Paenibacillus sp. TS12 glucosylceramidase: kinetic studies of a novel sub-family of family 3 glycosidases and identification of the catalytic residues

Krisztina PAAL^{*1}, Makoto ITO[†] and Stephen G. WITHERS^{*2}

^{*}Protein Engineering Network of Centres of Excellence of Canada and Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1, and

[†]Department of Bioscience and Biotechnology, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

GCase (glucosylceramidase) from *Paenibacillus* sp. TS12, a family 3 glycosidase, hydrolyses the β -glycosidic linkage of glucosylceramide with retention of anomeric configuration via a two-step, double-displacement mechanism. Two carboxyl residues are essential for catalysis, one functioning as a nucleophile and the other as a general acid/base catalyst. *p*-Nitrophenyl β -D-glucopyranoside [$K_m = 0.27 \pm 0.02$ mM and $k_{cat}/K_m = (2.1 \pm 0.2) \times 10^6$ M⁻¹ · s⁻¹] and 2,4-dinitrophenyl β -D-glucopyranoside [$K_m = 0.16 \pm 0.02$ mM and $k_{cat}/K_m = (2.9 \pm 0.4) \times 10^6$ M⁻¹ · s⁻¹] were used for continuous assay of the enzyme. The dependence of k_{cat} (and k_{cat}/K_m) on pH revealed a dependence on a group of $pK_a \leq 7.8$ in the enzyme–substrate complex which must be protonated for catalysis. Incubation of GCase with 2,4-dinitrophenyl 2-deoxy-2-fluoro- β -D-glucopyranoside caused time-dependent inactivation ($K_i = 2.4 \pm 0.7$ mM and $k_i = 0.59 \pm 0.05$ min⁻¹) due

to the accumulation of a trapped glycosyl–enzyme intermediate. Electrospray ionization MS analysis of the peptic digest of this complex showed that the enzyme was covalently labelled by the reagent at Asp-223, consistent with its role as nucleophile. A mutant modified at this residue (D223G) showed substantially reduced activity compared with the wild type ($> 10^4$), but this activity could be partially restored by addition of formate as an external nucleophile. Kinetic analysis of the mutant E411A indicated that Glu-411 serves as the general acid/base catalytic residue since this mutant was pH-independent and since considerable GCase activity was restored upon addition of azide to E411A, along with formation of a glycosyl azide product.

Key words: azide effect, general acid/base, glucosylceramidase, inactivation, nucleophile.

INTRODUCTION

Several different glycoside hydrolase enzymes act on glycosphingolipid substrates to degrade them to their constituents. Endoglycosylceramidase [1] (EC 3.2.1.123) cleaves the oligosaccharide moiety from the ceramide group in a single step, whereas galactosylceramidase (EC 3.2.1.46) and glucosylceramidase (GCase; β -glucocerebrosidase; EC 3.2.1.45), being exoglycosidases, catalyse the hydrolysis of the β -glycosidic linkage of the respective glycosylceramide to yield a monosaccharide and ceramide. GCase plays a central role in mammalian glycosphingolipid metabolism. Deficiency in its activity causes Gaucher's disease, a serious lysosomal storage disorder [2].

All GCases cloned and studied previously are of mammalian origin (e.g. human [3] and murine [4]) and belong to family 30 of the glycoside hydrolases (<http://afmb.cnrs-mrs.fr/CAZY/index.html>) [5]. The GCase cloned from *Paenibacillus* sp. TS12 and over-expressed in *Escherichia coli* recently [6] is the first bacterial GCase to be studied. This 91 kDa protein shows no sequence similarity to its mammalian counterparts and has been assigned to family 3 of the glycoside hydrolases. This family is a large and diverse group of enzymes including β -glucosidases, β -xylosidases and β -N-acetylhexosaminidases. Based on sequence alignments of 99 members of this family a phylogenetic tree was constructed, further dividing the family 3 glycosidases into six subfamilies [7]. PSI-BLAST search (<http://www.ncbi.nlm.nih.gov/BLAST/>) results indicated greatest sequence

identity and similarity (59 % identity, 73 % similarity) of the *Paenibacillus* GCase to the *Clostridium thermocellum* β -glucosidase (also a family 3 glycosidase); therefore *Paenibacillus* GCase was assigned to the same subfamily (subfamily 5).

Despite the size of family 3 of the glycosidases few of its members have been studied in detail. Only one three-dimensional structure is available at present: that of the barley exo-1,3-1,4-glucanase [8]. β -Glucosidases from *Flavobacterium meningosepticum* [9–11] and *Aspergillus wentii* [12] have undergone detailed kinetic investigations; however, the entire sequence of this latter enzyme is not known. Family 3 glycoside hydrolases are retaining β -glycosidases, catalysing glycosidic bond cleavage via a two-step, double-displacement mechanism [13] (Scheme 1), leading to retention of configuration at the anomeric centre of the glycon moiety. Two carboxyl residues are essential for catalysis, one acting as a nucleophile and the other as a general acid/base catalyst. Identification of the catalytic nucleophiles of several family 3 glycosidases by covalent labelling and MS analysis [10,14–16] has demonstrated the conserved nature of this residue across the family. Based on sequence alignments, Asp-223 of the *Paenibacillus* GCase was identified as the putative catalytic nucleophile and was in fact shown to be crucial for catalysis by site-directed mutagenesis [6].

