1072 *Biochemistry* **2010,** *49,* 1072–1081
DOI: 10.1021/bi902018y

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Divergence of Biochemical Function in the HAD Superfamily: D-*glycero*-D-*manno*-Heptose-1,7-bisphosphate Phosphatase (GmhB)†

Liangbing Wang,‡,⊥ Hua Huang,‡ Henry H. Nguyen,§ Karen N. Allen,*,§,∥ Patrick S. Mariano,*,‡ and Debra Dunaway-Mariano*,‡

‡*Department of Chemistry and Chemical Biology, University of New Mexico, Albuquerque, New Mexico 87131,* §*Department of Physiology and Biophysics, Boston University School of Medicine, Boston, Massachusetts 02118-2394, and* ∥*Department of Chemistry, Boston University, Boston, Massachusetts 02215.* ⊥*Current address: Chengdu Di'Ao Pharmaceutical Co., Institute of Medicine, Sichuan, China.*

Received November 24, 2009; Revised Manuscript Received December 31, 2009

ABSTRACT: D-*glycero*-D-*manno*-Heptose-1,7-bisphosphate phosphatase (GmhB) is a member of the histidinol-phosphate phosphatase (HisB) subfamily of the haloalkanoic acid dehalogenase (HAD) enzyme superfamily. GmhB supports two divergent biochemical pathways in bacteria: the D-*glycero*-D-*manno*-heptose-1α-GDP pathway (in S-layer glycoprotein biosynthesis) and the L-*glycero*-D-*manno*-heptose-1*β*-ADP pathway (in lipid A biosynthesis). Herein, we report the comparative analysis of substrate recognition in selected GmhB orthologs. The substrate specificity of the L-*glycero*-D-*manno*-heptose-1*β*-ADP pathway GmhB from *Escherichia coli* K-12 was evaluated using hexose and heptose bisphosphates, histidinol phosphate, and common organophosphate metabolites. Only D-*glycero*-D-*manno*-heptose 1*β*,7-bisphosphate ($k_{cat}/K_m = 7 \times 10^6$ M^{-1} s^{-1}) and D-*glycero*-D-*manno*-heptose 1α,7-bisphosphate ($k_{cat}/K_m = 7 \times 10^4$ M^{-1} s^{-1}) displayed physiologically significant substrate activity. ^{31}P NMR analysis demonstrated that *E. coli* GmhB selectively removes the C(7) phosphate. Steady-state kinetic inhibition studies showed that D-*glycero*-D-*manno*-heptose 1*β*-phosphate ($K_{is} = 60\ \mu M$, and $K_{ii} = 150\ \mu M$) and histidinol phosphate ($K_{is} = 1$ mM, and $K_{ii} = 6$ mM), while not hydrolyzed, do in fact bind to *E. coli* GmhB, which leads to the conclusion that nonproductive binding contributes to substrate discrimination. High catalytic efficiency and a narrow substrate range are characteristic of a well-evolved metabolic enzyme, and as such, *E. coli* GmhB is set apart from most HAD phosphatases (which are typically inefficient and promiscuous). The specialization of the biochemical function of GmhB was examined by measuring the kinetic constants for hydrolysis of the α- and *β*-anomers of D-*glycero*-D-*manno*-heptose 1*β*,7-bisphosphate catalyzed by the GmhB orthologs of the L-*glycero*-D-*manno*-heptose 1*β*-ADP pathways operative in *Bordetella bronchiseptica* and *Mesorhizobium loti* and by the GmhB of the D-*glycero*-D-*manno*-heptose 1α-GDP pathway operative in *Bacteroides thetaiotaomicron*. The results show that although each of these representatives possesses physiologically significant catalytic activity toward both anomers, each displays substantial anomeric specificity. Like *E. coli* GmhB, *B. bronchiseptica* GmhB and *M. loti* GmhB prefer the *β*-anomer, whereas *B. thetaiotaomicron* GmhB is selective for the α-anomer. By determining the anomeric configuration of the physiological substrate (D-*glycero*-D-*manno*-heptose 1,7-bisphosphate) for each of the four GmhB orthologs, we discovered that the anomeric specificity of GmhB correlates with that of the pathway kinase. The conclusion drawn from this finding is that the evolution of the ancestor to GmhB in the HisB subfamily provided for specialization toward two distinct biochemical functions.

D-*glycero*-D-*manno*-Heptose-1,7-bisphosphate phosphatase (GmhB)[1] is a member of a subfamily of the haloalkanoic acid dehalogenase (HAD) enzyme superfamily (*2*) named after the enzyme commonly known as "HisB"[2] (*3*). HisB is a bifunctional enzyme whose N-terminal domain catalyzes the dephosphorylation of histidinol phosphate and whose C-terminal domain catalyzes the dehydration of imidazole glycerol phosphate. These two reactions constitute the sixth and eighth steps of histidine biosynthesis, respectively. HisB has a limited biological range (viz. the *γ*-subdivision of proteobacteria wherein it coexists with GmhB), whereas GmhB is found in a wide variety of archea and bacteria (*1*). Fani et al. (*1*) proposed that the HisB subfamily ancestor functioned as a promiscuous phosphatase that ultimately specialized to form D-*glycero*-D-*manno*-heptose-1,7-bisphosphate

†This work was supported by National Institutes of Health Grant GM61099 to K.N.A. and D.D.-M.

*To whom correspondence should be addressed. D.D.-M.: e-mail, dd39@unm.edu; phone, (505) 277-3383; fax, (505) 277-2609. P.S.M.: e-mail, mariano@unm.edu; phone, (505) 277-6390; fax, (505) 277-2609. K.N.A.: e-mail, drkallen@bu.edu; phone, (617) 358-5544; fax, (617) 358-5554.

[1]Abbreviations: GmhB, D-*glycero*-D-*manno*-heptose-1,7-bisphosphate phosphatase; HAD, haloalkanoic acid dehalogenase; HisB, bifunctional histidinol-phosphate phosphatase/imidazole-glycerol-phosphate dehydratase; IPTG, isopropyl α-D-thiogalactopyranoside; ES-MS, electrospray mass spectrometry; NMR, nuclear magnetic resonance; KDOP-8-P, 2-keto-3-deoxyoctulosonate 8-phosphate.

[2]Recently, the name HisNB was recommended for use in place of HisB (*1*); however, the HisB name is used herein.

 Published on Web 01/05/2010

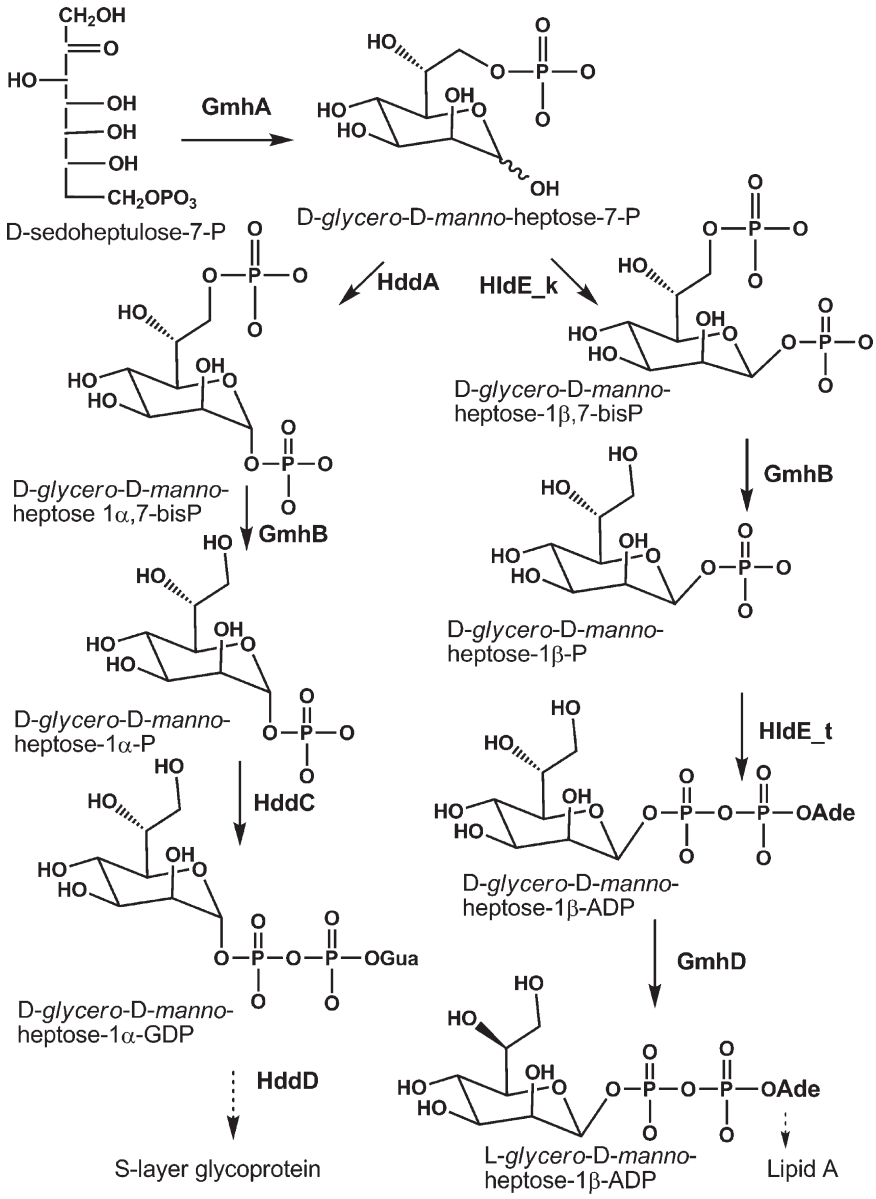

FIGURE 1: Chemical steps and enzyme catalysts of the *A. thermoaerophilus* D-*glycero*-D-*manno*-heptose 1α-GDP pathway (*4*) and the *E. coli* L-*glycero*-D-*manno*-heptose 1*β*-ADP pathway (*5*) .

phosphatase. Duplication of the ancestral gene in *γ*-proteobacteria allowed one copy to evolve to histidinol-phosphate phosphatase and eventually fuse with imidazole-glycerol-phosphate dehydratase to form HisB. In the study described herein, we examined the divergence of biochemical function in GmhB that has accompanied its colonization of bacteria.

