Mechanism-based fluorescent labeling of β-galactosidases: An efficient method in proteomics for glycoside hydrolases

Masaki Kurogochi,* Shin-Ichiro Nishimura,*a,b and Yuan Chuan Lee*c

aDivision of Biological Sciences, Graduate School of Science, Frontier Research Center for Post-Genomic Science and Technology, Hokkaido University, N21, W11, Sapporo 001-0021, Japan

bNational Institute of Advanced Industrial Science and Technology (AIST), Sapporo 062-8517, Japan

cBiology Department, Johns Hopkins University, 3400, N. Charles St., Baltimore, MD 21218, USA

*Corresponding authors: Dr. Shin-Ichiro Nishimura, Hokkaido University, e-mail: shin@glyco.sci.hokudai.ac.jp and Dr. Yuan Chuan Lee, Johns Hopkins University, e-mail: yclee@jhu.edu
ABSTRACT

(4-N-Dansyl-2-difluoromethylphenyl)-β-D-galactopyranoside was synthesized, and successfully tested on β-galactosidases from *Xanthomonas manihotis* [Wong-Madden, *Glycobiology* (1995) 5, 19-28; Taron, *Glycobiology* (1995) 5, 603-610], *E. coli* [Jacobson, *Nature* (1994) 369, 761-766], and *Bacillus circulans* [Fujimoto, *Glycoconjugate J.* (1988) 15, 155-160] for the rapid identification of the catalytic site. Reaction of the irreversible inhibitor with enzymes proceeded to afford a fluorescent labeled protein suitable for further high-throughput characterization by using anti-dansyl antibody and MALDI-TOF/TOF. Specific probing by a fluorescent aglycon greatly facilitated identification of the labeled peptide fragments from β-galactosidases. It was demonstrated by using *Xanthomonas manihotis* β-galactosidase that the Arg-58 residue, which is located within a sequence of \(^{56}\text{IPRAYWKD}^{63}\), was labeled by nucleophilic attack of the guanidinyl group. This sequence including Arg-58 [Leu-46 to Tyr-194] was similar to that [Met-1 to Tyr-151] of *Thermus thermophilus* A4 which is the first known structure of glycoside hydrolases family 42 [Hidaka, M. (2002) *J. Mol. Biol.* 322, 79-91]. A catalytic glutamic acid (Glu-537) of *E. coli* β-galactosidase was proved to be labeled by the same procedure, suggesting that the modification site with this irreversible substrate might depend both
on the nucleophilicity of the amino acids and their spatial arrangement in the individual catalytic cavity. Similarly, a Glu-259 in $^{257}\text{TLE}^{260}$ was selectively labeled using *Bacillus circulans* $\beta$-galactosidase, indicating that Glu-259 is one of the nucleophiles in the active site. The present method can be readily extended to other glycosidases, and should greatly aid the high-throughput proteomics of many glycoside hydrolases showing both retaining and inverting-type mechanisms.
INTRODUCTION

The fate of cell surface carbohydrates-bearing proteins and lipids are controlled by the systematic enzymatic reactions. Functional oligosaccharide chains of glycoconjugates such as glycoproteins, proteoglycans, and glycolipids produced by glycosyltransferases are eventually degraded by the specific individual metabolic systems (1). Glycoside hydrolases (GHs, carbohydrases, glycosidases; EC 3.2.1), take part in such degradation of the glycoconjugates (2-4). In addition, it is well known that glycoside hydrolases for α-D-mannopyranosides (α-mannosidases) have crucial roles as trimming enzymes in the glycoprotein synthetic pathways to control the eventual types of oligosaccharide structures such as high mannose-type, hybrid-type, and complex-type (5,6). Therefore, precise analysis of the structures, functions, and catalytic mechanism of GHs has become one of the most important processes in systematic biosynthesis and metabolism of glycoconjugates. Moreover, it is also evident that fine structural information on the active site of GHs permit rational design and efficient synthesis of potential modulators or inhibitors as sugar-based therapeutic reagents (7-9). Three-dimensional structures of GHs may be obtained by X ray crystallographic analysis of the crystallized proteins (10-13). However, large-scale expression and crystallization of proteins often become a bottle neck in the practical protein engineering as well as basic proteomics research.
and structural biology.

Recent progress in mass spectroscopy based on various ionization techniques such as matrix-associated laser desorption ionization (MALDI) and electrospray ionization (ESI) techniques greatly facilitated proteomic analysis useful for post-genomic science and technology. In fact, combined use of mass spectroscopy with high performance liquid chromatography (HPLC) or two-dimensional polyacrylamide gel electrophorasis (2D-PAGE) seems to be the most powerful tools for the high-throughput proteomics research (14-17). In the present paper, we report an efficient method for isolation and characterization of galactosidases based on fluorescence labeling by a suicide substrate. The value of this method for high-throughput proteomics study was demonstrated using Xanthomonas manihotis β-galactosidase as a model GH.

β-Galactosidases (EC 3.2.1.23) are exo-type GHs that release terminal D-galactopyranose residues from oligosaccharide chains in a variety of glycoconjugates. They are distributed widely in a number of microorganisms, plants, and animal tissues (18). Human isozymes of β-galactosidase have been investigated extensively from the standpoint of genetic disorders such as lactose intolerance (19) and mucolipidoses (20-24). Bacterial isozymes of β-galactosidases served as reporter molecules for the
evaluation of gene regulation (25). β-Galactosidases are divided into four distinct GH families, GH-1, GH-2, GH-35, and GH-42 based on the similarity in nucleotide and amino acid sequences (26-29). However, at present, only limited information on three dimensional structure, substrate binding site, catalytic site, and post-translational modifications are available. Considering the difficulties in isolation and characterization as well as expression of various isozymes of β-galactosidases, an alternative method to alleviate these time-consuming and tedious procedures for the functional proteomics of the GHs is highly desired.

