J. Biochem. 132, 237–243 (2002)

Molecular Cloning and Characterization of a Novel Glucocerebrosidase of *Paenibacillus* sp. TS12[1]

Tomomi Sumida, Noriyuki Sueyoshi, and Makoto Ito[2]

Department of Bioscience and Biotechnology, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581

Received May 2, 2002; accepted May 22, 2002

We report here the molecular cloning and characterization of a glucocerebrosidase [EC 3.2.1.45] from *Paenibacillus* sp. TS12. The open reading frame of the glucocerebrosidase gene consisted of 2,493 bp nucleotides and encoded 831 amino acid residues. The enzyme exhibited no sequence similarity with a classical glucocerebrosidase belonging to glycoside hydrolase (GH) family 30, but rather showed significant similarity with GH family 3 β-glucosidases from *Clostridium thermocellum*, *Ruminococcus albus*, and *Aspergillus aculeateus*. The recombinant enzyme, expressed in *Escherichia coli* BL21-(DE3)pLysS, had a molecular weight of 90.7 kDa and hydrolyzed NBD-labeled glucosylceramide, but not galactosylceramide, GM1a or sphingomyelin. The enzyme was most active at pH 6.5, and its apparent K_m and V_{max} values for NBD-labeled glucosylceramide and *p*-nitrophenyl-β-glucopyranoside were 223 μM and 1.60 μmol/min/mg of protein, and 593 μM and 112 μmol/min/mg of protein, respectively. Site-directed mutagenesis indicated that Asp-223 is an essential amino acid for the catalytic reaction and possibly functions a catalytic nucleophile, as in GH family 3 β-glucosidases. This is the first report of the molecular cloning and characterization of a glucocerebrosidase from a procaryote.

Key words: GH family 3 β-glucosidase, glucocerebrosidase, glucosylceramide, *Paenibacillus* sp., site-directed mutagenesis.

Glycosphingolipids (GSLs) are abundant in the outer leaflets of the plasma membranes of vertebrates, and could function as regulators of cellular differentiation, adhesion, and recognition (*1*). In mammals, GSLs are hydrolyzed by exo-glycosidases in lysosomes (*2*), while the microbial degradation of GSLs in natural habitats has not been fully elucidated. Very recently, we found a bacterium, *Paenibacillus* sp. TS12, that decomposed various GSLs by producing a series of exoglycosidases including glucocerebrosidase (GCase) (Sumida *et al.*, unpublished data).

GCase (EC 3.2.1.45, GH family 30, also known as glucosylceramidase) is an enzyme that hydrolyzes the β-glucosidic linkage between glucose and ceramide (Cer) in glucocerebroside (glucosylceramide, GlcCer), a precursor for almost all GSLs, and potentially important in axonal morphology and neuronal functions (*3, 4*). A genetic deficiency of GCase causes Gaucher disease, in which GlcCer is accumulated in lysosomes (*5*). GCase does not hydrolyze the same linkage if an oligosaccharide is linked with Cer, and thus is distinguishable from an endoglycoceramidase (EC 3.2.1.123, GH family 5), which hydrolyzes the β-glucosidic linkage between the oligosaccharide and Cer of various GSLs except cerebrosides. GCases have been cloned from *Homo sapiens* (*6*), *Mus musculus* (*7*), and *Pan troglodytes* (*8*). However, there have been no reports on the molecular cloning and characterization of a procaryote GCase, although *Salmonella typhimurium* possesses unknown proteins showing sequence similarity with GCase (*9*).

In this paper, we report the molecular cloning, expression and characterization of a novel GCase of *Paenibacillus* sp. TS12.

MATERIALS AND METHODS

Materials—*Escherichia coli* strains DH5α and BL21-(DE3)pLysS, and Pyrobest DNA polymerase were purchased from Takara Shuzo (Shiga). Plasmids pET23a and pBluescript II SK(+) were obtained from Novagen (Madison, WI) and Stratagene (La Jolla, CA), respectively. Restriction enzymes, T4 DNA ligase, and GlcCer were obtained from Wako Pure Chemical Industries (Osaka). Precoated Silica Gel 60 TLC plates were purchased from Merck (Germany). 4-Methylumbelliferyl-β-D-glucopyranoside (4MU-β-Glc) and various *p*-nitrophenyl-glycopyranosides (*p*NP-glycopyranosides) were from Sigma (St. Louis, MO). C12-NBD-Cer (NBD-C12:0/d18:1), C12-NBD-GM1a (NBD-C12:0/d18:1), and C12-NBD-sphingomyelin (NBD-SM, NBD-C12:0/d18:1) were prepared by use of sphin-

[1] Nucleotide sequence data reported are available in the DDBJ databases under accession number AB084154.

[2] To whom correspondence should be addressed. Phone: +81-92-642-2898, Fax: +81-92-642-2898 or 81-92-642-2907, E-mail: makotoi@agr.kyushu-u.ac.jp

Abbreviations: Cer, ceramide; GalCer, galactosylceramide; GCase, glucocerebrosidase; GlcCer, glucosylceramide; GTA buffer, 50 mM 3,3-dimethyl-glutaric acid, 50 mM Tris[hydroxymethyl]aminomethane, and 50 mM 2-amino-2-methyl-1,3-propanediol; 4MU-β-Glc, 4-methylumbelliferyl-β-D-glucopyranoside; NBD, 4-nitrobenz-2-oxa-1,3-diazole; pNP, *p*-nitrophenyl; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide; SM, sphingomyelin; TDC, taurodeoxycholate; TLC, thin-layer chromatography.

golipid ceramide *N*-deacylase (SCDase), as described in Refs. *10* and *11*. C12-NBD-GlcCer and C12-NBD-GalCer were synthesized as previously described (*12*). [^{14}C]GlcCer ([^{14}C]C12:0, [^{14}C]C16:0, [^{14}C]18:0, 1.85 GBq/mmol) was prepared as described in Ref. *13*. All other reagents were of the highest purity available.

Bacterial Strains—Paenibacillus sp. TS12 was isolated from soil from Fukuoka prefecture (Sumida *et al.*, unpublished results). TS12 is a short rod-shaped, Gram-negative, catalase-positive, oxidase-negative bacterium with peritrichous flagella. The G+C content of the DNA is 49 mol%. This bacterium was assigned to the genus *Paenibacillus* based on biochemical and physiological properties, and the results of 16S rDNA analysis. The strain was grown at 30°C in medium containing 0.5% polypepton, 0.1% yeast extract, 0.2% NaCl, 0.05% TDC, and 0.05% crude bovine brain gangliosides, pH 7.4. The medium was solidified with 1.5% agar if necessary. *E. coli* DH5α and BL21(DE3)pLysS, used as host cells for the expression of GCase, were grown in Luria-Bertani (LB) medium.

