Genes *malh* and *pagl* of *Clostridium acetobutylicum* ATCC 824 Encode \( \text{NAD}^{+} \)- and \( \text{Mn}^{2+} \)-dependent Phospho-\( \alpha \)-glucosidase(s)*

Received for publication, September 29, 2003, and in revised form, October 21, 2003
Published, JBC Papers in Press, October 21, 2003, DOI 10.1074/jbc.M310732300

John Thompson‡‡, Sonja Hess‡, and Andreas Pikis‡‡

From the Microbial Biochemistry and Genetics Unit, Oral Infection and Immunity Branch, NIDCR and the Proteomics and Mass Spectrometry Facility, Laboratory of Bioorganic Chemistry, NIDDK, National Institutes of Health, Bethesda, Maryland 20892

The genome of *Clostridium acetobutylicum* ATCC 824 contains two genes encoding \( \text{NAD}^{+} \)-, \( \text{Mn}^{2+} \)-, and dithiothreitol-dependent phospho-\( \alpha \)-glucosidases that can be assigned to family 4 of the glycosylhydrolase superfamily. The two genes, designated *malh* (maltose 6-phosphate hydrolase) and *pagl* (phospho-\( \alpha \)-glucosidase), respectively, reside in separate operons that also encode proteins of the phosphoenolpyruvate-dependent:sugar phosphotransferase system. *C. acetobutylicum* grows on a variety of \( \alpha \)-linked glucosides, including maltose, methyl-\( \alpha \)-D-glucoside, and the five isomers of sucrose. In the presence of the requisite cofactors, extracts of these cells readily hydrolyzed the chromogenic substrates, phenyl-\( \alpha \)-D-glucoside 4-\( \text{D-glucopyranoside} \) 6-phosphate, but whether hydrolysis reflected expression of enzymes encoded by the *malh* or *pagl* genes was not discernible by spectrophotometric analysis or polyacrylamide gel electrophoresis. Resolution of this question required the cloning of the *malh* and *pagl* genes, and subsequent construction of full-length, but catalytically inactive *MalH*. Of the two tested. Site-directed changes C169S and D170N yielded primarily maltose (4-\( \text{D-glucopyranosyl-}\alpha \)-D-glucopyranosyl 6-phosphate) and phosphate formation via the phosphoenolpyruvate-dependent phosphotransferase system. *C. acetobutylicum* has the capacity to metabolize a variety of carbohydrates, but starch, being inexpensive and readily available, is an industrially attractive source for microbial fermentation (3, 5). Prior to its metabolism by *C. acetobutylicum*, starch is first hydrolyzed by amylolytic enzymes (3) to yield primarily maltose (4-\( \text{O-}\alpha \)-D-glucopyranosyl-\( \alpha \)-D-glucopyranosyl). However, as emphasized by Tangney et al. (6), surprisingly little is known of the mechanism(s) by which this \( \alpha \)-linked disaccharide is transported into, and subsequently hydrolyzed within, the cell.

Hydrolysis of the glycosidic bond is catalyzed by a vast array of \( \text{O-glycosyl hydrolases} \) (EC 3.2.1.-), and substrate cleavage may be accompanied by either retention or inversion of configuration at the anomeric (C1) center (7, 8). An extremely helpful classification of these diverse and widespread enzymes has been developed during the past decade by Henrissat and associates (9, 10). This classification scheme is based on amino acid sequence similarity, and presently includes 91 distinct families that comprise the glycosylhydrolase superfamily of enzymes. Glycosylhydrolase family 4 (GHF4) includes 6-\( \text{phospho-}\alpha \)-glucosidases (13–15), 6-phospho-\( \beta \)-glucosidases (16), \( \alpha \)-glucosidases (17–19), and \( \alpha \)-galactosidases (20–22). The GHF4 enzymes have received considerable attention, in part, because members of this unique family are distinguished from all others by their requirements for reducing agent, divalent metal ion (\( \text{Mn}^{2+} \), \( \text{Co}^{2+} \), \( \text{Ni}^{2+} \), or \( \text{Fe}^{2+} \)) and dinucleotide (\( \text{NAD}^{+} \)) for activity. Whether these cofactors function in a catalytic or structural capacity in GHF4 enzymes has yet to be established (14, 18, 19).

Our interest in maltose dissimilation by *C. acetobutylicum* began in 1999 (see Fig. 2 of Ref. 16), by the finding that the (then partially sequenced) genome of this organism contained the organic solvents acetone and butanol (for historical reviews and current developments, see Refs. 1–3). The genome of *C. acetobutylicum* ATCC 824 has now been sequenced (4), and, in conjunction with recent advances in molecular biology and biotechnology, this knowledge has stimulated renewed interest in the genetic manipulation of the pathways and products of solven-

togenes. *C. acetobutylicum* has the capacity to metabolize a variety of carbohydrates, but starch, being inexpensive and readily available, is an industrially attractive source for microbial fermentation (3, 5). Prior to its metabolism by *C. acetobutylicum*, starch is first hydrolyzed by amylolytic enzymes (3) to yield primarily maltose (4-\( \text{O-}\alpha \)-D-glucopyranosyl-\( \alpha \)-D-glucopyranosyl). However, as emphasized by Tangney et al. (6), surprisingly little is known of the mechanism(s) by which this \( \alpha \)-linked disaccharide is transported into, and subsequently hydrolyzed within, the cell.

Hydrolysis of the glycosidic bond is catalyzed by a vast array of \( \text{O-glycosyl hydrolases} \) (EC 3.2.1.-), and substrate cleavage may be accompanied by either retention or inversion of configuration at the anomeric (C1) center (7, 8). An extremely helpful classification of these diverse and widespread enzymes has been developed during the past decade by Henrissat and associates (9, 10). This classification scheme is based on amino acid sequence similarity, and presently includes 91 distinct families that comprise the glycosylhydrolase superfamily of enzymes. Glycosylhydrolase family 4 (GHF4) includes 6-\( \text{phospho-}\alpha \)-glucosidases (13–15), 6-phospho-\( \beta \)-glucosidases (16), \( \alpha \)-glucosidases (17–19), and \( \alpha \)-galactosidases (20–22). The GHF4 enzymes have received considerable attention, in part, because members of this unique family are distinguished from all others by their requirements for reducing agent, divalent metal ion (\( \text{Mn}^{2+} \), \( \text{Co}^{2+} \), \( \text{Ni}^{2+} \), or \( \text{Fe}^{2+} \)) and dinucleotide (\( \text{NAD}^{+} \)) for activity. Whether these cofactors function in a catalytic or structural capacity in GHF4 enzymes has yet to be established (14, 18, 19).

