Biochimica et Biophysica Acta 1830 (2013) 2739–2749

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Identification of the acid/base catalyst of a glycoside hydrolase family 3 (GH3) β-glucosidase from *Aspergillus niger* ASKU28

Preeyanuch Thongpoo [a,b], Lauren S. McKee [b,c,1], Ana Catarina Araújo [b,1], Prachumporn T. Kongsaeree [a,d], Harry Brumer [b,c,e,*]

[a] *Interdisciplinary Graduate Program in Genetic Engineering, Faculty of Graduate School, Kasetsart University, Bangkok 10900, Thailand*
[b] *Division of Glycoscience, School of Biotechnology, Royal Institute of Technology (KTH), AlbaNova University Centre, 106 91 Stockholm, Sweden*
[c] *Wallenberg Wood Science Center, Royal Institute of Technology (KTH), Teknikringen 56-58, 100 44 Stockholm, Sweden*
[d] *Department of Biochemistry, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand*
[e] *Michael Smith Laboratories and Department of Chemistry, University of British Columbia, 2185 East Mall, Vancouver, BC, Canada V6T 1Z4*

ARTICLE INFO

Article history:
Received 10 July 2012
Received in revised form 7 November 2012
Accepted 19 November 2012
Available online 29 November 2012

Keywords:
β-Glucosidase
Acid/base catalyst
Aspergillus niger
Glycoside hydrolase family 3
Azide rescue

ABSTRACT

Background: The commercially important glycoside hydrolase family 3 (GH3) β-glucosidases from *Aspergillus niger* are anomeric-configuration-retaining enzymes that operate through the canonical double-displacement glycosidase mechanism. Whereas the catalytic nucleophile is readily identified across all GH3 members by sequence alignments, the acid/base catalyst in this family is phylogenetically variable and less readily divined.
Methods: In this report, we employed three-dimensional structure homology modeling and detailed kinetic analysis of site-directed mutants to identify the catalytic acid/base of a GH3 β-glucosidase from *A. niger* ASKU28.
Results: In comparison to the wild-type enzyme and other mutants, the E490A variant exhibited greatly reduced k_{cat} and k_{cat}/K_m values toward the natural substrate cellobiose (67,000- and 61,000-fold, respectively). Correspondingly smaller kinetic effects were observed for artificial chromogenic substrates *p*-nitrophenyl β-D-glucoside and 2,4-dinitrophenyl β-D-glucoside, the aglycone leaving groups of which are less dependent on acid catalysis, although changes in the rate-determining catalytic step were revealed for both. pH-rate profile analyses also implicated E490 as the general acid/base catalyst. Addition of azide as an exogenous nucleophile partially rescued the activity of the E490A variant with the aryl β-glucosides and yielded β-glucosyl azide as a product.
Conclusions and general significance: These results strongly support the assignment of E490 as the acid/base catalyst in a β-glucosidase from *A. niger* ASKU28, and provide crucial experimental support for the bioinformatic identification of the homologous residue in a range of related GH3 subfamily members.

1. Introduction

β-Glucosidases (EC 3.2.1.21) catalyze the hydrolysis of the glucosidic linkage of aryl- and alkyl-β-glucosides, and liberate terminal non-reducing β-glucosyl residues from oligosaccharides. Based on amino acid sequence and structural similarity, known β-glucosidases have been placed in Glycoside Hydrolase (GH) Families GH1, GH3, GH5, GH9, GH30, and GH116 [1–3]. Members of GH9 utilize a one-step, anomeric configuration-inverting mechanism, whereas the remainder operate through a double-displacement, configuration-retaining mechanism [1].

In the canonical "retaining" mechanism employed by GH3 β-glucosidases [4–6], the carboxylic acid side chains of two active-site residues function individually as a catalytic nucleophile and a general acid/base residue [7]. The first step of the mechanism involves the displacement of the leaving group (aglycone or saccharide) by the active-site nucleophile, assisted by proton donation to the nucleofuge from the conjugate acid form of the general acid/base residue. The resulting covalent glycosyl-enzyme intermediate is decomposed in the second step by an incoming water molecule, which is activated through proton abstraction by the basic form of the general acid/base residue. Alternatively, interception of the glycosyl-enzyme by an incoming saccharide can lead to transglycosylation in some enzymes [8].

The catalytic residues are, in general, highly conserved both within a particular GH Family and, more broadly, among a group of families comprising a GH Clan. As such, identification of a particular catalytic residue in one representative of a family allows facile prediction of the homologous

Abbreviations: pNP-Glc, *p*-nitrophenyl β-D-glucoside; 2,4-DNP-Glc, 2,4-dinitrophenyl β-D-glucoside; beta-Glc-N3, beta-D-glucosyl azide; GH, Glycoside Hydrolase; CAZy, carbohydrate-active enzymes from *Bacillus subtilis*; NagZ, β-*N*-acetylglucosaminidase; ExoP, exo-1,3/1,4-β-glucanase from *Pseudoalteromonas* sp.; Bgl3B, β-glucosidase B from *Thermotoga neapolitana*; ExoI, β-D-glucan exohydrolase isoenzyme from *Hordeum vulgare*; KmBgII, β-glucosidase I from *Kluyveromyces marxianus*

* Corresponding author at: Michael Smith Laboratories and Department of Chemistry, University of British Columbia, 2185 East Mall, Vancouver, BC, Canada V6T 1Z4.
E-mail addresses: g5089003@ku.ac.th (P. Thongpoo), mckee@kth.se (L.S. McKee), acaraujo@kth.se (A.C. Araújo), fscippt@ku.ac.th (P.T. Kongsaeree), harry@biotech.kth.se, brumer@msl.ubc.ca (H. Brumer).

[1] The equal contribution of A.C.A. and L.S.M. is highlighted.

0304-4165/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbagen.2012.11.014

residue in (nearly) all members [9]. Indeed, this is the case for the catalytic nucleophile of GH3, which has been conclusively identified as a conserved Asp residue via mechanism-based active-site labeling of a number of members [6,10–12]. The catalytic acid–base residue, on the other hand, has escaped general definition, as it appears to be presented by different protein motifs in various GH3 members from bacteria, fungi, and plants [12–18]. In its most divergent incarnation, a catalytic His–Asp dyad, in which the His side chain directly mediates catalysis, replaces the canonical carboxylate in some GH3 β-*N*-acetylglucosaminidases [13,19].

Aspergillus species, and *A. niger* in particular, are commercially important fungal sources of a wide range of carbohydrate-active enzymes, including β-glucosidases of relevance to biomass saccharification and other applications [20–22]. Homologous GH3 β-glycosidases from *Aspergillus* spp. have been biochemically scrutinized to various degrees, including the detailed analysis of transglycosylation potential [23]. In the current absence of a three-dimensional structural representative, some catalytically important active site residues in *Aspergillus* GH3 β-glucosidase have been pinpointed by seminal studies on the *A. niger* homologue: Residues W49 and W262 modulate the ratio of substrate hydrolysis to transglycosylation [24,25], while the neighboring D261 has been identified as the catalytic nucleophile by mechanism-based labeling with 2-deoxy-2-fluoro β-glucosyl fluoride and peptide mass spectrometry [11]. However, definitive biochemical and structural information on the nature of the general acid/base catalyst in GH3 β-glucosidases from *Aspergillus* spp. and other related fungi are currently lacking.

To address this need, we utilized site-directed mutagenesis guided by protein structure homology modeling, in harness with biochemical analysis, to identify glutamic acid 490 as the general acid/base residue in the GH3 β-glucosidase from *A. niger* strain ASKU28.

2. Materials and methods

2.1. Materials

Escherichia coli strains Top10 and DH5α were used for plasmid propagation. *Pichia pastoris* strain Y11430 (Invitrogen, Carlsbad, CA, USA) was used for protein expression. *A. niger* ASKU28 was from the existing stock collection at the Department of Microbiology, Faculty of Science, Kasetsart University, Thailand. The substrates used in this study were cellobiose, which was purchased from Fluka (St. Louis, MO, USA), *p*-nitrophenyl-β-D-glucopyranoside (*p*NP-Glc), which was purchased from Sigma-Aldrich (St. Louis, MO, USA), and 2,4-dinitrophenyl-β-D-glucopyranoside (2,4-DNP-Glc), which was kindly provided by Professor Stephen G. Withers, University of British Columbia, Vancouver, Canada. All other chemicals used in this study were of analytical grade.

2.2. Cloning of β-glucosidase from A. niger ASKU28

A. niger ASKU28 was grown in minimal medium [26] supplemented with 2% beechwood xylan for 2 days at 30 °C, 100 rpm. Total RNA was extracted from the fungal mycelia with an RNeasy Mini Kit (QIAGEN, Hilden, Germany). cDNA was produced according to the protocol for first-strand cDNA synthesis (Fermentas, Glen Burnie, MD, USA) by using oligo $(dT)_{18}$ as a primer. The coding sequence of mature β-glucosidase from *A. niger* ASKU28 (from position 19 of the amino acid sequence to the stop codon) was amplified from the first strand cDNA with 3 units of *Pfu* Ultra HF DNA polymerase (Promega, Madison, WI, USA) and specific primers, ASKU28_for and ASKU28_rev (Table 1), that were designed based on the nucleotide sequence of exons 2–7 of *A. niger bgl*1 gene (GenBank ID: AJ132386.1) [11]. The PCR products were ligated into pZErO™-2 cloning vector (Invitrogen, Carlsbad, CA, USA) at the *Eco*RV site, and transformed into *E. coli* strain Top10 by electroporation. The positive transformants were selected by plating onto LB medium containing 25 μg/ml kanamycin at 37 °C. The coding sequence of *A. niger* β-glucosidase was checked by DNA sequencing.

2.3. Protein expression and purification of wild-type β-glucosidase from A. niger ASKU28

The coding sequence of mature β-glucosidase from *A. niger* ASKU28 was subcloned into a pPICZαBNH8 expression vector [27] at the *Eco*RI-*Xba*I sites, and transformed into *E. coli* strain DH5α by electroporation. The positive transformants were selected by plating onto LB medium containing 25 μg/ml zeocin at 37 °C. After that, pPICZαBNH8-ASKU28 plasmid was extracted and linearized by *Pme*I before transforming into *P. pastoris* strain Y11430 by electroporation. The transformed clones were grown on YPDS agar containing 100 μg/ml zeocin for 3–5 days at 30 °C.

Fifty positive clones of *P. pastoris* harboring pPICZαBNH8-ASKU28 were selected for small scale expression. The transformants were grown in 5 ml BMGH medium overnight at 30 °C, 200 rpm, and expressed in 5 ml BMMH medium. Methanol was added every day to the final concentration of 0.5% (v/v) for 7 days. The level of β-glucosidase expression was checked by activity assay in the reaction containing 1 mM *p*NP-Glc in 0.1 M sodium acetate, pH 5.0, at 50 °C for 30 min. The *p*-nitrophenol released was determined by reading absorbance at 400 nm after stopping the assays by addition of 0.2 M Na_2CO_3. The *P. pastoris* clone that showed the highest activity was selected for large-scale expression in 1 L medium under the same culture conditions as above. Methanol was added every day to the final concentration of 0.5–0.8% (v/v) until the activity did not increase any further.

