Role of Glycoside Phosphorylases in Mannose Foraging by Human Gut Bacteria

Simon Ladevêze, Laurence Tarquis, Davide A. Cecchini, Juliette Bercovic, Isabelle André, Christopher M. Topham, Sandrine Morel, Elisabeth Laville, Pierre Monsan, Vincent Lombard, Bernard Henrissat, and Gabrielle Potocki-Véronèse

From the Institute National des Sciences Appliquées, Université Paul Sabatier, Institut National Polytechnique, Laboratoire d’Ingénierie des Systèmes Biologiques et des Procédés, Université de Toulouse, 135 Avenue de Rangueil, F-31077 Toulouse, the CNRS, UMR5504, F-31400 Toulouse, the Institut National de Recherche Agronomique, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, F-31400 Toulouse, and the Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille Université, CNRS UMR 7257, 163 Avenue de Luminy, F-13288 Marseille, France

Background: The relations between the gut microbiota, food, and host play a crucial role in human health.

Results: Prevalent bacterial glycoside phosphorylases are able to break down dietary carbohydrates and the N-glycans lining the intestinal epithelium.

Conclusion: GH130 enzymes are new targets to study interactions between host and gut microbes.

Significance: Glycoside phosphorylases are key enzymes of host glycans catabolism by gut bacteria.

To metabolize both dietary fiber constituent carbohydrates and host glycans lining the intestinal epithelium, gut bacteria produce a wide range of carbohydrate-active enzymes, of which glycosidase hydrolases are the main components. In this study, we describe the ability of phosphorylases to participate in the breakdown of human N-glycans, from an analysis of the substrate specificity of UhgbMP, a mannoside phosphorylase of the GH130 protein family discovered by functional metagenomics. UhgbMP is found to phosphorylize β-D-Mann-1,4-D-β-D-GlcNAc-1,4-D-β-D-GlcNAc and is also a highly efficient enzyme to catalyze the synthesis of this precious N-glycan core oligosaccharide by reverse phosphorylation. Analysis of sequence conservation within family GH130, mapped on a three-dimensional model of UhgbMP and supported by site-directed mutagenesis results, revealed two GH130 subfamilies and allowed the identification of key residues responsible for catalysis and substrate specificity. The analysis of the genomic context of 65 known GH130 sequences belonging to human gut bacteria indicates that the enzymes of the GH130 1 subfamily would be involved in mannann catabolism, whereas the enzymes belonging to the GH130 2 subfamily would rather work in synergy with glycosidase hydrolases of the GH92 and GH18 families in the breakdown of N-glycans. The use of GH130 inhibitors as therapeutic agents or functional foods could thus be considered as an innovative strategy to inhibit N-glycan degradation, with the ultimate goal of protecting, or restoring, the epithelial barrier.

The human gut microbiota is a dense and complex ecosystem, which plays a crucial role in maintaining human health. The many meta-omic studies carried out in the last few years have shown that certain metabolic diseases, such as obesity, and some inflammatory diseases, such as inflammatory bowel diseases (IBD) (1–4), are associated with structural and functional imbalances of the microbiota (5). The catabolism of complex carbohydrates by gut bacteria plays a key role in the microbial colonization and equilibrium of the digestive tract (6, 7). Indeed, the plant polysaccharides (cellulose, hemicelluloses, pectin, and resistant starch) that compose dietary fiber supply the main source of carbon for gut bacteria, in the form of monosaccharides produced by polysaccharide breakdown, such as glucose, xylose, arabinose, uronic acids, and to a lesser extent, galactose, mannose, and rhamnose (8). In particular, mannose is found in plant cell walls, in mannan and glucomannan, the backbone that consists of a homopolymer of β1,4-linked β-D-mannopyranosyl residues and a heterogeneous sequence of β1,4-linked β-D-glucopyranosyl and β-D-mannopyranosyl units (9). There is, however, a further source of complex carbohydrates in the gut, namely the heavily O- and N-glycosylated glycoproteins that line the intestinal epithelium to form a protective barrier against pathogens and chemical and mechanical aggression (10). Although host O-glycans are rich in N-acetylated hexosamines (GlcNAc, GalNAc, and Neu5Ac), galactose and fucose, many mature N-glycans contain eight α- and β-linked β-D-man- noppyranosyl residues linked to chitobiose (Fig. 1) (11). Human glycans-microbial interactions in the gastrointestinal tract play a crucial role in determining the outcome of relations of both commensals and pathogens with the host. Indeed, alterations in

The abbreviations used are: IBD, inflammatory bowel disease; GT, glycosyltransferase; GH, glycoside hydrolase; PL, polysaccharide lyase; GP, glycoside phosphorylase; DP, degree of polymerization; CE, carbohydrate esterase; BFMP, B. fragilis NCTC 9343 mannosylglucose phosphorylase; HPAEC-PAD, high performance anion exchange chromatography with pulsed amperometric detection; PDB, Protein Data Bank; pNP, p-nitrophenol.

* This work was supported by the French Ministry of Higher Education and Research and by the French National Institute for Agricultural Research (INRA). This article contains supplemental Figs. S1–S6 and Tables 1 and 2.

1 This article contains supplemental Figs. S1–S6 and Tables 1 and 2.
the structure and/or quantity of host glycans due to biosynthetic defects or microbial degradation alter their barrier function and are thought to be involved in the initiation and the maintenance of mucosal inflammation in IBDs and in the development of intestinal cancer (12).

To breakdown these complex carbohydrates of either plant or human origin, gut bacteria produce a full repertoire of carbohydrate-active enzymes (CAZymes, listed in the CAZy database (13)), of which glycoside hydrolases (GHs) are the main constituents, as revealed by Gill et al. (14). In particular, the degradation of dietary mannans requires both endo-β-mannanases and β-mannosidases to release mannose (15). The hydrolysis of N-glycans by a broad consortium of endo- and exo-β-glycosidases, in particular those produced by the prominent gut bacterium Bacteroides thetaiotaomicron (16–20), has also been described. Until August, 2013, only glycoside hydrolases (GHs) have been implicated in N-glycan breakdown.