**For almost a century, and to varying extent, *Clostridium acetobutylicum* has been used for the industrial production of**

---

1. P. M. Coutinho and B. Henrissat (1999) Carbohydrate-Active Enzymes server (afmb.cnrs-mrs.fr/CAZY/index.html); Wellcome Trust at Sanger Institute server (www.sanger.ac.uk/cgi-bin/Flamgetacat/PP02056).
2. The abbreviations used are: GHF4, glycosylhydrolase family 4; MalH, maltose 6-phosphate hydrolase; PagL, phospho-\( \alpha \)-glucosidase; pNPsGlc6P, \( \text{p-nitrophenyl-}\alpha \)-D-glucopyranoside 6-phosphate; DTT, dithiothreitol; PEP-PTS, phosphoenolpyruvate-dependent:sugar phosphotransferase system; 4MU-Ge6P, 4-methylumbelliferyl-\( \alpha \)-D-glucopyranoside 6-phosphate; ESI, electrospray ionization; MS, mass spectrometry; IPTG, isopropyl-\( \beta \)-D-thiogalactopyranoside; BSA, bovine serum albumin; G6PDH, glucose-6-phosphate dehydrogenase; BiTris, 2-[bis(2-hydroxyethyl)aminol]-2-[hydroxymethyl]propane-1,3-diol; MES, 4-morpholinethanesulfonic acid; HPLC, high performance liquid chromatography; LC, liquid chromatography.
two genes for which the deduced open reading frames exhibited extensive homology with an NAD\(^+\) and metal-dependent GHP4 phospho-\(\alpha\)-glucosidase that we had found in several bacterial species, including *Bacillus subtilis* (14), *Fusobacterium mortiferum* (13, 23, 24), and *Klebsiella pneumoniae* (15). For purposes of distinction, we refer to the two putative phospho-\(\alpha\)-glucosidases in *C. acetobutylicum* 824 as MalH (maltose-6-phosphate hydrolase) and PagL (\(\alpha\)-glucosidase), respectively. Completion of the *C. acetobutylicum* genome sequence in 2001 (4) revealed that the *malh* and *pagl* genes were located in separate operons (see Fig. 1) that also encoded components of the phosphoenolpyruvate-dependent sugar:phosphotransferase system (PEP-PTS, Refs. 25–27). Complementation studies conducted by Tangney et al. (6) established the PEP-PTS-catalyzed phosphorylation of maltose (4-\(\alpha\)-D-glucopyranosyl-D-glucopyranose) by cells of *C. acetobutylicum* grown previously on this disaccharide. Experiments in our laboratory also confirmed phospho-\(\alpha\)-glucosidase activity, not only in maltose-grown cells of *C. acetobutylicum*, but also in organisms grown previously on a variety of \(\alpha\)-D-glucosides, including maltotriol, methyl-\(\alpha\)-D-glucoside, and the five \(\alpha\)-D-glucosyl-\(\beta\)-D-fructose isomers of sucrose (28, 29) trivially designated: trehalulose \(\alpha\) (1–1), turanose \(\alpha\) (1–3), maltotriose \(\alpha\) (1–4), leucrose (1–5), and palatinose \(\alpha\) (1–6).

Although indicative of phospho-\(\alpha\)-glucosidase and PEP-PTS activities, it was unclear from the preceding observations whether the products of only one or both operons participated in the dissimilation of individual \(\alpha\)-glucosides by *C. acetobutylicum*. Two approaches were adopted in an attempt to resolve these questions. First, *malh* and *pagl* genes were cloned in high expression vectors, the enzymes were purified, and their substrate specificities were determined. As we show in this report, the similarity of the physical properties of MalH and PagL (e.g. molecular size, oligomeric structure, pl, and antibody cross-reactivity) precluded the separation and independent assay of the two enzymes in cell extracts, by either spectrophotometric or electrophoretic methods. Consequently, in a second approach, we used ESI/MS analysis of the microbial proteome to unequivocally identify the phospho-\(\alpha\)-glucosidase species induced by growth of *C. acetobutylicum* on different \(\alpha\)-glucosides.

**EXPERIMENTAL PROCEDURES**

**Materials**

The five isomers of sucrose were obtained from the following sources: maltotriol was from TCI America, leucrose was from Fluka; palatinose was from Wako Chemicals; trehalulose syrup and crystalline forms were generously provided by Sudzucker (Mannheim/Ochsenfurt, Germany), and Mitsu Sugar Co. (Kanagawa, Japan), respectively. Turanose and other high purity sugars, including glucose, sucrose (1-\(\alpha\)-D-glucopyranosyl-\(\beta\)-D-fructofuranoside), and maltose were purchased from Pfannstiel Laboratories. NAD\(^+\), NAD\(^-\) and its analogs, DEAE-TrisAcryl M anion-exchanger, Ultragels ACA54 (exclusion limit 90 kDa) and ACA44 (exclusion limit 200 kDa), p-nitrophenyl-\(\alpha\)-D- and \(\beta\)-D-glucopyranosides (pNPG/\(\beta\)GlCN), 4-methylumbelliferyl-\(\alpha\)-D-glucopyranoside (4MU Glc), and other reagents were obtained from Sigma. 5'-AMP Sepharose 4B was purchased from Amersham Biosciences. Glucose-6-phosphate dehydrogenase (G6PDH, EC 1.1.1.49) was from Roche Molecular Biochemicals. The phosphorylated chromogenic and fluorogenic substrates (pNPGluc6P and 4MUgluc6P, respectively) were prepared by selective phosphorylation of the parent compounds (at C6-GH) with phosphoglucomutase in tritium labeled phosphate containing small proportions of water (13). Phosphorylated \(\alpha\)-linked disaccharides, including sucrose 6-phosphate (sucrose-6P) and its five isomeric \(\beta\)-P derivatives were prepared as described previously (29). Phosphorylated \(\beta\)-linked disaccharides, including 4-\(\alpha\)-D-glucopyranosyl-\(\beta\)-D-glucopyranose 6-\(\beta\)-phosphate (celllobiose-6-P) and 6-\(\beta\)-D-glucopyranosyl-\(\alpha\)-D-glucopyranose 6-\(\beta\)-phosphate (gentiobiose-6-P), were prepared by phosphorylation of the parent compounds with purified ATP-dependent \(\beta\)-glucoside kinase (EC 2.7.1.85) from *K. pneumoniae* (30). Polyclonal rabbit antibody raised against purified phospho-\(\alpha\)-glucosidase (MalH) from F. mortiferum was prepared by Covance Research Products.