The culture supernatant from large-scale expression was added with 2 M ammonium sulfate before loading onto a hydrophobic interaction chromatography column containing phenyl sepharose beads (GE Healthcare, Sweden), that was previously equilibrated with 0.1 M sodium acetate, pH 5.0, containing 2 M ammonium sulfate. The recombinant β-glucosidase was eluted from the column with a linear gradient of 2–0 M ammonium sulfate in 0.1 M sodium acetate, pH 5.0. Fractions containing β-glucosidase activities were pooled and concentrated, and ammonium sulfate was removed by centrifugal ultrafiltration. The protein content of recombinant β-glucosidase was determined at absorbance 595 nm by using the Bio-Rad protein assay reagent kit (Bio-Rad, Hercules, CA, USA) and bovine serum albumin as the standard. 10% SDS–PAGE was performed to check the purity of the purified enzyme.

2.4. Sequence alignment and molecular modeling

To select candidates for site-directed mutagenesis, the amino acid sequence of the β-glucosidase from *A. niger* ASKU28 was aligned with those of the 2 reported *A. niger* β-glucosidases (GenBank ID: CAB75696 and GenBank ID: CAK48740), 14 enzymes from subfamily 4 of GH3 [28], β-glucosidase B (Bgl3B) from *Thermotoga neapolitana* [14], β-D-glucan exohydrolase isoenzyme (ExoI) from *Hordeum vulgare* (barley) [15], β-glucosidase I (KmBglI) from *Kluyveromyces marxianus* [16],

Table 1
Oligonucleotide primers used in this study. The mismatch sequences in the mutagenic primers are underlined.

Oligonucleotide primer	Nucleotide sequence (5′ to 3′)
ASKU28_for	TAG AAT TCT TGA TGA ATT GGC CTA CTC CCC
ASKU28_rev	CGT ATC TAG ATT AGT GAA CAG TAG GCA GAG A
D463A_f	GGT GTA TTC ACC GCC ACC GCT AAC TGG GCT ATC GAT C
D463A_r	GAT CGA TAG CCC AGT TAG CGG TGG CGG TGA ATA CAC C
D487A_f	CTT TGT CAA CGC CGC TTC TGG TGA GGG TTA C
D487A_r	GTA ACC CTC ACC AGA AGC GGC GTT GAC AAA G
E490A_f	GTC AAC GCC GAC TCT GGT GCT GGT TAC ATC AAT GTG GAC
E490A_r	GTC CAC ATT GAT GTA ACC AGC ACC AGA GTC GGC GTT GAC
D496A_f	GAG GGT TAC ATC AAT GTG GCT GGA AAC CTG GGT GAC C
D496A_r	GGT CAC CCA GGT TTC CAG CCA CAT TGA TGT AAC CCT C
D501A_f	GTG GAC GGA AAC CTG GGT GCT CGC AGG AAC CTG ACC CTG
D501A_r	CAG GGT CAG GTT CCT GCG AGC ACC CAG GTT TCC GTC CAC

β-glucosidase from *Flavobacterium meningosepticum* [17] and glucosylceramidase from *Paenibacillus* sp. TS12 [12] by using Clustal Omega (1.1.0) [29].

In parallel, the amino acid sequence of β-glucosidase from ASKU28 was submitted to the Phyre^2 server [30] to predict the fold of the enzyme. The final model from the server was then subjected to the Chiron server [31] to evaluate and resolve steric clashes in the model before further analysis. The stereochemical qualities of the energy-minimized model were assessed by the Procheck v3.5 [32] and Verify-3D [33] by submitting the atomic coordinates to the Swiss Model server (swissmodel.expasy.org) and the NIH MBI laboratory server (nihserver.mbi.ucla.edu), respectively. This three-dimensional model was aligned with the crystal structures of ExoI from barley in complex with glucose (PDB ID: 1ex1) [15] by using the "align" command in the PyMOL software program (Schrödinger LLC, Portland, OR, USA).

2.5. Site-directed mutagenesis and production of enzyme variants

The selected acidic residues, namely D463, D487, E490, D496, and D501, were replaced individually with alanine by site-directed mutagenesis by using 2 units of Phusion™ high-fidelity DNA polymerase (Finnzymes, Finland) according to the method published previously [34]. The residue numbers indicate their positions in the mature sequence of β-glucosidase from *A. niger* as reported previously [11]. The sequences of sense/antisense mutagenic primer pairs are listed in Table 1. The PCR products were treated with 10 units *Dpn*I at 37 °C overnight and transformed into the competent *E. coli* Top10F′ by heat shock method. The positive transformants were selected by plating onto LB agar plates containing 25 μg/mL zeocin at 37 °C overnight. Plasmids containing the correct mutation were identified by restriction digestion and DNA sequencing. Then, the mutant plasmids were linearized with *Pme*I, and transformed into *P. pastoris* Y11430 by electroporation. The positive transformed clones were selected on YPDS agar containing 100 μg/mL zeocin after incubation for 3–5 days at 30 °C. The positive clones from each mutation were selected for small scale expression as described in Section 2.3. The *P. pastoris* clones that showed the highest activities or the highest protein production from each mutation were selected for large-scale expression, and the enzyme variants were purified from the culture media of *P. pastoris* by the same method as described in Section 2.3.

2.6. Kinetic analysis

Kinetic parameters of the mutant forms of β-glucosidase from *A. niger* ASKU28 were determined toward cellobiose, *p*NP-Glc and 2,4-DNP-Glc, and compared to those of wild-type β-glucosidase. All kinetic measurements were performed in duplicate. The kinetic parameters were calculated by non-linear curve-fitting of the Michaelis–Menten equation with Origin, version 8 (OriginLab, Northampton, MA, USA).

The kinetic parameters of the purified enzymes toward cellobiose were determined in 0.1 M sodium citrate, pH 4.0, at 40 °C for 10 min (30 min for the E490A variant). The reactions were stopped by heating to 95 °C for 5 min, and the released glucose from the reaction was reacted with a glucose oxidase/peroxidase reagent (Sigma, St. Louis, MO, USA) for 15 min at 37 °C. The amount of glucose was determined by reading the absorbance at 400 nm, and comparing it to a standard curve of glucose.

The kinetic parameters of the purified enzymes toward *p*NP-Glc were determined in 0.1 M sodium citrate, pH 4.0, at 40 °C for 10 min to generate the data in Table 2. The pH dependence of the rate of enzyme-catalyzed *p*NP-Glc hydrolysis was determined in either 50 mM sodium citrate (pH 2.0–6.4) or 50 mM potassium phosphate (pH 6.0–8.0), at 40 °C for 10 min. In all cases, reactions were stopped by the addition of 0.2 M sodium carbonate, and the amount of *p*NP released was calculated from A_{400} values using an extinction coefficient of 18,300 M^{-1} cm^{-1} [35]. Measured extinction coefficients from standard samples of *p*-nitrophenol in all buffer systems after addition of the carbonate stop-solution were within experimental error of this published value.

The kinetic parameters of the recombinant wild-type β-glucosidase and the E490A variant toward 2,4-DNP-Glc were determined in 0.1 M sodium citrate, pH 3–6.4, and 0.1 M potassium phosphate, pH 6.0–8.0, at 40 °C by a continuous assay wherein the change in UV absorbance at 400 nm due to the release of 2,4-DNP was monitored spectrophotometrically (Cary 300 spectrophotometer, Varian, Australia). The rates of 2,4-DNP-Glc hydrolysis were calculated from the initial slopes of the change in A_{400} versus time (obtained by linear least-squares fitting) by using extinction coefficients determined for 2,4-dinitrophenol standards at each pH value.

2.7. Azide rescue

The kinetic parameters of the recombinant wild-type β-glucosidase and the E490A variant toward *p*NP-Glc in the presence of sodium azide were determined in 0.1 M sodium citrate, pH 6.0, containing 0–1 M sodium azide, for 10 min at 40 °C. Reactions were stopped by adding 2 M sodium carbonate. The concentration of *p*NP released from the reactions was measured as described in Section 2.6.

The kinetic parameters of the E490A variant toward 2,4-DNP-Glc in the presence of sodium azide were determined in 0.1 M sodium citrate, pH 6.0, containing 0–2 M sodium azide, at 40 °C. The rates of 2,4-DNP-Glc hydrolysis were determined by a continuous assay as described earlier. For this substrate and protein, a linear dependence of the initial rate of reaction (v_0) on substrate concentration [S] was observed, indicative of $[S] \ll K_m$ conditions; k_{cat}/K_m values were therefore determined by fitting the following equation to the data [36]: $v_0 = (k_{cat}/K_m)\,[E]_t\,[S]$.

^{1}H-NMR spectroscopy was utilized to determine the reaction products of the E490A variant with *p*NP-Glc and 2,4-DNP-Glc in the presence of sodium azide. ^{1}H-NMR data were recorded on a Bruker DMX 500 instrument at 500 MHz and 25 °C. Chemical shifts (δ) were quoted in parts per million (ppm) referenced to the solvent peak (HOD at 4.79 ppm). Determination of the anomeric configurations was based on the observed chemical shifts and coupling constants (*J* values, reported in Hz).

Table 2
Kinetic parameters of the recombinant wild-type β-glucosidase and mutant forms of *A. niger* ASKU28 β-glucosidase against cellobiose, *p*NP-Glc and 2,4-DNP-Glc in 0.1 M sodium citrate, pH 4.0, at 40 °C.

Substrate	Enzyme	K_m (mM)	*kcat* (s^{-1})	*kcat/Km* (s^{-1} mM^{-1})
Cellobiose	Wild-type	1.11 ± 0.09	67.4 ± 1.9	60.6 ± 5.4
	D463A	1.00 ± 0.05	78.4 ± 0.9	78.7 ± 4.4
	D487A	1.03 ± 0.01	87.7 ± 0.5	85.2 ± 0.8
	E490A	1.02 ± 0.05	0.001 ± 0.00003	0.001 ± 0.0001
	D496A	0.57 ± 0.02	78.3 ± 1.5	138 ± 6
	D501A	1.36 ± 0.04	47.3 ± 0.5	34.7 ± 1.0
*p*NP-Glc[a]	Wild-type	0.54 ± 0.04	141 ± 3	263 ± 18
	D463A	0.64 ± 0.03	160 ± 3	249 ± 12
	D487A	0.55 ± 0.02	163 ± 3	297 ± 14
	E490A	6.09 ± 0.15	0.32 ± 0.01	0.05 ± 0.002
	D496A	0.44 ± 0.03	141 ± 4	321 ± 21
	D501A	0.97 ± 0.02	104 ± 1	107 ± 2
2,4-DNP-Glc	Wild-type	0.49 ± 0.06	422 ± 25	863 ± 123
	E490A	0.10 ± 0.01	2.12 ± 0.13	21.2 ± 2.5
2,4-DNP-Glc (pH 6.0)	Wild-type	0.26 ± 0.04	71 ± 4.0	273 ± 44
	E490A	0.015 ± 0.003	0.34 ± 0.02	22.6 ± 4.7

[a] Non-Michealian kinetics were observed for pNP-Glc for the wild-type and most variants, which were indicative of substrate inhibition or transglycosylation at high [S] (Supplementary Fig. 7A). To simplify *comparative* kinetic analysis, *apparent kcat* and *Km* values are reported for pNP-Glc, which were obtained by fitting the standard Michaelis–Menten equation to data at [S]<2.5 mM (e.g. Supplementary Fig. 7B), except in the case of the E490A variant, where data in the range [S] = 0–20 mM were used (Supplementary Fig. 8).

