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

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

*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 Mg2+, 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 manno pyranose ring, suggesting its action as a catalytic nucleophile, either in the formation of a glycosyl enzyme intermediate according to the double-displacement SN2 reaction mechanism or in the stabilization of the oxocarbenium ion-like intermediate according to the $D_N^\alpha A_{N\alpha}$ 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) (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 α-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 synthase 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 $S_N^\alpha$-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 $S_N^2$-like reaction occurring in retaining glycoside hydrolases (9). Instead, a front-face $S_N^\alpha$-like mechanism was proposed for retaining GTs and later designated $D_N^\alpha A_{N\alpha}$ 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.

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, α-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.
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 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\textsubscript{ni} i-like mechanism by the position of Gln\textsuperscript{198} (13). Alternatively, studies on the bovine α-1,3-galactosyltransferase point toward formation of a glycosyl-enzyme intermediate, with Asp\textsuperscript{167} 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\textsubscript{ni} i-like cannot be totally disregarded for both inverting and retaining GTs. 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, Asp\textsuperscript{167} 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-μl drops (protein/crystalization 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 mM magnesium acetate, 0.1 mM sodium cacodylate, pH 6.5, 30–35% 2-methyl-2,4-pentanediol and 600 μM ZnCl\textsubscript{2} as additive.

Crystals of the binary complexes wild-type MpgS (wtMpgS)-GDP-Man-Mg\textsuperscript{2+} and H309A MpgS-GDP-Man-Mg\textsuperscript{2+} 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 μm 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-Mg\textsuperscript{2+} at 100 K on ESRF beamline ID-29. A data set from another wtMpgS-GDP-Man-Mg\textsuperscript{2+} 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-Mg\textsuperscript{2+} 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 Zn\textsuperscript{2+} heavy atom structure was determined with SHELXD (20), and the phase problem was solved with SHELXE (21). The best solution from SHELXD in 100 trials gave five Zn\textsuperscript{2+} 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 P\textsubscript{4}1\textsubscript{2}1\textsubscript{2}. 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
Structural Analysis of T. thermophilus HB27 MpgS

The structures of apo-MpgS and H309A MpgS 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-rotation-screw rigid body refinement (33) was carried out prior to restrained refinement of atomic positions and isotropic thermal motion parameters. One rigid body was defined for each of the two independent molecules in the asymmetric unit, including the 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, His314–Glu316, 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|Fc|−|Fo| and |Fc|−|Fo| 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 r2wksf for apo-MpgS, 2wvl and r2wlvsf for wtMpgS-GDP-Man-Mg²⁺, and 2wmv and 2wvmfsf 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 mphS 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 GGG GCG CAC GCC ATG AG-3’) and E251A-antisense (5’-CT

### Table 1

| Structure                        | Peak | Inflection | Remote |
|----------------------------------|------|------------|--------|
| wtMpgS-GDP-Man-Mg²⁺ MAD data set |      |            |        |
| wtMpgS-GDP-Man-Mg²⁺              |      |            |        |
| H309A MpgS-GDP-Man-Mg²⁺          |      |            |        |

*a R-merge = merging R-factor, (Σ[I(hkl)]²−(I(hkl))²)/((Σ[I(hkl)]²) x 100%.
*b R-meas = redundancy-independent R-factor (intensities), Σ[I(hkl)]²−(I(hkl))²/((Σ[I(hkl)]²) x 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 independent molecules in the asymmetric unit could be clearly discerned, and a better NCS operator was derived, correspondingly. 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 supplementary 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 were determined by molecular replacement with PHASER (32) using the refined protein chain coordinates of wtMpgS-GDP-Man-Mg²⁺.
CAT GGC GTG CGC CCC GGC GTT CG-3

