Structural and Biochemical Characterization of Glycoside Hydrolase Family 79 β-Glucuronidase from Acidobacterium capsulatum

Mari Michikawa, Hitomi Ichinose, Mitsuru Momma, Peter Biely, Seino Jongkees, Makoto Yoshida, Toshihisa Kotake, Yoichi Tsumuraya, Stephen G. Withers, Zui Fujimoto, and Satoshi Kaneko

Received for publication, January 25, 2012, and in revised form, February 19, 2012 Published, JBC Papers in Press, February 24, 2012, DOI 10.1074/jbc.M112.346288

From the Food Biotechnology Division, National Agriculture and Food Research Organization (NARO) Food Research Institute, 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan, Biomolecular Research Unit, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan, Institute of Chemistry, Slovak Academy of Sciences, Dubravcká cesta 9, SK-845 38 Bratislava 45, Slovak Republic, Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada, Department of Environmental and Natural Resource Science, Tokyo University of Agriculture and Technology, 3-5-8 Saiwai-cho, Fuchu, Tokyo 183-8509, Japan, and Graduate School of Science and Engineering, Saitama University, 255 Shimo-okubo, Sakura-ku, Saitama 338-8570, Japan

In the present study, we characterized a glycoside hydrolase family 79 β-glucuronidase from Acidobacterium capsulatum, which hydrolyzes PNP-β-GlcA approximately 300-fold slower than does the wild-type enzyme, whereas 4-O-methyl-GlcA-containing oligosaccharides are hydrolyzed only 7-fold slower.

β-Glucuronidases hydrolyze β-glucuronic acid (GlcA)-containing carbohydrates to release GlcA and are generally found in microorganisms, plants, and animals. GlcA is a component of proteoglycans, such as chondroitin sulfate proteoglycan, heparan sulfate proteoglycan, and hyaluronan from animals and arabinogalactan proteins from higher plants. Because β-glucuronidases are involved in the metabolism of proteoglycans, they are of biochemical, physiological, and medical interest. β-Glucuronidases have been used as reporter genes in many biological experiments in a manner similar to the use of green fluorescent protein and luciferase. β-Glucuronidases are classified into three glycoside hydrolase (GH) families, GH1, GH2, and GH79, according to their amino acid sequences in the Carbohydrate-Active EnZymes (CAZy) database (1, 2). The GH79 family includes heparanase (EC 3.2.1.166), baiacalin-β-d-glucuronidase (EC 3.2.1.167), 4-O-methyl-β-glucuronidase, and β-glucuronidase (3–7). GH79, along with GH2, belongs to the GH-A clan (1, 2). The first reported crystal structure of a β-glucuronidase was of the human GH2 enzyme (8), which was followed by the report of the Escherichia coli GH2 β-glucuronidase structure (9). The GH2 family consists of exo-acting enzymes with irregular Greek key motifs that is of unknown function.

Structural and Biochemical Characterization of Glycoside Hydrolase Family 79 β-Glucuronidase from Acidobacterium capsulatum

We present the first structure of a glycoside hydrolase family 79 β-glucuronidase from Acidobacterium capsulatum, both as a product complex with β-glucuronic acid (GlcA) and as its trapped covalent 2-fluoroglucuronyl intermediate. This enzyme consists of a catalytic (β/α)8-barrel domain and a β-domain with irregular Greek key motifs that is of unknown function. The enzyme showed β-glucuronidase activity and trace levels of β-glucosidase and β-xylanosidase activities. In conjunction with mutagenesis studies, these structures identify the catalytic residues as Glu173 (acid base) and Glu287 (nucleophile), consistent with the retaining mechanism demonstrated by 1H NMR analysis. Glu45, Tyr243, Tyr292–Gly294, and Tyr334 form the catalytic pocket and provide substrate discrimination. Consistent with this, the Y292A mutation, which affects the interaction between the main chains of Gln293 and Gly294 and the GlcA carboxyl group, resulted in significant loss of β-glucuronidase activity while retaining the side activities at wild-type levels. Likewise, although the β-glucuronidase activity of the Y334F mutant is ~200-fold lower (kcat/Km) than that of the wild-type enzyme, the β-glucosidase activity is actually 3 times higher and the β-xylanosidase activity is only 2.5-fold lower than the equivalent parameters for wild type, consistent with a role for Tyr334 in recognition of the C6 position of GlcA. The involvement of Glu355 in discriminating against binding of the O-methyl group at the C4 position of GlcA is revealed in the fact that the E45D mutant hydrolyzes PNP-β-GlcA approximately 300-fold slower (kcat/Km) than does the wild-type enzyme, whereas 4-O-methyl-GlcA-containing oligosaccharides are hydrolyzed only 7-fold slower.

Background: The three-dimensional structures of β-glucuronidase have been solved only for the GH2 enzymes.

Results: AcGlcA79A is composed of a (β/α)8-barrel domain and a β-domain.