Two reaction conditions were analyzed: (1) 36 μM of the E490A variant with 10 mM *p*NP-Glc in the presence of 50 mM sodium azide in 0.1 M sodium citrate, pD 6.0, at 37 °C, and (2) 31 nM of the E490A variant with 0.5 mM 2,4-DNP-Glc in the presence of 1 M sodium azide in 0.1 M sodium citrate, pD 6.1, at 37 °C. The E490A variant enzyme, sodium citrate (0.25 M, pH 6.0), sodium azide and substrates were each freeze-dried and subsequently dissolved in D_2O three times to reduce the proton concentration in the samples.

An initial ^{1}H-NMR spectrum of the mixture of sodium citrate buffer, sodium azide and the substrates was performed for comparison, prior to addition of the enzyme solution and subsequent NMR analysis over various time intervals. When *p*NP-Glc was used as substrate, each experiment was performed with 16 scans, whereas 128 scans acquired when 2,4-DNP-Glc was used as substrate. The following experimentally determined data for the substrates were used as spectral references:

*p*NP-Glc: ^{1}H NMR (500 MHz, D_2O): δ 8.24–8.21 (m, 2H, Ar–H), 7.23–7.20 (m, 2H, Ar–H), 5.27–5.22 (d, 1H, H1), 3.93 (dd, J = 12.4 and 2.1 Hz, 1H, H6), 3.75 (dd, J = 12.4 and 5.7 Hz, 1H, H6′), 3.69 (ddd, J = 9.7, 5.7 and 2.2 Hz, 1H, H5), 3.65–3.60 (m, 2H, H3, H4), 3.54–3.48 (m, 1H, H2).

2,4-DNP-Glc: ^{1}H NMR (500 MHz, D_2O): δ 8.90 (d, J = 2.8 Hz, 1H, Ar–H), 8.54 (dd, J = 9.4 and 2.8 Hz, 1H, Ar–H), 7.61 (d, J = 9.3 Hz, 1H, Ar–H), 5.43 (d, J = 7.6 Hz, 1H, H1), 3.94 (dd, J = 12.4 and 2.1 Hz, 1H, H6), 3.77 (dd, J = 12.3 and 5.6 Hz, 1H, H6′), 3.74–3.68 (m, 2H, H5, H3), 3.64 (at, 1H, J = 9.3 Hz, H4), 3.54 (at, 1H, J = 9.6 Hz, H2).

To detect the potential of the wild-type enzyme to utilize azide as a transglycosylation acceptor, analogous reactions were performed in 0.1 M sodium citrate, pH 6.0, at 37 °C and analyzed by thin-layer chromatography (TLC). 150 nM of the wild-type enzyme or 36 μM of the E490A variant were incubated with 10 mM *p*NP-Glc in the presence or absence of 50 mM sodium azide. Likewise, 6 nM of the wild-type enzyme or 31 μM of the E490A variant were incubated with 5 mM 2,4-DNP-Glc in the presence or absence of 1 M sodium azide. TLC was performed on silica gel 60 F_{254} aluminum plates (Merck, Darmstadt, Germany), followed by development in 7:2:1 (v/v/v) ethyl acetate/methanol/water and visualization with UV illumination and dipping into 10% sulfuric acid in methanol followed by charring.

3. Results and discussion

3.1. Cloning and expression of β-glucosidase from A. niger ASKU28

The coding sequence of mature β-glucosidase from *A. niger* ASKU28 (from position 19 of the amino acid sequence to the stop codon) was amplified with specific primers that were designed based on the nucleotide sequence of *A. niger* β-glucosidase (GenBank ID: CAB75696.1). The resulting RT-PCR product was 2546 bp long (GenBank ID: JX127252) and was found to be similar to the sequence of exons 2–7 of *A. niger bgl1* gene (GenBank ID: AJ132386.1) [11] with 32 positional differences. After translation to protein, the sequence of the recombinant β-glucosidase from *A. niger* ASKU28 was found to be nearly identical to that of β-glucosidase from *A. niger* B1 (GenBank ID: CAB75696.1) [11] with 1 amino acid difference of tyrosine at position 189 instead of phenylanine. Interestingly, the amino acid sequence of a homologous β-glucosidase from the genome sequence of *A. niger* CBS 513.88 (GenBank ID: CAK48740.1) has tyrosine at position 189, but differs from the *A. niger* ASKU28 sequence at 21 positions [22].

The coding sequence of the mature β-glucosidase from *A. niger* ASKU28 was then expressed as an N-terminal fusion protein with the α-mating factor and a polyhistidine tag. The expected size of the mature recombinant enzyme after the removal of the α-mating factor is 93 kDa. However, the recombinant enzyme showed an apparent molecular weight of 130 kDa by SDS–PAGE analysis (Supplementary Fig. 1), which decreased to 105 kDa after deglycosylation by Endo H_f enzyme (Supplementary Fig. 2), thus suggesting glycosylation by the yeast host, *P. pastoris* [37]. In keeping with this observation, the sequence of the mature β-glucosidase is predicted to contain 11 conserved N-glycosylation sites.

3.2. Sequence alignment, molecular modeling and production of mutant enzymes

The conserved acidic residues that could serve as the acid/base catalyst of the β-glucosidase were predicted by primary and tertiary structural alignments. Ninety-nine enzymes from GH3 were previously used to establish six subfamilies by phylogenetic analysis [28]; although the family now contains over 4000 members (http://www.cazy.org/GH3.html), these subfamilies are still broadly useful for descriptive purposes. In particular, the acid/base catalyst has been proposed to be conserved within the same subfamily but not conserved across subfamilies [17]. Sequence alignment of the *A. niger* ASKU28 β-glucosidase and β-glucosidases from *A. aculeatus* and *A. kawachii* in subfamily 4 showed 83% and 98% sequence identity, respectively, which confirmed assignment to this subfamily. The sequence of *A. niger* ASKU28 β-glucosidase was then aligned with those of the two previously reported β-glucosidases from *A. niger* [11,22], 14 enzymes from subfamily 4 of GH3 glycoside hydrolases [28], and 5 GH3 enzymes in which the identities of the acid/base catalysts are known (Supplementary Fig. 3).

From this alignment, the conserved catalytic nucleophile was identified as D261 within the VMSDW motif previously demonstrated for the *A. niger* B1 homologue [11]. However, potential candidate acid/base residues were not clearly identifiable, and moreover appeared to occur in a highly variable region of the sequence when considering other GH3 enzymes whose acid/base catalyst had been indicated by structural or kinetic analysis (Supplementary Fig. 3). In particular, it was only possible to achieve tenuous alignment of E458 of Bgl3B from *T. neapolitana* [14], E590 of KmBglI from *K. marxianus* [16], E473 of β-glucosidase from *F. meningosepticum* [17], E411 of glucosylceramidase from *Paenibacillus* sp. TS12 [12], and E491 of ExoI from barley [15] with E490 of *A. niger* ASKU28 β-glucosidase, which included significant gaps and mismatches in the neighboring residues. The acid/base catalysts of β-*N*-acetylglucosaminidase (NagZ) from *Bacillus subtilis* and β-N-acetylglucosaminidase 3A from *Clostridium paraputrificum* are located much further away from this region [13,18]. There are seven conserved acidic residues in *A. niger* ASKU28 β-glucosidase in the region flanking the reported acid/base catalysts, namely D463, E471, D487, E490, D496, D501 and D512. Of these, the residues E490 and D501 are fully conserved within subfamily GH3 4 enzymes.

In light of the ambiguity presented by sequence alignment, we then employed three-dimensional homology modeling to further inform our choice of target residues for site-directed mutagenesis. As of June 2012, the crystal structures of 6 bacterial and 2 eukaryotic GH3 members have been solved, as collected in the Carbohydrate Active EnZymes database (CAZy) [2]. These include the bacterial representatives SYNPCC7002_A0075 from *Synechococcus* sp. (unpublished, PDB ID: 3sql), DR1333 from *Deinococcus radiodurans* R1 (unpublished, PDB ID: 3tev), exo-1,3/1,4-β-glucanase (ExoP) from *Pseudoalteromonas* sp. [38], Bgl3B from *T. neapolitana* [14], β-N-acetylhexosaminidase from *Vibrio cholerae* [39,40], and NagZ from *B. subtilis* [13]. The two eukaryotic structures were those of the ExoI from barley [15] and KmBglI from *K. marxianus* [16]. The common structure of GH3 members includes an N-terminal $(\alpha/\beta)_8$ barrel which bears the catalytic nucleophilic residue. With the exception of certain bacterial β-*N*-acetylglucosaminidases, this domain is followed by a C-terminal $(\alpha/\beta)_6$ sandwich domain [1,13,28]. Rooted in seminal structural studies of the barley ExoI β-glucosidase [10,15] the catalytic acid/base residue appears to be generally borne on this second domain, although the His–Asp dyad implicated in one- and

two-domain bacterial β-GlcNAc'ases is contained within the first domain [13,19]. The structures of some GH3 enzymes are elaborated with additional domains. The extra C-terminal domain of ExoP from *Pseudoalteromonas* sp. appears to play a role in protein stability and may assist in substrate orientation [38]. KmBglI from *K. marxianus* has a PA14 domain inserted into the $(\alpha/\beta)_6$ sandwich domain that influences carbohydrate-binding, as well as a C-terminal fibronectin type III domain [16]. A similar additional C-terminal fibronectin type III domain is also found in Bgl3B from *T. neapolitana* and was suggested to be involved in substrate recognition and affect enzyme thermostability [14].