In 1990, Danzin et al. developed and reported an irreversible inhibitor (suicide substrate) of β-galactosidases based on a fluoromethylated aryl group as a novel aglycon (30). The reactive moiety generated from an aglycon conveniently labeled active site of the GHs for subsequent characterization, because nucleophilic attack of this aglycon by a nearby amino acid side chain at the active site gives rise to a tagged amino acid residue. Subsequently, applications of this principle have emerged. Ichikawa et al. developed a novel biotin-conjugated suicide substrates for isolation and characterization of N-acetyl-β-D-glucosaminidase (31). Janda and co-workers have reported a valuable application of this tagging strategy for chemical selection of catalytic antibody having β-galactosidase-like activity from a phage library (32).
Plückthun and co-workers have also succeeded in trapping a catalyst with alkaline phosphatase activity by using similar mechanism-based chemical selection strategy (33).

In adopting this principle, we chose fluorescent labeling of GH followed by peptide fingerprinting by mass spectroscopy in the high-throughput proteomics of GHs. When a glycosidase is expressed and produced using common bacterial systems, identification of the target protein in a cell lysate, a complex mixture of proteins and other biomolecules, requires rather tedious sequence of blotting, digestion and staining procedures prior to mass spectroscopic analyses (17). In addition, complexity of peptide fragments derived by treating the tagged proteins with specific peptidases still makes precise analyses of peptide sequences difficult. Therefore, we thought that irreversible fluorescence-labeling of GHs would facilitate isolation and identification of labeled active-site peptides as well as labeled GHs. We herein demonstrate that a mechanism-based tagging of β-galactosidases with a fluorescent-labeled suicide substrate (Figure 1) greatly expedited both an efficient chromatographic separation and an accurate sequencing of the labeled peptides by means of a tandem mass spectroscopy (MALDI-TOF-TOF).

Figure 1 (a) & (b)
EXPERIMENTAL PROCEDURES

Reagents and general method. Recombinant β-galactosidase from Xanthomonas manihotis was purchased from New England BioLabs. β-Galactosidases from E. coli (lacZ), Aspergillus oryzae, bovine liver, and endoproteinase Glu-C (EC 3.4.21.19) were obtained from Sigma-Aldrich Co. Ltd. β-Galactosidases from Bacillus circulans, trypsin (EC 3.4.21.4), and Aff-Gel 10 were purchased from Wako Pure Chemical Co. Ltd. Anti-dansyl rabbit IgG was obtained from Molecular Probes Inc. Donkey anti-rabbit IgG (H+L), alkaline phosphate conjugate and Western Blue® (stabilized substrate for alkaline phosphatase) were purchased from Promega Co. Ltd. All other chemical reagents were purchased from Wako Chemical Co. Ltd. All chemical reactions were performed in anhydrous solvents under nitrogen atmosphere in the dark. Reactions were monitored by thin-layer chromatography (t. l. c.) on precoated plates of Silica Gel 60F254 (E. Merck). Visualization was accomplished with UV light (254 nm or 365 nm) and anisaldehyde-methanol-sulfuric acid = 10/5/85 (v/v/v) solution or ninhydrin reagent (Wako Co.) followed by heating. Purifications of the synthetic products by column chromatography were carried out using silica gel 60N (KANTO Co, 63-210nm mesh). Melting points were measured and determined with a Yamato micro
melting-point apparatus.

NMR spectroscopy. $^1$H- and $^{13}$C-NMR spectra were recorded at 400 and 100 MHz, respectively, with Lambda 400 spectrometer (JEOL) using chloroform-$d$ and methanol-$d_4$ as solvents. All signals of the new compounds synthesized in this study were characterized with H-H COSY and HMQC techniques.

MALDI TOF mass spectrometry (34,35). Samples were desalted and concentrated using 10-$\mu$L C18 ZipTips$^\text{TM}$ (Millipore) according to the manufacturer’s instructions. Typically, samples were dissolved in 1 $\mu$L of 90% (v/v) acetonitrile containing trifluoroacetic acid (TFA) and mixed with the same volume of a saturated solution of 2,5-dihydroxy-benzoic acid (DHB) in 33% acetonitrile containing 0.1% TFA. The above mixtures (1 $\mu$L) were applied to a stainless steel target MALDI plate and air-dried before analysis in the mass spectrometer. All measurements were performed using an Ultraflex TOF/TOF mass spectrometer equipped with a reflector, and controlled by the Flexcontrol 1.2 software package (Bruker Daltonics GmbHS, Bremen, Germany). Ions generated by a pulsed UV laser beam (nitrogen laser, $\lambda = 337$ nm) were accelerated to a kinetic energy of 23.5 kV. Metastable ions generated by laser-induced decomposition of the selected precursor ions were analyzed without any additional collision gas. Precursor ions were accelerated to 8 kV and selected in a timed ion gate. The
fragments were further accelerated by 19 kV in the LIFT cell (LIFT means “lifting” the potential energy for the second acceleration of ion source), and their masses were analyzed after the ion reflector passage. Masses were automatically annotated by using XMASS 5.1.2 NT.

External calibration of MALDI mass spectra was carried out using singly charged monoisotopic peaks of a mixture of human angiotensin II (m/z 1046.542), bombesin (m/z 1619.823), ACTH (18-39) (m/z 2465.199), and somatostatin 28 (m/z 3147.472). The mixture of these peptides was measured on the central spot of a 3 x 3 square by using external calibration. To achieve mass accuracy better than 60 ppm, internal calibration was carried out by doping the matrix solution with a mixture of the calibration peptides. Calibration of these mass spectra was performed automatically by utilizing a customized macro command of the XMASS 5.1.2 NT software package. The macro command was used for the calibration of the monoisotopic singly charged peaks of the above mentioned peptides. LIFT-TOF/TOF spectra were annotated with the BioTools 2.1 software package.