In contrast to the catalytic nucleophile, the general acid/base residue does not appear to be well-conserved in glycoside hydrolase family 3. Covalent labelling and MS analysis identified Glu-473 as the general acid/base catalyst of the *F. meningosepticum*

Abbreviations used: CAPSO, 3-(cyclohexylamino)-2-hydroxy-1-propanesulphonic acid; DNPG, 2,4-dinitrophenyl β -D-glucopyranoside; ESI, electrospray ionization; 2FDNPG, 2,4-dinitrophenyl 2-deoxy-2-fluoro- β -D-glucopyranoside; GCase, glucosylceramidase; PNPG, *p*-nitrophenyl β -D-glucopyranoside.

¹ Present address: Department of Chemistry, Brock University, St. Catharines, Ontario, Canada L2S 3A1.

² To whom correspondence should be addressed (e-mail wITHERS@chem.ubc.ca).

Scheme 1 Mechanism of substrate hydrolysis by retaining glycosidases

1	387	EAKQAADS	SDVAVLFVGLPDRY	E SEGYDRT
2	396	EAKKAASS	SDVAVVFAGLPDEY	E SEGFDR
3	396	EAKQTALK	ADVAVIFAGLP	PEHY E CEGYDRT
4	399	QAVEAAMN	ADKVVFAGLP	DSF E SEGFDRK
5	469	SRRCVEHE	VSEAQAIIITIGRQ	AGEGLDRK
6	433	SEKEVHK	LAKKNDVAVIVIS	SRIS E CEGYDRK
7	434	SEKEIKKA	AKKNDVAVVVIS	SRIS E CEGYDRK
8	554	YAVRVAKS	VDVILCVGLTA	AEW E TEGEDRK
9	566	NAAELA	AAKHDKAVLIIGLNGEW	E TEGYDRE
10	535	EAVETARK	SDIVLLLVR	EGREWDTEGLDLP
11	562	RAAAAAQA	ADVAVVVG	LTEEE E TESVDKS
12	583	EAVRAAGA	ADAADVVG	TTEHGE E SEGYDRT
13	539	AAVERARD	SDVAVVFAYDDG	--- A ETADRT
14	501	RAVSIART	ARTAVVFAYDDG	--- S EDGDRT
15	456	SAVRAARA	ADVAVVYATQFT	--- F EGMDAP
16	470	RAARVAAG	ADVALVFANQWI	--- G EANDAQ
17	462	AAAALARQ	SDIVILFANQWM	--- S EGMDAY
18	418	EAVLAARR	ADIAIVFAIRAE	--- G EGFDSA
19	449	EAVATANK	ADVVLAIGETAELSG	E SSSRA
20	469	AEFVKSGG	FSYAIVAVGEHPYT	E TKGDNLN

Figure 1 Partial multiple sequence alignment of selected family 3 glycoside hydrolases

Names and GenBank accession numbers are as follows: 1, *Paenibacillus* sp. TS12 GCase, AB084154; 2, *Clostridium thermocellum* β -glucosidase, X15644; 3, *Clostridium stercorarium* β -glucosidase, Z94045; 4, *Ruminococcus albus* β -glucosidase, U92808; 5, *Prevotella ruminicola* β -glucosidase, U35425; 6, *Thermotoga neapolitana* β -glucosidase, Z77856; 7, *Thermotoga maritima* β -glucosidase, AE001690; 8, *Schizosaccharomyces pombe* β -glucosidase, AL355920; 9, *Kluyveromyces fragilis* β -glucosidase, X05918; 10, *Agrobacterium tumefaciens* β -glucosidase, M59852; 11, *Cellulomonas biazotea* β -glucosidase, AF005277; 12, *Streptomyces coelicolor* putative β -glucosidase, AL031013; 13, *Saccharopolyspora erythraea* β -glucosidase, Y14327; 14, *Streptomyces antibioticus* glycosidase OleR, AF055579; 15, *Acetobacter xylinus* β -glucosidase, AB010645; 16, *Cellvibrio gilvus* β -glucosidase, D14068; 17, *Aszospirillum irakense* β -glucosidase (SalA), AF090429; 18, *Mycobacterium tuberculosis* β -glucosidase, Z97050; 19, *F. meningosepticum* β -glucosidase, AF015915; 20, *Hordeum vulgare* (barley) exo-1,3-1,4-glucanase, AF102868. The second column indicates the numbers of the first amino acid in the cited sequence. Bold residues are (partially) conserved.

β -glucosidase [10]. In the X-ray crystal structure of the barley exo-1,3-1,4-glucanase [8] Glu-491 appears to be in an appropriate position to serve as the general acid/base catalyst. This residue, however, does not align with Glu-473 of the *F. meningosepticum* β -glucosidase (Figure 1). Selection of a putative general acid/base residue in the sequence of the *Paenibacillus*

GCase was therefore not straightforward. Multiple sequence alignment of the subfamily 5 enzymes (Figure 1, sequences 1–18) by using the program BioEdit (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>) identified two conserved carboxylates (Glu-411 and Asp-414 of *Paenibacillus* GCase) in the region where the acid/base residue is expected to be located. One of these (Glu-411) also aligned with Glu-473 of the *F. meningosepticum* β -glucosidase; therefore this residue seemed to be the most likely candidate for the role of the general acid/base catalyst.

In this study we describe detailed kinetic analysis of the wild-type *Paenibacillus* GCase and its active-site mutants (D223G and E411A) as well as covalent labelling followed by MS analysis to unequivocally confirm the identity of its catalytically important residues.