In this context, the three-dimensional structure of *A. niger* ASKU28 β-glucosidase was predicted by Phyre[2] to contain $(\alpha/\beta)_8$ barrel and $(\alpha/\beta)_6$ sandwich domains similar to the structure of BglI from the yeast *K. marxianus* (PDB ID: 3ac0, [16]); these enzymes share 33% amino acid identity. The quality of the model was further improved by an automated minimization by using the Chiron server. The Ramachandran plot (Supplementary Table 1), ProSa Z-scores, and Verify3D scores (Supplementary Fig. 4) indicated an acceptable model quality. Alignment of this model with the crystal structure of ExoI from barley in complex with glucose (PDB ID: 1ex1) revealed that residue E490 of the *A. niger* ASKU28 β-glucosidase was well-aligned with the predicted acid/base residue, E491, of barley ExoI (Fig. 1). To ensure completeness, five residues located in the non-conserved, flexible region of the *A. niger* β-glucosidase, namely D463, D487, E490, D496 and D501 (Fig. 1), were systematically subjected to site-directed mutagenesis to generate alanine variants for kinetic analysis. Two other residues, E471 and D512, were located in α-helices distal to the active site, and were not studied further.

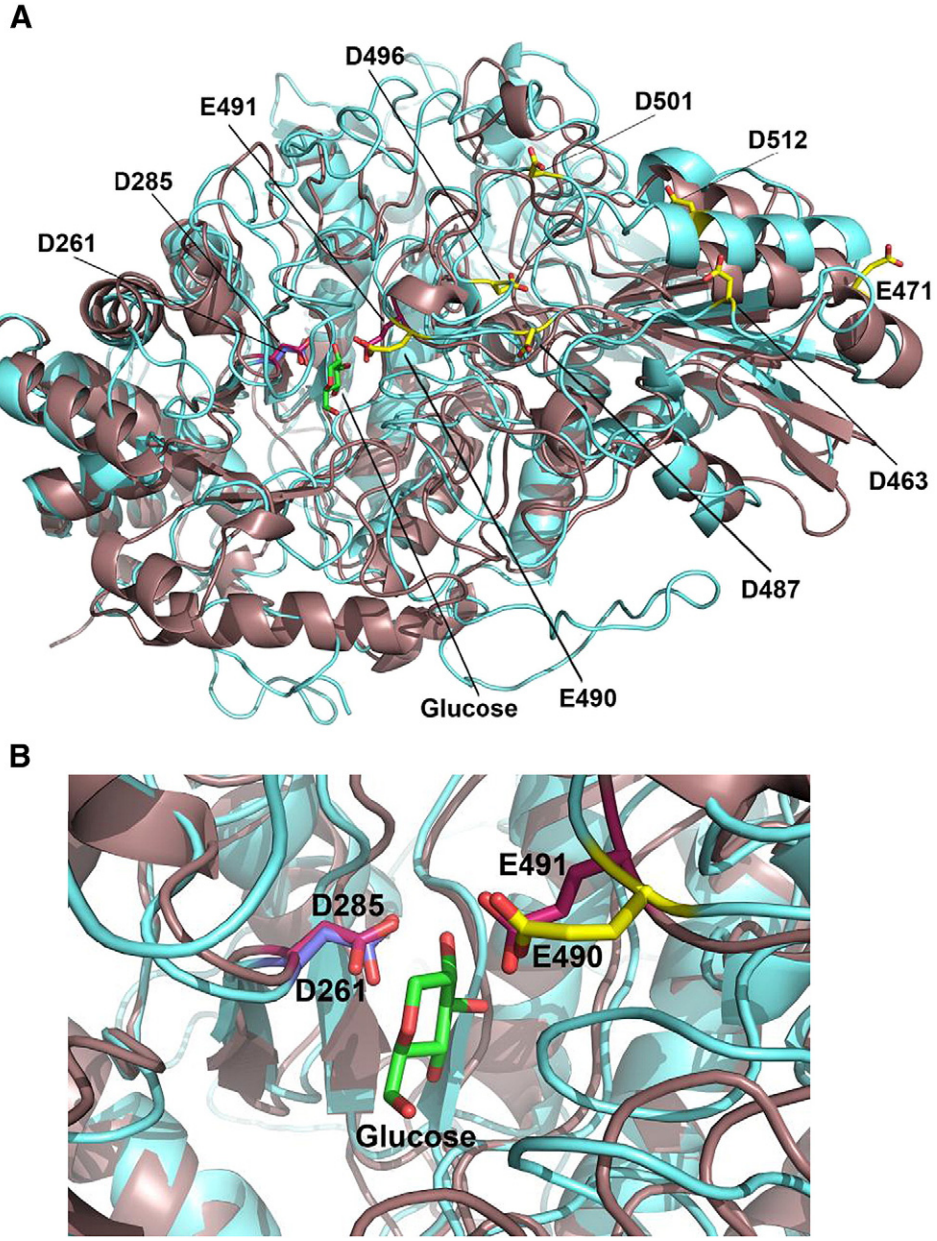


Fig. 1. Structural superimposition of the three-dimensional model of *A. niger* ASKU28 β-glucosidase (cyan) with the crystal structure of barley ExoI in complex with glucose (PDB ID: 1ex1, brown). (A) Overview (B) close-up of the active site region. Side chains of key residues are represented as stick models. The nucleophile and acid/base catalyst of ExoI (D285 and E491, respectively) are shown in magenta and the glucose molecule in the binding pocket of ExoI is shown in green. The side chain of the nucleophile (D261) in *A. niger* ASKU28 β-glucosidase is shown in purple. The side chains of the seven conserved acidic residues in *A. niger* ASKU28 β-glucosidase (D463, E471, D487, E490, D496, D501, D512) are shown in yellow. The picture was generated in PyMOL (Schrödinger LLC, Portland, OR, USA).

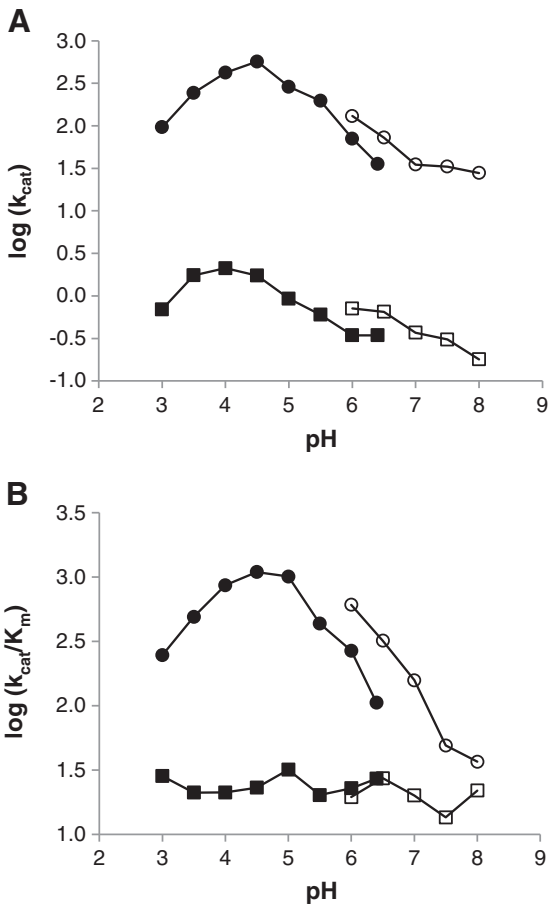


Fig. 2. pH dependence of kinetic parameters of the wild-type *A. niger* ASKU28 β-glucosidase (circles) and the E490A variant (squares) for the hydrolysis of 2,4-DNP-Glc in 0.1 M sodium citrate, pH 2.0–6.4 (filled symbols) and 0.1 M potassium phosphate, pH 6.0–8.0 (open symbols). (A) plot of log (*kcat*) vs pH; and (B) plot of log (*kcat*/*Km*) vs pH.

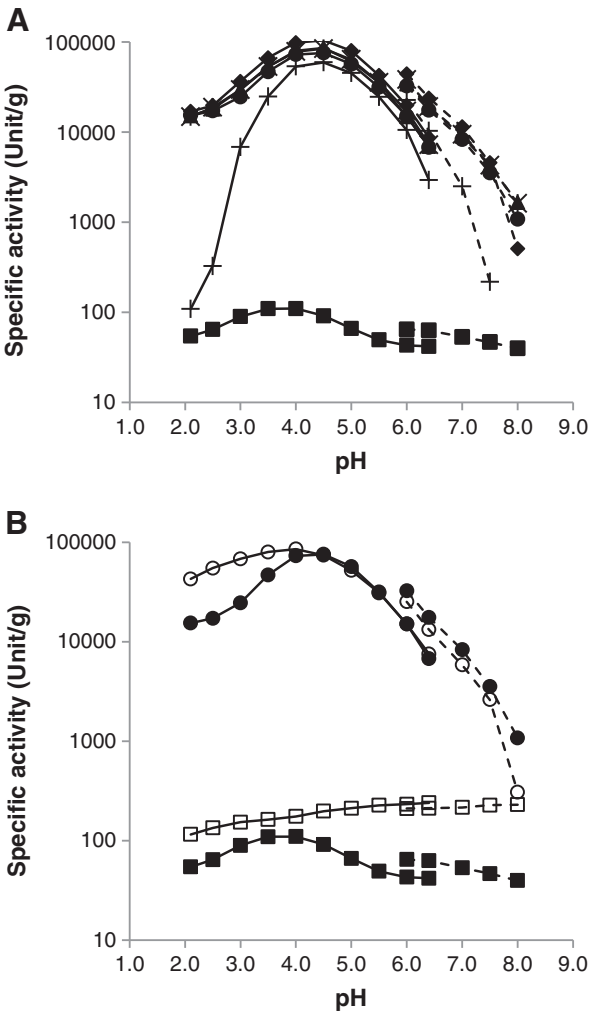


Fig. 3. pH dependence of the specific activities of the wild-type and all mutant forms of β-glucosidase from *A. niger* ASKU28 in the hydrolysis of *p*NP-Glc. The reactions were performed with *p*NP-Glc at 2 mM for the wild-type and three mutant enzymes, 4 mM for the D501A mutant, and 24 mM for the E490A mutant, in 50 mM sodium citrate, pH 2.0–6.4 (solid lines) and 50 mM potassium phosphate, pH 6.0–8.0 (dash lines). (A) pH profile of all enzymes; and (B) pH profile of the wild-type enzyme and the E490A mutant in the presence (open symbols) and absence (filled symbols) of 50 mM sodium azide. Wild-type (●), D463A (▲), D487A (♦), E490A (■), D496A (*) and D501A (+).

3.3. Kinetic analysis

The steady-state kinetic parameters of the recombinant wild-type and mutant forms of *A. niger* ASKU28 β-glucosidase against cellobiose, *p*NP-Glc and 2,4-DNP-Glc, were determined at the optimal pH, 4.0, and 40 °C (Table 2 and Supplementary Figs. 5–10). For cellobiose and *p*NP-Glc, the D463A, D487A, D496A, and D501A variants all exhibited kinetic constants that were similar to the wild-type enzyme, although small variations in k_{cat} and K_m values were observed, indicating subtle effects on catalysis (Table 2).

