BIOCHEMISTRY
including biophysical chemistry & molecular biology

ARTICLE

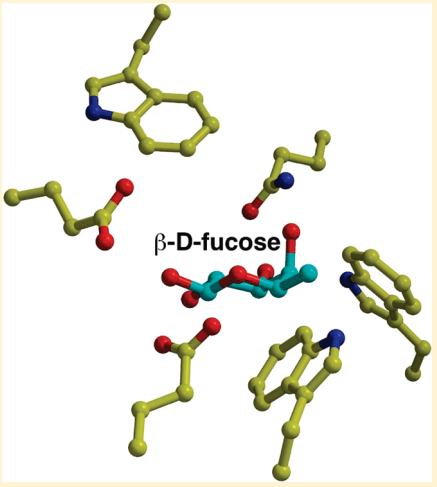
pubs.acs.org/biochemistry

Structural and Functional Analyses of a Glycoside Hydrolase Family 5 Enzyme with an Unexpected β-Fucosidase Activity

Shosuke Yoshida,[‡,§,#] David S. Park,[†,&,#] Brian Bae,[†] Roderick Mackie,[‡,§,∥] Isaac K. O. Cann,[*,‡,§,∥,⊥] and Satish K. Nair[*,†,§,&]

[†]Department of Biochemistry, [‡]Energy Biosciences Institute, [§]Institute for Genomic Biology, [∥]Department of Animal Sciences, [⊥]Department of Microbiology, and [&]Center for Biophysics and Computational Biology, University of Illinois, Urbana, Illinois 61801, United States

ABSTRACT: We present characterization of PbFucA, a family 5 glycoside hydrolase (GH5) from *Prevotella bryantii* B_14. While GH5 members typically are xylanases, PbFucA shows no activity toward xylan polysaccharides. A screen against a panel of *p*-nitrophenol coupled sugars identifies PbFucA as a β-D-fucosidase. We also present the 2.2 Å resolution structure of PbFucA and use structure-based mutational analysis to confirm the role of catalytically essential residues. A comparison of the active sites of PbFucA with those of family 5 and 51 glycosidases reveals that while the essential catalytic framework is identical between these enzymes, the steric contours of the respective active site clefts are distinct and likely account for substrate discrimination. Our results show that members of this cluster of orthologous group (COG) 5520 have β-D-fucosidase activities, despite showing an overall sequence and structural similarity to GH-5 xylanases.



The increasing cost and limiting resource of fossil fuels has resulted in increased research focus on renewable and environmental friendly energy sources. Much of the world's renewable carbon source is present in the organic components in the cell wall of plants, and harnessing this renewable energy source can mitigate reliance on fossil fuels.[1−3] One key interest has focused on the use of polysaccharide-hydrolyzing enzymes to harness the energies stored within the polymeric plant biomass composed of hemicellulose, lignin, and cellulose. Therefore, cellulolytic and hemicellulolytic enzymes that can deconstruct cellulose polymers into fermentable simple sugars may facilitate the utilization of a very large source of carbon, one that is plentiful and renewable (the chemical structures of representative sugars are shown in Figure 1).

Ruminant animals harbor a microbial consortium that governs enzymatic hydrolysis of plant cell wall polysaccharides to fermentable sugars. The released sugars are subsequently fermented to short-chain fatty acids that serve as the main energy source for the host animal.[4,5] Therefore, the genomes of plant cell wall degrading microbes in the rumen represent a rich source of enzymes in the form of glycoside hydrolases (GHs) and carbohydrate esterases (CE) that play critical roles in plant cell wall degradation. Glycoside hydrolases catalyze the cleavage of the glycosidic linkages located between adjacent carbohydrate residues, typically by utilizing either a configuration-inverting or configuration-retaining acid-catalyzed mechanism. Carbohydrate esterases aid the GH's by removing side chains in the complex polymers, thus allowing accessibility of the GH's to the glycosidic linkages in the backbone. Currently, there are 115 known families of GH's (organized into 14 clans) based on similarities in their amino acid sequences and three-dimensional folds.[6]

The majority of GH's in the publicly available databases have been assigned functions based on bioinformatic analysis and not through direct functional characterization. The literature, however, shows that glycoside hydrolases with very similar chemical sequences may exhibit different functional activities. An example of this observation was recently reported by Dodd and coworkers, who characterized four GH family 3 enzymes and showed that one of the four polypeptides functions as a cellodextrinase, while the other three proteins demonstrated β-xylosidase activities.[7]

A recent, massive metagenomic analysis of plant cell wall degrading enzymes in the cow rumen reported the existence of a large group of putative plant cell wall degrading enzymes awaiting functional characterization.[8] These observations suggest an enrichment of diverse catalytic activities that enable members

Received: January 14, 2011
Published: March 16, 2011

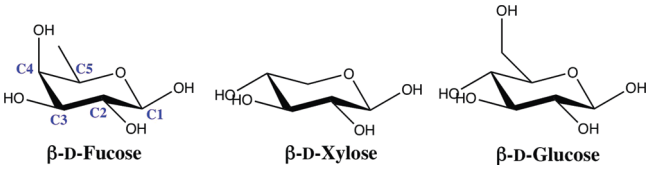


Figure 1. Chemical structures of β-D-fucose, β-D-xylose, and β-D-glucose.

of the rumen microbiome to capture nutrients, in the form of fermentable sugars locked up in plant cell wall polysaccharides. Thus, to gain insight into the strategies used by the rumen microbial consortium to deconstruct the energy-rich hemicellulosic and cellulosic components of plant cell walls, we have focused our attention on detailed biochemical analyses of the products of genes encoding glycoside hydrolase-like polypeptides from this environment.

