Active-site Peptide “Fingerprinting” of Glycosidases in Complex Mixtures by Mass Spectrometry

DISCOVERY OF A NOVEL RETAINING β-1,4-GLYCANASE IN CELLULOMONAS FIMI

Received for publication, August 1, 2005. Published, JBC Papers in Press, August 5, 2005, DOI 10.1074/jbc.M508434200

Omid Hekmat1, Young-Wan Kim1, Spencer J. Williams2, Shouming He, and Stephen G. Withers3

From the Protein Engineering Network of Centres of Excellence of Canada, Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada

New proteomics methods are required for targeting and identification of subsets of a proteome in an activity-based fashion. Here, we report the first gel-free, mass spectrometry-based strategy for mechanism-based profiling of retaining β-endoglycosidases in complex proteomes. Using a biotinylated, cleavable 2-deoxy-2-fluoro-oxorolysobiose inactivator, we have isolated and identified the active-site peptides of target retaining β-1,4-glycanases in systems of increasing complexity: pure enzymes, artificial proteomes, and the secreted proteome of the aerobic mesophilic soil bacterium Cellulomonas fimi. The active-site peptide of a new C. fimi β-1,4-glycanase was identified in this manner, and the peptide sequence, which includes the catalytic nucleophile, is highly conserved among glycosidase family 10 members. The glycanase gene (GenBankTM accession number DQ146941) was cloned using inverse PCR techniques, and the protein was found to comprise a catalytic domain that shares ~70% sequence identity with those of xylosans from Streptomyces sp. and a family 2b carbohydrate-binding module. The new glycanase hydrolyzes natural and artificial xylo-configured substrates more efficiently than their cello-configured counterparts. It has a pH dependence very similar to that of known C. fimi retaining glycanases.

With the completion of the genome sequences of many organisms, the field of proteomics faces several major tasks. One challenge is that of identification and assignment of structure/function to the tens of thousands of proteins encoded by prokaryotic and eukaryotic genomes. Another challenge is accurate quantitative analysis of changes in protein levels/activities that occur within a proteome as a response to biological perturbations that are due either to normal developmental and metabolic changes or to abnormalities associated with disease. Proteomic techniques such as comparative two-dimensional gel electrophoresis coupled with mass spectrometry, the isotope-coded affinity tagging approach(1), and variations of isotope-coded affinity tagging that identify sites of modification (2) one or more reporter groups that enable rapid detection (e.g. a fluorophore) and/or affinity isolation (e.g. biotin) (16). As such, ABPP methods can provide direct information on post-translational forms of protein regulation (17). However, most of the ABPP research has been limited to detection of the target enzymes, usually by Western blotting or in-gel fluorescence. Recently, a gel-free, mass spectrometry-based ABPP (MS-ABPP) method was used in which the labeled active-site peptides of target enzymes were affinity-isolated from the proteolytic digest of the proteome and analyzed by liquid chromatography/mass spectrometry (LC/MS) and tandem mass spectrometry (MS/MS) procedures to determine their amino acid sequences and the labeled catalytic residue (6). However, the ABPP probe used in that attempt targeted several mechanistically distinct enzyme classes in a non-directed fashion (6). The MS-ABPP approach relies on the assumption that the active-site peptide, containing the labeled catalytic residue, is highly conserved among members of the same enzyme family and therefore provides sufficient information to decipher the enzyme family to which the target enzyme belongs (18).

Here, we report the first MS-ABPP analysis of glycosidases in complex mixtures using the strategy shown in Fig. 1A. Biotin as the reporter group enables easy detection as well as affinity isolation using avidin/streptavidin resin (19). The affinity isolation process is facile and efficient because of the presence of a cleavable disulfide bond between the biotin tag and the mechanism-based inactivator (see Fig. 1, A and B). Fig. 1B shows the chemical structure of the ABPP probe 2,4-dinitrophenyl 4′-amino-N-7-[N-(o-biotinyl)-13-amino-4,7,10-trioxatridecan-ylamino]-4,5-dithiaheptanoyl-2,4′-dideoxy-2-fluoro-β-xylolysobiose (DNPF2X2SSB) and the mechanism of inactivation of target enzymes. The mechanism-based inactivator moiety of DNPF2X2SSB is a 2-de-
oxy-2-fluoroxylolobioside that targets retaining β-xylanases and mixed-function cellulases (20, 21). Endo-1,4-β-xylanases (EC 3.2.1.8) and cellulases (EC 3.2.1.4) are glycoside hydrolases (GHs) that catalyze the hydrolytic cleavage of β-1,4-linked polymers or oligomers of d-xylene and d-glucose, respectively. GHs catalyze the hydrolysis of glycosidic bonds with either retaining or inverting stereochemical outcome (22). Retaining glycosidases operate via a double-displacement mechanism that involves formation of a glycosyl-enzyme intermediate (glycosylation) and its breakdown (deglycosylation) (22). Two catalytic carboxylic acid residues (5 Å apart) act as the catalytic nucleophile and the general acid/base, respectively (see Fig. 1B) (22). The 2-deoxy-2-fluoroglycosides with a good aryl leaving group serve as excellent mechanism-based inactivators of retaining glycosidases (23, 24) because, with these molecules, the glycosylation step is usually much faster than the deglycosylation step, thus the fluoroglycosyl-enzyme intermediate accumulates and can be studied (24).