Construction of a Genomic DNA Library of Paenibacillus sp. TS12—Genomic DNA was prepared from *Paenibacillus* sp. TS12 by the method described in Ref. *14* and then partially digested with *Sau*3AI. The *Sau*3AI-fragments (2–10 kbp) were gel-purified and ligated to *Bam*HI-digested pBluescript II SK(+) DNA. The plasmids were used for the transformation of *E. coli* DH5α.

Expression Screening of the Gene Encoding GCase—E. coli DH5α cells transformed with the plasmids containing *Paenibacillus* DNA fragments were seeded (approximately 400 colonies/9.2-cm plate) on LB agar plates supplemented with 100 μg/ml of ampicillin and then incubated at 37°C for 16 h. Colonies were transferred from the plates onto nylon membranes (Biodyne A, Pall Co.), which were then incubated with 0.3 mM 4MU-β-Glc in 200 μl of 25 mM sodium acetate buffer, pH 5.5. Following incubation at 37°C for 30 min, positive colonies, visualized under a UV transilluminator, were picked up with sterilized toothpicks and transferred to 5 ml of LB medium. Following incubation at 37°C for 16 h with shaking, cells were harvested by centrifugation, suspended in 100 μl of 25 mM sodium acetate buffer, pH 5.5, and then lysed by sonication. The cell lysate was centrifuged at 8,000 ×*g* for 10 min, and the supernatant obtained was used as the crude enzyme solution. The activity of GCase was measured using C12-NBD-GlcCer as a substrate as described below. The positive clone was designated as pGlc4.

DNA Sequencing and Sequence Analysis—Nucleotide sequences were determined by the dideoxynucleotide chain termination method with a BigDye Terminator Cycle Sequencing Ready Reaction Kit Ver.3 (Applied Biosystems, USA) and a DNA sequencer (Applied Biosystems, model 377). Computer analyses were performed using DNASIS (Hitachi Software Engineering, Tokyo).

Construction of an Expression Vector—The following primers were used for PCR: UGlc4V (5′-ATA GCT AGC ATG GCG CAA CTC ACG CTT GAA GAA-3′) and LGlc4/2493 (5′-ATT CTC GAG TCC TCT GAC TAC ACT CAA GTC GGT-3′). UGlc4V and LGlc4/2493 contained an *Nhe*I site (underlined) and an *Xho*I site (double underlined), respectively. PCR was performed in 50 μl of a reaction mixture containing each primer at 0.2 μM, 50 ng of template DNA (pGlc4), 0.2 mM dNTPs (dATP, dCTP, dGTP, and dTTP), and 2 U of Pyrobest DNA polymerase using a T-personal 48 (Biometra, Germany) for 30 cycles (each consisting of denaturation at 98°C for 10 s and extension at 68°C for 2.5 min). PCR products were extracted from a 0.7% agarose gel, and the amplified products were digested with *Nhe*I and *Xho*I. The *Nhe*I/*Xho*I fragments were cloned into *Nhe*I/*Xho*I-digested pET23a. The recombinant plasmid was designated as pETGlc4.

Construction of Asp-223 Mutants—Mutagenesis was performed by the PCR overlap extension method (*15*), with UGlc4, LGlc4/2493 and the following oligonucleotide primers: D223E, 5′-GTC GTA TCG GAA TGG GGC GCC-3′ and 5′-GGC GCC CCA TTC CGA TAC GAC-3′: and D223N, 5′-GTC GTA TCG AAC TGG GGC GCC-3′ and 5′-GGC GCC CCA GTT CGA TAC GAC-3′ (underlining shows the locations of the mutations). PCR products were extracted from a 0.7% agarose gel, and the amplified products were digested with *Nhe*I and *Xho*I. The *Nhe*I/*Xho*I fragments were cloned into *Nhe*I/*Xho*I-digested pET23a. The recombinant Asp-223 mutant plasmids were designated as pETD223E and pETD223N.

Expression and Purification of Recombinant GCase and Asp-223 Mutants—E. coli BL21(DE3)pLysS cells transformed with pETGlc4 (or pETD223N or pETD223E) were grown at 25°C for 16 h in 5 ml of medium A (LB medium containing 100 μg/ml of ampicillin and 35 μg/ml of chloramphenicol) with shaking. The culture was then transferred to a 300-ml flask containing 100 ml of medium A and incubated at 25°C for 16 h with shaking. Then, isopropylthio-β-galactopyranoside (IPTG) was added to the culture to a final concentration of 0.1 mM to cause transcription. After an additional 1 h culture at 25°C, cells were harvested by centrifugation, and suspended in 10 ml of buffer A (10 mM Tris-HCl, pH 7.5, containing 150 mM NaCl, and 0.1% Triton X-100) containing protease inhibitors (5 μg/ml of leupeptin, chymostatin, and pepstatin). After sonication, cell debris was removed by centrifugation (8,000 ×*g* for 10 min), and the supernatant obtained was loaded on a HiTrap Chelating HP column (1 ml; Amersham Biosciences, Buckinghamshire, United Kingdom), pre-equilibrated with buffer A. The column was washed with 10 ml of buffer A and then the GCase was eluted with buffer B (10 mM Tris-HCl, pH 7.5, containing 150 mM NaCl, 0.1% Triton X-100, and 20 mM imidazole), buffer C (10 mM Tris-HCl, pH 7.5, containing 150 mM NaCl, 0.1% Triton X-100, and 50 mM imidazole), and buffer D (10 mM Tris-HCl, pH 7.5, containing 150 mM NaCl, 0.1% Triton X-100, and 200 mM imidazole). The active fractions were pooled and used for the characterization of the enzyme.