However, other types of CAZymes participate in the breakdown of complex carbohydrates, particularly those of plant origin, by working in synergy with GHs as follows: carbohydrate esterases (CE), polysaccharide lyases (PLs), and glycoside phosphorylases (GPs). The presence of CEs and PLs is readily detectable in gut bacterial genomes, metagenomes, and metatranscriptomes, because they are sufficiently divergent from GHs as to be classified in their own CAZyme families. This is not the case for GPs, which are found both in the glycosyltransferase (GT) and glycoside hydrolase families, depending on the sequence and catalytic mechanism similarities shared with GT and GH archetypes, respectively. The prevalence of GPs and their role in the metabolism of carbohydrates is therefore difficult to evaluate on the basis of sequence data alone, as a given family of GHs (or of GTs) may include both GHs (or GTs) as well as GPs. GPs catalyze the breakdown of a glycosidic linkage from oligosaccharide or polysaccharide substrates with concomitant phosphate glycosylation to yield a glycosyl-phosphate product and a sugar chain of reduced length. These enzymes are also able to perform reverse phosphorolysis (in the so-called “synthetic reaction”) to form a glycosidic bond between the glycosyl unit originating from the glycosyl-phosphate, which acts as the sugar donor, and a carbohydrate acceptor (23). Retaining GPs, for which phosphorolysis occurs via overall retention of the substrate anomeric configuration, are found in CAZy families GT4, GT35, and GH13. Inverting GPs are classified in families GH65, GH94, GH112, and GH130. Inverting GH-related GPs and hydrolytic enzymes use a similar single displacement mechanism, differing in the requirement of GPs for a single catalytic residue (the proton donor), and inverting GHs for two catalytic residues. In GP-catalyzed reactions, the reaction begins with the direct nucleophilic attack by phosphate to the glycosidic bond with the aid of the catalytic residue, which donates a proton to the glycosidic oxygen atom and then proceeds through an oxocarbenium cation-like transition state. In GH reaction mechanisms, the nucleophilic attack of the C-1 of the glycoside is performed by a water molecule activated by the catalytic base. The natural structural and functional diversity of GPs thus appears to be highly restricted because of the following: (i) they are found in only 7 of the 226 GH and GT families listed in the CAZy database (March 2013); (ii) approximately only 15 EC entries are currently assigned to GPs (24); (iii) their specificity toward glycosyl phosphates is limited to α- and β-D-glucopyranose 1-phosphate (25–29), which are the most prevalent substrates; α-D-galactopyranose 1-phosphate (30); N-acetyl-α-D-glucosamine 1-phosphate (31), and α-D-mannopyranose 1-phosphate (32). In July 2013, α-D-mannopyranose 1-phosphate specificity was described for just one enzyme produced by a human gut bacterium, the Bacteroides fragilis NCTC 9343 mannosylglucose phosphorylase (BfMP).
which converts β-ᴅ-mannopyranosyl-1,4-ᴅ-glucopyranose and phosphate into α-ᴅ-mannopyranose 1-phosphate and 1-ᴅ-glucose (32). This enzyme has been implicated in the catabolism of linear dietary mannans, assisted by a β-1,4-mannanase and a mannobiase 2-epimerase. During review of this study, Nihira et al. (33) reported the discovery of a metabolic pathway for N-glycans, which includes a β-1,4-mannosyl-N-acetyl-1,4-ᴅ-glucosamine phosphorylase named BT1033, produced by the human gut inhabitant B. thetaiotamicron. VPI-5482. BfMP and BT1033 are two of the four enzymes to be characterized in the recently created GH130 family, which includes a total of 447 entries, from archaea, bacteria, and eukaryotes. The other characterized GH130 enzymes are RaMP1 and RaMP2 from the ruminal bacterium Ruminococcus albus NE1 (34). These enzymes have also been proposed to participate in mannan catabolism in the bovine rumen and are assisted by an endomannanase and an epimerase, via RaMP2- and RaMP1-catalyzed phosphorylisis of β-1,4-manno-oligosaccharides and 4-O-β-ᴅ-mannopyranosyl-ᴅ-glucopyranose, respectively (34). X-ray crystallographic studies show that GH130 enzymes share a 5-fold propeller fold. Currently, atomic coordinates data sets for four protein structures are available in the RCSB Protein Data Bank (35) as follows: MSB8 (1VKD). However, no function has yet been attributed to these four proteins, thus limiting the understanding of structure specificity relations for GH130 enzymes and the investigation of their catalytic mechanism.

Recently, the sequence of another GH130 enzyme, which we refer to as UhgbMP (unknown human gut bacterium mannoside phosphorylase, GenBank™ accession number ADD61463.1), was discovered by functional metagenomics of the human gut microbiota (36). The 36.6-kbp metagenomic DNA fragment containing the UhgbMP-encoding gene was taxonomically assigned to an as yet unidentified bacterium belonging to the genus Bacteroides. However, UhgbMP itself presents 99% protein sequence identity with the hypothetical protein BACSTE_03540 from Bacteroides stercoris ATCC 43183 (accession number EDS13361.1).

Here, we present an integrative approach, based on analyses of UhgbMP substrate specificity and of metagenomic and genomic data at the level of the entire human gut ecosystem, to reveal the role of UhgbMP and of 64 other GH130 enzymes produced by known gut bacteria in the breakdown of host and dietary mannose-containing glycans. In addition, we establish the molecular basis of GH130 enzyme catalysis, supported by the experimental results of rational engineering of UhgbMP and the analysis of its three-dimensional molecular model. Finally, we discuss the potential of this enzyme for the development of therapeutic agents and functional foods to inhibit N-glycan degradation, with the ultimate goal of protecting the epithelial barrier in the IBD context.

**EXPERIMENTAL PROCEDURES**

**Recombinant UhgbMP Production and Purification**—First, the UhgbMP encoding gene was PCR-amplified from the Escherichia coli metagenomic clone (GenBank™ accession number GU942931) using primers forward 5'-AGTGATGAGTC-AAGTTTTATTTCCGGG-3' and reverse 5'-TCAGATGATGT-AGTTGTACTGTTTGGTATAATTCC-3', by using the Expand Long Template PCR kit (Roche Applied Science).

To allow heterologous UhgbMP production in E. coli with His, tag at the N-terminal extremity, the PCR product was purified and subsequently cloned into the pCR8/GW/TOPO entry vector (Invitrogen), and then into the pDESt17 destination vector (Invitrogen), according to the manufacturer's recommendations. E. coli BL21-AI cells (Invitrogen) harboring the UhgbMP-encoding plasmid were cultured at 20 °C for 24 h in ZYM-5052 autoinduction medium (37) supplemented with 100 μg/ml ampicillin, inoculated at A₆₀₀nm 0.1. Cells were harvested and resuspended in 20 mM Tris-HCl, pH 7.0, 300 mM NaCl, and lysed by sonication. Soluble lysate was applied to a TALON resin load with cobalt (GE Healthcare) equilibrated in 20 mM Tris-HCl, pH 7.0, 300 mM NaCl. After washing with 8 volumes of the same buffer supplemented with 10 mM imidazole, the protein was eluted in 20 mM Tris-HCl, pH 7.0, 300 mM NaCl, 150 mM imidazole. Finally, the protein sample was desalted on a PD-10 column (GE Healthcare) and eluted in 20 mM Tris-HCl, pH 7.0, 0.1% Tween 80 (v/v). In these conditions, 84% of UhgbMP remained soluble after 8 days at 4 °C, thus allowing further functional characterization. The purity of the purified wild-type UhgbMP and mutants was evaluated higher than 95% by SDS-PAGE using Any kD™ MiniPROTEAN® TGX™ Precast Gel (Bio-Rad) (supplemental Fig. 1). After migration, proteins were stained with the PageBlue Protein Staining Solution (Thermo Scientific) according to the manufacturer's recommendations. Protein concentrations were determined by spectrophotometry using a NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). The NanoDrop® measurement error was 5%. The calculated extinction coefficient of the purified UhgbMP fused to an N-terminal His₆ tag was 76 630 M⁻¹ cm⁻¹.

**UhgbMP Mutagenesis**—Site-directed mutagenesis was performed by using the pDESt17 plasmid harboring the wild-type UhgbMP encoding gene as PCR template and the primers 5'-AGTGGTGTGCAGATTACCCGGC-3' and 5'-TACG-AGCCTGATGCACACCCATTATG-3' for the D304N mutant, 5'-AGTGTTGTTCGGCAGATTATG-3' for the E273Q mutant, and 5'-AGCCTGATGCACACCCATTATG-3' for the E273Q mutant, and finally 5'-AGCCTGATGCACACCCATTATG-3' and 5'-TCCTCGCCATAACCCGAGTACAGTGC-3' for the Y103E mutant. The PCR products, amplified with the Phusion® High Fidelity DNA polymerase (New England Biolabs), were purified and digested by DpnI (New England Biolabs) before E. coli TOP10 transformation. Protein production and purification were identical for wild-type UhgbMP and its variants.

