

Full Paper

Functional expression and characterization of a glucuronoyl esterase from the fungus *Neurospora crassa*: identification of novel consensus sequences containing the catalytic triad

(Received March 4, 2016; Accepted March 29, 2016; J-STAGE Advance publication date: September 6, 2016)

Hung Hiep Huynh and Manabu Arioka*

Department of Biotechnology, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

The complete hydrolysis of lignocellulose requires the actions of a variety of enzymes, including those that cleave the linkage between lignin and hemicellulose. The enzyme glucuronoyl esterase (GE) that constitutes a novel family of carbohydrate esterases, CE15, has been shown to display a unique ability to cleave the ester linkage between lignin alcohols and xylan-bound 4-*O*-methyl-D-glucuronic acid of hemicellulose. We herein report identification, expression, and functional characterization of a new GE, NcGE, from the filamentous fungus *Neurospora crassa*. C-terminally c-myc and hexahistidine-tagged NcGE was heterologously expressed in the methylotrophic yeast *Pichia pastoris*. NcGE purified from the culture supernatant through Ni-NTA and anion exchange chromatographies showed the ability to hydrolyze the substrate 3-(4-methoxyphenyl) propyl methyl 4-*O*-methyl- α -D-glucopyranosiduronate, which mimics the ester linkage of 4-*O*-methyl-D-glucuronic acid in lignin-carbohydrate complexes (LCCs). This esterase showed the characteristic of a mesophilic enzyme with the temperature optimum at 40–50°C, and displayed the optimal activity at pH 7 and broad pH stability. Based on the alignment of NcGE with other GEs so far characterized, we propose novel consensus sequences for GEs containing the catalytic triad.

Key Words: carbohydrate esterase family 15; glucuronoyl esterase; hemicellulose; lignin-carbohydrate complex; *Neurospora crassa*

Abbreviations: AOX1, alcohol oxidase 1 gene; CBM, carbohydrate-binding module; EDTA, ethylenediaminetetraacetic acid; GE, glucuronoyl esterase; HPLC, high performance liquid chromatography; LCCs, lignin-carbohydrate complexes; PMSF, phenylmethylsulfonyl fluoride

Introduction

Plant cell wall is a complex mixture of cellulose, hemicellulose, and lignin. Cellulose, a linear polymer of β 1,4-linked D-glucose units, is considered to be the most abundant organic polymer on earth, and it is thus expected to be a highly potential resource for renewable energy. The complete degradation of lignocellulose requires, however, the hydrolysis of all of its chemical linkages. One of the intimate covalent linkages is found between lignin and hemicelluloses in the lignin-carbohydrate complexes (LCCs). Recent discovery of glucuronoyl esterase (GE) that specifically cleaves the ester linkage between aromatic lignin alcohol and 4-*O*-methyl-D-glucuronic acid residue of glucuronoxylan might be a promising solution for this issue.

GE activity was first discovered in the cellulose-spent culture fluid of the wood-rotting fungus *Schizophyllum commune* (Špáníková and Biely, 2006). Its partial amino acid sequences were used to identify an orthologue, named Cip2, in the genome of *Hypocrea jecorina* (anamorph: *Trichoderma reesei*). Cip2 was then cloned, heterologously expressed and characterized, which demonstrated that it displayed GE activity (Li et al., 2007). These findings led to the establishment of a new family of carbohydrate es-

*Corresponding author: Manabu Arioka, Department of Biotechnology, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan. Tel: +81-3-5841-8230 Fax: +81-3-5841-8033 E-mail: arioka@mail.ecc.u-tokyo.ac.jp

None of the authors of this manuscript has any financial or personal relationship with other people or organizations that could inappropriately influence their work.

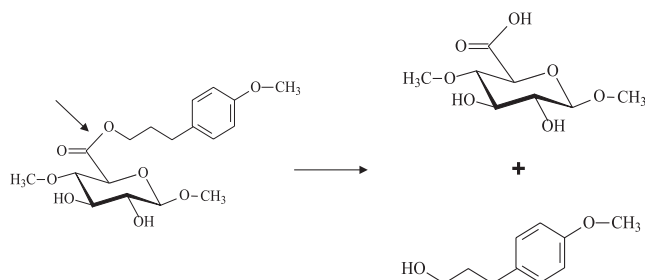


Fig. 1. GE substrate, 3-(4-methoxyphenyl) propyl methyl 4-*O*-methyl- α -D-glucopyranosiduronate, used in this study and the reaction products.