Prevotella bryantii B_14 is one of the dominant hemicellulolytic bacteria in the rumen.[9] To understand how this bacterium deconstructs the complex polymer that represents the most abundant sugar source next to cellulose, we have initiated a systematic biochemical analysis of the glycoside hydrolases found in its genome. In the present work, we analyzed two putative endoxylanases, encoded by two genes, Pb280 and Pb281, located in tandem on the genome. Screening of the gene products of Pb280 and Pb281 demonstrated that each protein functions as a β-fucosidase, instead of the predicted endoxylanase activity. To further understand the changes conferring this activity on the two enzymes, we determined the 2.2 Å resolution crystal structure of Pb280 and provide insights into the molecular determinants leading to the unexpected activity in these glycoside hydrolases.

EXPERIMENTAL PROCEDURES

Strains, Media, and Growth Condition. *Prevotella bryantii* B_14 was obtained from our culture collection at the Department of Animal Sciences, University of Illinois at Urbana−Champaign. *Prevotella bryantii* B_14 was grown under anaerobic conditions.[10] Gene manipulation and plasmid construction were performed in *Escherichia coli* JM109 (Stratagene, La Jolla, CA). *E. coli* BL21 (DE3) CodonPlus RIPL (Stratagene) was used for gene expression. The *E. coli* cells were grown aerobically at 37 °C in Luria−Bertani (LB) medium supplemented with appropriate antibiotics.

Gene Cloning, Expression, and Protein Purification. *P. bryantii* B_14 was grown, cells were harvested, and the genomic DNA was extracted using DNeasy Tissue kit (QIAGEN, Hilden, Germany). Genes were amplified from the genomic DNA by PCR using Prime STAR HS DNA polymerase (Takara Bio, Otsu, Japan). The coding regions for PbFucA and PbFucB, lacking the predicted signal peptide sequences, were amplified using the polymerase chain reaction. Primers for PCR were based on the sequences of each gene PCR (Table 2) and were engineered to incorporate NdeI and XhoI restriction sites at the 5′ and 3′ ends, respectively. The amplified fragments were cloned into pGEM-T vector (Promega, Madison, WI) by TA cloning and subcloned into a modified pET-28a expression vector (Novagen, San Diego, CA) that was engineered by replacing the kanamycin resistance gene with that for ampicillin resistance.

The expression vectors were introduced into *E. coli* BL21 (DE3) CodonPlus RIPL competent cells and grown in 10 mL of LB medium with ampicillin (100 μg/mL) and chloramphenicol (50 μg/mL) at 37 °C overnight. Each culture was transferred to a fresh LB medium with the same antibiotics and grown until the optical density at 600 nm reached ∼0.4. For each culture, the temperature for culturing was then decreased to 16 °C, and isopropyl β-D-thiogalactopyranoside (IPTG), at a final concentration of 0.1 mM, was added to the medium to induce production of the target protein. After 14 h, cells were harvested by centrifugation (5000 rpm, 4 °C, 15 min) and resuspended in 50 mL of lysis buffer (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 20 mM Imidazole). Cells were disrupted by a French press system (Avestin Inc., Ottawa, Canada), and the lysate was clarified by centrifugation (15 000 rpm, 4 °C, 30 min). The supernatant was filtered through a 0.22 μm pore size Durapore membrane (Millipore, Bedford, MA). The filtrate was applied to HisTrap FF 5 mL (GE Healthcare, Piscataway, NJ) column, and unbound proteins were washed with 20 column volumes of lysis buffer. The bound proteins were eluted with elution buffer (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 250 mM imidazole), and the buffer was exchanged to 50 mM Tris-HCl, pH 7.5, 300 mM NaCl by use of a desalting column (HiPrep 26/10 Desalting, GE Healthcare) for storage.

pH Profiles. The effects of pH for the activities of proteins were determined in the presence of 25 mM citrate−HCl (pH 3.5−5.5), 25 mM morpholine ethanesulfonic acid (MES)−NaOH (pH 5.5−7.0), and 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)−NaOH (pH 7.0−8.0). To determine the catalytic efficiency of enzymes, 2 μM of PbFucA or 0.5 μM of PbFucB was respectively incubated in 0, 5, 10, or 15 mM of *p*-nitrophenol (*p*NP) adducted β-D-fucopyranoside (Sigma-Aldrich, St. Louis, MO) in the buffers with different pHs. The release of *p*NP was monitored at the absorbance of 400 nm using Synergy 2 microplate reader (BioTek, Winooski, VT). The path length of the solution was determined by measuring the absorbance at 977 nm where water has a slight peak. All assays were carried out at 37 °C. Same concentration of bovine serum albumin (BSA) as PB280 or PB281 was utilized as a negative control. The pH-dependent k_{cat}/K_m was incorporated into the following equation that is for the enzyme utilizing two ionizable side chains for catalysis:

$$k_{cat}/K_m = (k_{cat}/K_m)_{max}/\{1 + 10^{(pK_{a1} - pH)} + 10^{(pH - pK_{a2})}\}$$

where k_{cat}/K_m is the experimental value, $(k_{cat}/K_m)_{max}$ is the highest k_{cat}/K_m at optimum pH, and pK_{a1} and pK_{a2} are the pK_as of two ionizable groups critical for enzyme activity. For PbFucA, values for k_{cat}/K_m could not be measured below pH 5.0 as the protein begins to precipitate below this pH range. The Graph Pad Prism v5.01 (GraphPad Software, San Diego, CA) was utilized for calculating these parameters.

