Mannose Foraging by Bacteroides thetaiotaomicron
STRUCTURE AND SPECIFICITY OF THE β-MANNOSIDASE, BtMan2A

The human colonic bacterium Bacteroides thetaiotaomicron, which plays an important role in maintaining human health, produces an extensive array of exo-acting glycoside hydrolases (GH), including 32 family GH2 glycoside hydrolases. Although it is likely that these enzymes enable the organism to utilize dietary and host glycans as major nutrient sources, the biochemical properties of these GH2 glycoside hydrolases are currently unclear. Here we report the biochemical properties and crystal structure of the GH2 B. thetaiotaomicron enzyme BtMan2A. Kinetic analysis demonstrates that BtMan2A is a β-mannosidase in which substrate binding energy is provided principally by the glycone binding site, whereas aglycone recognition is highly plastic. The three-dimensional structure, determined to a resolution of 1.7 Å, reveals a five-domain structure that is globally similar to the Escherichia coli LacZ β-galactosidase. The catalytic center is housed mainly within a (β/α)8 barrel although the N-terminal domain also contributes to the active site topology. The nature of the substrate-binding residues is quite distinct from other GH2 enzymes of known structure, instead they are similar to other clan GH-A enzymes specific for manno-configured substrates. Mutagenesis studies, informed by the crystal structure, identified a WDW motif in the N-terminal domain that makes a significant contribution to catalytic activity. The observation that this motif is invariant in GH2 mannosidases points to a generic role for these residues in this enzyme class. The identification of GH-A clan and GH2 specific residues in the active site of BtMan2A explains why this enzyme is able to harness substrate binding at the proximal glycone binding site more efficiently than mannan-hydrolyzing glycoside hydrolases in related enzyme families. The catalytic properties of BtMan2A are consistent with the flexible nutrient acquisition displayed by the colonic bacterium.

Sugar polymers represent the most abundant source of organic carbon in the biosphere (1). Whereas plant structural polysaccharides are important nutrients for organisms living in terrestrial and marine environments, symbiotic prokaryotes that inhabit the human large bowel utilize both dietary and host saccharides as valuable sources of energy and carbon (2). Commensurate with this flexibility in nutrient acquisition, the genomes of colonic bacteria, exemplified by the human symbiont Bacteroides thetaiotaomicron, have revealed a startling array of genes encoding glycoside hydrolases (3). Indeed, B. thetaiotaomicron expresses >200 glycoside hydrolases, more than any other prokaryote described to date (3). Significantly, in B. thetaiotaomicron, the vast majority of these enzymes are in glycoside hydrolase families (GHs) (4, 5) that contain mainly exo-acting enzymes, which may suggest that the organism is able to utilize the highly complex saccharide decorations appended to the backbone of plant structural polysaccharides, and host proteins presented on the surface of the intestinal mucosa. This expansion in exo-GHs is particularly evident in families 2 and 43, in both of which B. thetaiotaomicron has more than 30 members (3).

GH2 contains over 700 sequences encoding enzymes with a wide spectrum of different exo-acting β-glycosidase activities including β-galactosidase (EC 3.2.1.23), β-mannosidase (EC 3.2.1.25), β-glucuronidase (EC 3.2.1.31), and exo-β-glucosaminidase (EC 3.2.1.-). Thus far, three-dimensional structures have only been reported for β-galactosidases, notably the Escherichia coli LacZ protein (7) (but recently also the Arthrobacter sp. C2-2 LacZ homolog (8)) and for the β-glucuronidase from Homo sapiens (9). All these three-dimensional structures reveal a catalytic (β/α)8 barrel, typical of the wider clan GH-A glycosidases (10), in which the catalytic acid/base and nucleophile involved in catalysis with net retention of anomeric configuration lie on strands β-4 and β-7, respectively, of the (β/α)8 barrel. The vast majority of GH2 sequences are extended, reflecting a complex modular architecture in which a number of non-catalytic domains are appended to the (β/α)8 barrel catalytic module. Thus far, there have been no structural reports of classical β-mannosidases from family GH2, although the structure of a GH5 exo-β-mannanase, which targets manno-oligosaccharides has been reported by Dias and colleagues in 2004 (11) and the structures of several GH5 and GH26 endo-β-1,4-mannanases have been described (12–17).

The substrate specificities of B. thetaiotaomicron GH2 enzymes, exemplified by its β-mannosidases, are particularly

3 The abbreviations used are: GH, glycosidase hydrolase; DNP, 2,4-dinitrophenol; PNP, 4-nitrophenolate; CAPSO, 3-[(3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonic acid; Bicine, N,N-bis(2-hydroxyethyl)glycine; CAPS, 3-(cyclohexylamino)propanesulfonic acid; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)-1-propan-1-ol; BtMan2A, B. thetaiotaomicron Man2A; CmMan5A, C. mixtus Man5A.

2 To whom correspondence should be addressed. E-mail: h.j.gilbert@ncl.ac.uk.

1 These authors contributed equally to the results of this work.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The on-line version of this article (available at http://www.jbc.org) contains supplemental Table S1.

‡1

‡ The on-line version of this article (available at http://www.jbc.org) contains supplemental Table S1.