_Synthesis of a fluorescent-labeled suicide substrate._ A key synthetic intermediate, 2-difluoromethyl-4-nitro-phenyl 2,3,4,6-tetra-O-acetyl-β-D-galactopyranoside (1), was prepared from 2,3,4,6-tetra-O-acetyl-α-D-galactopyranosyl bromide according to the
procedure reported in a previous paper (32). Characterization data for 1 are as follows.

RF= 0.32 [toluene : ethyl acetate= 3 : 1(v/v)]; mp 132-134°C (crystallized from ethanol); ¹H NMR (400 MHz, CDCl₃) δ 2.01, 2.05, 2.07, 2.18 (all s, 12 H, 4 x -COCH₃), 4.07-4.25 (m, 3 H, H-5, H-6a, b), 5.12 (m, 1 H, H-3), 5.14 (d, 1 H, J = 7.78 Hz, H-1), 5.48 (m, 1 H, H-4), 5.55 (dd, 1 H, J = 2.75 Hz, H-2), 6.83 (t, 1 H, J = 54.65 Hz, CHF₂), 7.21 (d, 1 H, J = 9.16 Hz, aromatic) 8.31, and 8.47 (m, 2 H, aromatic). ¹³C NMR (100 MHz, CDCl₃) δ 20.48, 20.57, 20.61 (4 x CH₃) 61.30 (C-6), 66.56 (C-4), 67.73 (C-2), 70.16 (C-3), 71.73 (C-5), 99.29 (C-1), 109.90 (CHF₂), 114.90, 122.82, 127.68, 142.74, 143.65, 153.95 (6C, aromatic), 169.48, 169.94, 170.00, and 170.21 (4 x CO). Anal. calcd for C₂₁H₂₃F₂NO₁₂: C, 48.56; H, 4.46; N, 2.70. Found: C, 48.57; H, 4.45; N, 2.65.

To a solution of compound 1 (100 mg, 193 µmole) in dry MeOH (5 mL) was added NaOMe (15 mg, 280 µmole). The mixture was stirred at room temperature for 16 h. The solution was adjusted with Dowex 50W-X8 (H⁺) to pH 7.0, and evaporated after filtration with Celite® to afford a syrupy de-O-acetylated derivative. The crude product 2 [2-difluoromethyl-4-nitro-phenyl β-D-galactopyranoside; RF= 0.56 (CHCl₃ : MeOH : H₂O= 65 : 25 : 4)] (68 mg, 193 µmole) was dissolved in dry EtOAc (5 mL). To this solution was added 10 % palladium on activated carbon (30 mg), and the
mixture was stirred under H₂ atmosphere at room temperature for 16 h. The mixture was filtered with Celite® and the solution was evaporated to give an intermediate 3 (60 mg) in quantitative yield.

4-Amino-2-difluoromethyl phenyl β-D-galactopyranoside (3) [R₆= 0.40 (CHCl₃ : MeOH : H₂O= 65 : 25 : 4)] (60 mg, 190 µmole) was dissolved in dry MeOH (3 mL), and mixed with triethylamine (50 µL, 400 µmole), to which dansyl chloride (80 mg, 300 µmole) was added, and the mixture was stirred at room temperature in the dark for 16 h. The mixture was immediately separated with preparative thin layer chromatography (PTLC) using a solvent [CHCl₃ : MeOH : H₂O= 65 : 15 : 4] to yield a target compound 4 (51 mg, 69 %). R₆= 0.60 (CHCl₃ : MeOH : H₂O= 65 : 25 : 4); ¹H NMR (400 MHz, CD₃OD) δ 2.84 (s, 6 H, 2-CH₃), 3.49 (m, 1 H, H-3), 3.57 (m, 1 H, H-5), 3.68~3.73 (m, 3 H, H-2 and H-6), 3.83 (m, 1 H, H-4), 4.68 (d, 1 H, J = 7.64 Hz, H-1), 6.96 (t, 1 H, J = 55.20 Hz, CHF₂), 7.02~7.09 (m, 3 H, aromatic), 7.25 (d, 1 H, J = 7.40 Hz, aromatic), 7.45 (t, 1 H, aromatic), 7.59 (t, 1 H, aromatic), 8.09 (d, 1 H, J = 7.30 Hz, aromatic), 8.38 (d, 1 H, J = 8.30 Hz, aromatic), and 8.47 (d, 1 H, J = 8.30 Hz, aromatic). ¹³C NMR (100 MHz, CD₃OD) δ 45.76 (N-CH₃), 62.35 (C-6), 70.16 (C-4), 72.10 (C-2), 74.80 (C-3), 77.10 (C-5), 103.82 (C-1), 112.14 (CHF₂), 114.47, 116.48, 118.24, 120.26, 124.09, 126.06, 126.58, 129.34, 130.00, 131.10, 131.43, 131.64, 133.50,
135.76, 153.29, and 153.95 (16 C, aromatic). HR-FAB MS m/z; Anal calcd for C_{25}H_{28}F_{2}N_{2}O_{8}S (M): 554.1534, Found: 555.1600 [M+H]^+.