The GmhB substrate, D-*glycero*-D-*manno*-heptose 1,7-bisphosphate, is formed as an intermediate in the D-*glycero*-D-*manno*-heptose 1α-GDP (*4*) and L-*glycero*-D-*manno*-heptose 1*β*-ADP pathways (*5*) illustrated in Figure 1. The pathways provide activated *glycero-manno*-heptose units for incorporation into membrane surface glycoproteins and glycolipids (*4*−*12*). The two pathways begin with the isomerization of D-sedoheptulose 7-phosphate to D-*glycero*-D-*manno*-heptose 7-phosphate catalyzed by the heptose isomerase GmhA. D-*glycero*-D-*manno*-Heptose 7-phosphate exists as a mixture of C(1) α- and *β*-anomers. In *Aneurinibacillus thermoaerophilus*, the kinase HddA selects the α-anomer for phosphorylation with ATP, which generates D-*glycero*-D-*manno*-heptose 1α,7-bisphosphate (*4*). In *Escherichia coli*, the N-terminal "kinase" domain of a bifunctional kinase/nucleotidyl transferase (HldE_k/HldE_t) selects the *β*-anomer for phosphorylation with ATP to produce D-*glycero*-D-*manno*-heptose 1*β*,7-bisphosphate (*5*, *13*). The C(7) phosphate group is subsequently removed by GmhB to form the corresponding D-*glycero*-D-*manno*-heptose 1-phosphate, which is the substrate for the nucleotidyl transferase. In *A. thermoaerophilus*, the nucleotidyl transferase catalyzes the transfer of GMP from GTP to the C(1) phosphate of the D-*glycero*-D-*manno*-heptose 1α-phosphate, whereas in *E. coli*, the C-terminal domain (HldE_t) of the bifunctional kinase/nucleotidyl transferase catalyzes the transfer of AMP from ATP to the C(1) phosphate of the D-*glycero*-D-*manno*-heptose 1*β*-phosphate. Finally, in *A. thermoaerophilus*, a glycosyl transferase delivers the D-*glycero*-D-*manno*-heptose unit of the D-*glycero*-D-*manno*-heptose 1α-GDP to the carbohydrate core of the S-layer glycoprotein (*7*), whereas in *E. coli*, the D-*glycero*-D-*manno*-heptose 1*β*-ADP is first converted to the L-*glycero* epimer by the epimerase GmhD, before glycosyl transfer to the lipid A core (*14*).

Herein, we describe a comparative analysis of substrate recognition in four GmhB orthologs that function in one pathway or

the other. Our results show that *E. coli* GmhB displays high catalytic efficiency toward its physiological substrate (D-*glycero*-D-*manno*-heptose 1,7-bisphosphate) and a narrow substrate range, characteristic of a well-evolved metabolic enzyme. Furthermore, we demonstrate that an anomeric preference exists in each of the GmhB orthologs and show a correlation with the anomeric configuration of the physiological substrate. The results of this effort establish the foundation for a companion paper (*15*) in which we report and analyze the liganded structure of the GmhB from *E. coli* K-12 and the liganded structure of GmhB from *Bordetella bronchiseptica*. Together, these two studies define the structural elements of substrate recognition that underlie the divergence of GmhB function within the HAD HisB subfamily.

MATERIALS AND METHODS

Commercial Materials, Services, Spectral Determinations, Buffers, and General Procedures. Except where indicated, all chemicals were obtained from Sigma-Aldrich. Primers, T4 DNA ligase, and restriction enzymes were from Invitrogen. *Pfu* polymerase and the pET23 vector kit were from Stratagene. DEAE-Sepharose was from Amersham Biosciences. Host cells were purchased from Novagen. Genomic DNA (10798, 29148D, BAA-588D) was purchased from ATCC. The nucleotide sequence of each cloned gene was determined by the Center for Genetics in Medicine, University of New Mexico. Electrospray mass spectrometry (ES-MS) determinations were conducted by the University of New Mexico Mass Spectral Facility. ^{1}H NMR, ^{13}C NMR, and ^{31}P NMR (proton-decoupled) spectra were recorded on a Bruker Avance 500 NMR spectrometer at room temperature using D_2O or H_2O as the solvent. ^{1}H NMR and ^{13}C NMR data are reported as follows: chemical shift (parts per million) and multiplicity (s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet). Protein concentration determinations were conducted with the Bradford assay. SDS−PAGE chromatography was used to monitor protein purification. Plasmid DNA was purified using the Qiaprep Spin Miniprep Kit from Qiagen, and transformations were conducted with *E. coli* BL21(DE3) competent cells from Stratagene. The protein sample concentration was determined using a 10K Amicon Ultra Centrifugal filter (Millipore). All protein purifications were conducted at 4 °C using buffer A which consists of 50 mM K^+HEPES (pH 7.0), 5 mM $MgCl_2$, and 1 mM DTT.

Gene Cloning and Cell Lysate Preparation. The gene was amplified from genomic DNA by PCR using *Pfu* DNA polymerase and commercial oligonucleotide primers containing *Nde*I and *Bam*HI cut sites. The linearized pET-23b vector was ligated to the digested PCR product for cell transformation. The purified plasmid DNA was used to transform cells which were grown at 25 °C with agitation at 200 rpm in 10 L of Luria broth (containing 50 μg/mL ampicillin) to an OD_{600} of 0.6−1.0 and then induced for 4 h at 20 °C with 0.4 mM IPTG. The cells were harvested by centrifugation (6500 rpm for 15 min at 4 °C) to yield 3 g of cell paste/L of culture medium. The cell pellet was suspended in ice-cold buffer A and passed through a French press at 1200 psi. The lysate was centrifuged at 20000 rpm and 4 °C for 30 min, and the supernatant was immediately used in protein purification.

Recombinant Protein Purification. (*i*) *B. bronchiseptica RB50* D-*glycero*-D-*manno*-*Heptose-7-P Kinase* (*SwissProt accession number BB3463*). The cell supernatant was loaded onto a 40 cm × 5 cm DEAE-Sepharose 50−120 column, which was eluted with a 2 L linear gradient of KCl (from 0 to 0.5 M) in buffer A. The desired protein fractions were combined, adjusted to 15% $(NH_4)_2SO_4$ (w/v), and loaded onto an 18 cm × 3 cm Butyl-Sepharose column pre-equilibrated with buffer A containing 15% $(NH_4)_2SO_4$. The column was eluted with a 0.5 L linear gradient of $(NH_4)_2SO_4$ (from 15 to 0%) in buffer A. The desired protein fractions were combined, dialyzed against buffer A, concentrated, and stored at −80 °C: yield, 30 mg of protein/g of wet cells; mass, calcd 33846 Da, found 33846 Da.

(*ii*) *B. bronchiseptica* D-*glycero*-D-*manno*-*Heptose-1β,7-bisphosphate Phosphatase* (*SwissProt accession number Q7WG29*). The cell supernatant was chromatographed on a 40 cm × 5 cm DEAE-Sepharose column using a 2 L linear gradient of KCl (from 0 to 0.5 M) in buffer A as the eluant. The desired protein fractions were combined, adjusted to 20% $(NH_4)_2SO_4$ (w/v), and loaded onto an 18 cm × 3 cm Butyl-Sepharose column equilibrated with buffer A containing 20% $(NH_4)_2SO_4$. The column was eluted with a 0.5 L linear gradient of $(NH_4)_2SO_4$ (from 20 to 15%) in buffer A, followed by 0.5 L of buffer A containing 15% $(NH_4)_2SO_4$. The desired protein fractions were combined, concentrated, and loaded onto a Sephadex G-75 column. The column was eluted with buffer A, and the desired protein fractions were combined, dialyzed against buffer A, concentrated, and stored at −80 °C: yield, 6 mg of protein/g of wet cells; mass, calcd 19035 Da, found 19034 Da.

(*iii*) *B. bronchiseptica* D-*glycero*-D-*manno*-*Heptose-1β-phosphate Adenylyltransferase* (*SwissProt accession number Q7WF17*). The cell supernatant was chromatographed on a 40 cm × 5 cm DEAE-Sepharose 50−120 column using a 2 L linear gradient of NaCl (from 0 to 0.5 M) in buffer A as the eluant. The desired protein fractions were combined, adjusted to 15% $(NH_4)_2SO_4$ (w/v), and then loaded onto an 18 cm × 3 cm Butyl-Sepharose column equilibrated with buffer A containing 15% $(NH_4)_2SO_4$. The column was eluted at 4 °C with a 0.5 L linear gradient of $(NH_4)_2SO_4$ (from 15 to 0%) in buffer A. The desired protein fractions were combined, dialyzed against buffer A, concentrated, and stored at −80 °C: yield, 30 mg of protein/g of wet cells; mass, calcd 17440 Da, found 17322 Da. The N-terminal Met appears to have been removed by posttranslational modification.