1 These authors contributed equally to the results of this work.

2 To whom correspondence should be addressed. E-mail: h.j.gilbert@ncl.ac.uk.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
Mannose Foraging by \( B. \) thetaiotaomicron

To start to unravel, both structurally and biochemically, the complex glycomicrobiome of \( B. \) thetaiotaomicron we report here the cloning, expression, and enzymatic characterization of a GH2 \( \beta\)-mannosidase from \( B. \) thetaiotaomicron, hereafter \( Bt\)Man2A. The enzyme is shown to be a classical \( \beta\)-mannosidase, with high activity on mannobiose and aryl \( \beta\)-mannosides (in contrast to \( \text{exo-}\beta\)-mannanasases (11)), although it is unable to accommodate the galactosyl decorations present on plant-derived mannans. The enzyme is also able to remove \( \beta\)-mannosides from \( \text{Man-}\beta1,4\text{-GlcNAc} \), found in the core pentasaccharide of human \( N\)-glycans, consistent with the diversity of mannose-containing polymers presented to the bacterium in the large bowel. The three-dimensional structure of the enzyme at a resolution of 1.7 Å reveals a multidomain enzyme in which the active site pocket is housed within the TIM-barrel domain. Site-directed mutagenesis in conjunction with sequence comparisons of GH2 enzymes reveals active site residues that confer specificity for terminal mannoses.

MATERIALS AND METHODS

Cloning and Expression—Sequence comparisons indicated that open reading frame BT0458 encodes a GH2 mannosidase, defined hereafter as \( Bt\)Man2A. Using primers detailed under supplementary materials the region of the gene encoding mature \( Bt\)Man2A (residues 26–839) was amplified from the \( B. \) thetaiotaomicron genomic DNA by PCR. The DNA generated was restricted with Ndel and XhoI (sites introduced in the primers employed in the PCR) and cloned into appropriately digested pET28a (Novagen) to generate pLT001. \( Bt\)Man2A, encoded by pLT001, contains a C-terminal His\(_x\) tag. To produce recombinant \( Bt\)Man2A, \( E. \) coli strain BL21 (Novagen) containing pLT001 was cultured to mid-exponential phase at 37 °C in Luria-Bertani media (LB). The culture was then cooled to 16 °C and recombinant enzyme expression was induced by the addition of isopropyl \( \beta\)-D-thiogalactopyranoside to a final concentration of 100 \( \mu\)M followed by incubation for a further 16 h at this lower temperature. The mannosidase was then purified by immobilized metal ion affinity chromatography, using TALONT\textsuperscript{TM} as the column matrix, ion exchange chromatography using a Q12 column, and size exclusion chromatography employing an S200 column (as described in, for example, in Ref. 11).

To produce seleno-\( l\)-methionine \( Bt\)Man2A the \( E. \) coli methionine auxotroph B834(DE3) containing pLT001 was cultured as described following a similar protocol to that described by Charnock et al. (26) using the induction regime described above. The recombinant protein was purified as described above except all buffers were supplemented with 5 mM \( \beta\)-mercaptoethanol. All proteins were homogenous, as judged by SDS-PAGE, and matrix-assisted laser desorption ionization time-of-flight mass spectrometric analysis confirmed the incorporation of seleno-\( l\)-methionine into \( Bt\)Man2A and that the enzyme had not been processed.

Construction of Site-directed Variants—Site-directed mutagenesis, employing a QuikChange kit (Stratagene), primers listed in supplementary information, and pLT001 as the template DNA, was used to construct mutants of \( \text{man-}\)2A, encoding active site variants of \( Bt\)Man2A. All mutants were sequenced using T7

![Figure 1. Mannose foraging by B. thetaiotaomicron](image-url)
Mannose Foraging by B. thetaiotaomicron

forward and reverse and custom made primers to ensure that only the designed mutation had been introduced into the mannosidase gene.

Kinetic Analyses — The activity of BtMan2A was determined at 37 °C in 43.5 mM sodium phosphate, 10 mM citric acid (PC) buffer, pH 5.6, containing 1 mg/ml bovine serum albumin and the appropriate amount of the substrate. Appropriate concentrations of BtMan2A were added to start the reaction, which was monitored at 400 nm during the linear initial phase (<5% of substrate hydrolyzed). The release of 2,4-dinitrophenolate (DNP) from 2,4-dinitrophenyl-Man was monitored continuously. When using the substrate 4-nitrophenyl-β-D-mannopyranoside (PNP-Man) as the substrate, 4-nitrophenolate release (PNP) was measured discontinuously. At least four 200-μl aliquots were removed and the reaction was stopped by the addition of 800 μl of 1 M sodium carbonate that increased the pH to 12. The concentration of DNP and PNP produced was determined using a molar extinction coefficient of 15,000 and 10,000, respectively. For kinetic analysis at least six substrate concentrations that straddled the $K_m$ were employed and non-linear regression was used to estimate $K_m$ and $V_{max}$ of $Bt$ for mannosidase. The activity of BtMan2A against manno-oligosaccharides and mannose-containing polysaccharides (supplied by Megazyme International, Bray, County Wicklow, Ireland, except Man-β1,4-GlcNAc, which was provided by Dextra Laboratories, Reading, UK) was determined as described by Hogg et al. (14) except the reaction was carried out in PC buffer, pH 5.6. The rate of substrate hydrolysis was monitored by Dionex high performance anion-exchange chromatography (27) at a substrate concentration $<K_m$ (30 μM) enabling $K_m$ to be determined by measuring the rate of substrate hydrolysis. The $K_m$ for mannan hydrolysis was measured using a substrate concentration of 2 mg/ml. To determine the pH optimum of BtMan2A a discontinuous assay was used and 30 μM PNP-Man as substrate and the following buffers: 50 mM sodium citrate, pH 3.3–6.5; 50 mM sodium phosphate, pH 6.0–7.5; 50 mM Bicine, pH 6.5–9.0; 50 mM CAPSO, pH 8.0–10.0; 50 mM CAPS, pH 9.5–11. Temperature and proteinase stability was determined as described by Andrews et al. (28).

