured at the zinc absorption

edge K from a flash-cooled crystal of wtMpgS·GDP-Man·Mg²⁺ at 100 K on ESRF beamline ID-29. A data set from another wtMpgS·GDP-Man·Mg²⁺ crystal was measured to 2.8 Å at 100 K on ESRF beamline ID23-1. A data set from an apo-MpgS crystal was measured to 3.0 Å at 100 K on ESRF beamline ID14-4 (17). A data set from a H309A MpgS·GDP-Man·Mg²⁺ crystal was measured to 3.0 Å at 100 K on Diamond Light Source beamline I04. Diffraction images were processed with the XDS Program Package (18). A summary of the data collection statistics is presented in Table 1.

Structure Determination—Using the HKL2MAP (19) graphical user interface, the three-wavelength MAD data set was scaled and analyzed with SHELXC, the Zn²⁺ heavy atom substructure was determined with SHELXD (20), and the phase problem was solved with SHELXE (21). The best solution from SHELXD in 100 trials gave five Zn²⁺ sites with a correlation coefficient of 38.3%, two of them with apparently about half the occupancy of the top three. The SHELXE calculations gave a clear discrimination between the correct and the inverted substructure solutions and resolved the space group ambiguity in favor of P4₁2₁2. The phases derived from the MAD data were further improved, using the maximum likelihood heavy atom parameter refinement in SHARP (22), with resolution extension using the ID23-1 2.8 Å data set, and followed by density modification with SOLOMON (23). An optimizing density modification procedure using SOLOMON suggested a solvent content of 62%, which was used in all subsequent density modification calculations. Centroid SHARP phases were further improved by density modification with DM (24) using the

TABLE 1

Data collection statistics for the three MpgS crystals

	wtMpgS-GDP-Man·Mg ²⁺ MAD data set			wtMpgS-GDP-Man·Mg ²⁺	H309A MpgS-GDP-Man·Mg ²⁺
	Peak	Infection	Remote		
Beamline	ESRF ID 29			ESRF ID 23-1	DLS I04
Detector	ADSC Quantum Q315r				
Wavelength (Å)	1.2823	1.2831	1.2716	1.0672	0.9763
Data processing	XDS				
Space group	P 4 ₁ 2 ₁ 2				
Unit cell parameters (Å)	<i>a</i> = 113.1, <i>c</i> = 197.1	<i>a</i> = 113.5, <i>c</i> = 197.7	<i>a</i> = 114.2, <i>c</i> = 198.0	<i>a</i> = 113.7, <i>c</i> = 196.7	<i>a</i> = 112.7, <i>c</i> = 199.7
Resolution (Å)	45.18–2.99 (3.17–2.99)	45.31–2.99 (3.17–2.99)	45.41–2.99 (3.17–2.99)	45.1–2.80 (2.97–2.80)	48.90–2.98 (3.16–2.98)
No. of observations	481,922 (75,644)	486,946 (74,381)	196,686 (21,481)	197,746 (31,785)	69,101 (7084)
Unique reflections	49,479 (7960)	49,862 (7861)	50,076 (7601)	32,332 (5078)	25,598 (3441)
Completeness (%)	99.8 (99.4)	99.6 (97.8)	98.6 (92.6)	99.3 (98.8)	94.0 (80.6)
<i>R</i> -merge (%) ^a	9.4 (216.8)	10.1 (217.1)	10.4 (164.9)	6.8 (151.1)	4.5 (37.2)
<i>I</i> /σ(<i>I</i>)	16.18 (1.29)	14.90 (1.26)	8.33 (0.61)	15.62 (1.23)	15.41 (2.09)
<i>R</i> -meas (%) ^b	10 (229.2)	10.7 (229.4)	12.0 (198.9)	7.5 (165.1)	5.6 (47.9)
<i>Z</i> ^c	2				
Estimated solvent content (%) ^c	65.7				

^a *R*-merge = merging *R*-factor, $(\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|) / (\sum_{hkl} \sum_i I_i(hkl)) \times 100\%$. In the MAD data set, Bijvoet pairs were treated as separate observations during the scaling process.

^b *R*-meas = redundancy-independent *R*-factor (intensities), $\sum_{hkl} (N/(N-1))^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl) \times 100\%$ (56).

^c Number of molecules in the asymmetric unit according to the Matthews coefficient (57).

CCP4i graphical user interface (25, 26), and with non-crystallographic symmetry (NCS) averaging. No dimeric arrangement for the two MpgS molecules in the asymmetric unit could be initially perceived from the spatial arrangement of the five Zn²⁺ positions. Therefore, a preliminary NCS operator relating the two monomers was calculated with LSQKAB (27) by superposing two triangles derived from the five Zn²⁺ positions. At this point, a random 5% sample of the reflection data were flagged for *R*-free calculations (28) during model building and refinement. Initial model building and sequence docking was done automatically with Buccaneer (29), and a preliminary model was completed with Coot (30). At this stage, the dimeric arrangement of MpgS in the asymmetric unit could be clearly discerned, and a better NCS operator was derived, corresponding to a 2-fold NCS axis at the dimer interface. In parallel, and after inspection of the anomalous residual maps, a second SHARP/SOLOMON run was carried out based on the previous one, with anisotropic thermal motion *B*-factor correction for the Zn²⁺ sites, which gave better phases as judged by the figure of merit. Phases from this procedure were again improved in DM/CCP4i using the optimized 2-fold NCS operator. Phasing and phase refinement statistics, with and without anisotropic *B*-factor correction, are listed in supplemental Table S1. Using the new phases from DM, an automated model building, sequence docking, and refinement procedure was carried out with Buccaneer/REFMAC (29, 31). Finally, an electron density map, calculated using the 2.8 Å data set and the improved DM phases, was used for completing the model with Coot.

The structures of apo-MpgS and H309A MpgS-GDP-Man·Mg²⁺ were determined by molecular replacement with PHASER (32) using the refined protein chain coordinates of wtMpgS-GDP-Man·Mg²⁺.

Crystallographic Refinement—The three structures were refined using a maximum likelihood refinement procedure with REFMAC (31), as implemented through the CCP4i graphics user interface (26). Non-crystallographic symmetry restraints between the two independent molecules in the asymmetric unit were applied throughout the refinement, and a translation-libration-screw rigid body refinement (33) was carried out prior to restrained refinement of atomic positions and isotropic ther-

mal motion parameters. One rigid body was defined for each of the two independent molecules in the asymmetric unit, including its bound ligand GDP-Man. The electron density for three flexible loop regions designated as Loop 1, Loop 2, and Loop 3, respectively, in each monomer (see “Results”) and corresponding to the residue ranges Tyr²⁰³–Val²¹⁴, His³¹¹–Gly³¹⁶, and Gly³⁵²–Glu³⁵⁹, was not interpretable in some cases for Loops 1 and 3, as shown in Table 2, and therefore not included in the model. The same reflection indices retrieved from the wtMpgS-GDP-Man·Mg²⁺ data set were used for subsequent *R*-free calculations during the structural refinements of the apo-MpgS and H309A MpgS structures. The wtMpgS-GDP-Man·Mg²⁺ was refined using the ID23-1 2.8 Å data set to final values of *R*-factor and *R*-free of 0.176 and 0.234, respectively. The apo-MpgS structure was refined using the ID14-4 3.0 Å data set to final values of *R*-factor and *R*-free of 0.190 and 0.256. The H309A MpgS-GDP-Man·Mg²⁺ structure was refined using the I04 3.0 Å data set to final values of *R*-factor and *R*-free of 0.174 and 0.232. During the refinement process, the models were periodically inspected and corrected in a three-dimensional graphics work station with Coot (30) against 2|*F*_o| – |*F*_c| and |*F*_o| – |*F*_c| electron density maps, which were also used to include water molecules in the refinement. The final refinement statistics are presented in Table 2. Model stereochemical quality was assessed with PROCHECK (34). The secondary structure was analyzed with PROMOTIF version 3.0 (35) and PROCHECK (34).

Coordinates and structure factors were deposited in the RCSB Protein Data Bank (36) with accession numbers 2wvk and r2wvksf for apo-MpgS, 2wvl and r2wvlsf for wtMpgS-GDP-Man·Mg²⁺, and 2wvm and r2wvmsf for H309A MpgS-GDP-Man·Mg²⁺.