**Enzyme Assays**—All reactions were carried out at 37 °C (wild-type UhgbMP optimal temperature) in 20 mM Tris-HCl, pH 7.0 (wild-type UhgbMP optimal pH). Syntheses of manno-oligosaccharides from α-ᴅ-mannopyranose 1-phosphate were performed with 0.1 mg/ml purified UhgbMP during 24 h at 37 °C in 20 mM Tris-HCl, pH 7.0, with 10 mM of α-ᴅ-mannopy-
Phosphate and diacetylchitobiose (Dextra, UK, reference C8002). The apparent enzyme activity was determined by fitting the initial rates of 

**Reverse phosphorolysis**

mannan (Megazyme, Ireland, reference P-MANC), 1–20 mM D-β-mannopyranosyl-1,4-D-mannose (Megazyme, Ireland, reference O-MBI), 0.1–1 mM pNP-β-D-mannopyranosyl, or 0.05–0.5 mM of β-D-mannopyranosyl-1,4-N,N'-diacetylchitobiose (Dextra, UK, reference MC3020). Reverse phosphorolysis kinetic parameters were determined with 0.01 mg/ml purified enzyme by quantifying α-D-mannose 1-phosphate release from 0.5 to 10 mM inorganic phosphate (10 mM corresponding to the intracellular concentration of inorganic phosphate that was previously measured in bacteria (38)) and 1–10 mM of β-D-mannopyranosyl-1,4-D-glucose (Carbosynth, UK, reference OM04754), 0.4–4 mM β-1,4-D-mannan (Megazyme, Ireland, reference P-MANC), 1–20 mM β-D-mannopyranosyl-1,4-D-mannose (Megazyme, Ireland, reference O-MBI), 0.1–1 mM pNP-β-D-mannopyranosyl, or 0.05–0.5 mM of β-D-mannopyranosyl-1,4-N,N'-diacetylchitobiose (Dextra, UK, reference MC3020). Reverse phosphorolysis kinetic parameters were determined with 0.01 mg/ml purified enzyme by quantifying α-D-mannopyranosyl 1-phosphate consumption from 0.5 to 10 mM α-D-mannopyranosyl 1-phosphate and 5–40 mM of D-mannose, D-glucose, D-galactose, D-fructose, or N-acetyl-D-glucosamine, or 0.1–40 mM of N,N'-diacetylchitobiose (Dextra, UK, reference C8002). The apparent kinetic parameters for phosphorolysis and reverse phosphorolysis at fixed initial concentrations of inorganic phosphate and α-D-mannose 1-phosphate, respectively, were determined by fitting the initial rates of α-D-mannose 1-phosphate release and consumption to the Michaelis-Menten equation. The kinetic parameters for phosphorolysis and synthesis of β-D-mannopyranosyl-1,4,N,N'-diacetylchitobiose and β-D-mannopyranosyl-1,4-D-mannose were determined by varying carbohydrates, α-D-mannose 1-phosphate or inorganic phosphate concentrations, and by fitting the initial rates of α-D-mannose 1-phosphate release and consumption to the sequential random Bi Bi mechanism equation (39). Nonlinear regression was performed with SigmaPlot enzyme kinetics module, version 1.3 (Systat Software, Inc., San Jose, CA). UhgbMP-specific activity toward β-1,4-D-mannooligosaccharides, was determined with 0.1 mg/ml purified enzyme by quantifying α-D-mannopyranosyl 1-phosphate release rate from 10 mM inorganic phosphate and 10 mM β-D-mannooligosaccharides of polymerization degree 2–6 (Megazyme, Ireland). The percentage of UhgbMP inhibition by exogenous carbohydrates or polyols was measured with 0.1 mg/ml purified enzyme by quantifying α-D-mannopyranosyl 1-phosphate consumption rate from 10 mM α-D-mannopyranosyl 1-phosphate as glycosyl donor, and with and without 10 mM of L-rhamnose, D-altrose, D-alllose, D-fucose, D-mannitol, D-sorbitol, D-lyxose, xyitol, L-xylose, D-xylene, L-arabinose, or D-cellobiose.

The percentage of reverse phosphorolysis activity of the D104N, D304N, and E273Q variants, compared with that of the wild-type enzyme, was determined with 0.1 mg/ml of purified proteins by quantifying α-D-mannopyranosyl 1-phosphate consumption rate from 10 mM α-D-mannopyranosyl 1-phosphate as glycosyl donor and 10 mM D-mannose as acceptor. α-D-Mannopyranosyl 1-phosphate was quantified by using high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD). Carbohydrates and α-D-mannopyranosyl 1-phosphate were separated on a 4×250 mm Dionex CarboPak PA100 column. A gradient of sodium acetate (from 0 to 150 mM in 15 min) and an isotropic step of 300 mM sodium acetate in 150 mM NaOH was applied at a 1 ml min⁻¹ flow rate. Detection was performed using a Dionex ED40 module with a gold working electrode and a Ag/AgCl pH reference. Finally, the hydrolytic or phosphorolytic behavior of the wild-type UhgbMP and its Y103E variant was assessed by using 1 mM pNP-β-D-mannopyranosyl, in the absence or presence, respectively, of 10 mM inorganic phosphate. The pNP release was monitored at A₄₀₅ nm on a carry-100 UV-visible spectrophotometer (Agilent Technologies). Between three and five independent experiments were carried out to determine initial activity, kinetic constants, and percentage of inhibition by carbohydrates and polyols of wild-type UhgbMP and its mutants. For all reaction rate measurements, it was checked by HPAEC-PAD that less than 10% of substrate was consumed and that the amount of consumed or released α-D-mannopyranosyl 1-phosphate increased linearly with time.

**NMR Spectroscopy**—Freeze-dried reaction media were exchanged twice with 99.9 atom % D₂O and lyophilized. Deuterium oxide was used as the solvent, and sodium 2,2,3,3-tetra-2-deuterio-3-trimethylsilylpropionate was selected as the internal standard. ¹H and ¹³C NMR spectra were recorded on a Bruker Advance 500-MHz spectrometer using a 5-mm z-grad-ient TBI probe at 298 K, an acquisition frequency of 500.13 MHz, and a spectral width of 8012.82 Hz. Spectra were acquired and processed using TopSpin 3.0 software. The various signals were assigned by comparison with signals obtained from α-D-mannopyranose 1-phosphate, β-D-mannopyranosyl-1,4-D-mannose (Megazyme, Ireland, reference O-MBI), β-D-mannopyranosyl-1,4-D-glucose (Carbosynth, UK, reference OM04754), or β-D-mannopyranosyl-1,4-N,N'-diacetylchitobiose (Dextra, UK, reference MC3020), used as standards.