EXPERIMENTAL

Reagents and bacterial strains

Growth media components were obtained from Difco. *Pwo* DNA polymerase was from Roche. Zero Blunt PCR cloning kit and *E. coli* TOP 10 cells were obtained from Invitrogen. Deoxynucleoside triphosphates were obtained from Gibco. T4 DNA ligase (Rapid DNA ligation kit) was from Fermentas. Restriction endonucleases were from New England BioLabs. *E. coli* BL21(DE3) cells and His-Bind resin were from Novagen. Gel-extraction and plasmid-DNA-purification kits were from Qiagen. Oligonucleotide synthesis and DNA sequencing were performed at the Nucleic Acids and Peptide Service (NAPS) at the University of British Columbia, Vancouver, BC, Canada. Pepsin was from Boehringer Mannheim. The expression vector containing the gene encoding the His₆-tagged *Paenibacillus* sp. TS12 GCase (pETGlc4) was prepared in the laboratory of M. I. [6]. DNPG (2,4-dinitrophenyl β -D-glucopyranoside) was synthesized by Dr Michael Jahn and 2FDNPG (2,4-dinitrophenyl 2-deoxy-2-fluoro- β -D-glucopyranoside) was synthesized by Dr David Zechel in this laboratory previously. All other reagents were from Sigma-Aldrich Co.

Site-directed mutagenesis by PCR

The expression vector for the nucleophile mutant (D223G) was prepared by a four-primer mutagenic strategy. The following two PCR reactions were performed simultaneously in separate tubes. One reaction mixture contained a forward primer with a *SacII* restriction site (5'-CGT CCG CGG CGC ATA TTG CAG GCG TTC AAA GC-3'; the restriction site is underlined) and the reverse mutagenic primer (5'-CCC AGC CCG ATA CGA CGA TGC CCT CGT GG-3'; the mutant codon is shown in bold). Another reaction mixture contained the forward mutagenic primer (5'-CCA CGA GGG CAT CGT CGT ATC GGG CTG GG-3') and a reverse primer with a *ClaI* restriction site (5'-CCT ATC GAT TCG CGG GCG ATT TCG CGG GCA AG-3'). Both mixtures contained pETGlc4 as a template, 0.2 mM of each deoxynucleoside triphosphate and 0.025 unit/ μ l *Pwo* DNA polymerase. Reaction mixtures were heated to and kept at 94 °C for 3 min, then 30 cycles were performed (30 s at 94 °C; 45 s at 64 °C; 90 s at 72 °C), and the mixtures were kept at 72 °C for an additional 5 min (PerkinElmer GeneAmp PCR System2400). The desired products were purified by extraction from the agarose gel (QIAquick gel-extraction kit) following electrophoresis. These two fragments were then used as templates in a second round of PCR reaction with the forward primer containing the *SacII* restriction site and the reverse primer containing the *ClaI* restriction site. Conditions were the same as in the first round

of reactions. The sole product (≈ 550 bp) was isolated and ligated into a pCR-Blunt cloning vector (Zero Blunt PCR cloning kit) according to the manufacturer's protocol. Chemically competent TOP10 *E. coli* cells were transformed and selected by kanamycin resistance. Single colonies were selected and grown in LB_{kan} (Luria broth containing kanamycin) at 37 °C overnight, plasmids were isolated (QIAprep spin miniprep kit), and sequenced. The plasmid with the desired sequence was subjected to digestion with restriction endonucleases *Sac*II and *Cla*I along with a sample of pETGlc4 in a separate reaction vessel. Linearized pETGlc4 and the 550 bp insert from pCR-Blunt cloning vector were isolated (QIAquick gel extraction kit) and ligated using T4 DNA ligase (Rapid DNA ligation kit; Fermentas). Chemically competent TOP10 *E. coli* cells were transformed and selected by ampicillin resistance. Single colonies were selected and grown overnight in LB_{amp} (Luria broth containing ampicillin) at 37 °C, plasmids were isolated (QIAprep spin miniprep kit), and sequenced.

The expression vector for the acid/base mutant (E411A) was prepared by a two-primer approach. A forward primer containing an *Rsr*II restriction site and the codon for Ala (5'-CGT CGG ACC GTT ACG AAT CGG ATG GCT ACG-3') and a reverse primer with an *Afl*II restriction site (5'-GCT CTT AAG CCG TGT CCG AAC GGA AAC AGC-3') were used in the PCR reaction. The 410 bp product of the reaction was isolated and the expression vector was prepared as described for the nucleophile mutant above.

Over-expression and purification of recombinant GCCase and its mutants

All proteins were expressed as described by Sumida et al. [6]. Cells were subjected to lysis by BugBuster protein-extraction reagent from Novagen according to the manufacturer's protocol. Proteinase inhibitors leupeptin hemisulphate (1 μ g/ml), pepstatin (1 μ g/ml) and chymostatin (5 μ g/ml) were added to prevent protein degradation. The supernatant obtained after centrifugation was loaded on to an FPLC column packed with His-Bind resin equilibrated with the following buffer: 5 mM imidazole, 20 mM Tris and 0.5 M NaCl, pH 7.9. Elution was performed by using a gradient of imidazole reaching a final concentration of 200 mM in the same buffer. Fractions were analysed by SDS/PAGE and Western blotting (anti-His-tag; Novagen), and those containing the desired protein were pooled. The solvent was changed to a 50 mM sodium phosphate buffer, pH 6.5, by repeated dilution and concentration using a centrifugal filter device (Centriprep YM-50; Amicon). Protein concentration was measured by the Bio-Rad protein assay based on the Bradford dye-binding procedure using BSA as a standard.