Mutagenesis Studies—The construction of the E251A and H309A MpgS mutant proteins was done by using the QuikChange site-directed mutagenesis kit (Stratagene), according to manufacturer's instructions. The *mpgS* gene from *T. thermophilus* HB27, cloned in the pKK223-3 vector (17), was used as template for plasmid amplification by PCR. The following primers were used: E251A-sense (5'-CG AAC GCC GGG GCG CAC GCC ATG AG-3') and E251A-antisense (5'-CT

TABLE 2

Final refinement statistics for the three MpgS crystal structures

	wild-type apoMpgS	wild-type MpgS:GDP-Man:Mg ²⁺	H309A MpgS:GDP- Man:Mg ²⁺			
Resolution limits (Å)	42.9 - 2.97 (3.05 - 2.97)	45.1 - 2.81 (2.88 - 2.81)	98.2 - 2.98 (3.05 - 2.98)			
R-factor (%) ^a	0.190 (0.315)	0.176 (0.339)	0.174 (0.261)			
nr.reflections, no σ(F _o) cutoff	25502 (1558)	30259 (2142)	24263 (1155)			
Free R-factor (%) ^b	0.256 (0.369)	0.234 (0.426)	0.232 (0.299)			
nr. reflections, no σ(F _o) cutoff	1358 (81)	1620 (115)	1292 (51)			
Overall coordinate error estimate (Å) ^c	0.32	0.27	0.27			
<i>Model composition</i>						
non-hydrogen protein atoms	6076	6025	5881			
GDP-Man	-	78	78			
Zn ²⁺	5	5	3			
Mg ²⁺	-	2	2			
citrate	13	-	-			
Solvent molecules	80	36	15			
<i>Model r.m.s. deviations from ideality</i>						
Bond lengths (Å)	0.015	0.016	0.016			
Bond angles (°)	1.7	1.7	1.6			
Chiral centers (Å ³)	0.10	0.10	0.11			
Planar groups (Å)	0.008	0.007	0.008			
<i>Model completeness and validation</i>						
	Chain A	Chain B	Chain A	Chain B	Chain A	Chain B
Regions omitted	-	203-211	206-211	203-211	203-212 352-357	203-215 352-357
Mean B values (Å ²) ^d						
protein main-chain	107	102	108	105	89.9	85.2
protein side-chain	110	105	114	112	94.8	89.7
GDP-Man	-	-	115	108	85.1	88.4
Mg ²⁺ , Zn ²⁺	88.9	-	99.9	-	71.8	-
citrate	106.2	-	-	-	-	-
solvent	65.6	-	64.9	-	47.0	-
Ramachandran plot statistics. Residues in:						
most favored regions (%)	85.7	85.7	87.9	87.5	90.7	91.5
allowed regions (%)	13.1	13.1	11.2	11.3	8.4	8.5
generously allowed regions (%)	0.9	0.9	0.9	1.2	0.6	0
disallowed regions (%)	0.3	0.3	0.0	0.0	0.2	0
PROCHECK G-factor	-0.15	-0.19	-0.10	-0.08	-0.05	-0.03

^a R-factor = $\sum_{hkl} |F_o| - |F_c| / \sum_{hkl} |F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively.^b Free R-factor is the cross-validation R-factor computed from a randomly chosen subset of 5% of the total number of reflections, which were not used during the refinement. The same set of *hkl* indices was used for the three structures.^c Maximum likelihood estimate with REFMAC.^d Calculated from equivalent isotropic B values, including the translation-libration-screw contribution for the protein and GDP- α -D-mannose atoms.

CAT GGC GTG CGC CCC GGC GTT CG-3') for the E251A substitution and the primers H309A-sense (5'-G ACC CGG AAC CCC GCC CTC CAC GAG AAC-3') and H309A-anti-sense (5'-GTT CTC GTG GAG GGC GGG GTT CCG GGT C-3') for the H309A substitution. Confirmation of the mutations was done by DNA sequencing (AGOWA Genomics). Expression and purification were performed as previously described for wtMpgS (17). The stability of mutant proteins was assessed by differential scanning calorimetry. The melting temperatures were similar to that of wtMpgS ($T_m = 81^\circ\text{C}$): $T_m = 82^\circ\text{C}$ for the E251A mutant protein and $T_m = 81^\circ\text{C}$ for the H309A mutant protein.

Production and Purification of the *Rhodothermus marinus* MpgP for Kinetic Assays—Recombinant MpgP from *R. marinus* DSM4252 was produced as described previously (37) and used to dephosphorylate M-3-PG in the activity assays of wtMpgS and the mutant proteins H309A and E251A. Expression and purification was carried out as described for the *T. thermophilus* HB27 MpgS (17), but the last anionic exchange chromato-

graphic step (MonoQ) was replaced by a cationic exchange step (MonoS, Amersham Biosciences) using 20 mM MES-NaOH, pH 6.5, as the buffer system for protein elution through a linear NaCl gradient. The purest fractions eluted at about 200 mM NaCl and were used for activity assays.

Activity Assays—Prior to the activity assays, all protein samples were dialyzed against 20 mM Tris-HCl, pH 7.6, 25 mM NaCl, 10 mM EDTA, to ensure the removal of contaminant metal ions. A second dialysis step against 20 mM Tris-HCl, pH 7.6, was subsequently done to remove EDTA. Determination of the specific activities ($\mu\text{mol}/\text{min}/\text{mg}$) for the wild-type enzyme and the mutant proteins was done at 75°C (optimal temperature) and based on the release of inorganic phosphate from M-3-PG, using a protocol modified from that described by Empadinhas *et al.* (38). wtMpgS was used for establishing the experimental conditions and as positive control in the assays with the mutant proteins. The assay mixture contained 0.5 μg of pure wtMpgS (total volume of 50 μl) with an excess (5 mM) of GDP-Man and 3-PG in buffer

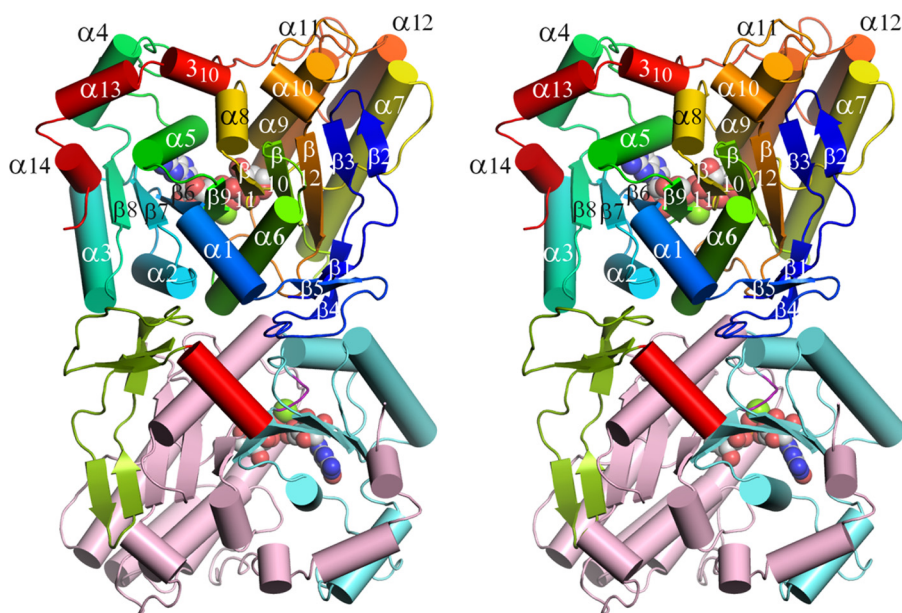


FIGURE 2. The structure of MpgS. Stereo schematic diagram representation of the overall structure and assembly of the MpgS-GDP-Man-Mg²⁺ homodimer. α -Helices are represented as cylinders, β -strands are shown as arrows, and GDP-Man and Mg²⁺ are shown in space-filling mode. Atom colors are as follows. Light gray, carbon; blue, nitrogen; red, oxygen; orange, phosphorus; green, Mg²⁺. The A monomer (bottom) is colored according to the domain regions mentioned under "Results." N-terminal β -region (light green), central α -helix (red), $\beta/\alpha/\beta$ Rossmann-like region (light blue), DAD motif (purple), and C-terminal α -helical region (light pink). In the B monomer (top), the secondary structure elements are colored from blue (N-terminal) to red (C-terminal). The two monomers are related by a 2-fold NCS axis perpendicular to the plane of the figure. This figure was prepared with PyMOL.

B1 (20 mM Tris-HCl, pH 7.6, 20 mM MgCl₂). An excess (5 μ g) of MpgP (specific for M-3-PG) was added to ensure the complete dephosphorylation of M-3-PG. The quantification of inorganic phosphate released was based on the Ames method (39). The effect of Zn²⁺ on the specific activity of wtMpgS was studied by performing the same assays, using 300 μ M of ZnCl₂; moreover, the competing effect of Zn²⁺ (300 μ M) was examined at a Mg²⁺ concentration of 20 mM. We verified that the MpgP activity was not affected by Zn²⁺ or Mg²⁺ at these concentrations. The specific activities of the MpgS mutant proteins were determined as for the wtMpgS, using instead 5 μ g of pure protein to ensure a measurable inorganic phosphate release from M-3-PG. All assays were done at least in duplicate. All solutions were prepared with ultrapure deionized water (Milli-Q, Millipore Corp.).