In stark contrast, the E490A variant exhibited a large depression of the k_{cat} and k_{cat}/K_m values (at least 67,000- and 61,000-fold, respectively) in comparison to the wild-type enzyme when cellobiose was used as a substrate, although the K_m value was within error of that of the wild-type enzyme (Table 2, Supplementary Figs. 5 & 6). With the aryl β-glucoside *p*NP-Glc, the E490A variant also exhibited significant decreases in the *apparent* k_{cat} and k_{cat}/K_m values (440- and 5300-fold, respectively) when compared to the wild-type enzyme (Table 2). The accompanying 11-fold increase in the *apparent* K_m value of the E490A variant for this substrate may be explained by elimination of substrate transglycosylation [23] and a resulting shift from non-Michaelian to classic Michaelis–Menten saturation kinetics [41], as reflected in $v_o/[E]_t$ vs. [S] plots (Supplementary Figs. 7 & 8).

When the activity of the E490A variant was tested on 2,4-DNP-Glc, which bears the good leaving group 2,4-dinitrophenol (pK_a 4.1), the k_{cat} and K_m values were ca. 200- and 5-fold lower relative to the wild-type enzyme, resulting in a ca. 40-fold decrease of the k_{cat}/K_m value at the pH optimum of the wild-type enzyme, 4.0 (Table 2). As a prelude to azide rescue studies and pH-rate profile measurements (vide infra), we also measured activity toward 2,4-DNP-Glc at pH 6.0. In this case, the k_{cat} and K_m values of the E490A variant were reduced ca. 210- and 17-fold relative to the wild-type enzyme, resulting in a ca. 12-fold decrease of the k_{cat}/K_m value (Table 2). At all pH values examined, plots of $v_o/[E]_t$ vs. [2,4-DNP-Glc] for both the wild-type and E490A variant were classically Michaelian over the substrate ranges

Table 3
The effect of sodium azide on kinetic parameters of the recombinant wild-type β-glucosidase toward *p*NP-Glc at pH 6.0, 40 °C for 10 min.

[NaN_3] (mM)	*K*m (mM)	*k*cat (s^{-1})	*k*cat/*K*m (s^{-1} mM^{-1})
0	0.35 ± 0.01	37.2 ± 0.2	105 ± 2
5	0.38 ± 0.02	37.7 ± 1.1	98.2 ± 6.5
10	0.39 ± 0.02	37.7 ± 1.3	97.6 ± 6.7
30	0.37 ± 0.003	36.1 ± 0.2	96.7 ± 1.1
50	0.38 ± 0.004	34.6 ± 0.1	91.4 ± 1.0
200	0.32 ± 0.01	24.2 ± 0.5	74.6 ± 3.0
500	0.31 ± 0.008	15.4 ± 0.2	48.7 ± 1.3

Table 4
The effect of sodium azide on kinetic parameters of the E490A mutant toward *p*NP-Glc at pH 6.0, 40 °C for 10 min.

[NaN_3] (mM)	*K*m (mM)	*k*cat (s^{-1})	*k*cat/Km (s^{-1} mM^{-1})
0	1.49 ± 0.06	0.15 ± 0.003	0.10 ± 0.005
1	9.80 ± 0.64	0.89 ± 0.03	0.09 ± 0.01
2	8.40 ± 0.55	0.79 ± 0.03	0.09 ± 0.01
5	5.72 ± 0.88	0.58 ± 0.07	0.10 ± 0.02
10	8.61 ± 0.67	0.84 ± 0.05	0.10 ± 0.01
30	9.26 ± 0.39	0.88 ± 0.03	0.09 ± 0.01
50	8.33 ± 0.51	0.85 ± 0.04	0.10 ± 0.01
200	7.92 ± 0.62	0.87 ± 0.05	0.11 ± 0.01
500	4.85 ± 0.23	0.56 ± 0.02	0.12 ± 0.01
800	3.52 ± 0.08	0.42 ± 0.004	0.12 ± 0.003
1000	7.01 ± 1.00	0.72 ± 0.06	0.10 ± 0.02

examined (≤1 mM for wild-type and ≤0.1 mM for the E490A variant; Supplementary Figs. 9–10 and data not shown).

Taken together, these results above are consistent with the assignment of the side chain of E490 as acid/base catalyst in the *A. niger* ASKU28 β-glucosidase. By analogy with the homologous enzyme from *A. niger* strain B1 [11], the GH3 β-glucosidase can be reliably predicted to use the canonical retaining mechanism of glycoside hydrolysis. As such, the hydrolytic breakdown of the common α-glucosyl-enzyme intermediate can be expected to be slowed equally for all three substrates examined here, due to removal of the acid/base catalyst in the mutant enzyme. For substrates with poor leaving groups, effects on the first chemical step, as revealed by k_{cat}/K_m values, are expected to be especially significant, as observed for the likely natural substrate, cellobiose (leaving-group pK_a ca. 16 [42]). The similar magnitude of reduction of both the k_{cat} and k_{cat}/K_m values for cellobiose in the E490A variant, because the K_m value is unaffected relative to the wild-type, implies that enzyme glycosylation (k_2) is rate-determining for this substrate.

In contrast, significant reductions of the K_m value are observed in the E490A variant at both optimal (4.0) and non-optimal (6.0) pH values on the artificial substrate 2,4-DNP-Glc (Table 2), the aglycone of which is a comparatively good leaving group (pK_a 4.1) that generally does not benefit greatly from proton assistance in glycosidases [43]. Given the complex dependence of K_m on both the rate of enzyme glycosylation (k_2) and deglycosylation (k_3) in two-step enzymes such as retaining glycosidases ($K_m = [k_3/(k_2 + k_3)] \cdot [(k_2 + k_{-1})/k_1]$; [44]), these observations suggest that k_3 is now significantly rate-limiting due to removal of the acid/base catalyst. Consequently, k_{cat} will reflect an increased dependence on, and may become equivalent to, k_3 ($k_{cat} = k_2 \cdot k_3/(k_2 + k_3)$; [44]). This effect is most pronounced at pH 6.0 and above (Table 2 and Fig. 2).

Indeed, we measured kinetic parameters for hydrolysis of 2,4-DNP-Glc by the wild-type enzyme and the E490A variant over the pH range 3–8 (Fig. 2). Plots of $\log(k_{cat})$ and $\log(k_{cat}/K_m)$ versus pH for the wild-type enzyme were curved as expected for the canonical retaining glycosidase mechanism involving two ionizable active-site residues. The observation that the plot of $\log(k_{cat})$ versus pH for the E490A variant was not strictly linear, as previously observed for other acid/base residue variants [12,45–47], may reflect a limited degree of buffer or specific acid catalysis in the deglycosylation step in the low pH regime. In contrast, the pH-independence of $\log(k_{cat}/K_m)$ observed for the E490A variant indicated the removal of catalytic assistance from general acid reside in the enzyme glycosylation step. A similar effect on k_{cat}/K_m values has been observed for the acid/base catalyst variant of the *Cellulomonas fimi* glycanase, Cex [48].

For *p*NP-Glc, the precise effects of E490A mutation on the pH dependence of the k_{cat} and k_{cat}/K_m values are difficult to divine, because of the complex kinetic profile exhibited by the wild-type enzyme (Supplementary Fig. 7A), which is indicative of substrate transglycosylation [41]. Contrasted with the simple Michaelian kinetics exhibited by the E490A variant, it is therefore difficult to draw meaningful conclusions regarding individual kinetic steps. Nonetheless, we measured specific activity values for the wild-type and all variants at either the substrate concentration yielding the maximum rate, or in the case of the E490A variant, at a concentration four-fold greater than the K_m value (Fig. 3A). Whereas the wild-type enzyme and all of the other variants exhibited curved pH-rate profiles as expected, that of the E490A was essentially linear, mirroring the results for 2,4-DNP-Glc (Fig. 2).

3.4. Chemical rescue by azide

Azide rescue experiments were undertaken to further confirm the identity of E490 as the catalytic acid/base of β-glucosidase from *A. niger* ASKU28. The kinetic parameters for *p*NP-Glc and 2,4-DNP-Glc were determined for the wild-type and E490A variant in the presence of increasing amounts of sodium azide at 40 °C and pH 6.0 (Tables 3–5). This non-optimal pH was chosen to minimize the amount of volatile, toxic hydrazoic acid (HN_3, pK_a 4.6) in the assay. High concentrations of azide had a slightly inhibitory effect on the wild-type enzyme (50% inhibition at 500 mM) due exclusively to the reduction of the k_{cat} value (Table 3); a similar inhibitory effect of azide was previously observed, for example, in a wild-type GH16 enzyme [44]. The addition of azide did not significantly affect the pH-rate profile of the wild-type *A. niger* ASKU28 β-glucosidase (Fig. 3B).

In contrast, the addition of sodium azide into assay mixtures of the E490A variant and *p*NP-Glc resulted in the direct activation of the enzyme reaction by increasing k_{cat} values, even at low concentrations (1 mM, Table 4). K_m values were simultaneously increased, possibly reflecting decreased accumulation of the glycosyl-enzyme due to a shift in the relative rates of formation and breakdown. Moreover, the effect of azide activation on specific activity values was pH independent over the range 3–8 (Fig. 3B).

Importantly, reaction product monitoring by NMR in a parallel reaction containing 50 mM sodium azide and 10 mM *p*NP-Glc (Fig. 4) directly indicated the production of β-glucosyl azide, the expected product from decomposition of the α-glucosyl-enzyme intermediate by nucleophilic attack (Scheme 1). After 1.5 h of incubation, 12% of the β-glucosyl azide product was produced, as indicated by the presence of a characteristic doublet at 4.73 ppm ($J = 8.8$ Hz), corresponding to the H1 peak [49,50]. Concomitantly, peaks due to the *p*-nitrophenol by-product were also detected at 8.18–8.14 ppm (m, Ar–H) and 6.94–6.90 ppm (m, Ar–H) as the intensity of the H1 peak of the starting material (5.27 ppm) decreased. Ultimately, characteristic doublets corresponding to the H1 peak of the hydrolysis products, α- and β-D-glucose (5.22 ppm ($J = 3.8$ Hz) and 4.63 ppm ($J = 8.0$ Hz), respectively [51]), were detected after 24 h and 48 h of incubation, at which time 75% of the substrate had been consumed. A control experiment containing 10 mM *p*NP-Glc, 50 mM sodium azide, and 0.1 M of sodium citrate, pD 6.0, but lacking the enzyme, indicated no background decomposition of the substrate after 24 h at 37 °C (data not shown).

Similar results were observed upon addition of sodium azide to the reaction of 2,4-DNP-Glc catalyzed by the E490A variant. Although an essentially linear dependence of $v_o/[E]_t$ on donor substrate concentration precluded extraction of individual kinetic parameters for all assays containing sodium azide (50 mM–2 M), increased k_{cat}/K_m values were indicative of activation (Table 5 and Supplementary Fig. 11). In parallel, reaction monitoring by ^1H NMR again revealed the generation of the

Table 5
The kinetic parameters of azide reactivation assays with 2,4-DNP-Glc at pH 6.0, 40 °C.