The $K_i$ for bis-Tris propane and isofagomine lactam was determined by measuring the $K_m$ for PNP-Man at three different concentrations of inhibitor. The enzyme was preincubated with the inhibitor for 1 h prior to carrying out the enzyme assays.

Crystallization and Data Collection — Pure proteins as judged by SDS-PAGE were concentrated to between 10 and 20 mg/ml and buffer exchanged into water (Sigma) using a Vivaspin 10-kDa cut-off concentrator. The protein was screened using the sitting drop vapor diffusion method together with the Hampton Crystal screen, Crystal screen 2, Hampton PEG/Ion (Hampton Research, Aliso Viejo, CA), the CSS Crystal screens I and II, and the PACT premier screen. The 100-nl droplets were formed by the mosquito® (TTP LabTech Ltd, Royston, Herts, UK) liquid handling robot, mixing 50 nl of protein with 50 nl of well solution. Initial crystals were obtained in the PACT premier screen conditions F2 and PEG/Ion screen G5. These conditions were optimized further to improve crystal quality resulting in conditions of either 0.15–0.35 M NaBr and 6–12% PEG 3350, 0.1 M bis-Tris propane, pH 7, or 0.08–0.2 M KSCN, and 8–15% PEG 3350,0.1 M bis-Tris propane, pH 7. A cryo-protectant solution was produced by supplementing the mother liquor with an additional 30% (v/v) glycerol. The lath-shaped crystals were harvested in rayon fiber loops then bathed in cryo-protectant solution prior to flash freezing in liquid N₂. Selenomethionine crystals were produced in the same manner.

Data were collected at the European Synchrotron Radiation Facility (ESRF) from single crystals at 100 K with a $\Delta \phi$ of 0.5°. Native data were collected at a wavelength of 0.97930 Å over an oscillation range of 200° on ID14-4 using an ADSC Q4R charged-coupled device detector. Selenium derivative data were collected at a wavelength of 0.9795 Å over an oscillation range of 200° with $\Delta \phi$ of 0.5° on ID23 using a Marmoset 225 charged-coupled device detector. The appropriate wavelength for data collection at the Se edge, optimized for the $f^*$ component of the anomalous scattering, was determined using a fluorescence scan.

Structure Solution and Refinement — All data were indexed and integrated in MOSFLM (29). All other computing was undertaken using the CCP4 suite unless otherwise stated. Native BtMan2A crystals were found to belong to the space group P2₁ with the approximate cell dimensions of $a = 91.5$ Å, $b = 116.0$ Å, $c = 99.2$ Å, and $β = 113.4°$ with two molecules in the asymmetric unit. The structure was solved using single-wavelength anomalous dispersion methods. Selenium positions were determined using SHELDX (30) and phases subsequently calculated with MLPHARE from the CCP4 suite (31). Solvent flattening and phase improvement were carried out in DM (32) also from the CCP4 suite. 5% of the data were set aside for cross-validation analysis and the behavior of $R_{free}$ was used to monitor and guide the refinement protocols. ARP/wARP (33), in conjunction with REFMAC (34), was used to automatically build the sequence into the electron density. Refinement was undertaken in REFMAC with manual correction to the model using COOT (35). Coordinates and observed structure factor amplitudes have been deposited with the Protein Data Bank. Figures were drawn with MOLSCRIPT (36) and BOBSCRIPT (37).

RESULTS

Selection and Expression of a GH2 β-Mannosidase — To identify a likely B. thetaiotaomicron GH2 β-mannosidase, the sequences of known, characterized β-mannosidases were used for BLAST searches of the UniProt sequence data base. Open reading frame BT0458 (hereafter designated man2A), one of the 32 family GH2 sequences in this organism (identified by the CAZy data base Refs. 4 and 5), displayed significant sequence similarity with both characterized and putative GH2 β-mannosidases, exhibiting 34 and 36% identity with the Thermotoga maritima and Thermatoga neopolitana β-mannosidases, respectively.