(*iv*) *E. coli K-12 Phosphoheptose Isomerase* (*NCBI accession number NP_414757*). The supernatant was chromatographed on a 40 cm × 5 cm DEAE-Sepharose 50−120 column using a 2 L linear gradient of KCl (from 0 to 0.5 M) in buffer A as the eluant. The desired fractions were combined, dialyzed against buffer A, and concentrated: yield, 10 mg of protein/g of wet cells; mass, calcd 20815 Da, found 20815 Da.

(*v*) *A. thermoaerophilus* D-*glycero*-D-*manno*-*Heptose-7-P Kinase* (*SwissProt accession number Q9AGY8*). The *E. coli* DH5α clone was a kind gift from P. Messner (Universitaet fuer Bodenkultur Wien Zentrom fuer NanoBiotechnologie, Wien, Austria). The gene was subcloned into destination vector pDEST17 by homologous recombination. The isolated plasmid was used to transform *E. coli* BL21 star (DE3) cells. The transformed cells were grown at 30 °C and 200 rpm in 4 L of LB medium (containing 50 μg/mL ampicillin) to an OD_{660} of 0.6−1.0 and then induced with 0.4 mM IPTG. The cells were harvested by centrifugation at 6500 rpm for 15 min, suspended in 150 mL of ice-cold lysis buffer [50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM imidazole (pH 8.0)], and passed through a French pressure cell press at 1500 psi. The lysate was clarified by centrifugation and chromatographed on a Ni-NTA column (equilibrated with lysis buffer). The column was washed with

buffer containing 50 mM NaH_2PO_4, 300 mM NaCl, and 20 mM imidazole (pH 8.0) before the protein was eluted with buffer containing 50 mM NaH_2PO_4, 300 mM NaCl, and 250 mM imidazole (pH 8.0). The desired fractions were combined, concentrated, and dialyzed against 50 mM K^+HEPES and 10 mM $MgCl_2$ (pH 7.5): yield, 13 mg of protein/g of wet cells.

(*vi*) *E. coli* D-*glycero*-D-*manno*-*Heptose-1,7-bisphosphate Phosphatase* (*SwissProt accession number P63228*). The cell supernatant was chromatographed on a 40 cm × 5 cm DEAE-Sepharose column using a 2 L linear gradient of KCl (from 0 to 0.5 M) in buffer A as the eluant. The desired protein fractions were combined, adjusted to 20% $(NH_4)_2SO_4$ (w/v), and loaded onto an 18 cm × 3 cm Butyl-Sepharose column equilibrated with buffer A containing 20% $(NH_4)_2SO_4$. The column was eluted with a 0.5 L linear gradient of $(NH_4)_2SO_4$ (from 20 to 15%) in buffer A and then with 0.5 L of buffer A containing 15% $(NH_4)_2SO_4$. The desired protein fractions were combined and concentrated to yield 15 mg of protein/g of wet cells: mass, calcd 21294 Da, found 21163 Da. The N-terminal Met appears to have been removed by posttranslational modification.

(*vii*) *Bacteroides thetaiotaomicron* D-*glycero*-D-*manno*-*Heptose-1,7-bisphosphate Phosphatase* (*SwissProt accession number Q8AAI7*). This enzyme was purified from the cell supernatant by using the purification procedure described for the *E. coli* enzyme, with the modification that the protein solution was adjusted to 15% $(NH_4)_2SO_4$ before being loaded onto the Butyl-Sepharose column, and a 15 to 0% $(NH_4)_2SO_4$ gradient in buffer A was used as the eluant: yield, 7 mg of protein/g of wet cells.

(*viii*) *Mesorhizobium loti* D-*glycero*-D-*manno*-*Heptose-1,7-bisphosphate Phosphatase* (*SwissProt accession number Q98I56*). The *E. coli* clone was obtained from the PSI Materials Repository (clone MlCD00090897). The *E. coli* cells were grown at 37 °C and 200 rpm in 6 L of LB medium (containing 50 μg/mL kanamycin) to an OD_{600} of 0.6−1.0, induced with 5 mM L-arabinose at 25 °C, and harvested by centrifugation upon reaching an OD_{600} of ∼1.7. The cell pellet (30 g) was suspended in 300 mL of ice-cold lysis buffer [50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM imidazole (pH 8.0)] and passed through a French pressure cell press at 1500 psi. The lysate was clarified by centrifugation and chromatographed on a Ni-NTA column equilibrated with lysis buffer. The column was washed with 50 mM NaH_2PO_4, 300 mM NaCl, and 20 mM imidazole (pH 8.0) before the protein was eluted with 50 mM NaH_2PO_4, 300 mM NaCl, and 250 mM imidazole (pH 8.0). The desired fractions were combined, concentrated, and dialyzed against 50 mM K^+HEPES and 10 mM $MgCl_2$ (pH 7.5): yield, 20 mg of protein/g of wet cells.

Preparation of GmhB Substrates and Substrate Analogues. (*i*) D-*glycero*-D-*manno*-*Heptose 1β,7-Bisphosphate*. Ribose 5-phosphate (0.310 g, 1.0 mmol), 0.30 g of fructose 6-phosphate (1.0 mmol), 1.10 g of ATP (2.0 mmol), 10 μL (100 mM stock, 1 μmol) of D-glyceraldehyde 3-phosphate, and 100 μL (100 mM stock, 10 μmol) of thiamine pyrophosphate were combined with 100 mL of 50 mM Tris containing 10 mM $MgCl_2$. The pH of the solution was adjusted to 7.5 with aqueous potassium hydroxide, and 10 mg of commercial *E. coli* transketolase [0.16 unit/mg (Sigma)], 60 mg of freshly prepared recombinant (*E. coli* K-12) D-sedoheptulose-7-phosphate isomerase (3 mg/mL), and 4 mg of freshly prepared recombinant *B. bronchiseptica* [β-C(1)OH specific] D-*glycero*-D-*manno*-heptose-7-P kinase were added. The solution was incubated overnight at 25 °C and chromatographed on a Dowex 1X8-100 (HCO_3^-) column. After being washed with 600 mL of deionized water, the column was eluted with a 2 L linear gradient of NH_4HCO_3 (from 0 to 1 M). The column fractions were monitored using an acid-labile phosphate assay (*16*). The D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate-containing fractions (eluted at 0.4−0.5 M NH_4HCO_3) were combined and concentrated to provide 110 mg of D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate (28% yield): ^{1}H NMR δ 4.86 (d, 1H, J = 9.0 Hz), 4.00 (s, 1H), 3.80 (d, 2H, J = 12 Hz), 3.66 (m, 1H), 3.60 (t, 1H, J = 9.0 Hz), 3.47 (d, 1H, J = 9.5 Hz), 3.30 (d, 1H, J = 10.0 Hz); ^{13}C NMR δ 94.9 (bs), 77.1, 72.5, 70.9 (d, J = 4.4 Hz), 69.7 (d, J = 5.9 Hz), 66.2, 63.5 (d, J = 3.9 Hz); ^{31}P NMR δ 4.80 and 3.01; mass, calcd for $C_7H_{15}O_{13}P_2$ 368.9988 Da, found 368.9965 Da.

(*ii*) D-*glycero*-D-*manno*-*Heptose 1α,7-Bisphosphate*. D-Ribose 5-phosphate (0.5 g, 1.6 mmol), 0.5 g of D-fructose 6-phosphate (1.6 mmol), 1.70 g of ATP (3.0 mmol), 10 μL (100 mM stock, 1 μmol) of D-glyceraldehyde 3-phosphate, and 100 μL (100 mM stock, 10 μmol) of thiamine pyrophosphate were combined with 100 mL of 50 mM Tris containing 10 mM $MgCl_2$. The pH of the solution was adjusted to 7.5 with aqueous potassium hydroxide, and 25 mg of commercial *E. coli* transketolase [0.16 unit/mg (Sigma)], 120 mg of freshly prepared recombinant (*E. coli* K-12) D-sedoheptulose-7-phosphate isomerase, and 120 mg of recombinant *A. thermoaerophilus* [α-C(1)OH specific] D-*glycero*-D-*manno*-heptose-7-P kinase were added. The solution was incubated overnight at 25 °C and subjected to the same workup and purification procedure described above to provide 600 mg of D-*glycero*-D-*manno*-heptose 1α,7-bisphosphate (84% yield): ^{1}H NMR δ 5.20 (d, 1H, J = 8.0 Hz), 4.03 (dd, 1H, J = 6.75 Hz, J = 3.0 Hz), 3.90 (td, 1H, J = 7.8 Hz, J = 3.5 Hz), 3.81 (m, 2H), 3.77 (t, 1H), 3.55 (t, 1H, J = 3.0 Hz), 3.49 (td, 1H, J = 10.0 Hz, J = 3.0 Hz); ^{13}C NMR δ 94.9 (d, J = 3.8 Hz), 73.1, 70.6 (d, J = 6.9 Hz), 70.4 (d, J = 6.6 Hz), 70.1, 66.9, 64.5 (d, J = 3.4 Hz); ^{31}P NMR δ 3.42, 1.47; mass, calcd for $C_7H_{15}O_{13}P_2$ 368.9988 Da, found 368.9981 Da.