Enzyme Assay—The activity of GCase was measured by the following three methods. Assay I (*p*NP-β-Glc being used as a substrate); the reaction mixture contained 100 nmol of *p*NP-β-Glc and an appropriate amount of the enzyme in 100 μl of 25 mM sodium phosphate buffer, pH 6.5. Following incubation at 37 °C for a specified period, the reaction was stopped by adding 100 μl of 1 N NaOH, and then the absorbance was measured at 405 nm. One unit of the enzyme was defined as the amount that catalyzed the release of 1 μmol of *p*-nitrophenol per min from *p*NP-β-Glc under the conditions used. Assay II (C12-NBD-GlcCer being used as a substrate); the reaction mixture contained 200 pmol of C12-NBD-GlcCer and an appropriate amount of the enzyme in 20 μl of 25 mM sodium phosphate buffer,

(A)

```
-184  GATCCCATCTCTCCCGGTTGAATGATTTGCTTTAGCTTTTATCATACCTAAACAAAGATTTAAAATCCAATAAAAACATTTAAATTATCATATTAGTGAAAACGCTAACCTGATAAACTG  -63
 -64  GCGTTACGGTAATGCGTTACTGATAAAGGAGGAGAGTCTAAATTGAAAAATATCGGTGAAATTGTGGCGCAACTCACGCTTGAAGAAAAAGCCGGCCTCTGTTCGGGGGAAAGCTTTTGG   57
                                                                  M  A  Q  L  T  L  E  E  K  A  G  L  C  S  G  E  S  F  W    19

  58  AGGACCAAAGCAATTGATCGTCTGGGGATTCCGTCCATCATGATGACAGACGGACCTCACGGCTTGCGCAAGCAAGCGGGGGAAGCGGACCATCTGGGACTGAACGAGAGCATTCCGGCA   177
  20  R  T  K  A  I  D  R  L  G  I  P  S  I  M  M  T  D  G  P  H  G  L  R  K  Q  A  G  E  A  D  H  L  G  L  N  E  S  I  P  A    59

 178  ACGTGCTTTCCGACCGCCGCCGGGCTTGCGAGCTCCTGGGACCGCGAACTGGTGCGAAAGGTAGGAGAAGCGCTGGGAAAGGAAAGCCAGGCAGAGAACGTCTCCATCCTGCTGGCACCT   297
  60  T  C  F  P  T  A  A  G  L  A  S  S  W  D  R  E  L  V  R  K  V  G  E  A  L  G  K  E  S  Q  A  E  N  V  S  I  L  L  G  P    99

 298  GGCGCGAATATTAAACGTTCGCCACTGTGCGGGAGGAACTTCGAGTATTTCTCGGAAGATCCGTATCTGACGGGCGAGTTGGCCGCGGCGCATATTGCAGGCGTTCAAAGCCAGGGTGTC   417
 100  G  A  N  I  K  R  S  P  L  C  G  R  N  F  E  Y  F  S  E  D  P  Y  L  T  G  E  L  A  A  A  H  I  A  G  V  Q  S  Q  G  V   139

 418  GGCACGTCGCTGAAGCATTTCGCTGTCAACAACCAGGAGCATCGCCGGATGACGACGGATGCTGTGGTGGACGAACCGACGCTGCGCGAAATTTATTTGACCGGCTTCGAGATTGCCGTG   537
 140  G  T  S  L  K  H  F  A  V  N  N  Q  E  H  R  R  M  T  T  D  A  V  V  D  E  R  T  L  R  E  I  Y  L  T  G  F  E  I  A  V   179

 538  AAGAAATCGCAGCCATGGACGGTCATGTCGGCGTACAACCCGATGAACGGAACCTACTGCTCCGAAAACGAAACGTTGCTGACCCGCATTCTGAAGGAGGAATGGGGCCACGAGGGCATC   657
 180  K  K  S  Q  P  W  T  V  M  S  A  Y  N  R  M  N  G  T  Y  C  S  E  N  E  T  L  L  T  R  I  L  K  E  E  W  G  H  E  G  I   219

 658  GTCGTATCGGACTGGGGCGCCGTCAACGAAGCGGCTGCGAGCGTGGCGGCCGGCATGGAGCTGGAGATGCCGTCCAGCCATGGCATCGGCCAAAGGAAAATCGTGGCGGCGGTGGAAAGC   777
 220  V  V  S  D  W  G  A  V  N  E  A  A  A  S  V  A  A  G  M  E  L  E  M  P  S  S  H  G  I  G  Q  R  K  I  V  A  A  V  E  S   259

 778  GGAGAACTGTCCGTCGAGGCGCTGGATCGGGCAGTGACGCGGCTTTTGACTGTGATTTTCAAAGCTGTCGACAGCCGGAAGACGGACGCCACTTACGACAAGGAAGCGCATCACTTACTT   897
 260  G  E  L  S  V  E  A  L  D  R  A  V  T  R  L  L  T  V  I  F  K  A  V  D  S  R  K  T  D  A  T  Y  D  K  E  A  H  H  L  L   299

 898  GCCCGCGAAATCGCCCGCGAATCGATGGTGCTGCTCAAAAATGAAGGCAATCTGCTCCCGCTGGCAAAGACGGGCAAACTGGCGATCATCGGAGCCATGGCTGAAGCAGGTTCGATACCAA  1017
 300  A  R  E  I  A  R  E  S  M  V  L  L  K  N  E  G  N  L  L  P  L  A  K  T  G  K  L  A  I  I  G  A  M  A  E  Q  V  R  Y  Q   339

1018  GGTGGCGGAAGCTCCCACATCAAGCCGACAAAGCTGGATAGCATCAGGGACGAGATCGAAAAATCCGCCAGAAGTGCGGAAATCCGTTATTCGAAAGGGTATCTTCTCGAAAGCGACGAG  1137
 340  G  G  G  S  S  H  I  K  P  T  K  L  D  S  I  R  D  E  I  E  K  S  A  R  S  A  E  I  R  Y  S  K  G  Y  L  L  E  S  D  E   379

1138  AGCGACGAGTCTTTGCTGAACGAGGCGAAGCAAGCCGCAGCTGACTCTGATGTCGCGGTGCTGTTCGTCGGGCTGCCGGACCGTTACGAATCCGAAGGCTACGATCGGACGCATCTGAAT  1257
 380  S  D  E  S  L  L  N  E  A  K  Q  A  A  A  D  S  D  V  A  V  L  F  V  G  L  P  D  R  Y  E  S  E  G  Y  D  R  T  H  L  N   419

1258  TTGCCGGCTAACCACATCGAACTGATCGAGCGGATCGCATCCGTTCAGCCGAACGTCGTTGTGATCTTGAGCAACGGTTCTCCCGTCGTTATGCCGTGGCTGGGTCATGCGAAGGCCGTG  1377
 420  L  P  A  N  H  I  E  L  I  E  R  I  A  S  V  Q  P  N  V  V  V  I  L  S  N  G  S  P  V  V  M  P  W  L  G  H  A  K  A  V   459

1378  CTCGAAGCTTACCTGGGCGGTCAGGCTGCGGCGGAGCGATCGCCGACCTGTTGTTCGGCGACGCCAATCCGAGCGGCAAGCTGGCGAGACGTTCCCGCATAGCCTGAAGCACAATCCG  1499
 460  L  E  A  Y  L  G  G  Q  A  A  G  G  A  I  A  D  L  L  F  G  D  A  N  P  S  G  K  L  A  E  T  P  P  H  S  L  K  H  N  P   499

1498  TCCCATCCTTTTTATCCTGGCGAGGGCGATCGGACGGAATACCGGGAAGGCATTTTTGTCGGTTATCGCTATTTCGACGCGAAGGATATAGAGCCGCTGTTTCCGTTCGGACACGGCTTA  1619
 500  S  H  P  F  Y  P  Q  E  G  D  R  T  E  Y  R  E  G  I  F  V  G  Y  R  Y  F  D  A  K  D  I  E  P  L  F  P  F  G  H  G  L   539

1618  AGCTATACGGCGTTTTCCTATTCCGGATTGAAGCTGGACAAAAGCGAGATGACAGACCGGGACATCGTGCAAGTCCGCGTCAACGTGAAGAACACGGGGGGACGGTTCGGCAAGGAAACC  1739
 540  S  Y  T  A  F  S  Y  S  G  L  K  L  D  K  S  E  M  T  D  R  D  I  V  Q  V  R  V  N  V  K  N  T  G  G  R  F  G  K  E  T   579

1738  GTTCAGCTTTACGTCCACAGTCGGAATTCCAGCGTCATTCGTCCGGAAAAAGAGCTGAAAGGCTTTGCGAAGGTATCGTTAAACCCGGAGGAAGAACAGACGGTTACGTTCGCGCTTGAT  1859
 580  V  Q  L  Y  V  H  S  R  N  S  S  V  I  R  P  E  K  E  L  K  G  F  A  K  V  S  L  N  P  E  E  E  Q  T  V  T  F  A  L  D   619

1858  AAACGAAGCTTTGCCTATTACAACGCGGAATTGAAAGAGTGGCATGCTGAAACGGGCGAATATGAAATATTGATCGGCAGCTCTTCGCGCGGATATCGCGCTTCGGACGGCATTGACGGTC  1979
 620  K  R  S  F  A  Y  Y  N  A  E  L  K  E  W  H  A  E  T  G  E  Y  E  I  L  I  G  S  S  S  R  D  I  A  L  R  T  A  L  T  V   659

1978  CAGTCCACGACCGAAATCGTCCCAACATTTCATCGGAATACGACACTCGGAGAGCTGATGGAAAATCCGGCAACGCTCCCGATTCTTGCGCACTTGCAGAGCATGGCGCCGCAACAGCAG  2099
 660  Q  S  T  T  E  I  V  P  T  F  H  R  N  T  T  L  G  E  L  M  E  N  P  A  T  L  P  I  L  A  H  L  Q  S  M  A  P  Q  Q  Q   699

2098  GCGCAATCGGACTCGGTGTCCCCAGACATGATGATGGCGATGATGCGATACATGCCGCTGCGCGCGCTGCTTCCCTTTACCGGCGGCGCGATGACGGAAGAGACGCTTGGCATGTTGCTG  2219
 700  A  Q  S  D  S  V  S  P  D  M  M  M  A  M  M  R  Y  M  P  L  R  A  L  L  P  F  T  G  G  A  M  T  E  E  T  L  G  M  L  L   739

2218  GAGCAGTTTAATCAGGCCGTTCGCGGCGAAAAGAATCAACCTCATGCAAGCGAGGGAAGTTCTGCGGCTTTTAACGAATACTCGACGCTGGGCGACCTCTTGGCTCACGAAGCAGCTGTT  2339
 740  E  Q  F  N  Q  A  V  R  G  E  K  N  Q  P  H  A  S  E  G  S  S  A  A  F  N  E  Y  S  T  L  G  D  L  L  A  H  E  A  A  V   779

2338  GCTGTATTAGAAAAGCATCTCCCCGGCATATCGACGAATCCGATGATCAGCATGGGGAAAGGACTCACTCTCAAGCAACTGGCCGGCATTCCGCAAGCGAATATACCCGAGGAGTTAATC  2459
 780  A  V  L  E  K  H  L  P  G  I  S  T  N  P  M  I  S  M  G  K  G  L  T  L  K  Q  L  A  G  I  P  Q  A  N  I  P  E  E  L  I   819

2458  TCTACAATTGTGACCGACTTGAGTGTAGTCAGAGGATAATTGCCCGATCGTACAGGAGGGCTTGCCCTTTTGTACGATTTTTATCAAGGCTGTTTTTGTTAAAATAAACGTGTAACCGGC  2579
 820  S  T  I  V  T  D  L  S  V  V  R  G  *                                                                                    831

2578  GAGGATAGAGGGGCTCCAGAAAGATAAAACGATAAATGGAGGAATAAAAAAAAATGGCATTCAACGAAAACTCAACATTAGGCGACCTTTTGGCGAATGAAGCGGCTGTTGCCGTTCTGG  2699
2698  AAAAGCATCTTCCCGGCATTTCGACGAATCCGATGGCCAGTATGGGGAAATTATTGTCACTTAAGCAACTGGCCGCCATTCCGCAAGCGAACATGCCGCAGGATTTAATAAACGCGGTCG  2819
2818  TTGCCGACTTGGCAGGAATCGAAGAGAAAGGCGTTCAGGCGACTTCCGAGGAAAGCCCTTTGGCATCTGTGGCGGGGGCATCCGCGAATAGGGATGAGACGCTGCCGGAAGGCGCCGTTG  2939
2838  CTATCGTAACGGGAGCGGGTAGCGGGATC
```