Two products are generated in the reaction: methyl 4-*O*-methyl-D-glucuronic acid and 3-(4-methoxyphenyl) propyl alcohol. The arrow indicates where the enzyme cleaves the substrate.

terase, CE15, in the carbohydrate-active enzymes (CAZy) database (www.cazy.org; Cantarel et al., 2009). GEs from *Phanerochaete chrysosporium* (GE1 and GE2) and *Sporotrichum thermophile* (StGE2) were subsequently identified and successfully expressed (Đuranová et al., 2009b; Topakas et al., 2010). From the alignment of homologous sequences, a novel consensus sequence G-C-S-R-X-G conserved in CE15 family members was suggested (Topakas et al., 2010). This consensus was further supported by the crystal structures of two glucuronoyl esterases, Cip2 and StGE2 (Charavgi et al., 2013; Pokkuluri et al., 2011), which confirmed the presence of this sequence in the catalytic center and the role of central serine residue as a nucleophile in the substrate cleavage. Later, the cloning of ScGE from *Schizophyllum commune* (Wong et al., 2012) was performed, followed by two other GEs from *Podospira anserine* and *Cerrena unicolor* (d'Errico et al., 2015; Katsimpouras et al., 2014).

Kinetic analysis using the synthetic substrates which mimic the ester bond in the plant cell wall indicated that glucuronoyl esterases specifically recognize the alkyl and arylalkyl esters of 4-*O*-D-glucuronic acid (Špáníková et al., 2007; Wong et al., 2012) and prefer 4-*O*-methyl-D-glucuronic acid esters rather than D-glucuronic acid esters (Đuranová et al., 2009a). In addition, the carbohydrate part to which the uronic acid is linked does not play a decisive role in the mechanism of hydrolysis (Špáníková et al., 2007). The aromatic alcohol group of lignin in nature often entails a very large molecule, and the preference of GEs from *C. unicolor* and some other fungi for the ester of bulky alcohol was also suggested (d'Errico et al., 2015). For these reasons, 3-(4-methoxyphenyl) propyl methyl 4-*O*-methyl- α -D-glucopyranosiduronate (Fig. 1) that mimics the ester linkage between hemicellulose and lignin is often used as a typical substrate for GEs. This substrate contains 4-*O*-methyl group in D-glucuronic acid and also methoxyphenyl propyl group of alcohol representing the bulky alcohol.

In the present work, we report the identification, functional expression, and characterization of a putative GE, NcGE, from *Neurospora crassa*. BLAST search shows that significant amino acid sequence homology is found between NcGE and Cip2. In addition, the alignment with Cip2 revealed that the consensus sequence G-C-S-R-X-G

that features the characteristic serine residue involved in the catalytic mechanism is also conserved in NcGE (Topakas et al., 2010). We cloned the gene from *N. crassa* genome and expressed it in the methylotrophic yeast *Pichia pastoris*. NcGE purified from the culture supernatant was capable of hydrolyzing synthetic compound that mimics the ester linkage described in the lignin-carbohydrate complexes (LCCs). From the alignment of NcGE with other characterized GEs, we propose two additional consensus sequences, besides the one described above, containing the glutamic acid and histidine residues that constitute the catalytic triad of GEs.

Materials and Methods

Strains and growth media. *Escherichia coli* DH5 α and *Pichia pastoris* KM71H strains were used in this study. *E. coli* was cultured in LB low salt medium (1% peptone, 0.5% yeast extract, and 0.5% NaCl) containing zeocin (40 μ g/ml) as a selection marker for pPICZ α vector. For methanol induction of *P. pastoris*, YPG medium containing 1% yeast extract, 2% peptone, 1% glycerol, and zeocin (100 μ g/ml) was used, whereas glycerol and zeocin were omitted in YP medium.

Cloning of NcGE gene. The gene encoding NcGE (NCU09445.7) was amplified from the genomic DNA of *N. crassa* strain FGSC987 by polymerase chain reaction using PrimeStar DNA polymerase (Takara, Japan). Two PCR primers used to amplify the putative mature region (amino acids 21-395) were designed as follows: 5'-AAGCAGAATTCGCCCTGCTTCTCAAATCT-3' (*Eco*R I site is underlined) and 5'-GCTCTAGATACGACAAAGCCGGCACATC-3' (*Xba*I site is underlined). The PCR cycle includes denaturation (98°C, 10 sec), annealing (55°C, 10 sec), and extension (72°C, 2 min). The amplified fragment was digested with the restriction enzymes and then gel-purified to be cloned into pPICZ α vector. After verification of the nucleotide sequence, this recombinant plasmid pPICZ α /NcGE was used in the subsequent studies.