We have demonstrated this MS-ABPP methodology at three levels of increasing complexity. First, as a test of the strategy, we used this approach to isolate and identify the labeled active-site peptides of an endoxylanase (Bcx from Bacillus circulans) and a mixed-function endoxylanase/cellulase (Cex from Cellulomonas fimi) from their peptic digests. As retaining endoxylanases, Bcx (GH family 11) and Cex (GH family 10) were ideal candidates for targeting by our probe and had been investigated extensively (21). Second, we applied the method to several model enzyme mixtures (artificial proteomes) including Bcx, Cex, or the family 10 glycanases with acarbohydrate-binding domains. The secreted proteome of the aerobic mesophilic xylanolytic soil bacterium C. fimi secretes a complex array of xylanases and cellulases concurrently (25, 26); however, only four C. fimi retaining β-1,4-glycanases have been characterized to date (see Table One) (21, 25, 27–30). Our MS-ABPP analysis, gene cloning, and kinetic characterization of the enzyme revealed that C. fimi secretes a previously unidentified family 10 glycanase with a carbohydrate-binding module (CBM) from CBM family 2b.

**EXPERIMENTAL PROCEDURES**

*General Materials and Methods—*Abg (β-glucosidase from Agrobacterium sp.; EC 3.2.1.21; 51,169 Da), Bcx (endo-1,4-β-xylanase from B. circulans; EC 3.2.1.8; 20,396 Da), Cex (xylanase/cellulase from C. fimi; EC 3.2.1.8; 47,121 Da), GlvA (maltose-6-phosphate glucosidase from Bacillus subtilis; EC 3.2.1.122; 50,513 Da), and LgtC (lipo polysaccharid lyase from C. fimi; EC 2.4.1; 35,743 Da) were purified as described previously (30–34). Hen egg white lysozyme (HEWL; EC 3.2.1.17; 14,315 Da) was purchased from Sigma (Oakville, Ontario, Canada). Mutant E78C Bcx was prepared as described (35). The substrates 2,5-dinitrophenyl (DNPX2) and 2,4-dinitrophenyl (DNPC) were synthesized at Sigma (Oakville, Ontario, Canada). Mutant E78C Bcx was prepared as described (35). The substrates 2,5-dinitrophenyl (DNPX2) and 2,4-dinitrophenyl (DNPC) were synthesized at Sigma (Oakville, Ontario, Canada). Mutant E78C Bcx was prepared as described (35). The substrates 2,5-dinitrophenyl (DNPX2) and 2,4-dinitrophenyl (DNPC) were synthesized at Sigma (Oakville, Ontario, Canada).

**Inactivation Kinetics**—All kinetic studies were performed using a Unicam 8700 UV-visible spectrophotometer equipped with a circulating water bath. The time-dependent inactivation of each enzyme or proteome by DNP2FX2SSB was monitored by measuring the residual xylanase/cellulase activity over time. This was accomplished by the addition of an aliquot of the inactivation mixture at appropriate time intervals to a solution of the substrate (DNPX2 or DNPC) in the appropriate buffer. The initial rates were then measured under steady-state conditions by spectrophotometric monitoring of the release of the 2,4-dinitrophenol(ate) or 2,5-dinitrophenol(ate) at a wavelength of either 440 nm (DNPX2) or 400 nm (DNPC). Pseudo-first-order rate constants (kobs = k(1)/(K + [I])) for the decay of activity were determined in each case from fitting the decay curve to a single exponential decay equation using nonlinear regression with the program GraFit. For Bcx, the inactivation mixture contained 0.20 mg/ml Bcx and 0.500 mM DNP2FX2SSB in BSA-free buffer A. Both inactivation and control samples were incubated at 40 °C, and aliquots were taken at appropriate time intervals and added to a cuvette containing DNPX2 substrate so that the final assay contained 2.8 μg/ml Bcx (active + inactive) and 2.46 mM DNPX2 in buffer A at 40 °C. For Cex, the inactivation mixture contained 57 μg/ml Cex and 22.9 μM DNP2FX2SSB in BSA-free buffer B. Both inactivation and control samples were incubated at 37 °C, and aliquots were taken at appropriate time intervals and added to a cuvette containing DNPC substrate so that the final assay contained 0.81 μg/ml Cex (active + inactive) and 0.493 mM DNPC in buffer B at 37 °C. For C. fimi secreted proteome, the inactivation mixture contained 5.5 mg/ml protein (from C. fimi culture medium) and 0.100 mM DNP2FX2SSB in BSA-free buffer B. Both inactivation and control samples were incubated at 37 °C, and aliquots were taken at appropriate time intervals and added to the cuvette containing either DNPC or DNPC substrate so that the final assay contained 0.11 mg/ml protein and 1.00 mM DNPC or 2.00 mM DNPC in buffer B at 37 °C. For enzyme inactivation prior to proteolysis for affinity isolation, the desired amount of enzyme or proteome was incubated in the presence of 0.100–0.500 mM DNP2FX2SSB for 2–6 h at the temperatures and conditions mentioned above for each enzyme or proteome. Where necessary, samples were dialyzed at room temperature using a non-cellulose 5-kDa cutoff membrane centrifugal filter unit (Millipore Corp.) to exchange buffer A or B with buffer C and to remove excess inactivator. If no dialysis was necessary, enough buffer C was added to the sample to bring it to pH 2.0.