Qualitative Binding Assessment with Insoluble Polysaccharides. Binding studies were carried out using a slightly modified protocol from our previously established method. 1 mL of 2 μM PbFucA in 50 mM sodium citrate-HCl (pH 5.5), containing 300 mM NaCl (buffer A), was mixed with 20 mg of either insoluble Oat-spelt xylan (is-OSX), insoluble Birchwood xylan (is-BWX), or Avicel PH-101(Avicel). The reaction mixture was gently mixed at 4 °C for 1 h. The insoluble polysaccharide was then precipitated by centrifugation (13 000 rpm, 4 °C, 2 min). The precipitate was washed with buffer A to remove nonspecific binding proteins and boiled in the SDS-PAGE sample buffer. The collected supernatant (unbound protein) and eluted fractions (bound protein) were resolved by SDS-PAGE. The control sample was prepared by

incubating the protein without insoluble polysaccharide in the reaction buffer. Band intensity was measured using the GeneTools software (Syngene, Frederick, MD).

Site-Directed Mutagenesis. Site-directed mutagenesis was carried out using the QuikChange Multi Site-Directed Mutagenesis Kit (Stratagene), according to the manufacturer's instructions. Primers used in the site-directed mutagenesis study are summarized in Table 1. For the specific activity determination of PbFucA and PbFucB for *p*NP-β-D-fucopyranoside, 2 μM of PbFucA or 0.5 μM of PbFucB was respectively incubated in 15 mM of *p*NP-β-D-fucopyranoside in citrate−HCl pH 5.5 containing 150 mM NaCl, and the initial rate that the *p*NP is released was determined. The same concentration of BSA as PbFucA or PbFucB was utilized as a negative control.

Crystallization and X-ray Data Collection. Prior to crystallization, the polyhistidine affinity tag was cleaved using thrombin, and the resultant sample was further purified using size-exclusion chromatography (Sephacryl 200, GE Healthcare) in a buffer consisting of 20 mM HEPES (pH 7.5) and 100 mM KCl. Initial crystallization conditions were established with commercial screens from Hampton Research and Emerald Biosystems. Refinement of promising conditions yielded large crystals suitable for diffraction analysis. Typically, a 2 μL drop containing PbFucA at 10 mg/mL (in 100 mM KCl, 10 mM HEPES, pH = 7.5) was added to 2 μL of precipitant (0.2 M $(NH_4)_2SO_4$, 0.1 M MES monohydrate (pH 6.5), 30% poly(ethylene glycol) monomethyl ether 5000) and equilibrated over a well containing the precipitant solution at 8 °C. Crystals grew over a period of 3 weeks and, following transient immersion into a precipitant solution supplemented with 25% glycerol, were vitrified by direct immersion into liquid nitrogen.

Crystals of PbFucA occupied space group *C*222 with unit cell parameters a = 229.14 Å, b = 224.48 Å, c = 229.8 Å and contained eight molecules of the complex in the crystallographic asymmetric unit. Although the unit cell parameters of each cell edge are suspiciously close, attempts to scale the diffraction data in higher symmetry space groups resulted in unexpectedly high merging R values. Analysis of the distribution of diffraction intensities reveals that these crystals are merohedrally twinned with a twin fraction of 0.217. Diffraction data were collected to 2.2 Å resolution at an insertion device line (LS-CAT-Sector 21ID-D, Advanced Photon Source, Argonne, IL) and integrated and scaled using the HKL2000 package.[11] A 4-fold redundant data set was collected with an overall R_{merge} 0.119, and $I/\sigma(I)$ = 2.2 in the highest resolution shell. Relevant data collection statistics are given in Table 1.

Crystallographic phases were determined by the molecular replacement method[12,13] using the refined coordinates of a hypothetical protein from *Bacteroides fragilis* NCTC 9343 on unknown function whose structure has been determined by the New York Center for Structural Genomics (PDB ID: 3CLW, 29% sequence identity over 496 residues). Following rigid body refinement of the initial molecular replacement solution, the atomic model was subject to automatic rebuilding using either ARP/wARP[14] or the PHENIX[15] program suite, resulting in placement of roughly 70% of the main chain and 40% side chain atoms. The remainder of the model was fitted using XtalView[16] and further improved by rounds of refinement with REFMAC5.[17,18] Cross-validation used 5% of the data in the calculation of the free R factor.[19,20] The stereochemistry of the models was routinely monitored throughout the course of refinement using PROCHECK.[21] Crystal parameters, data collection parameters, and refinement statistics for each of the structures are summarized in Table 2. The refined coordinates have been deposited in the PDB.