(*iii*) D-*glycero*-D-*manno*-*Heptose 1β-Phosphate and* D-*glycero*-D-*manno*-*Heptose 1α-Phosphate*. D-*glycero*-D-*manno*-Heptose 1β,7-bisphosphate (30 mg) and D-*glycero*-D-*manno*-heptose 1α,7-bisphosphate (80 mg) were separately dissolved in 1 mL of a 10 mM $MgCl_2$/10 mM K^+HEPES mixture (pH 7.5) to which an aliquot (1 or 20 μL) of 5 mM *E. coli* GmhB (vide supra) was then added. Following a 12 h incubation period at 25 °C, each reaction mixture was subjected to anion exchange column chromatography (vide supra) to provide D-*glycero*-D-*manno*-heptose 1β-phosphate in 83% yield and D-*glycero*-D-*manno*-heptose 1α-phosphate in 92% yield.

(a) D-*glycero*-D-*manno*-Heptose 1β-phosphate: ^{1}H NMR δ 4.90 (d, 1H, J = 8.5 Hz), 3.84 (t, 1H, J = 3.5 Hz), 3.80 (s, 1H), 3.59 (m, 2H), 3.49 (m, 2H), 3.30 (t, 1H); ^{13}C NMR δ 94.9, 76.6, 72.3, 70.9, 70.4 (d), 66.7, 61.0; ^{31}P NMR δ 0.965; mass, calcd for $C_7H_{14}O_{10}P$ 289.0325 Da, found 289.0300 Da.

(b) D-*glycero*-D-*manno*-Heptose 1α-phosphate: ^{1}H NMR δ 4.98 (d, 1H, J = 8.5 Hz), 3.68 (p, 1H, J = 3.0 Hz), 3.58 (m, 3H), 3.42 (m, 3H); ^{13}C NMR δ 94.4 (d, J = 4.5 Hz), 71.9, 71.5, 70.3 (d, J = 7.2 5 Hz), 69.9, 67.5, 61.3; ^{31}P NMR δ 2.36 (d, J = 7.3 Hz); mass, calcd for $C_7H_{14}O_{10}P$ 289.0325 Da, found 289.0330 Da.

(*iv*) *Preparation of Glucose 1β,6-Bisphosphate*. Maltose (0.36 g, 1.0 mmol) and 0.4 g (3.0 mmol) of KH_2PO_4 were dissolved in 100 mL of a 50 mM Tris/10 mM $MgCl_2$ mixture. The pH of the solution was adjusted to 7.5, and freshly prepared recombinant *Lactococcus lactis* maltose phosphorylase (100 units) (*17*) was added. Following a 12 h incubation period

at 25 °C, ATP (1.0 g, 1.9 mmol) was added. The pH of the solution was adjusted to 9.0 before addition of commercial *Bacillus stearothermophilus* fructose-6-phosphate kinase (100 units). Following a 12 h incubation period at 25 °C, the reaction mixture was subjected to the same purification procedure described above to yield 180 mg of glucose 1β,6-bisphosphate: ^{1}H NMR δ 4.67 (d, 1H, J = 14 Hz), 3.82 (dd, 1H, J = 11 Hz, J = 5 Hz), 3.76 (s, 1H), 3.31 (s, 3H), 3.10 (t, 1H); ^{13}C NMR δ 94.9 (d, J = 4.25 Hz), 73.2 (d, J = 6.75 Hz), 72.7, 72.2 (d, J = 5.75 Hz), 66.7, 60.7; ^{31}P NMR δ 4.80, 2.96; mass, calcd for $C_6H_{13}O_{12}P_2$ 338.9882 Da, found 338.9883 Da.

(*v*) *Preparation of Mannose 1β,6-Bisphosphate and Mannose 1α,6-Bisphosphate*. In two separate solutions, mannose (0.30 g, 1.67 mmol) and 2.0 g (3.6 mmol) of ATP were dissolved in 100 mL of 50 mM Tris containing 10 mM $MgCl_2$. The pH of the solutions was adjusted to 7.5 with aqueous potassium hydroxide before addition of 1000 units of commercial hexokinase to each. Following a 12 h incubation period at 25 °C, 5 mL of 2 mM *B. bronchiseptica* D-*glycero*-D-*manno*-heptose-7-P kinase and *A. thermoaerophilus* D-*glycero*-D-*manno*-heptose-7-P kinase (vide supra) were added to the individual solutions. After 36 h, the reaction solutions were chromatographed as described for D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate (vide supra) to give 170 mg of mannose 1β,6-bisphosphate and 470 mg of mannose 1α,6-bisphosphate, respectively.

(a) Mannose 1β,6-bisphosphate: ^{1}H NMR δ 4.86 (d, 1H, J = 9 Hz), 3.78 (m, 2H), 3. (dd, 1H, J = 11.5 Hz, J = 6 Hz), 3.48 (d, 2H, J = 7.5 Hz), 3.24 (s, 1H); ^{13}C NMR δ 94.8, 75.5 (d, J = 7.13 Hz), 72.0, 71.1 (d, J = 4.25 Hz), 65.8, 62.7 (d, J = 2.75 Hz); ^{31}P NMR δ 4.71, 2.81; mass, calcd for $C_6H_{13}O_{12}P_2$ 338.9882 Da, found 338.9870 Da.

(b) Mannose 1α,6-bisphosphate: ^{1}H NMR δ 5.06 (d, 1H, J = 8 Hz), 3.72 (m, 4H), 3.64 (d, 1H, J = 9.50 Hz), 3.52 (t, 1H, J = 10.0 Hz); ^{13}C NMR δ 94.6 (d, J = 3.5 Hz), 71.9 (d, J = 7.0 Hz), 70.5 (d, J = 6.5 Hz), 69.2, 65.9, 62.6 (d, J = 2.75 Hz); ^{31}P NMR δ 4.30, 1.98; mass, calcd for $C_6H_{13}O_{12}P_2$ 338.9882 Da, found 338.9888 Da.

(*vi*) *Preparation of Sedoheptulose 1,7-Bisphosphate*. Erythrose 4-phosphate (25 mg, 0.1 mmol) and dihydroxyacetone phosphate (20 mg, 0.1 mmol) were dissolved in 100 mL of 50 mM Tris containing 10 mM $MgCl_2$. The pH of the solution was adjusted to 7.5 with aqueous potassium hydroxide before addition of 50 μL of 200 μM recombinant *Giardia lamblia* fructose-1,6-bisphosphate aldolase (*18*). After 15 h, the reaction solution was chromatographed as described for D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate (vide supra) to yield 30 mg of sedoheptulose 1,7-bisphosphate (*19*): mass, calcd for $C_7H_{14}O_{10}P$ 368.9988 Da, found 368.9961 Da.

(*vii*) *L-Histidinol Phosphate*. This compound was prepared from L-histidinol and polyphosphoric acid using a modification of the published procedure (*20*). A mixture of L-histidinol dihydrochloride (1 g, 4.5 mmol) and polyphosphoric acid (4 g) was heated on a steam bath for 1 h, after which time 50 mL of ice-cold water was added. The pH of the solution was adjusted to 7.5 with aqueous barium hydroxide, and the resulting precipitant was removed by centrifugation. The supernatant was chromatographed as described for D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate (vide supra) to yield 300 mg of L-histidinol phosphate: ^{1}H NMR δ 8.16 (s, 1H), 7.18 (s, 1H), 3.89 (s, 1H), 3.77 (s, 1H), 3.61 (s, 1H), 3.04 (m, 2H); ^{13}C NMR δ 134.7, 129.3, 117.4, 62.4, 51.3, 24.8; ^{31}P NMR δ 5.3; mass, calcd for $C_6H_{11}N_3O_4P$ 220.0487 Da, found 220.0468 Da.

(*viii*) *D-glycero-D-manno-Heptose 1β-ADP and D-Glucose 1β-ADP*. D-*glycero*-D-*manno*-Heptose 1β-phosphate (0.01 g, 0.038 mmol) and 0.033 g of ATP (0.06 mmol) or D-glucose 1β-phosphate (100 mg, 0.3 mmol) (vide supra) and 0.17 g of ATP (0.3 mmol) were dissolved in 5 mL of 50 mM Tris containing 10 mM $MgCl_2$. The pH of the solution was adjusted to 7.5 with aqueous potassium hydroxide before addition of a 1 mL aliquot from a 1 mg/mL stock solution of *B. bronchiseptica* D-*glycero*-D-*manno*-heptose-1β-phosphate adenylyltransferase (vide supra). The solution was incubated for 1 h at room temperature and then chromatographed on a Sephadex G-10 column using deionized water as the eluant. The desired fractions were combined and lyophilized to give 20 mg of D-*glycero*-D-*manno*-heptose 1β-ADP or 52 g of D-glucose 1β-ADP.

(a) D-*glycero*-D-*manno*-Heptose 1β-ADP: ^{1}H NMR δ 8.48 (s, 1H), 8.21 (s, 1H), 6.12 (d, 1H, J = 5.5 Hz), 5.19 (d, 1H, J = 8.5 Hz), 4.52 (t, 1H, J = 4.5 Hz), 4.39 (s, 1H), 4.22 (s, 2H), 4.06 (d, 1H, J = 2.5 Hz), 3.99 (s, 1H), 3.74 (d, 2H, J = 3.5 Hz), 3.68 (t, 1H, J = 9.5 Hz), 3.62 (dd, 1H, J = 9.5 Hz, J = 2.5 Hz), 3.43 (dd, 1H, J = 9.5 Hz, J = 2.5 Hz); ^{13}C NMR δ 154.9, 152.1, 148.4, 139.2, 117.9, 95.2 (d, J = 3.6 Hz), 86.2, 83.2 (d, J = 8.9 Hz), 76.3, 72.0, 71.1, 69.9 (d, J = 7.0 Hz), 69.7, 66.4, 64.6 (d, J = 6.1 Hz), 61.0; ^{31}P NMR δ −10.50 (d, J = 17.4 Hz), −12.51 (dd, J = 16.97 Hz, J = 6.46 Hz); mass, calcd for $C_{17}H_{26}N_5O_{16}P_2$ 618.0850 Da, found 618.0845 Da.

