Journal of Structural Biology 173 (2011) 46–56

Contents lists available at ScienceDirect

Journal of Structural Biology

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

Structural and functional analysis of three β-glucosidases from bacterium *Clostridium cellulovorans*, fungus *Trichoderma reesei* and termite *Neotermes koshunensis*

Wen-Yih Jeng [a,c], Nai-Chen Wang [a,c], Man-Hua Lin [a,c], Cheng-Tse Lin [a], Yen-Chywan Liaw [b], Wei-Jung Chang [a], Chia-I Liu [a,c,d], Po-Huang Liang [a,d], Andrew H.-J. Wang [a,c,d,*]

[a] Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan
[b] Institute of Molecular Biology, Academia Sinica, Taipei 115, Taiwan
[c] Core Facility for Protein Production and X-ray Structural Analysis, Academia Sinica, Taipei 115, Taiwan
[d] Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan

ARTICLE INFO

Article history:
Received 7 May 2010
Received in revised form 17 July 2010
Accepted 29 July 2010
Available online 1 August 2010

Keywords:
Cellulases
Enzyme kinetics
Glycosyl hydrolase
Manganese enhancement
Tris inhibition

ABSTRACT

β-Glucosidases (EC 3.2.1.21) cleave β-glucosidic linkages in disaccharide or glucose-substituted molecules and play important roles in fundamental biological processes. β-Glucosidases have been widely used in agricultural, biotechnological, industrial and medical applications. In this study, a high yield expression (70–250 mg/l) in *Escherichia coli* of the three functional β-glucosidase genes was obtained from the bacterium *Clostridium cellulovorans* (*Cc*BglA), the fungus *Trichoderma reesei* (*Tr*Bgl2), and the termite *Neotermes koshunensis* (*Nk*Bgl) with the crystal structures of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl, determined at 1.9 Å, 1.63 Å and 1.34 Å resolution, respectively. The overall structures of these enzymes are similar to those belonging to the β-retaining glycosyl hydrolase family 1, which have a classical $(\alpha/\beta)_8$-TIM barrel fold. Each contains a slot-like active site cleft and a more variable outer opening, related to its function in processing different lengths of β-1,4-linked glucose derivatives. The two essential glutamate residues for hydrolysis are spatially conserved in the active site. In both *Tr*Bgl2 and *Nk*Bgl structures, a Tris molecule was found to bind at the active site, explaining the slight inhibition of hydrolase activity observed in Tris buffer. Manganese ions at 10 mM exerted an approximate 2-fold enzyme activity enhancement of all three β-glucosidases, with *Cc*BglA catalyzing the most efficiently in hydrolysis reaction and tolerating Tris as well as some metal inhibition. In summary, our results for the structural and functional properties of these three β-glucosidases from various biological sources open important avenues of exploration for further practical applications.

1. Introduction

D-Glucose is widely utilized in the food industry and is also employed as a precursor in the production of ethanol and other chemicals. Sugar cane, starch and cellulose are the primary sources for glucose production. As such, it requires the exploitation of large tracts of agricultural land to satisfy the global demand for these products. In order to avoid this growing infringement on the amount of land available for food production, the use of agricultural and forestry wastes have been investigated as a preferable alternative to producing cellulosic ethanol (Service, 2007). The β-1,4 linked glucose polymer, cellulose, can be enzymatically hydrolyzed into glucose by three types of cellulase: endoglucanase (EC 3.2.1.4), which randomly digests the crystalline cellulose, exoglucanase or cellobiohydrolase (EC 3.2.1.91), which cleaves cellobiose off from the non-reducing ends of the cellulose polymer chain, and β-glucosidase (EC 3.2.1.21), which breaks the β-1,4 glucosidic bonds of cellobiose to produce glucose monomers (Sticklen, 2008).

Enzymatic hydrolysis of cellulose is a major limitation in lignocellulose conversion to glucose (Lynd et al., 2002; Zhang and Lynd, 2004). During enzymatic hydrolysis, insufficient β-glucosidase activity causes an accumulation of cellobiose with the consequence

Abbreviations: Bis-Tris, bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane; *Cc*BglA, *Clostridium cellulovorans* β-glucosidase A; EDTA, ethylenediaminetetraacetic acid; Hepes, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; *Nk*Bgl, *Neotermes koshunensis* β-glucosidase; PCR, polymerase chain reaction; PEG, polyethylene glycol; *p*NPG, *p*-nitrophenyl-β-D-glucopyranoside; RMSD, root-mean-square deviation; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; SSCF, simultaneous saccharification and co-fermentation; SSF, simultaneous saccharification and fermentation; *Tr*Bgl2, *Trichoderma reesei* β-glucosidase 2; Tris, tris(hydroxymethyl)aminomethane.

* Corresponding author. Address: Institute of Biological Chemistry, Academia, Sinica, 128 Academia Rd., Sec 2, Taipei 115, Taiwan. Fax: +886 2 2788 2043.
E-mail address: ahjwang@gate.sinica.edu.tw (A.H.-J. Wang).

1047-8477/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2010.07.008

that both endoglucanase and exoglucanase activities are severely inhibited (Bommarius et al., 2008). In the overall degradation route of cellulose, glucose accumulation reduces the efficiency of enzymatic saccharification. The conversion processes from lignocellulose to bioethanol, such as simultaneous saccharification and fermentation (SSF) or simultaneous saccharification and co-fermentation (SSCF) present possible strategies for reducing the product inhibition of glucose (Lynd et al., 2002).

β-Glucosidases, a heterogeneous group of exo-type glycosyl hydrolases, cleave β-glucosidic linkages in disaccharide or glucose-substituted molecules. β-Glucosidases occur in all kingdoms of life and play fundamental biological roles in processes such as degradation of cellulose and other carbohydrates for nutrient uptake (Beguin, 1990; Bhatia et al., 2002) and developmental regulation or chemical defense against pathogen attack (Brzobohaty et al., 1993; Leah et al., 1995; Zagrobelny et al., 2008). According to the classification of Carbohydrate-Active enZymes (CAZy) (http://www.cazy.org) (Cantarel et al., 2009; Henrissat, 1991), β-glucosidases have been classified into glycosyl hydrolase families 1 and 3. In a structural view, glycosyl hydrolase family 1 enzymes fold into an $(\alpha/\beta)_8$ barrel structure and act as a β-retaining mechanism using Glu as the catalytic nucleophile (Wang et al., 1995); whereas glycosyl hydrolase family 3 β-glucosidases, with only a few structures solved, exploit an Asp residue in nucleophile attacks.

In this study, we report the crystallization and X-ray analysis of three β-glucosidases: *Cc*BglA from *Clostridium cellulovorans* (anaerobic bacterium), *Tr*Bgl2 from *Trichoderma reesei* (Ascomycota fungus), and *Nk*Bgl from *Neotermes koshunensis* (termite), all highly expressed in *Escherichia coli* using the pET system (Novagen). A Tris molecule in the active site of both *Tr*Bgl2 and *Nk*Bgl was located, and this may be responsible for the inhibition of the hydrolase activity. Interestingly, Mn^{2+} in the 2–10 mM concentration, significantly enhanced the hydrolase activity of all three β-glucosidases, *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl at a rate of 150% to 230%. According to our results, *Cc*BglA exhibits the highest potential for use in the process of lignocellulosics ethanol production, especially for the processes of SSF and SSCF due to its high enzyme activity at a moderate reaction temperature near neutral pH.

2. Materials and methods

2.1. Enzymes, chemicals and bacterial strains

E. coli strains used in this study were DH5α, BL-21 (DE3) (Novagen), and XL1-Blue (Stratagene) cells. The enzymes for DNA manipulation were purchased from New England Biolabs. Novozyme 188, salicin, glucose, 3,5-dinitrosalicylic acid, *p*-nitrophenyl-β-D-glucopyranoside (*p*NPG), resorufin-β-D-glucopyranoside and other major chemicals were purchased from Sigma–Aldrich. The β-glucosidase of Novozyme 188 was further fractionated by size exclusion chromatography using a Sephacryl S-200 column equilibrated with 100 mM NaCl in 50 mM Hepes buffer pH 8.0 (Boisset et al., 2000). The active fractions were pooled and concentrated using 30 kDa cut-off size membranes of Amicon-Ultra-15 (Millipore).

2.2. Gene cloning

Genes from *C. cellulovorans* (*CcbglA*; NCBI accession number AY268940), *T. reesei* (*Trbgl2*; AB003110) and *N. koshunensis* (*Nkbgl*; AB073638) were synthesized and optimized to the favored codon usage for *E. coli* by the Geneart company (Regensburg, Germany). The *Cc*BglA gene fragment was amplified by polymerase chain reaction (PCR) with forward 5′ GGAATTCCATATGGAAAAACTG CGCTTTCCG 3′ (NdeI site underlined) and reverse 5′ CCGCTC GAGTTTGTTGCTGCGTTCGATCAG 3′ (XhoI site underlined) primers. The *Tr*Bgl2 gene fragment was amplified by PCR with forward 5′ GGAATTCCATATG*CATCACCATCATCATCAC*ATGTTGCCCAAGGACTTT CAG 3′ (NdeI site underlined) and reverse 5′ CGGAATTC **TCA**CGCCGCCGCAATCAGCTC 3′ (EcoRI site underlined) primers. The *Nk*Bgl gene fragment was amplified by PCR with forward 5′ GGAATTCCATATGGATGTTGCTTCCTCCGACAC 3′ (NdeI site underlined) and reverse 5′ CCGCTCGAGGTCTCTGAATCTCTCTGGGATC 3′ (XhoI site underlined) primers. Each PCR product of the three genes was subcloned into the expression vector of pET-21a (Novagen) to generate pET-21-*CcbglA*, pET-21-*Trbgl2* and pET-21-*Nkbgl*. The N-terminal 20 residues of authentic *Nk*Bgl were predicted to be an N-terminal signal peptide and were obviated in the construction of pET-21-*Nkbgl*. For easier purification, a His_6 extension was added on the N-terminus of *Tr*Bgl2 and the C-termini of *Cc*BglA and *Nk*Bgl. The inactive E193D mutant of *Nk*Bgl was generated by using the PCR-based QuikChange site-directed mutagenesis kit (Stratagene) using pET-21-*Nkbgl* as the template.