**Three-dimensional Molecular Modeling**—The UhgbMP sequence was submitted to the I-TASSER server for automated protein structure and function prediction (40). The homologous *T. maritima* TM1225 structure (PDB accession code 1VKD) was used to provide spatial restraints. The three-dimensional model of UhgbMP predicted by I-TASSER was then further refined by energy minimization using the CFF91 force field implementation in the DISCOVER module of the InsightII software suite (Accelrys, San Diego). The CFF91 cross-terms, a harmonic bond potential, and a dielectric constant of 1.0 were specified in the energy function. An initial minimization was performed with positional restraints on the protein backbone using a steepest descent algorithm followed by conjugated gradient minimization until the maximum RMS energy gradient was less than 0.5 kcal mol⁻¹ Å⁻¹. The system was then fully relaxed without positional restraints. α-D-Mannopyranosyl 1-phosphate, β-D-manneheptaose, and β-D-mannopyranosyl-1,4-N,N'-diacetylchitobiose were manually docked into the active site of UhgbMP. The ligand complexes were then optimized according to the minimization protocol described above with the ligand molecules free to move. Molecular graphics images were produced using PyMOL software (Schrödinger, LLC).
**N-Glycan Degradation by Bacterial Glycoside Phosphorylases**

**GH130 Multiple Sequence Alignment Analyses**—The 369 public sequences of GH130 enzymes listed in the CAZY database in January, 2013, were aligned with MUSCLE version 3.7 (41). A distance matrix was generated from the multiple sequence alignment using the BLOSUM62 amino acid residue substitution matrix. The output result file was subjected to hierarchical clustering using Ward’s method (39), and the resulting tree was visualized using DENDROSCOPE 3 (42). Two sequence clusters were clearly apparent, members of which were accordingly assigned to the GH130_1 or GH130_2 subfamilies.

Position-dependent amino acid residue variation in multiple sequence alignment data were analyzed using the Shannon information entropy measure ($H_X$), calculated using SEQUESTER software. The Shannon entropy ($H_X$) at residue alignment position ($X$), corrected for the normalized frequency of residue type occurrence, is computed as shown in Equation 1,

\[
H_X = \frac{\sum_{i=1}^{20} p_i \ln p_i}{\sum_{i=1}^{20} p_i \ln p_i}
\]  

where $p_{i|X}$ is the conditional probability of residue type $i$ occurrence at alignment position $X$, and $p_i$ is the normalized probability of residue type $i$ occurrence at any position. To minimize sampling bias, normalized residue type probability values were taken as those documented by Ranganathan and co-workers (43),³ garnered from sequence data for all natural proteins. Values of $H_X$ lie in the range 0–1; a zero value corresponds to a fully conserved residue position, and a value of unity represents a distribution in which each residue type has an equal chance of occurrence. SEQUESTER appends Shannon entropies at the genomic sequence contained in the recombinant E. coli clone (accession number GU942931), we subcloned the UhgbMP encoding gene to produce a soluble protein tagged with a His$_6$ tag at the N terminus, with a yield of 32 mg of purified protein per liter of culture. We first characterized the substrate specificity of UhgbMP for carbohydrate phosphorylization. In contrast to RaMP1 and BfMP, but similar to RaMP2 and BT1033, UhgbMP exhibits a relaxed specificity toward carbohydrate substrates (Table 1). UhgbMP is able to phosphorylate β-D-mannopyranosyl-1,4-D-glucose, β-D-mannopyranosyl-1,4-D-mannose, and 5 mM D-glucose, D-mannose, D-galactose, D-fructose, or N-acetyl-D-glucosamine, and 5 mM α-D-mannopyranosyl-1-phosphate and 0.1–1 mM N,N-diacytchitobiose.

**TABLE 1**

<table>
<thead>
<tr>
<th>Phosphorylase</th>
<th>$K_{ap}$ (mM)</th>
<th>$K_{up}$ (mM)</th>
<th>$K_{up} / K_{ap}$ (mM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>β-D-mannopyranosyl-1,4-D-mannose</td>
<td>1.1 ± 0.04</td>
<td>1.1 ± 0.01</td>
<td>0.94</td>
</tr>
<tr>
<td>β-D-mannopyranosyl-1,4-D-glucose</td>
<td>47 ± 6</td>
<td>61 ± 6</td>
<td>1.27</td>
</tr>
<tr>
<td>β-D-Mann</td>
<td>2.2 ± 0.4</td>
<td>3.1 ± 0.3</td>
<td>1.41</td>
</tr>
<tr>
<td>β-D-mannopyranosyl-1,6-N-acetylchitobiose</td>
<td>0.018 ± 0.003</td>
<td>0.072 ± 0.002</td>
<td>4.0</td>
</tr>
<tr>
<td>N,N-diacytchitobiose</td>
<td>0.4 ± 0.1</td>
<td>0.009 ± 0.0001</td>
<td>0.02</td>
</tr>
</tbody>
</table>

³ O. Rivoire, and R. Ranganathan, unpublished data.

or GH130_2 sequences, respectively, were counted. Prediction of transmembrane topology and signal peptides was performed using PHOBOS.

Analysis of GH130 Encoding Gene Prevalence in the Human Gut Microbiome—The 369 GH130 sequences referenced in the CAZY database in January, 2013, were searched by TBLASTN analysis ($E$-value = 0, identity ≥ 90%) in the catalog of prevalent genes (as defined by Qin et al. (3)) encoded on contigs of metagenomic sequences obtained from fecal samples of 162 individuals from the MetaHit cohort and of 139 individuals from the National Institutes of Health HMP cohort in the United States.

**RESULTS**

**Uhgb_MP Substrate and Product Specificity**—From the metagenomic sequence contained in the recombinant E. coli clone, we subcloned the UhgbMP encoding gene to produce a soluble protein tagged with a His$_6$ tag at the N terminus, with a yield of 32 mg of purified protein per liter of culture. We first characterized the substrate specificity of UhgbMP for carbohydrate phosphorylization. In contrast to RaMP1 and BfMP, but similar to RaMP2 and BT1033, UhgbMP exhibits a relaxed specificity toward carbohydrate substrates (Table 1). UhgbMP is able to phosphorylate β-D-mannopyranosyl-1,4-D-glucose, β-D-mannopyranosyl-1,4-D-mannose, and 5 mM D-glucose, D-mannose, D-galactose, D-fructose, or N-acetyl-D-glucosamine, and 5 mM α-D-mannopyranosyl-1-phosphate and 0.1–1 mM N,N-diacytchitobiose.
products obtained during phosphorolysis reactions, irrespective of the carbohydrate substrate. UhgbMP is thus an exo-acting enzyme, able to breakdown only the first β-mannosidic linkage at the nonreducing end of oligosaccharides.

α-D-Mannopyranose 1-phosphate was therefore tested as a glycosyl donor for reverse phosphorolysis. First, even without any carbohydrate acceptor, UhgbMP produces mannose and, furthermore, manno-oligosaccharides of DP ranging from 1 to 12, indicating that water itself plays the role of first acceptor at the beginning of the reaction. The β-1,4 regio-specific synthesis of manno-oligosaccharides was characterized by 1H and 13C NMR (supplemental Fig. 2). Various carbohydrates or polyols were tested as acceptors (Table 1). D-GlcNAc and β-D-GlcNAc-1,4-D-GlcNAc were the best recognized acceptors, given that the UhgbMP $K_v$ values for these compounds are 4- and 55-fold lower than for D-mannose, respectively. Starting from α-D-mannopyranose 1-phosphate and D-GlcNAc or β-D-GlcNAc-1,4-D-GlcNAc, UhgbMP generates series of manno-oligosaccharides containing D-GlcNAc or β-D-GlcNAc-1,4-D-GlcNAc at their reducing end, with a DP of up to 4. D-Glucose, D-mannose, D-galactose, and D-fructose are also recognized as acceptors.