(b) D-Glucose 1β-ADP: ^{1}H NMR δ 8.49 (s, 1H), 8.24 (s, 1H), 6.13 (d, 1H, J = 6.0 Hz), 5.00 (t, 1H, J = 7.5 Hz), 4.53 (s, 1H), 4.39 (s, 1H), 4.22 (s, 2H), 3.87 (d, 1H, J = 12.5 Hz), 3.67 (dd, 1H, J = 12.0 Hz, J = 6.0 Hz), 3.51 (dd, 2H, J = 19 Hz, J = 8.5 Hz), 3.36 (t, 2H, J = 8.0 Hz); ^{13}C NMR δ 155.0, 152.3, 148.5, 139.3, 118.0, 97.5 (d, J = 5.5 Hz), 86.5, 83.4 (d, J = 8.5 Hz), 76.0, 74.7, 73.2 (d, J = 8.0 Hz), 70.0, 69.0, 64.9 (d, J = 5.3 Hz), 60.3; ^{31}P NMR δ −10.39 (d, J = 20.0 Hz), −12.14 (dd, J = 20.0 Hz, J = 7.27 Hz); mass, calcd for $C_{16}H_{24}N_5O_{15}P_2$ 588.0744, found 588.0734.

Substrate Activity Screens. (*i*) *^{31}P NMR Analysis*. *E. coli* GmhB-catalyzed hydrolysis reactions of the sugar bisphosphates D-glucose 1α,6-diphosphate, D-glucose 1β,6-bisphosphate, D-mannose 1β,6-bisphosphate, D-mannose 1α,6-bisphosphate, D-fructose 1α,6-bisphosphate, D-sedoheptulose 1,7-bisphosphate, D-*glycero*-D-*manno*-heptose 1α,7-bisphosphate, and D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate were examined by using ^{31}P NMR in the proton-decoupled mode. The reaction solutions initially contained 10 mM sugar bisphosphate, 1 or 50 μM GmhB, 2 mM $MgCl_2$, and 50 mM K^+HEPES (pH 7.5, 25 °C) and were incubated for 5−30 h. ^{31}P NMR spectra of inorganic phosphate and the anticipated sugar monophosphate products were measured for use as chemical shift standards.

(*ii*) *Spectrophotometric Analysis*. The rate of *E. coli* GmhB-catalyzed hydrolysis of phosphate esters was measured spectrophotometrically by using the Enzcheck phosphate assay kit (Invitrogen) to continuously monitor phosphate formation (at 360 nm, $\Delta\varepsilon$ = 11 mM^{-1} cm^{-1}). The 1 mL reaction solutions initially contained 50 mM Tris-HCl (pH 7.5, 25 °C), 1 mM $MgCl_2$, 1 mM substrate, 10−20 μM *E. coli* GmhB, 0.2 mM MESG, and 1 unit of purine nucleoside phosphorylase. The background level of phosphate release was measured using the corresponding control reaction, which excluded the GmhB. In cases where the catalyzed and control reactions produced similar absorbance versus time traces, the phosphate ester was judged not to be a substrate and the k_{cat} defined as $< 1 \times 10^{-4}$ s^{-1}.

Steady-State Kinetic Constant Determination. Steady-state kinetic constant determinations were conducted at 25 °C using reaction solutions initially containing GmhB (0.001−10 μM), varying concentrations of phosphate ester (0.5−5K_m), 1 mM $MgCl_2$, 0.2 mM 2-amino-6-mercapto-7-methylpurine ribonucleoside, 1 unit of purine nucleoside phosphorylase, and 50 mM Tris-HCl (pH 7.5). Data were fitted to the equation $V_0 = V_{max}[S]/(K_m + [S])$ with KinetAsyst I, where V_0 is the initial velocity, V_{max} is the maximum velocity, [S] is the substrate concentration, and K_m is the Michaelis−Menten constant for substrate. The k_{cat} value was calculated from V_{max} and [E] according to the equation $k_{cat} = V_{max}/[E]$, where [E] is the enzyme concentration. The inhibition constants for D-*glycero*-D-*manno*-heptose 1β-phosphate and histidinol phosphate were determined by measuring the initial velocities of *E. coli* GmhB-catalyzed hydrolysis of D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate as a function of substrate and inhibitor concentration [I]. The initial velocity data were fitted to the equation $V = (V_{max}[S])/[K_m(1 + [I]/K_{is}) + [S](1 + [I]/K_{ii})]$ to define the slope and intercept inhibition constants, K_{is} and K_{ii}, respectively.

RESULTS AND DISCUSSION

GmhB Substrate Specificity Profile Determination. The participation of the GmhB orthologs from *A. thermoaerophilus* and *E. coli* in the respective biosynthetic pathways leading to D-*glycero*-D-*manno*-heptose 1α-GDP and L-*glycero*-D-*manno*-heptose 1β-ADP (Figure 1) was previously demonstrated by reconstitution of the pathways in vitro (*4*, *5*). However, the catalytic efficiencies and anomeric specificities of the two phosphatases were not reported. More recently, *E. coli* GmhB (along with 22 other *E. coli* HAD phosphatases) was screened for activity against an 80-compound library of common organophosphate metabolites representing a variety of structural classes (*21*). The *E. coli* GmhB was found to display activity toward only one library compound, fructose 1,6-bisphosphate (k_{cat}/K_m = 450 M^{-1} s^{-1}). The first objective of this study was to measure the catalytic efficiency of *E. coli* GmhB toward its physiological substrate and further define its substrate specificity profile. This was accomplished by preparing and measuring the substrate activities of natural hexose 1,6-bisphosphates and heptose 1,7-bisphosphates, including the α- and β-anomers of D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate.

With the exceptions of fructose 1,6-bisphosphate and glucose 1α,6-bisphosphate, the hexose 1,6-bisphosphates and heptose 1,7-bisphosphates were prepared by enzymatic synthesis as illustrated in Figure 2 (see Materials and Methods for details). Fructose 1,6-bisphosphate proved to be a poor substrate for *E. coli* GmhB [k_{cat}/K_m = 30 M^{-1} s^{-1} (Table 1)]. ^{31}P NMR spectral analysis of the reaction between 10 mM fructose 1,6-bisphosphate and 50 μM GmhB in 50 mM Tris-HCl and 1 mM $MgCl_2$ (pH 7.5, 25 °C, 30 h) revealed that the substrate C(1)P was selectively removed. Specifically, whereas the spectrum of the fructose 1,6-bisphosphate contains resonances at 5.0 ppm [C(1)P] and 4.7 ppm [C(6)P], the spectrum of the product mixture exhibited resonances at 4.7 ppm (fructose 6-phosphate) and 3.3 ppm (orthophosphate). Fructose 6-phosphate and fructose 1-phosphate were subsequently shown to have no detectable substrate activity ($k_{cat} < 1 \times 10^{-4}$ s^{-1}) by using a spectrophotometric assay for orthophosphate. Glucose 1α,6-bisphosphate, glucose 1β,6-bisphosphate, mannose 1α,6-bisphosphate, and mannose 1β,6-phosphate (Figure 2) were also found to have no detectable substrate activity with *E. coli* GmhB; however, D-sedoheptulose 1,7-bisphosphate (Figure 2) displayed a modest level of activity [k_{cat}/K_m = 800 M^{-1} s^{-1} (Table 1)]. ^{31}P NMR spectral analysis of the reaction of 10 mM D-sedoheptulose 1,7-bisphosphate with 50 μM GmhB in 50 mM Tris-HCl and 1 mM $MgCl_2$ (pH 7.5, 25 °C, 30 h) revealed that one of the two phosphate groups (resonances coincide at 5.5 ppm) was removed (indicated by the appearance of an orthophosphate resonance at 3.3 ppm) but did not identify which phosphate group was removed. Together, these findings, along with the kinetic constants reported in Table 1, show that *E. coli* GmhB is not active with monophosphorylated hexoses and heptoses and that it has limited activity with ketohexose and ketoheptose bisphosphates.

E. coli GmhB was shown to dephosphorylate its physiological substrate D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate with high efficiency [$k_{cat}/K_m = 7 \times 10^6$ M^{-1} s^{-1} (Table 1)] and the corresponding α-anomer, D-*glycero*-D-*manno*-heptose 1α,7-bisphosphate, with reduced, yet physiologically significant, efficiency [$k_{cat}/K_m = 7 \times 10^4$ M^{-1} s^{-1} (Table 1)]. ^{31}P NMR analysis of the reaction between 10 μM *E. coli* and 10 mM D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate [C(1)P δ = 3.0 ppm; C(7)P δ = 4.8 ppm] or D-*glycero*-D-*manno*-heptose 1α,7-bisphosphate [C(1)P δ = 2.9 ppm; C(7)P δ = 5.2 ppm] in 50 mM Tris-HCl and 1 mM $MgCl_2$ (pH 7.5, 25 °C, 5 h) revealed complete removal of the substrate C(7) phosphate group and no detectable hydrolysis of the corresponding D-*glycero*-D-*manno*-heptose 1β-phosphate [C(1)P δ = 3.0 ppm] or D-*glycero*-D-*manno*-heptose 1α-phosphate [C(1)P δ = 2.9 ppm] product. Thus, *E. coli* GmhB is specific in targeting removal of the C(7)P of D-*glycero*-D-*manno*-heptose 1,7-bisphosphate.