Transformation of *P. pastoris*. The plasmid pPICZ α /NcGE was linearized with *Bgl* II and transformed to *P. pastoris* by electroporation according to the instruction manual of EasySelect *Pichia* expression kit (Invitrogen). The transformants were selected on the plate containing zeocin (100 μ g/ml). Colony PCR using KOD FX Neo (Toyobo, Japan) was applied to screen the correct transformants. Positive clones were kept frozen in 25% glycerol at -80°C.

Expression of NcGE in *P. pastoris*. The *P. pastoris* transformant was pre-cultured in 10 ml YPG medium at 30°C for 24 h with strong shaking. Then the cells were transferred to the main culture containing 190 ml YPG medium in 500 ml flask and grown for another 24 h with shaking at 150 rpm. The cells were harvested by centrifugation (1,500 g, 5 min) and re-suspended in 40 ml YP medium for methanol induction. The medium was supplemented daily with 0.75% (v/v) methanol and the culture was continued for 4 days at 26°C. The supernatant was collected to analyze the protein expression by SDS-



Fig. 2. Alignment between NcGE and Cip2.

The alignment between NcGE (*N. crassa* NCU09445.7) and Cip2 (*H. jecorina* AAP57749) is shown. The underlines show the signal peptides, and the shaded letters indicate the CBM of Cip2. The consensus GCSR sequence (Topakas et al., 2010) is boxed.

PAGE and Western blotting. In the Western blotting, anti-c-myc mouse monoclonal antibody (Clontech) and peroxidase-labeled anti-mouse IgG antibody (Vector Laboratories) were used as the primary and secondary antibodies, respectively. Western Lightning Plus system (PerkinElmer) and LAS-4000mini EPUV luminescent image analyzer (Fujifilm) were used to detect c-Myc tag in the recombinant fusion protein.

Purification of NcGE by Ni^{2+} -NTA affinity chromatography. A sufficient amount of Ni^{2+} -nitriloacetic acid (NTA) agarose was added to the column for binding of the secreted NcGE protein in the supernatant. Purification of NcGE was performed based on the instruction of QIAGEN manual. The contaminant proteins were removed by washing the column with 50 mM Tris-HCl buffer (pH 8), 200 mM NaCl containing 0 or 10 mM imidazole. The target protein was eluted with the same buffer containing 100 mM imidazole. To elute NcGE, one ml of elution buffer was used in the first fraction, and 0.5 ml each of elution buffer was applied in the subsequent three fractions to obtain all NcGE enzyme.

Purification of NcGE by ion exchange chromatography. To achieve higher purity, NcGE partially purified by Ni^{2+} -NTA chromatography was applied to ÄKTA chromatography system (GE Healthcare) with PrimerView 5.0 program. Enzyme purification was carried out using HiTrap SP XL column (1 ml) with a flow rate of 1 ml/min at room temperature. The starting buffer was 10 mM Tris-HCl, pH 7, while the elution buffer contained 1 M NaCl. A linear gradient of 0% to 100% elution buffer was applied to collect the target protein. Purified NcGE was kept in 10 mM Tris-HCl (pH 7) containing 20% glycerol for storage.

Substrate for glucuronoyl esterase. In order to check the enzyme activity, the synthetic compound 3-(4-methoxyphenyl) propyl methyl 4-*O*-methyl- α -D-glucopyranosiduronate was used (Fig. 1) (Sasagawa et al., 2011; Špáníková et al., 2007). This compound has 4-*O*-