RESULTS

The Structure of MpgS—Both MpgS-GDP-Man-Mg²⁺ binary complexes from *T. thermophilus* HB27 (wild type and H309A mutant), as well as the apo-MpgS, crystallized in tetragonal space group P4₁2₁2, with a homodimer in the asymmetric unit. The dimeric assembly, represented in Fig. 2, is consistent with the prediction by size exclusion chromatography (17) and probably corresponds to the biological unit in solution. Each monomer displays an overall globular α/β topology with an extended central β -sheet composed of nine β -strands in the order $\beta_2, \beta_3, \beta_{12}, \beta_{10}, \beta_{11}, \beta_9, \beta_6, \beta_7$, and β_8 , with β_2 and β_{11} running in an anti-parallel direction, and its connectivity (40) can be described as 0 1 7X 1X 1X -3X -2X 1 -2, considering strand β_2 as the origin. The architecture of each monomer (supplemental Fig. S4) suggests the presence of three distinct

structural regions. The N-terminal domain comprises two of these regions: a mixed 5-stranded β -region (residues 1–43) composed by strands β_1 – β_5 , and a Rossmann-like $\alpha/\beta/\alpha$ sandwich region (residues 55–172), both linked to a central helix α_1 (residues 45–52), which is oriented obliquely to the plane of the central β -sheet. The C-terminal domain (residues 173–391) begins immediately after the DXD motif (Asp¹⁶⁷-Ala¹⁶⁸-Asp¹⁶⁹), typically present among the metal-dependent GT-A enzymes (12); includes the second half of the β -sheet as well as a 4-helical parallel bundle composed by helices $\alpha_7, \alpha_9, \alpha_{11}$, and α_{12} , which encloses part of the catalytic pocket; and ends with a long C-terminal tail that folds back onto the Rossmann-like N-terminal region. This is a representative structure of the GT55 family, together with the recently available selenomethionine derivative of *Pyrococcus horikoshii* MpgS (PDB code 2ZU7). Secondary structure

matching against all of the PDB using the DALI (41) search engine, revealed a general structural preservation with many other nucleoside-diphospho-sugar (NDP-sugar) transferases, including those involved in the synthesis of lipopolysaccharides, as well as some pyrophosphorylases present at the early stages of peptidoglycan biosynthesis. However, as expected, the highest Z-score (52.9) was observed for its orthologous family member *P. horikoshii* MpgS, followed by members of the GT81 family: *Mycobacterium tuberculosis* glucosyl-3-phosphoglycerate synthase (GpgS) (42) (PDB 3E26; Z-score = 21.9), the putative GpgS from *Mycobacterium avium paratuberculosis* encoded by MAP2569c (PDB 3CKN; Z-score = 21.8) (43), and the novel MpgS/GpgS from *Rubrobacter xylanophilus* that is able to use GDP-Man as well as GDP-glucose as substrates (PDB 3F1Y; Z-score = 21.2) (44). This last enzyme is suggested to be at the evolutionary split from which all MpgSs and GpgSs have evolved (38). *R. marinus* MgS, representative of the GT78 family and involved in the alternative single step pathway for mannosylglycerate synthesis (7, 9), is also one of the most closely related enzymes (PDB 2BO8; Z-score = 19.0). Indeed, these can all be included in the MgS-like fold family (SCOP v.1.75; c.68.1.18) (46), suggesting a common structural core composed of the central β -sheet and defined by strands $\beta_{12}, \beta_{10}, \beta_{11}, \beta_9, \beta_6, \beta_7$, and β_8 (MpgS nomenclature; supplemental Fig. S4) as a reminiscent feature from a common ancestor.

Unlike the $\alpha/\beta/\alpha$ conserved subdomain, the α -rich C-terminal region is highly variable. It is known to be associated with the sugar acceptor binding specificity and is also likely related to the protein solubility (9). The helical bundle is linked to the

β -sheet core through flexible loops, Loop 1 (Tyr²⁰³–Val²¹⁴) and Loop 2 (His³¹¹–Gly³¹⁶), which are involved in the substrate binding-mediated response (9, 42, 43). In addition to these two highly mobile regions, a third (Gly³⁵²–Glu³⁵⁹) is located at the beginning of the long C-terminal tail.

The two MpgS monomers in the asymmetric unit are related by a 2-fold NCS axis at the dimer interface, which is defined by the tight packing of the N-terminal β -region (residues 22–32) with the flexible loop regions of the $\alpha/\beta/\alpha$ Rossmann-like sandwich, comprising mainly the loops connecting $\alpha 2$ to $\beta 7$ (residues 70–79) and $\alpha 3$ to $\beta 8$ (residues 107–109). The dimerization interaction is mostly electrostatic, and its effect is potentiated through the solvent-accessible residues of the two interlocking $\alpha 1$ helices (residues 45–50) and also by π -stacking between NCS-related pairs of Trp¹⁷⁷ and Phe¹⁷² residues. The dimer interaction surfaces were analyzed with PISA (47), where the contributions from bound GDP-Man and metal ions were excluded. Upon dimer formation, an average of 1670 Å² is buried in each monomer, representing about 10% of its total solvent-accessible surface area. In addition, an estimated total of 32 hydrogen bonds and 36 salt bridges are formed across the dimer interface. These interactions very likely contribute to its stabilization and play an essential role in the assembly of the biological unit of *T. thermophilus* HB27 MpgS, similarly to *P. horikoshii* MpgS. On the other hand, the biological unit proposed for GpgS and MAP2569c (42, 43) involves a 2-fold crystallographic axis at the C-terminal subdomain, which is also seen in the biological dimer of *R. xylanophilus* MpgS/GpgS, whereas the MgS biological unit is believed to be a tetramer (9).

The Catalytic Pocket—In the crystal structure of GDP-Man·Mg²⁺ binary complexes of wtMpgS and H309A mutant protein, the ligand lies in the catalytic pocket, located near the bottom of the central β -sheet (with reference to Fig. 2), and in a cleft formed between the Rossmann-like and the C-terminal regions, bordered by several loops. Here, three binding regions can be described according to Fig. 3 and [supplemental Table S2](#): the typically conserved NDP binding pocket, formed by the Rossmann-like subdomain; the sugar binding pocket, located at the boundary of the cleft region; and the less conserved acceptor binding region, located between the α -helical bundle and the C-terminal region of the central β -sheet. At the pocket bottom, and only observed in the crystal structures of apo-MpgS and wtMpgS·GDP-Man·Mg²⁺, a second metal ion position partially occupied by Zn²⁺ lies about 6 Å below the β -face of the mannose ring.

A number of hydrophilic residues participate in GDP-Man·Mg²⁺ recognition and binding. These are conserved in *P. horikoshii* MpgS (PDB 2ZU8), establishing a common network and scaffold environments representative of the GT55 family. The evolutionary relationship among the four diverging enzymes MpgS, MgS, GpgS, and MpgS/GpgS was assessed, outlining key residues and structural constraints relevant to catalysis ([supplemental Fig. S5](#)). The conservation of a global scaffold in the catalytic pocket is shown by their solvent-exposed NDP-sugar binding pockets allowing ready access of the NDP-sugar. From here, as represented in Fig. 4A, two tunnels cross the protein bulk leading back to the surface; the first (tunnel I) lies more or less along the orientation of the NDP-sugar binding

pocket, whereas the second (tunnel II) is roughly perpendicular to it and begins near the bottom of the sugar ring binding region. Comparing the structures of MpgS, GpgS, and MgS, tunnel II (Fig. 4B) emerges as the most likely common candidate for acceptor entrance into the active site of MpgS as well as of MgS and GpgS, in agreement with previous proposals (9, 42, 43).

The pocket bottom is located at the boundary of the cleft region, formed by residues Pro⁶⁰–Glu⁶⁴ and by the DXD motif. It forms a type IV β -turn, and its influence spans two moieties in the NDP-sugar substrate because it is involved in sugar recognition and binding through Asp¹⁶⁷ as well as in the stabilization of the ribose moiety. In addition, through the *syn*-coordination of the catalytic Lewis acid Mg²⁺ by Asp¹⁶⁹, it is also indirectly interacting with the pyrophosphate. The significance of this motif in catalysis has been extensively studied by mutagenesis, which has always resulted in an activity drop below detectable levels (9, 13, 48). Although the second aspartate is prevalently associated with Lewis acid coordination, the first is found to interact with the sugar moieties or their respective analogues.