D-*glycero*-D-*manno*-Heptose 1β-phosphate and D-*glycero*-D-*manno*-heptose 1α-phosphate were synthesized (Figure 2) and shown by using a spectrophotometric assay for orthophosphate to have no detectable substrate activity with *E. coli* GmhB. This observation agrees with the findings of the ^{31}P NMR analysis. The D-*glycero*-D-*manno*-heptose 1β-phosphate was shown to be a noncompetitive inhibitor. On the basis of the measured inhibition constants (K_{is} = 60 ± 10 μM, and K_{ii} = 150 ± 50 μM), it is evident that D-*glycero*-D-*manno*-heptose 1β-phosphate binds to the GmhB active site tightly, yet nonproductively.

Histidinol phosphate was tested as a substrate for *E. coli* GmhB, but no detectable activity was observed. Although histidinol phosphate bears little resemblance to the physiological substrate of GmhB, the proposed evolutionary link between GmhB and the histidinol-phosphate phosphatase (*1*, *3*, *22*) made some amount of cross reactivity a possibility. Histidinol phosphate proved to be a noncompetitive inhibitor of *E. coli* GmhB with a modest binding affinity (K_{is} = 1 ± 0.1 mM, and K_{ii} = 6 ± 1 mM). Therefore, like D-*glycero*-D-*manno*-heptose 1β-phosphate and D-*glycero*-D-*manno*-heptose 1α-phosphate, histidinol phosphate appears to escape GmhB catalysis by binding in a nonproductive mode.

Together, these results characterize *E. coli* GmhB as a HAD phosphatase with a narrow substrate range and a high catalytic efficiency toward its physiological substrate. Because these two properties are characteristic of a highly evolved metabolic enzyme, they indicate that the evolution of GmhB has occurred under strong selection pressure. As HAD phosphatases primarily assume the roles of repair enzymes, housekeeping enzymes, or phosphate scavengers (*23*), they tend to be inefficient and lax in their substrate specificities. The 80-compound screen applied to the 22 HAD phosphatases of *E. coli* revealed a general tendency toward substrate promiscuity (*21*). Focused substrate activity

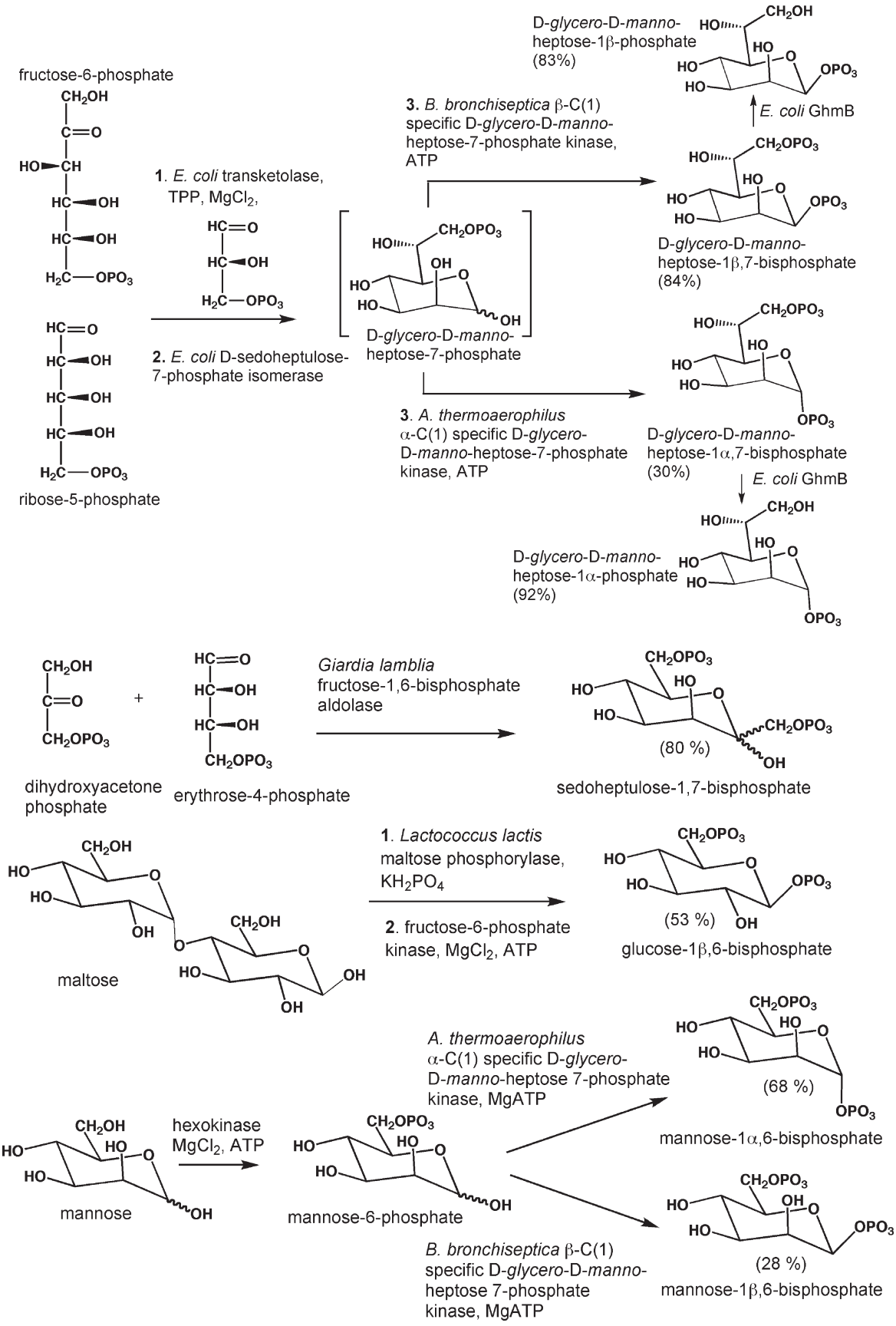

FIGURE 2: Chemical steps and enzyme catalysts used in the synthesis of the hexose and heptose bisphosphates used in the substrate profile determination of GmhB.

screening designed to explore structural space closely tied to the respective physiological substrates of two HAD housekeeping enzymes, *E. coli* NagD (a nucleoside monophosphatase) and BT4131 from *B. thetaiotaomicron* (a hexose-6-phosphate phosphatase), showed that although the two enzymes do not hydrolyze organophosphates from other structural classes,

Table 1: Steady-State Kinetic Constants for *E. coli* GMHB-Catalyzed Hydrolysis of Phosphate Esters in 50 mM Tris-HCl (pH 7.5, 25 °C) Containing 1 mM $MgCl_2$[a]

substrate	P[b]	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate	C(7)P	35.7 ± 0.2	5.0 ± 0.1	7×10^6
D-*glycero*-D-*manno*-heptose 1α,7-bisphosphate	C(7)P	4.6 ± 0.1	67 ± 1	7×10^4
sedoheptulose 1,7-bisphosphate	C(7) or C(1)	0.51 ± 0.01	610 ± 20	8×10^2
fructose 1,6-bisphosphate	C(1)P	0.039 ± 0.001	1501 ± 5	3×10^1

[a]Compounds that were tested with the *E. coli* or *B. bronchiseptica* GmhB but failed to display substrate activity ($k_{cat} < 1 \times 10^{-4}$ s^{-1}) include fructose 6-phosphate, mannose 6-phosphate, fructose 1-phosphate, glucose 1α,6-bisphosphate, glucose 1β,6-bisphosphate, mannose 1α,6-bisphosphate, mannose 1β, 6-phosphate, histidinol phosphate, phosphoenolpyruvate, uridine 5-diphosphate, adenosine 5-monophosphate, and dimyristoyl-*sn*-glycero-3-phosphate. [b]Phosphate removed in the catalyzed reaction.

organophosphates that belong to the same structural class as the physiological substrates are hydrolyzed with modest efficiency ($k_{cat}/K_m = 1 \times 10^2$ to 1×10^4 M^{-1} s^{-1}) and with little distinction (*24*, *25*). *E. coli* GmhB does not fit this profile but instead mirrors that of *E. coli* 2-keto-3-deoxyoctulosonate-8-phosphate (KDO-8-P) phosphatase, a member of the HAD "8KDO" subfamily (*26*, *27*). Like *E. coli* GmhB, KDO-8-P phosphatase functions in the Lipid A core glycan biosynthetic pathway. Although these two phosphatases support different branches of the Lipid A core pathway, we propose that their evolution may have been subject to a common selection pressure.

Tuning GmhB Anomeric Specificity for Biochemical Function. An interesting feature of the specialization of GmhB is its anomeric specificity. Whereas the physiological substrate for the *E. coli* enzyme is the β-anomer of D-*glycero*-D-*manno*-heptose 1,7-bisphosphate, the α-anomer is the physiological substrate of the *A. thermoaerophilus* GmhB (Figure 1). A priori, one might have expected that GmhB would lack the ability to discriminate the spatial disposition of the substrate C(1) phosphate. However, this work has shown that *E. coli* GmhB prefers the β-anomer over the α-anomer by a factor of 100:1 (Table 1). If this specificity is a consequence of selection pressure, then the reverse might hold true for the *A. thermoaerophilus* GmhB; i.e., it would show a preference for the α-anomer over the β-anomer. Placed in a broader context, the important question is "whether the anomeric specificities of the pathway kinase and phosphatase co-evolved". To address this issue of anomeric specificity and physiological substrate, comparisons were made for four GmhB orthologs derived from distantly related bacterial species. The selection of orthologs for study was guided by technical concerns that included the availability of the GmhB clone (or the genomic DNA for gene cloning) and protein solubility, as well whether the physiological substrate could be assigned with confidence using bioinformatics analysis. In addition, the availability of the X-ray crystal structure, which would allow anomeric specificity to be analyzed in the context of active site structure, was also a factor guiding ortholog selection.