**Preparation of Cell-free Extracts of *C. acetobutylicum***

Washed cells were resuspended in 2.5 volumes of TMND buffer to yield a homogeneous thick suspension of cells. Lysozyme (0.6 mg) was added to 3-mI volumes of each cell suspension, and the preparations were incubated for 30 min at 37 °C. The organisms were disrupted (six 15-s periods of sonic oscillation, at 0 °C in a Branson model 185 fine probe sonifier, and the samples were clarified by centrifugation (14,000 rpm for 30 min at 5 °C) in an Eppendorf bench top centrifuge model 5417R). The supernatant fluids, containing 25–30 mg of protein ml\(^-1\), were removed for enzyme assays and for the SDS-PAGE, Western blot, and activity stain experiments presented in Fig. 6.

**Cloning of *malh* and *pagl* Genes from *C. acetobutylicum***

Based on the complete genome sequence of *C. acetobutylicum* 824 (4), the following pairs of primers were designed to amplify the genes *malh* and *pagl*, respectively. For amplification of *malh*, the forward primer was 5'-GGGGCCCATGCGATGAAAAATTTCTAGTTAATAGGCGG-3' (the *malh* sequence is in boldface, and the Ncol site is underlined) and reverse primer was 5'-CCCGGGGTACCTCAATTATAAATTTAATTCCGGCCATGAA-3' (the sequence complementary to the downstream region of *malh* is in bold face, and the PstI site is underlined). For amplification of *pagl*, the forward primer was 5'-GGGCCCATGGGATATGATTTTGAATTGTTTGCGTCAGG-3' (the *pagl* sequence is in boldface, and the Ncol site is underlined) and reverse primer was 5'-CCGGGGATCTCAATTATAAATTTAATTCCGGCCATGAA-3' (the sequence complementary to the downstream region of *pagl* is in boldface, and the EcoRI site is underlined). The *malh* and *pagl* genes were amplified by the high fidelity *Pfu* polymerase (Stratagene). The components of the amplification mixtures (100 \(\mu\)l) were as follows: 5 units of *Pfu* DNA polymerase, 1 \(\times\) reaction buffer provided by the manufacturer, 20 nm each of the four DNTPs, 250 ng of each primer, and 100 ng of DNA, amplifications were carried out in a thermal cycler (GeneAmp PCR System 9700, PE Applied Biosystems). After an initial 2-min incubation at 95 °C, the mixtures were subjected to 30 cycles of amplification under the following conditions: denaturation at 95 °C, 1 min; annealing at 50 °C, 1 min; and extension at 72 °C for 2 min/kilobase of insert. These procedures were followed by a 10-min run-off at 72 °C. The amplicons were digested with restriction endonucleases (maltose was digested with Ncol and PstI; *pagl* with Ncol and EcoRI), electrophoresed in 1% agarose, and purified (Qiagen gel extraction kit). The purified ~1.3-kb DNA fragments were ligated to the similarly digested (and purified) high expression vector *pTrcHis2B* to form the two recombinant plasmids *pmalh* and *ppagl*. *pmalh* and *ppagl* were introduced into *E. coli* and *K. pneumoniae*, respectively. In these constructs, *malh* and *pagl* are under control of the trc promoter, which is also regulated by the lacO operator and the product of the lacB gene. Expression of the products of *malh* and *pagl* are induced in the presence of isopropyl-\(\beta\)-D-thiogalactopyranoside (IPTG). Recombinant plasmids were introduced into *Escherichia coli* pEP43 by electroporation, and transformants were selected on LB agar plates supplemented with 15 \(\mu\)g/ml ampicillin. (Note that the mutant strain *E. coli* pEP43 lacks all phospho-\(\beta\)-glucosidase activities normally resident in this organism, and this strain was generously provided by Dr. B. G. Hall, Biology Department, University of Rochester, Rochester, NY.).

**Site-directed Mutagenesis of *malh***

The role of individual amino acid residues in the CDMP motif of MalH was assessed by site-directed mutagenesis of the *malh* gene. This method requires the use of *Pfu* Turbo DNA polymerase and a temperature cycler, and the appropriate reagents were obtained as the QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA).
Plasmids of pTrcHis2Bmalh containing the desired mutations (C169S, D170N, M171V, and P172A) were introduced into cells of *E. coli* PEP 43 by electroporation. Verification of the relevant base changes was performed by DNA sequence analysis. Sequencing was performed by the dideoxynucleotide chain termination method using modified T7 DNA polymerase and [α-35S]dATP labeling. The sequencing reagents were obtained as a kit (Sequenase version 2.0, U. S. Biochemical Corp., Cleveland, OH). The MacVector sequence analysis package (version 7.0, Genentech Computer Group, Madison, WI) was used to assemble, edit, and analyze the results.

**Growth of Recombinant Cells and Preparation of Extracts Containing MalH and PagL**

Cells of *E. coli* PEP 43 containing plasmids encoding either malh or pagl genes were grown on a rotary shaker at 37 °C in LB medium supplemented with ampicillin (150 μg/ml). At *A*₅₆₀ nm = 0.5 units, a solution of IPTG was added to a final concentration of 1 mM, and growth was continued for 4 h. The cells were harvested by centrifugation (15,000 × g for 10 min at 5 °C), and the organisms (−3.1 g/liter) were disrupted (at 0 °C) by two 1.5-min periods of sonic oscillation in a Branson model 350 sonifier operating at maximum power, and the cell extract was clarified by ultracentrifugation (180,000 × g for 2 h at 5 °C). The high speed supernatant fluid was dialyzed against 4 liters of TMND buffer in a cold room overnight. MalH was purified by chromatography in three stages by low pressure chromatography. Column flow rates were maintained by use of a P-1 peristaltic pump interfaced with a Frac-100 collector. Elution of proteins was monitored at 280 nm by a UV-1 optical control unit connected to a single-channel chart recorder.