(B)

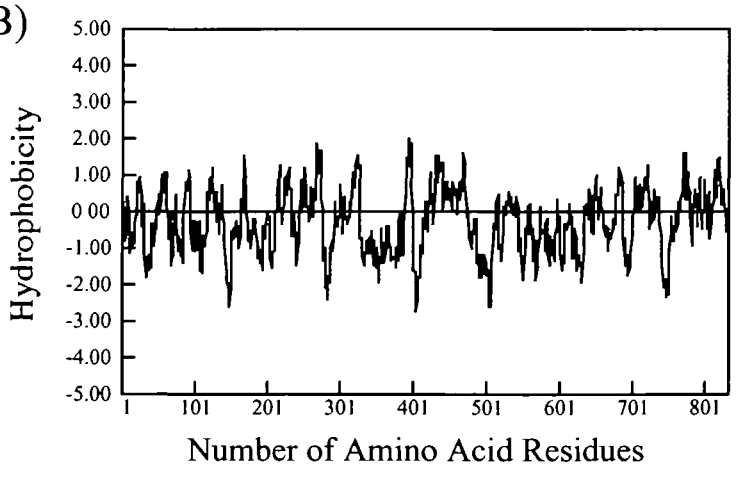

Fig. 1. **DNA and deduced amino acid sequences (A), and hydropathy plot (B) of the TS12 GCase.** (A) The deduced amino acid sequence is shown as a one-letter code below the nucleotide sequence. The possible Shine-Dalgarno sequence is underlined. The termination codon is indicated by an asterisk (*). (B) Hydropathy analysis of the coding region was performed based upon the deduced amino acids according to Kyte and Doolittle (*23*).