**Proteolysis**—For Bcx, a solution of inactivated Bcx in buffer C was first heat-denatured (boiling for 2 min) and then proteolyzed in the presence of pepstatin (50.1 mass ratio) at room temperature for 3 h. For Cex, a solution of inactivated Cex in buffer C was proteolyzed in the presence of pepstatin (100:1 mass ratio) at room temperature for 6 h. For model proteomes A–C, a solution of the inactivated enzyme mixture in buffer C was first heat-denatured (boiling for 2 min) and then proteolyzed in the presence of pepstatin (50:1 mass ratio) at room temperature for 6–8 h. For C. fimi, the inactivated C. fimi secreted proteome in buffer C was first heat-denatured (boiling for 2 min) and then proteolyzed in the presence of pepstatin (50:1 mass ratio) at room temperature for 8–10 h. In all of the above cases, the proteolytic digest was either immediately subjected to the affinity isolation procedure or freeze-dried first.

---

5 S. J. Williams, O. Hekmat, and S. G. Withers, submitted for publication.
Functional Proteomics of Glycosidases by Mass Spectrometry

Affinity Isolation of the Labeled Peptides—The peptic digest of the inactivated enzyme or proteome in buffer C was incubated with streptavidin resin at a ratio of 1.0 ml of streptavidin resin/60 nmol of biotin-equivalent at room temperature for 1–2 h with shaking. The unlabeled peptides were washed away in buffer C using a spin filter (Novagen). The labeled peptides bound to the streptavidin resin were incubated with 30 mM TCEP in buffer D for 40 min at room temperature with shaking and eluted using a spin filter. The reduction process was repeated several times. The eluted peptide samples were then freeze-dried, redissolved in H2O, and analyzed by electrospray ionization (ESI) MS.

LC/ESI-MS and ESI-MS/MS Analyses—ESI-MS spectra were recorded on a PE-Sciex API 300 triple quadrupole mass spectrometer (Sciex, Thornhill, Ontario) equipped with an ion spray ion source. The ion source voltage was 5 kV, and the orifice voltage was 50 V. All spectra were obtained in the positive-ion single quadrupole scan mode (LC/MS) or the tandem MS fragment ion scan mode (MS/MS). The LC/MS mode, the quadrupole mass analyzer was scanned over an m/z range of 300–2000 atomic mass units with a step size of 0.5 Da and a dwell time of 1.0 ms/step. In the MS/MS mode, the spectra were obtained by selectively introducing the precursor peptide with the m/z value of interest from the first quadrupole (Q1) into the collision cell (Q2) and observing the fragment ions in the third quadrupole (Q3). Q3 scan range was 50–1900 atomic mass units with the same step size and dwell time as above. Protein or peptide samples were introduced by reverse-phase HPLC on an Ultrafast microprotein analyzer (Michrom Bioresources, Inc., Auburn, CA) directly interfaced with the mass spectrometer. For protein analysis, the sample was loaded onto a PLRP-S C18 column (1.0 × 50 mm; Michrom Bioresources Inc.) and eluted with a linear gradient of 5–90% solvent B over 3 min, followed by 90% solvent B over 5 min at a flow rate of 50 μl/min (solvent A, 0.05% trifluoroacetic acid and 2% CH3CN in water; and solvent B, 0.045% trifluoroacetic acid and 80% CH3CN in water). For peptide analysis, the sample was loaded onto a Jupiter C18 column (1.0 × 150 mm; Phenomenex Inc., Torrance, CA) and eluted with a linear gradient of 0–60% solvent B over 60 min at a flow rate of 50 μl/min.

C. fimi Culture Growth and Secreted Protein Preparation—C. fimi cells (ATCC 484) were grown in minimal medium containing 0.1% xylan (from birchwood) as the inducer at 30 °C for 5 days. Cells were then removed by centrifugation at 4 °C, and the culture supernatant was concentrated and exchanged with buffer E under N2 pressure in an Amicon ultrafiltration cell (Millipore, Beverly, MA) using a non-cellulose ultrafiltration membrane disc (Pall Life Sciences, Ann Arbor, MI). The secreted protein preparation in buffer E was then filtered through a 0.45-μm non-cellulose membrane (Millipore Corp.) and analyzed by SDS-PAGE. Protein concentration was determined using the Micro BCA™ protein assay reagent kit with BSA standards (Pierce). A measure of the xylanase and cellulase activities of the secreted protein preparation was obtained by assaying with DNPC and DNPS, and 80% CH3CN in water). For peptide analysis, the sample was loaded onto a Jupiter C18 column (1.0 × 150 mm; Phenomenex Inc., Torrance, CA) and eluted with a linear gradient of 0–60% solvent B over 60 min at a flow rate of 50 μl/min.

C. fimi Culture Growth and Secreted Protein Preparation—C. fimi cells (ATCC 484) were grown in minimal medium containing 0.1% xylan (from birchwood) as the inducer at 30 °C for 5 days. Cells were then removed by centrifugation at 4 °C, and the culture supernatant was concentrated and exchanged with buffer E under N2 pressure in an Amicon ultrafiltration cell (Millipore, Beverly, MA) using a non-cellulose ultrafiltration membrane disc (Pall Life Sciences, Ann Arbor, MI). The secreted protein preparation in buffer E was then filtered through a 0.45-μm non-cellulose membrane (Millipore Corp.) and analyzed by SDS-PAGE. Protein concentration was determined using the Micro BCA™ protein assay reagent kit with BSA standards (Pierce). A measure of the xylanase and cellulase activities of the secreted protein preparation was obtained by assaying with DNPC and DNPS, and 80% CH3CN in water). For peptide analysis, the sample was loaded onto a Jupiter C18 column (1.0 × 150 mm; Phenomenex Inc., Torrance, CA) and eluted with a linear gradient of 0–60% solvent B over 60 min at a flow rate of 50 μl/min.