Conclusion: The substrate binding site of AcGlcA79A is adapted for recognition of GlcA as a substrate.

Significance: This is the first report describing the crystal structure, mechanism, and catalytic residues of a GH79 enzyme.

The atomic coordinates and structure factors (codes 3VNY, 3VNZ, and 3VO0) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

1. These authors contributed equally to this work.
2. Supported by the Natural Sciences and Engineering Research Council, Canada.
3. To whom correspondence may be addressed. Tel./Fax: 81-29-838-7877; E-mail: zui@affrc.go.jp.
4. To whom correspondence may be addressed. Tel.: 81-29-838-8022; Fax: 81-29-838-7996; E-mail: sakaneko@affrc.go.jp.

The abbreviations used are: GlcA, glucuronic acid; AcGlcA79A, β-glucuronidase from Acidobacterium capsulatum; GH, glycoside hydrolase; MeGlcA, 4-O-methyl-GlcA; PNP, p-nitrophenol; PNP-β-GlcA, PNP-β-glucopyranosyl; PNP-β-Xyl, PNP-β-xyloroxyranosyl; PNP-β-GlcA, PNP-β-glucuronide; MeGlcA-β-1,6-Gal, MeGlcA-β-1,6-Gal-β-1,6-Gal; DNP-2FGlcA, 2′,4′-dinitrophenyl 2-deoxy-2-fluoro-β-o-glucuronide.

This article contains supplemental Experimental Procedures and Figs. S1–S3.

The atomic coordinates and structure factors (codes 3VNY, 3VNZ, and 3VO0) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).
**β-Glucuronidase from A. capsulatum**

enzymes, such as β-galactosidases, β-mannosidases, and β-glucuronidases, whereas the GH79 family consists of both *exo*-acting (β-glucuronidase) and *endo*-acting (heparanase) enzymes. Human heparanase plays a decisive role in disease-related processes, such as cell invasion, angiogenesis, and cancer metastasis. Several structural models of heparanase have been generated by computer programs using known clan GH-A structures as templates (10–12). However, to date, no experimental structure or even a preliminary crystallization report of a GH79 enzyme is available; the three-dimensional structure of a GH79 enzyme is eagerly anticipated.

For many years, bacterial β-glucuronidases were believed to be restricted to *E. coli* and closely related Enterobacteriaceae. Recently, however, the enzyme activity has been found among non-enterobacterial, anaerobic residents of the digestive tract, such as *Bacteroides* and *Clostridium* (13–15). Although these species exhibit lower β-glucuronidase activity than *E. coli*, they are ~100-fold more abundant. Hence, it is suggested that they make a significant overall contribution to the enterohepatic circulation. Other bacterial genera with β-glucuronidase activity are *Streptococcus*, *Staphylococcus*, *Corynebacterium*, *Alcaligenes*, and many soil bacteria (16). In these bacteria, the enzyme is involved in carbohydrate and energy metabolism and may contribute to invasion during bacterial pathogenesis.

In contrast, GH79 β-glucuronidases from *Arabidopsis thaliana*, *Aspergillus niger*, and *Neurospora crassa* have been reported to act on plant proteoglycans (5–7). The *A. niger* and *N. crassa* enzymes have been shown to release both GlcA and 4-O-methyl-GlcA (MeGlcA) from arabinogalactan proteins (6, 7). No GH2 β-glucuronidases are known in plants, and no GH2 β-glucuronidase homologous sequences have been found in plants (17). Therefore, GH79 enzymes appear to play important roles in the metabolism of plant proteoglycans.

Thus, β-glucuronidases play an important role in plants, animals, and probably microorganisms. In this study, we present the three-dimensional structure of β-glucuronidase from *Acidobacterium capsulatum* (AcGlcA79A); this is the first structure of a GH79 enzyme. We also performed mutagenesis studies based on the structure of the enzyme-GlcA complex. Our results clearly demonstrate the structure-function relationship of GH79 β-glucuronidase and will help guide studies of other GH79 enzyme systems.

**EXPERIMENTAL PROCEDURES**

*Substrates*—*p*-Nitrophenyl (PNP)-glycosides including PNP-β-glucopyranoside (PNP-β-Glc), PNP-β-xylpyranoside (PNP-β-Xyl), and PNP-β-glucuronide (PNP-β-GlcA) were purchased from Sigma. MeGlcA-β-1,6-Gal-β-1,6-Gal (MeGlcA-β-1,6-Gal) was prepared as described previously (6), and 2′,4′-dinitrophenyl-2-deoxy-2-fluoro-β-D-glucuronide (DNPGlcA) was synthesized according to a method described previously (18).