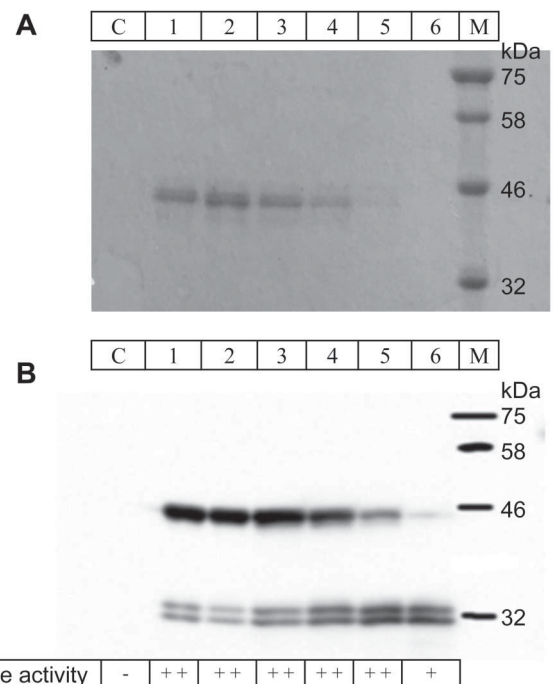


Fig. 3. Production of NcGE in *P. pastoris*.

CBB staining (A) and Western blot analysis (B) of the culture supernatants of *P. pastoris* pPICZαA/NcGE transformant. Lanes 1–6 indicate the samples from day 1–6, respectively. C, control (empty vector) transformant. M, protein marker.

methyl group in glucuronic acid which is suggested to play an important role in the enzyme-substrate interaction (Đuranová et al., 2009a; Špáníková et al., 2007). Besides that, the substrate contains the aromatic UV-absorbing chromophore which can be used to detect the de-esterification of substrate by high performance liquid chromatography (HPLC).

Determination of enzyme activity. Enzyme reaction was conducted in 50 mM sodium phosphate buffer, pH 6.0 at

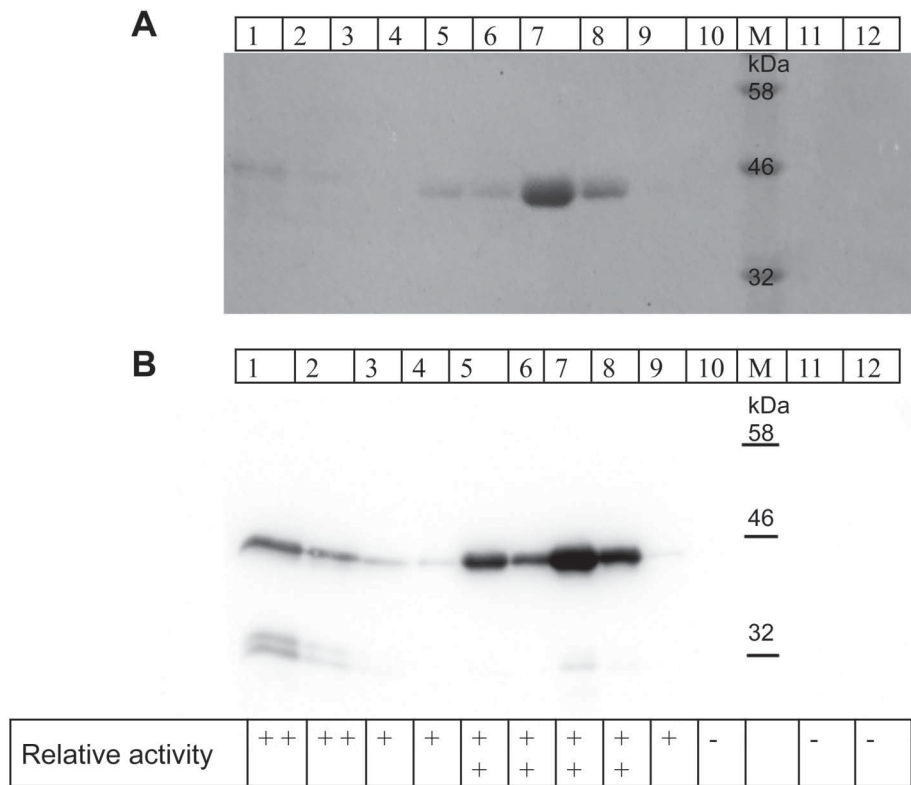


Fig. 4. Purification of NcGE by Ni²⁺-NTA column chromatography. CBB staining (A) and Western blot analysis (B) of Ni²⁺-NTA column chromatography fractions. Lane 1, the culture supernatant of pPICZαA/NcGE transformant; lane 2, flow through, lanes 3 and 4, wash with 0 mM imidazole; lanes 5 and 6, wash with 20 mM imidazole; lanes 7–10, elution with 100 mM imidazole; lanes 11 and 12, elution with 250 mM imidazole; M, protein marker.