2.3. Protein expression and purification

All DNA constructs were confirmed by nucleotide sequencing. The correct constructs were transformed into *E. coli* BL-21 (DE3) competent cells for protein expression. An overnight culture of a single transformant (100 ml) was used to inoculate six liters of fresh Luria–Bertani (LB) medium containing 100 μg/ml ampicillin, and was subsequently incubated at 30 °C with shaking until the medium reached an $A_{600\ nm}$ of 0.8–1.0. The temperature was then reduced to 13 °C (for *Cc*BglA and *Tr*Bgl2) or 16 °C (for *Nk*Bgl), and the bacteria were allowed to cool for one hour before adding the inducer, isopropyl β-thiogalactopyranoside (IPTG), to a final concentration of 0.2 mM (for *Cc*BglA), 0.5 mM (for *Tr*Bgl2) or 0.4 mM (for *Nk*Bgl). After further growth for 16 h, the bacterial cells were harvested by centrifugation at 7000*g* and immediately resuspended in a lysis buffer containing 20 mM Tris–HCl, 400 mM NaCl, 10 mM imidazole, pH 7.5. The cells were then destroyed by using Cell Disruption System (Constant Systems Ltd.) and cell debris was removed by centrifugation at 17,000*g*. The cell-free extract was loaded onto a Ni^{2+}-nitrilotriacetic acid column, which had been equilibrated with the lysis buffer. The column was washed using the same buffer, and the His_6-tagged proteins were subsequently eluted by a linear gradient of 10–300 mM imidazole. After checking by SDS–PAGE, the purified His_6-tagged *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl proteins were concentrated and changed to enzyme storage buffer (50 mM Tris–HCl, 100 mM NaCl, pH 8.0) using 30 kDa cut-off size membranes with Jumbosep (Pall) and Amicon-Ultra-15 (Millipore). As the Tris molecule was found to occupy each active site of *Tr*Bgl2 and *Nk*Bgl in our crystal structures and acted as an inhibitor, the Tris–HCl was subsequently replaced by Hepes buffer for enzyme purification and storage. The expression and purification of *Nk*Bgl E193D mutant was performed using the same protocol established for native *Nk*Bgl. About 70 mg/l of *Cc*BglA, 240 mg/l of *Tr*Bgl2 and 80 mg/l of *Nk*Bgl were obtained, with a purity of greater than 98% as determined by SDS–PAGE.

2.4. Optimum pH and temperature determination of enzyme activity

The *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl were prepared and stored by using Hepes buffer in the production procedure to avoid Tris inhibition in enzyme assay. Enzyme activity of β-glucosidase toward salicin was measured according to the method of Miller (Miller, 1959) using glucose as the standard. A total 200 μl reaction mixture containing 0.5% salicin and enzyme (10 μl enzyme stored and diluted in 50 mM Hepes and 100 mM NaCl, pH 8.0) in 100 mM phosphate-citrate buffer under different pH conditions (3.0–7.0) was incubated at temperatures from 30 to 70 °C for 30 min. The 3,5-dinitrosalicylic acid reagent (200 μl) was added

and the solution was incubated at 95 °C for 5 min. Finally, 600 μl distilled water was added and the absorbance was measured at 540 nm by Lambda Bio 40 spectrometer (Perkin–Elmer). One unit of enzyme activity corresponds to one μmol of glucose per minute in the reaction. The commercial Novozyme 188 is a crude protein mixture produced by *Aspergillus niger*. In addition to β-glucosidase, it contains another major protein which appears to be α-amylase with a molecular weight of about 65 kDa. The size of the enzyme was determined by zymogram and protein identification using mass spectrometry. Protein concentrations were determined as described by Bradford (Bradford, 1976) using the Bio-Rad Protein Assay Kit with BSA as the standard. All experiments were performed in triplicate using a single prepped protein.

2.5. Effect of chemical reagents and metal ions on β-glucosidase activity

The effect of chemicals and metal ions on each β-glucosidase activity was assayed in 100 mM phosphate–citrate buffer using 0.5% salicin as the substrate at its optimal pH and temperature for 30 min. Various chemicals, such as Tris, EDTA, dithiothreitol (DTT), and 2-mercaptoethanol (2-ME), and metal ions such as $CaCl_2$, $CoCl_2$, $CuSO_4$, $MgCl_2$, $MnCl_2$, NaN_3, and $ZnSO_4$ were added to the reaction mixture to determine their effect at different concentrations. To avoid changing the pH values of the reaction mixture, the stock solution of Tris was adjusted to the optimal assay pH, i.e. pH 5.5 for *Nk*Bgl and pH 6.0 for both *Cc*BglA and *Tr*Bgl2. The relative activities were defined using the activity ratio to each β-glucosidase activity without chemicals added. All experiments were performed in triplicate using a single prepped protein.

2.6. Enzyme kinetic determination

Two substrates, *p*NPG and resorufin-β-D-glucopyranoside, were used for kinetic analysis for *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl. Both reactions were performed in 100 mM phosphate–citrate buffer in pH 5.5 or 6.0, depending on the optimal condition of each β-glucosidase. The former substrate was assayed at 40 or 45 °C using a temperature controller system (Jasco EHC-716), while the latter substrate was investigated at 25 °C. Reactions were carried out using 0.01–15 mM *p*NPG in a total volume of 2 ml. The reaction was initiated by adding 20 μl of *Cc*BglA, *Tr*Bgl2 or *Nk*Bgl to give a final concentration of 10–100 nM of enzyme. The released *p*-nitrophenol was monitored continuously at 410 nm by Jasco V-630 Bio spectrophotometer. Reactions were carried out using a final concentration of 0.1–150 μM resorufin-β-D-glucopyranoside in a total volume of 1 ml. The reaction was initiated by the addition of 10 μl of *Cc*BglA, *Tr*Bgl2 or *Nk*Bgl to give a final concentration of 0.1–4 nM of enzyme. The resorufin release was monitored continuously at excitation and emission wavelengths of 535 nm and 595 nm, respectively by Fluorolog-3 spectrofluorometer (Horiba). K_m and k_{cat} values were calculated by fitting the data to the Michaelis–Menten equation using GraphPad Prism 5.0.

2.7. Crystallization

All proteins were crystallized by the sitting-drop vapor diffusion method at 25 °C, with each sitting-drop prepared by mixing 2 μl each of protein solution and the reservoir solution. The CombiClover Junior sitting plate (Emerald BioSystems) and 250 μl each well of reservoir solution were used in the crystallization. The concentration of both *Cc*BglA and *Tr*Bgl2 is 10 mg/ml in 50 mM Tris–HCl buffer (pH 8.0) containing 100 mM NaCl and 5 mM dithiothreitol, whereas *Nk*Bgl is 25 mg/ml. The crystals of *Cc*BglA were obtained with a reservoir solution containing 21–23% (w/v) PEG 3350 (J.T. Baker) and 0.3–0.45 M Li_2SO_4 (J.T. Baker) in 0.1 M Hepes (J.T. Baker) buffer at pH 7.5. The crystals of *Tr*Bgl2 were obtained with a reservoir solution containing 20–22% (w/v) PEG 3350 and 0.23 M $MgCl_2$ (Fluka) and 12.5 mM sodium formate (Fluka) in 0.1 M Tris buffer at pH 8.0. The crystals of *Nk*Bgl were obtained with a reservoir solution containing 18–21% (w/v) PEG 3350 and 0.1–0.25 M $MgCl_2$ in 0.1 M Bis-Tris (Fluka) buffer at pH 6.5. Crystals of *Nk*Bgl-E193D mutant were obtained at similar conditions to that of native *Nk*Bgl though the Tris–HCl buffer was replaced by Hepes buffer.

2.8. Diffraction data collection and processing

All crystals were flash-cooled to 100 K in a stream of cold nitrogen prior to data collection. Prior to flash cooling, the crystals were soaked briefly in a cryoprotectant solution. The cryoprotectant solution for *Cc*BglA crystals contains 38% (w/v) PEG 3350 and 0.45 M Li_2SO_4 in 0.1 M Hepes buffer at pH 7.5, while for *Tr*Bgl2 and *Nk*Bgl crystals it contains 8% (w/v) glycerol, 23% (w/v) PEG 3350, 0.12 M NaCl and 0.25 M $MgCl_2$. The data of *Nk*Bgl-E193D-*p*NPG complex was collected from the inactivate *Nk*Bgl-E193D crystal after soaking for one minute with the cryoprotectant solution of *Nk*Bgl containing saturated *p*NPG. X-ray diffraction data were collected at SPXF beamlines BL13B1 and BL13C1 at the National Synchrotron Radiation Research Center (NSRRC) in Taiwan and beamline BL5A at the Photon Factory (PF) in Japan. The diffraction images were processed using the program package of HKL2000 (Otwinowski and Minor, 1997). The crystal data and data collection statistics of *Cc*BglA, *Tr*Bgl2, *Nk*Bgl and *Nk*Bgl-E193D-*p*NPG complex are summarized in Table 1.