The kinetic mechanism of UhgbMP was investigated by determining the initial reaction velocities for β-D-Manp-1,4-D-Manp and β-D-Manp-1,4-β-D-GlcNAc-1,4-D-GlcNAc phosphorolysis and synthesis at various initial concentrations of substrates. First, double-reciprocal plots of 1/v versus 1/[β-D-Manp-1,4-D-Manp] at various concentrations of inorganic phosphate and of 1/v versus 1/[D-Manp] at various concentrations of α-D-mannopyranose 1-phosphate crossed at a certain point, on the left of the 1/v axis (Fig. 3). UhgbMP thus catalyzes the β-D-Manp-1,4-D-Manp phosphorolysis and synthesis through a sequential Bi Bi mechanism involving the formation of a ternary complex. The equilibrium constant was calculated from the Haldane relationship for a sequential random Bi Bi mechanism (39) (the most probable mechanism fitting with this kinetic dataset) as shown in Equation 2,

$$K = \frac{(k_{cat} \times K_{S, \alpha-D-Manp-1-P} \times K_{M, D-Manp})}{(K_{cat} \times K_{S, \beta-D-Manp-1,4-D-Manp} \times K_{M, D-Manp})}$$

(Eq. 2)

$k_{cat}$, $K_S$, and $K_M$ being $k_{cat}$ for phosphorolytic and synthetic reactions, respectively (Table 2). The value of 2.1 indicates that reaction equilibrium favors phosphorolysis. The kinetic parameters calculated for β-D-Manp-1,4-β-D-GlcNAc-1,4-D-GlcNAc phosphorolysis and synthesis are clearly different. β-D-Manp-1,4-β-D-GlcNAc-1,4-D-GlcNAc synthesis is inhibited by β-D-GlcNAc-1,4-D-GlcNAc concentrations higher than 1 mM (Fig. 3). For β-D-GlcNAc-1,4-D-GlcNAc and β-D-Manp-1,4-β-D-GlcNAc-1,4-D-GlcNAc concentrations lower than 1 and 0.5 mM, respectively, the most probable mechanism is also a mixed-type sequential random Bi Bi mechanism. The $K_v$ value calculated in these conditions is 0.32, a value similar in magnitude to the value 0.658 obtained for D-Manp-1,4-β-D-Glc phosphorolysis by RaMP2 (34).

Further investigations of UhgbMP mechanism will be necessary to determine the order of substrate binding and product release and to better understand how this enzyme works on its natural substrates, including β-D-Manp-1,4-D-GlcNAc.

Finally, with all other compounds tested as acceptors, UhgbMP synthesizes the same (Man)$_n$ series as when α-D-mannopyranose 1-phosphate was the sole substrate. This shows that none of the other tested molecules were good acceptors, because water and additional mannose and manno-oligosaccharide units were used preferentially by the enzyme. However, the presence of cellobiose, D-fucose, L-rhamnose, L- and D-xylose, N-acetyl-D-galactosamine (D-GalpNAc), D-altrose, D-allose, xylitol, D-lyxose, or D-mannitol decreased enzyme-specific activity, in some cases quite markedly (Table 3). But concentrations of these compounds did not decrease during reaction, and no significant additional product was produced compared with reaction in the presence of α-D-mannopyranose 1-phosphate as sole substrate. These carbohydrates thus act as UhgbMP inhibitors. Various sugar-phosphates were also tested as glycosyl donors for reverse phosphorolysis as follows: α-D-fructose-1- and -6-phosphate, D-ribose 1-phosphate, α-D-galactosamine 1-phosphate, α-D-glucosamine 1-phosphate, D-mannose 6-phosphate, and α-D-glucopyranosyl 1- and -6-phosphates. None of them was consumed by UhgbMP, and no reaction product appeared on HPAEC-PAD chromatograms after 24 h at 37°C.

Key Residues Involved in Mannoside Phosphorolysis Mechanism—To investigate the UhgbMP catalytic mechanism, we first built a three-dimensional model of the enzyme using the atomic coordinates of the TM1225 protein from T. maritima MSB8 (PDB accession code 1VKD) as a structural template. The enzyme, which adopts a five-bladed β-propeller fold (Fig. 4A), was identified from structural genomics initiatives. Of the
four proteins of known structure that share the same fold, TM1225 aligns with the highest sequence identity to UhgbMP (59%), but it has yet to be functionally characterized.

To identify putative catalytic amino acid residues, we analyzed the multiple alignment of the 369 protein sequences referenced in CAZy family GH130 (January 2013) (supplemental Fig. 3). The most conserved amino acid residue positions were then mapped onto the three-dimensional model of UhgbMP. Of these, Asp-104 and Asp-304 were found to be part of a narrow groove putatively considered to be the active site. Another conserved amino acid residue, Glu-273, present in a short helix turn section, was also identified in the groove.

These three amino acid residues were individually mutated to investigate their potential role in the UhgbMP catalytic activity. FIGURE 3. Double-reciprocal plots for β-d-mannopyranosyl-1,4-N,N′-diacetylchitobiose synthesis (A) and phosphorolysis (B) and for β-d-mannopyranosyl-1,4-α-mannose synthesis (C) and phosphorolysis (D). The initial velocities of α-β-mannose 1-phosphate consumption and release were measured at various concentrations of substrates. A and C, concentrations of α-β-mannose 1-phosphate were 0.5 mM (filled square), 1 mM (filled circle), 2 mM (filled triangle), and 5 mM (open square). B and D, concentrations of inorganic phosphate were 0.5 mM (filled square), 1 mM (filled circle), 2 mM (filled triangle) and 5 mM (open square). Data are the mean values ± S.D. for three to five independent experiments.

TABLE 2
Kinetic parameters for the phosphorolysis and synthesis of β-d-mannopyranosyl-1,4-α-mannose and β-d-mannopyranosyl-1,4-N,N′-diacetylchitobiose

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Phosphorolysis</th>
<th>Synthesis</th>
<th>β-d-Mann-1,4-β-d-Mann</th>
<th>β-d-Mann-1,4-β-d-GlcpNAc-1,4-d-GlcpNAc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kinetic parameters</td>
<td>kcat (s⁻¹)</td>
<td>Km (mM)</td>
<td>Vmax (mM s⁻¹)</td>
<td>Km (mM)</td>
</tr>
<tr>
<td>Phosphorolysis</td>
<td>1.17 ± 0.02</td>
<td>0.8 ± 0.2</td>
<td>1.1 ± 0.2</td>
<td>0.8 ± 0.2</td>
</tr>
<tr>
<td>Synthesis</td>
<td>24 ± 6</td>
<td>4.2 ± 0.2</td>
<td>1.8 ± 0.5</td>
<td>27 ± 1</td>
</tr>
</tbody>
</table>

TABLE 3
Percentage of UhgbMP inhibition by carbohydrates and polyols, quantified by measuring α-β-mannopyranose 1-phosphate consumption rate from 10 mM α-β-mannopyranose 1-phosphate as glycosyl donor, with and without 10 mM of carbohydrates or polyols

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>% of inhibition</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-Arabinose</td>
<td>2 ± 0.4</td>
</tr>
<tr>
<td>D-Cellobiose</td>
<td>37 ± 1</td>
</tr>
<tr>
<td>D-Fucose</td>
<td>12 ± 2</td>
</tr>
<tr>
<td>L-Fucose</td>
<td>11 ± 4</td>
</tr>
<tr>
<td>L-Rhamnose</td>
<td>31 ± 2</td>
</tr>
<tr>
<td>Xyitol</td>
<td>29.4 ± 0.4</td>
</tr>
<tr>
<td>D-Lyxose</td>
<td>19 ± 2</td>
</tr>
<tr>
<td>L-Xylose</td>
<td>58 ± 1</td>
</tr>
<tr>
<td>Mannitol</td>
<td>37 ± 4</td>
</tr>
<tr>
<td>D-Altrrose</td>
<td>84 ± 3</td>
</tr>
<tr>
<td>D-Xylose</td>
<td>86.2 ± 0.2</td>
</tr>
<tr>
<td>D-Allose</td>
<td>94 ± 3</td>
</tr>
</tbody>
</table>
mechanism. Results showed that mutation of Asp-104 into an asparagine completely abolished UhgbMP activity, although D304N and E273Q mutations retained only 3.9 ± 0.5 and 0.2 ± 0.1% of the native catalytic activity, respectively. These three amino acids were thus considered as candidate catalytic residues. To further elucidate the functional roles of these residues, molecular modeling techniques were used to dock α-D-mannopyranose 1-phosphate into the putative active site groove.