The β -face of the mannosyl moiety docks against the main chain of the conserved Gly²⁵⁰ at the back of the sugar binding pocket. The α -face is sheltered by a hydrophobic platform, which is also a topologically conserved feature among the retaining GT-A enzymes (9). Hydrophobic contacts are provided by Tyr²⁶⁸ and to a lesser extent by Met³²². In both the wild-type and H309A mutant binary complexes, the mannopyranosyl moiety was found in the typically stable ⁴C₁ chair conformation (Fig. 5, A and B) and with an orientation similar to that found in *P. horikoshii* MpgS (PDB 2ZU8), regardless of a Lewis acid absence in this structure. In contrast, in both MgS (PDB 2BO8) and GpgS (PDB 3CKQ), their sugar moieties were oriented in a different binding mode. The most notable consequence is that in MpgS, Asp¹⁶⁷ O⁸² accepts a hydrogen bond from O2' and is also within van der Waals distance (~3.5 Å) from the anomeric carbon C1' (Fig. 3A) or alternatively from the O5' oxygen (~3.1 Å) as observed for *P. horikoshii* (Fig. 3B). This environment differs from that observed in MgS and GpgS, involving the O3' and O2' oxygen atoms instead (Fig. 3, C and D). These differences may result from different torsion angles adopted by the pyrophosphate backbone, depending on the Lewis acid metal (none, Mg²⁺, or Mn²⁺) present. The GDP-Man orientation found in the MpgS structures has mechanistic implications, which will be discussed below.

MpgS has shown catalytic specificity for 3-PG (49), whereas acceptor plasticity has been reported for MgS (9) and to a lesser extent for GpgS (38). Because the three enzymes share a global catalytic scaffold, 3-PG is also likely to bind to MpgS in a structurally equivalent position to that proposed for 3-PG in GpgS and D-glycerate in MgS (9, 42, 43). Although confirming tunnel II as the most likely entry point for 3-PG into the catalytic site, our docking calculations ([supplemental data](#)) failed to yield a meaningful solution in terms of catalysis ([supplemental Fig. S2](#)). This may be due to a limitation of the current docking methods, which cannot fully account for induced fit (*i.e.* the binding site is not rigid but is formed during the ligand binding process). The binding mechanism of 3-PG probably acts con-

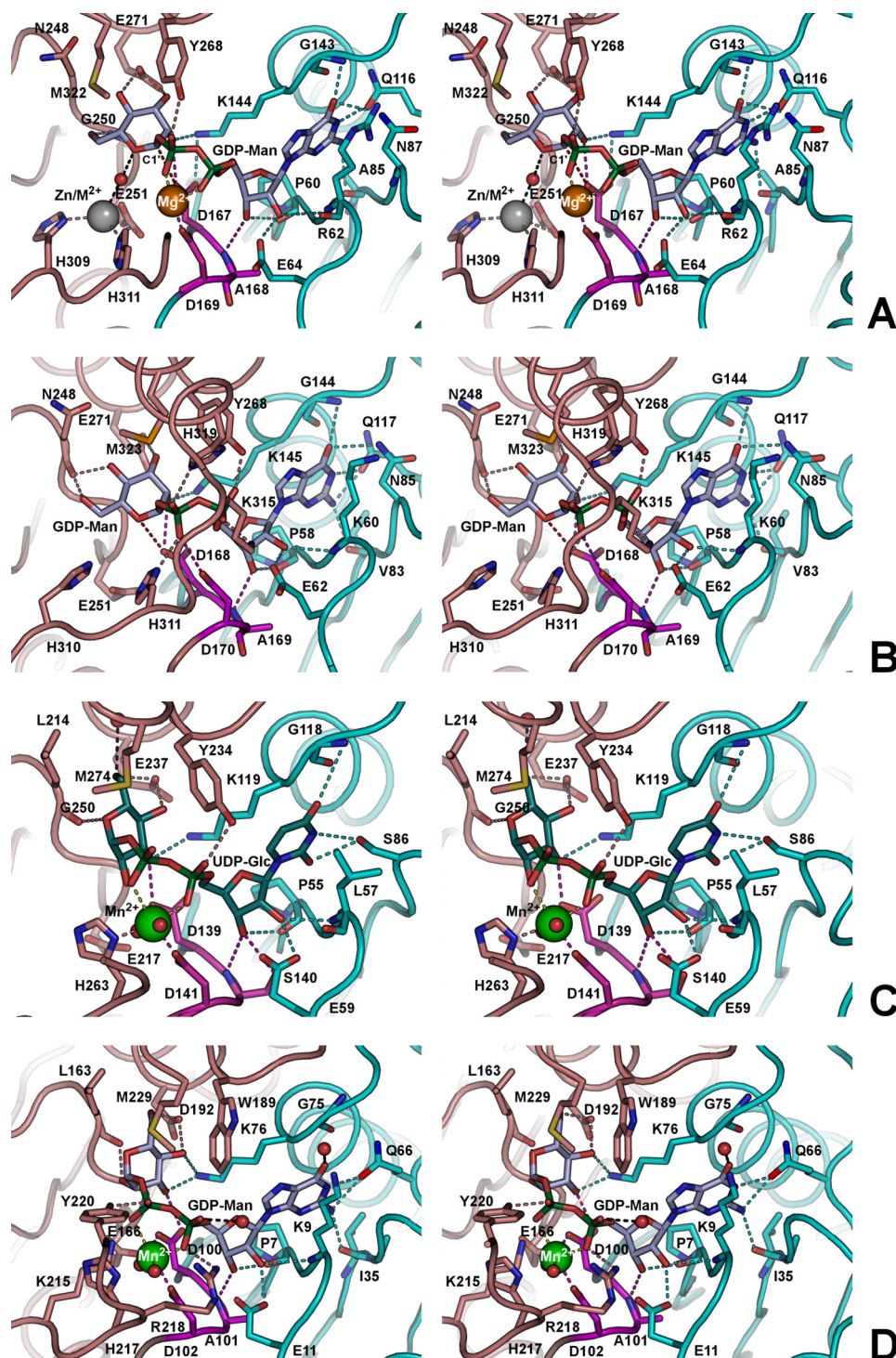


FIGURE 3. The NDP-sugar binding pocket in MpgS and structurally related enzymes. In these stereo images, the protein backbone is represented as a C^{α} tube diagram, the side chains and NDP-sugar ligands are drawn as sticks, and the metal ions and the water molecule are drawn as spheres. The protein backbone and the side-chain carbon atoms are colored according to the protein regions described under "Results." Rossmann-like region (cyan), DAD motif (purple) and C-terminal α -helical region (salmon). Other atom colors are blue for nitrogen, red for oxygen, yellow for sulfur, gold for selenium, green for phosphorus, orange for Mg^{2+} , green for Mn^{2+} , and gray for Zn/M^{2+} . The dashed lines indicate the hydrogen bonds (proposed for MpgS, considering the limited resolution of the diffraction data) or electrostatic interactions established between protein side-chain atoms and the bound NDP-sugar and metal ion ligands and are colored according to the respective protein regions. A, *T. thermophilus* HB27 MpgS; B, *P. horikoshii* MpgS (PDB 2ZU8). M323 is a selenomethionine. A dashed red line represents the van der Waals proximity of Asp¹⁶⁷ to the anomeric C1' carbon in *T. thermophilus* (A) and of Asp¹⁶⁸ to O5' oxygen in *P. horikoshii* (B) of the mannosyl moieties. C, the putative GpgS from *M. avium paratuberculosis* encoded by MAP2569c (PDB 3CKQ). D, *R. marinus* MgS (PDB 2BO8). This figure was created with PyMOL.

certainly and/or synergistically upon GDP-Man· Mg^{2+} binding, similarly to what has been shown for the inverting β -1,4-galactosyltransferase-1 and retaining α -1,3-galactosyltransferase enzymes (14, 50, 51). Also, optimal catalysis occurs at $\sim 75^{\circ}\text{C}$ in MpgS; therefore, thermally induced protein dynamics must be considered.

The Second Metal Ion Coordination Site—The finding of an internal Zn^{2+} site located 6 Å below the β -face of the mannosyl moiety (Fig. 5A) led us to consider the existence of a second catalytic metal ion and its role in the glycosyl transfer mechanism. As seen from the SHELXD results and the heavy atom parameter refinement in SHARP, this site is only partly occupied by Zn^{2+} in both monomers of the wtMpgS·GDP-Man· Mg^{2+} and apo-MpgS structures. *In vivo*, a divalent metal ion other than Zn^{2+} (probably Mg^{2+} or Mn^{2+}) would be bound, and this binding site was therefore designated as Zn/M^{2+} . The metal is coordinated by Glu²⁵¹ O^{ε1}, His³⁰⁹ N^{δ1}, and His³¹¹ N^{ε2}, and an axial water molecule completes the coordination sphere, although it is not visible in all monomers. These residues belong to the sequence motif NXGEX_nHXX, conserved in all MpgS enzymes with known sequence.