2.9. Structure determination and refinement

The structures of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl were solved by molecular replacement with CNS (Brunger et al., 1998) using the structures of the PDB entries of 1OD0 (*Thermotoga maritima* β-glucosidase A), 2E3Z (*Phanerochaete chrysosporium* β-glucosidases 1A) and 1CBG (*Trifolium repens* cyanogenic β-glucosidase) as the search models, respectively. The structural models were built and refined using Xtalview (McRee, 1999), Coot (Emsley and Cowtan, 2004), CNS, and Refmac (Winn et al., 2003). The structure of *Nk*Bgl-E193D-*p*NPG was solved by molecular replacement using CNS with the refined structure of *Nk*Bgl as a search model. A subset of 5% randomly selected reflections was excluded from computational refinement to calculate the R_{free} factor throughout the refinement (Brunger, 1993). All final refinements were carried out using Refmac with TLS group tensor and anisotropic *B* factor without NCS restrains. The *Cc*BglA crystals contained four molecules in an asymmetric unit and the models were refined to a final R_{work} of 14.6% and R_{free} of 19.8%. The *Tr*Bgl2 crystals contained four molecules in an asymmetric unit and the models were refined to a final R_{work} of 16.4% and R_{free} of 21.7%. Both *Nk*Bgl and *Nk*Bgl-E193D-*p*NPG crystals contained one molecule in an asymmetric unit. The *Nk*Bgl model was refined to a final R_{work} of 11.5% and R_{free} of 14.4%, while a R_{work} of 12.2% and R_{free} of 15.3% was obtained for *Nk*Bgl-E193D-*p*NPG. The stereochemistry and structure of the final models were analyzed by RAMPAGE (Lovell et al., 2003) and SFCHECK (Vaguine et al., 1999) of the CCP4 program suite (Bailey, 1994). The refinement statistics are summarized in Table 1. The molecular figures were produced by using PyMOL (DeLano Scientific, http://www.pymol.org).

2.10. PDB accession codes

The atomic coordinates and structure factors for β-glucosidases from *C. cellulovorans*, *T. reesei*, and *N. koshunensis* have been deposited in the Protein Data Bank (PDB accession code: 3AHX, 3AHY, 3AHZ). The *Nk*Bgl-E193D in complex with *p*NPG was deposited as 3AI0.

Table 1
Data collection and refinement statistics of β-glucosidases.

Crystals PDB code	*Cc*BglA 3AHX	*Tr*Bgl2 3AHY	*Nk*Bgl 3AHZ	*Nk*Bgl-E193D-*p*NPG 3AI0
Data collection				
Radiation source	NSRRC BL13B1	PF BL5A	NSRRC BL13C1	NSRRC BL13C1
Wavelength (Å)	1.00000	1.00000	0.97315	0.97315
Space group	$P4_12_12$	$P2_1$	$C2$	$C2$
Unit cell parameter				
a (Å)	128.50	92.83	92.95	93.09
b (Å)	128.50	103.51	68.45	68.76
c (Å)	264.06	94.80	75.22	75.75
β(°)		105.37	95.53	95.68
Resolution (Å)	30–1.90 (1.97–1.90)[a]	30–1.63 (1.69–1.63)[a]	30–1.34 (1.39–1.34)[a]	30–1.40 (1.45–1.40)[a]
Number of reflections	171,402 (16,713)	210,766 (20,610)	105,222 (10,431)	93,220 (9108)
Completeness (%)	99.2 (98.0)	96.9 (95.3)	99.9 (99.7)	99.5 (97.4)
Redundancy	6.7 (3.8)	5.6 (5.6)	5.7 (5.5)	4.5 (4.0)
R_{merge} (%)	6.2 (32.3)	8.8 (50.3)	5.4 (46.9)	4.1 (42.8)
I/σ (I)	27.5 (3.8)	20.9 (4.5)	29.7 (3.9)	31.2 (3.0)
Overall Wilson *B* factor (Å^2)	31.0	23.9	16.2	18.6
Refinement[b]				
Reflections (work)	151,131 (19,522)	192,806 (26,234)	95,986 (13,134)	84,715 (10,986)
Reflections (free)	7980 (1015)	10,155 (1434)	5061 (726)	4491 (575)
R_{work} (%)	14.6 (15.8)	16.4 (16.9)	11.5 (15.6)	12.2 (17.6)
R_{free} (%)	19.8 (23.3)	21.7 (24.4)	14.4 (19.7)	15.3 (22.2)
Geometry deviations				
Bond length (Å)	0.006	0.005	0.006	0.007
Bond angles (°)	1.4	1.4	1.5	1.4
Mean *B*-values (Å^2)/No.				
Protein atoms	18.4/14,530	12.1/14,791	8.9/3811	10.3/3810
Water molecules	30.6/2147	26.7/3100	25.4/806	26.5/781
Other compound atoms	32.6/42	19.2/32	17.0/14	13.7/27
Ramachandran plot (%)[c]				
Favored	98.2	96.9	97.9	97.9
Allowed	1.8	3.1	2.1	2.1

[a] Values in the parentheses are for the highest resolution shells.
[b] All positive reflections are used in the refinements.
[c] Categories were defined by RAMPAGE.

3. Results

3.1. Enzyme activities

The effect of pH and temperature on the β-glucosidases' activity toward salicin substrate was characterized. Both *Cc*BglA and *Tr*Bgl2 were optimally active at pH 6, while *Nk*Bgl was optimal at pH 5.5 (Fig. 1). The optimum pH for Novozyme 188, a commercial β-glucosidase enzyme, is pH 4, at which condition our enzymes show less activity. However, at their optimum pH, *Tr*Bgl2, *Nk*Bgl, and *Cc*BglA show at least 2-, 5-, and 10-fold increases in activity in comparison to the commercial enzyme Novozyme 188 in equal protein quantity by weight. The temperature for maximum activity was found to be 40 °C for *Tr*Bgl2 and 45 °C for *Cc*BglA and *Nk*Bgl at

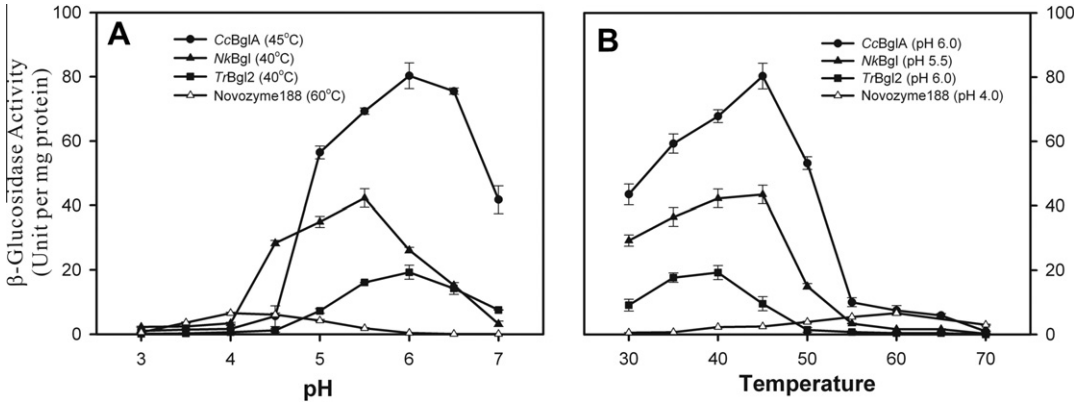

Fig. 1. Effect of pH and temperature on the activity of β-glucosidases. The optimal conditions of (A) pH and (B) temperature affected profiles of β-glucosidases were determined by salicin as substrate. The optimal temperature of *Cc*BglA (45 °C), *Tr*Bgl2 (40 °C), *Nk*Bgl (40 °C) and Novozyme 188 (60 °C) were used to determine the pH effect in (A). The optimal pH of *Cc*BglA (pH 6.0), *Tr*Bgl2 (pH 6.0), *Nk*Bgl (pH 5.5) and Novozyme 188 (pH 4.0) were used to determine the temperature effect in (B). The enzyme activities were measured in 0.5% salicin in 100 mM phosphate–citrate buffer at pH 3.0–7.0 and 30–70 °C for 30 min. One unit of β-glucosidase activity is defined as 1 μmol of glucose produced from salicin per minute. The functional protein fraction separated by Sephacryl S-200 size exclusion chromatography of Novozyme 188 was used in the assay. The *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl were prepared and stored by using Hepes buffer to avoid the inhibited effect of Tris in the assay. All measurements were analyzed with triplicates with error bars indicated. Error bars represent the standard deviation of three replicates from a single prepped protein.

their optimum pH. At temperatures higher than the optimum, *Cc*BglA activity declines drastically and shows similar activity to Novozyme 188 at 60 °C.

To determine the affinity of β-glucosidases for β-1,4 linked glucose derivatives, the kinetics properties with *p*NPG and resorufin-β-D-glucopyranoside were investigated. All three of these β-glucosidases show higher affinity toward the latter substrate, especially *Tr*Bgl2, which is 121-fold higher toward resorufin-β-D-glucopyranoside than *p*NPG (Table 2). In every situation examined, the bacterial *Cc*BglA demonstrated the highest catalytic efficiency among the three species.

3.2. Effect of chemical reagents and metal ions

To investigate the effect of metal ions and reducing agents on enzyme activity, β-glucosidases were assayed in the presence of 2, 5, or 10 mM of metal ions or chemicals as shown in Fig. 2. Among the metal ions, Mn^{2+} significantly enhanced the β-glucosidase activity by 150–230% in all three tested concentrations. As to the Cu^{2+} ion, it notably reduced up to 90% of *Tr*Bgl2 and 80% of *Nk*Bgl activity at 10 mM Cu^{2+}, although the same concentration was tolerable for *Cc*BglA where 60% activity remained. A higher concentration of Zn^{2+} (10 mM) severely inhibited more than 90% of *Cc*BglA and *Tr*Bgl2 activities; however, no obvious inhibition was observed in *Nk*Bgl. The chelating agent EDTA exerts 11–41% inhibition on enzyme activity, suggesting the presence of some trace elements in the original purified enzymes or reaction buffer in addition to those that we studied, i.e., divalent cations other than Ca^{2+}, Mg^{2+}, etc., which contributed to enzyme activation. In the presence of reducing agent, dithiothreitol had no effect on β-glucosidase activity; however, 2-mercaptoethanol resulted in 30–50% inhibition of β-glucosidase activity. With regard to the metals and chemicals used in this study, *Tr*Bgl2 is more affected by their presence than *Cc*BglA and *Nk*Bgl.