30°C to avoid the self-degradation of substrate under high alkaline condition (Li et al., 2007; Špáníková and Biely, 2006). Hydrolysis of substrate was analyzed by HPLC system (Hitachi High-Tech Science, Japan) using the reverse-phase column (SSC-2300, Senshu Scientific Co., Ltd, Japan) and acetonitrile:water = 2:1 (v/v) at a flow rate 1 ml/min as a mobile phase.

Enzymatic characterization of NcGE. To quantify the enzyme activity, the amount of the reaction product, 3-(4-methoxyphenyl)-1-propanol, was measured by the standard curve drawn using the commercially obtained same compound (Sigma-Aldrich, Japan; cat. No. 142328). One unit of GE activity was defined as the amount of the enzyme required to release 1 μmol of alcohol product per min at 30°C. The kinetic parameters of NcGE were obtained by the Lineweaver-Burk plot analysis. For the quantification of proteins, the standard procedure of Bradford method (Bio-Rad, Japan) was applied using the bovine serum albumin as the standard. The optimum pH was determined by performing the assay at 30°C for 10 min over the pH range 3.0–9.0 using 0.1 M citrate-phosphate buffer (pH 3.0–7.0) and 0.1 M Tris-HCl buffer (pH 7.0–9.0). The pH stability was evaluated by pre-incubating the enzyme with these buffers at 30°C for 30 min and then testing the residual activity at 30°C, pH 7.0. The optimum temperature was determined by performing the assay for 10 min at various temperatures from 30°C to 70°C in 0.1 M citrate-phosphate buffer (pH 7.0). For the thermostability, the

remaining activity was checked after incubation at different temperatures (30–70°C) for 30 min at pH 7.0.

Effect of metal ion and other reagents. Purified NcGE was mixed with metal ions (1 mM each of Mg²⁺, Mn²⁺, Ca²⁺, Fe²⁺, Fe³⁺, Cu²⁺, Zn²⁺, Al³⁺, Ni²⁺, and Li⁺) or reagents (10 mM (final concentration) of glycerol, Tween 80, acetic anhydride, ascorbic acid, dimethyl sulfoxide, dithiothreitol, 2-mercaptoethanol, SDS; 0.5–1 mM phenylmethylsulfonyl fluoride (PMSF); 5 mM ethylenediaminetetraacetic acid (EDTA)) and incubated at 30°C for 10 min. Then the residual GE activity was measured. The control enzyme activity without cations/reagents was taken as 100%. The experiments were done in triplicates and the statistical difference was analyzed by Student's *t* test.

Homology analysis. The amino acid sequences of GEs were collected from NCBI database and aligned by ClustalX program (Larkin et al., 2007). Evolutional analyses were conducted by MEGA6 (Tamura et al., 2013) using the Neighbor-Joining method. The bootstrap consensus tree was inferred from 1000 replicates. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The data were computed using the pairwise deletion and *p*-distance algorithm with out-group sequences including two different acetyl xylan esterases (AXEs) from *H. jecorina* (CAA93247.1; CE family 5) and *Aspergillus niger* (CAA01634.1; CE family 1), and one feruloic acid esterase (FAE) from *N. crassa* (CAC05587.1; CE family 1).