Enzyme	[NaN_3] (mM)	*k*cat/Km (s^{-1} mM^{-1})
Wild-type	0	197 ± 29
E490A	0	18.4 ± 5.5
	50	66.1 ± 7.4
	500	174.3 ± 5.7
	1000	217.1 ± 9.2
	2000	249.2 ± 7.2

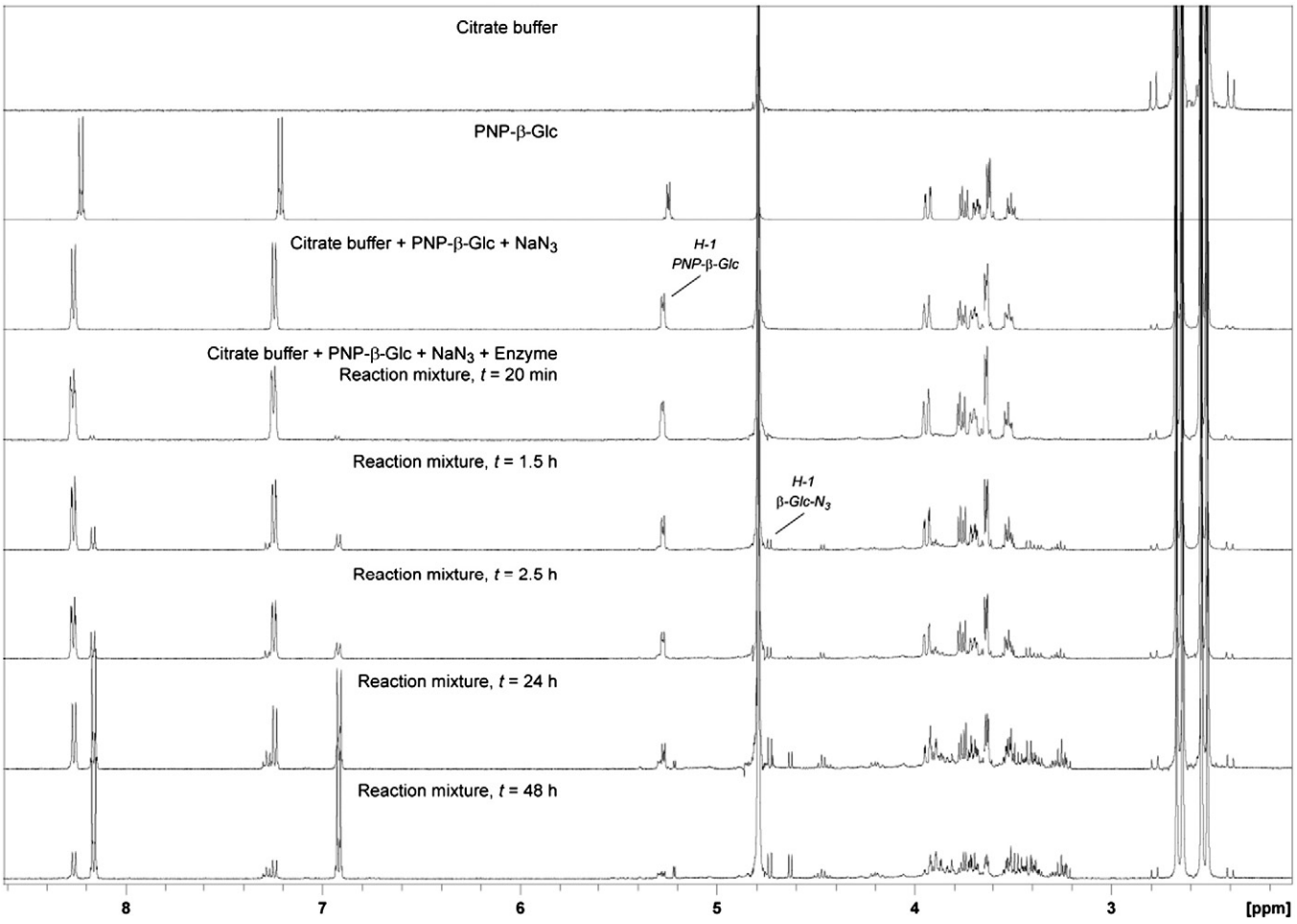
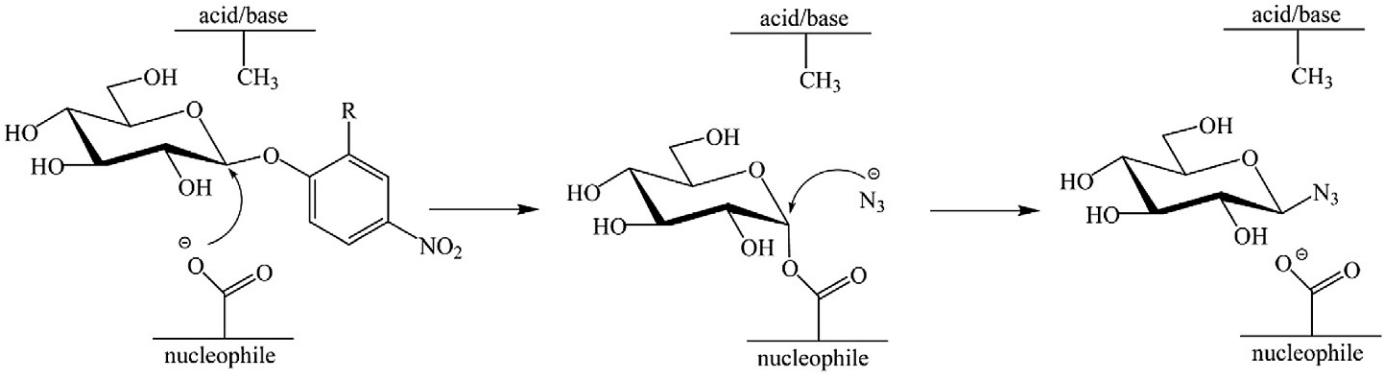


Fig. 4. [1]H-NMR spectra of the reaction between the E490A mutant with 10 mM *p*NP-Glc in the presence of 50 mM sodium azide in 0.1 M sodium citrate, pD 6.0, at 37 °C.

expected β-glucosyl azide from 0.5 mM 2,4-DNP-Glc and 1 M sodium azide in the presence of the E490A variant (Fig. 5). After 1.5 h of incubation, the characteristic doublet of the H1 of β-glucosyl azide was observed at 4.83 ppm ($J = 8.8$ Hz), and the signal increased over time. The doublet at 5.53 ppm ($J = 7.6$ Hz), which corresponds to H1 of 2,4-DNP-Glc, continually decreased and was completely absent after 24 h of incubation. No hydrolysis products (α- and β-D-glucose) were detected. The slight shift of spectral peaks compared to reported values [49–51] is ascribed to solute concentration effects due to complex sample preparation.

Thin-layer chromatographic analysis of analogous reaction mixtures containing either phenyl β-glucoside revealed that wild-type *A. niger* ASKU28 β-glucosidase produced only the expected phenol plus glucose, both in the presence and absence of the competitive nucleophile azide (data not shown, see Section 2.7 for experimental



pNP-Glc, R = H
2,4-DNP-Glc, R = NO_2

Scheme 1. Azide rescue mechanism of an alanine mutant of the acid/base catalyst β-glucosidase toward *p*NP-Glc or 2,4-DNP-Glc. In the absence of sodium azide, the catalytic acid/base mutant enzyme was trapped in the glucosyl-enzyme intermediate form. The additional sodium azide then acted as an exogenous base to attack the glucosyl-enzyme intermediate and reactivated the mutant enzyme yielding β-glucosyl azide product.

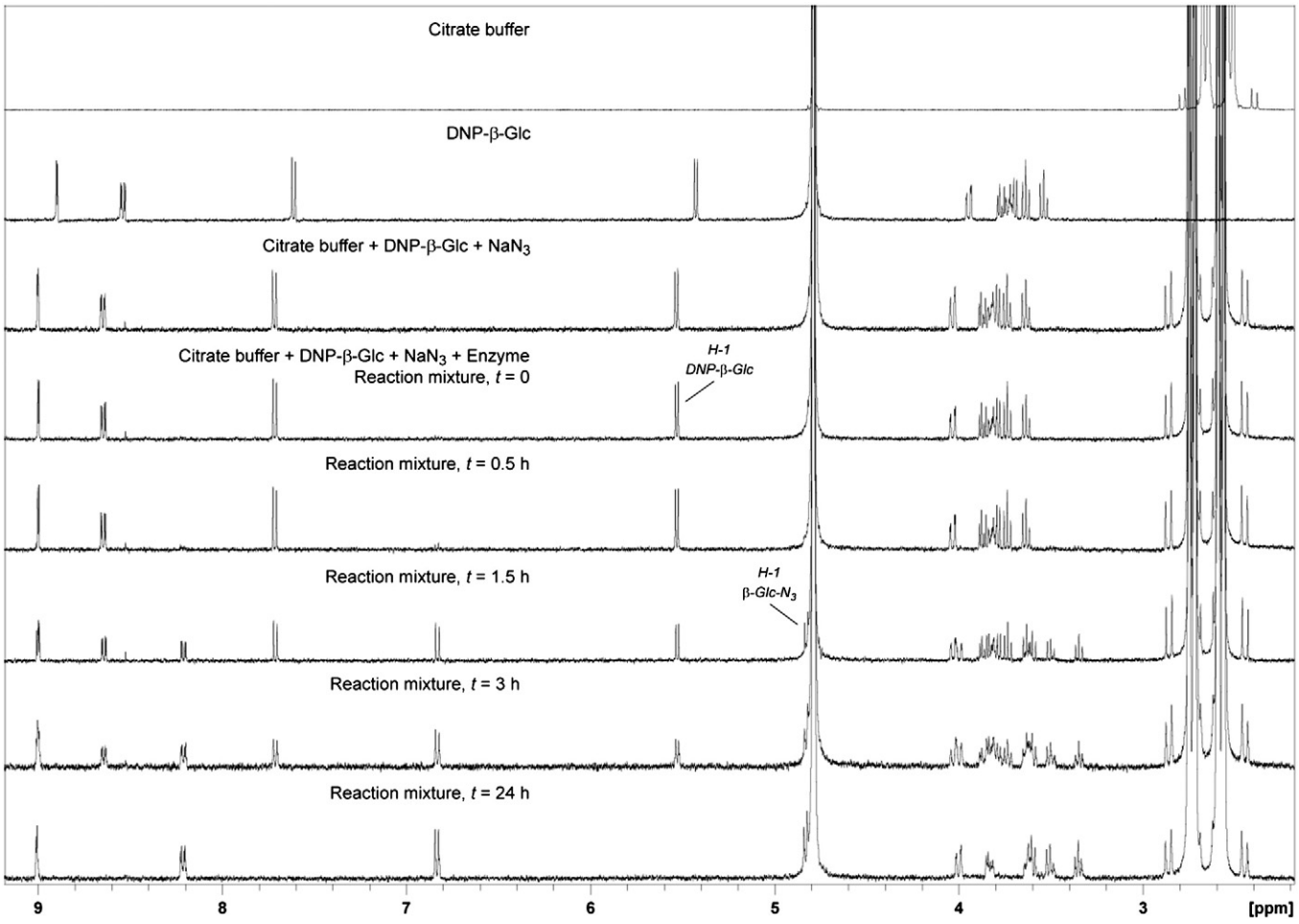


Fig. 5. ^{1}H-NMR spectra of the reaction between the E490A mutant with 0.5 mM 2,4-DNP-Glc in the presence of 1 M sodium azide in 0.1 M sodium citrate, pD 6.1, at 37 °C.

details). This indicated that transglycosylation to azide, which is mechanistically possible, is kinetically insignificant in the wild-type enzyme and a unique property of the E490A variant.