The steady-state kinetic constants were measured for GmhB-catalyzed hydrolysis of the α- and β-anomers of D-*glycero*-D-*manno*-heptose 1,7-bisphosphate (Table 2). The orthologs from *E. coli*, *B. bronchiseptica*, and *M. loti* were shown to prefer the β-anomer (by 100-, 150-, and 18-fold, respectively), whereas the GmhB from *B. thetaiotaomicron* prefers the α-anomer (6-fold). To correlate the observed GmhB anomeric preference with the anomeric configuration of the physiological substrate, we investigated the anomeric specificity of the kinase and nucleotidyl transferase of the corresponding biosynthetic pathway (Figure 1).

In the case of GmhB from *B. bronchiseptica* (Q7WG29), the genes encoding the pathway kinase and nucleotidyl transferase were cloned and the substrate specificity of the purified enzymes was determined. The genes were located within the *B. bronchiseptica* genome via Blast searches using the sequences of the *E. coli* and *A. thermoaerophilus* pathway kinases and nucleotidyl transferases as queries. The sequence of bifunctional kinase/adenyl transferase (HldE_k/HldE_t) from *E. coli* identified the putative *B. bronchiseptica* kinase Q7WGU8 (51% identical) and nucleotidyl transferase Q7WF17 (48% identical). The purified HldE_k homologue Q7WGU8 was shown to selectively produce D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate when incubated with ATP and D-*glycero*-D-*manno*-heptose 7-phosphate in buffered solution (Figure 2) (see Materials and Methods for details of product isolation and spectral properties). Likewise, the purified HldE_t homologue Q7WF17 was shown to form D-*glycero*-D-*manno*-heptose 1β-ADP upon incubation with ATP and D-*glycero*-D-*manno*-heptose 7-phosphate in buffered solution (see Materials and Methods for details of product isolation and spectral properties). No product was obtained when GTP, UTP, or CTP was substituted for ATP. Moreover, no product was obtained when D-*glycero*-D-*manno*-heptose 1α-phosphate was substituted for D-*glycero*-D-*manno*-heptose 1β-phosphate. On the other hand, Q7WF17 was shown to catalyze ATP phosphorylation of glucose 1β-phosphate to form glucose 1β-ADP, as demonstrated by spectral analysis of the purified compound. Thus, Q7WF17 is specific for ATP as the phosphoryl donor, and for the β-anomer of hexose or heptose 1-phosphate as the adenyl group acceptor.

On the basis of the findings of a Blast search of the *B. bronchiseptica* genome, we conclude that only one set of pathway kinase−phosphatase−transferase activities is present and that the three enzymes (Q7WGU8, Q7WG29, and Q7WF17) convert the D-*glycero*-D-*manno*-heptose 7-phosphate to D-*glycero*-D-*manno*-heptose 1β-ADP.

Table 2: Steady-State Kinetic Constants for Hydrolysis of D-*glycero*-D-*manno*-Heptose 1α,7-Bisphosphate and D-*glycero*-D-*manno*-Heptose 1β,7-Bisphosphate by the Selected D-*glycero*-D-*manno*-Heptose-1,7-bisphosphate Phosphatases (GmbBs)[a]

enzyme source	anomer	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})	β/α
E. coli	α	4.6 ± 0.1	67 ± 1	6.9×10^4	100
E. coli	β	35.7 ± 0.2	5.0 ± 0.1	7.1×10^6	
B. bronchiseptica	α	5.9 ± 0.2	280 ± 30	2.1×10^4	150
B. bronchiseptica	β	22 ± 0.5	6.9 ± 0.5	3.2×10^6	
M. loti	α	4.4 ± 0.1	58 ± 1	7.6×10^4	18
M. loti	β	18 ± 0.2	13 ± 0.2	1.4×10^6	
B. thetaiotaomicron	α	29 ± 1	33 ± 1	8.8×10^5	0.17
B. thetaiotaomicron	β	13 ± 1	86 ± 1	1.5×10^5	

[a]The reactions were conducted in 50 mM Tris-HCl (pH 7.5, 25 °C) containing 1 mM $MgCl_2$.

The pathway phosphoheptose isomerase (Q7WEW5) and D-*glycero*-D-*manno*-heptose-1β-ADP epimerase (Q7WGU9) were identified using the sequences of the known *E. coli* enzymes as queries (40 and 54% identical, respectively). Thus, all enzymes of the L-*glycero*-D-*manno*-heptose 1β-ADP pathway are accounted for. It can be concluded that the physiological substrate of the *B. bronchiseptica* GmhB is the β-anomer and that it, like the *E. coli* GmhB, displays significantly higher reactivity with the β-anomer versus the α-anomer.

The biochemical function of GmhB from *M. loti* was also examined. Blast searches of the *M. loti* genome identified a single GmhB gene (Q98I56), single matches to the *E. coli* bifunctional kinase/adenyl transferase (HldE_k/HldE_t) (*M. loti* Q98I54, 42% identical sequence), and no matches[3] to the *A. thermoaerophilus* C(1)OH α-anomer specific D-*glycero*-D-*manno*-heptose-7-phosphate kinase (HddA) or the *A. thermoaerophilus* D-*glycero*-D-*manno*-heptose-1α-phosphate guanidyl transferase HddC. The genes encoding HldE_k/HldE_t and GmhB are located in a gene cluster, which includes genes that encode heptose-phosphate isomerase GmhA (Q98I55; 35% identical with *E. coli* GmhA P63224) and the D-*glycero*-D-*manno*-heptose-1β-ADP epimerase, GmhD (Q98I52; 61% identical with *E. coli* GmhA P67910). The presence of the epimerase, GmhD, is a strong indicator that the end product of the GmhB pathway is L-*glycero*-D-*manno*-heptose 1β-ADP and thus that the physiological substrate of the *M. loti* GmhB is D-*glycero*-D-*manno*-heptose 1β-phosphate.

The biochemical functions of the GhmB orthologs in *E. coli*, *B. bronchiseptica*, and *M. loti* are identical in that they support the synthesis of L-*glycero*-D-*manno*-heptose 1β-ADP for the production of the Lipid A glycan core. The three bacterial species derive from different subdivisions of the kingdom of proteobacteria (gamma, beta, and alpha, respectively). Even though the sequences of three GmhB orthologs have diverged to ~70% nonidentity (as deduced from pairwise alignments), the catalytic efficiencies toward their physiological substrate D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate [k_{cat}/K_m = 7.1, 3.3, and 1.4×10^6 M^{-1} s^{-1}, respectively (Table 2)] are conserved as are their lower catalytic efficiencies toward D-*glycero*-D-*manno*-heptose 1β,7-bisphosphate [k_{cat}/K_m = 6.6, 2.1, and 7.6×10^4 M^{-1} s^{-1}, respectively (Table 2)]. However, the anomeric specificity of the *M. loti* GmhB is noticeably lower (18:1) than those of the *E. coli* GmhB (100:1) and *B. bronchiseptica* GmhB (150:1). The X-ray structures of all three enzymes are now known and will be discussed in the context of the catalytic efficiency and substrate specificity in the following paper (*15*).

The final objective of this investigation was to demonstrate the biochemical function of GmhB from *B. thetaiotaomicron*. Blast searches were conducted to show that there is a single copy of a GmhB encoding gene and that it is contained in a gene cluster that encodes the homologues of all four of the *A. thermoaerophilus* D-*glycero*-D-*manno*-heptose 1α-GDP pathway enzymes: GmhA (Q8AA19; 58% identical sequence), HddA (Q8AAJ0; 58% identical sequence), GmhB (Q8AAI7; 34% identical sequence), and HddC (Q8AAI8; 47% identical sequence) (Figure 1). The gene cluster is located in one of the numerous capsular polysaccharide biosynthetic loci of the *B. thetaiotaomicron* genome. Thus, the likely function of the D-*glycero*-D-*manno*-heptose 1α-GDP pathway of *B. thetaiotaomicron* is to generate activated D-*glycero*-D-*manno*-heptose units for incorporation into the extracellular capsular polysaccharide. The physiological substrate of the *B. thetaiotaomicron* GmhB is D-*glycero*-D-*manno*-heptose 1α,7-bisphosphate, which is shown to have an activity constant ($k_{cat}/K_m = 8.8 \times 10^5$ M^{-1} s^{-1}) 6-fold higher than that measured for the β-anomer ($k_{cat}/K_m = 1.5 \times 10^5$ M^{-1} s^{-1}) (Table 2).

[3]The C(1)OH β-anomer specific D-*glycero*-D-*manno*-heptose-7-phosphate kinase (HldE_k) and the C(1)OH α-anomer specific D-*glycero*-D-*manno*-heptose-7-phosphate kinase (HddA) derive from different enzyme superfamilies (ribokinase and GHMP, respectively) and thus are unambiguously distinguished at the amino acid sequence level. Likewise, the nucleotidyl transferases HldE_t and HddC of the respective pathways have evolved within separate clades of the cytidyltransferase superfamily and are thus easily distinguished at the amino acid sequence level.