Enzyme kinetics

Kinetic studies were performed in 25 mM sodium phosphate buffer, pH 6.5, at 37 °C with different concentrations of PNPG (*p*-nitrophenyl β -D-glucopyranoside) or DNPG as substrates ranging from $0.14 \times K_m$ to $7 \times K_m$ where applicable. Reaction mixtures were preincubated in the cell holder and the reaction was initiated by addition of the enzyme. Substrate hydrolysis was followed by monitoring absorbance at 400 nm ($\epsilon = 10\,910\text{ M}^{-1} \cdot \text{cm}^{-1}$ for 2,4-dinitrophenolate and $\epsilon = 5350\text{ M}^{-1} \cdot \text{cm}^{-1}$ for *p*-nitrophenolate). Hydrolysis rates were corrected for spontaneous substrate hydrolysis. Assays were repeated three times. Parameters were determined by fitting initial velocity data to the Michaelis–Menten equation by GraFit 4.0.

Stability of GCCase

GCCase (3.5 μ g/ml) was incubated at 37 °C in 25 mM buffer solutions of various pH values {pH 5.5–6.0, sodium citrate; 6.5–8.0, sodium phosphate; 8.5–9.0, CAPSO [3-(cyclohexylamino)-2-hydroxy-1-propanesulphonic acid]}. Aliquots were removed at various time points and assayed for GCCase activity with 2 mM PNPG under the incubation conditions. Residual activity was plotted against time and data were fitted to a first-order rate equation using GraFit 4.0 to determine pseudo-first-order rate constants for the loss of enzyme activity.

Dependence of wild-type and E411A activity on pH

Michaelis–Menten parameters for the hydrolysis of PNPG by the wild-type enzyme were determined in the pH range of 6.0–8.0 as described in the Enzyme kinetics section. Buffers were the same as those used for the stability study. The molar absorption coefficient of *p*-nitrophenolate was determined at each pH value in the appropriate buffer. The plot of k_{cat} versus pH was analysed by fitting the data to a single titration curve using GraFit 4.0.

Values of k_{cat} for the enzymic hydrolysis of PNPG by the acid/base mutant (E411A) were determined at 37 °C in the pH range of 6.0–9.5 in the buffers used for the stability study and 25 mM CAPSO at pH 9.5. Concentrations of the enzyme and the substrate were 0.16 and 300 μ M, respectively.

Enzyme-inactivation kinetics

GCCase (5 μ g/ml) was incubated at 25 °C in the presence of various concentrations (0–5 mM) of 2FDNPG in 50 mM sodium phosphate buffer, pH 6.5, containing 0.1 % (w/v) Triton X-100. Aliquots were removed and assayed for GCCase activity with 1 mM DNPG in 25 mM phosphate buffer, pH 6.5, at 25 °C at appropriate time points. Residual activity was plotted against time for each inactivator concentration and data were fitted to a first-order rate equation using GraFit 4.0 to determine pseudo-first-order rate constants of inactivation ($k_{\text{i,obs}}$). Values of $k_{\text{i,obs}}$ were then plotted against inactivator concentration and fitted to eqn. (1) to determine K_i and k_i .

$$k_{\text{i,obs}} = \{k_i[I]/(K_i + [I])\} + k_d \quad (1)$$

where k_d was necessary to account for spontaneous, time-dependent denaturation of the enzyme.

Labelling of GCCase with 2FDNPG and ESI (electrospray ionization)-MS analysis of the proteolytic digest

GCCase (0.14 mg/ml in 50 mM sodium phosphate, pH 6.5) was incubated with 5 mM 2FDNPG for 15 min at 25 °C and then analysed by ESI-MS as described previously [10]. Samples of labelled and unlabelled GCCase were subjected to proteolysis by incubation with pepsin (0.3 mg/mg GCCase) in 100 mM sodium phosphate, pH 2.0, at 25 °C for 2 h. Samples were then frozen until MS analysis. Peptide separation and ESI-MS/MS analysis of the labelled peptide were carried out as described previously [10].

Rescue of the nucleophile mutant D223G and the general acid/base mutant E411A

D223G (0.6 μ M) was assayed for activity with 1 mM DNPG in 25 mM sodium phosphate, pH 6.5, at 37 °C in the presence of various concentrations of sodium formate (0–4 M). Wild-type GCCase activity was also measured under the same conditions.

Table 1 Kinetic parameters for substrate hydrolysis by *Paenibacillus* GCCase and its active-site mutants

Conditions were: 25 mM sodium phosphate, pH 6.5, at 37 °C. wt, wild type.

Enzyme	Substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1} \cdot s^{-1}$)	$K_m(wt)/K_m$	$k_{cat}(wt)/k_{cat}$	$[k_{cat}/K_m(wt)]/[k_{cat}/K_m]$
wt	DNPG	0.16 ± 0.02	470 ± 15	$(2.9 \pm 0.4) \times 10^6$	—	—	—
wt	PNPG	0.27 ± 0.02	560 ± 11	$(2.1 \pm 0.2) \times 10^6$	—	—	—
D223G	DNPG	0.15 ± 0.03	0.012 ± 0.004	79 ± 31	1.1	3.9×10^4	3.7×10^4
D223G	PNPG	0.33 ± 0.09	0.012 ± 0.002	36 ± 12	0.8	4.6×10^4	5.8×10^4
E411A	DNPG	≈ 0.0003	0.067 ± 0.002	$\approx 2 \times 10^5$	≈ 500	7.0×10^3	≈ 10
E411A	PNPG	0.018 ± 0.001	0.099 ± 0.002	$(5.5 \pm 0.3) \times 10^3$	15	5.6×10^3	380

E411A (0.16 μ M) was assayed for activity with 1 mM DNPG in 25 mM sodium phosphate, pH 6.5, at 37 °C in the presence of various concentrations of sodium azide (0–2 M). Wild-type GCCase activity was also measured under the same conditions. Products of the reaction (1.6 μ M E411A, 1 mM DNPG, 0.5 M sodium azide and 25 mM sodium phosphate, pH 6.5, 37 °C) were analysed by TLC and ESI-MS.