The man2A gene comprises 2592 bp encoding a 864-amino acid protein (BtMan2A) with a $M_r$ of 96,922 in which the 26 N-terminal residues are predicted to comprise a cleavable signal peptide. The man2A gene encoding mature BtMan2A was successfully expressed in E. coli, and the protein was purified by immobilized metal ion affinity chromatography to electrophoretic homogeneity (data not shown). The protein had a $M_r$,
Mannose Foraging by B. thetaiotaomicron

**TABLE 1**

Kinetic properties of wild type and mutants of BtMan2A

| BtMan2A | Substrate | $K_m$ | $k_{cat}$ | $k_{cat} / K_m$ |
|---------|-----------|-------|----------|-----------------|
| Wild type | DNP-Man | $8.8 \times 10^{-5}$ | 12.867 | $1.5 \times 10^{-1}$ |
| Wild type | PNP-Man | $8.3 \times 10^{-5}$ | 7.689 | $4.0 \times 10^{-1}$ |
| Wild type | Man-$\beta$-1,4-GlcNAc | $8.3 \times 10^{-5}$ | 5.2 | $4.0 \times 10^{-1}$ |
| Wild type | Mannobiose | $8.3 \times 10^{-5}$ | 6.0 | $4.0 \times 10^{-1}$ |
| Wild type | Mannotriose | $8.3 \times 10^{-5}$ | 1.3 | $4.0 \times 10^{-1}$ |
| Wild type | Mannopentaose | $8.3 \times 10^{-5}$ | 1.9 | $4.0 \times 10^{-1}$ |
| Wild type | Gal-mannobiose$^a$ | $8.3 \times 10^{-5}$ | 5.1 | $4.0 \times 10^{-1}$ |
| Wild type | Gal-mannotriose$^a$ | $8.3 \times 10^{-5}$ | 9.3 | $4.0 \times 10^{-1}$ |
| Wild type | $\beta$-Mannan$^b$ | $8.3 \times 10^{-5}$ | 2.8 | $4.0 \times 10^{-1}$ |
| E462A | PNP-Man | $8.3 \times 10^{-5}$ | 5.47 | $4.0 \times 10^{-1}$ |
| E555A | PNP-Man | $8.3 \times 10^{-5}$ | 0.508 | $4.0 \times 10^{-1}$ |
| E462A | DNP-Man | $8.3 \times 10^{-5}$ | 0.048 | $4.0 \times 10^{-1}$ |
| E555A | DNP-Man | $8.3 \times 10^{-5}$ | 0.011 | $4.0 \times 10^{-1}$ |
| N461A | DNP-Man | $8.3 \times 10^{-5}$ | 51 | $4.0 \times 10^{-1}$ |
| Q546A | PNP-Man | $8.3 \times 10^{-5}$ | 2066 | $4.0 \times 10^{-1}$ |
| YS57A | PNP-Man | $8.3 \times 10^{-5}$ | 1611 | $4.0 \times 10^{-1}$ |
| W198A | PNP-Man | $8.3 \times 10^{-5}$ | 1223 | $4.0 \times 10^{-1}$ |
| N461A | PNP-Man | $8.3 \times 10^{-5}$ | 14.98 | $4.0 \times 10^{-1}$ |
| W395A | PNP-Man | $8.3 \times 10^{-5}$ | 53.18 | $4.0 \times 10^{-1}$ |
| W200A | PNP-Man | $8.3 \times 10^{-5}$ | 1.2 | $4.0 \times 10^{-1}$ |
| W645A | PNP-Man | $8.3 \times 10^{-5}$ | 2.8 | $4.0 \times 10^{-1}$ |
| D199A | PNP-Man | $8.3 \times 10^{-5}$ | 2.72 | $4.0 \times 10^{-1}$ |

$^a$ Gal-mannobiose, $^b$-$\alpha$-D-galactosyl-mannobiose; mannotriose, $^c$-$\alpha$-D-galactosyl-mannotriose.

The activity of BtMan2A against the polysaccharides was evaluated using 0.5% substrate and enzyme at a concentration of 40 nM.

![Kinetics of aryl-\(\beta\)-mannoside hydrolysis by wild type and mutants of BtMan2A.](image)

The enzyme displays no activity against \(\alpha\)-D configured glycosides or any dyed polysaccharides. The data were entirely consistent with the designation of BtMan2A as a \(\beta\)-mannosidase displaying no endo-activity. The kinetic parameters of BtMan2A (Table 1 and Fig. 2) show that the enzyme is \(~100,000\)-fold more active against manno- compared with gluco-configured substrates. Analysis of the reaction products released by BtMan2A from a range of \(\beta\)-linked manno-oligosaccharides showed that mannose and a manno-oligosaccharide with a degree of polymerization of \(n-1\), where \(n\) is the degree of polymerization of the substrate, were the initial reaction products demonstrating that the enzyme displays classical \(exo\)-activity, releasing mannose from the non-reducing end of polymeric substrates. Even at a manno-oligosaccharide concentration of 1 mM BtMan2A produced equal amounts of mannose and the \(n-1\) product; no oligosaccharides larger than the substrate were produced. Thus, under the conditions used, BtMan2A does not display significant transglycosylating activity, in contrast to many retaining glycosidases (extensively reviewed in Ref. (38)). This may indicate that the +1 subsite (39) does not bind mannose tightly and thus an activated water is more likely to be generated in the

determined by matrix-assisted laser desorption ionization time-of-flight mass spectrometric analysis, consistent with the predicted size of the processed enzyme (data not shown) indicating that the enzyme was not proteolytically degraded during its expression or subsequent purification.