(4-N-Dansyl-phenyl)-β-D-galactopyranoside was prepared from a commercially available p-nitrophenyl-β-D-galactopyranoside by employing a procedure similar to that for the conversion of compound 2 to 4. ^1H NMR for (4-N-dansyl-phenyl)-β-D-galactopyranoside: (400 MHz, CD_{3}OD) δ 2.84 (s, 6 H, 2 x CH_{3}), 3.49 (dd, 1 H, H-3), 3.57 (t, 1 H, H-5'), 3.66-3.72 (m, 3 H, H-2, H-6), 3.83 (m, 1 H, H-3'), 4.66 (d, 1 H, J=7.8 Hz, H-1'), 6.83-6.86 (m, 4 H, aromatic), 7.25 (d, 1 H, aromatic), 7.43 (d, 1 H, aromatic), 7.58 (t, 1 H, aromatic), 8.06 (d, 1 H, aromatic), 8.39 (d, 1 H, aromatic), and 8.48 (d, 1 H, aromatic).

(4-N-Dansyl-2-difluoromethylphenyl)-α-maltoside was synthesized from maltose peracetate and all intermediates were characterized by ^1H-NMR and high resolution mass spectroscopy. Detail synthetic conditions and analytical data for this compound will be reported in the following paper in combination with proteomics study of α-amylases: HR-FAB MS m/z; Anal calcd for C_{31}H_{38}F_{2}N_{2}O_{13}S (M): 716.2063, Found: 717.2134 [M+H]^+.

**Inhibitory effect by a suicide substrate.** Irreversible inhibitions of β-galactosidases by a suicide substrate 4 were measured by fluorometric spectrometry.
in the presence of 4-methylumbelyferyl β-D-galactopyranoside (36) as a competitive substrate ($\lambda_{\text{ex}} = 365$ nm, $\lambda_{\text{em}} = 440$ nm). The inhibition constants ($K_i$) and the half-life for inhibition ($t_{1/2}$) were determined according to the procedure described in the literature (37). Kinetic data were obtained from each of the five enzymes and were shown in Figure 2.

**Mechanism-based labeling of β-galactosidases.** β-Galactosidase (300-1000 µg) was incubated in the presence or absence of the suicide substrates 4 (5 µg, final concentration: 45 µM) at 37 ºC, in 200 µL of 100 mM citrate phosphate buffer (pH 5.0) or 100 mM Hepes-NaOH (pH 7.5). The reaction mixture was subjected directly to the purification by gel permeation chromatography (GPC) using GS510 7E column (Shodex) with 100 mM Tris-HCl buffer (pH 7.5) containing 100 mM NaCl and 0.1% (v/v) SDS as eluent (flow rate: 0.3 mL/min). Chromatography was monitored by means of a HITACHI L-6200 HPLC system equipped with a HITACHI L-7405 UV detector and a HITACHI F-1050 fluorescence spectrophotometer. Efficacy of fluorescent labeling of enzymes under above condition was estimated to be in a range from 10 to 30% by elucidating fluorescence peak area derived from the labeled proteins. The mixture of labeled and unlabeled proteins was directly used for further enzymatic digestion and isolation procedures of the target labeled peptide fragment. Considering
$K_i$ values obtained from the kinetic analysis (Figure 2), it is clear that a large excess of suicide substrate 4 would be required for complete labeling of enzymes. In this study, however, this was settled on 45 µM of 4 because of its limited solubility.

**Western blotting of dansylated β-galactosidases.** Each labeled β-galactosidase (1-50 µg) was applied to 5-20% polyacrylamide gels (ATTO. Co. Ltd, JAPAN) followed by SDS-PAGE. All samples were heated prior to application to SDS-PAGE wells. Proteins were blotted to PVDF membranes (Immun-Blot™ PVDF membrane, Bio-Rad), and the dansylated proteins were detected with the Donkey anti-rabbit IgG AP conjugate (Promega. Co. Ltd) and anti-dansyl rabbit IgG complex using Western Blue® (Promega. Co. Ltd) as a substrate for alkaline-phosphatase. After blocking using 0.3% BSA in TBS buffer, the primary and secondary anti-body were used at a dilution of 1:5000 (about 0.2 µg/mL) and exposed to the membranes for 30 min at 37°C, and washed with TBS containing 0.01% Tween 20 at each step. The final development with alkaline-phosphatase was performed at 37°C for 20 min.

**Enzymatic digestion of β-galactosidase from Xanthomonas manihotis, E. coli, or Bacillus circulans.** After GPC separation of the reaction mixture, the labeled enzymes were concentrated by Centricon YM-30 (Millipore) to approximately 100 µg / 300 µL in 100 mM ammonium bicarbonate (pH 8.3) containing 100 mM NaCl. To this

$15$
solution was added trypsin (2.5 µg / 15 µL) and the mixture was incubated overnight at 37 °C, mixed with an endoproteinase Glu-C or pepsin (2.5 µg / 15 µL) and further incubated overnight at 37 °C.

*Preparation and use of anti-dansyl antibody column.* Anti-dansyl rabbit IgG (0.5 mL, 1 mg/mL) was dialyzed at 4°C for 10 min against a solution of 0.1 M sodium bicarbonate containing 0.1 M NaCl using Centricon YM-30 (Millipore) (5 x 3 mL, 10,000 rpm). Then, the retentate was mixed with 0.5 mL of Affi-Gel 10 (Bio-Rad) according to the manufacture’s instructions and the mixture was packed into a 0.5 x 2 cm column. It was estimated by using Bradford method (38) that 0.98 mg of IgG was coupled with 1.0 mL of polymer support. This antibody column seemed to show a capacity to bind with 1.18 nmol of (4-N-dansyl-phenyl)-β-D-galactopyranoside in a preliminary test. A typical procedure for the isolation of fluorescent-labeled peptide fragment generated by peptidases treatment was as follows: A sample solution (50 µL) was allowed to pass through the anti-dansyl rabbit IgG column at room temperature with a flow rate of 0.1 mL/min. After washing with 3 mL of 100 mM Hepes-NaOH solution (pH 7.5) containing 100 mM NaCl, peptides adsorbed onto the antibody column were eluted with 10% (v/v) acetic acid-water and the fractions containing the desired materials were concentrated by using centrifugal evaporation.
Fluorescent labeling of living cells by irreversible inhibitor. B16 cells were maintained in Dulbecco’s modified Eagle’s medium (DMEM) from Sigma Chem. Co. Ltd. These cells (2 x 10⁴ cells/well) in 200 µL of DMEM were cultured in eight-well plastic chamber slides (IWAKI Co.) at 37°C for 24 h. To the cultured cells in each cell was added 3.0 µg of suicide substrate 4 (final concentration: 27 µM) or (4-N-dansyl-phenyl)-β-D-galactopyranoside and (4-N-dansyl-2-difluoromethylphenyl)-α-maltoside as negative controls and the mixtures were incubated for 6 hr. Each well was then washed with PBS twice and fixed by treating with 10% formaldehyde solution. The plates were air dried and examined with a PROVIS AX80 microscope under excitation at 360 nm (Olympus Co.).