3.5. Updated comparison of acid/base catalysts among GH3 enzymes

As discussed above, global multiple sequence alignments of GH3 enzyme often fail to successfully identify the homologous catalytic acid/base residue across diverse family members, due to low conservation of primary protein structure in the flanking regions. The likely acid/base of *A. niger* ASKU28 β-glucosidase, E490, occurs within the sequence motif SG**E**GY (Fig. 6). E490 is highly conserved within the consensus sequence (S/A)(G/S/T/R)**E**(G/N)(Y/V/F/E/T) of GH3 subfamily 4 enzymes (see also Supplementary Fig. 3). Conservation of the proposed acid/base also extends to the (S/E)(G/T/S)**E**(G/S)(Y/S) motif of GH3 subfamily 5 enzymes whose acid/base catalysts have been reported [12,14,16,17]. The sequence surrounding the acid/base catalyst in the distantly related barley ExoI of subfamily 1 [28] is often not readily aligned with the corresponding region in these microbial GH3 members using automated alignment algorithms (especially ClustalW, data not shown; also ClustalΩ, Supplementary Fig. 3). This specifically highlights the challenge in broadly predicting this residue across those GH3 subfamilies whose members are comprised of two domains [28] based on sequence data alone. In such cases, three-dimensional structure homology modeling (Fig. 1) can specifically inform manual sequence alignment (Fig. 6).

4. Conclusions

The data presented here provide essential experimental support for the identification of glutamic acid 490 as the acid/base catalyst in the β-glucosidase of *A. niger* ASKU28, and, by analogy, the assignment of this catalytically important residue in a range of fungal GH3 members. Furthermore, this analysis highlighted a clear link regarding catalytic sequence conservation between these fungal members of subfamily 4 and bacterial/yeast members of subfamily 5. Given the vast expansion in the number of members of GH3 since the turn of the millenium, a revised, robust subfamily classification [52–54] is sorely needed to allow extension of this analysis to the family at-large. We anticipate that our present work will inform future bioinformatic and structural enzymology studies concerning GH3 enzymes in the contexts of rapidly emerging (meta)genomic sequence data and the industrial utility of these enzymes in biomass modification.

Acknowledgments

The authors especially thank Assoc. Prof. Vichien Kitpreechavanich (Kasetsart University, Thailand) for providing *Aspergillus niger* ASKU28 and Prof. Stephen G. Withers (University of British Columbia, Canada) for providing 2,4-DNP-Glc. H.B. also thanks Prof. Withers and Prof. Gideon J. Davies (University of York, UK) for enlightening discussions. P.T. is a recipient of the Ph.D. scholarship from the Office of the Higher Education Commission, Thailand. A.C.A. was supported by the Swedish Research Council Formas (via "CarboMat — The KTH Advanced

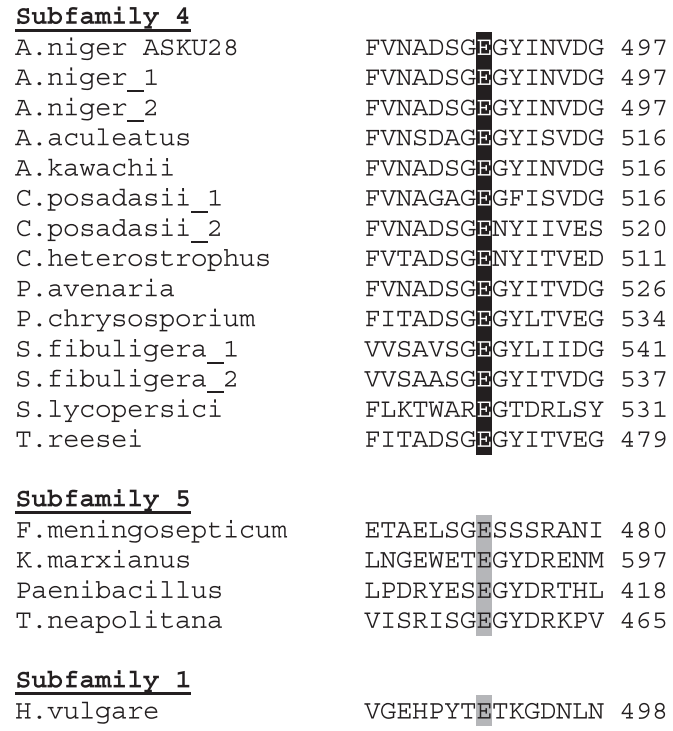


Fig. 6. Manual alignment of partial sequences of GH3 subfamily 4 members including subfamily 1 and 5 members with identified catalytic acid/base residues. The conserved glutamic residues are shown as white letters against black background, except for those that have been experimentally determined previously, which are highlighted in gray. Full species names and Genbank IDs of the enzymes are as follows, *Aspergillus niger* ASKU28 β-glucosidase, JX127252 (this study); *A. niger*_1 β-glucosidase, CAB75696.1; *A. niger*_2 β-glucosidase *bgl*I, CAK48740.1; *A. aculeatus* β-glucosidase precursor, BAA10968.1; *A. kawachii* β-glucosidase, BAA19913.1; *Coccidioides posadasii*_1 β-glucosidase, AAB67972.1; *C. Posadasii*_2 β-glucosidase precursor, AAF21242.1; *Cochliobolus heterostrophus* β glucosidase homolog, AAB84005.1; *Phaeosphaeria avenaria* β-glucosidase, CAB82861.1; *Phanerochaete chrysosporium* β-glucosidase, BAB85988.1; *Saccharomycopsis fibuligera*_1 β-glucosidase 1 precursor, AAA34314.1; *S. fibuligera*_2 β-glucosidase 2 precursor, AAA34315.1; *Septoria lycopersici* β-1,2-D-glucosidase, AAB08445.1; *Trichoderma reesei* β-D-glucoside glucohydrolase, AAA18473.1; *Flavobacterium meningosepticum* β-glucosidase, AAB66561.1; *Kluyveromyces marxianus* KmBgll, ACY95404.1; *Paenibacillus* sp. TS12 glucosylceramidase, BAC16750.1; *Thermotoga neapolitana* Bgl3B, ABI29899.1; *Hordeum vulgare* Exol, AAD23382.1.

Carbohydrate Materials Consortium") and L.S.M. was supported by the Knut & Alice Wallenberg Foundation (via the Wallenberg Wood Science Centre). Additional funding from the Michael Smith Laboratories, the National Research Council of Thailand and the Office of the Higher Education Commission, Thailand, is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbagen.2012.11.014.