pH 6.5 or Tris-HCl buffer, pH 7.5, containing 0.1% (w/v) TDC. Following incubation at 37°C for a given period, the reaction was stopped by heating in a boiling water bath for 5 min. The sample was dried with a Speed Vac SC110 (Savant Ins), and the residue was dissolved in 10 μl of chloroform/methanol (2:1, v/v) and applied to a TLC plate, which was then developed with chloroform/methanol/25% ammonia (90:20:0.5, v/v, solvent I). C12-NBD-Cer (NBD-dodecanoylsphingosine) released through the action of the enzyme and the remaining C12-NBD-GlcCer were visualized under UV illumination, and then analyzed and quantified with a Shimadzu CS-9300PC chromatoscanner (excitation 475 nm, emission 525 nm). Assay III ([^{14}C]GlcCer being used as a substrate); the reaction mixture contained 100 pmol of [^{14}C]GlcCer and an appropriate amount of the enzyme in 20 μl of 25 mM sodium phosphate buffer, pH 6.5, containing 0.05% (w/v) TDC. Following incubation at 37°C for a specified period, the reaction was stopped by heating in a boiling water bath for 5 min. The sample was dried, dissolved in 10 μl of chloroform/methanol (2:1, v/v), and then applied to a TLC plate, which was then developed with solvent I. The [^{14}C]Cer released through the action of the enzyme and the remaining [^{14}C]GlcCer were analyzed with an imaging analyzer (BAS1500 model, Fuji Film).

Protein Assay and Polyacrylamide Gel Electrophoresis—The protein content was determined by the bicinchononic acid method (Pierce, Rockford) or SDS-PAGE using bovine serum albumin as the standard. SDS-PAGE was carried out according to the method of Laemmli (*16*). The proteins on the SDS-PAGE gel were visualized by staining with Coomassie Brilliant Blue.

Western blotting—The proteins, separated by 10% SDS-PAGE, were transferred to a PVDF membrane using a semi-dry blotter (BIO-RAD, Hercules, CA). The membrane was then incubated with anti-polyhistidine tag mouse IgG monoclonal antibodies (Invitrogen Japan, Tokyo) for 6 h at room temperature. The bands were visualized with HRP-labeled anti-mouse IgG antibodies and a peroxidase-staining kit (Nacalai Tesque, Kyoto).

RESULTS

Molecular Cloning, Sequencing, and Alignment of GCase of Paenibacillus sp. TS12—A clone (pGlc4) carrying a 3,149-bp *Sau*3AI fragment was isolated from a TS12 genomic library by expression cloning, as described under "MATERIALS AND METHODS." One open reading frame (ORF), 2,493 bp long and encoding 831 amino acids, was identified in the clone (Fig. 1A). A possible Shine-Dalgarno ribosome binding sequence started 7 bases upstream from the initiation codon, GTG. The molecular weight and pI of the enzyme were estimated to be 90,707 and 5.34, respectively, from the deduced amino acid sequence. No hydrophobic region was found in the N-terminal region, which was also clearly indicated by hydrophobicity plots (Fig. 1B). The deduced amino acid sequence of TS12 GCase exhibited 58, 51, and 29% identity to that of GH family 3 β-glucosidases from *Clostridium thermocellum, Ruminococcus*

Fig. 2. **Alignment of the TS12 GCase and other bacterial β-glucosidases.** TS12 GCase is aligned with other bacterial β-glucosidases. Residues conserved in more than 75% of the strains are shown on a black background, while those found in all the strains are indicated by asterisks (*). Gaps inserted into the sequences are indicated by points (•). Alignment was performed using CLUSTAL W (*24*). The conserved catalytic Asp residue (⋆) is indicated below the sequence. Patterns conserved throughout the GH family 3 β-glucosidases are shown by underlining.

albus, and *Aspergillus aculeatus*, respectively (Fig. 2). The SDW motif, the Asp of which would function as a nucleophile (*17*), and typical patterns conserved throughout the GH family 3 β-glucosidases were found in the sequence of the TS12 enzyme. The identity of *Paenibacillus* GCase to the human GCase is only 6% at the amino acid level.

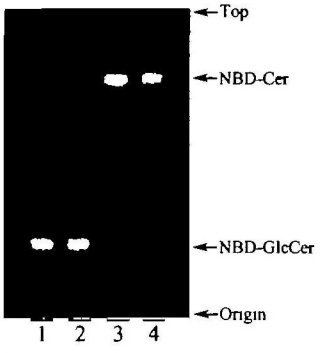

Fig. 3. **Hydrolysis of NBD-GlcCer by the recombinant GCase.** TLC showing the hydrolysis of NBD-GlcCer. Aliquots of 200 pmol of NBD-GlcCer were incubated with 10 mU of the enzyme at 37°C for 30 min in 20 μl of 25 mM sodium phosphate buffer, pH 6.5, containing 0.1% TDC. The hydrolysis of NBD-GlcCer was examined as described under "MATERIALS AND METHODS." Lane 1, NBD-GlcCer; lane 2, cell lysate of mock transfectant + NBD-GlcCer; lane 3, cell lysate of transfectant carrying pETGlc4 + NBD-GlcCer; lane 4, NBD-Cer.