PCR Amplification of the Internal Gene Probe—The internal gene probe encoding the part of the new C. fimi glycanase (named Cfx) lying between the acid/base and nucleophile regions was PCR-amplified using 100 ng of NotI-digested C. fimi genomic DNA, 1 μM each primer (Cfx-AB, 5'-TTC GAC GTS RTC AAC GAG-3'; and Cfx-Nu, 5'-GTC SAG CTC SG TAT CTC SAC-3'), 20 mM Tris (pH 7.5), 0.2 mM each dNTP, 5% dimethyl sulfoxide, 4 mM MgSO4, and 2.5 units of Pwo polymerase in a 50-μl reaction volume. Thirty PCR cycles (30 s at 94 °C, 30 s at 60 °C, and 45 s at 72 °C) were performed in a thermal cycler (GeneAmp PCR System 2400, PerkinElmer Life Sciences). Both ends of the resulting PCR product were phosphorylated with T4 polynucleotide kinase, and then the DNA fragment was subcloned into Smal-digested pUC18.

Inverse PCR—We used the inverse PCR method (iPCR) as described by Ochman et al. (36) to clone the entire sequence of the cfx gene. Genomic DNA was digested with BamHI, NotI, SacI, or XhoI, and then the self-ligated libraries were used as templates for iPCR. A total of three
DNA fragments (Cfx-Sc, Cfx-Xh, and Cfx-Bm) were amplified using three primer sets as follows: a set of the iPCR-rev primer (5′-CCA GTC GTT GCC GGT CTC CAG GGT TTA AAC TCC GGC-3′) and the iPCR-fw primer (5′-TGG GTC GGC TTC CAG GGG CAC TTC AAG TCC GGC-3′), a set of the iPCR-Sac-rev primer (5′-CCG AGC GTC CGT CAG CCA TCT GCC AGT-3′) and the iPCR-fw primer, and a set of the iPCR-Bam-rev primer (5′-TCA GGG GGG CAG GTC GAA GCT GGC GAC-3′) and the iPCR-Xho-fw primer (5′-GAC CGT CCG GCC GCA GAA GAT CAT CCG-3′), respectively. After analysis of the sequences of the three fragments, primers for amplification of the full-length analysis of the sequences of the three fragments, primers for amplification of the sequences were designed as follows: Cfx-TOP-Nc primer, 5′-GGC AGT CAT GGC CAC GAA ACT CCA CGC GAC-3′; and Cfx-END-Nt primer, 5′-GTC ACG TGC GGC GCG GCA TGA CCC GCT GAC-3′. The products of PCR with these primers were subcloned into plasmid pR2TK, which has a B. subtilis α-amylase promoter (amyR2) from pAR2 (37) as well as multiple cloning sites, a histidine tag sequence, and the T7 terminator from pET28a. The resulting plasmid was designated pR2Cfx-His6.

Expression of the cfx Gene in Escherichia coli—A single colony was cultured overnight in 50 ml of LB medium containing 20 μg/ml kanamycin at 37°C. The pre-culture was then re-inoculated in 1 liter of medium and grown for an additional 15 h. The cell pellet was harvested and resuspended in 50 mM Tris-HCl (pH 8.0) containing 5 mM imidazole and 300 mM NaCl. The cell suspension was twice through a French press at 4°C and centrifuged at 10,000 × g for 30 min, and the soluble extract was purified by nickel affinity chromatography. The protein was then dialyzed against 50 mM phosphate buffer (pH 7.0) and stored at 4°C. Protein concentration was determined using the Micro BCA™ protein assay reagent kit with BSA standards.

Cfx Kinetics—Initial rates of hydrolysis of p-nitrophenyl β-xiloside (pNPX), p-nitrophenyl β-glucoside (pNPG), p-nitrophenyl β-xyloligoside (pNPGx), and p-nitrophenyl β-cellubioside (pNPC) catalyzed by Cfx (the new xylanase from C. fim) were determined at 37°C by monitoring the reactions spectrophotometrically at 400 nm, where Δε400 = 7280 M−1 cm−1 (32). Assays were carried out in buffer F at 37°C, and reactions were initiated by the addition of the enzyme. The substrate and enzyme concentrations in the assays were as follows: 1.0–10 mM pNPX and 1.1 mM Cfx; 1.0–10 mM pNPX and 6.8 mM Cfx; 0.02–2.0 mM pNPX with 0.14 mM Cfx; and 0.2–10 mM pNPX and 2.8 mM Cfx. For pNPXs and pNPC, the plots of initial rates versus substrate concentrations were hyperbolic, and the values of Km and Vmax were determined by fitting the initial velocity curves to the standard Michaelis-Menten equation using nonlinear regression with the program GraFit. For pNPX and pNP, no saturation occurred over the practical concentration ranges, and the values of Vmax/Km were determined by linear regression analysis of the data.