Of the identified docking modes, only one appeared compatible with spatial constraints provided by the UhgbMP reaction mechanism (Fig. 4A) as follows: (i) provision for the specific recognition of mannose at the −1 subsite; (ii) substrate interaction with at least one of the (Asp-104, Asp-304, or Glu-273) putative catalytic acidic residue side chains; (iii) presence of a favorable phosphate-binding site allowing reverse phosphorolysis to take place; (iv) presence of binding subsites able to accommodate β-D-Manp-1,4-β-D-GlcNAc-1,4-β-D-GlcNAc and β-1,4-D-mannan chains in catalytically productive binding modes with respect to active site residue(s) implicated in the chemical reaction mechanism (Fig. 4, B and C).

In this binding mode, where the α-D-mannopyranosyl 1-phosphate ring structure is stabilized in the −1 subsite through stacking interactions with Tyr-103, the Asp-104 residue located on the β-face of the catalytic chiral center can act as the unique proton donor during catalysis, whereas Arg-150, Arg-168, and Asn-151 amino acid residues can favorably assist phosphate group positioning consistent with inversion of configuration at C-1 of mannose (Fig. 4A). Given the high conservation of these amino acid residues within the GH130 family alignment, we suggest that enzymes contained in this family should share the same single displacement mechanism described for GH-like inverting phosphorylases. In such a mechanism, the aspartic acid corresponding to Asp-104 in UhgbMP would be the sole catalytic residue, having the role of proton donor (Fig. 5). The phosphate group, stabilized through ionic interactions with highly conserved Arg-150, Arg-168, and Asn-151 residues in UhgbMP, would then act as the nucleophile (Fig. 4A).

**Molecular Basis of Substrate Specificity of GH130 Enzymes**

The biochemical data from the kinetic characterization of BfMP, RaMP1, RaMP2, BT1033, and UhgbMP revealed marked specificity differences among the five enzymes. The BfMP and RaMP1 + 1 subsite is highly specific for glucose, whereas RaMP2, BT1033, and UhgbMP display looser specificity both toward carbohydrate substrates for phosphorolysis and acceptors for synthetic reactions. We therefore investigated whether the five characterized proteins could represent different GH130 subfamilies. Indeed, in most of the cases, enzymes classified in a CAZy subfamily share the same substrate and/or product specificity, reflecting a high degree of conservation in their active site (46, 47). A phylogenetic tree was constructed, based on the multiple alignments of the GH130 enzyme sequences. Three clusters of sequences clearly appeared. BfMP and RaMP1 are contained in subfamily GH130_1 (79 sequences, Fig. 6 and supplemental Table 1). UhgbMP, RaMP2, and BT1033 are contained in subfamily GH130_2 (42 sequences), together with
TM1225 (PDB accession code 1VKD). The GH130_NC cluster contains 248 sequences of as yet uncharacterized proteins that are too heterogeneous to permit the creation of univocal subfamilies. This cluster contains the other three proteins of known structure, namely the BDI_3141 protein from P. distasonis ATCC 8503, the BT_4094 protein from B. thetaiotaomicron VPI-5428, and the BACOVA_03624 protein from B. ovatus ATCC 8483. However, a key residue position occupied by Tyr-103 in UhgbMP allows the discrimination of the GH130_NC sequences from those of the two other subfamilies. The tyrosine residue at position 103 in UhgbMP, which lies close to the phosphate group binding site in the three-dimensional model of the docked complex of the enzyme with α-D-mannopyranosyl-1,4-N,N'-diacetylchitobiose phosphorolysis by UhgbMP, with Asp-104 acting as the proton donor.
are hydrolases and not phosphorylases. The Glu residue corresponding to Tyr-103 in UhgbMP is expected to act as the second catalytic residue, taking on the role of base. To validate this hypothesis, we attempted to transform UhgbMP, which possesses low intrinsic hydrolase activity when assayed in the absence of inorganic phosphate, into a hydrolase, through the replacement of Tyr-Y103 by a glutamic acid. In the presence of $10 \text{mM Pi}$, the hydrolytic activity of the wild-type enzyme was $6.3 \times 10^{-3} \pm 3 \times 10^{-4} \text{mol min}^{-1} \text{mg}^{-1}$, a value similar to that obtained for mutant Y103E ($5.7 \times 10^{-3} \pm 1 \times 10^{-3} \text{mol min}^{-1} \text{mg}^{-1}$). In contrast, without any phosphate, the hydrolytic activity of the wild-type enzyme was $1.4 \times 10^{-3} \pm 1 \times 10^{-4} \text{mol min}^{-1} \text{mg}^{-1}$, although it was $3.8 \times 10^{-3} \pm 3 \times 10^{-4} \text{mol min}^{-1} \text{mg}^{-1}$ for Y103E. Nevertheless, the Y103E mutant pNP release curve versus time reached a plateau after only 10 min reaction at 37 °C with and without phosphate, although less than 5% of substrate was consumed (supplemental Fig. 4). This phenomenon was also observed for the wild-type enzyme without phosphate. This indicates that inorganic phosphate is probably involved in maintaining the UhgbMP active site conformation and that the Y103E mutation may alter this conformation. Because the product yields are so weak without phosphate, it remains difficult to conclude that the Y103E mutant is really a hydrolase. Functional and structural investigations will thus be needed to confirm that the GH130_NC cluster contains mannoside hydrolases.

Our attempts to check correct folding of the Y103E variant, as well as that of the wild-type enzyme and of variants D104N, D304N, and E273Q by circular dichroism using a J-815 UV Spectro orientation (Jasco) failed, because of the high absorbance of Tween 80 at wavelengths between 200 and 290 nm (48).

To study the active site conservation in family GH130, we projected the results of the multiple alignments of the GH130 protein sequences onto the UhgbMP three-dimensional model, in which $\beta$-d-Manp-$1,4$-$\beta$-d-GlcNAc-$1,4$-$\beta$-d-GlcNAc or $\beta$-1,4-linked d-mannoheptaose (Fig. 7) was docked to enable the mapping of putative binding carbohydrate subsites. Clearly, residues lining the UhgbMP catalytic furrow are highly conserved only among the GH130_2 sequences. The +1 putative subsite appears delimited by Tyr-103, Asp-304, His-174, Tyr-240, and Phe-283 residues that are specifically conserved in the GH130_2 family (Fig. 4B and supplemental Fig. 5). In the presence of $\beta$-d-Manp-$1,4$-$\beta$-d-GlcNAc-$1,4$-$\beta$-d-GlcNAc, Tyr-103 and Asp-304 residues can form H-bond interactions with the N-acetyl-d-glucosamine positioned in the +1 subsite. In the presence of $\beta$-1,4-linked d-mannoheptaose, the Tyr-103 can interact with the mannosyl residue located in the +1 subsite, whereas Asp-304 establishes a hydrogen bond with the mannosyl moiety at +2 subsite. Asp-304 thus probably plays a key role in substrate binding at +1 and +2 subsites, explaining why its mutation dramatically alters UhgbMP activity.