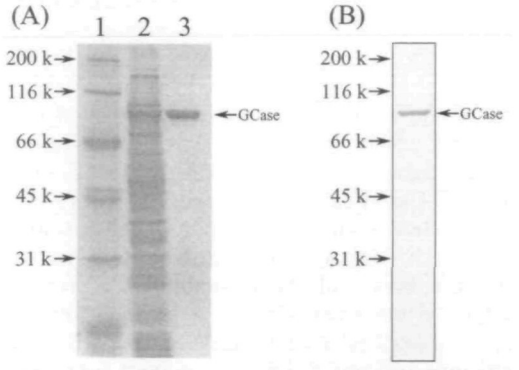

Fig. 4. **SDS-PAGE of the recombinant purified GCase.** (A) *E. coli* BL21(DE3)pLysS cells transformed with pETGlc4 were cultured at 25°C for 1 h after the addition of IPTG to a final concentration 0.1 mM. The cell lysate was analyzed by 10% SDS-PAGE. Proteins were stained with Coomassie Brilliant Blue. Lane 1, marker proteins (molecular weights in parentheses): rabbit muscle myosin (200,000), *E. coli* β-galactosidase (116,250), bovine serum albumin (66,200), ovalbumin (45,000), carbonic anhydrase (31,000), soybean trypsin inhibitor (21,500), and lysozyme (14,400); lane 2, cell lysate of the transfectant carrying pETGlc4; lane 3, purified recombinant GCase. (B) Western blotting of the recombinant GCase. The cell lysate was analyzed by 10% SDS-PAGE and then transferred to a PVDF membrane for detection with anti-His tag (C-term) monoclonal antibodies as described under "MATERIALS AND METHODS."

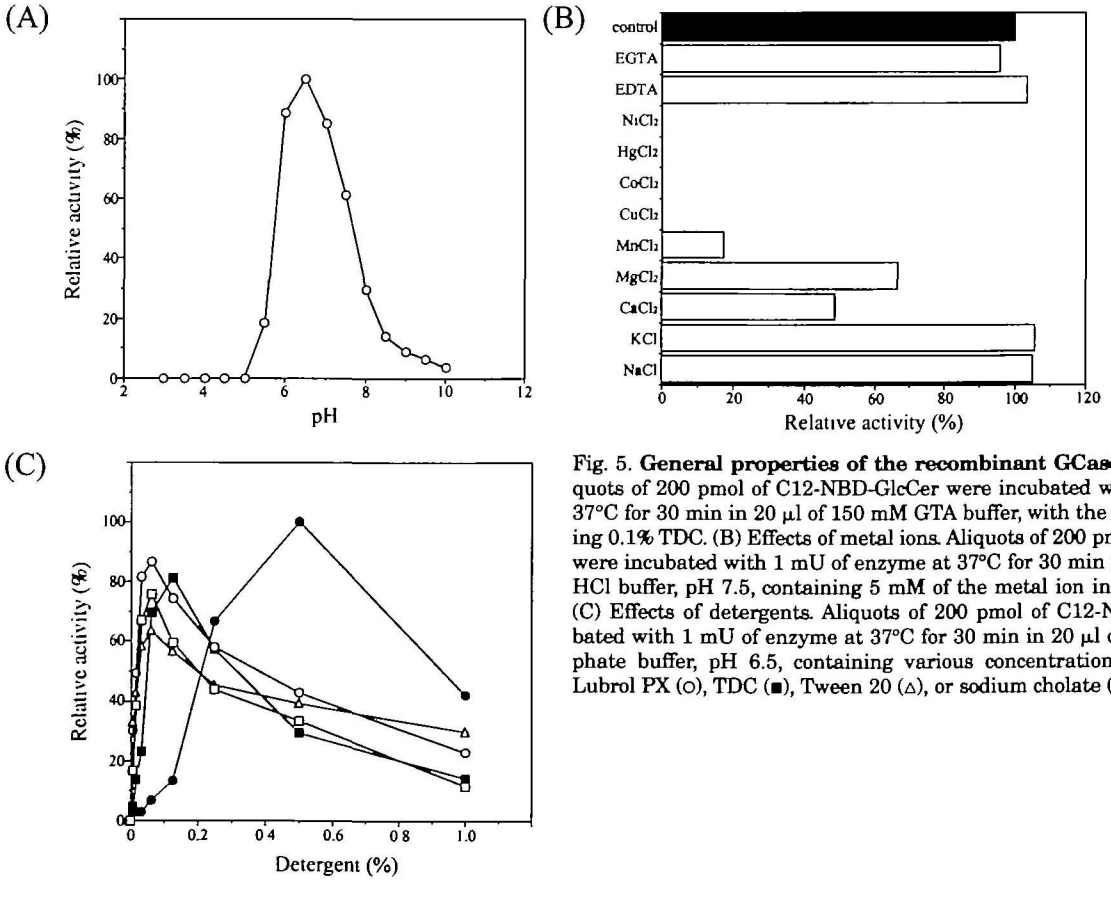

Fig. 5. **General properties of the recombinant GCase.** (A) Effect of pH. Aliquots of 200 pmol of C12-NBD-GlcCer were incubated with 1 mU of enzyme at 37°C for 30 min in 20 μl of 150 mM GTA buffer, with the indicated pHs, containing 0.1% TDC. (B) Effects of metal ions. Aliquots of 200 pmol of C12-NBD-GlcCer were incubated with 1 mU of enzyme at 37°C for 30 min in 20 μl of 25 mM Tris-HCl buffer, pH 7.5, containing 5 mM of the metal ion indicated and 0.1% TDC. (C) Effects of detergents. Aliquots of 200 pmol of C12-NBD-GlcCer were incubated with 1 mU of enzyme at 37°C for 30 min in 20 μl of 25 mM sodium phosphate buffer, pH 6.5, containing various concentrations of Triton X-100 (□), Lubrol PX (○), TDC (■), Tween 20 (△), or sodium cholate (●).

Expression and Purification of the Recombinant GCase—Expression vector pETGlc4 was constructed by inserting a fragment of the coding sequence of GCase between the *Nhe*I and *Xho*I sites of plasmid pET23a. In pETGlc4, transcription of the recombinant gene is controlled by the T7 promoter and can be induced by IPTG. A cell lysate of the transfectant carrying pETGlc4 hydrolyzed NBD-GlcCer to produce NBD-Cer (Fig. 3, lane 3), whereas that of the mock transfectant showed no activity (lane 2), indicating that the gene encodes the GCase. The recombinant GCase was purified from the lysate of the recombinant *E. coli* carrying pETGlc4 by chromatography on a HiTrap Chelating HP column, as described under "MATERIALS AND METHODS." The final preparation of the enzyme gave a single protein band corresponding to a molecular weight of 90.7 kDa on SDS-PAGE with Coomassie brilliant blue staining (Fig. 4A, lane 3), which exactly corresponded to the band obtained on Western blotting using anti-polyhistidine tag (C-term) monoclonal antibodies (Fig. 4B). The molecular weight of the enzyme estimated by SDS-PAGE well agreed with that deduced from the nucleotide sequence.