*Expression of AcGlcA79A and Mutant Generation*—*A. capsulatum* NBRC15755 was obtained from the National Institute of Technology and Evaluation (Kazusa, Japan). The gene encoding a putative β-glucuronidase (ACP_2665; GenBank™ accession number ACO32043) was amplified from *A. capsulatum* genomic DNA by PCR using Phusion DNA polymerase (Finnzymes, Espoo, Finland) and the following primers: for-

| Data collection and structure refinement statistics of AcGlcA79A |  |
|---|---|
| **AcGlcA79A complex** |  |
| **Space group** |  |
| **a** | 217.9 |
| **b** | 101.2 |
| **c** | 217.8 |
| **Wavelength (Å)** | 1.0000 |
| **Resolution (Å)** | 50.0-1.50 |
| **Completeness (%)** | 100.0 |
| **Multiplicity** | 14.6 |
| **Average I/σ(I)** | 47.0 |
| **Unique reflections** | 90,146 |
| **Observed reflections** | 1,312,592 |
| **Resolution (Å)** | 30.6-1.50 |
| **R**<sub>i</sub> | 0.172 |
| **R**<sub>f</sub> | 0.190 |
| **Allowed region (%)** | 1.3 |
| **No. of water molecules** | 505 |
| **Average B** | 16.5 |
| **Allowed region (%)** | 1.9 |
| **No. of water molecules** | 351 |
| **Average B** | 22.1 |
| **Allowed region (%)** | 1.7 |
| **No. of water molecules** | 366 |
| **Average B** | 24.5 |

*Note:* The table data were obtained from the Protein Data Bank (PDB) entry 4122.
ward, 5′-CAT ATG GCT TTT GCC CGC GGC GGA CTG GCT-3′; reverse, 5′-CAT ATG GCT TTT GCC CGC GGA TTC GAG CAA TGC GCC GGA-3′. The amplified DNA was cloned into pET30 (+) (Novagen, Darmstadt, Germany) at Ndel and HindIII restriction enzyme sites (underlined). Recombinant enzymes were expressed using the T7 expression system in E. coli BL21-Gold(DE3) (Novagen) and purified with a C-terminal histidine tag (supplemental Fig. S1). Amino acid substitutions in AcGlcA79A were generated by inverse PCR using the expression vector pET30/AcGlcA79A as template DNA and the appropriate primers. Expression of mutants and their purification were performed in the same way as for wild-type AcGlcA79A. Further details can be found in the supplemental Experimental Procedures.

Crystalization, Data Collection, and Structural Determination—Initial crystallization screening of AcGlcA79A was conducted by the sitting drop, vapor diffusion method at 20 °C, mixing 0.3 μl of the protein solution (4.6 mg ml⁻¹) and an equal volume of precipitant against 50 μl of reservoir solution using the following commercially available kits: JCSG + Suite (Qiagen, Düsseldorf, Germany), Crystal Screen HT, and Index HT (Hampton Research, Aliso Viejo, CA). After refinement of the crystallization conditions, AcGlcA79A was crystallized by the sitting drop, vapor diffusion method with a precipitant solution composed of 2.0 M sodium phosphate monobasic monohydrate/potassium phosphate dibasic (0.5:9.5 (v/v), pH not adjusted) and with a protein concentration of 2.5 mg ml⁻¹. Crystals with maximum dimensions of 200 × 200 × 500 μm were consistently obtained within a few days at 20 °C. Selenomethionine (Se-Met)-labeled AcGlcA79A was produced using the E. coli B834(DE3) methionine auxotroph and was crystallized under the same conditions as for the native enzyme. The GlcA complex was prepared by soaking the AcGlcA79A crystals in crystallization liquor containing GlcA powder for 5 min. The fluorinated glucuronide (2FGlcA) intermediate complex was prepared by adding the crystallization liquor containing 0.3% (w/v) DNP-2FGlcA into the AcGlcA79A crystallization drop and incubating for 5 min.

Diffraction experiments were conducted at the Photon Factory (PF) or the PF-Advanced Ring (PF-AR), High Energy

### TABLE 2

| Substrates | PNP-β-Glc | PNP-β-Xyl | PNP-β-Glc | PNP-β-Xyl |
|------------|-----------|-----------|-----------|-----------|
| WT         | 5.6 ± 0.1 | 2.3 ± 0.2 | 4.3 ± 0.2 | 2.0 ± 0.1 |
| E173G      | 5.6 ± 0.1 | 2.3 ± 0.2 | 4.3 ± 0.2 | 2.0 ± 0.1 |
| E173A      | 5.6 ± 0.1 | 2.3 ± 0.2 | 4.3 ± 0.2 | 2.0 ± 0.1 |
| E287G      | 5.6 ± 0.1 | 2.3 ± 0.2 | 4.3 ± 0.2 | 2.0 ± 0.1 |
| Y334F      | 5.6 ± 0.1 | 2.3 ± 0.2 | 4.3 ± 0.2 | 2.0 ± 0.1 |
| Y292A      | 5.6 ± 0.1 | 2.3 ± 0.2 | 4.3 ± 0.2 | 2.0 ± 0.1 |
| H327K      | 5.6 ± 0.1 | 2.3 ± 0.2 | 4.3 ± 0.2 | 2.0 ± 0.1 |
| H327T      | 5.6 ± 0.1 | 2.3 ± 0.2 | 4.3 ± 0.2 | 2.0 ± 0.1 |
| E45Q       | 5.6 ± 0.1 | 2.3 ± 0.2 | 4.3 ± 0.2 | 2.0 ± 0.1 |

Each set of results is based on the average of three independent measurements. Results are means ± S.E. ND, not detectable.