The Tyr-242 and Phe-283 residues provide additional stacking platforms at +1 subsite to stabilize the bound glycosyl unit. At the putative +2 subsite, the N-acetyl-d-glucosamine residue at the reducing end of the $\beta$-d-Manp-$1,4$-$\beta$-d-GlcNAc-$1,4$-$\beta$-d-GlcNAc is found stabilized through hydrogen bonding interactions with Tyr-242, Pro-279, Asn-280, and Asp-304, while establishing van der Waals interactions with Val-271. Interestingly, in our three-dimensional model of the active site, the N-acetyl groups of the GlcNAc moieties docked at +1 and +2 subsites are found to fit nicely into a pocket formed by the loop Pro-271–Pro-284, which is locked in a “closed conformation” via a salt bridge ionic interaction between Asp-277 and Arg-65 residues (Fig. 4B and supplemental Fig. 5). The conformation of this loop may be modified by the E273Q mutation, thus altering oligosaccharide accommodation in the active site. Sequence analysis indicates that such a loop is very well conserved within the subfamily GH130_2 but less within GH130_1 (supplemental Fig. 3). This could suggest an involvement of this loop in the specificity determination of GH130_2 enzymes toward manno-oligosaccharides longer than DP2.

The +3 to +6 putative subsites have been mapped using the docked mannoheptasaccharide (Figs. 4C and 7 and supplemental Fig. 5). Surprisingly, very few aromatic residues able to provide stacking interactions with the manno-oligosaccharide chain have been found. However, a dense network of likely hydrogen bonding interactions and van der Waals contacts could explain the increase of UhgbMP catalytic efficiency with the polymerization degree of manno-oligosaccharides. In the +3 putative subsite, the mannosyl residue can be stabilized by
interactions with Tyr-240, Asn-238, and Val-241 residues, although at the putative +4 subsite, the sugar moiety is rather engaged in interactions with Asn-280 and Asn-238 residues. Residues Tyr-264, Trp-208, and Arg-269 are seen to establish interactions with the mannosyl moiety located at subsite +5. The length of the funnel binding site is found to accommodate up to +5 subsites. Outside this funnel, the +6 subsite has been putatively defined as sitting on the aromatic Trp-208 residue, which is almost perfectly conserved among the GH130_2 subfamily and not in GH130_1.

Finally, two important differences between GH130_1 and GH130_2 subfamilies, identified by sequence alignment analysis, could explain why GH130_1 enzymes seem to exhibit a very narrow specificity toward β-D-Manp-1,4-D-GlcP, as observed for RaMP1 and BfMP, although GH130_2 enzymes would be able to act on longer manno-oligosaccharides. First, the UhgbMP loop Gly-121–Gly-125, which defines the extremity of the −1 subsite, is very well conserved within the GH130_2 subfamily, whereas an insertion of 12 residues is observed in all GH130_1 sequences (supplemental Fig. 3). This longer loop could thus prevent the accommodation of long oligosaccharides in the negative subsite of GH130_1 enzymes. Second, we observed that the His-174 residue, which is not conserved in the GH130_1 subfamily and is contained in the UhgbMP Pro-169–Asp-179 loop, interacts with the sugar moiety bound at the +1 subsite. Interestingly, a motif containing five glycine residues was rather found within the GH130_1 subfamily in place of His-174 observed in GH130_2 enzymes that could play a major role in substrate accommodation within the active site (supplemental Fig. 3).

**Genomic Context of GH130 Encoding Genes, Focus on Human Gut Bacteria—** Differences in substrate specificity of enzymes belonging to the GH130_1 and GH130_2 subfamilies may illustrate different roles in the microbial ecosystem in which they are produced. 23 of the 79 GH130_1 sequences are mainly surrounded by GH18 and GH92 and, to a lesser extent, GH97 sequences (Fig. 6 and supplemental Fig. 6). GH18 and GH92 contain the endo-β-N-acetylglucosaminidases and α1,2-, α1,3-, or α1,6-mannosidases that are implicated in the hydrolysis of human N-glycans (16, 17, 19, 20), although the only three characterized GH97 enzymes present α-glucosidase (EC 3.2.1.20) or α-galactosidase (EC 3.2.1.22) activity. Finally, only two multigenic clusters were found to encode both a GH130_2 enzyme and a GH2 enzyme, the GH2 family containing β-mannosidases capable of hydrolyzing the disaccharide β-D-Manp-1,4-D-GlcNAc (17).

The multigenic cluster containing UhgbMP encoding gene and that from *B. stercoreis* ATCC 43183 containing its homolog BACSTE_03540 (accession number EDS13361.1) are clear examples of this gene organization with several GH92, GH18, and GH97 sequences surrounding a GH130_2 encoding gene (Fig. 1 and supplemental Fig. 6). In these cases, the results of *in silico* detection of signal peptide or transmembrane topology indicate that UhgbMP and BACSTE_03540 are probably intracellular, although the glycoside-hydrolases encoded by the same multigenic clusters would rather be secreted.

**DISCUSSION**

The biochemical characterization of UhgbMP, a representative of the GH130 enzyme family, revealed its flexibility toward carbohydrate substrates. UhgbMP is capable of catalyzing the phosphorylation of β-D-mannopyranosyl-1,4-β-glucopyranosyl, β-1,4-d-manno-oligosaccharides of DP >5 and mannans, as well as the N-glycan core oligosaccharide β-D-Manp-1,4-β-1,4-D-GlcNAc-1,4-β-D-GlcPNAc. Based on a sequence analysis of the GH130 family and on a structural model of UhgbMP, supported by the experimental findings presented here and else-
where by Senoura et al. (32), Nihira et al. (33), and Kawahara et al. (34), we propose the creation of two GH130 subfamilies, of which the members probably share the same single displacement mechanism involving a single catalytic acidic residue (corresponding to UhgbMP D104) acting as the proton donor. Subfamily GH130_1 gathers together enzymes (including BfMP and RaMP1) exhibiting narrow specificity toward \( \beta-D-Manp-1,4-D-Glc \). Conversely, enzymes of the GH130_2 subfamily show high promiscuity toward their substrates and products. Their active sites would be sufficiently extensible to accommodate complex carbohydrate structures such as \( \beta-D-mannan \) and \( \beta-D-manno-oligosaccharides \), with or without \( \beta-D-Glc:pNAc-1,4-D-GlcNAc, D-Glc:pNAc, \) or \( D-glucose \) at their reducing end. The resolution of crystallographic structures of representatives of each of GH130 clusters in complex with substrates and products will, however, be necessary to confirm the role of the key residues identified here and to deepen our understanding of the reaction mechanisms operating in these enzymes.

Among the 369 enzymes archived in the GH130 family (January, 2013), 65 belong to human gut bacteria. The strong prevalence of the corresponding genes, in particular that coding for UhgbMP, in the human gut metagenome suggests that these enzymes have a major role for foraging mannose in this ecosystem. Here, we have demonstrated that UhgbMP, and probably also the 16 other enzymes in the proposed GH130_2 subfamily produced by gut bacteria (Supplemental Table 1), including BT1033, are able to participate in breaking down the host mannosylated glycoproteins lining the intestinal epithelium.