The initial rates of hydrolysis of birchwood xylan and Cm-cellulose catalyzed by Cfx were measured following the p-hydroxybenzoic acid hydrazide method (38), which quantitates the production of reducing sugar. Saturating solutions of xylan (<1% (w/v)) and Cm-cellulose (<2% (w/v)) in buffer F were prepared and preincubated at 37°C. Reactions were initiated by the addition of Cfx (0.14 μM in xylan solution and 1.4 μM in Cm-cellulose solution). Aliquots of 100 μl were removed at appropriate time intervals and added to 900 μl of 50 mM p-hydroxybenzoic acid hydrazide and 0.5 M NaOH in H2O in glass test tubes. Samples were then heated in a 75°C water bath for 30 min and cooled to room temperature. The absorbance at 420 nm was measured, and the amount of reducing sugar released was determined using a glucose standard curve generated under the same conditions. To ensure the validity of the assays for enzyme activity, enzyme concentrations were chosen so that the production of reducing sugar in the assay fell within the linear response range of the standard curve. Assuming that saturating solutions of xylan and Cm-cellulose were also saturating for the enzyme, the Vmax value was determined by linear regression analysis of the data in each case.

Cfx was found to be stable over a pH range of 4.0–9.0 for several hours. The pH dependence of kcat/Km for Cfx was determined by the substrate depletion method. The time course of depletion of 0.1 mM pNPX2 (<Km) was followed for two or more half-lives in the presence of 0.11 mM Cfx in 50 mM buffer at the appropriate pH at 37°C by monitoring the increase in A400 with time. Reactions were initiated by the addition of Cfx. The buffers used were sodium citrate for pH 4.0–6.0, buffer F for pH 6.5–7.8, and Tris for pH 8.0–9.0. The depletion curve in each case was fitted to an exponential decay equation using the program GraFit, and the kcat/Km values were determined using the equation kobs = (kcat/Km)[E]0.

The effect of temperature on Cfx stability was determined by incubating aliquots of Cfx (0.16 mg/ml) in buffer F at 20, 30, 40, 50, and 70°C for 10 h. Residual activities were then determined by measuring the initial rates as described above. The assays contained 1.0 mM pNPX2 and 57 mM Cfx in buffer F at 37°C.

RESULTS

Design of the ABPP Probe DNP2FX2SSB—DNP2FX2SSB (Fig. 1B) consists of a biotinylated arm containing a disulfide bond (cleavable by reduction) and a mechanism-based inactivator (2-deoxy-2-fluorooxylbiosoide) that targets retaining β-xylanases and mixed-function cellobiases (20, 21). Detailed synthesis of DNP2FX2SSB will be reported elsewhere.5

Time-dependent Inactivation of Bcx and Cex by DNP2FX2SSB—Incubation of both Bcx and Cex with DNP2FX2SSB resulted in time-dependent inactivation of each enzyme according to pseudo first-order kinetics (Fig. 2A), as confirmed by following aliquots of the inactivation mixture with the substrate DNPXx for Bcx or DNPC for Cex and monitoring the release of 2,4-dinitrophenol(ate) or 2,5-dinitrophenol(ate) in solution over time. For each enzyme, the inactivation data fit well into a single exponential decay equation: (initial rate)i = (initial rate)i0(exp(-ki obs)), where kobs = k0/[I + Ki]. Inactivation of Bcx in the presence of 0.5 mM DNP2FX2SSB occurred with an apparent rate constant (kobs) of (2.01 ± 0.08) × 10−3 s−1, which is similar to the value of (3 ± 1) × 10−3 s−1 calculated for the parent non-biotinylated inactivator, 2,4-dinitrophenyl-2-deoxy-2-fluoro-β-xyloligoside (DNP2FXx) (39). Likewise, inactivation of Cex in the presence of 22.9 μM DNP2FX2SSB occurred with kobs = (6.7 ± 0.1) × 10−4 s−1, whereas the corresponding value for the armless DNP2FXx is (8.2 ± 0.2) × 10−4 s−1 (20). These results clearly demonstrate that the presence of the biotin arm does not significantly alter the inactivation behavior.