### TABLE 3

| Activity of wild type (WT) and mutants of AcGlcA79A against PNP-β-Glc | Relative to WT |
|------------------------------------------------------------------------|--------------|
| kₘ (nM) | kₐ (s⁻¹) | kₐ/kₘ (nM⁻¹s⁻¹) | Relative to WT |
|---------|---------|-----------------|----------------|
| WT      | 0.015 ± 0.001 | 34 ± 1 | 2303 | 1 |
| E173G   | 0.5 ± 0.1   | 2.0 × 10⁻⁴  | 1  |
| E173A   | 0.1²     | 5.6 × 10⁻⁵  | 1  |
| E287G   | 0.2²     | 9.1 × 10⁻⁵  | 1  |
| Y334F   | 5.7 ± 0.5  | 5.2 × 10⁻⁴  | 1  |
| Y292A   | 0.55²   | 2.4 × 10⁻⁴  | 1  |
| H327N   | 0.030 ± 0.001 | 31 ± 1 | 1026 | 0.5 |
| H327S   | 5.5 ± 0.4  | 3.5 × 10⁻⁴  | 1  |
| E45D    | 4.8 ± 0.6  | 2.5 × 10⁻⁴  | 1  |

The enzyme (0.1 μM) was incubated with PNP-β-GlcA at a concentration of 1.5 μM. Kinetic parameters were not determined because mutants were too inactive to obtain individual kinetic constants. The activity of the rest of mutants such as E173Q, E287A, Y334W, Y292A, H327K, H327T, and E45Q was not detected under this condition.

The enzyme (0.1 μM) was incubated with PNP-β-Glc at a concentration of 1.5 μM. Kinetic parameters were not determined because mutants were too inactive to obtain individual kinetic constants. The activity of the rest of mutants such as E173Q, E287A, Y334W, Y292A, H327K, H327T, and E45Q was not detected under this condition.

Each set of results is based on the average of three independent measurements. Results are means ± S.E. ND, not detectable.

[^1]: The enzyme (0.1 μM) was incubated with PNP-β-Glc at a concentration of 1.5 μM. Kinetic parameters were not determined because mutants were too inactive to obtain individual kinetic constants. The activity of the rest of mutants such as E173Q, E287A, Y334W, Y292A, H327K, H327T, and E45Q was not detected under this condition.
Accelerator Research Organization, Tsukuba, Japan. The crystals were moved into the mother liquor containing 20% (v/v) glycerol as a cryoprotectant, and a single crystal was scooped in a nylon loop and flash frozen in a nitrogen gas stream at −178 °C. Diffraction data were collected with the Quantum 270 CCD detector (Area Detector Systems Corp., Poway, CA). Data were integrated and scaled using the programs DENZO and Scalepack in the HKL2000 program suite (19). 

Structural analysis of AcGlcA79A was conducted through exchange chromatography with pulsed amperometric detection (20); this corresponded to the 55% solvent content of the crystals.

Structural analysis of AcGlcA79A was conducted through the multiwavelength anomalous dispersion method using SeMet-labeled AcGlcA79A crystals. Seven selenium atom positions were determined, and initial phases were calculated using the program AutoSol and AutoBuild wizards in the PHENIX program suite (21). The obtained initial model was further improved using Coot and Refmac5 (23, 24). One phosphate ion and several glycerol molecules were added into the model.

For the analyses of GlcA- and 2FGlcA-binding structures of AcGlcA79A, structural determination was conducted twice from D2O before use. A solution of 15 mM PNP-β-GlcA molecule, the Matthews coefficient was calculated to be 2.7 Å³ Da⁻¹ (20); this corresponded to the 55% solvent content of the crystals.

The substrate specificity of AcGlcA79A toward various PNP-glycosides was determined. The assay method was identical to that described for PNP-β-GlcA. The kinetic parameters of the wild type and mutants of AcGlcA79A for PNP-β-GlcA, PNP-β-Glc, and PNP-β-Xyl were determined as follows. The reactions were performed in McIlvaine buffer (pH 3.5) containing 0.01–10 mM substrates, 0.1% (w/v) bovine serum albumin, and 0.9 nm–10.0 mM enzyme at 37 °C for up to 10 min. The amount of PNP released was determined from the A₄₀₀ data. The kinetic parameters kcat and Km were determined by Lineweaver-Burk plot from three independent experiments and at five substrate concentrations. The substrate specificity and the catalytic efficiency of the wild type and mutants of AcGlcA79A were analyzed using MeGlcA-β,1,6-Gal₂. Briefly, the enzyme (5 μM) was incubated with the substrate (10 μM) in McIlvaine buffer (pH 3.5) at 37 °C. At regular time intervals, the amount of degradation of each substrate was quantified by high performance anion exchange chromatography with pulsed amperometric detection (29). The assay was performed in duplicate.