**Step 1: DEAE-TrisAcryl M (Anion Exchange) Chromatography**—The dialyzed solution (−56 ml) was transferred at a flow rate of 0.6 ml/min to a column (2.6 × 14 cm) previously equilibrated with TMND buffer. Nonadsorbed materials were removed by washing with buffer, and MalH was eluted with 500 ml of a linear, increasing concentration gradient of NaCl (0–0.2 M) in TMND buffer. Fractions of 5 ml were collected, and MalH activity was detected by the intense yellow color formed upon addition of fraction samples (5 μl) to microtiter wells containing 100 μl of pNPGlc6P assay solution. Fractions with highest activity (fractions 37–44) were pooled and concentrated to 5.7 ml by filtration in an Amicon pressure unit (PM-10 membrane; 45 p.s.i.).

**Step 2: Ultrogel AcA-54 (Molecular Sieve) Chromatography**—5.4 ml of the two GHF4 phospho-α-glucosidases were aligned using the ClustalW algorithm (33). Identical amino acids are highlighted (black background), and numbers at left and right denote residue positions. The arrows above the sequences at the N terminus indicate a portion of the nucleotide-binding domain, including the conserved GlxXS motif of GHF4 proteins (14, 16, 18, 19). Arrows beneath the sequences indicate residues that constitute the signature pattern of proteins in this family. 1 Glutamyl residues that may participate in catalysis (14, 19) are indicated by filled circle.

**Fig. 2.** Alignment of the deduced amino acid sequences encoded by *malh* and *pagl* genes of *C. acetobutylicum* 824. The sequences of the two GHF4 phospho-α-glucosidases were aligned using the ClustalW algorithm (33). Identical amino acids are highlighted (black background), and numbers at left and right denote residue positions. The arrows above the sequences at the N terminus indicate a portion of the nucleotide-binding domain, including the conserved GlxXS motif of GHF4 proteins (14, 16, 18, 19). Arrows beneath the sequences indicate residues that constitute the signature pattern of proteins in this family. 1 Glutamyl residues that may participate in catalysis (14, 19) are indicated by filled circle.

**Fig. 1.** Schematic representation and molecular organization of the putative α-glucoside PEP-PTS operons *mal* and *pagl* of *C. acetobutylicum* 824. *Arrows* indicate direction of transcription. Gene designations are given above, and chromosomal gene numbers are shown within the arrows. Values in parentheses represent the number of residues encoded by the designated open reading frame. (The genome sequence is available under GenBank® accession no. AE001437. Graphical representation of the genome and detailed annotation may be found at www.genomecorp.com/programs/sequence_data_clost.shtml.)
of concentrate from step 1 was applied at 0.3 ml/min to a column of Ultrogel AcA-54 (2.8 × 94 cm), previously equilibrated with TMND buffer containing 0.1 M NaCl. Fractions of 3 ml were collected, and those fractions with greatest activity (fractions 56–62) were pooled and concentrated to 2 ml in a 10-ml Amicon filtration cell. Analysis of this preparation by SDS-PAGE revealed one major polypeptide (mass ~ 50 kDa) and trace amounts of three other polypeptides (mass ~ 70–100 kDa).

Step 3: Ultrogel AcA-44 (Molecular Sieve) Chromatography—The 2-ml concentrate from step 2 was applied (flow rate, 0.15 ml/min) to a column of Ultrogel AcA-44 (1.6 × 94 cm) equilibrated with the buffer solution used in step 2. Fractions of 2 ml were collected, and a single peak of enzyme activity was eluted at approximately the void volume of the column. These fractions (fractions 55–59) were concentrated to a volume of 3.8 ml containing 5.9 mg of protein/ml (specific activity 4.9 μmol of pNPGlucose hydrolyzed min⁻¹ mg of protein⁻¹). SDS-PAGE analysis of this preparation revealed a single polypeptide (MalH, mass ~ 50 kDa). The same procedures were used for the purification of PagL.

Electrophoresis Procedures

Native (nondenaturing) electrophoresis and SDS-PAGE were performed in the Novex X-Cell mini system, using reagents and conditions according to the instructions from the manufacturer (Invitrogen). In SDS-PAGE experiments, denatured samples, and Mark212® wide range protein standards were electrophoresed in precast NuPage (4–12%) BisTris gels with MES-SDS (pH 7.3) as the running buffer. Proteins were visualized by staining with Coomassie Brilliant Blue R-250. For Western (immuno) blots, the separated proteins and SeeBlueTM molecular weight markers were transferred to nitrocellulose membranes using NuPage transfer buffer (pH 7.2). Immunodetection of phospho-α-glucosidases (MalH or PagL) was achieved by sequential incubation of the membrane with polyclonal antibody to MalH from F. mortiferum and goat anti-rabbit horseradish peroxidase-conjugated antibody, as described previously (13). Electrophoresis of cell extracts for in situ detection of enzyme activity was carried out under native (nonreducing) conditions (at ~10 °C) in precast Triis-glycine (4–20%) gels. Triis-glycine (pH 8.3) supplemented with 1 mM MnCl₂, 0.1 mM NAD, and 1 mM dithiothreitol was used as the running buffer. For detection of MalH and PagL activities, the gel was immediately immersed in 30 ml of 25 mM Tris-HCl (pH 7.5) containing 1 mM MnCl₂, 1 mM NAD⁺, 1 mM DTT, and 0.1 mM fluoroencephor substrate 4MU-Glc6P. After 2–5 min of incubation, the gel was photographed under long wave UV light (Ektapan (Eastman Kodak Co.) film, green filter, 2-min exposure).