References

[1] J.R. Ketudat Cairns, A. Esen, β-Glucosidases, Cell. Mol. Life Sci. 67 (2010) 3389–3405.
[2] B.L. Cantarel, P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics, Nucleic Acids Res. 37 (2009) D233–D238.
[3] S. Sansenya, R. Opassiri, B. Kuaprasert, C.J. Chen, J.R. Ketudat Cairns, The crystal structure of rice (*Oryza sativa* L.) Os4BGlu12, an oligosaccharide and tuberonic acid glucoside-hydrolyzing β-glucosidase with significant thioglucohydrolase activity, Arch. Biochem. Biophys. 510 (2011) 62–72.
[4] M. Hrmova, A.J. Harvey, J. Wang, N.J. Shirley, G.P. Jones, B.A. Stone, P.B. Høj, G.B. Fincher, Barley β-D-glucan exohydrolases with β-D-glucosidase activity, J. Biol. Chem. 271 (1996) 5277–5286.
[5] G. Legler, K.-R. Roeser, H.-K. Illig, Reaction of β-D-glucosidase A3 from *Aspergillus wentii* with D-Glucal, Eur. J. Biochem. 101 (1979) 85–92.
[6] D.J. Vocadlo, C. Mayer, S. He, S.G. Withers, 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-alpha-L-idopyranosyl fluoride, Biochemistry 39 (2000) 117–126.
[7] D.L. Zechel, S.G. Withers, Glycosidase mechanisms: anatomy of a finely tuned catalyst, Acc. Chem. Res. 33 (1999) 11–18.
[8] M.L. Sinnott, Catalytic mechanism of enzymic glycosyl transfer, Chem. Rev. 90 (1990) 1171–1202.
[9] G.J. Davies, M.L. Sinnott, Sorting the diverse: the sequence-based classifications of carbohydrate-active enzymes, Biochem. J. (2008), http://dx.doi.org/10.1042/bj20080382.
[10] M. Hrmova, J.N. Varghese, R. De Gori, B.J. Smith, H. Driguez, G.B. Fincher, Catalytic mechanisms and reaction intermediates along the hydrolytic pathway of a plant β-D-glucan glucohydrolase, Structure 9 (2001) 1005–1016.
[11] S. Dan, I. Marton, M. Dekel, B.A. Bravdo, S. He, S.G. Withers, O. Shoseyov, Cloning expression characterization and nucleophile identification of family 3, *Aspergillus niger* β-glucosidase, J. Biol. Chem. 275 (2000) 4973–4980.
[12] K. Paal, M. Ito, S.G. Withers, *Paenibacillus* sp. TS12 glucosylceramidase: kinetic studies of a novel sub-family of family 3 glycosidases and identification of the catalytic residues, Biochem. J. 378 (2004) 141–149.
[13] S. Litzinger, S. Fischer, P. Polzer, K. Diederichs, W. Welte, C. Mayer, Structural and kinetic analysis of *Bacillus subtilis* N-acetylglucosaminidase reveals a unique Asp–His dyad mechanism, J. Biol. Chem. 285 (2010) 35675–35684.
[14] T. Pozzo, J.L. Pasten, E.N. Karlsson, D.T. Logan, Structural and functional analyses of β-glucosidase 3B from *Thermotoga neapolitana*: a thermostable three-domain representative of glycoside hydrolase 3, J. Mol. Biol. 397 (2010) 724–739.
[15] J.N. Varghese, M. Hrmova, G.B. Fincher, Three-dimensional structure of a barley β-D-glucan exohydrolase, a family 3 glycosyl hydrolase, Structure 7 (1999) 179–190.
[16] E. Yoshida, M. Hidaka, S. Fushinobu, T. Koyanagi, H. Minami, H. Tamaki, M. Kitaoka, T. Katayama, H. Kumagai, Role of a PA14 domain in determining substrate specificity of a glycoside hydrolase family 3 β-glucosidase from *Kluyveromyces marxianus*, Biochem. J. 431 (2010) 39–49.
[17] J. Chir, S. Withers, C.F. Wan, Y.K. Li, 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 (2002) 857–863.
[18] H. Li, G. Zhao, H. Miyake, H. Umekawa, T. Kimura, K. Ohmiya, K. Sakka, Identification of a catalytic residue of *Clostridium paraputrificum N*-acetyl-β-D-glucosaminidase Nag3A by site-directed mutagenesis, Biosci. Biotechnol. Biochem. 70 (2006) 1127–1133.
[19] J.-P. Bacik, G.E. Whitworth, K.A. Stubbs, D.J. Vocadlo, B.L. Mark, Active Site Plasticity within the Glycoside Hydrolase NagZ Underlies a Dynamic Mechanism of Substrate Distortion, Chem. Biol. 19 (2012) 1471–1482.
[20] J. van den Brink, R.P. de Vries, Fungal enzyme sets for plant polysaccharide degradation, Appl. Microbiol. Biotechnol. 91 (2011) 1477–1492.
[21] M.H. Jaime Eyzaguirre, A. Leschot, β-Glucosidases from filamentous fungi: properties structure and applications, in: K.J. Yarema (Ed.), Handbook of Carbohydrate Engineering, CRC Press, Florida, 2005, pp. 645–685.
[22] H.J. Pel, J.H. de Winde, D.B. Archer, P.S. Dyer, G. Hofmann, P.J. Schaap, G. Turner, R.P. de Vries, R. Albang, K. Albermann, M.R. Andersen, J.D. Bendtsen, J.A.E. Benen, M. van den Berg, S. Breestraat, M.X. Caddick, R. Contreras, M. Cornell, P.M. Coutinho, E.G.J. Danchin, A.J.M. Debets, P. Dekker, P.W.M. van Dijck, A. van Dijk, L. Dijkhuizen, A.J.M. Driessen, C. d'Enfert, S. Geysens, C. Goosen, G.S.P. Groot, P.W.J. de Groot, T. Guillemette, B. Henrissat, M. Herweijer, J.P.T.W. van den Hombergh, C.A.M.J.J. van den Hondel, R.T.J.M. van der Heijden, R.M. van der Kaaij, F.M. Klis, H.J. Kools, C.P. Kubicek, P.A. van Kuyk, J. Lauber, X. Lu, M.J.E.C. van der Maarel, R. Meulenberg, H. Menke, M.A. Mortimer, J. Nielsen, S.G. Oliver, M. Olsthoorn, K. Pal, N.N.M.E. van Peij, A.F.J. Ram, U. Rinas, J.A. Roubos, C.M.J. Sagt, M. Schmoll, J. Sun, D. Ussery, J. Varga, W. Vervecken, P.J.J. van de Vondervoort, H. Wedler, H.A.B. Wosten, A.-P. Zeng, A.J.J. van Ooyen, J. Visser, H. Stam, Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88, Nat. Biotechnol. 25 (2007) 221–231.
[23] C. Bohlin, E. Praestgaard,M. Baumann, K. Borch, J. Praestgaard, R.Monrad, P.Westh, A comparative study of hydrolysis and transglycosylation activities of fungal β-glucosidases, Appl. Microbiol. Biotechnol. (in press), http://dx.doi.org/10.1007/s00253-012-3875-9.
[24] H.F. Seidle, K. McKenzie, I. Marten, O. Shoseyov, R.E. Huber, Trp-262 is a key residue for the hydrolytic and transglucosidic reactivity of the *Aspergillus niger* family 3 β-glucosidase: Substitution results in enzymes with mainly transglucosidic activity, Arch. Biochem. Biophys. 444 (2005) 66–75.
[25] H.F. Seidle, S.J. Allison, E. George, R.E. Huber, Trp-49 of the family 3 β-glucosidase from *Aspergillus niger* is important for its transglucosidic activity: creation of novel β-glucosidases with low transglucosidic efficiencies, Arch. Biochem. Biophys. 455 (2006) 110–118.
[26] C. Riou, J.-M. Salmon, M.-J. Vallier, Z. Günata, P. Barre, Purification, characterization, and substrate specificity of a novel highly glucose-tolerant β-glucosidase from *Aspergillus oryzae*, Appl. Environ. Microbiol. 64 (1998) 3607–3614.
[27] P. Toonkool, P. Metheenukul, P. Sujiwattanarat, P. Paiboon, N. Tongtubtim, M. Ketudat-Cairns, J.R. Ketudat-Cairns, J. Svasti, Expression and purification of dalcochinase, a β-glucosidase from *Dalbergia cochinchinensis* Pierre, in yeast and bacterial hosts, Protein Expr. Purif. 48 (2006) 195–204.
[28] A.J. Harvey, M. Hrmova, R. De Gori, J.N. Varghese, G.B. Fincher, Comparative modeling of the three-dimensional structures of family 3 glycoside hydrolases, Proteins 41 (2000) 257–269.
[29] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol. 7 (2011).

[30] L.A. Kelley, M.J.E. Sternberg, Protein structure prediction on the Web: a case study using the Phyre server, Nat. Protoc. 4 (2009) 363–371.
[31] S. Ramachandran, P. Kota, F. Ding, N.V. Dokholyan, Automated minimization of steric clashes in protein structures, Proteins 79 (2011) 261–270.
[32] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[33] D. Eisenberg, R. Lüthy, J.U. Bowie, [20] VERIFY3D: assessment of protein models with three-dimensional profiles, R.M.S. Charles, W. Carter Jr. (Eds.), Methods Enzymol. 277 (1997) 396–404, (Academic Press).
[34] J. Braman, C. Papworth, A. Greener, Site-directed mutagenesis using double-stranded plasmid DNA templates, Methods Mol. Biol. 57 (1996) 31–44.
[35] R.J. Bryant, D.R. Rao, S. Ogutu, α and β-galactosidase activities and oligosaccharide content in peanuts, Plant Foods Hum. Nutr. 58 (2003) 213–223.
[36] J. Wicki, J. Schloegl, C.A. Tarling, S.G. Withers, Recruitment of both uniform and differential binding energy in enzymatic catalysis: Xylanases from families 10 and 11, Biochemistry 46 (2007) 6996–7005.
[37] J. Cregg, J. Cereghino, J. Shi, D. Higgins, Recombinant protein expression in *Pichia pastoris*, Mol. Biotechnol. 16 (2000) 23–52.
[38] Y. Nakatani, S.M. Cutfield, N.P. Cowieson, J.F. Cutfield, Structure and activity of exo-1,3/1,4-β-glucanase from marine bacterium *Pseudoalteromonas* sp. BB1 showing a novel C-terminal domain, FEBS J. 279 (2012) 464–478.
[39] M.D. Balcewich, K.A. Stubbs, Y. He, T.W. James, G.J. Davies, D.J. Vocadlo, B.L. Mark, Insight into a strategy for attenuating AmpC-mediated β-lactam resistance: Structural basis for selective inhibition of the glycoside hydrolase NagZ, Protein Sci. 18 (2009) 1541–1551.
[40] K.A. Stubbs, M. Balcewich, B.L. Mark, D.J. Vocadlo, Small molecule inhibitors of a glycoside hydrolase attenuate inducible AmpC-mediated β-lactam resistance, J. Biol. Chem. 282 (2007) 21382–21391.
[41] H. Brumer, P.F.G. Sims, M.L. Sinnott, Lignocellulose degradation by *Phanerochaete chrysosporium*: purification and characterization of the main alpha-galactosidase, Biochem. J. 339 (1999) 43–53.
[42] H.G. Damude, V. Ferro, S.G. Withers, R.A.J. Warren, Substrate specificity of endoglucanase A from *Cellulomonas fimi*: fundamental differences between endoglucanases and exoglucanases from family 6, Biochem. J. 315 (1996) 467–472.
[43] H.D. Ly, S.G. Withers, Mutagenesis of glycosidases, Annu. Rev. Biochem. 68 (1999) 487–522.
[44] J.L. Viladot, E. de Ramon, O. Durany, A. Planas, Probing the mechanism of *Bacillus* 1,3-1,4-β-D-glucan 4-glucanohydrolases by chemical rescue of inactive mutants at catalytically essential residues, Biochemistry 37 (1998) 11332–11342.
[45] T. Bravman, V. Belakhov, D. Solomon, G. Shoham, B. Henrissat, T. Baasov, Y. Shoham, Identification of the catalytic residues in family 52 glycoside hydrolase, a beta-xylosidase from *Geobacillus stearothermophilus* T-6, J. Biol. Chem. 278 (2003) 26742–26749.
[46] Y.-K. Li, J. Chir, S. Tanaka, F.-Y. Chen, Identification of the general acid/base catalyst of a family 3 β-glucosidase from *Flavobacterium meningosepticum*, Biochemistry 41 (2002) 2751–2759.
[47] Q. Wang, D. Trimbur, R. Graham, R.A.J. Warren, S.G. Withers, Identification of the acid/base catalyst in *Agrobacterium faecalis* beta-glucosidase by kinetic analysis of mutants, Biochemistry 34 (1995) 14554–14562.
[48] A.M. MacLeod, D. Tull, K. Rupitz, R.A.J. Warren, S.G. Withers, Mechanistic consequences of mutation of active site carboxylates in a retaining β-1,4-glycanase from *Cellulomonas fimi*, Biochemistry 35 (1996) 13165–13172.
[49] M. Moracci, A. Trincone, G. Perugino, M. Ciaramella, M. Rossi, Restoration of the activity of active-site mutants of the hyperthermophilic β-glycosidase from *Sulfolobus solfataricus*: dependence of the mechanism on the action of external nucleophiles, Biochemistry 37 (1998) 17262–17270.
[50] U. Sirion, H.J. Kim, J.H. Lee, J.W. Seo, B.S. Lee, S.J. Lee, S.J. Oh, D.Y. Chi, An efficient F-18 labeling method for PET study: Huisgen 1,3-dipolar cycloaddition of bioactive substances and F-18-labeled compounds, Tetrahedron Lett. 48 (2007) 3953–3957.
[51] J.E. Gurst, NMR and the structure of D-glucose, J. Chem. Educ. 68 (1991) 1003.
[52] H. Aspeborg, P.M. Coutinho, Y. Wang, H. Brumer, B. Henrissat, Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5), BMC Evol. Biol. 12 (2012) 186.
[53] V. Lombard, T. Bernard, C. Rancurel, H. Brumer, P.M. Coutinho, B. Henrissat, A hierarchical classification of polysaccharide lyases for glycogenomics, Biochem. J. 432 (2010) 437–444.
[54] M.R. Stam, E.G.J. Danchin, C. Rancurel, P.M. Coutinho, B. Henrissat, Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins, Protein Eng. Des. Sel. 19 (2006) 555–562.