In both the *Tr*Bgl2 and *Nk*Bgl crystal structures, a Tris molecule was observed in the active site. In order to gain an insight into the effect of Tris on β-glucosidases, different concentrations of Tris molecules were applied. At all concentrations tested, Tris inhibits the three β-glucosidases as shown in Fig. 2. Severe inhibition was detected both in *Tr*Bgl2 and *Nk*Bgl when the Tris concentration was higher than 20 mM.

3.3. Overall structures of CcBglA, TrBgl2 and NkBgl

The primary sequences of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl, belonging to the glycoside hydrolase family 1, are composed of 445, 465 and 498 residues, respectively (Supplementary Fig. S1). The N-terminal signal peptide of 20 residues of *Nk*Bgl is obviated in the recombinant protein. The overall structures of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl are similar to the PDB-deposited structures of glycosyl hydrolase family 1 enzymes, which have a classical $(\alpha/\beta)_8$-TIM barrel fold (Fig. 3A–C). Amino acid sequence alignment of nine glycosyl hydrolase family 1 enzymes and the secondary structure of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl are shown in Supplementary Fig. S1.

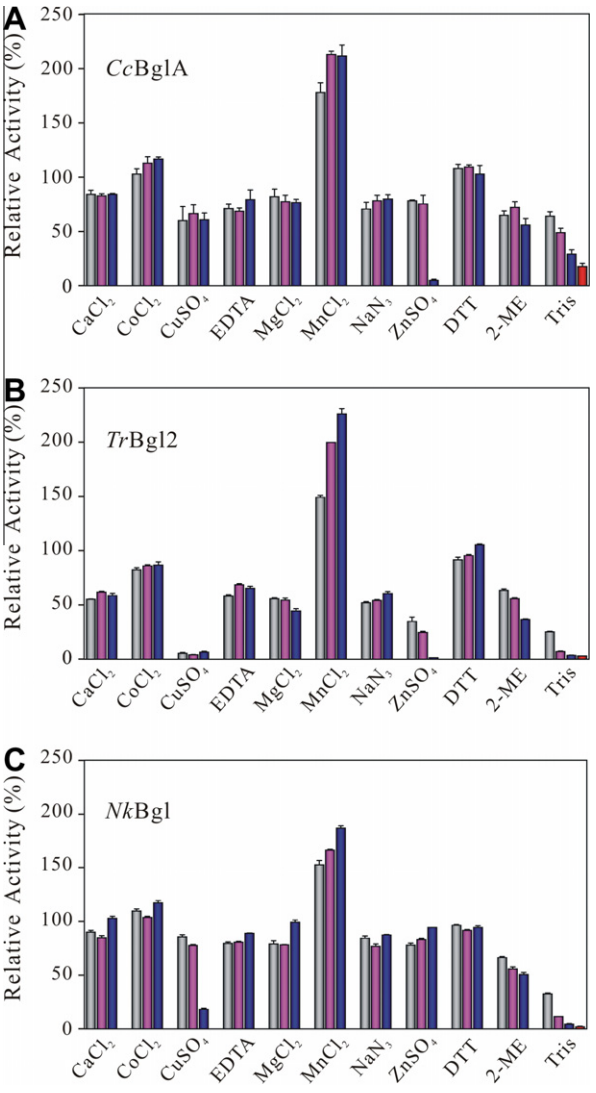

Fig. 2. Effect of chemical reagents and metal ions on the β-glucosidase activity. Concentrations of various chemicals (except Tris) ranged from 2 mM (gray), 5 mM (magenta) to 10 mM (blue). Concentrations of Tris ranged from 5 mM (gray), 20 mM (magenta), 50 mM (blue) to 100 mM (red). The relative activities were defined using the activity ratio to each β-glucosidase activity without chemicals added. The optimal pH and temperature of each β-glucosidase were used in the assay. All measurements were analyzed with triplicates with error bars indicated. Error bars represent the standard deviation of three replicates from a single prepped protein.

The crystal structure of *Cc*BglA was determined at 1.9 Å. There are four protein molecules and two PEG fragments in an asymmetric unit. The structures of four chains of *Cc*BglA with root-mean-square deviations (RMSDs) ranged from 0.23 to 0.30 Å over 441–444 Cα atoms. The two PEG fragments were located on the packing surface of chains A–C and B–D.

Table 2
Catalytic activity of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl.

Substrate	*p*-Nitrophenyl-β-D-glucopyranoside			Resorufin-β-D-glucopyranoside		
Protein	Parameters k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($s^{-1}\cdot mM^{-1}$)	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($s^{-1}\cdot mM^{-1}$)
*Cc*BglA	50.67 ± 1.00	0.15 ± 0.01	340 ± 27	97.45 ± 4.17	0.0087 ± 0.0012	11,418 ± 1906
*Tr*Bgl2	6.91 ± 0.16	0.86 ± 0.07	8.1 ± 0.8	31.38 ± 0.85	0.0071 ± 0.0008	4477 ± 599
*Nk*Bgl	1.39 ± 0.03	0.29 ± 0.02	4.8 ± 0.4	9.01 ± 0.30	0.0116 ± 0.0014	788 ± 114

Values are shown with standard deviations calculated from three replicates using a single prepped protein.

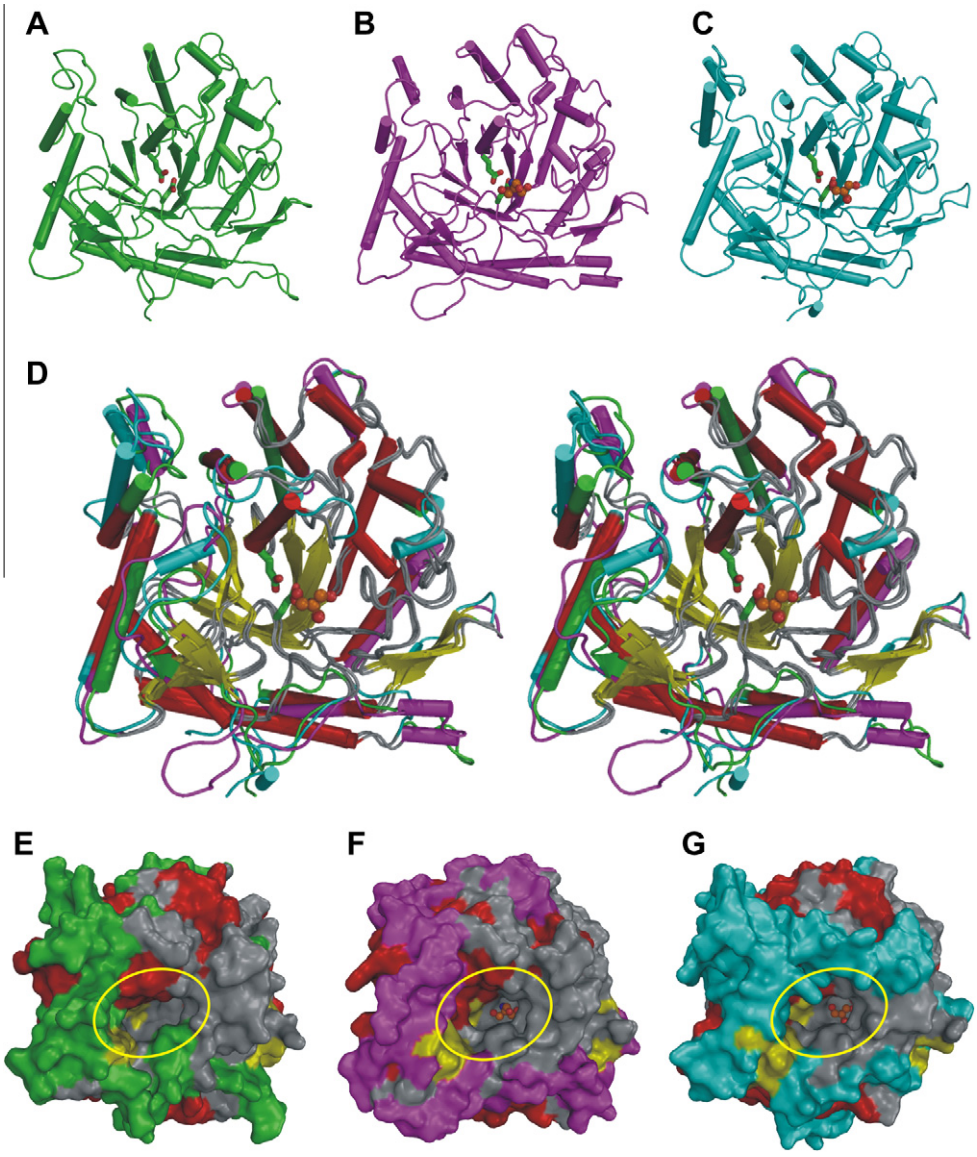

Fig. 3. Comparison of the structures of β-glucosidases from bacteria, fungus and termite. Ribbon representation of the structures of (A) *Cc*BglA, (B) *Tr*Bgl2, (C) *Nk*Bgl and (D) stereoview of the three superimposed β-glucosidases. The overall structures of β-glucosidases adopt the typical TIM barrel fold. The highly superimposed helices are shown in red, whereas strands in yellow, loops in gray. Variant regions of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl different from the three superimposed structures are shown in green, magenta and cyan, respectively. The two catalytic glutamates of *Nk*Bgl and the Tris molecule located in the active site of *Nk*Bgl are shown as stick and ball models. Molecular surface representation of (E) substrate-free *Cc*BglA, (F) *Tr*Bgl2 and (G) *Nk*Bgl in complex with Tris moleclules. The view and color painting in (E–G) are the same as in (D). The active sites of (E–G) are indicated by the yellow circles. For displayed glutamates and Tris molecules, oxygen atoms are in red, nitrogen atoms are in blue, and carbon atoms are in green (for glutamates) or orange (for Tris).