**RESULTS AND DISCUSSION**

**Expression and Characterization of AcGlcA79A**—The cloning and expression of AcGlcA79A are described in the supplemental Experimental Procedures. Recombinant AcGlcA79A activities were tested using various PNP-glycosides as substrates. AcGlcA79A showed significant activity only for PNP-β-GlcA with negligibly weak or no activity for other substrates. Using PNP-β-GlcA as the substrate, maximal enzyme activity was detected at pH 3.0 and 50 °C. The enzyme was stable between pH 2.8 and 4.5 at 30 °C for 1 h and at pH 3.0 at <55 °C for 1 h. The Km and kcat values of AcGlcA79A for PNP-β-GlcA were 0.015 ± 0.001 mM and 34 ± 1 s⁻¹, respectively (Tables 2 and 3). These values of AcGlcA79A are similar to the previously reported values of GH79 enzymes (7). The Km values of the A. niger and N. crassa recombinant GH79 enzymes were 0.030 and 0.038 mM, respectively, and the kcat values were 26.9 and 13.5 s⁻¹, respectively (7). Because the A. niger and N. crassa enzymes showed 4-O-methyl-β-glucuronidase activity, the activity of AcGlcA79A for the MeGlcA-containing oligosaccharide MeGlcA-β,1,6-Gal₂ was tested. AcGlcA79A barely hydrolyzed MeGlcA-β,1,6-Gal₂, and its activity for MeGlcA-β,1,6-Gal₂ was 10³ times lower than that for PNP-β-GlcA (Table 4). These results suggest that AcGlcA79A is a genuine β-glucuronidase.

| TABLE 4 |
|Activity of AcGlcA79A wild type and E45D mutant for PNP-β-GlcA and MeGlcA-β,1,6-Gal₂ |
|---|---|---|---|---|
| | PNP-β-GlcA | MeGlcA-β,1,6-Gal₂ | Ratio* (MeGlcA/GlcA) | Relative to WT |
| Wild type | 2.303 ± 355 | 7.7 × 10⁻¹ | 1.4-Å resolution (space group A_4₁2₂) |
| E45D | 7.5 ± 1.7 | 1.1 × 10⁻¹ | 3.0 × 10⁻⁶ | 43 |

*Ratio (MeGlcA/GlcA) = (kcat/Km) for MeGlcA-β,1,6-Gal₂)/(kcat/Km for PNP-β-GlcA).
Three-dimensional Structure of AcGlcA79A—The crystal structure of AcGlcA79A was determined by the multiwavelength anomalous dispersion method using Se-Met derivative data. Native and GlcA complex structures were determined successively. Structure refinement statistics are summarized in Table 1. The quality and accuracy of the final structures were further demonstrated to show that all residues fall within the common regions of the Ramachandran stereochemistry plot. Recombinant AcGlcA79A is composed of a single polypeptide chain of 488 amino acids with the additional C-terminal residues (476KLAAALEHHHHHH488) derived from the expression vector and purification tag. Seventeen N-terminal residues (Met1–Ser17) and five C-terminal residues (His484–His488) were not identified because of a lack of electron density. Three cis-peptide bonds were found at Gly214-Pro215, Gly246-Pro247, and Ser457-Gly458. The final model consists of one AcGlcA79A.

**FIGURE 1. Structure of AcGlcA79A.** A, stereoview of the ribbon model of AcGlcA79A-GlcA complex in rainbow-ordered colors. Two catalytic residues are displayed in red. The bound GlcA molecule and phosphate ion are shown as stick models. B, schematic topological diagram of AcGlcA79A. α-Helices are shown as cylinders, α3_10-helices are shown as shaded cylinders, and β-strands are shown as arrows. The colors correspond to those in A. Two catalytic residues are labeled in red.
molecule with one or two phosphate ions and one or four glucose molecules. The overall structures of the ligand-free and GlcA complexes were almost the same, and the calculated root mean square difference was 0.2 Å for the Ca atom pairs.

The protein consists of domains A and C (Fig. 1). The N-terminal β-strand (Val820–Ile835) is inserted into the domain C, and the peptide folds into the (β/α)8-barrel of domain A (Gly41–Ala360) and finally enforces the β-domain of domain C. The secondary structure elements of AcGlcA79A are numbered in Fig. 1B.