Table 1. Data Collection, Phasing, and Refinement Statistics

	PbFucA
data collection	
space group	*C*222
cell dimensions	
a, *b*, *c* (Å)	229.14, 229.48, 229.80
resolution (Å)	50−2.2 (2.28−2.2)[a]
R_{sym} (%)	11.9 (43.0)
$I/\sigma(I)$	11.1 (2.2)
completeness (%)	94.9 (82)
redundancy	3.9 (2.5)
refinement	
resolution (Å)	25.0−2.2
number of reflections	274 192
R_{work}/R_{free}[b]	24.4%/27.9%
number of atoms	
protein	32128
solvent	1569
average *B* value	
protein	25.1
solvent	25.5
rms deviations	
bond angles (Å)	1.43
bond lengths (deg)	0.014

[a] Highest resolution shell is shown in parentheses. [b] R factor = $\Sigma(|F_{obs}| - k|F_{calc}|)/\Sigma|F_{obs}|$, and R_{free} is the R value for a test set of reflections consisting of a random 5% of the diffraction data not used in refinement.

Table 2. Primers Used in This Study

primers	sequence	purpose
ORF0280-F	5′-CATATGAATACTAAGGTAACGATTAACCCTTCTAAAACTTATC-3′[a]	cloning
ORF0280-R	5′-CTCGAGTTAATGGATAATCATTTTTTTACCTTGTTTG-3′[a]	cloning
PbFucA-E205A	5′-CATCGACCCCGTTAACGCACCGGCTTTCAACTGG-3′[b]	mutagenesis
PbFucA-E329A	5′-GAAACAGCACAGACAGCATGGAGCATGCTCGATGC-3′[b]	mutagenesis
ORF0281-F	5′-GGAATTCCATATGCAGCAGACACGCATCTATCACATTGATACCG-3′[a]	cloning
ORF0281-R	5′-CTCGAGTTACTGCGCAACAAAAGTCGTAATCG-3′[a]	cloning
PbFucB-E202A	5′-CCTCTGTCCGGTTAACGCACCGGATGGGCACTGG-3′[b]	mutagenesis
PbFucB-E328A	5′-GGTAGACTATTGGCAGACCGCAACCTGTATCATGGGGAATG-3′[b]	mutagenesis

[a] Nucleotides incorporated for restriction enzyme digestion are underlined. [b] Nucleotides corresponding to the substituted amino acids are underlined.

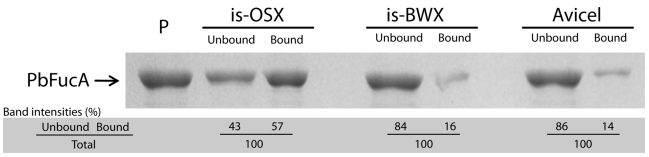


Figure 2. Qualitative polysaccharide binding analysis of PbFucA. For each experiment 2 μM PbFucA was incubated with the respective insoluble polysaccharide. Lane P represents the same amount of protein incubated in the same buffer, but without substrate. PbFucA demonstrates strong binding affinity to Oat-spelt xylan (OSX) and weak, but discernible, affinity to both Brichwood xylan (BWX) and Avicel PH-101.

RESULTS AND DISCUSSION

The ORF0280 gene (Genbank accession number: ZP_07060000) on the genome of *P. bryantii* B_14 is annotated to encode a putative xylanase gene. According to the NCBI Conserved Domain Database (CDD), the corresponding protein is assigned to a cluster of orthologous group of proteins COG5520, whose functional annotation is "O-Glycosyl hydrolase". ORF0280 is located upstream of a second putative xylanase (ORF0281; ZP_07060001). The predicted gene products of ORF0280 and ORF0281 (Pb280 and Pb281, respectively) are homologous, showing 34% identity and 49% similarity.

In order to characterize the enzymatic activities of these proteins, the proteins were overexpressed in *E. coli* and purified to near homogeneity. In order to determine the carbohydrate binding activity of PbFucA, we utilized insoluble polysaccharides to carry out quantitative pull-down assays. The highest binding activity of PbFucA was observed against Oat-spelt xylan, although modest activity could also be observed with Birchwood xylan and Avicel PH-101 (Figure 2). The interactions with these polysaccharides are presumed to occur through interactions with the β sandwich domain that structurally resembles canonical stand-alone carbohydrate binding modules (vide infra).

We further tested the capacity of Pb280 and Pb281 proteins to degrade xylans but detected no activities against such polysaccharides. To elucidate the substrate preferences for these proteins, we investigated their activities against *p*-nitrophenol (*p*NP) adducts of α- or β-linked sugars including arabinose, fucose, xylose, mannose, maltose, rhamnose, galactose, glucose, and cellobiose (Figure 3). Both PB280 and PB281 showed the highest activities against *p*NP-β-D-fucopyranoside and are thus designated as PbFucA and PbFucB, respectively.