Detection of the Covalent Biotinylated Fluoroglycosyl-enzyme Intermediate and Demonstration of Reductive Cleavage of the Biotinylated Arm by ESI-MS—The stoichiometry of full inactivation of Bcx and Cex by DNP2FX2SSB was determined using LC/ESI-MS analysis. Reconstructions of the ESI mass spectra of the native enzymes and those of the fully inactivated enzymes (Fig. 2B) showed that inactivation resulted in a mass increase of 885 ± 2 Da, which corresponds to the molecular mass of 2FX2SSB. Therefore, in each case, one enzyme molecule reacted with one DNP2FX2SSB molecule, forming the biotinylated fluoroglycosyl-enzyme intermediate and releasing 2,4-dinitrophenol(ate) (Fig. 1B). Furthermore, we demonstrated that the biotinylated arm could be cleaved from the fluoroglycosyl-enzyme intermediate via reduction of the disulfide bond in the presence of the water-soluble phosphine TCEP.
FIGURE 2. Application of the MS-ABPP strategy to Bcx and Cex. A, time-dependent inactivation kinetics. ▲, experimental data representing the hydrolytic activities of control samples (without DNP2FX2SSB) toward DNPX2 (Bcx) or DNPC (Cex) over time. The lines are the linear fits. ●, experimental data representing the hydrolytic activities of samples toward DNPX2 (Bcx) or DNPC (Cex) in the presence of DNP2FX2SSB over time. The curves are the nonlinear fits of the data to a single exponential decay equation. B, reconstructs of ESI mass spectra demonstrating the covalent labeling of Bcx and Cex by DNP2FX2SSB to form the biotinylated fluoroxylobiosyl-enzymes and cleavage of the biotin arm via reduction of the disulfide bond by TCEP. calc., calculated; DNP, 2,4-dinitrophenol(ate). C, ESI mass spectra showing the affinity-isolated, labeled active-site peptides of Bcx and Cex (full label, 2FX2SH; label fragment, 2FX). All peaks represent [M + H]⁺ or [M + 2H]²⁺ species. The spectra represent the reverse-phase HPLC elution time of 15–45 min. D, ESI-MS/MS fragment ion spectra showing the amino acid sequences of the affinity-isolated peptides bearing the labeled catalytic nucleophile. HSB, the cleared biotin arm.
without destroying the ester bond in the intermediate. Reconstructions of the ESI mass spectra of the inactivated enzymes that had been incubated with TCEP (Fig. 2B) showed that the reduction resulted in a mass decrease of 536 ± 2 Da, which corresponds to the molecular mass of the reduced biotinylated arm (HSB). The choice of TCEP over dithiothreitol is based upon the fact that reduction by dithiothreitol is not effective at pH < 7 (40) and also upon the known reaction of dithiothreitol with the ester linkage in the glycosyl-enzyme intermediate (41). TCEP was found to reduce the disulfide bond effectively at 5 < pH < 7 (42), in which range the fluoroglycosyl-enzyme intermediate has a longer lifetime (24) without cleavage of the ester linkage.

Detection of the Affinity-isolated, Labeled Active-site Peptides by ESI-MS and Their Identification by ESI-MS/MS—The labeled active-site peptides of Bcx and Cex were isolated from the peptic digests of the pure labeled enzymes. The enzymes were first incubated with DNP2FX2SSB until fully inactivated; excess DNP2FX2SSB was removed; and samples were subjected to peptic proteolysis at pH 2.0. Samples were kept at acidic pH throughout the affinity isolation process.
because the fluoroglycosyl-enzyme intermediate is more stable at acidic pH (24). The peptic digest was incubated with streptavidin resin; unbound peptides were washed away, and biotinylated, labeled active-site peptides were eluted via reduction of the disulfide bond with TCEP. The isolated peptides were then analyzed by LC/ESI-MS as shown in Fig. 2C. The m/z values of the isolated labeled peptides were exactly as expected based on previous studies with the non-biotinylated label (39, 43). In the case of Cex, two additional fragments were also observed, most likely due to orifice collisions (Fig. 2C). The sequences of the isolated peptides were unambiguously confirmed by MS/MS fragmentation analysis (Fig. 2D) to be $[^{35}S]$YGWRSPLELY and $[^{258}$VQITEL for Bcx and Cex, respectively, containing the catalytic nucleophiles Glu$^{78}$ (Bcx) and Glu$^{233}$ (Cex) bearing the label as the ester (39, 43).

**Extension to Artificial Proteomes**—To confirm the specificity of the ABPP probe and to demonstrate the mechanism-based nature of the strategy, studies were performed in which one or more xylanases were included in mixtures of other sugar-modifying enzymes. Three model proteomes were constructed: model proteome A included 10 nmol each of native Abg, Bcx, GlvA, and HEWL, and lipopolysaccharyl-$\alpha$-galactosyltransferase C from *N. meningitidis* (LgtC), model proteome B included 10 nmol each of native Abg, Bcx, GlvA, and HEWL, and lipopolysaccharyl-$\alpha$-galactosyltransferase C from *N. meningitidis* (LgtC), model proteome C included 10 nmol each of native Abg, Bcx, GlvA, LgtC, and mutant E78C Bcx; and model proteome D included 10 nmol each of native Abg, Bcx, GlvA, and HEWL. Each of these mixtures was reacted with DNP2FX$_2$SSB and treated exactly as described for the pure enzymes and model proteomes. Of the isolated peptides shown in Fig. 4B, the species at m/z 1084.0, 864.8, and 731.0 (+1) clearly arose from Cex because these are the same peptides seen in Fig. 2C. The labeled peptide at m/z 1055.7 (+1) was a completely new species and did not match the expected labeled active-site peptides of any of the four known retaining $\beta$-1,4-glycanases from *C. fimi* (TABLE ONE). The species at m/z 1055.7 (+1) was therefore subjected to sequence analyses both by MS/MS fragmentation (Fig. 4C) and by Edman degradation. These analyses identified the sequence as VQITEL, with the label attached to the Glu residue. This sequence matches the highly conserved region of the primary amino acid sequence surrounding the catalytic nucleophile in GH family 10 enzymes.