Domain A consists of a (β/α)8-barrel, which is the catalytic domain in many glycoside hydrolases. The GlcA binding site is located at the C-terminal end of the central β-barrel and was designated as the catalytic pocket. GH79 and GH2 enzymes are classified into the same GH-A clan, and the overall folding of their structures shows certain similarities. The root mean square difference for the core Ca atoms between the catalytic domains of AcGlcA79A and E. coli GH2 β-glucuronidase was calculated to be 2.6 Å. In addition to the core secondary structures of the (β/α)8-barrel, several loops fold into some small substructures and contribute to the formation of the catalytic pocket (Fig. 1).

Domain C consists of nine β-strands with strands Cβ0 and Cβ4 aligned parallel and the others in interacting antiparallel configurations (Fig. 1B). The eight strands other than Cβ0 comprise a typical antiparallel β-domain structure containing Greek key motifs (Fig. 1). This C-terminal domain structure is observed in many glycoside hydrolases, although in many cases its function remains unknown.

Active Site and Catalytic Mechanism—The active site of AcGlcA79A is located at the C-terminal end in the central (β/α)8-barrel and is represented by a pocket shape (Fig. 1). In the GlcA complex structure, Glu173 is located near the O1 atom of GlcA (2.3–2.4 Å), and Glu287 is located close to the C1 atom (3.0 Å). The distance between Glu173 and Glu287 of AcGlcA79A is 5.25 Å. The location of these acidic amino acids suggests that Glu173 and Glu287 are the catalytic residues of AcGlcA79A, functioning as acid/base and nucleophile, respectively, consistent with the expected retaining mechanism (Fig. 2). It has been reported that the distance between two catalytic residues is 7–10 Å in inverting glycosyl hydrolases and 5–6 Å in retaining glycosyl hydrolases (30). This is also apparent from the rule of the clan GH-A (4/7 superfamily) catalytic module in which the two catalytic residues are located posteriorly in Aβ4 and Aβ7 (31). The locations of Glu173 and Glu287 of AcGlcA79A are the same as those of the catalytic residues of GH2 enzymes and correspond to the 4/7 superfamily rule.

To verify the identities of the catalytic residues of AcGlcA79A, Gly, Ala, and Gln mutants of Glu173 and Glu287 were constructed. As expected, the activities of all mutants were at least 104 times lower than those of the wild-type enzyme (Table 2). The kcat/Km values of E173G, E173A, and E287G mutants were 0.5, 0.1, and 0.2 mM−1s−1, respectively. Chemical rescue experiments were subsequently performed for E173G, E173A, E287G, and E287A using formic acid and acetic acid as exogenous nucleophiles (see supplemental Experimental Procedures). The β-glucuronidase activities of the mutants were measured in McIlvaine buffers (0.1 M citric acid and 0.2 M Na2HPO4) of various pH values with or without 90 mM formic acid and 90 mM acetic acid (supplemental Fig. S2). The pH activity profiles of E173G and E173A were the same with or without the chemicals, and the enzyme activities of the mutants remained significantly lower than that of the wild type. However, the pH activity profiles of E287G and E287A with formic acid were different in the pH range 2–5. The enzyme activity of E287G with formic acid near the optimum pH (2.6–3.5) recovered completely to the same level as that of the wild type. For E287A, 25% of the enzyme activity was recovered. In contrast, acetic acid did not affect the enzyme activities of E287G and E287A most likely because acetic acid is too large to enter the pocket created by the mutation. These data further suggest that Glu287 is the nucleophile.

Because the catalytic mechanism of GH79 enzymes had not been verified experimentally, the anomeric configuration of the hydrolysis product was determined (Fig. 3). The hydrolysis of PNP-β-GlcA with AcGlcA79A in D2O was followed by 1H NMR analysis. The enzyme load was sufficient to complete the hydrolysis of 15 mM substrate in about 4 min. As shown in Fig. 3, the disappearance of the H1 signal of the glucuronide at 5.3 ppm was accompanied by parallel changes of the PNP aglycone signals into PNP signals (signals between 6.8 and 8.3 ppm). Free uronic acid was generated as the β-anomer as evidenced by the rapid appearance of the H1β resonance at 4.7 ppm. The α-anomer signal only appeared later as a result of mutarotation (Fig. 4 and supplemental Fig. S3). These results show that AcGlcA79A is a retaining enzyme utilizing a double displacement mechanism of hydrolysis. It is very probable that all enzymes belonging to this family will prove to be retaining. Consistent with this mechanism, we detected transglycosylation products generated by AcGlcA79A (data not shown). Similar transglycosylation reactions catalyzed by GH79 β-glucuronidase from A. niger have been reported previously (6, 7).
Ligand-binding Structure—The catalytic center of AcGlcA79A is pocket-shaped as is typically found in exo-acting enzymes (Figs. 2, A and B, and 4). In the GlcA complex structure, the aromatic ring of Tyr243 lies parallel to the pyranosyl ring of GlcA, providing stabilizing hydrophobic interactions. The O2 and O3 atoms of GlcA are recognized by hydrogen bonds, such as those between the O2 and Asn172-N at O2 atoms and between the O3 and both Glu45-O at O2 and Asn80-N atoms. The O4 atom of GlcA is surrounded by Pro104 and His327 and forms a hydrogen bond with Asp105. Recognition of the GlcA carboxyl group is achieved by hydrogen bonds between O6A and Gln293-N and between O6B and both Gly294-N and Tyr334-O (Fig. 2). Fig. 4C shows a cross-sectional view of the active site pocket of the GlcA complex structure and clearly indicates that there is no extra space around the substrate. This was also indicated by a sugar soaking experiment in which no residual sugar was observed when the AcGlcA79A crystals were soaked with Glc, Xyl, or GaA.