In order to further characterize the molecular basis for substrate specificity, we determined the crystal structure of PbFucA to 2.2 Å resolution. The final refined model consists of residues Thr3 through Thr511 (numbering based on the mature enzyme lacking the signal sequence) minus the C-terminal 48 residues. The overall topology of Pb280 consists of a central catalytic $(\alpha/\beta)_8$ barrel (composed of residues Thr14 through Phe413) flanked by amino- and carboxy-terminal extensions that consists of sheets that form an independent β sandwich domain (Figure 5). This small β domain associate with the backside of the catalytic domain, and two domains asscoiate through hydrophobic interactions. By analogy with other polypeptides that are topologically similar, the small β sandwich domain is carbohydrate-binding domain derivative.[22,23]

A structure-based DALI search against the Protein Data Bank identifies the GH5 family xylanase from the organism *Erwinia chrysanthemi*, XynA (PDB number: 1NOF)[22] as the closest characterized homologue (rmsd of 2.8 Å between 369 aligned

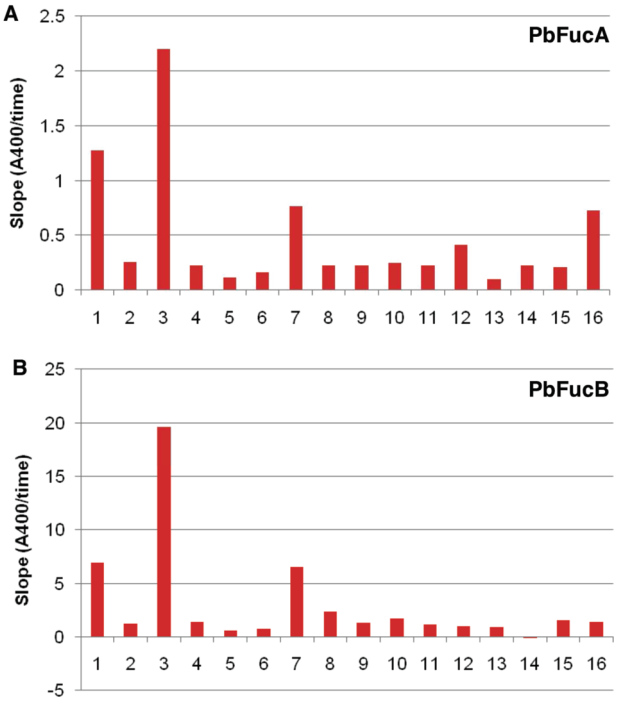


Figure 3. Substrate preference for (A) PbFucA and (B) PbFucB. The substrates set utilized consist of paranitrophenol conjugates of (lane 1) α-L-arabinopyranoside, (lane 2) α-L-arabinofuranoside, (lane 3) β-D-fucopyranoside, (lane 4) α-L-fucopyranoside, (lane 5) β-L-fucopyranoside, (lane 6) α-D-galactopyranoside, (lane 7) β-D-galactopyranoside, (lane 8) α-D-glucopyranoside, (lane 9) β-D-glucopyranoside, (lane 10) β-D-maltopyranoside, (lane 11) α-D-maltopyranoside, (lane 12) α-D-mannopyranoside, (lane 13) β-D-mannopyranoside, (lane 14) α-L-rhamnopyranoside, (lane 15) β-D-xylopyranoside, (lane 16) β-D-cellobiose.

Cα atoms), which shares 24% identity and 40% similarity with PbFucA. The respective catalytic $(\alpha/\beta)_8$ barrel domains can be aligned well, but the β strands of the smaller domain deviate significantly. A structure-based DALI search using only the coordinates of this β sandwich domain reveals structural similarities with similar modules in polypeptides with a range of activities, including the CelM2 glucanase (PDB number: 3FW6; rmsd of 2.1 Å between 89 aligned Cα atoms), the family 51 L-α-furanosidase (PDB number: 1QW8; rmsd of 2.4 Å between 88 aligned Cα atoms), and family 44 endoglucanase (PDB number: 2EO7; rmsd of 2.1 Å between 88 aligned Cα atoms).

The identities of the PbFucA/B active site residues were determined based on earlier studies of the GH5 endoglucanase (Cel5) from *E. chrsyanthemi* and GH10 xylanase from *Penicillium simplicissimum*, where Glu253 is the nucleophile and Glu165 serves as the general acid or base during the hydrolysis of the substrate. The Glu253/Glu165 in XynA structurally corresponds to Glu309 (PbFucA)/Glu311 (PbFucB) and Glu185 (PbFucA)/Glu185 (PbFucB). To demonstrate that these residues in PbFucA and PbFucB contribute to catalysis, we mutated these residues to alanine by site-directed mutagenesis and measured the enzymatic activities against *p*NP-β-D-fucopyranoside. Each of the mutations abolished enzymatic activity (Figure 4), clearly showing that these residues that we predicted are the catalytic residues directly contribute to the catalysis and that β-D-fucosidase activities are derived from PbFucA and PbFucB. PbFucA