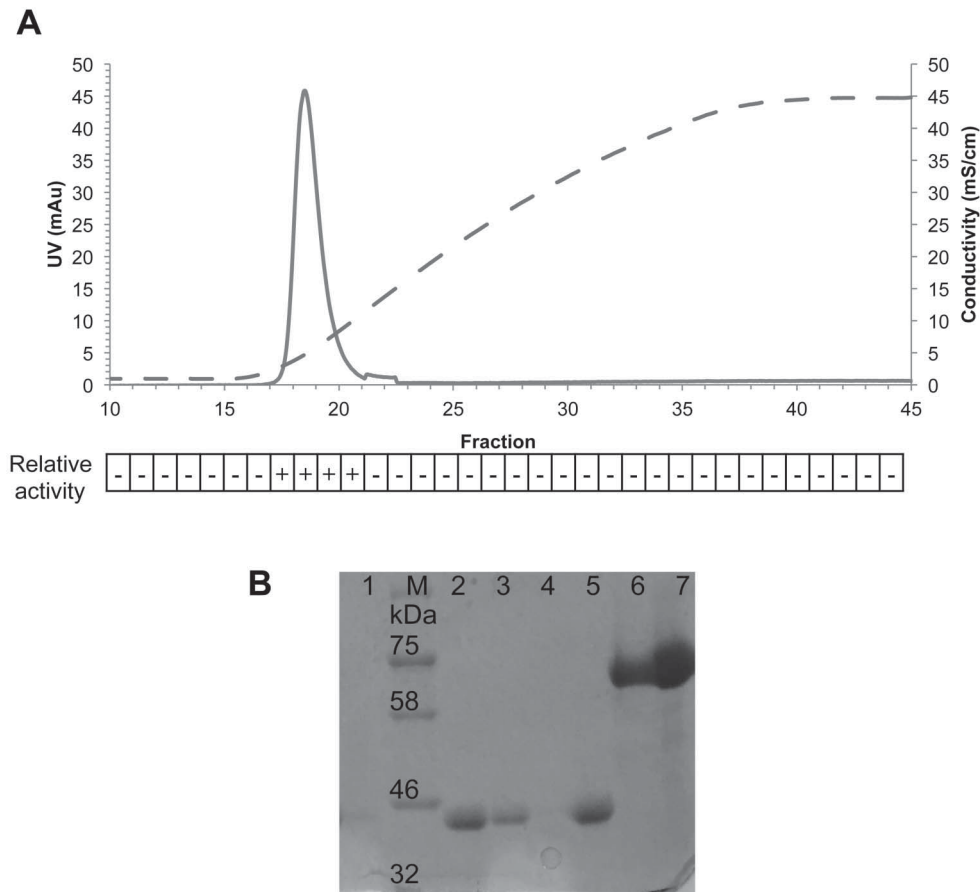


Fig. 5. Cation exchange chromatography.

(A) A_{280} (solid line), conductivity (dashed line), and GE activity of each fraction of cation exchange chromatography. (B) Lanes 1–4, fractions 18–21; lane 5, NcGE sample before loading; lanes 6 and 7, bovine serum albumin (5 and 10 μ g, respectively); M, protein marker.

Results

Sequence analysis of NcGE

A BLAST search of *N. crassa* genome database using the amino acid sequence of Cip2 from *H. jecorina* resulted in the identification of a putative protein, NCU09445.7, displaying a significant sequence similarity to Cip2 (Fig. 2) and other CE15 GEs. This protein, named NcGE, is comprised of 395 amino acids with the predicted N-terminal signal peptide. The genomic sequence encoding NcGE does not contain any intron and the deduced protein lacks the carbohydrate binding module (CBM). The calculated molecular mass of mature polypeptide is around 41.3 kDa with the predicted isoelectric point of about 8.8 (www.expasy.org; Gasteiger et al., 2005). NcGE has a consensus G-C-S-R-X-G motif conserved in the CE15 family where the serine residue serves as the catalytic nucleophile (Pokkuluri et al., 2011; Topakas et al., 2010).

Cloning, expression, and purification of NcGE

An *E. coli*/*P. pastoris* shuttle vector pPICZ α A was chosen to produce recombinant NcGE. This vector has a strong alcohol oxidase 1 (*AOX1*) promoter which drives the expression of a foreign gene when methanol is added to the culture medium. In addition, NcGE was produced with C-terminal c-myc and hexahistidine tags. SDS-PAGE analy-

sis of the culture supernatant clearly showed that a protein of about 44 kDa was produced in the pPICZ α A/NcGE transformant, whereas no protein was detected in the vector-transformed control strain (Fig. 3A). Production of NcGE was further confirmed by Western blot analysis (Fig. 3B) using anti-c-Myc antibody. Smaller-sized doublet bands at around the position of 32 kDa were also detected, suggesting that NcGE was partially degraded. Semi-quantitative measurement of GE activity using the substrate shown in Fig. 1 demonstrated that GE activity was detected in the fractions containing the 44 kDa protein, but not in the culture supernatant of the control strain. Since the highest amount of NcGE protein was obtained after 2 days of methanol induction with the approximate protein concentration of 0.06 mg/ml, methanol induction was performed for 2 days to achieve the highest yield of secreted NcGE protein.