Enzyme Activity Assays

The activities of MalH and PagL in cell extracts, and in column fractions during enzyme purification, were measured (at 37 °C) in a discontinuous assay that contained (in 2 ml): 50 mM Tris-HCl buffer (pH 7.5), 1 mM MnCl₂, 1 mM NAD⁺, 1 mM DTT, and 0.5 mM pNPGlucose. After enzyme addition, samples of 0.25 ml were withdrawn (20-s intervals over a 2-min period) and added immediately to 0.75 ml of a solution containing 0.5 M Na₂CO₃ and 0.1 M EDTA to stop the reaction. The A₅₉₀ nm of the solution was measured, and the rates of substrate hydrolysis (pNP formation) were calculated from the progress plots, assuming a molar extinction coefficient for the (yellow) p-nitrophenoxide anion ε = 18,300 M⁻¹ cm⁻¹. One unit of enzyme activity is the amount of MalH or PagL that catalyzes the formation of 1 μmol of pNP min⁻¹ at 37 °C. A continuous spectrophotometric assay was used in kinetic analyses, and determination of substrate specificity of MalH and PagL with respect to hydrolysis of the α phospho- and β-phospho-disaccharides. This G6PDH/NAD⁺-coupled assay monitors the formation of Glc6P produced upon hydrolysis of the glucoside-6P substrate(s). The standard assay contained (in 1 ml): 0.1 M HEPES buffer (pH 7.5), 1 mM MgCl₂, 1 mM MnCl₂, 1 mM NAD⁺, 1 mM NADP⁺, 10 mM DTT, 3 units of G6PDH, 0.5–3 mM disaccharide-6P, and 50–150 μg of purified enzyme. Reactions were initiated by enzyme addition, and the increase in A₄₉₀ nm was followed in a Beckman DU 640 recording spectrophotometer. Initial rates of reaction were determined with the kinetics program of the instrument, and a molar extinction coefficient ε = 6,220 M⁻¹ cm⁻¹ was assumed for calculation of NADPH (Glc6P) formed. Kinetic parameters were determined from Eadie-Hofstee plots generated by the dogStar software (D. G. Gilbert) version 1.0.1 kinetics program.

Analytical Methods

Protein concentrations were either measured by the BCA protein assay kit (Pierce) or calculated using the theoretical molar absorption coefficient, as described previously (13). Electroacrophoresis was performed in the Novex X-Cell mini system, using reagents and conditions according to the instructions from the manufacturer (Invitrogen). In SDS-PAGE experiments, denatured samples, and Mark212® wide range protein standards were electrophoresed in precast NuPage (4–12%) BisTris gels with MES-SDS (pH 7.3) as the running buffer. Proteins were visualized by staining with Coomassie Brilliant Blue R-250. For Western (immuno) blots, the separated proteins and SeeBlueTM molecular weight markers were transferred to nitrocellulose membranes using NuPage transfer buffer (pH 7.2). Immunodetection of phospho-α-glucosidases (MalH or PagL) was achieved by sequential incubation of the membrane with polyclonal antibody to MalH from F. mortiferum and goat anti-rabbit horseradish peroxidase-conjugated antibody, as described previously (13). Electrophoresis of cell extracts for in situ detection of enzyme activity was carried out under native (nonreducing) conditions (at ~10 °C) in precast Triis-glycine (4–20%) gels. Triis-glycine (pH 8.3) supplemented with 1 mM MnCl₂, 0.1 mM NAD, and 1 mM dithiothreitol was used as the running buffer. For detection of MalH and PagL activities, the gel was immediately immersed in 30 ml of 25 mM Tris-HCl (pH 7.5) containing 1 mM MnCl₂, 1 mM NAD⁺, 1 mM DTT, and 0.1 mM fluoroencephor substrate 4MU-Glc6P. After 2–5 min of incubation, the gel was photographed under long wave UV light (Ektapan (Eastman Kodak Co.) film, green filter, 2-min exposure).

FIG. 3. Purification and Mₜ determination of the two GHF4 phospho-α-glucosidases from C. acetobutylicum: MalH (A) and PagL (B). Samples from each stage of purification were denatured, resolved by SDS-PAGE, and stained with Coomassie Brilliant Blue R-250. Numbers represent approximate molecular mass (kDa) of protein standards (Std) in the left lanes. C. Western (immuno) blot showing the highly specific cross-reactivity of MalH and PagL with polyclonal rabbit antibody prepared against phospho-α-glucosidase from F. mortiferum.
FIG. 4. Estimation of the solution-state $M_\text{r}$ of MalH and PagL by gel permeation HPLC. Details of chromatography are provided under “Experimental Procedures.” Numbers refer to protein standards used for column calibration: 1, carboxy anhydrase (29 kDa); 2, ovalbumin (44 kDa); 3, bovine serum albumin (BSA monomer, 66 kDa); 4, BSA dimer (130 kDa); and 5, aldolase (150 kDa). The column was calibrated with the N termini of MalH and PagL was obtained using an ABI 477A instrument was performed with a standard Agilent ES tuning mix. The HP1100 LC-mass selective detector was used for these studies. Data were derived from the plot of log $M_\text{r}$ versus retention time (min).

The basal assay (2 ml) contained 50 mm Tris-HCl (pH 7.5), 0.5 mm pNP-Glc6P and additions were made to the following final concentrations: 1 mm MnCl2, 1 mm NAD+, and 20 mm DTT. Assays were started by addition of either MalH (59.4 μg) or PagL (109 μg).

Activity expressed as micromoles of pNP-Glc6P hydrolyzed min$^{-1}$ mg of protein$^{-1}$.

NDH, no detectable activity.

Substrate specificity and kinetic parameters of purified phospho-α-glucosidase(s): MalH and PagL

TABLE III

| Test substrate | **$K_m$** | $V_{max}$ | **$K_m$** | $V_{max}$ |
|---------------|----------|----------|----------|----------|
| **MalH** | | | | |
| Trehalulose-6-P | 1.92 | 0.43 | NDH | |
| Sucrose-6-P | NDH | NDH | | |
| Turanose-6-P | 0.82 | 0.82 | NDH | NDH |
| Maltulose-6-P | 1.11 | 0.65 | NDH | |
| Leucrose-6-P | 1.87 | 0.12 | NDH | |
| Palatulose-6-P | 2.47 | 0.60 | NDH | |
| Maltose-6-P | 1.95 | 0.49 | NDH | |
| pNP-Glc6P | 0.028 | 5.22 | 0.065 | 4.43 |
| Cellubiose-6-P (β) | NDH | NDH | | |
| Gentibioreose-6-P (β) | NDH | NDH | | |
| Methyl-β-Glc6P | NDH | NDH | | |

Expressed as millimolar concentration.