RESULTS

Synthesis. Fluorescent-labeled suicide substrate (4) for high throughput proteomics of β-galactosidases was prepared from a known 2-difluoromethyl-4-nitro-phenyl 2,3,4,6-tetra-O-acetyl-β-D-galactopyranoside (1) by de-O-acetylation, reduction of 4-nitro group, and subsequent N-dansylation reaction as indicated in Scheme 1. Purity and chemical structure of compound 4 were ascertained by NMR and HR-mass spectrometry.
According to the method and condition previously described (30), we examined the effect of compound 4 on the hydrolysis of 4-methylumburyferyl β-D-galactopyranoside by commercially available β-galactosidases from Aspergillus oryzae (K_i=1.44 mM; t_{1/2}=10.5 min), Bacillus circulans (K_i=1.38 mM; t_{1/2}=11.2 min), bovine liver (K_i=0.22 mM; t_{1/2}=4.22 min), Xanthomonas manihotis (K_i=38.5 mM; t_{1/2}=21.1 min), and E. coli (K_i=3.5 mM; t_{1/2}=4.3 min) as indicated in Figure 2. These results suggest that compound 4 is a versatile and general suicide substrate for inhibition of β-galactosidases activities. Irreversibility of the inhibition by compound 4 was demonstrated by elucidating kinetic data according to the previous reports (30, 37).

We also examined efficacy of this inhibitor as a fluorescent labeling reagent for these enzymes. All β-galactosidases used in this study were successfully labeled by this reagent, because enzymes exhibited significant fluorescence emission attributable to dansyl group (λ_em = 520 nm) only after mixed with the suicide substrate. Western blotting of these enzymes also indicated covalent labeling by compound 4 as shown in Figure 3. Xanthomonas manihotis, Aspergillus oryzae, and E. coli exhibited clear
single bands at 58K and 114K, respectively. However, significant degradation of the enzymes were observed in cases of bovine liver and *Bacillus circulans*. Thus, the labeled proteins were treated with proteases to give complex peptide fragments observed in MALDI-TOF mass spectra as shown in Figure 4. It is evident that a signal due to fluorescent labeled peptides (mass=1404.67, 1371.70, and 861.13) can not be directly identified if the peptide mixture was analyzed, owing to the complexity of signals of other peptides found in the upper spectra. However, the MALDI-TOF mass spectra of the peptides readily purified by anti-dansyl antibody column showed a clear and simple signals corresponding to the labeled peptide fragments.

Figure 3

Figure 4

*Characterization of catalytic domain of β-galactosidases from Xanthomonas manihotis, E. coli, and Bacillus circulans.* We firstly utilized β-galactosidase from *Xanthomonas manihotis* (39,40) because this enzyme showed the lowest affinity (K<sub>i</sub>=38.5 mM, t<sub>1/2</sub>=21.1 min) for the suicide substrate 4 among enzymes used herein as indicated in Figure 2(d). Thus, the labeled protein was treated with trypsin and endoproteinase Glu-C to give complex peptide fragments observed in MALDI-TOF mass spectra. However, the MALDI-TOF mass spectrum of the peptide readily purified by anti-dansyl antibody column showed a clear and simple signal
corresponding to the labeled peptide fragment (mass=1404.67).

Figure 5

The parent ion of 1404.67 m/z from the fluorescent labeled peptide was subjected to LIFT TOF/TOF analysis to yield a spectrum having a series of b-ions (1271, 957, 723, 211, and 114) and y-ions (682, 611, 448, 262 and 134) as indicated in Figure 5. The internal fragment ions without labeled amino acid residues, in which b-ions with masses of 1049, 916, 788, 602, 439, 368 and y-ions with masses of 836, 939 were also observed. In addition, the fragmentation from a dansylated arginine was detected, allowing us to conclude that this peptide (the calculated molecular mass; [M+H] 1404.68) was sequence as 56IPRAYWKD53 in which Arg-58 was dansylated.

Interestingly, it was found by homology analysis (LALIGN, Local Alignments) of amino acid sequences of three β-galactosidases listed in PDB (10-12) that the sequence from Leu-46 to Tyr-194 including the Arg-58 dansylated sequence of Xanthomonas manihotis β-galactosidase seems to be quite similar to the sequence corresponding to the active site of Thermus thermophilus β-galactosidase (12), the sequence from Met-1 to Tyr-151 (data not shown). Although these sequences of 151 amino acids between Thermus thermophilus and Xanthomonas manihotis β-galactosidases have low homology (25.3%), one may speculate on the commonality
that Glu-184 can be the catalytic nucleophile, and Arg-147 and Asn-183 may provide hydrogen bonding with OH groups of the galactose residue. In fact, a plausible model of three-dimensional structure clearly predicts that the Arg-58 residue of \textit{Xanthomonas manihotis} β-galactosidase seems to locate quite similar position in the active site to the corresponding Lys-13 residue in the TIM barrel structure of the \textit{Thermus thermophilus} β-galactosidase known as a member of GH-42 family (12). This similarity of the amino acid sequence was not found in other two β-galactosidases from \textit{E. coli} (10) and hyperthermophilic archeon \textit{Sulfolobus solfataricus} (11).