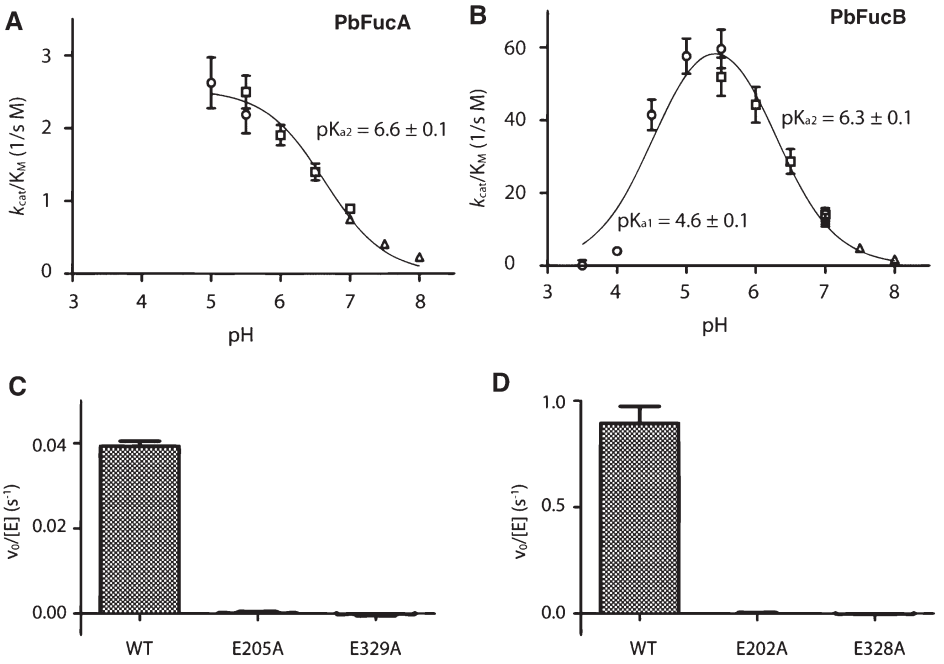


Figure 4. Biochemical characterization of PbFucA and PbFucB. (A, B) The pH dependence of the kinetic parameters (k_{cat}/K_m) for the hydrolysis of *p*NP-β-D-fucopyranoside catalyzed by PbFucA and PbFucB. Measurements were carried out in the following buffers at a concentration of 25 mM of citrate−HCl (open circles), MES−NaOH (open squares), and HEPES−NaOH (open triangles). The pK_a values of the ionizable groups are shown in this panel. The error bar shows standard error. (C, D) Specific activities of PbFucA, PbFucB, and site-directed mutants in which the catalytic nucleophile or the general acid/base catalyst has been changed. Activities were determined under 15 mM of *p*NP-β-D-fucopyranoside.

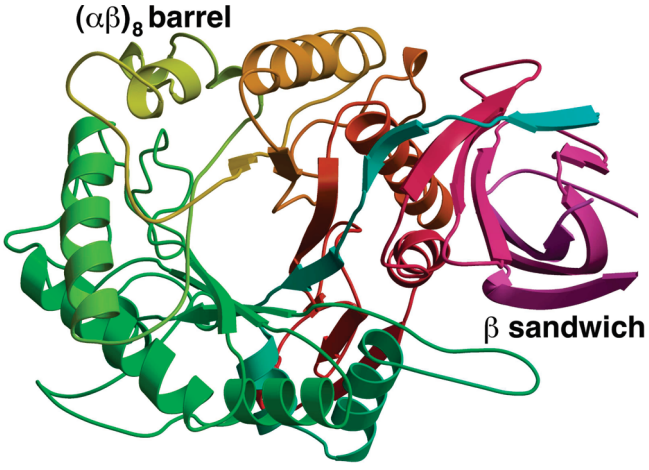


Figure 5. Overall structure of PbFucA. Ribbon diagram derived from the 2.2 Å resolution crystals structure of PbFucA showing the disposition of the $(\alpha/\beta)_8$ barrel and the β sandwich domains. The diagram is colored from the N to C terminus with a color ramp from cyan to red.

likely utilizes the Koshland double-displacement mechanism of retaining glycosidases utilizing two acidic amino acid residues to carry out general acid−base chemistry[24,25] (Figure 7).

Given the fact that PbFucA demonstrates β-D-fucosidase activity despite sharing sequence and structural homology with family 5 xylanases and family 51 α-L-arabinofuranosidases, we wished to identify the molecular bases for this differential specificity. As crystals of inactive variants of PbFucA in complex with substrate analogues proved recalcitrant to diffraction analysis, we generated a model for bound substrate fucose based on structural superpositions with *G. stearothermophilus* GH-51 α-L-arabinofuranosidase AbfA (11% sequence identity; rmsd of 3.5 Å between 374 aligned Cα atoms) and *B. agaradhaerens* GH-5 β-glucanase Cel5A (14% sequence identity; rmsd of 3.4 Å between 250 aligned Cα atoms). A superposition of the structures of PbFucA with AbfA shows that the five-membered L-furanose ring of the substrate in the latter cocrystal structure can largely be accommodated within the PbFucA. However, the C5 and 5-hydroxyl of L-arabinofuranose would clash with Trp358 of PbFucA (the C5-5-OH bond would be situated about 1.75 Å away from Trp358 of PbFucA). The equivalent residue in AbfA is Gln351, which is preceded by a cis-Ala350, resulting in the significant movement of Gln351 in AbfA (relative to Trp358 in PbFucA) that is sufficient to accommodate the L-arabinofuranosidic substrate (Gln351 Nε2-5-OH distance of 3.1 Å). Similarly, the superposition with Cel5A provides a structural rationale for discrimination against glucopyranosidic substrates. The six-membered D-glucose ring fits within the active site of PbFucA, and the C6 and 6-OH of the substrate are situated away from Trp358 and do not result in any steric clashes with the active site residues of PbFucA (Figure 6). However, the ^4E conformation of D-glucose, with C4 above the plane and an equatorial 4-OH, results in steric clashes between the 4-OH and both Trp358 and Trp25 in PbFucA. The residue located at the equivalent position to Trp25 in PbFucA is Gly36 in Cel5A. Although both PbFucA and Cel5A contain a conserved aromatic residue at the position of Trp358 (Trp262 in Cel5A), this residue in PbFucA is situated closer to the active site cavity. This shift is a consequence of the bulky aromatic Tyr357 as the preceding residue in PbFucA, whereas in Cel5A the residue preceding Trp262 is the smaller Asn261. As a result of both these changes, the active site cleft of Cel5A is larger than that in PbFucA, and consequently, the proximity of the equatorial 4-OH of a glucopyranosidic substrate in PbFucA is precluded by steric clashes with Trp25 and Trp358.