General Properties of the Recombinant GCase—The activity of the enzyme was found to be optimal at around pH 6.5 with 150 mM GTA buffer when C12-NBD-GlcCer was used as a substrate (Fig. 5A). The enzymic activity was strongly inhibited by various metal ions; complete inhibition was observed with Cu^{2+}, Co^{2+}, Hg^{2+}, and Ni^{2+}, and strong inhibition with Mn^{2+}, Mg^{2+}, and Ca^{2+} at 5 mM, whereas K^+, Na^+, EDTA, and EGTA did not have a significant effect on the activity (Fig. 5B). This enzyme was activated on the addition of detergents when C12-NBD-GlcCer, but not *p*NP-β-Glc, was used as a substrate. For Triton X-100, Lubrol PX, TDC and Tween 20, the optimum concentrations were found to be 0.05–0.15%, whereas it was 0.5% for sodium cholate (Fig. 5C).

Substrate Specificity of the Recombinant GCase—First, the substrate specificity of the enzyme was examined using various (glyco)sphingolipids. The enzyme hydrolyzed NBD-GlcCer efficiently, whereas NBD-labeled GalCer, GM1a and SM were completely resistant to hydrolysis by the enzyme. Among ^{14}C-labeled GlcCers containing different fatty acid molecules, the enzyme most preferred the GlcCer containing lauric acid (C12:0), followed by those containing palmitic acid (C16:0) and stearic acid (C18:0) in this order. The initial reaction velocity of the enzyme for NBD-GlcCer (NBD-C12:0) was 5-times faster than that for [^{14}C]GlcCer ([^{14}C]C12:0), suggesting that the susceptibility of the GlcCer to the enzyme increases with the attachment of NBD to the fatty acid moiety of ceramide. It should be noted that the same tendency was observed for the neutral ceramidase (*18*). To elucidate the specificity of the enzyme in more detail, various *p*NP-glycopyranosides were employed. It was found that the enzyme hydrolyzed *p*NP-β-Glc most efficiently, but did not hydrolyze *p*NP-β-Gal, *p*NP-β-GlcNAc, or *p*NP-α-Glc, indicating that the enzyme is a β-glucosidase. However, the cellobioside, a glucose dimer linked by a β1,4 linkage, is strongly resistant to the enzyme. Interestingly, the enzyme showed weak activity toward *p*NP-β-Xyl (Table I). The apparent K_m and V_{max} values for NBD-GlcCer and *p*NP-β-Glc were 223 μM and 1.60 μmol/min/mg of protein, and 593 μM and 112 μmol/min/mg of protein, respectively.

Analysis of Mutants D223E and D223N—We found the SDW motif, which is a putative catalytic nucleophile of GH family 3 β-glucosidases, in the deduced amino acid sequence of the TS12 GCase. Thus, to elucidate whether or not Asp223 is crucial for the catalytic reaction of the enzyme, two mutant enzymes, in which Asp-223 was replaced by glutamate (D223E) or asparagine (D223N), were produced by site-directed mutagenesis. Fig. 6 shows the time courses of hydrolysis of NBD-GlcCer by the wild-type and

TABLE I. **Substrate specificity of rGCase.**

Substrate	Relative activity (%)
NBD-GlcCer	100
NBD-GalCer	0
NBD-GM1a	0
NBD-SM	0
[^{14}C]GlcCer (C12:0)	100
(C16:0)	82.6
(C18:0)	40.1
*p*NP-β-Glc	100
*p*NP-β-Gal	
*p*NP-β-GlcNAc	0
*p*NP-β-GalNAc	0
*p*NP-β-Glcβ1-4Glc(Cellobioside)	0.9
*p*NP-β-Xyl	7.3
*p*NP-α-Glc	0
*p*NP-α-Gal	0
*p*NP-α-GlcNAc	0
*p*NP-α-GalNAc	0

Aliquots of 200 pmol of C12-NBD-substrates and 100 pmol of [^{14}C]GlcCer were incubated with 4 mU of enzyme at 37°C for 30 min in 20 μl of sodium phosphate buffer, pH 6.5, containing an appropriate amount of TDC, as described under "MATERIALS AND METHODS." *p*NP-glycosides (100 nmol) were incubated with 0.2 mU of enzyme at 37°C for 30 min in 100 μl of sodium phosphate buffer, pH 6.5. Values are the means for triplicate determinations.

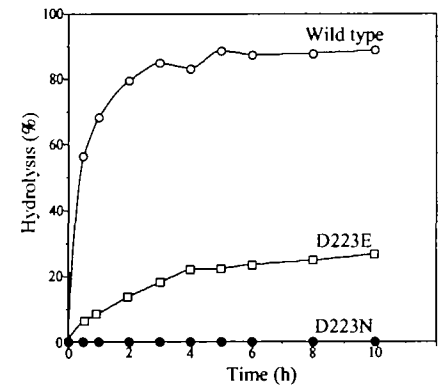

Fig. 6. **Time courses of the hydrolysis of NBD-GlcCer by the wild-type and mutant GCases.** Aliquots of 200 pmol of NBD-GlcCer were incubated with 10 ng of the wild-type enzyme, 600 ng of D223E, or 3 μg of the D223N mutant enzyme at 37°C for the periods indicated in 20 μl of 25 mM sodium phosphate buffer, pH 6.5, containing 0.1% TDC. The hydrolysis of NBD-GlcCer was examined as described under "MATERIALS AND METHODS."

TABLE II. **Mutation of catalytic Asp in the TS12 enzyme.**

Enzyme	K_m (μM)	V_{max} (μmol/min/mg protein)
Wild type	593	111.8
D223E	1,350	8.4
D223N	ND	ND

ND, not determined. Values are the means for triplicate determinations.

mutant GCases. It was found that the replacement of Asp-223 by Glu greatly reduced the reaction velocity whereas that by Asn completely abolished the activity. The apparent K_m and V_{max} were also determined for the wild-type and mutant CGase (D223E) using *p*NP-β-Glc as a substrate. Replacement of Asp223 with Glu increased the K_m by 2.3-fold whereas it decreased the V_{max} by one-thirteenth (Table II). These results indicate that Asp-223 in the SDW motif of the GCase is crucial for hydrolysis of the β-glucosidic linkage of GlcCer as well as *p*NP-β-Glc, suggesting that the residue functions as a catalytic neucleophile in β-glucosidases belonging to GH family 3.