The crystal structure of *Tr*Bgl2 was determined at 1.63 Å and contained four protein and Tris molecules in the asymmetric unit. The structures of four chains of *Tr*Bgl2 with RMSDs ranged from 0.23 to 0.59 Å over 464–465 Cα atoms.

The crystal structures of *Nk*Bgl were determined at 1.34 Å and contained one protein, Tris, and glycerol molecule in the asymmetric unit. The crystal structure of *Nk*Bgl-E193D-*p*NPG was determined at 1.40 Å and contained one protein, *p*NPG, and glycerol molecule in the asymmetric unit. The RMSD was 0.15 Å over 472 Cα atoms of structures of *Nk*Bgl and *Nk*Bgl-E193D-*p*NPG.

The accessible surface areas of the monomer to solvent of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl are about 17,650, 17,000 and 17,300 $Å^2$, respectively. Although four protein molecules of *Cc*BglA and *Tr*Bgl2 were present in the asymmetric unit, they did not reveal any specific interactions that could result in the formation of stable quaternary structures by the Protein Interfaces, Surfaces, and Assemblies (PISA) server (Krissinel and Henrick, 2007). Interestingly, one protein molecule of *Nk*Bgl was present in the asymmetric unit, but two monomers interacted to form a dimer through crystallographic packing. However, *Nk*Bgl exists as a monomeric form in solution as determined by gel filtration chromatography (data not shown). The results suggest that *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl likely function as a monomer.

The active sites of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl formed a 15–20 Å deep slot-like cleft and were located on connecting loops at the C-terminal end of the β-sheets of the TIM barrel. The bottom of each active site is surrounded by negatively charged residues (Supplementary Fig. S2A–C). The catalytic acid/base of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl are Glu166, Glu165 and Glu193, respectively, located on the TXNEP motif at the end of β-strand 4 (where X is a hydrophobic amino acid residue). The catalytic nucleophile of *Cc*BglA is Glu352, located on the I/VTENG motif at the end of β-strand 7;

whereas for *Tr*Bgl2 it is Glu367, and for *Nk*Bgl it is Glu402 (Supplementary Fig. S1) (Wang et al., 1995). Two *cis*-peptide bonds were found at Ala181-Pro182 and Trp399-Ser400 of *Cc*BglA, at Ala180-Pro181 and Trp417-Ser418 of *Tr*Bgl2, at Ala208-Pro209 and Trp444-Ser445 of *Nk*Bgl. The presence of such bonds is typical for the glycosyl hydrolase family 1 (Seshadri et al., 2009). In addition, one more *cis*–peptide bond was found between Glu350 and Pro351 of *Nk*Bgl.

3.4. Comparison of structures of CcBglA, TrBgl2 and NkBgl

The one chain crystal structure of *Nk*Bgl was compared with the four chain structures of *Cc*BglA and *Tr*Bgl2 with RMSDs ranging from 1.27 to 1.32 Å over 441–444 Cα atoms, and 1.15–1.21 Å over 465–466 Cα atoms, respectively. A comparison of the four chain crystal structures of *Cc*BglA and *Tr*Bgl2 produced RMSDs ranging from 1.17 to 1.34 Å over 441–444 Cα atoms. The ribbon represen-

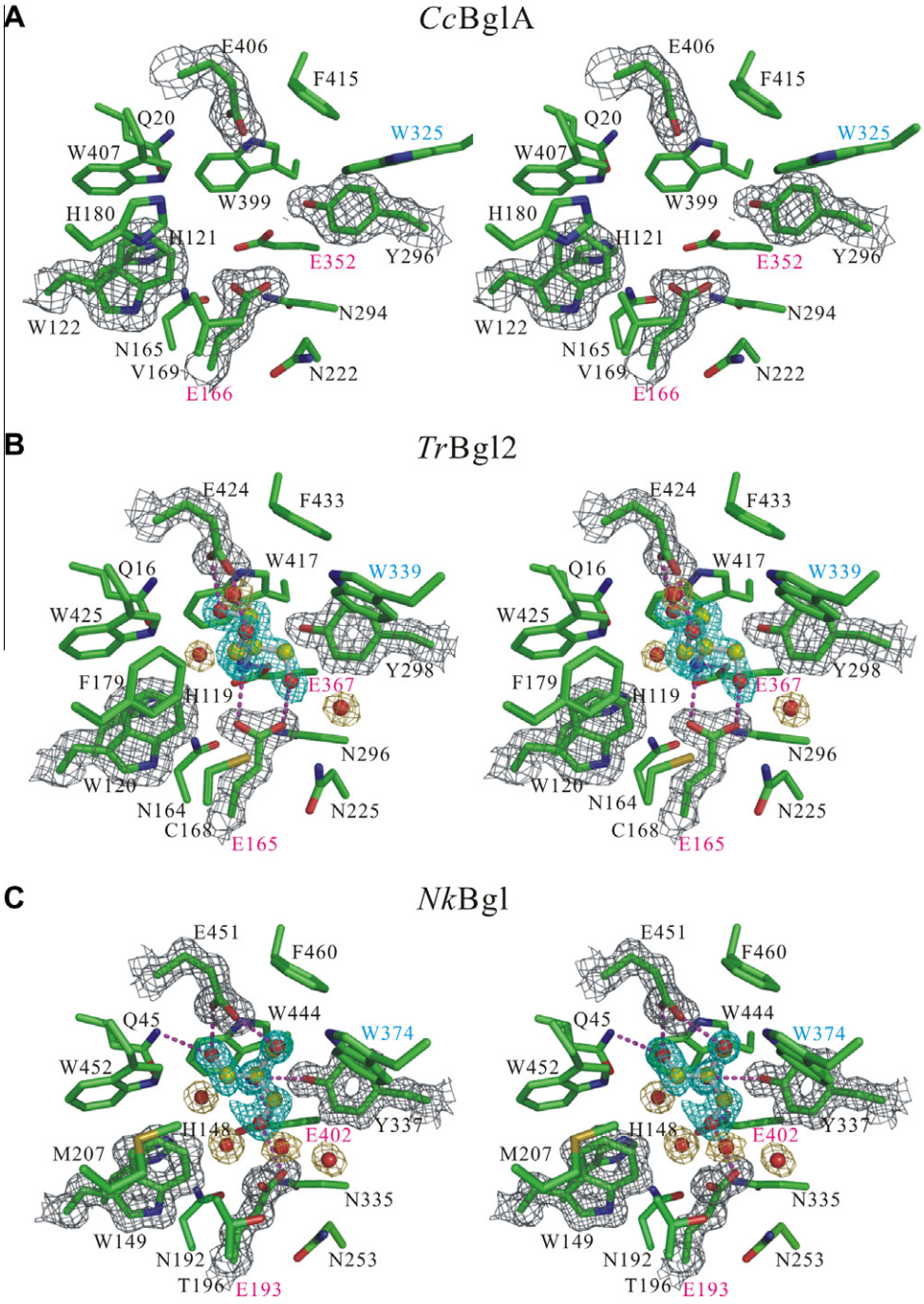

Fig. 4. The active sites of β-glucosidases from bacteria, fungus and termite. Stereoview of the surrounding residues of the active site of (A) substrate-free *Cc*BglA, (B) *Tr*Bgl2 and (C) *Nk*Bgl in complex with Tris moleclules. |Fo| − |Fc| difference Fourier maps for the Tris molecules contoured at 3.0σ are shown in cyan. 2|Fo| − |Fc| maps for the side chains of selective residues are shown in gray, whereas water molecules contoured at 1.0σ are in yellow-orange. The side chains of residues around the active site of the β-glucosidases are shown as stick models. The Tris molecules are shown as ball-and stick models. The glutamates labeled in magenta are the catalytic nucleophile and acid–base residues of these β-glucosidases. Significant variance of the side chain of Trp325 of *Cc*BglA was observed. The hydrogen-bond interactions of the residues and Tris molecules are labeled in magenta dotted lines. The carbon atoms of proteins and Tris molecules are shown in green and yellow, respectively. The oxygen atoms are shown in red, nitrogen in blue and sulfur in gold.

tation of superimposed structures of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl is shown in Fig. 3D. The Cα traces of superimposed structures of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl are shown in Supplementary Fig. S3. The structures of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl share many similarities. The core structures of the TIM barrel forming at the bottom of its active site are especially highly conserved (Fig. 3E–G and Supplementary Fig. S2D–F) where the glycone (subsite − 1) and aglycone (subsite + 1) binding sites are located (Supplementary Fig. S1). However, with the exception of the right lobe, numerous variances do exist in the structures of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl in the outer opening regions around their active sites (Fig. 3E–G).

3.5. The active site of TrBgl2 and NkBgl is occupied by a Tris molecule

In the structures of *Tr*Bgl2 and *Nk*Bgl a Tris molecule was found to bind at each active site. The electron density of the difference Fourier map, calculated including all refined atoms except the Tris, clearly defined all atoms of the Tris molecules (Fig. 4B–C). Although no Tris molecule was observed in the active site of *Cc*BglA, some small, irregular and unidentified extra electron density was observed. The inhibition of Tris on the hydrolase activity of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl were in concert with that observed in the crystal structures. Among residues around the active sites of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl, only the side chain of Trp325 of *Cc*BglA shows a significant variance (Fig. 4A). The Trp339 of *Tr*Bgl2 and Trp374 of *Nk*Bgl have similar chi-1 (∼58°) and chi-2 (∼−91°) angles; however, the Trp325 of *Cc*BglA has different chi-1 (∼−56°) and chi-2 (∼−22°) angles. The location and orientation of Tris molecules and side chains of the three key glutamates (Glu165, Glu367 and Glu424) are highly superimposed in each chain of *Tr*Bgl2 (Supplementary Fig. S4A). Although the side chains of three key glutamates (Glu165/193, Glu367/402 and Glu424/451) of *Tr*Bgl2 and *Nk*Bgl were located at similar positions, the Tris molecules of *Tr*Bgl2 and *Nk*Bgl were bound in different orientations (Supplementary Fig. S4B). A total of three residues of *Tr*Bgl2 formed five direct hydrogen bonds with Tris, and five residues of *Nk*Bgl formed seven hydrogen bonds (Supplementary Fig. S5).