Expressed as micromoles of substrate hydrolyzed min$^{-1}$ mg of protein$^{-1}$.

Compounds in boldface are sucrose isomers.

NDH, no detectable hydrolysis.

RESULTS

The Two Putative α-Glucoside:PTS Operons of C. acetobutylicum 824—The genetic and organizational similarities of the two putative α-glucoside:PTS operons are illustrated in Fig. 1. For convenience, they are designated the maltose (mal) and phospho-α-glucoside (pagl) operons, respectively. Genes in each operon encode proteins representing: (i) a helix-turn-helix protein from the RpiR family of transcriptional regulators, (ii) a fused EII(CB) membrane-localized transporter of the PEP-PTS, and (iii) putative phospho-α-glucosidase(s), designated maltose 6'-phosphate hydrolase (MalH) and phospho-α-glucosidase (PagL), respectively. It is of interest that the pagl operon also includes a gene that encodes an EIIA component of the PTS. A corresponding gene is not present in the mal operon.
of C. acetobutylicum, nor is a comparable EIIA gene found in any of the α-glucoside:PTS operons described recently in other species (15, 24, 32).

**Sequence Alignment of MalH and PagL from C. acetobutylicum 824**—Comparative sequence alignment reveals 50% residue identity between MalH and PagL (Fig. 2). Significantly, both proteins contain an N-terminal NAD⁺-binding domain (14, 15, 18, 19), and they exhibit the conserved motif (Glx(X)NH) and signature sequence (P-X-[SA]-(X)(L|MV|F)Y)₂-[QN]-X₂-N-P-X₁-[QR](TA)-X₁₀-[QR](KD)-X₁-[LV]-[GN]-X₁-C) that are characteristic of all GHF4 members. ¹ All family 4 phospho-α-glucosidases described thus far contain the amino acid motif CDMP characteristic of all GHF4 members. ¹ All family 4 phospho-α-glucosidases described thus far contain the amino acid motif CDMP characteristic of all GHF4 members. ¹ All family 4 phospho-α-glucosidases described thus far contain the amino acid motif CDMP characteristic of all GHF4 members. ¹ All family 4 phospho-α-glucosidases described thus far contain the amino acid motif CDMP characteristic of all GHF4 members.

**Cloning, Expression, and Purification of MalH and PagL**—Genes malh and pagl were cloned into the vector pTrcHis2B, and the recombinant plasmids were transformed into cells of *E. coli* PEP 43 for IPTG-induced expression of the two proteins. The same three-stage procedure was used for the chromatographic purification of MalH and PagL (see “Experimental Procedures”). Analysis of denatured samples of the purified proteins by SDS-PAGE, revealed a single polypeptide of ~50 kDa for both MalH and PagL (Fig. 3, A and B, respectively). The two proteins were indistinguishable by their migration positions, and both MalH and PagL cross-reacted strongly with polyclonal antibody raised against phospho-α-glucosidase (MalH) from *F. mortiferum* (Fig. 3C).

**Physical Properties of PagL and MalH**—In contrast with the molecular mass estimates obtained by SDS-PAGE, HPLC gel filtration yielded molecular mass estimates of ~200 kDa for both MalH and PagL (Fig. 4). The failure to detect species of intermediate molecular mass (~50 or ~100 kDa) suggests that, in their native solution states, both MalH and PagL exist in tetrameric form. The homogeneity of purified PagL was confirmed by microsequence analysis that provided the unambiguous determination of 28 of the first 30 residues from the N terminus of the protein: GMKKS(Y)IVGGGSRYTPDMLAM(L)EQKER. A glycine residue derived from the vector DNA precedes the start Met residue, and amino acids were not recovered in cycles 8 and 25. With the exception of the initial glycine and the two “missing” residues (Cys), the sequence obtained in agreement with that deduced by translation of the *pagl* gene. The results of ESI/MS yielded a molecular mass for PagL of 51,205 Da, which is greater than the theoretical value (51,155 Da) by the mass of a single glycyll residue. Microsequence analysis also revealed an additional glycine residue at the N terminus of MalH, but otherwise the sequence of the next 29 amino acids agreed precisely with that expected from translation of the *malh* gene: GMKKFVSIVAGGGSFETPGIVLMMLDNMDK. Consistent with these data, the mass of the protein determined experimentally by ESI/MS (50,938 Da) was greater than the theoretically predicted mass of 49,972. The mass difference was attributed to an additional Gly residue and an oxidation.

**Cofactor Requirements and Substrate Specificity of PagL and MalH**—Previous studies of proteins assigned to GHF4 (14, 16, 17, 20–22), have shown that NAD⁺, Mn²⁺ ion, and DTT are prerequisites for enzyme activity. Data presented in Table I show that both MalH and PagL are dependent upon these same cofactors for hydrolysis of the chromogenic substrate, pNP-α-glucoside. In studies of nucleotide specificity (Table II), it was found that only β-NAD⁺ (and, to a lesser degree, its deamino derivative) served as activator(s) of MalH and PagL. Interestingly, in the case of the latter enzyme, considerable activation was elicited by the guanine analog, β-NGD⁻. In the presence of saturating concentrations of cofactors (1 mM Mn²⁺, 1 mM NAD⁺, and 20 mM DTT), kinetic analyses yielded comparable parameters for the hydrolysis of pNP-α-glucoside by the two enzymes: MalH, $K_m = 28.4 ± 5.9 \mu M$ and $V_{\text{max}} = 5.2 ± 0.4 \mu M$ mol hydrolyzed min⁻¹ mg of protein⁻¹; and PagL, $K'_m = 58.3 ± 5.6 \mu M$ and $V'_{\text{max}} = 4.4 ± 0.2 \mu M$ mol hydrolyzed min⁻¹ mg of protein⁻¹. However, the two enzymes differed markedly in their capacity to hydrolyze the “natural” phosphorylated α-glucoside products of the PEP-dependent PTS (Table III). The failure of sucrose-6-P, MalH catalyzed the hydrolysis of all O-α-linked disaccharide 6-phosphates tested, including maltose-6-P and all five phosphorylated isomers of sucrose. Surprisingly, under the same conditions, PagL failed to cleave any of these compounds. None of the β-linked disaccharide 6-phosphates tested (cellobiose-6-P, gentiobiose-6-P, and methyl-β-glucoside) were substrates for either MalH or PagL.