Kinetic parameters for the hydrolysis of DNPG by the mutant E411A (0.16 μ M) were determined in the presence of various concentrations of sodium azide (0–500 mM) in 25 mM sodium phosphate, pH 6.5, at 37 °C. Substrate concentration was in the range of $0.14 \times K_m$ to $7 \times K_m$.

RESULTS AND DISCUSSION

Kinetic studies and covalent labelling of the wild-type enzyme

Glucosylceramide, the natural substrate of GCCase, is not a convenient substrate for continuous assay and therefore for kinetic studies of the enzyme. However, PNPG [6] and DNPG are also readily hydrolysed by GCCase, yielding products that strongly absorb at 400 nm. Initial velocity data were plotted against substrate concentration and kinetic parameters were obtained by fitting the data to the Michaelis–Menten equation. Results are summarized in Table 1. The Michaelis constant for PNPG hydrolysis ($K_m = 270 \mu$ M) is in reasonable agreement with the value of 593 μ M reported earlier [6].

The stability of GCCase was examined at 37 °C in the pH range of 5.5–9.0. Incubation at pH 5.5 resulted in rapid loss of activity, whereas the enzyme was modestly stable at pH 6.0–8.0, with increasing rates of loss of activity under more basic conditions. Pseudo-first-order rate constants (k_{denat}) for the loss of GCCase activity as a function of pH are shown in Figure 2.

Kinetic parameters for PNPG hydrolysis by GCCase were determined at a series of pH values in the range where the enzyme proved to be sufficiently stable during the course of the assay (pH 6.0–8.0). Plotting k_{cat} (or k_{cat}/K_m) against pH, a bell-shaped profile is expected due to the participation of two ionizable groups in catalysis. Unfortunately, due to the irreversible denaturation at pH ≤ 5.5 , pK_{a1} (the acid ionization constant of the catalytic nucleophile) could not be determined. The acid ionization constant of the general acid/base catalytic residue (pK_{a2}) was estimated to be 7.8 by fitting the k_{cat} versus pH data points to a single titration curve (Figure 3a). However, as data are not available below pH 6.0, this value should rather be regarded as the upper limit of pK_{a2} for the enzyme–substrate complex. The intrinsic pK_a of the general acid/base catalytic residue could not be estimated in a similar fashion from the k_{cat}/K_m -versus-pH plot (Figure 3b), as this graph appears to be shifted towards the more acidic pH region where kinetic data could not be obtained due to the instability of the enzyme.

Figure 2 Stability of *Paenibacillus* GCCase as a function of pH

Plot of k_{denat} for the loss of activity versus pH; incubation in 25 mM buffer solutions at 37 °C as described in the Experimental section; assay with 2 mM DNPG under conditions identical to those of the incubation. Standard error was within 7 %, except for pH 5.5 (13 %).

2-Deoxy-2-fluoro- β -D-glycosides are powerful tools for the identification of the catalytic nucleophile of retaining β -glycosidases [17]. Substitution of the hydroxyl group by fluorine at the C-2 position causes a considerable decrease in the rate of both the glycosylation and deglycosylation steps (Scheme 1) through destabilization of the oxocarbenium-ion-like transition states. Due to the presence of a good leaving group (2,4-dinitrophenolate), however, the rate of the glycosylation step is less affected, leading to accumulation of the covalent glycosyl–enzyme intermediate.

GCCase was incubated in the presence of various concentrations of 2FDNPG and aliquots were assayed for activity with DNPG at appropriate time points. Pseudo-first-order rate constants ($k_{i,obs}$) for the loss of enzyme activity at each inactivator concentration were determined by fitting the data to a first-order rate equation (Figure 4a). Values of the obtained rate constants were then plotted against the inactivator concentration and the data fitted to eqn (1), giving $K_i = 2.4 \pm 0.7$ mM and $k_i = 0.59 \pm 0.05$ min^{-1} (Figure 4b). A pseudo-first-order rate constant for loss of activity greater than 0 was detected without added inactivator ($k_d = 0.01$ min^{-1}) due to denaturation of the enzyme during the time of inactivation.

ESI mass spectra of intact GCCase and that labelled with 2FDNPG were compared. Molecular mass values were determined to be 92 018 Da (the calculated value for GCCase expressed with a C-terminal His₆ tag is 92 098 Da) and 92 188 Da, respectively. The observed mass difference of 170 Da indicated the incorporation of a single 2-deoxy-2-fluoroglucosyl moiety (theoretical value, 164 Da) into GCCase upon treatment with 2FDNPG.

Unlabelled and labelled GCCase were subjected to proteolytic digestion with pepsin. The resulting peptides were separated

Figure 3 Dependence of PNPG hydrolysis by *Paenibacillus* GCCase on pH

(a) k_{cat} versus pH plot (●, wild-type; ○, 1000 × k_{cat} , E411A; standard error was within 5 %). (b) k_{cat}/K_m versus pH plot (standard error was within 10 %). Conditions were 25 mM buffer, 37 °C.