SUMMARY AND CONCLUSIONS

The results of the studies described above demonstrate that the bacterial HAD phosphatase GmhB has a narrow substrate range and a high catalytic efficiency toward its physiological substrate D-*glycero*-D-*manno*-heptose 1,7-bisphosphate. Consequently, GmhB is a highly evolved metabolic enzyme that functions in the L-*glycero*-D-*manno*-heptose 1β-ADP and D-*glycero*-D-*manno*-heptose 1α-GDP pathways. The anomeric specificity of GmhB correlates with that of the pathway kinase. Thus, the evolutionary path leading from an ancient ancestor to modern GmhB diverged to support two distinct biochemical functions. Notably, each of the four GmhB orthologs examined has retained a physiologically significant level of activity toward the anomeric antipode of the physiological substrate. Anomer selectivity may have evolved after catalytic efficiency, or alternatively, structural changes made to enhance the activity toward the physiological substrate might have also enhanced the activity toward the glycoside epimer, but to a lesser degree. On a practical level, a GmhB that dephosphorylates both D-*glycero*-D-*manno*-heptose 1,7-bisphosphate anomers might be capable of multitasking, i.e., supporting both biosynthetic pathways within the same bacterium. Indeed, the expansion of GmhB biological range via horizontal gene transfer might be facilitated if the GmhB activity acquired by the gene transfer is applicable to both D-*glycero*-D-*manno*-heptose 1,7-bisphosphate anomers, rather than to a single anomer. Nevertheless, a bioinformatics-based search for evidence of GmhB multitasking revealed that the co-occurrence of the two heptose pathways in a single bacterium is rare. Moreover, the two bacterial species that are found to contain both pathways [viz. *E. coli* 052 (*11*) and *Yersinia pseudotuberculosis* (*9*)] also contain two GmhB phosphatases, one for each pathway. Thus, the occurrence of one GmhB specialist or the other (differing in anomeric specificity) among bacterial species appears to be tied to the occurrence of the corresponding (L-*glycero*-D-*manno*-heptose 1β-ADP vs D-*glycero*-D-*manno*-heptose 1α-GDP) pathway.

REFERENCES

1. Fani, R., Brilli, M., Fondi, M., and Lio, P. (2007) The role of gene fusions in the evolution of metabolic pathways: The histidine biosynthesis case. *BMC Evol. Biol. 7* (Suppl. 2), S4.
2. Koonin, E. V., and Tatusov, R. L. (1994) Computer Analysis of Bacterial Haloacid Dehalogenases Defines a Large Superfamily of Hydrolases with Diverse Specificity: Application of an Iterative Approach to Database Search. *J. Mol. Biol. 244*, 125–132.
3. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D., and Aravind, L. (2006) Evolutionary genomics of the HAD superfamily: Understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. *J. Mol. Biol. 361*, 1003–1034.
4. Kneidinger, B., Marolda, C., Graninger, M., Zamyatina, A., McArthur, F., Kosma, P., Valvano, M. A., and Messner, P. (2002)

Biosynthesis pathway of ADP-L-*glycero*-β-D-*manno*-heptose in *Escherichia coli*. *J. Bacteriol. 184*, 363–369.
5. Kneidinger, B., Graninger, M., Puchberger, M., Kosma, P., and Messner, P. (2001) Biosynthesis of nucleotide-activated D-*glycero*-D-*manno*-heptose. *J. Biol. Chem. 276*, 20935–20944.
6. Valvano, M. A., Messner, P., and Kosma, P. (2002) Novel pathways for biosynthesis of nucleotide-activated *glycero-manno*-heptose precursors of bacterial glycoproteins and cell surface polysaccharides. *Microbiology 148*, 1979–1989.
7. Kosma, P., Wugeditsch, T., Christian, R., Zayni, S., and Messner, P. (1995) Glycan structure of a heptose-containing S-layer glycoprotein of *Bacillus thermoaerophilus*. *Glycobiology 5*, 791–796.
8. Schaffer, C., and Messner, P. (2004) Surface-layer glycoproteins: An example for the diversity of bacterial glycosylation with promising impacts on nanobiotechnology. *Glycobiology 14*, 31R–42R.
9. Bruneteau, M., and Minka, S. (2003) Lipopolysaccharides of bacterial pathogens from the genus *Yersinia*: A mini-review. *Biochimie 85*, 145–152.
10. Kondakova, A. N., Ho, N., Bystrova, O. V., Shashkov, A. S., Lindner, B., Creuzenet, C., and Knirel, Y. A. (2008) Structural studies of the O-antigens of *Yersinia pseudotuberculosis* O:2a and mutants thereof with impaired 6-deoxy-D-*manno*-heptose biosynthesis pathway. *Carbohydr. Res. 343*, 1383–1389.
11. Feng, L., Senchenkova, S. N., Yang, J., Shashkov, A. S., Tao, J., Guo, H., Cheng, J., Ren, Y., Knirel, Y. A., Reeves, P. R., and Wang, L. (2004) Synthesis of the heteropolysaccharide O antigen of *Escherichia coli* O52 requires an ABC transporter: Structural and genetic evidence. *J. Bacteriol. 186*, 4510–4519.
12. Masoud, H., Perry, M. B., Brisson, J. R., Uhrin, D., Li, J., and Richards, J. C. (2009) Structural elucidation of the novel core oligosaccharide from LPS of *Burkholderia cepacia* serogroup O4. *Glycobiology 19*, 462–471.
13. McArthur, F., Andersson, C. E., Loutet, S., Mowbray, S. L., and Valvano, M. A. (2005) Functional analysis of the *glycero-manno*-heptose 7-phosphate kinase domain from the bifunctional HldE protein, which is involved in ADP-L-*glycero*-D-*manno*-heptose biosynthesis. *J. Bacteriol. 187*, 5292–5300.
14. Raetz, C. R., Guan, Z., Ingram, B. O., Six, D. A., Song, F., Wang, X., and Zhao, J. (2009) Discovery of new biosynthetic pathways: The lipid A story. *J. Lipid Res. 50* (Suppl.), S103–S108.
15. Nguyen, H., Wang, L., Huang, H., Peisach, E., Dunaway-Mariano, D., and Allen, K. N. (2010) Structural Determinants of Substrate Recognition in the HAD Superfamily Member D-*glycero*-D-*manno*-Heptose-1,7-bisphosphate Phosphatase, GmhB. *Biochemistry* DOI 10.1021/bi902019q.
16. Chen, P. S., Toribara, T. Y., and Warner, H. (1956) Microdetermination of Phosphorus. *Anal. Chem. 28*, 1756–1758.
17. Nilsson, U., and Radstrom, P. (2001) Genetic localization and regulation of the maltose phosphorylase gene, malP, in *Lactococcus lactis*. *Microbiology 147*, 1565–1573.
18. Galkin, A., Kulakova, L., Melamud, E., Li, L., Wu, C., Mariano, P., Dunaway-Mariano, D., Nash, T. E., and Herzberg, O. (2007) Characterization, kinetics, and crystal structures of fructose-1,6-bisphosphate aldolase from the human parasite, *Giardia lamblia*. *J. Biol. Chem. 282*, 4859–4867.
19. Smyrniotis, P. Z., and Horecker, B. L. (1956) The preparation of sedoheptulose diphosphate. *J. Biol. Chem. 218*, 745–752.
20. Ames, B. N., and Horecker, B. L. (1956) The biosynthesis of histidine: Imidazoleacetol phosphate transaminase. *J. Biol. Chem. 220*, 113–128.
21. Kuznetsova, E., Proudfoot, M., Gonzalez, C. F., Brown, G., Omelchenko, M. V., Borozan, I., Carmel, L., Wolf, Y. I., Mori, H., Savchenko, A. V., Arrowsmith, C. H., Koonin, E. V., Edwards, A. M., and Yakunin, A. F. (2006) Genome-wide analysis of substrate specificities of the *Escherichia coli* haloacid dehalogenase-like phosphatase family. *J. Biol. Chem. 281*, 36149–36161.
22. Rangarajan, E. S., Proteau, A., Wagner, J., Hung, M. N., Matte, A., and Cygler, M. (2006) Structural snapshots of *Escherichia coli* histidinol phosphate phosphatase along the reaction pathway. *J. Biol. Chem. 281*, 37930–37941.
23. Allen, K. N., and Dunaway-Mariano, D. (2009) Markers of fitness in a successful enzyme superfamily. *Curr. Opin. Struct. Biol. 19*, 658–665.
24. Tremblay, L., Zhang, G., Dai, J., Dunaway-Mariano, D., and Allen, K. N. (2006) Structure and Activity Analyses of *E. coli* K-12 NagD Provide Insight Into the Evolution of Biochemical Function in the HAD Enzyme Superfamily. *Biochemistry 45*, 1183–1193.
25. Lu, Z., Dunaway-Mariano, D., and Allen, K. N. (2005) HAD superfamily phosphotransferase substrate diversification: Structure and function analysis of HAD subclass IIB sugar phosphatase BT4131. *Biochemistry 44*, 8684–8696.
26. Parsons, J. F., Lim, K., Tempczyk, A., Krajewski, W., Eisenstein, E., and Herzberg, O. (2002) From structure to function: YrbI from *Haemophilus influenzae* (HI1679) is a phosphatase. *Proteins 46*, 393–404.
27. Wu, J., and Woodard, R. W. (2003) *Escherichia coli* YrbI is 3-deoxy-D-*manno*-octulosonate 8-phosphate phosphatase. *J. Biol. Chem. 278*, 18117–18123.