3.6. Analysis of the crystal structure of NkBgl-E193D in complex with the pNPG substrate

The structure of *Nk*Bgl-E193D in complex with the *p*NPG substrate (Fig. 5B and Supplementary Fig. S4C) was determined at 1.40 Å resolution from crystals that were soaked with the cryoprotectant solution containing saturated *p*NPG. The *p*NPG is a chromogenic substrate for β-glucosidases and has the Michaelis–Menten constant (k_m) of 0.29 mM for *Nk*Bgl in 100 mM phosphate-citrate buffer with pH 5.5 and at 40 °C (Table 2). The electron density of the Fourier difference map, calculated by including all refined atoms except *p*NPG, clearly defined all atoms of the *p*NPG molecule (Supplementary Fig. S4C). Comparing the structures of Tris bound *Nk*Bgl with *p*NPG bound *Nk*Bgl-E193D, significant difference in these structures were observed at the side chains of the residues of Glu/Asp193, Asn253, Trp374, Trp444, Asn449 and Glu451. The residues of Asn253 and Trp374 were located within the aglycone binding pocket and the residues of Trp444, Asn449 and Glu451 were located within the glycone binding pocket. In addition, the catalytic acid/base of *Nk*Bgl was mutated from Glu193 to the smaller Asp193 which creates room in the substrate binding pocket and a subsequent rotation of the Asn253 side chain toward the substrate binding pocket was observed (Fig. 5). Furthermore, it seems that the β-carboxyl group of Asp193 possesses the ability to attack the O1 atom of *p*NPG through a water molecule (Supplementary Figs. S4C and S6).

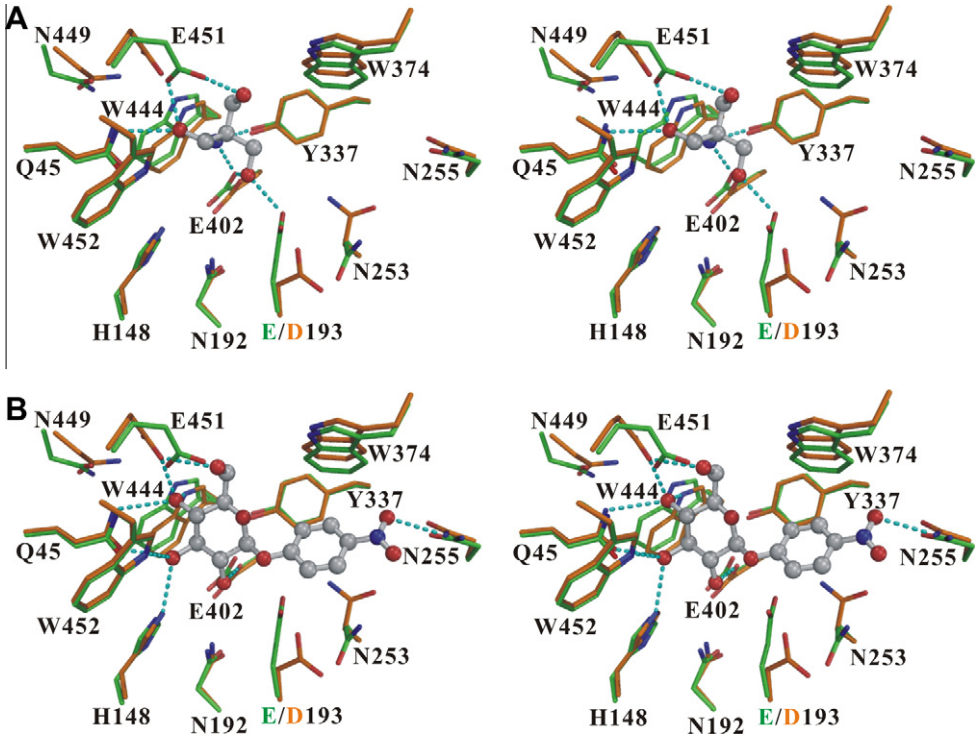

Fig. 5. Comparison of the active site in *Nk*Bgl complexes. Stereoview of the superimposed surrounding residues of the active site of (A) *Nk*Bgl-Tris and (B) *Nk*Bgl-E193D-*p*NPG complexes. The side chains of residues around the active site of *Nk*Bgl are shown as line models. Carbon atoms of residues are shown in green for *Nk*Bgl-Tris and orange for *Nk*Bgl-E193D-*p*NPG. The Tris and *p*NPG molecules are shown as ball-and stick models (carbon atoms in gray). Oxygen atoms are shown in red and nitrogen in blue. Cyan dotted lines indicate the hydrogen-bond interactions between residues and Tris/*p*NPG molecules.

The residues Gln45, His148, Asn255, Glu402, Trp444, Glu451 and Trp452 of *Nk*Bgl-E193D formed direct hydrogen bonds with *p*NPG (Fig. 5B and Supplementary Fig. S6). A total of seven residues of *Nk*Bgl formed nine direct hydrogen bonds with *p*NPG. Residues Asp193, Asn253, Asn255, Asn335, Tyr337, Trp374 and Glu458 of *Nk*Bgl-E193D formed water-mediated hydrogen bonds with *p*NPG and the residues Trp149, Asn253, Tyr337, Trp374, Trp444 and Phe460 of *Nk*Bgl-E193D interacted with *p*NPG through hydrophobic contacts (Supplementary Fig. S6). As mentioned above, the residues Gln45, His148, Trp149, Asn335, Trp444, Glu451, Trp452, Glu458 and Phe460 were located within the glycone binding pocket and the residues of Asn253, Asn255, Tyr337 and Trp374 were located within the aglycone binding pocket (Fig. 6A and Supplementary Fig. S1). Fig. 6B represents a retaining β-glucosidase mechanism in accordance with the structure of the *Nk*Bgl-E193D-*p*NPG complex. Furthermore, the Glu193 residue of *Nk*Bgl, the catalytic acid/base, was mutated into Asp193 to reduce enzyme activity in order to investigate the enzyme structure in complex with *p*NPG. Notably, the binding mode of the hydroxyl groups with Tris was similar to that of the glucoside geometry in the structure of the *Nk*Bgl-E193D-*p*NPG complex.

4. Discussion

C. cellulovarans, one of the most efficient cellulolytic organisms, produces an extracellular non-cellulosomal β-glucosidase A to degrade small oligosaccharides cooperatively with cellulosomes. The C-terminal His_6-tagged recombinant *Cc*BglA has been expressed in *E. coli* BL-21 (DE3) by Kosugi et al. (2006). The recombinant *Cc*BglA in our study shows a higher affinity for the substrates