/H11032

) for the E251A substitution and the primers H309A-sense (5′/H11032-ACC CGG AAC CCC GCC CTC CAC GAG AAC-3′/H11032) and H309A-anti-sense (5′/H11032-GTT CTC GTG GAG GGC GGG GTT CCG GGT-3′/H11032) 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 (\(T_m = 81°C\)): \(T_m = 82°C\) for the E251A mutant protein and \(T_m = 81°C\) 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 chromatographic 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 (\(\mu\)mol/min/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 \(\mu\)g of pure wtMpgS (total volume of 50 \(\mu\)l) with an excess (5 mM) of GDP-Man and 3-PG in buffer

| TABLE 2 |
|-----------------|-----------------|-----------------|
| Final refinement statistics for the three MpgS crystal structures |
| | wild-type apoMpgS | wild-type MpgS:GDPM-Man:Man \(\text{GDP}^2\) | H309A MpgS:GDPM-Man:Man \(\text{GDP}^2\) |
| 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 (%) | 0.190 (0.315) | 0.176 (0.339) | 0.174 (0.261) |
| nR: reflections, no \(\sigma(F_o)\) cutoff | 25502 (1558) | 30259 (2142) | 24263 (1155) |
| Free R-factor (%) | 0.256 (0.369) | 0.234 (0.426) | 0.232 (0.299) |
| nR: reflections, no \(\sigma(F_o)\) cutoff | 1358 (81) | 1620 (115) | 1292 (51) |
| Overall coordinate error estimate (Å) | 0.32 | 0.27 | 0.27 |

Model composition

- non-hydrogen protein atoms: 6076
- GDP-Man: 78
- Zn\(^{2+}\): 5
- Mg\(^{2+}\): 2
- citrate: 13
- Solvent molecules: 80

Model r.m.s. deviations from ideality

- Bond lengths (Å): 0.015
- Bond angles (°): 1.7
- Chiral centers (Å\(^2\)): 0.10
- Planar groups (Å): 0.008

Model completeness and validation

| Regions omitted | Chain A | Chain B |
|-----------------|---------|---------|
| protein main-chain | 107 | 102 |
| protein side-chain | 110 | 105 |
| GDP-Man | - | 115 |
| Mg\(^{2+}\), Zn\(^{2+}\) | 88.9 | 99.9 |
| citrate | 106.2 | - |
| solvent | 65.6 | 64.9 |

Ramachandran plot statistics. Residues in:

- most favored regions (%): 85.7
- allowed regions (%): 13.1
- generously allowed regions (%): 0.9
- disallowed regions (%): 0.3

PROCHECK G-factor:

- Chain A: -0.15
- Chain B: -0.19

\(\text{R-factor} = \sum_{i<j} |F_{ij} - |F_{ij}\| \Sigma_{i,j} |F_{ij}|\), where \(|F_{ij}|\) and \(|F_{ij}|\) are the observed and calculated structure factor amplitudes, respectively.

\(\text{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.}

\(\text{Maximum likelihood estimate with REFMAC.}

\(\text{Calculated from equivalent isotropic } B\text{ values, including the translation-libration-screw contribution for the protein and GDP-\(\alpha\)-\(\delta\)mannose atoms.}

\(\text{Structural Analysis of } T. \text{ thermophilus HB27 MpgS}\)
We verified that the MpgP activity was not affected by Zn^{2+}

The effect of Zn^{2+} on the specific activity of wtMpgS was studied by performing the same assays, using 300 μM of ZnCl_2; moreover, the competing effect of Zn^{2+} (300 μM) was examined at a Mg^{2+} (3 mM) concentration of 20 mM. We verified that the MpgP activity was not affected by Zn^{2+} or Mg^{2+} at these concentrations. The specific activities of the MpgS mutant proteins were determined as for the wtMpgS, using instead 5 μg 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-Mg^{2+} 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 β2, β3, β12, β10, β11, β9, β6, β7, and β8, with β2 and β11 running in an anti-parallel direction, and its connectivity (40) can be described as 0 1 7 X 1 X 1 X −3 X −2 X 1 −2, considering strand β2 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 β1–β5, and a Rossmann-like α/β/α sandwich region (residues 55–172), both linked to a central helix α1 (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 (Asp^{167}–Ala^{168}–Asp^{169}), 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 α7, α9, α11, and α12, 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 22U7). 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 MsS-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 β12, β10, β11, β9, β6, β7, and β8 (MpgS nomenclature; supplemental Fig. S4) as a reminiscent feature from a common ancestor.