In vivo, incorporation of GH reaction products into metabolic pathways is energy-consuming, although direct metabolism of the glycoside phosphates synthesized by GPs does not require ATP. Phosphorylation reactions mediated by GPs would thus be more advantageous than hydrolysis under certain physiological conditions, for example during intensive use of carbohydrate resources or in anoxic environments such as the gastrointestinal tract, where ATP cannot be efficiently produced by the respiratory chain-linked phosphorylation process (51).

In common with other GPs, especially those previously identified in microorganisms having a facultatively anaerobic lifestyle like *Bifidobacterium* sp., *Lactobacillus* sp., or *Clostridium* sp (50), enzymes in the proposed GH130_1 and GH130_2 families can therefore be considered as catabolic enzymes, working in tandem with catabolic GHS in dietary and host mannoside breakdown.

Based on the analysis of the genomic context of GH130_2 encoding genes in gut bacteria, and accordingly with Nihira et al. (33), we propose that mannoside phosphorylation catalyzed by GH130_2 enzymes acts in concert with GH activities of enzymes in the GH18 and GH92 families, to completely breakdown \( N \)-glycans, as proposed in Fig. 1. The role of the GH97s whose genes belong to the same multigenic systems as those coding for the proposed GH130_2 enzymes (for example, in *B. stercoris* ATCC 43183) is less clear. Family GH97 contains only three characterized enzymes, of which two are \( \alpha \)-glucosidases with \( \alpha-1,4 \)-link specificities (52, 53). However, the functional diversity of GH97 enzymes has not yet been thoroughly explored, and if these enzymes were previously thought to contribute to dietary carbohydrates in the human gut, they may well also be active in the breakdown of \( N \)-glycans, for example by liberating, like the GH99 enzymes (18), the \( \alpha-1,3 \)-linked glucosyl residue at the nonreductive end of the immature \( N \)-glycans. In the particular case of the unknown gut bacterium producing UhgbMP and of *B. stercoris* ATCC 43183 producing its homolog BACSTE_03540, and based on the results of in silico detection of signal peptides, we assume that the UhgbMP and BACSTE_03540 physiological role would be the intracellular phosphorylation of short oligosaccharides that can be internalized in the cell (di- or trisaccharides (54)), like the \( \beta-D-Manp-1,4-D-Glc:pNAc \) disaccharide or the \( \beta-D-Manp-1,4-\beta-D-Glc:pNAc-1,4-D-Glc:pNAc \) trisaccharide resulting from extracellular hydrolysis of \( N \)-glycans by glycoside hydrolases belonging to the GH18, GH92, and maybe also GH97 families. This deglycosylation arsenal thus allows one to deprotect the glycoproteins of the intestinal epithelium. Their protein part could also thus be degraded by the endopeptidases, similarly to those putatively encoded by the genes ADD61465.1 and ADD61469.1, which belong to the same metagenomic DNA fragment than the UhgbMP encoding gene, so as to perforate the intestinal epithelium (Fig. 1). However, further studies will be needed to confirm the physiological role of these enzymes, like metabolic and transcriptomic analyses of gut bacteria producing GH130_2 enzymes (like *B. stercoris* ATCC 43183) in the presence of \( N \)-glycans as carbon source.

The results presented here show that in vitro, UhgbMP is also able to effectively degrade manno-oligosaccharides and mannan itself. This is also probably the case for other gut bacterial enzymes of the proposed GH130_2 subfamily, which show a similar active site topology, allowing the accommodation of long oligosaccharides. The intestinal bacteria that produce GH130_2 enzymes would thus be able to use dietary hemicelluloses and their hydrolysis products as carbon sources if needed, which would give them a competitive advantage with respect to the other bacteria, to colonize and maintain themselves in the intestinal tract. The use of manno-achs carbohydrates (plant or yeast mannans, linear \( \beta-1,4 \)-linked manno-oligosaccharides, or \( \beta-D-Manp-1,4-\beta-D-Glc:pNAc-1,4-D-Glc:pNAc \)) as functional foods could thus allow the metabolism of the \( N \)-glycan-degrading bacteria to be bypassed and, eventually, may reduce the degradation of the epithelial barrier for therapeutic applications, especially for patients suffering from IBD. Indeed, in their gut metagenome, the known GH130_2 and GH130_NC encoding genes seem highly prevalent, even if enzyme assays on fecal samples, as well as metatranscriptomic and metaproteomic studies, should be performed as evidence that GH130 could be biomarkers of IBDs. However, this approach presents a risk of overfeeding this specific type of bacteria and increasing their prevalence in the gut, and thus the potential for \( N \)-glycan breakdown, in the case of reduced intake of exogenous manno-achs carbohydrates. The utilization of GH130 enzyme inhibitors, like \( \Delta \)-altrose, \( \Delta \)-xylose, and \( \Delta \)-allose for UhgbMP or suicide inhibitors that could be specifically designed thanks to the available or future structural data, would thus be the preferred strategy in therapeutic contexts.

Finally, we have shown here that the relaxed specificity displayed by UhgbMP toward acceptor substrates can be used for the stereo- and region-selective synthesis of original glyco-con-
jugates and oligosaccharides. In particular, this enzyme is able to synthesize β-1,4-linked α-mannosidases, which have been reported to present prebiotic properties (55), even if their use as functional food may be risky, from our point of view. UhgbMP is also highly effective for production of N-glycan core oligosaccharides, such as β-1,4-Manp-1,4-α-GlcNAc and β-1,4-Manp-1,4-β-D-GlcNAc-1,4-β-D-GlcNAc, whose commercial price today exceeds $10,000 per mg. UhgbMP-based β-mannoside synthesis processes appear as highly attractive compared with those based on mannosyltransferases, which use expensive activated sugar nucleotides as donors (56, 57), or on transmannosylation catalyzed by native or engineered mannosidases or mannanases (44, 58–62). Indeed, a two-step UhgbMP-based nosylation catalyzed by native or engineered mannosidases or activated sugar nucleotides as donors (56, 57), or on transmammalian glycans toward different classes of acceptors to synthesize mannosylated products. UhgbMP is thus an enzymatic tool with high potential for synthesizing molecules like N-glycan core oligosaccharides, of major importance for studying, and potentially also for controlling interactions between host and gut microbes.

Acknowledgments—We cordially thank Nelly Monties for technical assistance. MetaToul (Metabolomics & Fluxomics Facilities, Toulouse, France) is grateful acknowledged for access to NMR facilities. MetaToul is supported by grants from the Région Midi-Pyrénées, the European Regional Development Fund, the SICOVAL, the Infrastructures en Biologie Sante et Agronomie (IBiSa, France), the CNRS, and the Institut National de la Recherche Agronomique.

REFERENCES
21. Deleted in proof
22. Deleted in proof
Role of Glycoside Phosphorylases in Mannose Foraging by Human Gut Bacteria
Simon Ladevèze, Laurence Tarquis, Davide A. Cecchini, Juliette Bercovici, Isabelle André, Christopher M. Topham, Sandrine Morel, Elisabeth Laville, Pierre Monsan, Vincent Lombard, Bernard Henrissat and Gabrielle Potocki-Véronèse

J. Biol. Chem. 2013, 288:32370-32383.
doi: 10.1074/jbc.M113.483628 originally published online September 16, 2013

Access the most updated version of this article at doi: 10.1074/jbc.M113.483628

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2013/09/16/M113.483628.DC1

This article cites 57 references, 15 of which can be accessed free at http://www.jbc.org/content/288/45/32370.full.html#ref-list-1