NcGE was purified from the culture supernatant firstly by Ni^{2+} -NTA affinity chromatography (Fig. 4). It seemed that the target protein did not strongly bind to the column, since a small amount of protein was detected in the flow through (Fig. 4, lane 2) and wash fractions containing no or 20 mM imidazole (Fig. 4, lanes 3–6). The majority of NcGE was eluted by 100 mM imidazole (Fig. 4, lanes 7 and 8). Further purification of NcGE was conducted by cation exchange chromatography at pH 7.0, since the pre-

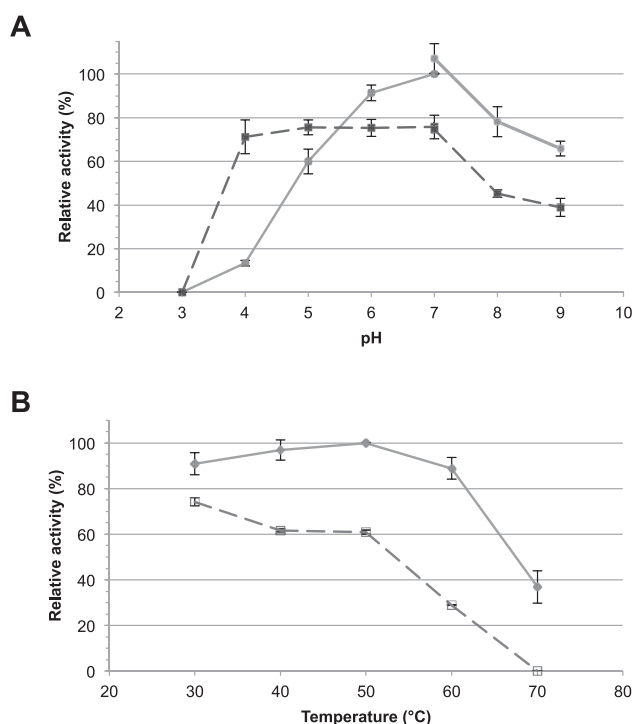


Fig. 6. Effect of pH (A) and temperature (B) on the activity (solid line) and stability (dashed line) of purified NcGE.

The optimum pH was determined by performing the assay at 30°C for 10 min over the pH range 3.0–9.0. The pH stability was evaluated by pre-incubating the enzyme at different pH at 30°C for 30 min and then testing the residual activity at 30°C, pH 7.0. The optimum temperature was determined by performing the assay for 10 min at various temperatures from 30°C to 70°C. For the thermostability, the remaining activity was checked after incubation at different temperatures (30–70°C) for 30 min at pH 7.0.

dicted pI of NcGE is about 8.8. The enzyme was successfully purified (Fig. 5).

Enzymatic characterization of NcGE

The optimum pH and temperature for NcGE was pH 7.0 and 40–50°C, respectively (Fig. 6). NcGE was stable in the relatively broad pH range (pH 4–7). The enzyme was stable at 30°C, but it completely lost its activity when pre-incubated at 70°C for 30 min. There are no available data which examined the temperature profiles of GEs employing the same substrate used in this study. However, in Cip2 and StGE (both of which were also heterologously expressed in *P. pastoris* using pPICZα vector), the pH and temperature profiles were determined using 4-nitrophenyl 2-*O*-(methyl-4-*O*-methyl-α-D-glucopyranosyluronate)-β-D-xylopyranoside as a substrate (Li et al., 2007; Topakas et al., 2010). Both GEs showed higher optimum temperature (40–60°C and 50–55°C for Cip2 and StGE, respectively). In addition, Cip2 was quite stable at 40°C (Li et al., 2007) and StGE displayed significant stability up to 50°C (Topakas et al., 2010). This is in contrast to NcGE which quickly lost its activity at higher temperatures, especially above 50°C, despite that NcGE is most closely related to StGE (72% identity) in the phylogenetic tree (Fig. 7).

Next, the kinetic constants of NcGE were determined. K_m was around 15 mM and the catalytic efficiency ($k_{cat}/$

Table 1. Effect of metal ions and other reagents on the activity of NcGE.