The functional relevance of Zn^{2+} was assessed from measurements of maximal specific activity (V_{max}) at 75°C . When Mg^{2+} was used as single metal co-factor at 20 mM (concentration for maximal activity) the MpgS specific activity was 45 ± 3 $\mu\text{mol}/\text{min}/\text{mg}$. When measured solely in the presence of Zn^{2+} (300 μM), the activity decreased about 100-fold (0.4 ± 1.2 $\mu\text{mol}/\text{min}/\text{mg}$), but when 20 mM Mg^{2+} was added in combination with Zn^{2+} (300 μM), the MpgS specific activity recovered substantially (5.3 ± 0.1 $\mu\text{mol}/\text{min}/\text{mg}$). These findings suggest that Zn^{2+} can displace Mg^{2+} at the second metal binding site; therefore, the partial occupation of Zn^{2+} in the wtMpgS·GDP-Man· Mg^{2+} and apo-MpgS structures may also be

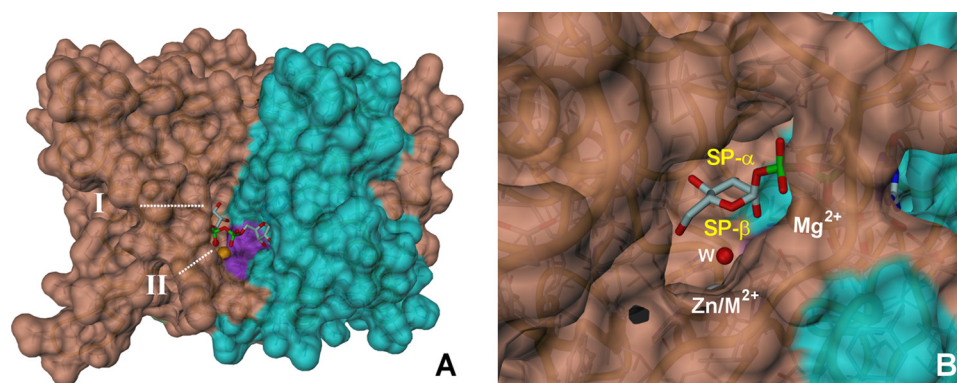


FIGURE 4. **Two tunnels lead from the NDP-sugar binding pocket to the protein surface of MpgS.** The solvent-accessible molecular surface was calculated with MSMS (55) using the coordinates of protein chain A in wtMpgS-GDP-Man-Mg²⁺ with a probe radius of 1.4 Å. The protein regions described under “Results” that are visible in this orientation are colored cyan (Rossmann-like region, purple (DAD motif) and salmon (C-terminal α -helical region). The protein backbone is represented as a gold-colored C α tube diagram, the side chains and GDP-Man are drawn as sticks, and the metal ions and water molecules are drawn as spheres. Atom colors are light blue for carbon, blue for nitrogen, red for oxygen, green for phosphorus, orange for Mg²⁺, and gray for Zn/M²⁺. In this figure, part of Loop 1 (residues 205–215) not visible in the x-ray structure of MpgS-GDP-Man-Mg²⁺ was built using apo-MpgS chain A as a template. A, overall view showing the NDP-sugar binding pocket and the entrance to tunnels I and II; the orientation of the tunnels is indicated by white dotted lines; B, detailed view showing the tunnel II entrance. SP- α , sugar pocket α -face; SP- β , sugar pocket β -face. This figure was created with DINO.

interpreted in terms of a fully occupied metal site with a mixed Zn²⁺/Mg²⁺ population. On the other hand, the anomalous diffraction data provided no evidence for Zn²⁺ occupation at the catalytic Lewis acid position of Mg²⁺, coordinated to GDP-Man.

Kinetic assays of the MpgS mutants E251A and H309A showed negligible specific activities; none was detected for E251A, and only 0.3 ± 0.2 $\mu\text{mol}/\text{min}/\text{mg}$ was measured for H309A. These results, combined with the structural evidence for the absence of this metal site in the H309A MpgS mutant, unequivocally show the existence of a second catalytic metal ion in MpgS.

DISCUSSION

The Conserved Catalytic Scaffold and Its Chemical Significance for Glycosyl Transfer—The structure of *T. thermophilus* MpgS and comparison with its closest homologues highlight the molecular features underlying the functional specialization, namely for the synthesis of glucosylglycerate and mannosylglycerate. Despite sharing a common catalytic scaffold with other GT-A enzymes, discrete structural changes at the NDP binding pocket are relevant for nucleoside specificity, as exemplified by the steric discrimination for UDP-Glu over GDP-Glu in GpgS (42, 43) provided by the local secondary structure and the orientation of the Ser⁸¹ side chain (Fig. 3C). On the other hand, acceptor specificity is more complex because it is dependent on the scaffold plasticity that accompanies catalysis, as suggested by the different conformations of the flexible regions equivalent to Loops 1 and 2 in MpgS and by point mutations (9).

The previously assigned structural motif at the β -face of the sugar moiety for the retaining GT-A enzymes (9) is represented in MpgS by Lys¹⁴⁴ and Glu²⁷¹. In MpgS, Glu²⁷¹ is located near GDP-Man and interacts with the mannosyl moiety (Fig. 3A). In fact, two charged residues at the β -face of the sugar pocket

region are topologically conserved in both inverting and retaining GTs, binding either the sugar donor or its acceptor (Fig. 6).

Asn²⁴⁸, structurally equivalent to Gln¹⁸⁹ in α -1,4-galactosyltransferase, is thought to play a central role in the D_N^{*}A_{Nss} mechanism, where it is presumed to provide the mild nucleophilic character for the stabilization of the developing oxocarbenium ion transition state, without an effective covalent bond formation. However, in *T. thermophilus* and *P. horikoshii* MpgS, its side chain is not properly positioned for such a role, at least in the observed conformation of the sugar donor. Still, the role of Asn²⁴⁸ in catalysis should not be disregarded because it is part of a conserved sequence motif NXGEX_nHXH present in all annotated MpgS of the

GT55 family and may play a role in acceptor binding.

The unique orientation of the sugar moiety found in both MpgS structures places its anomeric carbon in close proximity to the conserved Asp¹⁶⁷ of the DXD motif (Fig. 6A). Such orientation was independent of the presence of the catalytic metal and may be representative of the GT55 family because it was also seen in the structure of *P. horikoshii* MpgS, despite the absence of a catalytic Lewis acid. Thus, it may be speculated that different strategies for oxocarbenium ion stabilization may exist for different GT families, leading to alternative pathways with the preservation of the net reaction stereochemistry. Herein we propose a more complete structural signature that tailors the stereo-specific outcome, comprising Asp¹⁶⁷-Lys¹⁴⁴-Glu²⁷¹, combined with the orientation (*in versus out*) of the glycosidic bond linking the β -phosphate to the anomeric carbon relative to the DXD motif, as summarized in Table 3, the mechanistic implications of which are discussed below.

Asp¹⁶⁷ as the Putative Catalytic Nucleophile in the Retaining Catalysis—In contrast with the relative high specificity observed for the NDP group (Fig. 3 and supplemental Table S2), the sugar pocket dimensions allow for binding of other monosaccharyl moieties as well as changes in orientation for any given sugar ring. The natural availability of these moieties is a restricting factor (52), and catalytic proficiency is likely to require a well defined sugar ring stereochemistry as well as a proper orientation of its scissile glycosidic bond toward the incoming proton donor, in a similar fashion to the stereospecific protonation in glycoside hydrolases (53). In the retaining GTs, the incoming acceptor is likely to act as a Brønsted acid, prior to its nucleophilic attack at the anomeric carbon center (12). In the MpgS structures, a lone pair of O^{1B} oxygen in GDP-Man is efficiently oriented toward the entrance of tunnel II, the expected entry point of 3-PG. Its C2(OH) group, once activated, attacks the mannose anomeric carbon C1' by nucleophilic substitution, either via the double displacement S_N2-like