DISCUSSION

Recently, *O*-glycoside hydrolases, which hydrolyze the *O*-glycosidic linkage between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety, have been classified into 87 GH families based on their amino acid similarities but not substrate specificity (*19*). The TS12 GCase can be placed in GH family 3 based on the following observations; (i) the deduced amino acid sequence of the TS12 enzyme shows high similarity to β-glucosidases belonging to GH family 3 (*20*) such as *C. thermocellum*, *R. albus*, but *A. aculeatus*, and little similarity to the known GCases belonging to GH family 30. (ii) Although exo-type β-glucosidases [EC 3.2.1.21] mainly belong to GH family 1 or 3, the TS12 enzyme showed no similarity to the GH family 1 enzymes. (iii) Furthermore, the TS12 enzyme was found to behave as a GH family 3 enzyme, being active against *p*NP-β-xyloside but not *p*NP-β-galactoside (Table I). In contrast, GH family 1 β-glucosidases can hydrolyze *p*NP-β-galactoside but not *p*NP-β-xyloside. In conclusion, the TS12 enzyme can be classified into GH family 3 based on sequence similarity, and as a member of EC 3.2.1.45 given the substrate specificity for GlcCer. It remains to be elucidated whether or not the known β-glucosidases belonging to GH family 3 can hydrolyze GlcCer.

On resolution of the three-dimensional structure of GH family 3 β-glucosidases, Asp-285 and Glu-491 in the barley enzyme were suggested to function as a nucleophile and a general acid/base catalyst, respectively (*17*). The function of Asp-285 in the SDW motif as a nucleophile was also supported by active-site affinity labeling (*21*) and site-directed mutagenesis (*22*). This study clearly indicates that Asp-223 in the SDW motif of the TS12 enzyme is crucial for the reaction in which the β-glucosidic linkage of GlcCer as well as *p*NP-substrate is hydrolyzed, and suggests that Asp-223 functions as a catalytic nucleophile.

REFERENCES

1. Hakomori, S., Handa, K., Iwabuchi, K., Yamamura, S., and Prinetti, A. (1998) New insights in glycosphingolipid function: 'glycosphingolipid domain' a cell surface assembly of glycosphingolipids with signal transducer molecules, involved in cell adhesion coupled with signaling. *Glycobiology* **8**, xi–xix
2. Pshezhetsky, A.V. and Ashmarina, M. (2001) Lysosomal multienzyme complex: biochemistry, genetics, and molecular pathophysiology. *Prog. Nucleic Acid Res. Mol. Biol.* **69**, 81–114
3. Schwarz, A. and Futerman, A.H. (1997) Distinct roles for ceramide and glucosylceramide at different stages of neuronal growth. *J. Neurochem.* **17**, 2929–2938
4. Boldin, S.A. and Futerman, A.H. (2000) Up-regulation of glucosylceramide synthesis upon stimulation of axonal growth by basic fibroblast growth factor. *J. Biol. Chem.* **275**, 9905–9909
5. Korkotian, E., Schwarz, A., Pelled, D., Schwarzmann, G., Segal, M., and Futerman, A.H. (1999) Elevation of intracellular glucosylceramide levels results in an increase in endoplasmic reticulum density and in functional calcium stores in cultured neurons. *J. Biol. Chem.* **274**, 21673–21678
6. Tsuji, S., Choudary, P.V., Martin, B.M., Winfield, S., Barranger, J.A., and Ginns, E.I. (1986) Nucleotide sequence of cDNA containing the complete coding sequence for human lysosomal glucocerebrosidase. *J. Biol. Chem.* **261**, 50–53
7. Carstea, E.D., Murray, G.J., and O'Neill, R.R. (1992) Molecular and functional characterization of the murine glucocerebrosidase gene. *Biochem. Biophys. Res. Commun.* **184**, 1477–1483
8. Martinez-Arias, R., Calafell, F., Mateu, E., Comas, D., Andres, A., and Bertranpetit, J. (2001) Sequence variability of a human pseudogene. *Genome Res.* **11**, 1071–1085
9. McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du *et al.* (2001) Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* **413**, 852–856
10. Tani, M., Kita, K., Komori, H., Nakagawa, T., and Ito, M. (1998) Enzymatic synthesis of ω-amino-ceramide: preparation of a sensitive fluorescent substrate for ceramidase. *Anal. Biochem.* **263**, 183–188
11. Nakagawa, T., Tani, M., Kita, K., and Ito, M. (1999) Preparation of fluorescence-labeled GM1 and sphingomyelin by the reverse hydrolysis reaction of sphingolipid ceramide *N*-deacylase as substrates for assay of sphingolipid-degrading enzymes and for detection of sphingolipid-binding proteins. *J. Biochem.* **126**, 604–611
12. Morigaki, M., Nagao, M., Miura, Y., Takahara, K., Tada, M., Nakajima, S., Baba, N., and Shimizu, S. (1998) Syntheses of sphingomyelins and ceramide bearing a docosahexanoyl or arachidonoyl group. *Biosci. Biotechnol. Biochem.* **62**, 2070–2072
13. Ito, M., Mitsutake, S., Tani, M., and Kita, K. (1999) Enzymatic synthesis of [^{14}C]ceramide, [^{14}C]glycosphingolipids and ω-aminoceramide. *Methods Enzymol.* **311**, 682–689
14. Saito, H. and Miura, K. (1963) Preparation of transforming deoxyribonucleic acid by phenol treatment. *Biochem. Biophys. Acta* **72**, 619–629
15. Ho, S.N., hunt, H.D., Horton, R.N., Pullen, J.K., and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using polymerase chain reaction. *Gene* **77**, 51–59
16. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685
17. Harvey, A.J., Hrmova, M., Gori, R.D., Varghese, J.N., and Fincher, G.B. (2000) Comparative modeling of the three-dimensional structures of family 3 glycoside hydrolases. *Proteins* **41**, 257–269
18. Tani, M., Okino, N., Mitsutake, S., and Ito, M. (1999) Specific and sensitive assay for alkaline and neutral ceramidases involving C12-NBD-ceramide. *J. Biochem.* **125**, 746–749
19. Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280**, 309–316
20. Rojas, A. and Romeu, A. (1996) A sequence analysis of the β-glucosidase sub-family B. *FEBS Lett.* **378**, 93–97
21. Bause, E. and Legler, G. (1980) Isolation and structure of a tryptic glycopeptide from the active β-glucosidase A_3 from *Aspergillus wentii*. *Biochem. Biophys. Acta* **626**, 459–465
22. Keresztessy, Z., Brown, K., Dunn, M.A., and Hughes, M.A. (2001) Identification of essential active-site residues in the cyanogenic β-glucosidase (linamarase) from cassava (*Manihot esculenta* Crantz) by site-directed mutagenesis. *Biochem. J.* **353**, 199–205
23. Kyte, J. and Doolittle, R.F. (1982) A simple methods for displaying the hydropathic character of protein. *J. Mol. Bol.* **157**, 105–132
24. Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680