**Cloning the Gene Encoding the New Retaining $\beta$-1,4-Glycanase (Cfx)**—To obtain partial information on the gene (cfx) encoding the new retaining $\beta$-1,4-glycanase (Cfx), we first amplified an internal gene fragment using two primers: the Cfx-AB primer, encoding peptide WDV(V/I)NE, a conserved region around the catalytic acid/base residue of GH family 10 enzymes; and the Cfx-Nu primer, encoding peptide VQITELD (VQITEL identified by the MS-ABPP method plus a conserved Asp residue). The nucleotide sequences of both primers were chosen according to the codon usage of *C. fimi* based on the 16 reported open reading frames. Restriction-digested genomic DNA (rather than the whole genomic DNA) was used as a template because *C. fimi* genes are >70% GC-rich. Indeed, 350-bp DNA fragments were PCR-amplified successfully in four reaction mixtures containing BamHI-, SalI-, NotI-, or XhoI-digested genomic DNA as a template. The PCR product obtained from the NotI-digested genomic DNA template (a NotI site exists in the corresponding gene fragment of Cex) was re-amplified under the same PCR conditions, and the amino acid sequence (deduced from the DNA sequence) confirmed that the PCR fragment encoded part of a GH family 10 xylanase. Two regions (iPCR-1 and iPCR-2) were designed for iPCR based on the deduced amino acid sequence showing no close similarity to Cex and *C. fimi* xylanase C. A PCR fragment from the region upstream of the cfx gene to the middle of the cfx gene was...
amplified from the SacI library only. The full nucleotide sequence was obtained by two more sequential iPCRs using the second primer set (iPCR-3 and iPCR-4) and the third primer set (iPCR-5 and iPCR-6) (Fig. 5). The cfx gene sequence has been deposited in the GenBank™ Data Bank (accession number DQ146941).

**Kinetic Characterization of Cfx**—To investigate the catalytic properties of Cfx, recombinant Cfx was produced in *E. coli* and purified by nickel affinity chromatography. Cfx stability was constant for 10 h over 20–40°C, but rapidly decreased at higher temperatures (Fig. 6A). The steady-state kinetic parameters for hydrolysis of several aryl-β-glyco-

was investigated using pNPX as the substrate over a pH range of 4.0–9.0. The $k_{cat}/K_m$ profile was found to be a classical bell-shaped curve (Fig. 6B), indicating two catalytically essential ionizable groups ($pK_{a1} = 4.5 \pm 0.1$ and $pK_{a2} = 7.6 \pm 0.1$).

**DISCUSSION**

**Active Site-directed Nature of the MS-ABPP Strategy and Endo-type Nature of the Target Enzymes**—Three separate experimental methods point to the active site-directed and mechanism-based nature of the MS-ABPP approach: kinetic analysis showing single exponential decay of the hydrolytic activities (Figs. 2A and 4A), ESI-MS analysis showing 1:1 stoichiometry for the reaction between DNP2FX$_2$SSB and its target enzymes (Fig. 2B), and ESI-MS/MS analysis showing the sequences of the affinity-isolated peptides bearing the catalytic nucleophile glutamic acid derivatized as its ester (Figs. 2D and 4C) (24, 39, 43).

In addition to being the first MS-based ABPP strategy applied to a biological proteome for glycosidase profiling, the approach described here is superior, in terms of specificity and efficacy, to those described previously for glycosidases (15, 46–48). A previous method that made use of quinone methide chemistry produced cross-labeling when carried out in a mixture of proteins because, after initial reaction at the active site of a glycosidase, the glycoside that provided specificity was lost, and a reactive quinone methide was generated, which is known to leave the active site and to react with exposed nucleophilic residues anywhere on the protein or indeed on any protein in the mixture (46, 47). A very recent attempt at proteomic analysis of exoglycosidases using a ligation strategy was more specific, but lacked the desired efficacy (15), with the tagged substrate reacting $10^4$ times more slowly than the parent substrate. By contrast, our reagent (DNP2FX$_2$SSB) reacted at rates (Fig. 2A) that are almost identical to those of the parent armless inactivator (DNP2FX$_2$) (20, 39). This indicates that the biotin arm, attached to the 4'-carbon of the xylitol, does not sterically hinder its access into the active site to any significant extent. This is consistent with the fact that Bcx and Cex are both endoxylanases, with their active sites located in an open cleft (31, 49, 50). In contrast, exo-acting enzymes have a pocket-shaped active site that will not easily accommodate modified sugars, especially those with a bulky biotin arm (15, 50).

## Table One

| Enzyme name (designation) | GH family | CBM family |
|---------------------------|-----------|------------|
| 1,4-β-D-Glucanase D (CenD$^a$) | 5 | 2a |
| Xylanase B (Cex)$^b$ | 10 | 2a |
| Xylanase C (XylC)$^c$ | 10 | 9, 22 |
| Xylanase D (XylD)$^c$ | 11 | 2b |

$^a$ An endoelucelulase with very low xylanase activity (1500-fold less compared with Cex) (27).

$^b$ An endoxylanase with low cellulase activity (25).

$^c$ An endoxylanase.

---

**Figure 4.** Analysis of the *C. fim* secreted proteome by the MS-ABPP strategy. A, time-dependent inactivation kinetics. Experimental data representing the hydrolytic activities of control samples (without DNP2FX$_2$SSB) toward DNPX$_2$ or DNPX over time. The lines are the linear fits. Experimental data representing the hydrolytic activities of samples toward DNPX$_2$ or DNPX over time. The curves are the nonlinear fits of the data to a single exponential decay equation with or without offset. B, ESI mass spectrum showing the affinity-isolated, labeled active-site peptides from the peptic digest of the *C. fim* secreted proteome after treatment with DNP2FX$_2$SSB. All peaks represent [M + H]$^+$ species. The spectrum represents the reverse-phase HPLC elution time of 23.5–23.9 min. calc., calculated. C, ESI-MS/MS fragment ion spectrum arising from the labeled active-site peptide of the new GH family 10 enzyme that was isolated from the *C. fim* secreted proteome.
Identification of a Previously Undiscovered GH Family 10 Glycosidase in the C. fimi Secreted Proteome—The xylanolytic/cellulolytic soil bacterium C. fimi, which produces and secretes a complex array of xylanases and cellulases (25, 26), is an ideal system for the discovery of new retaining β-1,4-glycanases. To date, four such C. fimi retaining enzymes (TABLE ONE) have been cloned and characterized (21, 25, 27–30). However, given the nature of the genetic cloning strategies, it is conceivable that some enzymes were missed (26).