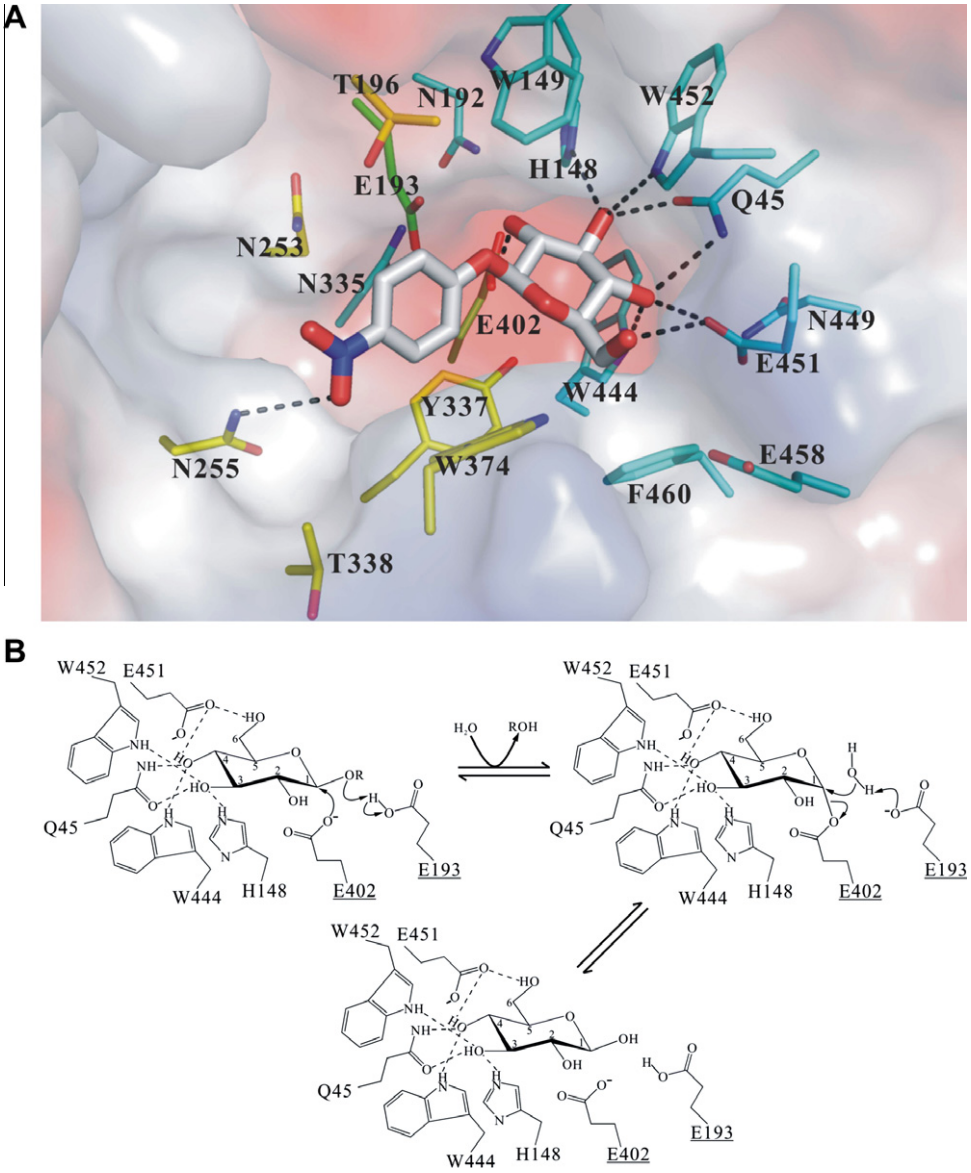

Fig. 6. Interactions of substrate in the active site of *Nk*Bgl. (A) View into the active site, in particular the glycone- and aglycone-binding pockets. Positive charges on the electrostatic surfaces are drawn in blue and negative charges in red. The *p*NPG molecule and the side chains of residues around the active site of *Nk*Bgl are shown as stick models. The carbon atoms of *p*NPG are in white. The carbon atoms of the catalytic residues of *Nk*Bgl are in green, those of the glycone-binding site in cyan, those of the aglycone-binding site in yellow. The hydrogen-bond interactions of the residues and *p*NPG molecule are labeled in black dotted lines. All oxygen atoms are in red and nitrogen in blue. The coordination atoms of Asp193 and Asn253 of *Nk*Bgl-E193D-*p*NPG complex were replaced by the coordination atoms of Glu193 and Asn253 of *Nk*Bgl-Tris complex in this illustration. (B) Scheme of the retaining β-glucosidase mechanism and the residues within the glycone-binding site of *Nk*Bgl. Only residues linked by directed hydrogen bond and catalytic glutamates are shown, and the catalytic glutamates are highlighted. The hydrogen bonds are indicated by dashed lines.

*p*NPG (K_m = 0.15 mM) and resorufin-β-D-glucopyranoside (K_m = 0.0087 mM) than the cellobiose (K_m = 2.9 mM) previously studied by Kosugi et al. (2006).

Some *Trichoderma* species exhibit strong cellulose-degrading activities. The *Tr*Bgl2 gene has been previously isolated from asexual fungus *Trichoderma reesei* (teleomorph *Hypocrea jecorina*), and the recombinant *Tr*Bgl2 has been produced by *Aspergillus oryzae* (Takashima et al., 1999). Using salicin as the substrate instead of the *p*NPG, the recombinant *Tr*Bgl2 demonstrates the same pH and temperature optima during catalysis with the fungus-expressed recombinant *Tr*Bgl2. As for the kinetic parameters, *Tr*Bgl2 expressed in *E. coli* shows approximately 2.5-fold higher affinity to *p*NPG than the one expressed in *A. oryzae*. Interestingly, there is another β-glucosidase II previously purified from *T. reesei* by Chen et al. (1992), which optimally catalyzes the degradation of *p*NPG in more acidic and thermophilic conditions (pH 4, 60 °C) and shows 100-fold higher catalytic efficiency. However, there is no structural information about this extra long β-glucosidase II, which has an amino acid sequence almost twice the length of *Tr*Bgl2.

With regard to β-glucosidases in insects, since β-glycans are key dietary sources for some insects, the substrate specificity of insect β-glucosidases seems to be broad. The recombinant *Nk*Bgl we studied catalyzed the degradation of *p*NPG more efficiently than β-glycosidase from *Spodoptera frugiperda* larval midgut (Marana et al., 2001). The *Nk*Bgl has not only a higher affinity toward the substrate but a 2.5-times higher turnover number. Compared with the native *Nk*Bgl purified from *N. koshunensis* salivary glands, our recombinant *Nk*Bgl exhibits similar pH and temperature optima using salicin as the substrate, and shows a higher affinity toward *p*NPG (K_m = 0.29 mM) and resorufin-β-D-glucopyranoside (K_m = 0.00116 mM) than the cellobiose (K_m = 2.5 mM) previously studied by Tokuda et al. (2002). To our knowledge, *Nk*Bgl is the first insect β-glucosidase structure to have been determined.

Comparing the primary sequences and tertiary structures of *Cc*BglA, *Tr*Bgl2 and *Nk*Bgl with the structures deposited in PDB by Blast (http://blast.ncbi.nlm.nih.gov/) and PDBeFold (also known as SSM) (Krissinel and Henrick, 2004), it is no surprise to find the structure most similar to *Cc*BglA is β-glucosidase A from a bacterial source – *Thermotoga maritima*, which shows 49% identity, 73% similarity, and an RMSD of 0.97 Å over 430 Cα atoms (Zechel et al., 2003). The *Tr*Bgl2 demonstrates most similarity with a fungal β-glucosidase structure from basidiomycete *Phanerochaete chrysosporium*, which has 54% identity, 73% similarity and an RMSD of 0.90 Å over 451 Cα atoms for (Nijikken et al., 2007). The insect *Nk*Bgl shows 45% identity, 66% similarity, and 0.90 Å over 446 Cα atoms for aphid *Brevicoryne brassicae* myrosinase (Husebye et al., 2005). Aphid *B. brassicae* myrosinase (EC 3.2.1.147), hydrolyzing glucosinolates into isothiocyanates, showed considerable similarity in gene sequence to glycosyl hydrolase family 1 members (Husebye et al., 2005). According to the protein sequence alignment and the secondary structure analysis, the linker regions between α4–β5 and βB–βC of bacterial β-glucosidases are shorter than those of the fungal and insect proteins (Supplementary Fig. S1). Compared to bacterial and fungal β-glucosidases, both *Nk*Bgl and myrosinase in insects show shorter linker regions between αF–αG and longer ones for β7–α7 (Supplementary Fig. S1). All the above regions are located on the surfaces of β-glucosidases. These distal regions seem not to affect the active site directly; however, we cannot obviate their indirect effects on enzyme properties among proteins from different species.

The activity of the three β-glucosidases was inhibited by most cations we examined; however, Mn^{2+} significantly enhanced their enzyme activity by at least 50%, which implies that this cation may be involved in enzyme function. In previous studies, strong activation by Mn^{2+} was observed in two ascomycetes strains; 69% toward *p*NPG in *Hanseniaspora vineae* (Vasserot et al., 1989) and 77% in *Aspergillus oryzae* (Riou et al., 1998). Nevertheless, our study demonstrates the highest stimulation, a 226% increase in relative activity, in *Tr*Bgl2-catalyzed degradation of salicin in the presence of 10 mM $MnCl_2$. These results suggest that Mn^{2+} ions might assist the attacking reaction of the catalytic residues (Dismukes, 1996; Riou et al., 1998). However, as the Mn^{2+}-coordinated β-glucosidase structure is yet to have been solved, it is hard to clarify the real role of Mn^{2+} in enzyme function.

In the active sites of the three β-glucosidases, Cys168 of *Tr*Bgl2 is supposed to be more reactive with metal ions than *Cc*BglA Val169 and *Nk*Bgl Thr196, whose residues are located at a similar position as shown in Fig. 4. We speculated that the thiol-group of Cys168 of *Tr*Bgl2 may play an important role in the inhibition of heavy metal ions. Uzun et al. (2008) have reported the following metal ion affinity order of a cysteine functionalized poly(hydroxyethyl methacrylate) monolith: $Cd^{2+} > Cu^{2+} > Hg^{2+} > Pb^{2+} > Zn^{2+} >> Co^{2+}$. This is consistent with our findings that $CuSO_4$ and $ZnSO_4$ caused a drastic inhibition of *Tr*Bgl2, while $CoCl_2$ did not (Fig. 2B). Additionally, 10 mM $ZnSO_4$ caused severe inhibition of *Cc*BglA but not in *Nk*Bgl (Fig. 2). It is supposed that Zn^{2+} ions at high concentration may coordinate to two His ligands (His121 and His180) in the active site of *Cc*BglA (Fig. 4) (Tainer et al., 1992). This metal attaching, Cu^{2+} or Zn^{2+} with Cys in *Tr*Bgl2 as well as Zn^{2+} with His in *Cc*BglA, may hinder the substrate binding and inhibit catalytic reaction in the active site.

In previous reports, Tris served as an inhibitor to reduce the hydrolase activity of β-glucosidases (Ait et al., 1982; Ferreira and Terra, 1983; Park et al., 2001). Here, the crystal structures of *Tr*Bgl2 and *Nk*Bgl in complex with the Tris, and *Nk*Bgl-E193D in complex with the *p*NPG, show the Tris molecule interacting directly with the residues on the glycone-binding site (Fig. 6A and Supplementary Figs. S5 and 6). Furthermore, some chemical reagents containing the poly-hydroxyl group, such as Tris, glycerol and tartaric acid, were also observed to bind to the active site of glycosyl hydrolase family 1 β-glucosidases (Nam et al., 2008; Sanz-Aparicio et al., 1998; Seshadri et al., 2009). Although there is no obvious interaction between Tris molecules and the Trp339 of *Tr*Bgl2 or Trp374 of *Nk*Bgl as shown in the crystal structures, it is supposed that the Tris tolerance of *Cc*BglA is relevant to the different orientation of the side chain of Trp325 in *Cc*BglA (Figs. 2 and 4).

In the conversion processes of lignocellulose to bioethanol, SSF and SSCF offer promising strategies for reducing the product inhibition of glucose. In both SSF and SSCF processes, the hydrolysis of cellulose by different kinds of cellulases proceeds simultaneously with the microbial fermentation to produce ethanol. However, the optimum conditions for cellulase catalysis and microbial fermentation are usually not in concert. Reaction conditions at a near neutral pH and moderate temperature are favorable for microbial growth and fermentation in SSF and SSCF processes. In our study, the bacterial *Cc*BglA exhibits the highest activity and shows the most tolerance to various metal ions. Synergism between *Cc*BglA and Celluclast 1.5L (with endoglucanase and exoglucanase) was also observed on the saccharification of Avicel (microcrystalline cellulose) or acid-pretreated rice straw (data not shown). Furthermore, its optimal reaction parameters demonstrate that *Cc*BglA should be a good candidate for utilization in the SSF and SSCF processes (Lynd et al., 2002). In order to further the industrial application of lignocellulosics ethanol production, we have employed a fermentor to produce active enzymes and increase production yield. Further research into enzyme yield improvement is ongoing.