on a reversed-phase HPLC column and were detected by ESI-MS as described previously [10]. The two chromatograms were compared to find the labelled peptide. A singly charged peptide of 1140.8 Da was found in the chromatogram of the labelled digest (Figure 5a) that was not present in the unlabelled digest (Figure 5b). Subtraction of the mass of the label (164 Da) indicates an unlabelled peptide of 976.8 Da, which was indeed present in the chromatogram of the unlabelled digest. The labelled peptide (1140.8 Da peak) was then isolated by reversed-phase HPLC and subjected to further analysis by ESI-MS/MS as described in previous studies [10]. The mass spectrum of the fragments obtained by collision-induced dissociation of the parent ion of m/z 1140.4 (Figure 6a) contained a series of b -ions (N-terminal fragments produced upon cleavage of the CO-NH bond) corresponding to the unlabelled fragments VSD (m/z 302.1), VSDW (m/z 488.2), VSDWGA (m/z 616.2), VSDWGAV (m/z 715.3), VSDWGAVN (m/z 829.3) and VSDWGAVNE (m/z 958.4). b -ions carrying a label of 164 Da (2-deoxy-2-fluoroglucosyl moiety*) detected were as follows: VSD* (m/z 466.2), VSD*W (m/z 652.2), VSD*WG (m/z 709.2), VSD*WGA (m/z 780.3), VSD*WGAV (m/z 879.3), VSD*WGAVN (m/z 993.4) and VSD*WGAVNE (m/z 1122.4). The spectrum also contained a series of y -ions (C-terminal fragments produced upon cleavage of the CO-NH bond) corresponding to the unlabelled fragments NE (m/z 262.1), VNE (m/z 361.2), AVNE (m/z 432.2), GAVNE (m/z 489.2), WGAVNE (m/z 675.3), DWGAVNE (m/z 790.3) and VSDWGAVNE (m/z 976.4). y -ions carrying a label of 164 Da detected were: D*WGAVNE (m/z

Figure 4 Inactivation of *Paenibacillus* GCCase with 2FDNPG

(a) V/V_0 (where V_0 is V at time 0) over time in the presence of various concentrations of 2FDNPG: ○, 5 mM; △, 3 mM; ▽, 2 mM; ◇, 1 mM; star, 0.32 mM; ●, 0.176 mM. (b) Plot of $k_{i,obs}$ against [2FDNPG]. Conditions were 50 mM sodium phosphate, pH 6.5, and 0.1 % (w/v) Triton X-100, at 25 °C. Standard error was within 16 %.

954.3) and VSD*WGAVNE (m/z 1140.4). These results show unequivocally that the sequence of the labelled peptide of m/z 1140.4 is VSDWGAVNE, corresponding to amino acids 221–229 of the *Paenibacillus* GCCase, and that the residue labelled with the 2-deoxy-2-fluoroglucosyl moiety is Asp-223 (Figure 6b). Our results therefore confirm that this is indeed the catalytic nucleophile of GCCase, as suspected from sequence alignments and previous studies [6].

Kinetic studies of the nucleophile mutant D223G

Kinetic parameters for the hydrolysis of DNPG and PNPG by the nucleophile mutant D223G were determined (Table 1). A decrease of nearly 5 orders of magnitude in the apparent first-order rate constant (k_{cat}) and the catalytic efficiency (k_{cat}/K_m) was observed for the cleavage of both substrates, confirming the crucial role of Asp-223 in catalysis [6]. Very similar trends were observed for DNPG hydrolysis by the D247G (nucleophile) mutant of the family 3 β -glucosidase from *F. meningosepticum* [11]. Furthermore, at high substrate concentrations, inhibition of the D223G mutant of GCCase by the substrate PNPG was observed ($K_i = 1.7 \pm 0.5$ mM), but not with the wild-type enzyme. The origin of this behaviour is not clear, but it strongly suggests that the activity detected for the nucleophile mutant was not likely to be caused by contamination with wild-type enzyme.

Figure 5 Detection of the 2-deoxy-2-fluoroglucosyl-labelled peptide by ESI-MS

(a) ESI mass spectrum of the labelled digest eluted at 25 min. (b) ESI mass spectrum of the unlabelled digest eluted at 24 min.

It has been demonstrated that the activity of a mutant glycosidase missing the essential catalytic nucleophile can be partially restored by the addition of an external nucleophile that is able to fit into the active site in the presence of bound substrate [18,19]. Indeed, addition of formate, an excellent mimic of the missing catalytic residue, substantially restored activity. A 4 M concentration of this exogenous nucleophile resulted in a 100-fold rate enhancement for the hydrolysis of 1 mM DNPG (Figure 7) by the mutant D223G. By contrast, wild-type GCase was inhibited by sodium formate in a concentration-dependent fashion as seen in the inset to Figure 7.

Results of the kinetic analysis of the mutant D223G and the formate-rescue experiment are fully consistent with Asp-223 being the catalytic nucleophile of *Paenibacillus* GCase.

Kinetic studies of the general acid/base mutant E411A

Michaelis–Menten parameters for DNPG and PNPG hydrolysis by the general acid/base mutant E411A were determined (Table 1). A reduction of nearly 4 orders of magnitude in the values of k_{cat} was observed relative to the wild type, indicating the important role of Glu-411 in catalysis. This rate reduction

was significantly smaller than that observed for the nucleophile mutant D223G, consistent with previous observations on other glycosidases [19]. Michaelis constants (K_m) were greatly reduced for E411A in comparison to the wild-type enzyme, presumably as a result of significant accumulation of the covalent glycosyl–enzyme intermediate. This effect was considerably larger (nearly 3 orders of magnitude) for DNPG, whereas the K_m for PNPG only decreased by a factor of 15. As a consequence, the K_m value in this case was extremely difficult to measure accurately, and the value of 0.3 μM is simply an estimate. This difference in the relative magnitudes of K_m for the two substrates can easily be explained by the difference in the leaving group abilities of *p*-nitrophenolate and 2,4-dinitrophenolate. This latter, by being a far better leaving group ($\text{p}K_a = 3.96$), requires significantly less assistance from a general acid catalyst in the first (glycosylation) step than does the former ($\text{p}K_a = 7.18$). As a consequence, the rate of the glycosylation step for DNPG hydrolysis is less affected by the removal of the general acid/base catalytic residue (while the deglycosylation step is slowed equally for both substrates), causing greater accumulation of the covalent intermediate. Apparent second-order rate constants (k_{cat}/K_m) for substrate hydrolysis by E411A were much less affected than the apparent first-order rate constants (k_{cat}), indicating less