Unlike the α/β/α 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
The tight packing of the N-terminal by a 2-fold NCS axis at the dimer interface, which is defined by 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 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 beginning of the long C-terminal tail.

The Catalytic Pocket—In the crystal structure of GDP-Man-Mg\textsuperscript{2+} 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-Mg\textsuperscript{2+}, a second metal ion position partially occupied by Zn\textsuperscript{2+} lies about 6 Å below the β-face of the mannose ring.

A number of hydrophilic residues participate in GDP-Man-Mg\textsuperscript{2+} 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 Pro\textsuperscript{60}–Glu\textsuperscript{64} 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 Asp\textsuperscript{167} as well as in the stabilization of the ribose moiety. In addition, through the syn-coordination of the catalytic Lewis acid Mg\textsuperscript{2+} by Asp\textsuperscript{169}, 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 Gly\textsuperscript{250} 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 Tyr\textsuperscript{268} and to a lesser extent by Met\textsuperscript{322}. In both the wild-type and H309A mutant binary complexes, the mannopyranosyl moiety was found in the typically stable 4C\textsubscript{1} 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, Asp\textsuperscript{167} O\textsuperscript{2-} accepts a hydrogen bond from O\textsuperscript{2-} and is also within van der Waals distance (∼3.5 Å) from the anomeric carbon C1’ (Fig. 3A) or alternatively from the OS’ 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, Mg\textsuperscript{2+}, or Mn\textsuperscript{2+}) 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-
certainly and/or synergistically upon GDP-ManMg\(^{2+}\) 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 ∼75 °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\(^{-}\), His309 N\(^{1}\), and His311 N\(^{2}\), and an axial water molecule completes the coordination sphere, although it is not visible in all monomers. These residues belong to the sequence motif N\(X\)GE\(X_{n}\)H, conserved in all MpgS enzymes with known sequence.

The functional relevance of Zn\(^{2+}\) was assessed from measurements of maximal specific activity (\(V_{\text{max}}\)) 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 µmol/min/mg. When measured solely in the presence of Zn\(^{2+}\) (300 µM), the activity decreased about 100-fold (0.4 ± 1.2 µmol/min/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 µmol/min/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 wtMpgSGDP-ManMg\(^{2+}\) and apo-MpgS structures may also be

![FIGURE 3. The NDP-sugar binding pocket in MpgS and structurally related enzymes.](image-url)
interpreted in terms of a fully occupied metal site with a mixed Zn$^{2+}$/Mg$^{2+}$ population. On the other hand, the anomalous difference data provided no evidence for Zn$^{2+}$ occupation at the catalytic Lewis acid position of Mg$^{2+}$.

Kinetic assays of the MpgS mutants E251A and H309A showed negligible specific activities; none was detected for E251A, and only 0.3 ± 0.2 μmol/min/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 Ser$^{248}$ 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 Lys$^{144}$ and Glu$^{271}$. In MpgS, Glu$^{271}$ 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).

Asn$^{248}$, structurally equivalent to Gln$^{189}$ in α-1,4-galactosyltransferase, is thought to play a central role in the D$_{1,2}$A$_{	ext{NNa}}$ 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 Asn$^{248}$ in catalysis should not be disregarded because it is part of a conserved sequence motif NXXGEX$_{X}$HXH 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 Asp$^{167}$ 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 Asp$^{167}$-Lys$^{144}$-Glu$^{271}$, 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.