dx.doi.org/10.1021/bi200222u |*Biochemistry* 2011, 50, 3369–3375

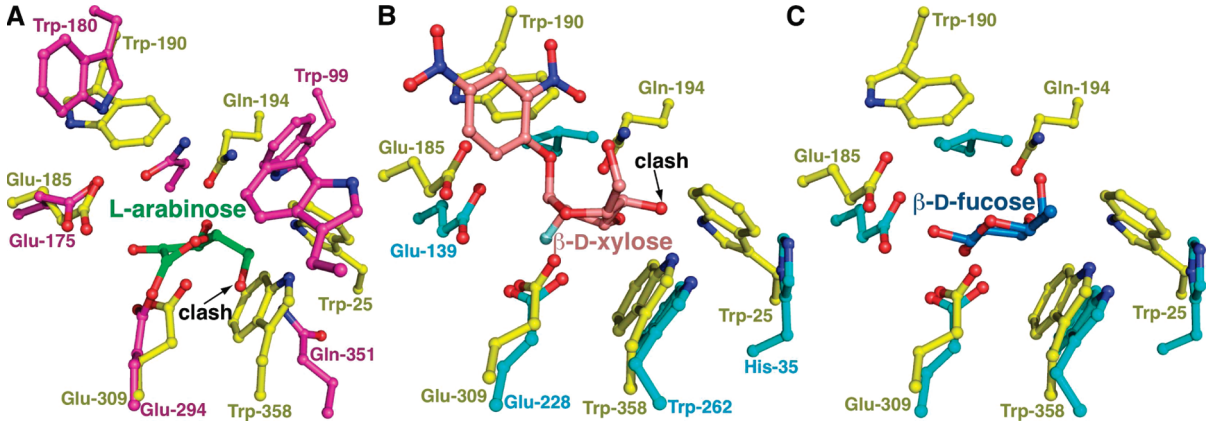


Figure 6. Active site comparisons. A comparison of the active site of PbFucA (shown in yellow) with that of (A) the GH-51 L-α-arabinofuranosidase AbfA (in purple) in complex with L-arabinose (in green) and (B) GH-5 β-glucanase Cel5A (in cyan) in complex with β-D-xylose (in pink) shows the basis for substrate specificity of PbFucA. (C) A model for binding of β-D-fucose (in blue) at the active site of PbFucA, based on the structure of GH-5 β-glucanase Cel5A (in cyan).

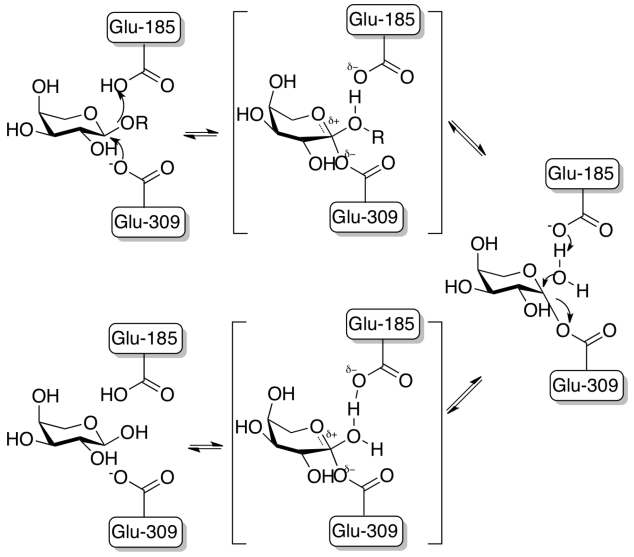


Figure 7. Reaction mechanism for retaining glycoside hydrolases. Double-displacement mechanism for the hydrolysis of β-D-fucopyranoside sugars by PbFucA. Structures of the oxocarbenium-like transition states are shown in brackets.

Prior mechanistic studies establish the importance of a conserved tyrosine residue that is proposed to donate a hydrogen bond to the endocyclic O4 during the formation of the oxocarbenium-like transition state.[26,27] In PbFucA, the equivalent tyrosine (Tyr278) is shifted due to steric occlusion by Trp310 and Ser311. However, the phenolic hydroxyl is situated at the same location as in other retaining glycosidases, in the proximity of the nucleophilic Glu309 (Oη−Oε2 distance of 2.8 Å).