Biochemical Properties of BtMan2A—BtMan2A was screened for catalytic activity using both a series of aryl-glycosides and dyed polysaccharides. The results showed that the enzyme displays no activity against xylose, arabinose, or galactose containing substrates, trace activity against 4-nitrophenyl-\(\beta\)-D-glucopyranoside but significant activity against PNP-Man. The enzyme displays no activity against \(\alpha\)-D configured glycosides or any dyed polysaccharides. The data were entirely consistent with the designation of BtMan2A as a \(\beta\)-mannosidase displaying no endo-activity. The kinetic parameters of BtMan2A (Table 1 and Fig. 2) show that the enzyme is \(~100,000\)-fold more active against manno- compared with gluco-configured substrates. Analysis of the reaction products released by BtMan2A from a range of \(\beta\)-linked manno-oligosaccharides showed that mannose and a manno-oligosaccharide with a degree of polymerization of \(n-1\), where \(n\) is the degree of polymerization of the substrate, were the initial reaction products demonstrating that the enzyme displays classical \(exo\)-activity, releasing mannose from the non-reducing end of polymeric substrates. Even at a manno-oligosaccharide concentration of 1 mM BtMan2A produced equal amounts of mannose and the \(n-1\) product; no oligosaccharides larger than the substrate were produced. Thus, under the conditions used, BtMan2A does not display significant transglycosylating activity, in contrast to many retaining glycosidases (extensively reviewed in Ref. (38)). This may indicate that the +1 subsite (39) does not bind mannose tightly and thus an activated water is more likely to be generated in the
Mannose Foraging by B. thetaiotaomicron

proximal glycone region of the active site. The enzyme displays extremely low activity against 6'-α-D-galactosyl-mannobiose and 6'-α-D-galactosyl-mannotriose (~40-fold lower activity than against mannobiose and mannotriose, respectively), indicating that decoration of the aglycone region of the substrate located in the +2 or +1 subsite can be tolerated, albeit poorly whereas no detectable hydrolysis of the doubly substituted 6',6'-α-D-galactosyl-mannopentaose (Fig. 3 and Table 1) was evident. The specificity of BtMan2A for linear unsubstituted mannans (Fig. 1) and its inability to exhibit any endo-activity explains why the enzyme is only able to produce small amounts of mannose from galactomannan or glucomannan (data not shown).

Analysis of the biophysical properties of BtMan2A showed that the enzyme has a pH optimum of 5.6, typical of many GH2 enzymes, is inactivated at temperatures in excess of 58 °C and is completely resistant to proteolytic attack. This latter feature is typical of extracellular glycoside hydrolases (40), and is consistent with the highly proteolytic environment of the large bowel where numerous anaerobic bacteria secrete a range of different proteinases.

Kinetic analysis of manno-oligosaccharide hydrolysis showed that BtMan2A hydrolyzed mannotriose ~20-fold more efficiently than mannobiose, whereas the activity of the enzyme against mannotetraose and mannopentaose was similar to the trisaccharide (Fig. 3 and Table 1). These data indicate that BtMan2A has one glycone (~1) and two aglycone (+1 and +2) subsites that can accommodate β1,4-linked mannose residues. BtMan2A also hydrolyzes Man-β1,4-GlcNAc (kcat/Km 5.2 × 10⁴ min⁻¹ M⁻¹) at a rate similar to mannobiose (kcat/Km 6.0 × 10⁴ min⁻¹ M⁻¹), suggesting that the +1 subsite displays considerable plasticity in substrate specificity (Table 1 and Fig. 3). The biochemical properties of BtMan2A against aryl-β-glycosides are similar to the corresponding Cellulomonas fimii enzyme (41, 42), although the Bacteroides glycoside hydrolase does not display the substrate inhibition observed in the Cellulomonas β-mannosidase.

The much higher activity (~500-fold) of BtMan2A for aryl-mannosides compared with disaccharides substrates reflects both the better leaving group abilities of PNP and DNP and perhaps also that the +1 subsite binds sugar molecules weakly and nonspecifically. The +2 subsite contributes a modest 1.5 kcal/mol to productive substrate binding. By contrast the Cellulibrio mixtus exo-mannanase, CmMan5A, is ~30,000-fold less active than BtMan2A against PNP-Man highlighting the pivotal role of the glycone (~1) binding site in the Bacteroides enzyme, whereas the C. mixtus glycoside hydrolase derives much of its catalytic power from substrate binding to the aglycone subsites (11). Indeed, in considerable contrast to BtMan2A the GH5 enzyme is 100- and 10,000-fold more active against the β1,4-linked manno-oligosaccharides mannosidase and mannotriose, respectively, than PNP-Man. The importance of the ~1 subsite in BtMan2A is further emphasized by its inhibition, relative to CmMan5A, by monosaccharide inhibitors. Isofagomine lactam, displays a Ks of 400 μM on CmMan5A (43) but binding is ~36 times stronger to BtMan2A with a Ks value of 11 μM, essentially identical for the Ks of 9 μM reported for the classical snail β-mannosidase (44).
Mannose Foraging by B. thetaiotaomicron

Structure of BtMan2A—The structure of BtMan2A was solved using selenomethionine single wavelength anomalous dispersion data to 2.8 Å in harness with native data to 1.7 Å (see experimental statistics in Table 2). The P2₁ crystal form contains a dimer, consistent with dynamic light scattering experiments and size exclusion chromatography (not shown) indicating that the protein is also a dimer in solution. BtMan2A consists of five clear domains (Fig. 4A) with the catalytic center the central of these. Such a modular organization is similar to that seen in some other GH2 enzymes such as the E. coli (7) and Arthrobacter sp. GH2 β-galactosidase (8). Domain 1 (colored blue in Fig. 4A) consisting of residues 28–218, is a β sheet domain containing a five-stranded antiparallel β sheet, a four-stranded antiparallel β sheet, and an α helical region, reminiscent of carbohydrate-binding modules (reviewed in Ref. 45).