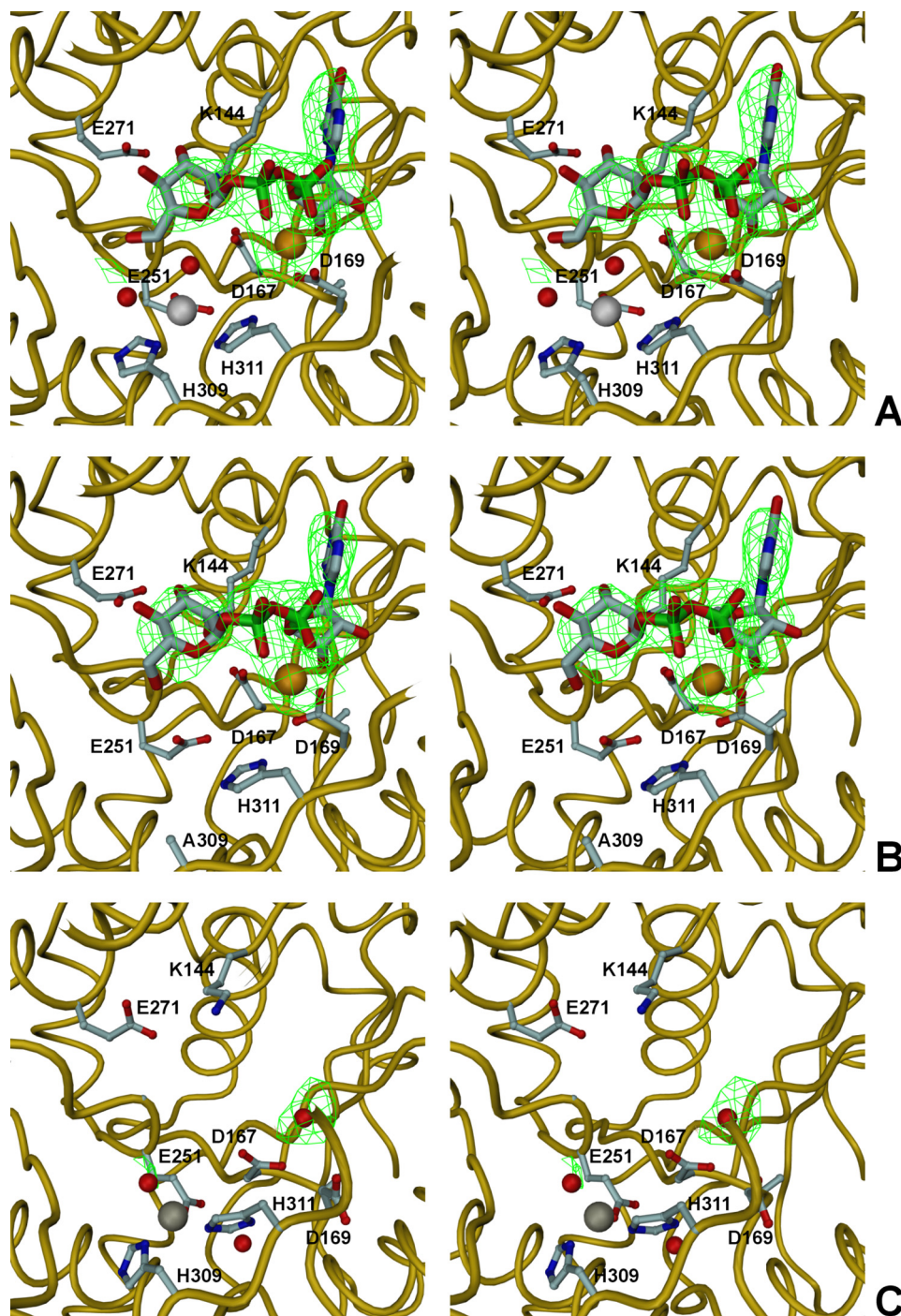


FIGURE 5. The Zn/Mg^{2+} binding site in the MpgS structures. Stereoviews showing the NDP-sugar binding pocket and the Zn/Mg^{2+} binding site in wtMpgS·GDP-Man· Mg^{2+} (A), H309A MpgS·GDP-Man· Mg^{2+} (B), and apo-MpgS (C). The protein backbone is represented as a gold-colored C^{α} tube diagram, GDP-Man is drawn as sticks, the side chains of the residues mentioned throughout are labeled and drawn as ball-and-stick representations, and the metal ions and water molecules are drawn as spheres. Atom colors are light blue for carbon, blue for nitrogen, red for oxygen, green for phosphorus, orange for Mg^{2+} , and gray for Zn/Mg^{2+} . In A and B, the $|F_o| - |F_c|$ electron density maps covering GDP-Man· Mg^{2+} after the first round of refinement, including only the protein chain and the Zn^{2+} ions, are colored green and drawn at the map 3 root mean square deviation level. In C, a similar $|F_o| - |F_c|$ electron density map reveals only a water molecule in the same region. This figure was prepared with DINO.

mechanism or by the $\text{D}_\text{N}^*\text{A}_\text{Nss}$ mechanism (Fig. 6A). In general, the breaking of the glycosidic bond occurs together with the formation of several oxocarbenium ion-like transition states and short-lived intermediaries (12). In the MpgS structures, this charge development can be stabilized by Asp¹⁶⁷ at C1' in

T. thermophilus or alternatively at O5' as observed in *P. horikoshii*. This residue is hydrogen-bonded to O2' in the mannopyranosyl ring as well as to N $^{\zeta}$ of the conserved Lys¹⁴⁴, and these interactions effectively fix the side-chain conformation of Asp¹⁶⁷ in a suitable orientation to act either as a nucleophilic stabilizer in the retaining $\text{D}_\text{N}^*\text{A}_\text{Nss}$ mechanism or as an effective catalytic nucleophile according to the double displacement $\text{S}_\text{N}2$ -like mechanism. In the latter case, the formation of a glycosyl-enzyme intermediate would imply a reorientation of the C2'(OH) epimer (most likely to a pseudoequatorial position) as the sugar-puckering evolves into a transition state conformation analogous to that proposed for several mannosidases (54).

Despite sharing a common scaffold with MpgS, the equivalent Asp residue in MgS and GpgS is hydrogen-bonded to O3' instead, and it can be argued that in those crystal structures, the sugar donor is not in an orientation leading to efficient catalysis. This may have been an artifact of co-crystallization with Mn^{2+} . However, the structural changes that very likely accompany entry of the acceptor into the active site pocket may drive the sugar ring back into an orientation suitable to catalysis according to one of the two proposed mechanisms.

An analysis of the orientation of the scissile bond in relation to the DXD motif found in both inverting and retaining GT-A enzymes showed that the *in* orientation is prevalent in the former, whereas in the latter, it was mostly observed in the *out* orientation (Table 3). The *in* orientation favors the presence of an internal proton donor, whereas the *out* orientation satisfies the stereochemistry required for an in-line protonation from the sugar acceptor. For example, in the inverting β -1,4-galactosyltransferase-1 (GalT1), the β -phosphate was shown to be in the *in* orientation, properly placed to be protonated by a conserved water molecule that interacts with the DXD motif and the metal-Lewis acid (Fig. 6B). Such protonation is thought to be mandatory for the irreversibility of the reaction, by preventing the reforming of

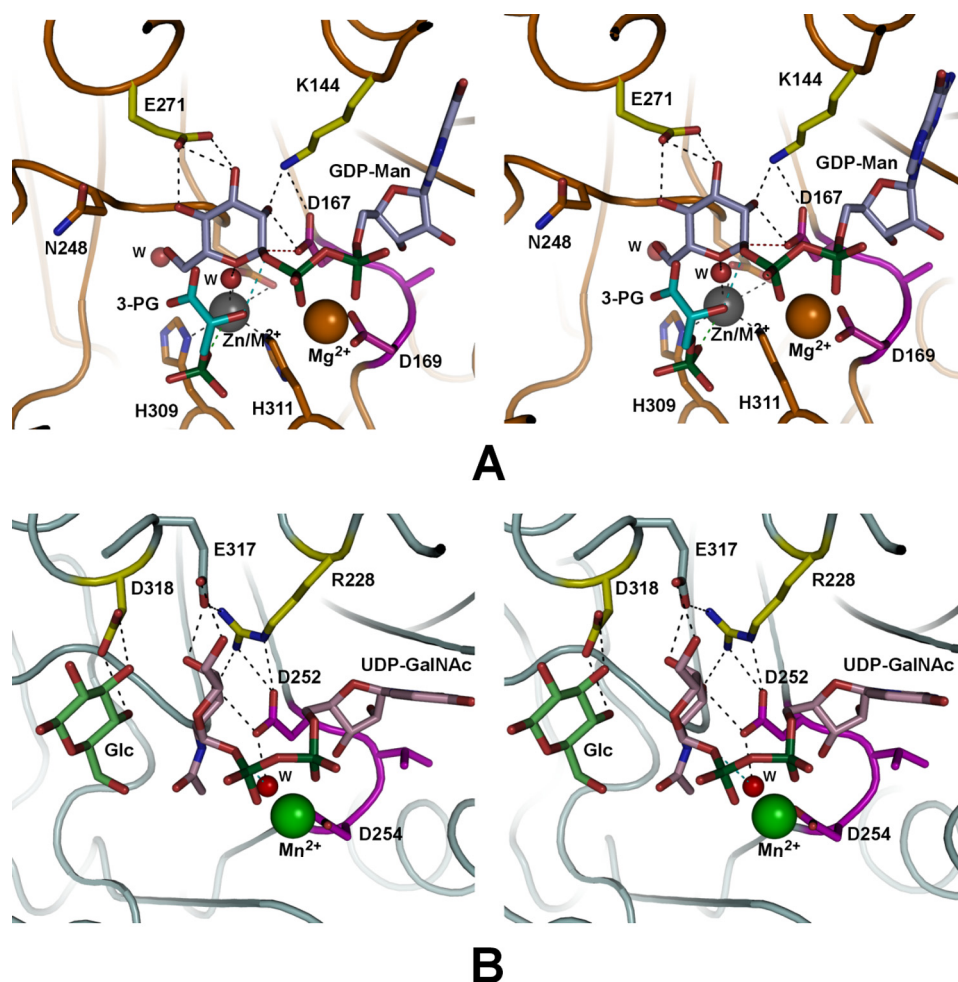


FIGURE 6. Structural signatures in retaining and inverting glycosyltransferases with GT-A fold. Shown is a comparison of the structural signatures present in two representative structures of the retaining and inverting GTs with GT-A fold, bound to their sugar donor. *A*, the predicted catalytic complex of *T. thermophilus* MpgS. The scissile glycosidic bond is oriented in the *out* configuration with respect to the DXD motif favoring an *in-line* protonation (cyan dashed line) from the incoming 3-PG, prior to its nucleophilic attack at the anomeric carbon center C1'. *B*, merged coordinates of the inverting Gal-T1 with its UDP-GalNAc donor (from PDB 1OQM) and Glc acceptor (from PDB 2FYD) oriented as MpgS. The scissile bond adopts the *in* configuration positioning the glycosidic oxygen to be protonated by the structural water molecule (W; cyan dashed line). The protein backbones are represented as C^α tube diagrams, bright orange for MpgS and light cyan for Gal-T1. The DXD motif is colored magenta in both molecules. The substrate molecules and the side chains of labeled protein residues are drawn as sticks, and the metal ions and water molecules are represented by spheres. Carbon atoms are colored yellow in the signature residues, magenta in the DXD motif, light blue in GDP-Man, cyan in 3-PG, light pink in UDP-GalNAc, and light green in Glc. Other atoms are colored red for oxygen, blue for nitrogen, dark green for phosphorus, orange for Mg²⁺, gray for Zn/Mg²⁺, and green for Mn²⁺. Hydrogen bonds (proposed for MpgS, considering the limited resolution of the diffraction data) are represented as dashed lines. This figure was prepared with PYMOL.