**Requirement(s) for DTT and a Conserved Cysteine Residue for Enzyme Activity**—In the presence of the requisite NAD⁺ and Mn²⁺ cofactors, almost all GHF4 enzymes require a comparatively high concentration of DTT for optimum activity (17, 20, 21), but the role(s) of this reducing agent have yet to be...
Phospho-α-glucosidases of *C. acetobutylicum* 824

defined (for discussion, see Ref. 19). For both MalH and PagL, 5–10 mM DTT elicited maximum rates of pNPβGlc6P hydrolysis (data not shown). All GHF4 proteins contain a conserved cysteine residue centered around amino acids 160–180 of their sequences. Crystallographic data obtained during the structural determination of α-glucosidase (AgIA) from *Thermotoga maritima* (19) provide evidence that loss of activity of this GHF4 enzyme may result from oxidation of the conserved Cys-SH-174 residue to the sulfinic acid (Cys-SO2H) form. Of GHF4 enzyme may result from oxidation of the conserved Cys-169 and the adjacent residue Asp-170, as essential residues for MalH activity in *C. acetobutylicum*. Changes C169S, D170N yielded catalytically inactive proteins (Fig. 5).

Expression of Phospho-α-glucosidase Activity in Cells of *C. acetobutylicum* 824—Cells of *C. acetobutylicum* were grown on a variety of carbohydrates, including glucose, β-glucosides (cellobiose, gentiobiose), and α-glucosides (maltose, maltitol, methyl-α-glucoside, sucrose, and its five O-α-linked isomers). Cell-free extracts were prepared and assayed for phospho-α-glucosidase activity using pNPβGlc6P as substrate (Table IV). No enzyme activity was detectable in cells grown previously on glucose, sucrose, or β-glucosides. However, high levels of phospho-α-glucosidase activity were induced during growth of the organism on all other α-glucosides tested.

Analysis of Cell Extracts by PAGE Procedures—In an attempt to determine whether the phospho-α-glucosidase activity was the result of the presence of MalH or PagL, (or both enzymes), samples of the extracts were examined by three different PAGE procedures. SDS-PAGE showed that extracts containing phospho-α-glucosidase activity were induced for high level expression of a polypeptide of ~50 kDa (Fig. 6A), which also cross-reacted specifically with antibody prepared against phospho-α-glucosidase (MalH) from *F. mortiferum* (Fig. 6B). Finally, native PAGE and *in situ* staining for enzyme activity revealed a single zone of fluorescence formed by the hydrolysis of the fluorogenic substrate, 4MUβGlc6P (Fig. 6C). Although these findings confirmed the presence of phospho-α-glucosidase activity, the similarity in molecular masses of MalH and PagL (49,972 and 51,155 Da, respectively), precluded identification of the catalytically active species in the extracts. We therefore resorted to electrospray ionization-mass spectrometry (ESI/MS) in an attempt to answer these questions.

Identification of MalH in the Proteome of *C. acetobutylicum* by ESI/MS—An extract prepared from maltulose-grown cells of *C. acetobutylicum* was used in our initial ESI/MS analysis. An excellent separation of components of the microbial proteome was achieved by reverse phase (C8) chromatography using a shallow gradient of 5% acetic acid/acetonitrile. The mass selective detection spectrum revealed a prominent polypeptide at an elution time of 31.5 min (Fig. 7A), and the ESI spectrum exhibited an envelope of multiply charged molecules ranging from +30 to +72 charge states (Fig. 7B). After deconvolution (Fig. 7C), the average molecular mass of the protein (49,973 Da) was in perfect agreement with the theoretical value calculated from the amino acid composition of MalH. Subsequently, ESI/MS was used for the proteomic analysis of cells of *C. acetobutylicum* grown previously on 10 different sugars. Consistent with the enzymatic analyses (Table IV), the phospho-α-glucosidase was not detectable in the proteome of glucose or sucrose-grown organisms (Table V). However, extracts from cells grown previously on all α-glucosides, including the five separate isomers of sucrose, contained MalH. (It may be noted parenthetically that further ESI/MS studies also provided evidence for expression of PagL during growth of *C. acetobutylicum* on maltose and methyl-α-D-glucoside. The significance of these findings is presently unclear.)

DISCUSSION

The aims of our investigation were twofold: first, to determine the physical and enzymatic properties of the proteins encoded by the *malh* and *pagl* genes; and second, to study the expression of the two gene products during growth of *C. acetobutylicum* on maltose and related α-linked glucosides.

(i) Comparative Properties of MalH and PagL from *C. acetobutylicum*—Single genes encoding NAD⁺ and metal-dependent phospho-α-glucosidase are present in *B. subtilis* (glcA; Refs. 14 and 32), *F. mortiferum* (malH; Refs. 23 and 24), and *K. pneumoniae* (aglB; Ref. 15). However, to our knowledge this is the first report of a microbial genome that encodes genes (*malh* and *pagl*) for two GHF4 phospho-α-glucosidases. The isoforms exhibit 50% residue identity, and in solution both MalH and PagL are tetramers consisting of similarly sized subunits (~50 kDa).

![Fig. 6. Expression of GHF4 phospho-α-glucosidase(s) in cells of *C. acetobutylicum* 824 during growth on different sugars.](https://example.com/fig6.png)
The two phospho-α-glucosidases of *C. acetobutylicum* are antigenically similar, and both proteins cross-react strongly with polyclonal rabbit antibody raised against phospho-α-glucosidase (MalH) from *F. mortiferum*. In addition to their specific requirements for NAD⁺ and Mn²⁺ ion, almost all GHF4 enzymes, including MalH and PagL, are dependent upon relatively high concentrations of reducing agent (DTT, β-mercaptoethanol) for activation or optimum activity. In all GHF4 enzymes, a cysteine residue is conserved at amino acid positions 160–180 (14). Lodge et al. (19) suggest that this critical residue (Cys-174) may participate in substrate cleavage by α-glucosidase (AglA) from the hyperthermophilic bacterium, *T. maritima*. An electron density map of the crystalline but inactive form of GHF4 α-glucosidase, provides evidence for the oxidation of Cys-174 to cysteine sulfinic acid (Cys-SO₂H). This observation prompted Lodge et al. (19) to suggest that DTT (or β-mercaptoethanol) may be necessary to maintain Cys-174 in a catalytically active reduced (-SH) state.