2-Deoxy-2-fluoroglycosides, which act as mechanism-based inhibitors to form covalent intermediates, are powerful chemical tools for identifying the active site nucleophile in retaining glycosidases as reported by Withers and Aebersold (18). By adding DNP-2F-GlcA to the AcGlcA79A crystal, the fluorinated glucuronide residue (2F-GlcA) was observed to be covalently bound to Glu287 (Fig. 2C) through an α-con-
**β-Glucuronidase from A. capsulatum**

figured linkage. The positions of the O2 and O3 atoms of 2FGlcA are almost the same as with GlcA, but the C1 atom is shifted by 1.3 Å toward Glu287 because of covalent bond formation (Fig. 2D).

**Mutagenesis Study of AcGlcA79A**—To investigate the basis of substrate specificity of AcGlcA79A, appropriate mutations were introduced at the residues that interact with GlcA, and the properties of the constructed mutants were characterized (Tables 2 and 3). Although GlcA and Glc are structurally different only at the C6 position (−COOH for GlcA and −CH2OH for Glc), AcGlcA79A discriminates greatly between GlcA and Glc. To address this discrimination, a Y334F mutant was constructed along with Y334W and Y292A mutants to alter the space around the C6 position of GlcA and change the position of the main chain of Gln293 and Gly294. Around the C6 position of GlcA, the main chain atoms of Gln293 and Gly294 and Tyr334−Oη interact with the carboxyl group of GlcA (Fig. 2).

As shown in Table 3, the catalytic efficiency of the Y334F mutant for hydrolysis of PNP-β-GlcA is lower than that of the wild type due to an increased $K_m$ value. In contrast, the activity for PNP-β-Glc is close to that of the wild type. Consequently, the activity ratio for PNP-β-Glc to PNP-β-GlcA (Glc/GlcA ratio) is 567 times higher than that of the wild type, strongly suggesting that Glu45 is one of the key residues by which AcGlcA79A distinguishes the 4-O-methyl group of GlcA.

The substrate binding site of AcGlcA79A is specialized for recognition of GlcA as a substrate. A mutagenesis study revealed that Tyr334 and Tyr292 interact with the C6 position of GlcA and that Glu45 recognizes the C4 position of GlcA. Single amino acid mutations did not drastically change the substrate specificity because part of the substrate binding site was formed by the main chain atoms of the protein. These results will help guide studies of other GH79 enzyme systems.

**Acknowledgments**—We thank the beamline researchers at Photon Factory for assistance with synchrotron radiation. We also thank Dr. Ivetta Uhlířiová from the Institute of Slovak Academy of Sciences for assistance with the NMR experiment and Yuan Liu from the Tokyo University of Agriculture and Technology for assistance with the constructions of expression vectors for the mutants.