the glycosidic bond and thus facilitating the nucleophilic attack from the incoming acceptor at the β -face (51). Alternatively, our results suggest that, at least for some GT-A retaining enzymes, Asp¹⁶⁷ in the DXD motif can work as the catalytic nucleophile.

A Second Catalytic Metal Ion in MpgS Enzymes—In addition to the Lewis acid metal coordinated to GDP-Man, found in the *T. thermophilus* MpgS·GDP-Man·Mg²⁺ binary complex, our mutagenesis studies highlighted the requirement of a second catalytic metal ion for efficient catalysis in *T. thermophilus* MpgS. The coordination site is formed by Glu²⁵¹, His³⁰⁹, His³¹¹, and an axial water molecule. Equivalent residues are present in *P. horikoshii* MpgS; however, no metal was found in any of its structures (PDB 2ZU7, 2ZU8, and

2ZU9). This can be explained by the absence of any metal ion with chemistry equivalent to Zn²⁺ in the crystallization buffer, contrary to *T. thermophilus* MpgS.

The coordinating His³⁰⁹ is replaced by a lysine in MgS and by an arginine in both GpgSs and the MpgS/GpgS from *R. xylanophilus*. These residues are less likely to act as metal ion ligands, but their positively charged side chains may play a role parallel to that of the second catalytic metal ion in *T. thermophilus* MpgS.

Elimination of the Zn/Mg²⁺ binding site via an E251A or H309A mutation effectively destroyed enzyme activity, by abolishing the metal coordination ability, as seen in the H309A structure. This effect can be rationalized in terms of an interaction between one of the negatively charged phosphate oxygens of 3-PG and the positively charged second catalytic metal ion. Indeed, a simple manual modeling of a 3-PG molecule in its expected binding region shows that not only this interaction is possible, but also the C2(OH) group becomes oriented toward the glycosidic oxygen (less than 3 Å distant), as expected during general acid catalysis (Fig. 6A). Therefore, without this metal center, the inability for 3-PG binding in the correct orientation will hinder catalysis. This notion is supported by the structure of the putative GpgS from *M. avium paratuberculosis* (PDB 3CKJ), where a citrate molecule occupies a position similar to that of 3-PG modeled in *T. thermophilus* (Fig. 6A) and interacts

with Arg²⁶¹ through its γ -carboxylate group. This binding position could not be reproduced by our docking calculations (supplemental Fig. S2) because it requires a local rearrangement of the protein structure to eliminate close contacts (2.5–3 Å) with 3-PG.

Zn²⁺ was shown to act as an inhibitor of M-3-PG synthesis, competing with Mg²⁺ in the occupation of the second metal ion binding site. Because the crystallization buffer contained 0.4–0.6 mM Zn²⁺ and 200 mM Mg²⁺, the affinity of the second metal binding site for Zn²⁺ is clearly much higher than that for Mg²⁺. The inhibitory effect of Zn²⁺ may have two origins: 1) due to its acidity, which, via its axial water molecule destabilizes the oxocarbenium ion transition state, similarly to the 5-fluoro-sugar inhibitors or 2) due to a higher affinity for one of the

TABLE 3

Representative structures of inverting and retaining classes of glycosyltransferases with GT-A fold

α -GalT, α -1,3-galactosyltransferase; LgtC, α -1,4-galactosyltransferase; Gal-T1, β -1,4-galactosyltransferase-1; GnT1, *N*-acetylglucosaminyltransferase 1; GlcAT-I, β -1,3-glucoronyltransferase; Gal-T1, β -1,4-galactosyltransferase; UDP-2F-Gal, UDP-2-deoxy-2-fluorogalactose; GAL, β -D-galactose; DEL, 4-deoxygalactose; Bo, *Bos taurus*; Nm, *Neisseria meningitidis*; Tt, *T. thermophilus*; Ph, *P. horikoshii*; Oc, *Orytolagus cuniculus*; Rm, *R. marinus*; Map, *M. avium paratuberculosis*; Hs, *Homo sapiens*.

Enzyme (source)	Family	PDB	GT class ^a	DXD	Sugar donor in crystal structure	Acceptor in crystal structure	Proposed nucleophile	Structure Resolution Å	Glycosidic oxygen ^b	Reference/Source
α -GalT (Bo) ^c	GT6	1G93	ret	225DAD ²²⁷	UDP-2F-Gal	GAL	E317 ^d	2.4	<i>in</i>	Ref. 14
LgtC (Nm)	GT8	1GA8	ret	103DID ¹⁰⁵	UDP-2F-Gal	DEL	Q189 ^e	2.0	<i>out</i>	Ref. 13
Glycogenin 1 (Oc)	GT8	1LL2	ret	102DAD ¹⁰⁴	UDP-Glc		D163 ^d	1.9	<i>out</i>	Refs. 12 and 48
MpgS (Tt)	GT55	2WVL	ret	167DAD ¹⁶⁹	GDP-Man		D167 ^e	2.8	<i>out</i>	This work
MpgS (Ph)	GT55	2ZU8	ret	168DAD ¹⁷⁰	GDP-Man			2.4	<i>out</i>	T. Kawamura, unpublished data
MgS (Rm)	GT78	2BO8	ret	100DAD ¹⁰²	GDP-Man		Leu ^{163e}	2.8	<i>out</i>	Ref. 9
GpgS (Map)	GT81	3CKQ	ret	139DSD ¹⁴¹	GDP-Man		Leu ^{214e}	3.0	<i>out</i>	Ref. 43
GnT1 (Oc)	GT13	1FOA	inv	211EDD ²¹³	UDP-GlcNAc			1.8	<i>in</i>	Ref. 58
GlcAT-I (Hs)	GT43	1KWS	inv	194DDD ¹⁹⁶	UDP-GlcUA			2.1	<i>in</i>	Ref. 59
Gal-T1 (Bo)	GT7	1OQM	inv	252DVD ²⁵⁴	UDP-GalNAc			2.3	<i>in</i>	Ref. 45

^a ret, retaining GT; inv, inverting GT.

^b Orientation of the electron lone pairs in the glycosidic oxygen (linking the β -phosphate with the anomeric C1' carbon) with respect to the DXD motif: *in*, one pair toward DXD; *out*, the other pair toward the acceptor.

^c Behaves as an inverting GT in the presence of small nucleophile molecules, like azide and water (14).

^d In double displacement S_N2-like mechanism.

^e In front face D_N*A_{Nss} mechanism.

reaction products, preventing its release from the catalytic pocket and therefore slowing down the reaction turnover. Also, under physiological conditions, the second metal ion binding site may be labile. Indeed, the degree of structural flexibility in this region composed of flexible Loops 1 and 2 may transiently shape this site for binding a second metal ion during the catalytic cycle. In particular, His³¹¹ in *T. thermophilus* MpgS is positioned in the vicinity of Loop 2 and may coordinate either one of the two catalytic metal ions. This is reinforced by the alternate conformations of the equivalent residue His³¹² in two *P. horikoshii* structures, interacting with either GDP-Man (PDB 2ZU8; Fig. 3B) or GDP-Co²⁺ (PDB 2ZU9). In the latter structure, the imidazole group of His³¹² moves away from the second catalytic metal site and coordinates instead the Lewis acid metal of the GDP-Co²⁺ leaving group, possibly destabilizing the second catalytic metal center, leading to the release of its bound metal ion. This release most likely follows that of the M-3-PG product, whereas the GDP-M²⁺ leaving group is stabilized by residues equivalent to His³¹¹ in *T. thermophilus* MpgS.