Asp$^{167}$ 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 Bronsted acid, prior to its nucleophilic attack at the anomeric carbon center (12). In the MpgS structures, a lone pair of O$^{18}$ 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 manno anomeric carbon C1’ by nucleophilic substitution, either via the double displacement $S_N$2-like
mechanism or by the $D{\text{N}}^+.A_{\text{Nss}}$ 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 Asp$^{167}$ at C1’ in $T.\text{thermophilus}$ or alternatively at O5’ as observed in $P.\text{horikoshii}$. This residue is hydrogen-bonded to O2’ in the mannopyranosyl ring as well as to N6 of the conserved Lys$^{144}$, and these interactions effectively fix the side-chain conformation of Asp$^{167}$ in a suitable orientation to act either as a nucleophilic stabilizer in the retaining $D{\text{N}}^+.A_{\text{Nss}}$ mechanism or as an effective catalytic nucleophile according to the double displacement $S_N2$-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 $\beta$-1,4-galactosyltransferase-1 (Gal-T1), the $\beta$-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
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-Mg\(^{2+}\) 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 Zn\(^{2+}\) 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/Mg\(^{2+}\) 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.

Zn\(^{2+}\) was shown to act as an inhibitor of M-3-PG synthesis, competing with Mg\(^{2+}\) in the occupation of the second metal ion binding site. Because the crystallization buffer contained 0.4–0.6 mM Zn\(^{2+}\) and 200 mM Mg\(^{2+}\), the affinity of the second metal binding site for Zn\(^{2+}\) is clearly much higher than that for Mg\(^{2+}\). The inhibitory effect of Zn\(^{2+}\) 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
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, His\(^{311}\) in *T. thermophilus* MpgS is positioned in the vicinity of Loop 2 and may coordinate either of the two catalytic metal ions. This is reinforced by the alternate conformations of the equivalent residue His\(^{312}\) in two *P. horikoshii* structures, interacting with either GDP-Man (PDB 2ZU8; Fig. 3B) or GDP-Co\(\text{e}^{2-}\) (PDB 2ZU9). In the latter structure, the imidazole group of His\(^{312}\) moves away from the second catalytic metal site and coordinates instead the Lewis acid metal of the GDP-Co\(\text{e}^{2-}\) 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-Man leaving group is stabilized by residues equivalent to His\(^{311}\) 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 Asp\(^{167}\) in the DXD motif as the catalytic nucleophile in retaining GTs, in light of both the S\(_{N}2\)-like double displacement and the D\(_{N}A_{\text{gas}}\) 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 Lys\(^{144}\), Asp\(^{167}\), and Glu\(^{271}\). 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.

**REFERENCES**

1. 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.
2. Borges, N., Ramos, A., Raven, N. D., Sharp, R. J., and Santos, H. (2002) *Extremophiles* 6, 209—216
3. Faria, T. Q., Mingote, A., Siopa, F., Ventura, R., Maycock, C., and Santos, H. (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
4. Alarico, S., Empadinhas, N., Mingote, A., Simões, C., Santos, M. S., and da Costa, M. S. (2007) *Extremophiles* 11, 833—840
5. Cattarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) *Nucleic Acids Res.* 37, D233—D238
6. Martins, I. O., Empadinhas, N., Marugg, J. D., Miguel, C., Ferreira, C., da Costa, M. S., and Santos, H. (1999) *J. Biol. Chem.* 274, 35407—35414
7. Empadinhas, N., Marugg, J. D., Borges, N., Santos, H., and da Costa, M. S. (2001) *J. Biol. Chem.* 276, 43580—43588
8. Flint, J., Taylor, E., Yang, M., Bolam, D. N., Taillöf, 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
9. Wither, S. G., and Aebbersold, R. (1995) *Protein Sci.* 4, 361—372
10. Zechel, D. L., and Wither, S. G. (2000) *Acc. Chem. Res.* 33, 11—18
11. Lairson, L. L., Henrissat, B., Davies, G. J., and Wither, S. G. (2008) *Ann. Rev. Biochem.* 77, 521—555
12. Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Wither, S. G., and Strynadka, N. C. (2001) *Nat. Struct. Mol. Biol.* 8, 166—175
13. Jamaluddin, H., Tumbale, P., Wither, S. G., Acharya, K. R., and Brew, K.