On the other hand, when \textit{E. coli} (lacZ) β-galactosidase whose three-dimensional structure had been elucidated as a member of GH-2 family by X ray crystallography (10) was employed for the same proteomics analysis using the present strategy (Figure 6), it was revealed that one of the two catalytic glutamic acids, Glu-537 located in a sequence of $^{530}\text{TRPLILCE}^{537}$ (the calculated molecular mass: [M+H]$^+$ 1371.64), involved in the active site was directly labeled with a fluorescent group as indicated in its TIM barrel structure (Figure 7). This result suggests that the modification site with this irreversible substrate might deeply depend on the reaction mechanism (velocity), nucleophilicity of the catalytic and neighboring amino acid residues, and their spatial arrangement (vicinity) in the individual catalytic cavity. It is also noteworthy that
MALDI-LIFT-TOF/TOF is a nice tool for the identification of such (post translational) modification sites with unstable functional groups such as hemiacetal and/or acyl related linkages found in this case as well as labile O-glycans (41). As for *Bacillus circulans* (42), it was suggested that one of the two catalytic amino acids, Glu-259 located in a sequence of $^{257}$TLE$^{260}$ (the calculated molecular mass; $[M+H]^+$ 861.326), involved in the active site was labeled by a dansyl group through the same hemiacetal linkage found in the case of *E. coli* β-galactosidase. At present, neither amino acid sequences nor DNA sequences have been reported for *Aspergillus oryzae* and bovine liver β-galactosidases, although these two enzymes are commercially available.

Mapping of the fluorescent labeled amino acids with neighboring peptide sequences were summarized in Figure 9.

![Figure 6](#)

![Figure 7](#)

![Figure 8](#)

![Figure 9](#)

In addition, we have examined the selective labeling of the living cells by using the present irreversible inhibitor. The fact that β-galactosidases are one of the most important lysosomal-enzymes commonly found in mammalian cells, strongly prompted
us to examine the mechanism-based labeling of B16 as a model. Preliminarily, cytoplasm of B16 cells were found to be stained only when the cells were cultured with the irreversible inhibitor 4, while (4-N-dansyl-phenyl)-β-D-galactopyranoside or (4-N-dansyl-2-difluoromethylphenyl)-α-maltoside used as negative controls did not show any capacity to tag the enzymes with the fluorescent group (data not shown).

DISCUSSION

A number of irreversible inhibitors have been utilized for elucidating the mechanism and active sites of GHs. For example, Withers and coworkers have already established practical and efficient methods for the identification of the catalytic nucleophile of the active site in GHs using an irreversible inhibitor, 2,4-dinitrophenyl 2-deoxy-2-fluoro-β-D-galactopyranoside (43-45), for the identification of the catalytic amino acid residues of a variety of β-galactosidases. Actually, this type of inhibitors has been expected to be useful reagents for trapping the intermediate formed by the action of the retaining-type GHs (18). However, it should be noted that an intermediate generated by inverting-type reaction mechanisms can not be captured by employing these mechanism-based labeling strategy. Moreover, isolation of the labeled proteins derived from treating with these irreversible inhibitors and peptide
fragments generated by proteolytic digestion still requires laborious and tedious
purification processes, impeding its application to practical and speedy high throughput
proteomics of GHs.

The advantage of the present reagent lies in that compound 4 will react with
both retaining and inverting-type GHs. Since compound 4 generates the same reactive
intermediate through either of the hydrolytic mechanisms, a catalytic amino acid
(Glu-537 of E. coli and Glu-259 of Bacillus circulans β-galactosidase) or a nucleophile
nearby the active site (Arg-58 of Xanthomonas manihotis β-galactosidase) must attack
this intermediate to yield dansylated enzymes. In addition to labeling the enzyme
dansyl group greatly facilitates further isolation and characterization of the peptide mass
fingerprinting of GHs because of availability of anti-dansyl antibody. Two
dimensional SDS-PAGE and automated sampling of the modified peptide would also
accelerate high throughput analysis by MALDI-LIFT-TOF/TOF apparatus. Labeling
of the cytoplasmic enzymes of the living cells will also become an efficient method for
staining, quantifying, and isolating the target/unknown GHs. Starting with
4-amino-2-difluoromethyl phenyl glycosides of other sugars, a variety of
mechanism-based fluorescent labeling reagents of exo-glycosidases can be readily
prepared. We are currently developing such reagents for β-hexosaminidases,
α-mannosidases, and neuraminidases.

ACKNOWLEDGMENT

This work was partly supported by a grant for “Research project of the functional glycomaterials” from the Ministry of Education, Culture, Science, and Technology, Japan and by a grant for “Glycocluster project” from NEDO, Japan. The authors wish to express their gratitude to Dr. Y. Ichikawa of Johns Hopkins University (now at Optimer Ltd.) for providing us some unpublished results and timely advice. We also thank Dr. R. T. Lee of Johns Hopkins University, Dr. K. Iwata, Dr. H. Hoshi, and Dr. J. Inokuchi of Hokkaido University for their helpful discussion and technical assistance.