Figure 6 ESI-MS/MS analysis of the 2-deoxy-2-fluoroglucosyl-labelled peptide

(a) MS/MS daughter-ion spectrum of the labelled peptide. (b) Rationalization of the observed singly charged *y*- and *b*-ions, defined as above. Masses shown in parentheses are those of expected peptides that were not found in the mass spectra.

Figure 7 Effect of formate on DNPG hydrolysis catalysed by *Paenibacillus* GCase

Conditions were 25 mM sodium phosphate and 1 mM DNPG, pH 6.5, at 37 °C. ●, Wild type; ○, D223G. V_0 is the activity of the respective enzyme in the absence of sodium formate. The inset shows the effect of formate on wild-type activity on an expanded scale.

impairment of the glycosylation step (reflected in k_{cat}/K_m) than the deglycosylation step (reflected in k_{cat}). These changes in the kinetic parameters for substrate hydrolysis by GCase upon

mutation of Glu-411 to Ala are those expected as a result of removing the general acid/base catalyst and are comparable with those observed for the E374G (general acid/base) mutant of the family 3 β -glucosidase from *F. meningosepticum* [11]. A second indication that Glu-411 is indeed the acid/base catalyst comes from measurement of the pH-dependence of catalysis. Elimination of Glu-411 of GCase resulted in essentially pH-independent catalysis of PNPG hydrolysis in the basic limb of the pH/rate profile (Figure 3a), much as would be expected if the catalytic acid/base residue was removed.

The activity of a mutant glycoside hydrolase missing the general acid/base catalytic residue can also be restored by the addition of an external nucleophile that is able to react with the covalent glycosyl-enzyme intermediate without the need for general base catalysis [20]. The effect is more pronounced for substrates with a relatively good leaving group (such as 2,4-dinitrophenolate) requiring less assistance from a general acid catalyst for leaving-group departure in the glycosylation step. Addition of 0.5–1 M azide resulted in a 40-fold increase in catalytic activity of the mutant E411A when assayed with 1 mM DNPG (Figure 8). By contrast, wild-type enzyme was slightly activated when < 0.5 M sodium azide and inhibited when > 0.5 M sodium azide was added to the assay mixture. Kinetic parameters for DNPG hydrolysis by E411A were determined in the presence of various concentrations of sodium azide (Figure 9). The apparent first-order rate constant (k_{cat}) and the Michaelis constant (K_m) were significantly increased upon addition of sodium azide to

Figure 8 Effect of azide on DNPG hydrolysis catalysed by *Paenibacillus* GCase

Conditions were 25 mM sodium phosphate and 1 mM DNPG, pH 6.5, at 37 °C. ●, Wild type; ○, E411A. V_0 is the activity of the respective enzyme in the absence of sodium azide. The inset shows the effect of azide on wild-type activity on an expanded scale.

Figure 9 Effect of azide on the kinetic parameters of DNPG hydrolysis catalysed by *Paenibacillus* GCase mutant E411A

△, k_{cat} ; ○, K_m ; ▽, k_{cat}/K_m . Constant(o) refers to the respective constant measured in the absence of sodium azide. Conditions were 25 mM sodium phosphate, pH 6.5, at 37 °C.

the assay mixture. By contrast, the apparent second-order rate constant (k_{cat}/K_m) was affected only to a very small extent; it was essentially unchanged in the concentration range of 100–500 mM sodium azide. These results indicate that the increase in enzymic activity of E411A upon addition of an external nucleophile is the consequence of a significant enhancement of the rate of the deglycosylation step (reflected in k_{cat}) with no substantial effect on the glycosylation step (reflected in k_{cat}/K_m). These findings are fully consistent with earlier results obtained for the β -glucosidase from *F. meningosepticum* [9].

TLC analysis of the reaction mixture containing the mutant E411A, DNPG and sodium azide revealed the formation of an additional product besides 2,4-dinitrophenolate and glucose, which was identified by ESI-MS as glucosyl azide [measured, $m/z = 227.1$; calculated for $\text{GlcN}_3 - \text{H} + \text{Na}^+$ $m/z = 227.2$ (where GlcN_3 is glucosyl azide)]. This product was not observed when

wild-type enzyme was used, nor was it seen in the absence of any enzyme. Since glucosyl azide was not detected in the presence of the wild-type enzyme it can be deduced that its formation in the reaction catalysed by the mutant E411A could only be a result of the elimination of the general acid/base catalyst, allowing turnover of the covalent intermediate by the nucleophilic azide ion.

Conclusions

Catalytic residues of GCCase from *Paenibacillus* sp. TS12 were identified by covalent labelling and detailed kinetic analysis of both the wild-type enzyme and site-directed mutants. The catalytic nucleophile was confirmed to be Asp-223, underlining the fully conserved nature of this residue across subfamilies of glycoside hydrolase family 3. By identifying Glu-411 as the general acid/base catalyst of the *Paenibacillus* GCCase, the general acid/base residue of a further subfamily within the glycoside hydrolase family 3 was determined. Glu-411 aligns with the catalytic acid/base residue of the *F. meningosepticum* β -glucosidase, but not with that of the barley α -1,3-1,4-glucanase. An additional conserved active-site carboxylate in the sequence of the subfamily 5 enzymes (Asp-414 of the *Paenibacillus* GCCase; Figure 1) may play an important role in tuning the pK_a values of one or both of the catalytic residues or perhaps in substrate binding.