**REFERENCES**

1. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* **37**, D233–D238
2. Henrissat, B., and Davies, G. (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7**, 637–644
3. Parish, C. R., Freeman, C., and Hulett, M. D. (2001) Heparanase: a key enzyme involved in cell invasion. *Biochim. Biophys. Acta* **1471**, M99–M108
4. Sasaki, K., Taura, F., Shoyama, Y., and Morimoto, S. (2000) Molecular characterization of a novel β-glucuronidase from *Scutellaria baicalensis* Georgi. *J. Biol. Chem.* **275**, 27466–27472
5. Eudes, A., Mouille, G., Thévenin, J., Goyallon, A., Minic, Z., and Jouanin, L. (2008) Purification, cloning and functional characterization of an endogenous β-glucuronidase in *Arabidopsis thaliana*. *Plant Cell Physiol.* **49**, 1331–1341
6. Kuroyama, H., Tsutsui, N., Hashimoto, Y., and Tsumuraya, Y. (2001) Purification and characterization of a β-glucuronidase from *Aspergillus niger*. *Carbohydr. Res.* **333**, 27–39
7. Konishi, T., Kotake, T., Soraya, D., Matsuoka, K., Koyama, T., Kaneko, S., Igarashi, K., Samejima, M., and Tsumuraya, Y. (2008) Properties of family 79 β-glucuronidases that hydrolyze β-glucuronosyl and 4-O-methyl-β-glucuronosyl residues of arabinogalactan-protein. *Carbohydr. Res.* **343**, 1191–1201
8. Jain, S., Drendel, W. B., Chen, Z. W., Mathews, F. S., Sly, W. S., and Grubb, J. H. (1996) Structure of human β-glucuronidase reveals candidate lypo-
somal targeting and active-site motifs. Nat. Struct. Biol. 3, 375–381
9. Wallace, B. D., Wang, H., Lane, K. T., Scott, J. E., Orans, J., Koo, J. S., Venkatash, M., Jobin, C., Yeh, L. A., Mani, S., and Redinbo, M. R. (2010) Alleviating cancer drug toxicity by inhibiting a bacterial enzyme. Science 330, 831–835
10. Hulett, M. D., Hornby, J. R., Ohms, S. J., Zuegg, J., Freeman, C., Gready, J. E., and Parish, C. R. (2009) Identification of active-site residues of the pro-metastatic endoglycosidase heparanase. Biochemistry 39, 15659–15667
11. Sapay, N., Cabannes, E., Petitou, M., and Imberty, A. (2012) Molecular model of human heparanase with proposed binding mode of a heparan sulfate oligosaccharide and catalytic amino acids. Biopolymers 97, 21–34
12. Gandhi, N. S., Freeman, C., Parish, C. R., and Mancera, R. L. (2000) Identification of active-site residues of the pro-metastatic endoglycosidase heparanase. Glycobiology 25, 35–55
13. McBay, A. J., and Macfarlane, G. T. (1998) Ecological and physiological studies on large intestinal bacteria in relation to production of hydrolytic and reductive enzymes involved in formation of genotoxic metabolites. J. Med. Microbiol. 47, 407–416
14. Nakamura, J., Kubota, Y., Miyaoka, M., Saitoh, T., Mizuno, F., and Benno, Y. (2002) Comparison of four microbial enzymes in Clostridia and Bacteroides isolated from human feces. Microbiol. Immunol. 46, 487–490
15. Davé, M., McCrae, S. I., Stevens, V. J., Duncan, S. H., and Louis, P. (2008) Distribution of β-glucosidase and β-glucuronidase activity and of β-glucuronidase gene gus in human colonic bacteria. FEMS Microbiol. Ecol. 66, 487–495
16. Ritz, K., Dighton, J., and Giller, K. E. (1994) Beyond the Biomass: Compositional and Functional Analysis of Soil Microbial Communities, pp. 149–156, John Wiley and Sons Ltd., Chichester, UK
17. Arul, L., Benita, G., Sudhakar, D., Thayumanavan, B., and Balasubramanian, P. (2008) β-Glucuronidase of family-2 glycosyl hydrolase: a missing member in plants. Bioinformation 3, 194–197
18. Withers, S. G., and Aebersold, R. (1995) Approaches to labeling and identification of active site residues in glycosidases. Protein Sci. 4, 361–372
19. Orwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326
20. Matthews, B. W. (1968) Solvent content of protein crystals. J. Mol. Biol. 33, 491–497
21. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Elofsson, A., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221
22. Cohen, S. X., Morris, R. J., Fernandez, F. I., Ben Jelloul, M., Kukowski, M., Parthasarathy, V., Lamzin, V. S., Kleywegt, G. J., and Perrakis, A. (2004) Towards complete validated models in the next generation of ARP/wARP. Acta Crystallogr. D Biol. Crystallogr. 60, 2222–2229
23. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132
24. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. 53, 240–255
25. Schüttelkopf, A. W., and van Aalten, D. M. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr. D Biol. Crystallogr. 60, 1355–1363
26. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Ca geometry: ϕ, ψ and Cβ deviation. Proteins 50, 437–450
27. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242
28. Ichinose, H., Yoshida, M., Fujimoto, Z., and Kaneko, S. (2008) Characterization of a modular enzyme of exo-1,5-α-arabinofuranosidase and arabinan binding module from Streptomyces avermitilis NRBC14893. Appl. Microbiol. Biotechnol. 80, 399–408
29. Ichinose, H., Yoshida, M., Kotake, T., Kuno, A., Igarashi, K., Tsurumaya, Y., Samejima, M., Hirabayashi, J., Kobayashi, H., and Kaneko, S. (2005) An exo-β-1,3-galactanase having a novel β-1,3-galactan-binding module from Phanerochaete chrysosporium. FEBS Lett. 579, 323–328
30. Davies, G., and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. Structure 3, 853–859
31. Jenkins, J., Lo Leggio, L., Harris, G., and Pickersgill, R. (1995) β-Glucosidase, β-galactosidase, family A cellulases, family F xylanases and two barley γ-glycerinases form a superfamily of enzymes with 8-fold βα architecture and with two conserved glutamates near the carboxyl-terminal ends of β-strands four and seven. FEBS Lett. 328, 281–285