REFERENCES

1. Dwek, R. A. (1996) Chem Rev 96, 683-720
2. Sears, P., and Wong, C. H. (1998) Cell Mol Life Sci 54, 223-252
3. Herscovics, A. (1999) Biochim Biophys Acta 1473, 96-107
4. Parodi, A. J. (2000) Biochem J 348, 1-13
5. Roth, J., Zuber, C., Guhl, B., Fan, J. Y., and Ziax, M. (2002) Histochem Cell Biol 117, 159-169
6. Akama, T. O., Nakagawa, H., Sugihara, K., Narisawa, S., Ohyama, C., Nishimura, S.-I, O’Brien, D. A., Moremen, K. W., Millan, J. L., and Fukuda, M. N. (2002) Science 295, 124-127

7. von Itzstein, M., Wu, W. Y., Kok, G. B., Pegg, M. S., Dyason, J. C., Jin, B., Van Phan, T., Smythe, M. L., White, H. F., Oliver, S. W., and et al. (1993) Nature 363, 418-423

8. Taylor, G. (1993) Nature 363, 401-402

9. Babu, Y. S., Chand, P., Bantia, S., Kotian, P., Dehghani, A., El-Kattan, Y., Lin, T. H., Hutchison, T. L., Elliott, A. J., Parker, C. D., Ananth, S. L., Horn, L. L., Laver, G. W., and Montgomery, J. A. (2000) J Med Chem 43, 3482-3486

10. Jacobson, R. H., Zhang, X. J., DuBose, R. F., and Matthews, B. W. (1994) Nature 369, 761-766

11. Carlos, F. A., Ian, S., Marco, M., Maria, C., Roberto, N., Mose, R., and Laurence, H. P. (1997) J. Mol. Biol. 271, 289-802

12. Masafumi, H., Shinya, F., Naomi, O., Hidemasa, M., Hiroshi, M., Hirofumi, S., and Takayoshi, W. (2002) J. Mol. Biol. 322, 79-91

13. Carsten, R., Leila, L. L., Pedro, M. C., Bernard, H., and Sine, L. (2002) Biochemistry 41, 15135-15143
14. Keough, T., Lacey, M. P., Fieno, A. M., Grant, R. A., Sun, Y., Bauer, M. D., and Begley, K. B. (2000) *Electrophoresis* **21**, 2252-2265

15. Mann, M., Hendrickson, R. C., and Pandey, A. (2001) *Annu Rev Biochem* **70**, 437-473

16. Ziegler, Z. (2001) *Anal Chem* **73**, 4063-4070

17. Henzel, W. J., Watanabe, C., and Stults, J. T. (2003) *J Am Soc Mass Spectrom* **14**, 931-942

18. Sinnott, M. L. (1990) *Chem Rev* **90**, 1171 - 1202

19. Montgomery, R. K., Buller, H. A., Rings, E. H., and Grand, R. J. (1991) *Faseb J* **5**, 2824-2832

20. Kobayashi, T., Ohta, M., Goto, I., Tanaka, Y., and Kuroiwa, Y. (1979) *J Neurol* **221**, 137-149

21. Verheijen, F., Brossmer, R., and Galjaard, H. (1982) *Biochem Biophys Res Commun* **108**, 868-875

22. Mueller, O. T., and Shows, T. B. (1982) *Hum Genet* **60**, 158-162

23. Pshezhetsky, A. V., and Ashmarina, M. (2001) *Prog Nucleic Acid Res Mol Biol* **69**, 81-114

24. Callahan, J. W. (1999) *Biochim Biophys Acta* **1455**, 85-103
25. Lapierre, L., Mollet, B., and Germond, J. E. (2002) *J Bacteriol* **184**, 928-935
26. Henrissat, B. (1991) *Biochem J* **280**, 309-316
27. Henrissat, B., and Bairoch, A. (1993) *Biochem J* **293**, 781-788
28. Henrissat, B., and Davies, G. (1997) *Curr Opin Struct Biol* **7**, 637-644
29. Henrissat, B., and Davies, G. J. (2000) *Plant Physiol* **124**, 1515-1519
30. Halazy, S., Berges, V., Ehrhard, A., and Danzin, C. (1990) *Bioorg Chem* **18**, 330-344
31. Ichikawa, M., and Ichikawa, Y. (2001) *Bioorg Med Chem Lett* **11**, 1769-1773
32. Janda, K. D., Lo, L. C., Lo, C. H., Sim, M. M., Wang, R., Wong, C. H., and Lerner, R. A. (1997) *Science* **275**, 945-948
33. Cesaro-Tadic, S., Lagos, D., Honegger, A., Rickard, J. H., Partridge, L. J., Blackburn, G. M., and Plückthun, A. (2003) *Nat Biotechnol* **21**, 679-685
34. Schnaible, V., Wefing, S., Resemann, A., Suckau, D., Bucker, A., Wolf-Kummeth, S., and Hoffmann, D. (2002) *Anal Chem* **74**, 4980-4988
35. Suckau, D., Resemann, A., Schuergenber, M., Hufnagel, P., Franzen, J., and Holle, A. (2003) *Anal Bioanal Chem* **376**, 952-965
36. Cheetham, P. S., and Dance, N. E. (1976) *Biochem J* **157**, 189-195
37. Tipton, K. F. (1996) *Enzymology* LABFAX (Engel, P. C., Ed.), Academic Press,
p151-171.