We thank Shouming He for obtaining and analysing the mass spectra and Emily Kwan for technical assistance. We also thank PENCE, the Protein Engineering Network of Centres of Excellence of Canada, and Neose Technologies for financial support.

REFERENCES

- Ito, M. and Yamagata, T. (1989) Purification and characterization of glycosphingolipid-specific endoglycosidases (endoglycoceramidases) from a mutant strain of *Rhodococcus* sp. Evidence for three molecular species of endoglycoceramidase with different specificities. *J. Biol. Chem.* **264**, 9510–9519
- Roebber, D., Achari, A., Manavalan, P., Edmonds, T. and Scott, D. L. (2003) Crystallization and preliminary X-ray analysis of recombinant human acid β -glucocerebrosidase, a treatment for Gaucher's disease. *Acta Crystallogr.* **59**, 343–344
- Grace, M. E., Newman, K. M., Scheinker, V., Berg-Fussman, A. and Grabowski, G. A. (1994) Analysis of human acid β -glucosidase by site-directed mutagenesis and heterologous expression. *J. Biol. Chem.* **269**, 2283–2291
- 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
- Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280**, 309–316
- Sumida, T., Sueyoshi, N. and Ito, M. (2002) Molecular cloning and characterization of a novel glucocerebrosidase of *Paenibacillus* sp. TS12. *J. Biochem.* **132**, 237–243
- Harvey, A. J., Hrmova, M., De Gori, R., Varghese, J. N. and Fincher, G. B. (2000) Comparative modeling of the three-dimensional structures of family 3 glycoside hydrolases. *Proteins* **41**, 257–269
- Varghese, J. N., Hrmova, M. and Fincher, G. B. (1999) Three-dimensional structure of a barley β -D-glucan exohydrolase, a family 3 glycosyl hydrolase. *Structure* **7**, 179–190
- Li, Y.-K., Chir, J., Tanaka, S. and Chen, F.-Y. (2002) Identification of the general acid/base catalyst of a family 3 β -glucosidase from *Flavobacterium meningosepticum*. *Biochemistry* **41**, 2751–2759
- Chir, J., Withers, S. G., Wan, C.-F. and Li, Y.-K. (2002) Identification of the two essential groups in the family 3 β -glucosidase from *Flavobacterium meningosepticum* by labelling and tandem mass spectrometric analysis. *Biochem. J.* **365**, 857–863
- Li, Y.-K., Chir, J. and Chen, F.-Y. (2001) Catalytic mechanism of a family 3 β -glucosidase and mutagenesis study on residue Asp-247. *Biochem. J.* **355**, 835–840
- Roeser, K. R. and Legler, G. (1981) Role of sugar hydroxyl groups in glycoside hydrolysis. Cleavage mechanism of deoxyglucosides and related substrates by β -glucosidase A3 from *Aspergillus wentii*. *Biochim. Biophys. Acta* **657**, 321–333
- Zechele, D. and Withers, S. G. (2000) Glycosidase mechanisms: anatomy of a finely tuned catalyst. *Acc. Chem. Res.* **33**, 11–18

- 14 Dan, S., Marton, I., Dekel, M., Bravdo, B.-A., He, S., Withers, S. G. and Shoseyov, O. (2000) Cloning, expression, characterization, and nucleophile identification of family 3, *Aspergillus niger* β -glucosidase. *J. Biol. Chem.* **275**, 4973–4980
- 15 Hrmova, M., Varghese, J. N., De Gori, R., Smith, B. J., Driguez, H. and Fincher, G. B. (2001) Catalytic mechanism and reaction intermediates along the hydrolytic pathway of a plant β -D-glucan glucohydrolase. *Structure* **9**, 1005–1016
- 16 Vocadlo, D. J., Mayer, C., He, S. and Withers, S. G. (2000) Mechanism of action and identification of Asp242 as the catalytic nucleophile of *Vibrio furnisii* *N*-acetyl- β -D-glucosaminidase using 2-acetamido-2-deoxy-5-fluoro- α -L-idopyranosyl fluoride. *Biochemistry* **39**, 117–126
- 17 Wicki, J., Rose, D. R. and Withers, S. G. (2002) Trapping covalent intermediates on β -glycosidases. *Methods Enzymol.* **354**, 84–105
- 18 MacLeod, A. M., Tull, D., Rupitz, K., Warren, R. A. and Withers, S. G. (1996) Mechanistic consequences of mutation of active site carboxylates in a retaining β -1,4-glycanase from *Cellulomonas fimi*. *Biochemistry* **35**, 13165–13172
- 19 Ly, H. D. and Withers, S. G. (1999) Mutagenesis of glycosidases. *Annu. Rev. Biochem.* **68**, 487–522
- 20 MacLeod, A. M., Lindhorst, T., Withers, S. G. and Warren, R. A. (1994) The acid/base catalyst in the exoglucanase/xylanase from *Cellulomonas fimi* is glutamic acid 127: evidence from detailed kinetic studies of mutants. *Biochemistry* **33**, 6371–6376

Received 9 July 2003/2 October 2003; accepted 16 October 2003

Published as BJ Immediate Publication 16 October 2003, DOI 10.1042/BJ20031028