TABLE 2
X-ray data quality and structure refinement statistics

|                     | BtMan2A selenomethionine | BtMan2A native |
|---------------------|--------------------------|-----------------|
| Data                |                          |                 |
| Beam line (Å)       | ID23                     | ID14-4          |
| Wavelength (Å)      | 0.9795                   | 0.97930         |
| Resolution of data (Å) | 91.28–2.800             | 39.6–1.70 (1.79–1.7) |
| Rmerge              | 0.065 (0.115)            | 0.063 (0.272)   |
| Mean I/sI           | 16.6 (9.7)               | 15.3 (5.3)      |
| Completeness %      | 100 (100)                | 100 (100)       |
| Multiplicity        | 4.2 (4.2)                | 4.1 (4.1)       |
| Refinement*         |                          |                 |
| Rmerge/Rfree        | 0.16/0.19                |                 |
| Number protein atoms| 694/7069                 |                 |
| Number water molecules/ glycerol atoms/Cl ions | 1856/54/4 |                 |
| Root mean square deviation |                       | 0.016         |
| 1–2 bonds (Å)       | 1.59                     |                 |
| Root mean square deviation angles (°) |                                 |
| Mean main chain B (Å²) | Mol A 16                | Mol B 14        |
| Mean B (Å²) water/ glycerol/Cl |                          | 27/29/20  |
| PDB code            | 2JE8                     |                 |

*a Of the 1674 modeled protein residues Cys-398, Ile-194, and Trp-510 lie in "disallowed" regions of the Ramachandran plot with clear and unambiguous density. A total of 31 side chains have been modeled with alternative conformations.

Domains 2 (cyan) and 4 (yellow) are structurally very similar to domain 1, again displaying immunoglobulin/CBM-like folds, which, in this case, comprised residues 219–331 and 676–780, respectively. Both these domains consist of two antiparallel β sheets one containing four β strands and the other three. Domain 3 (green) contains the catalytic center and has the classical (B/a)₆ “TIM” barrel fold consisting of residues 332–675, discussed below, whereas domain 5 (red) comprises residues 786–860. The first two residues of the C-terminal His tag are evident in one but not both subunits of the homodimer. The structure of BtMan2A differs from that of the E. coli LacZ β-galactosidase most significantly in the orientation of domain 5, which in LacZ is bent away from the active site giving the protein a more globular shape. The dimerization interface of BtMan2A is formed by interactions between domain 5 of both monomers, whereas in the E. coli β-galactosidase the equivalent domain also forms the basis of the interactions which lead to tetrameric oligomerization.

A DALI (46) search on the whole protein shows that BtMan2A is most similar to the H. sapiens β-glucuronidase and the E. coli β-galactosidase with an root mean square deviation of 2.9 and 3.9 Å and Z scores of 24.5 and 23.0, respectively. Taking each of the five domains separately, reveals that domain 1 (residues 2–192) has the highest similarity to the N-terminal domains of both these proteins with a root mean square deviation of 2.8 Å and a Z score of 12.4 with the E. coli β-galactosidase and root mean square deviation = 2.9 Å and Z = 11.8 for the H. sapiens β-glucuronidase. Domains 2 (219–331) and 4 (676–780) both show most similarity with the equivalent domains of E. coli β-glucuronidase and root mean square deviation = 2.9 Å and Z = 11.8 for the H. sapiens β-glucuronidase. Despite the general similarities between the E. coli and BtMan2A proteins the fifth domain is clearly significantly different. A DALI search carried out only on residues in this domain (758–834) results in none of the first hits being for the GH2 enzymes. The most similar structure (Z = 6.1) found during this search is an appendage domain in mice involved in endocytosis. With almost the same Z score (5.9) an “immunoglobulin-like domain” in a sialidase from Micromonospora viri-

FIGURE 4. The structure of BtMan2A. A shows a divergent (wall-eyed) stereo schematic of the structure of BtMan2A, drawn with the program MOLSCRIPT (36), and colored according to the domain. The bis-Tris propane molecule is drawn as gray shaded spheres and gives an indication of the active center. B displays the observed electron density for a molecule of Bis-Tris propane (gray) in the active center of BtMan2A, with the acid-base and nucleophile residues, Glu-462 and Glu-555, shown in green. The compound binds with its positive nitrogen center reflecting the positive charge on the anomeric carbon of the substrate at the transition state(s) that is reflected in its Kᵣ value of 100 μM. The map shown is a maximum-likelihood weighted 2Fᵣ-obs − Fᵣ calc synthesis contoured at 1σ (~0.36 electrons Å⁻³) and the figure was drawn with BOBSCRIPT (37).
difficilis shows distinct similarities in the β sheet regions although the intermediate loops are significantly different.