Interestingly, the hydrolytic activity of the C. fimi secreted proteome toward DNPX2 was completely abolished in the presence of DNPG, whereas the hydrolytic activity toward DNPC plateaued at ~16% of its initial value (Fig. 4A). This is consistent with the fact that all known xylanases are retainers, whereas many cellulases are inverters (25). Indeed, many of the known C. fimi cellulosomal enzymes are glyco-

![FIGURE 5. Multiple sequence alignments of the catalytic domains of three GH family 10 enzymes. The domains are from Cfx (the new C. fimi glycanase discovered in this study), S. lividans Xyl10A(Slxyl10A); and S. lividans xylanase; Swiss-Prot accession number P26514), and Cex (a C. fimi glucanase; Swiss-Prot accession number Q59277). Catalytic acid/base and nucleophile residues are marked with inverted triangles, and the amino acid sequences corresponding to the synthesized primers are overlaid.

![FIGURE 6. Temperature stability and pH dependence of Cfx. A, 10-h temperature stability of Cfx; B, pH dependence of kcat/Km for Cfx-catalyzed hydrolysis of pNPsx2.

A few major secreted enzymes are responsible for the observed retaining β-1,4-glycanase activity. One of those is indeed Cex, but the other enzyme with an active-site peptide sequence of VQITEL is a new GH family 10 β-1,4-glycanase previously unidentified in C. fimi, as further confirmed by cloning (Fig. 5). Of the other three known but undetected β-1,4-glycanases (TABLE ONE), Cend has extremely low xylanase activity and is not expected to react with DNPG (27), and XylC is likely intracellular because it lacks the leader sequence typical of secreted prokaryotic proteins (28). XylD is extracellular, but it is probably secreted in minute amounts (below the detection limit of the ABPP methodology) under the growth conditions employed.

Characteristics of Cfx—The deduced amino acid sequence of the Cfx catalytic domain shares ~70% sequence identity with those of xylanases from Streptomyces sp., 52% with Cex, 32% with C. fimi XylC, and 28% with Cellvibrio japonicus XylA. Like Cex and many other GH family 10 enzymes, Cfx has a C-terminal CBM (see afmb.cnrs-mrs.fr/~cazy/
CAZY). Interestingly, the CBM of Cfx displays a high degree of identity to the xylan-binding modules from CBM family 2b, which have been previously found only in family 11 xylanases. Therefore, Cfx is the first example of a GH family 10 enzyme with CBM2b.

Cfx exhibits a substrate specificity expected of GH family 10 enzymes, viz., largely xylanolytic and, to a much lesser extent, also cellulolytic. This is in contrast to GH family 11 enzymes, which are exclusively xylanolytic. As shown in TABLE TWO, comparison of the kinetic parameters of Cfx with those of Cex and *Streptomyces lividans* Xyl110A shows that the kinetic behavior of Cfx and *S. lividans* Xyl110A is very similar and somewhat different from that of Cex. This is not surprising given the high amino acid sequence identity between Cfx and xylanases from *Streptomyces* sp., yet another testament to the fact that high primary sequence similarity often reflects structural and functional similarities.

Finally, the $k_{cat}/K_m$ for Cfx depends on two ionizable groups (in the free enzyme) of $pK_{a1} = 4.5$ and $pK_{a2} = 7.6$, very similar to those of Cex ($pK_{a1} = 4.1$ and $pK_{a2} = 7.7$) (30).

In summary, the specific MS-ABPP strategy described here is a particularly valuable approach for the identification of new β-1,4-glycanases from diverse proteomes or even from metaproteomes. β-1,4-Glycanases are of considerable utility in the pulp and paper industry, in particular valuable approach for the identification of new endoglycosidases.

Acknowledgments—We thank Emily Kwan, Dr. Harry Brumer III, and Dr. R. A. J. Warren for valuable assistance and discussions. We also acknowledge R. A. J. Warren for valuable assistance and discussions. We also acknowledge R. A. J. Warren for valuable assistance and discussions. We also acknowledge R. A. J. Warren for valuable assistance and discussions.
Active-site Peptide "Fingerprinting" of Glycosidases in Complex Mixtures by Mass Spectrometry: DISCOVERY OF A NOVEL RETAINING β-1,4-GLYCANASE IN CELLULOMONAS FIMI

Omid Hekmat, Young-Wan Kim, Spencer J. Williams, Shouming He and Stephen G. Withers

J. Biol. Chem. 2005, 280:35126-35135.
doi: 10.1074/jbc.M508434200 originally published online August 5, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M508434200

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 52 references, 16 of which can be accessed free at http://www.jbc.org/content/280/42/35126.full.html#ref-list-1