38. Bradford, M. M. (1976) *Anal Biochem* **72**, 248-254

39. Taron, C. H., Benner, J. S., Hornstra, L. J., and Guthrie, E. P. (1995) *Glycobiology* **5**, 603-610

40. Wong-Madden, S. T., and Landry, D. (1995) *Glycobiology* **5**, 19-28

41. Kuroguchi, M., Matsushita, T., and Nishimura, S. -I. (2004) *Angew. Chem. Int. Ed.* **43**, in press (Z460020)

42. Fujimoto, H., Miyasato, M., Ito, Y., Sasaki, T., and Ajisaka, K. (1988) *Glycoconjugate J.* **15**, 155-160

43. Gebler, J. C., Aebersold, R., and Withers, S. G. (1992) *J Biol Chem* **267**, 11126-11130

44. McCarter, J. D., Burgoyne, D. L., Miao, S., Zhang, S., Callahan, J. W., and Withers, S. G. (1997) *J Biol Chem* **272**, 396-400

45. Dan, S., Marton, I., Dekel, M., Bravdo, B. A., He, S., Withers, S. G., and Shoseyov, O. (2000) *J Biol Chem* **275**, 4973-4980
Legends for Scheme and Figures

**Scheme 1.** Synthetic scheme of a specific irreversible inhibitor for β-galactosidases.

Conditions of each reaction and spectral data were described in the experimental part of the text.

**Figure 1.** (a) Chemical structure of a specific irreversible inhibitor having a fluorescent group.  (b) A plausible mechanism of the mechanism-based fluorescent labeling by compound 4.

**Figure 2.** Kinetic behavior of the specific irreversible inhibitor 4. To a solution of β-galactosidase and 0 to 1.5 mM of fluorescent-labeled substrate was added 4-methylumberyferyl β-D-galactopyranoside (final concentration: 3 mM) and changes in the hydrolytic activity of the enzyme were determined by measuring the relative fluorescence intensity.  (a) β-galactosidase from *Aspergillus oryzae*, (b) β-galactosidase from bovine liver, (c) β-galactosidase from *Bacillus circulans*, (d) β-galactosidase from *Xanthomonas manihotis*, (e) β-galactosidase from *E. coli*.

**Figure 3.** Western immunoblot of β-galactosidases labeled by a dansyl group reacted with anti-dansyl rabbit IgG. Donkey anti-rabbit IgG conjugated with alkaline phosphatase and Western Blue® (Stabilized substrate for alkaline phosphatase) were used for the staining the primary antibody.  CBB means coomassie brilliant blue.
Figure 4. MALDI-TOF mass spectra of fluorescent labeled peptides. Peptide fragments produced by digestion with some proteases were analyzed. Upper; mixtures of peptides derived from labeled enzymes, bottom; isolated peptides by using anti-dansyl antibody column chromatography.

Figure 5. MALDI LIFT-TOF/TOF mass spectrum of the dansylated peptide fragment (m/z 1404.67) and assignment of the identified sequence 56-63: IPRAWKD, which contains a dansylated Arg-58 (R) residue.

Figure 6. MALDI LIFT-TOF/TOF mass spectrum of the dansylated peptide fragment (m/z 1371.64) and assignment of the identified sequence 530-537: TRPLILCE, which contains a dansylated Glu-537 (E) residue.

Figure 7. TIM barrel view and topology of the fluorescent labeled Glu-537 in the three-dimensional structure of \textit{E. coli} \(\beta\)-galactosidase in PDB.

Figure 8. MALDI LIFT-TOF/TOF mass spectrum of the dansylated peptide fragment (m/z 861.326) and assignment of the identified sequence 257-260: TLEE, which contains a dansylated Glu-259 (E) residue.

Figure 9. Sequences surrounding the labeled nucleophile region of \(\beta\)-galactosidases.
Specific substrate

Fluorescence probe

Reactive group

Figure 1a
Figure 1b

O-Gal bond cleavage

- HF
**Aspergillus oryzae** β-galactosidase

\[
y = 10.478x + 15.121 \\
R^2 = 0.9867 \\
\]

Ki = 1.44 mM  
\( t_{1/2} = 10.5 \text{ min} \)

\( 1/\text{[suicide substrate]} \ (1/\text{mM}) \)

---

**Bovine liver** β-galactosidase

\[
y = 27.971x + 6.0854 \\
R^2 = 0.9564 \\
\]

Ki = 0.22 mM  
\( t_{1/2} = 4.22 \text{ min} \)

\( 1/\text{[suicide substrate]} \ (1/\text{mM}) \)

---

**Bacillus circulans** β-galactosidase

\[
y = 11.64x + 16.099 \\
R^2 = 0.9989 \\
\]

Ki = 1.38 mM  
\( t_{1/2} = 11.2 \text{ min} \)

\( 1/\text{[suicide substrate]} \ (1/\text{mM}) \)

---

**Xanthomonas manihotis** β-galactosidase

\[
y = 0.7904x + 30.455 \\
R^2 = 0.9745 \\
\]

Ki = 38.5 mM  
\( t_{1/2} = 21.1 \text{ min} \)

\( 1/\text{[suicide substrate]} \ (1/\text{mM}) \)

---

**E. coli** β-galactosidase

\[
\text{Ki} = 3.5 \text{ mM} \\
\text{t}_{1/2} = 4.3 \text{ min} \\
R^2 = 0.9645 \\
\]

\( 1/\text{[suicide substrate]} \ (1/\text{mM}) \)

---

**Xanthomonas manihotis** β-galactosidase

**Bacillus circulans**

**Aspergillus oryzae**

**E. coli**

Bovine liver

---

**Figure. 2**
Figure 4
Figure 5
Figure 6
Figure 7
Figure 8
Xan. 41 GKYPLSGAIHFQRPAYWKDLQKARALGLNTVETYV
E. coli. 521 IKKWLSLPGETRPLILCEYAHAMGNSLGGFAKYWQAHRQY
Bac. c. 241 HWGEHHTRSAESVVETLEELIKQNGSVCNFMAHGGTFNG

Figure 9
Scheme 1
Mechanism-based fluorescent labeling of beta-galactosidases: An efficient method in proteomics for glycoside hydrolases
Masaki Kurogochi, Shin-Ichiro Nishimura and Yuan Chuan Lee

J. Biol. Chem. published online August 12, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M401718200

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