The Catalytic Center of BtMan2A—The catalytic domain (residues 332–675) is structurally similar to a number of clan GH-A glycoside hydrolases (clan GH-A consists of glycoside hydrolase families with a (β/α)8-fold that contain a conserved catalytic apparatus and hydrolyze the glycosidic bond by a double displacement or “retaining” mechanism) with DALI Z-scores of Z = 23.5 to the human GH2 β-glucuronidase, Z = 19.7 with the exo-mannanase CmMan5A, and Z = 17.7 to the E. coli β-galactosidase. The catalytic residues, implicated as Glu-462 and Glu-555, lie on strands 4 and 7, respectively, of the (β/α)8 barrel as expected, and the active site comprises a deep pocket. Comparison of the active site of BtMan2A with that of the E. coli β-galactosidase shows that in the case of the former domain 2 plays no role in the formation of the catalytic site, whereas in the β-galactosidase a loop from domain 2 forms part of the β barrel. There are substantial differences between the active sites of the two proteins where strands 7, 8, 1, and 2 lie in similar orientations and the remaining strands in BtMan2A are shifted such that they occupy the same regions as occupied by the domain 2 loop in the E. coli β-galactosidase. Notably, in both E. coli β-galactosidase and BtMan2A, aromatic residues derived from domain 1 contribute to the base of the active site pocket and thus to the “exa” activity of the enzymes (discussed further below).

The active center of BtMan2A reveals unambiguous electron density for a single molecule of bis-Tris propane in the active site from the crystallization conditions (Fig. 4B). Tris itself has been observed in a number of glycosidase structures (19) and bis-Tris propane binds similarly here, with one of the Tris moieties occupying the −1 subsite such that its amino group lies approximately in the position expected for the positive charge on the transition state during catalysis and with the hydroxyl groups lying close to the positions expected for the mannoside hydroxyls of the true substrate. Bis-Tris propane indeed acts as a competitive inhibitor of the enzyme with a $K_i$ determined to be 100 μM at pH 7 (see “Materials and Methods). Thus far, we have been unable to crystallize the protein in a bis-Tris propane-free form for ligand-binding studies.

Comparison of the active site of BtMan2A with the GH5 exo-mannanase CmMan5A reveals significant similarity between the −1 subsite of these two enzymes (Fig. 5) suggesting that there is some conservation in mannosyl binding machinery in these proteins. The position and orientation of the two catalytic carboxylates in the retaining mechanism, Glu-462 and Glu-555 in BtMan2A, are essentially identical to the corresponding residues, Glu-215 and Glu-330, respectively, in CmMan5A (Fig. 5). Consistent with the roles of these residues, the E462A and E555A mutants display catalytic efficiencies ~10^3- and 10^5-fold lower, respectively, than the wild type enzyme (Table 1 and Fig. 2). Indeed, the acid-base catalytic function of Glu-462 is supported by the $K_i$ of E462A for DNP-Man, which is ~30-fold lower than the wild type enzyme (Fig. 2 and Table 1). The decrease in this kinetic parameter indicates that the mutation has had a larger influence on $k_3$, the deglycosylation step, than glycosylation ($k_2$) in which protonation of the glycosidic oxygen is not critical when the leaving group has a $pK_a$ significantly lower than the pH of the reaction ($pK_a$ of DNP is 3.5). Detailed biochemical analysis of wild type and site-directed mutants of the related C. fimii GH2 β-mannosidase (41, 42), and other clan GH-A glycoside hydrolases, including GH10 xylanases, GH26 endo-β-1,4-mannanases, and GH53 endo-β1,4-galactanses demonstrate that the equivalent residues to BtMan2A Glu-462 and Glu-555, in these related enzymes, fulfill the same catalytic functions.

In addition to the “catalytic” groups, the aromatic tryptophan platform “below” the −1 subsite is maintained in almost identical orientation between BtMan2A and CmMan5A (Fig. 5). In the case of BtMan2A Trp-645 provides this function, whereas the equivalent in CmMan5A is Trp-376. The importance of this residue in substrate binding, presumably via hydrophobic contacts with the mannosyl residue at the −1 subsite, is demonstrated by the W645A mutation causing a substantial increase in $K_m$ and a 10^4-fold decrease in catalytic efficiency (Table 1). Mutagenesis studies also indicate that three other aromatic residues in the glycone binding site, Trp-198, Trp-200, and Trp-395 (the latter residue corresponds to Trp-137 in many GH-A clan enzymes), contribute to substrate binding as the respective mutations W198A, W200A, and W395A significantly increase the $K_m$ and decrease the catalytic efficiency of the enzyme. By analogy to Trp-137 in CmMan5A, Trp-395 makes a major contribution to the pocket topology of the active site and hence the exo-activity of the enzyme, and is predicted to hydrogen bond to O-3 of the substrate (Fig. 5) (based upon a similar interaction mediated by a histidine, which is equivalent to Trp-395 in many GH-A clan enzymes). Significantly neither Trp-198 nor Trp-200 are in the central catalytic domain but are present in domain 1 that displays a CBM-like fold and contributes to the base of the active site pocket. Similarly the domain 1 residue Asp-199 also points
Mannose Foraging by *B. thetaiotaomicron*