CONCLUSIONS

The three structures of *T. thermophilus* HB27 MpgS reported herein constitute the first published structural representatives of a glycosyltransferase from the GT55 family. Their sugar-binding mode suggests a possible role for Asp¹⁶⁷ in the DXD motif as the catalytic nucleophile in retaining GTs, in light of both the S_N2-like double displacement and the D_N*A_{Nss} mechanisms. Structural and biochemical evidence is presented, for the first time, for a second catalytic metal ion, which is likely to be relevant for catalysis in all MpgS enzymes. Finally, structural comparison with retaining and inverting GTs led to the proposal of an extended structural signature, formed by residues Lys¹⁴⁴, Asp¹⁶⁷, and Glu²⁷¹. These residues very likely belong to a common structural scaffold between retaining and inverting GTs, which, combined with the orientation of the glycosidic bond linking the β -phosphate to the anomeric sugar carbon in relation to

the DXD motif, can determine the stereo-specific outcome of the glycosyl transfer reaction.

Acknowledgments—We thank the ESRF for support with data collection and M. S. da Costa (University of Coimbra) for the generous gift of a plasmid containing the gene encoding *T. thermophilus* HB27 MpgS. Helpful suggestions from Clemens Vornrhein (Global Phasing Ltd.) and Kevin Cowtan (University of York) regarding the SHARP and BUCCANEER calculations are also gratefully acknowledged. Figures were prepared with PyMOL and DINO.

REFERENCES

- Santos, H., Lamosa, P., Faria, T. Q., Borges, N., and Neves, C. (2007) in *Physiology and Biochemistry of Extremophiles* (Gerday, C., and Glansdorff, N., eds) pp. 86–103, American Society for Microbiology Press, Washington, D. C.
- Borges, N., Ramos, A., Raven, N. D., Sharp, R. J., and Santos, H. (2002) *Extremophiles* 6, 209–216
- Faria, T. Q., Mingote, A., Siopa, F., Ventura, R., Maycock, C., and Santos, H. (2008) *Carbohydr. Res.* 343, 3025–3033
- Santos, H., Lamosa, P., Faria, T. Q., Pais, T. M., de La Paz, M. L., and Serrano, L. (2008) in *Thermophiles: Biology and Technology at High Temperatures* (Robb, F., Antranikian, G., Grogan, D., and Driessen, A., eds) pp. 9–24, CRC Press, London
- Alarico, S., Empadinhas, N., Mingote, A., Simões, C., Santos, M. S., and da Costa, M. S. (2007) *Extremophiles* 11, 833–840
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) *Nucleic Acids Res.* 37, D233–D238
- Martins, L. O., Empadinhas, N., Marugg, J. D., Miguel, C., Ferreira, C., da Costa, M. S., and Santos, H. (1999) *J. Biol. Chem.* 274, 35407–35414
- Empadinhas, N., Marugg, J. D., Borges, N., Santos, H., and da Costa, M. S. (2001) *J. Biol. Chem.* 276, 43580–43588
- Flint, J., Taylor, E., Yang, M., Bolam, D. N., Tailford, L. E., Martinez-Fleites, C., Dodson, E. J., Davis, B. G., Gilbert, H. J., and Davies, G. J. (2005) *Nat. Struct. Mol. Biol.* 12, 608–614
- Withers, S. G., and Aebersold, R. (1995) *Protein Sci.* 4, 361–372
- Zechel, D. L., and Withers, S. G. (2000) *Acc. Chem. Res.* 33, 11–18
- Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) *Ann. Rev. Biochem.* 77, 521–555
- Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G., and Strynadka, N. C. (2001) *Nat. Struct. Biol.* 8, 166–175
- Jamaluddin, H., Tumbale, P., Withers, S. G., Acharya, K. R., and Brew, K.

- (2007) *J. Mol. Biol.* **369**, 1270–1281
15. Coutinho, P. M., Deleury, E., Davies, G. J., and Henrissat, B. (2003) *J. Mol. Biol.* **328**, 307–317
16. Lairson, L. L., and Withers, S. G. (2004) *Chem. Commun.* **40**, 2243–2248
17. Goncalves, S., Borges, N., Santos, H., and Matias, P. M. (2009) *Acta Crystallogr. F* **65**, 1014–1017
18. Kabsch, W. (1993) *J. Appl. Crystallogr.* **26**, 795–800
19. Pape, T., and Schneider, T. R. (2004) *J. Appl. Crystallogr.* **37**, 843–844
20. Schneider, T. R., and Sheldrick, G. M. (2002) *Acta Crystallogr. D* **58**, 1772–1779
21. Sheldrick, G. M. (2002) *Z. Kristallogr.* **217**, 644–650
22. La Fortelle, E. d., and Bricogne, G. (1997) *Methods Enzymol.* **276**, 472–494
23. Abrahams, J. P., and Leslie, A. G. (1996) *Acta Crystallogr. D* **52**, 30–42
24. Cowtan, K. (1994) in *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, Vol. 31, pp. 34–38, Daresbury Laboratory, Warrington, UK
25. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D* **50**, 760–763
26. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) *Acta Crystallogr. D* **59**, 1131–1137
27. Kabsch, W. (1976) *Acta Crystallogr. A* **32**, 922–923
28. Brünger, A. T. (1992) *Nature* **355**, 472–475
29. Cowtan, K. (2006) *Acta Crystallogr. D* **62**, 1002–1011
30. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D* **60**, 2126–2132
31. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D* **53**, 240–255
32. Storoni, L. C., McCoy, A. J., and Read, R. J. (2004) *Acta Crystallogr. D* **60**, 432–438
33. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) *Acta Crystallogr. D* **57**, 122–133
34. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283–291
35. Hutchinson, E. G., and Thornton, J. M. (1996) *Protein Sci.* **5**, 212–220
36. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res.* **28**, 235–242
37. Borges, N., Marugg, J. D., Empadinhas, N., da Costa, M. S., and Santos, H. (2004) *J. Biol. Chem.* **279**, 9892–9898
38. Empadinhas, N., Albuquerque, L., Mendes, V., Macedo-Ribeiro, S., and da Costa, M. S. (2008) *FEMS Microbiol. Lett.* **280**, 195–202
39. Ames, B. N. (1966) *Methods Enzymol.* **8**, 115–118
40. Richardson, J. S. (1977) *Nature* **268**, 495–500
41. Holm, L., Kääriäinen, S., Rosenström, P., and Schenkel, A. (2008) *Bioinformatics* **24**, 2780–2781
42. Pereira, P. J., Empadinhas, N., Albuquerque, L., Sá-Moura, B., da Costa, M. S., and Macedo-Ribeiro, S. (2008) *PLoS ONE* **3**, e3748
43. Fulton, Z., McAlister, A., Wilce, M. C., Brammananth, R., Zaker-Tabrizi, L., Perugini, M. A., Bottomley, S. P., Coppel, R. L., Crellin, P. K., Rossjohn, J., and Beddoe, T. (2008) *J. Biol. Chem.* **283**, 27881–27890
44. Sá-Moura, B., Albuquerque, L., Empadinhas, N., da Costa, M. S., Pereira, P. J., and Macedo-Ribeiro, S. (2008) *Acta Crystallogr. F* **64**, 760–763
45. Ramakrishnan, B., and Qasba, P. K. (2002) *J. Biol. Chem.* **277**, 20833–20839
46. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) *J. Mol. Biol.* **247**, 536–540
47. Krissinel, E., and Henrick, K. (2007) *J. Mol. Biol.* **372**, 774–797
48. Gibbons, B. J., Roach, P. J., and Hurley, T. D. (2002) *J. Mol. Biol.* **319**, 463–477
49. Empadinhas, N., Albuquerque, L., Henne, A., Santos, H., and da Costa, M. S. (2003) *Appl. Environ. Microbiol.* **69**, 3272–3279
50. Ramakrishnan, B., Boeggeman, E., Ramasamy, V., and Qasba, P. K. (2004) *Curr. Opin. Struct. Biol.* **14**, 593–600
51. Ramakrishnan, B., Ramasamy, V., and Qasba, P. K. (2006) *J. Mol. Biol.* **357**, 1619–1633
52. Bülter, T., and Elling, L. (1999) *Glycoconj. J.* **16**, 147–159
53. Tom, D. H., and Andrea, T. V. (1999) *Angew. Chem. Int. Ed.* **38**, 750–770
54. Vasella, A., Davies, G. J., and Böhm, M. (2002) *Curr. Opin. Chem. Biol.* **6**, 619–629
55. Sanner, M. F., Olson, A. J., and Spehner, J. C. (1996) *Biopolymers* **38**, 305–320
56. Diederichs, K., and Karplus, P. A. (1997) *Nat. Struct. Biol.* **4**, 269–275
57. Matthews, B. W. (1968) *J. Mol. Biol.* **33**, 491–497
58. Unligil, U. M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H., and Rini, J. M. (2000) *EMBO J.* **19**, 5269–5280
59. Pedersen, L. C., Darden, T. A., and Negishi, M. (2002) *J. Biol. Chem.* **277**, 21869–21873