Acknowledgments

Portions of this research were carried out at the National Synchrotron Radiation Research Center (NSRRC), a national user facil-

ity supported by the National Science Council, Taiwan. The Synchrotron Radiation Protein Crystallography Facility is supported by the National Research Program for Genomic Medicine. We thank NSRRC of Taiwan, SPring-8 and the Photon Factory of Japan for beam time allocations. This work was supported by Academia Sinica and the National Science Council of Taiwan (Grant Numbers NSC 96-3112-B-001-014, NSC 96-3114-P-004-001, NSC 97-3112-B-001-017, and NSC 97-3114-P-001-001 to A.H.J.W.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jsb.2010.07.008.

References

Ait, N., Creuzet, N., Cattaneo, J., 1982. Properties of β-glucosidase purified from *Clostridium thermocellum*. J. Gen. Microbiol. 128, 569–577.

Bailey, S., 1994. The Ccp4 suite – programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. 50, 760–763.

Beguin, P., 1990. Molecular biology of cellulose degradation. Annu. Rev. Microbiol. 44, 219–248.

Bhatia, Y., Mishra, S., Bisaria, V.S., 2002. Microbial β-glucosidases: cloning, properties, and applications. Crit. Rev. Biotechnol. 22, 375–407.

Boisset, C., Fraschini, C., Schulein, M., Henrissat, B., Chanzy, H., 2000. Imaging the enzymatic digestion of bacterial cellulose ribbons reveals the endo character of the cellobiohydrolase Cel6A from *Humicola insolens* and its mode of synergy with cellobiohydrolase Cel7A. Appl. Environ. Microbiol. 66, 1444–1452.

Bommarius, A.S., Katona, A., Cheben, S.E., Patel, A.S., Ragauskas, A.J., Knudson, K., Pu, Y., 2008. Cellulase kinetics as a function of cellulose pretreatment. Metab. Eng. 10, 370–381.

Bradford, M.M., 1976. Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. Anal. Biochem. 72, 248–254.

Brunger, A.T., 1993. Assessment of phase accuracy by cross validation: the free *R* value. Methods and applications. Acta Crystallogr. D Biol. Crystallogr. 49, 24–36.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., Warren, G.L., 1998. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. 54, 905–921.

Brzobohaty, B., Moore, I., Kristoffersen, P., Bako, L., Campos, N., Schell, J., Palme, K., 1993. Release of active cytokinin by a β-glucosidase localized to the maize root meristem. Science 262, 1051–1054.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B., 2009. The carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. 37, D233–D238.

Chen, H., Hayn, M., Esterbauer, H., 1992. Purification and characterization of two extracellular β-glucosidases from *Trichoderma reesei*. Biochim. Biophys. Acta 1121, 54–60.

Dismukes, G.C., 1996. Manganese enzymes with binuclear active sites. Chem. Rev. 96, 2909–2926.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.

Ferreira, C., Terra, W.R., 1983. Physical and kinetic properties of a plasma-membrane-bound β-D-glucosidase (cellobiase) from midgut cells of an insect (*Rhynchosciara americana* larva). Biochem. J. 213, 43–51.

Henrissat, B., 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 280, 309–316.

Husebye, H., Arzt, S., Burmeister, W.P., Hartel, F.V., Brandt, A., Rossiter, J.T., Bones, A.M., 2005. Crystal structure at 1.1 Angstroms resolution of an insect myrosinase from *Brevicoryne brassicae* shows its close relationship to β-glucosidases. Insect Biochem. Mol. Biol. 35, 1311–1320.

Kosugi, A., Arai, T., Doi, R.H., 2006. Degradation of cellulosome-produced cello-oligosaccharides by an extracellular non-cellulosomal β-glucan glucohydrolase, BglA, from *Clostridium cellulovorans*. Biochem. Biophys. Res. Commun. 349, 20–23.

Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. 60, 2256–2268.

Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797.

Leah, R., Kigel, J., Svendsen, I., Mundy, J., 1995. Biochemical and molecular characterization of a barley seed β-glucosidase. J. Biol. Chem. 270, 15789–15797.

Lovell, S.C., Davis, I.W., Arendall 3rd, W.B., de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., Richardson, D.C., 2003. Structure validation by Cα geometry: φ, ψ and Cβ deviation. Proteins 50, 437–450.

Lynd, L.R., Weimer, P.J., van Zyl, W.H., Pretorius, I.S., 2002. Microbial cellulose utilization: fundamentals and biotechnology. Microbiol. Mol. Biol. Rev. 66, 506–577.

Marana, S.R., Jacobs-Lorena, M., Terra, W.R., Ferreira, C., 2001. Amino acid residues involved in substrate binding and catalysis in an insect digestive β-glycosidase. Biochim. Biophys. Acta 1545, 41–52.

McRee, D.E., 1999. XtalView/Xfit – a versatile program for manipulating atomic coordinates and electron density. J. Struct. Biol. 125, 156–165.

Miller, G.L., 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal. Chem. 31, 426–428.

Nam, K.H., Kim, S.J., Kim, M.Y., Kim, J.H., Yeo, Y.S., Lee, C.M., Jun, H.K., Hwang, K.Y., 2008. Crystal structure of engineered β-glucosidase from a soil metagenome. Proteins 73, 788–793.

Nijikken, Y., Tsukada, T., Igarashi, K., Samejima, M., Wakagi, T., Shoun, H., Fushinobu, S., 2007. Crystal structure of intracellular family 1 β-glucosidase BGL1A from the basidiomycete *Phanerochaete chrysosporium*. FEBS Lett. 581, 1514–1520.

Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. Meth. Enzymol. 276, 307–326.

Park, S.Y., Bae, E.A., Sung, J.H., Lee, S.K., Kim, D.H., 2001. Purification and characterization of ginsenoside R_{b1}-metabolizing β-glucosidase from *Fusobacterium* K-60, a human intestinal anaerobic bacterium. Biosci. Biotechnol. Biochem. 65, 1163–1169.

Riou, C., Salmon, J.M., Vallier, M.J., Gunata, Z., Barre, P., 1998. Purification, characterization, and substrate specificity of a novel highly glucose-tolerant β-glucosidase from *Aspergillus oryzae*. Appl. Environ. Microbiol. 64, 3607–3614.

Sanz-Aparicio, J., Hermoso, J.A., Martinez-Ripoll, M., Lequerica, J.L., Polaina, J., 1998. Crystal structure of β-glucosidase A from *Bacillus polymyxa*: insights into the catalytic activity in family 1 glycosyl hydrolases. J. Mol. Biol. 275, 491–502.

Service, R.F., 2007. Cellulosic ethanol. Biofuel researchers prepare to reap a new harvest. Science 315, 1488–1491.

Seshadri, S., Akiyama, T., Opassiri, R., Kuaprasert, B., Cairns, J.K., 2009. Structural and enzymatic characterization of Os3BGlu6, a rice β-glucosidase hydrolyzing hydrophobic glycosides and (1 → 3)- and (1 → 2)-linked disaccharides. Plant Physiol. 151, 47–58.

Sticklen, M.B., 2008. Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. Nat. Rev. Genet. 9, 433–443.

Takashima, S., Nakamura, A., Hidaka, M., Masaki, H., Uozumi, T., 1999. Molecular cloning and expression of the novel fungal β-glucosidase genes from *Humicola grisea* and *Trichoderma reesei*. J. Biochem. 125, 728–736.

Tainer, J.A., Roberts, V.A., Getzoff, E.D., 1992. Protein metal-binding sites. Curr. Opin. Biotechnol. 3, 378–387.

Tokuda, G., Saito, H., Watanabe, H., 2002. A digestive β-glucosidase from the salivary glands of the termite, *Neotermes koshunensis* (Shiraki): distribution, characterization and isolation of its precursor cDNA by 5′- and 3′-RACE amplifications with degenerate primers. Insect Biochem. Mol. Biol. 32, 1681–1689.

Uzun, L., Turkmen, D., Yilmaz, E., Bektas, S., Denizli, A., 2008. Cysteine functionalized poly(hydroxyethyl methacrylate) monolith for heavy metal removal. Colloids Surf. A Physicochem. Eng. Asp. 330, 161–167.

Vaguine, A.A., Richelle, J., Wodak, S.J., 1999. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. Acta Crystallogr. D Biol. Crystallogr. 55, 191–205.

Vasserot, Y., Christiaens, H., Chemardin, P., Arnaud, A., Galzy, P., 1989. Purification and properties of a β-glucosidase of *Hanseniaspora Vineae* Van Der Walt and Tscheuschner with the view to its utilization in fruit aroma liberation. J. Appl. Bacteriol. 66, 271–279.

Wang, Q., Trimbur, D., Graham, R., Warren, R.A., Withers, S.G., 1995. Identification of the acid/base catalyst in *Agrobacterium faecalis* β-glucosidase by kinetic analysis of mutants. Biochemistry 34, 14554–14562.

Winn, M.D., Murshudov, G.N., Papiz, M.Z., 2003. Macromolecular TLS refinement in REFMAC at moderate resolutions. Meth. Enzymol. 374, 300–321.

Zagrobelny, M., Bak, S., Moller, B.L., 2008. Cyanogenesis in plants and arthropods. Phytochemistry 69, 1457–1468.

Zechel, D.L., Boraston, A.B., Gloster, T., Boraston, C.M., Macdonald, J.M., Tilbrook, D.M.G., Stick, R.V., Davies, G.J., 2003. Iminosugar glycosidase inhibitors: Structural and thermodynamic dissection of the binding of isofagomine and 1-deoxynojirimycin to β-glucosidases. J. Am. Chem. Soc. 125, 14313–14323.

Zhang, Y.H., Lynd, L.R., 2004. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. Biotechnol. Bioeng. 88, 797–824.