In MalH of *C. acetobutylicum*, the positionally equivalent residue Cys-169 and the adjacent amino acid Asp-170 are essential for pNP-Glc₆P hydrolysis and the site-directed changes C₁⁶⁹S and D₁⁷⁰N yield catalytically inactive proteins (Fig. 5). The first crystal structure of a GHF4 phospho-α-glucosidase (GlvA from *B. subtilis* (Ref. 14)) has recently been determined to a resolution of 2.05 Å. This elegant study shows that Cys-171 of GlvA is co-ordinately linked to Mn²⁺, whereas Asp-172 hydrogen-bonds with the 2'-OH of the Glc₆P moiety of the substrate (maltose-6'P). The results of our site-directed mutagenesis of the corresponding residues Cys-169 and Asp-170 in MalH of *C. acetobutylicum* confirm the catalytic importance of these two amino acids. It is of interest to note that, in all members of the phospho-α-glucosidase subgroup, the critical Cys residue is present in the motif CDMP. However, in the phospho-β-glucosidase subgroup of GHF4, the amino acids

---

### Table V

| Growth substrate | MalH (*M*ᵣ = 49,973) | PagL (*M*ᵣ = 51,156) |
|------------------|------------------------|----------------------|
| Glucose          | −                      | −                    |
| Methyl-α-D-glucoside | +                      | +                    |
| Malose           | −                      | +                    |
| Maltitol         | −                      | +                    |
| Trehalulose*     | −                      | +                    |
| Sucrose          | −                      | +                    |
| Turanose         | +                      | −                    |
| Maltulose        | +                      | −                    |
| Leucrose         | +                      | −                    |
| Palatinose       | +                      | −                    |

* Symbols + and − indicate presence or absence of enzyme.

* Compounds in boldface are isomers of sucrose.

---

3 S. S. Rajan, X. Yang, F. R. Collart, L. Y. Yip, S. G. Withers, A. Varrot, J. Thompson, G. J. Davies, and W. F. Anderson, submitted for publication.

4 S. G. Withers, personal communication.
CN(VI)DP are invariably present. It remains to be determined whether or not these Cys-containing motifs also play a role in discrimination of phosphoglucomutases with respect to their α- and β-linked substrates.

(ii) Expression and Substrate Specificity of MalH and PagL.—In light of their many physical similarities, it was not unexpected to find that MalH and PagL exhibited comparable kinetic parameters with respect to hydrolysis of pNPaGlc6P. It was, however, surprising to discover that the two enzymes differed markedly in their ability to cleave “natural” phospho-α-glucosides produced via the PEP-PTS. MalH was expressed during growth of C. acetobutylicum on all α-glucosides except sucrose (Table V) and, in the presence of the requisite cofactors, the purified enzyme hydrolyzed all phospho-α-glucosides tested except sucrose-6P (Table III). From these data, it is reasonable to suggest that proteins encoded by the mal operon (EII(CB)mal and MalH) promote the vectorial phosphorylation, and intracellular cleavage of α-glucosides by cells of C. acetobutylicum. As described previously for the homologous α-glucoside-PTS operons in B. subtilis (32), F. mortiferum, and K. pneumoniae (24), the mal operon of C. acetobutylicum lacks an EIIA gene for which the polypeptide product is essential for PTS function. In B. subtilis, neither the sucrose-PTS (35) nor trehalose-PTS (36) operons contain the necessary EIIAα or EIIAα genes. For these two PTSs it is believed that the constitutive EIIAglc can substitute for the “missing” components. We propose that a similar complementation between EIIAglc and EII(CB)mal allows the PTS-mediated transport and phosphorylation of α-glucosides in C. acetobutylicum.

Metabolism of Sucrose Isomers—C. acetobutylicum has the capacity to metabolize a variety of carbohydrates (3, 5) including sucrose (37), as energy sources for growth. However, to our knowledge, this is the first report of the fermentation of the five isomers of sucrose by this Gram-positive anaerobic organism.

Growth of C. acetobutylicum on the α-D-glucosyl-α-fructose elicits expression of proteins encoded by the mal operon, but sucrose itself is not an inducer. The utilization of sucrose by the oral streptococci, Streptococcus mutans and Streptococcus sobrinus, is a major factor in the etiology of dental caries (38, 39). Remarkably, oral streptococci are unable to metabolize the isomers of sucrose (34, 40, 41). Three of the five α-D-glucosyl-α-fructoses (palatinose, trehalulose, and leucrose) are now produced on the industrial scale, and, because they are noncariogenic and ~50% as sweet as sucrose, these isomers attract attention as potential substitutes for dietary sucrose (11, 12). Until recently, it had been assumed that bacteria in general lacked the capacity to utilize the isomers of sucrose. However, studies in our laboratory have shown that several species of bacteria from both Gram-positive (B. subtilis) and Gram-negative genera (F. mortiferum and K. pneumoniae) readily metabolize these unusual disaccharides. Furthermore, the PTS transport proteins and the phospho-α-glucosidase(s) of these species exhibit 65–85% residue identity with the corresponding proteins encoded by the mal operon of C. acetobutylicum (6, 24). The finding that C. acetobutylicum also grows on the five disaccharides suggests that the dissimilation of sucrose isomers by microorganisms may be more widespread than presently envisaged.
Genes \textit{malh} and \textit{pagl} of \textit{Clostridium acetobutylicum} ATCC 824 Encode NAD$^+$- and Mn$^{2+}$-dependent Phospho-\(\alpha\)-glucosidase(s)

John Thompson, Sonja Hess and Andreas Pikis

\textit{J. Biol. Chem.} 2004, 279:1553-1561. doi: 10.1074/jbc.M310733200 originally published online October 21, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M310733200

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

This article cites 38 references, 13 of which can be accessed free at http://www.jbc.org/content/279/2/1553.full.html#ref-list-1