CONCLUSION

The combined biochemical studies presented here expand the substrate scope of GH5 family enzyme to include fucopyranoside substrates. We show that there are changes in the active site contours that likely account for this change. Analysis of substrate deformation by PbFucA, the position of the nucleophile, and the trajectory of the acid/base catalyst relative to the substrate[28] will require additional studies of with substrates or substrate analogues.

AUTHOR INFORMATION

Corresponding Authors

*Address correspondence to: Isaac Cann: 1105 Institute for Genomic Biology, 1206 West Gregory Drive, Tel: 217-333-2090, E-mail: icann@illinois.edu or Satish K. Nair: Department of Biochemistry, 600 S. Mathews Ave, Urbana, IL 61801. Tel: 217-333-0641, E-mail: snair@illinois.edu. Fax: 217-244-5858.

Author Contributions

#These authors contributed equally to this work.

ACKNOWLEDGMENT

We thank Keith Brister and Joseph Brunzelle at LS-CAT (21-ID at APS) for facilitating data collection. This work was funded in part by a grant from the Energy Biosciences Institute to IKOC.

REFERENCES

(1) Carroll, A., and Somerville, C. (2009) *Annu. Rev. Plant Biol. 60*, 165–182.
(2) Solomon, B. D. (2010) *Ann. N.Y. Acad. Sci. 1185*, 119–134.
(3) Weng, J. K., Li, X., Bonawitz, N. D., and Chapple, C. (2008) *Curr. Opin. Biotechnol. 19*, 166–172.
(4) Flint, H. J., Bayer, E. A., Rincon, M. T., Lamed, R., and White, B. A. (2008) *Nat. Rev. Microbiol. 6*, 121–131.
(5) Krause, D. O., Denman, S. E., Mackie, R. I., Morrison, M., Rae, A. L., Attwood, G. T., and McSweeney, C. S. (2003) *FEMS Microbiol. Rev. 27*, 663–693.
(6) Yoeman, C. J., Han, Y., Dodd, D., Schroeder, C. M., Mackie, R. I., and Cann, I. K. (2010) *Adv. Appl. Microbiol. 70*, 1–55.
(7) Dodd, D., Kiyonari, S., Mackie, R. I., and Cann, I. K. (2010) *J. Bacteriol. 192*, 2335–2345.
(8) Hess, M., Sczyrba, A., Egan, R., Kim, T. W., Chokhawala, H., Schroth, G., Luo, S., Clark, D. S., Chen, F., Zhang, T., Mackie, R. I., Pennacchio, L. A., Tringe, S. G., Visel, A., Woyke, T., Wang, Z., and Rubin, E. M. (2011) *Science 331*, 463–467.
(9) Miyazaki, K., Martin, J. C., Marinsek-Logar, R., and Flint, H. J. (1997) *Anaerobe 3*, 373–381.
(10) Griswold, K. E., and Mackie, R. I. (1997) *J. Dairy Sci. 80*, 167–175.

(11) Otwinowski, Z., Borek, D., Majewski, W., and Minor, W. (2003) *Acta Crystallogr., Sect. A 59*, 228–234.
(12) McCoy, A. J. (2007) *Acta Crystallogr., Sect. D: Biol. Crystallogr. 63*, 32–41.
(13) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) *J. Appl. Crystallogr. 40*, 658–674.
(14) Perrakis, A., Sixma, T. K., Wilson, K. S., and Lamzin, V. S. (1997) *Acta Crystallogr., Sect. D: Biol. Crystallogr. 53*, 448–455.
(15) Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J., and Adams, P. D. (2008) *Acta Crystallogr., Sect. D: Biol. Crystallogr. 64*, 61–69.
(16) McRee, D. E. (1999) *J. Struct. Biol. 125*, 156–165.
(17) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr., Sect. D: Biol. Crystallogr. 53*, 240–255.
(18) Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999) *Acta Crystallogr., Sect. D: Biol. Crystallogr. 55*, 247–255.
(19) Brunger, A. T. (1992) *Nature 355*, 472–475.
(20) Kleywegt, G. J., and Brunger, A. T. (1996) *Structure 4*, 897–904.
(21) Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477–486.
(22) Larson, S. B., Day, J., Barba de la Rosa, A. P., Keen, N. T., and McPherson, A. (2003) *Biochemistry 42*, 8411–8422.
(23) Varrot, A., Schulein, M., and Davies, G. J. (2000) *J. Mol. Biol. 297*, 819–828.
(24) Koshland, D. E. (1953) *Biol. Rev. Cambridge Philos. Soc. 28*, 416–436.
(25) Vocadlo, D. J., Davies, G. J., Laine, R., and Withers, S. G. (2001) *Nature 412*, 835–838.
(26) Hovel, K., Shallom, D., Niefind, K., Belakhov, V., Shoham, G., Baasov, T., Shoham, Y., and Schomburg, D. (2003) *EMBO J. 22*, 4922–4932.
(27) Sidhu, G., Withers, S. G., Nguyen, N. T., McIntosh, L. P., Ziser, L., and Brayer, G. D. (1999) *Biochemistry 38*, 5346–5354.
(28) Vasella, A., Davies, G. J., and Bohm, M. (2002) *Curr. Opin. Chem. Biol. 6*, 619–629.