into the active site (Fig. 5) and the carboxylic acid clearly makes a substantial contribution to both substrate recognition and transition state stabilization as the D199A mutation displays greatly reduced activity. Asn-461, whose orientation is stabilized through a bidentate hydrogen bond with Arg-391, is highly conserved in clan GH-A enzymes (although in GH26 it is a histidine), and by making a hydrogen bond with O-2 of the glycone sugar plays a key role in substrate binding and transition state stabilization. The low activity displayed by the mutant N461A is entirely consistent with the prominent role Asn-461 plays in catalysis contributing around 4.3 kcal/mol to catalysis, similar to that observed for the same interaction for enzymes acting on gluco-configured enzymes (6).

Comparison of *BtMan2A* with *CmMan5A* in complex with the inhibitor isofagomine lactam (43) allows identification of other residues in *BtMan2A* that will interact with substrate. Thus, Tyr-537 is predicted to make a hydrogen bond with substrate O-6 and, indeed, the Y537A mutation causes a 100-fold loss of activity. Tyr-537 is predicted to make a hydrogen bond with sub-

The plasticity in substrate recognition displayed by *BtMan2A* is consistent with the colonic ecosystem of the bacterium. The organism utilizes both dietary and host glycans as major sources of nutrients. The substrate specificity of *BtMan2A*, which releases mannose from both β1,4-linked manno-oligosaccharides and from the Manβ1,4GlcNAcβ1,4GlcNAc stem of the core region of N-linked human glucans, reflects the nutrient sources presented to the prokaryote. Indeed *man2A* is part of an operon containing a sialidase, a sialic acid-specific 9-O-acetylatedase, and three β-hexosaminidase genes, which degrade mammalian glycans (2), and this genetic locus is up-regulated when host mice are fed diets lacking fermentable polysaccharides, enabling the bacterium to scavenge carbohydrates from the mammals glycoproteins (2). Thus the most common natural substrate for *B. thetaiotaomicron* is likely to be Man-GlcNAc derived from the host glycoproteins. By contrast, in *CmMan5A* the aglycone binding sites play a key role in substrate recognition and thus the enzyme displays much tighter specificity for mannose-containing homopolymers. Although it is currently unclear...
whether the plasticity in aglycone recognition displayed by \( \text{BtMan2A} \) is a general feature of GH2 \( \beta \)-mannosidases (the activity of these enzymes are generally assessed using aryl-glycosides as substrates), it is interesting to note that the genome of \( B. \text{thetaiotaomicron} \) contains 32 GH2 open reading frames, and it is possible that tighter aglycone substrate recognition may confer more specialized roles for at least some of these enzymes. Indeed, the glycosobiology of \( B. \text{thetaiotaomicron} \), in which there has been a dramatic expansion in GH2 and GH43, which to date contain primarily \( \text{exo} \)-acting enzymes, but a paucity of enzymes in the families known to contain primarily \( \text{endo} \)-acting plant structural polysaccharidases, with no obvious candidate \( \text{endo} \)-\( \beta \)-mannanases, may suggest that the bacterium specializes in scavenging sugars both from other oligosaccharides produced by \( \text{endo} \)-acting glycoside hydrolases, presumably expressed by colonic microorganisms, and from the diverse human glycans in the colonic milieu.

REFERENCES

1. Coughlan, M. P. (1985) in Biotechnology and Genetic Engineering Reviews (Russell, G. E., ed) pp. 1–54, Intercept, Paris
2. Sonnenburg, J. L., Xu, J., Leip, D. D., Chen, C. H., Westover, B. P., Weatherford, J., Buhler, J. D., and Gordon, J. I. (2005) \( \text{Science} \) 307, 1955–1959
3. Skalova, T., Dohnalek, J., Spiwok, V., Lipovova, P., Vondrackova, E., Jacobson, R. H., Zhang, X. J., Dubose, R. F., and Matthews, B. W. (1994) \( \text{Biochemistry} \) 33, 761–766
4. T. Hogg, D., Pell, G., Dupree, P., Goubet, F., Martin-Orue, S. M., Armand, S., and Gilbert, H. J. (2003) \( \text{Biochemistry} \) 40, 54369–54379
5. Hogg, D., Pell, G., Dupree, P., Goubet, F., Martin-Orue, S. M., Armand, S., and Gilbert, H. J. (2003) \( \text{Biochemistry} \) 39, 5013–5021
6. Andrews, S. R., Taylor, E. J., Pell, G., Vincent, F., Ducsos, V. M., Davies, G. J., Lakey, J. H., and Gilbert, H. J. (2004) \( \text{J. Biol. Chem.} \) 279, 54369–54379
7. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) \( \text{Nat. Struct. Biol.} \) 6, 458–463
8. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) \( \text{Cell Walls} \), Chapman and Hall, London
9. Cell Walls, Chapman and Hall, London
10. Brzozowski, A. M., Lawson, D. M., Turkenburg, J. P., Bisgaard-Frantzen, H., Svendsen, A., Borchert, T. V., Dauter, Z., Wilson, K. S., and Davies, G. J. (2000) \( \text{Biochemistry} \) 39, 9099–9107