Cation or reagent	Relative activity (%)
Control	100 ± 1.65
MnCl ₂	93.6 ± 2.34
MgCl ₂	118.4 ± 0.91
CaCl ₂	127.0 ± 7.11
FeSO ₄	103.6 ± 6.72
FeCl ₃	114.9 ± 2.37
CuCl ₂	63.0 ± 1.58
ZnSO ₄	84.1 ± 0.31
AlCl ₃	118.0 ± 3.46
NiSO ₄	59.7 ± 3.52
LiCl	75.9 ± 5.65
Glycerol	103.8 ± 6.37
Tween 80	149.6 ± 1.38
EDTA	102.5 ± 1.33
PMSF	40.8 ± 5.36
Acetic anhydride	93.2 ± 5.19
Ascorbic acid	97.5 ± 4.05
Dimethyl sulfoxide	86.0 ± 6.70
Dithiothreitol	71.1 ± 2.63
2-Mercaptoethanol	83.9 ± 3.47
Urea	85.5 ± 1.59
SDS	0.0 ± 0.0

K_m) was 1.12 mM⁻¹s⁻¹. Similar k_{cat}/K_m ratio (1.1 mM⁻¹s⁻¹) was observed in Cip2 (Li et al., 2007). In the case of ScGE, a GE from *S. commune*, K_m was much lower (1.78 mM) (Špáníková et al., 2007), indicating that ScGE has higher affinity for this substrate, which is reflected in the higher catalytic efficiency (k_{cat}/K_m was reported to be 4.38 mM⁻¹s⁻¹ in ScGE).

Effect of metal ions and other reagents

The effect of cations and chemical compounds on the activity of NcGE is shown in Table 1. The metal ions such as Mg²⁺, Fe³⁺, and Al³⁺ increased the enzyme activity with the highest effect being observed with Ca²⁺ and Tween 80, whereas significant inactivation (≈40%) was caused by Cu²⁺ and Ni²⁺. The enzyme was affected neither by glycerol and ascorbic acid, nor by two potential inhibitors acetic anhydride and dimethyl sulfoxide. As expected, the activity of NcGE was significantly inhibited by the compounds influencing the enzyme structure, such as dithiothreitol, 2-mercaptoethanol, and urea, while the enzyme was completely inhibited by SDS.

As mentioned above, NcGE has the consensus sequence for GEs containing the nucleophile serine residue. In agreement, NcGE activity was decreased by 60% when the enzyme was incubated with 0.5 mM PMSF, and the activity was totally lost in the presence of 1 mM PMSF (data not shown), demonstrating that this enzyme is truly a serine type-esterase. In contrast, NcGE activity was not affected by 5 mM EDTA, suggesting that it is not a metalloenzyme.

Discussion

N. crassa has the ability to metabolize both cellulose

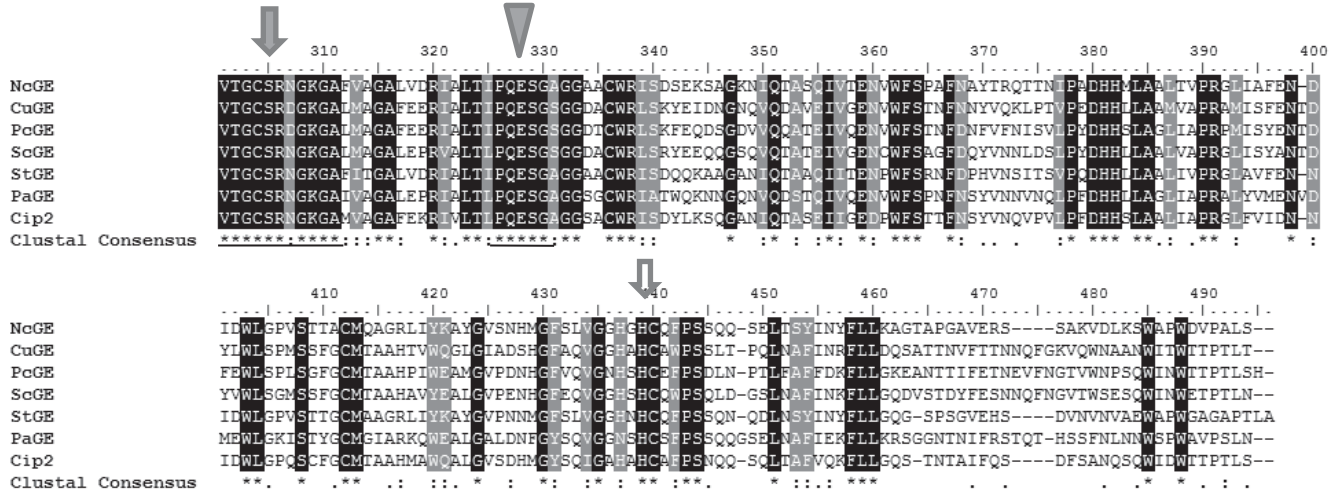
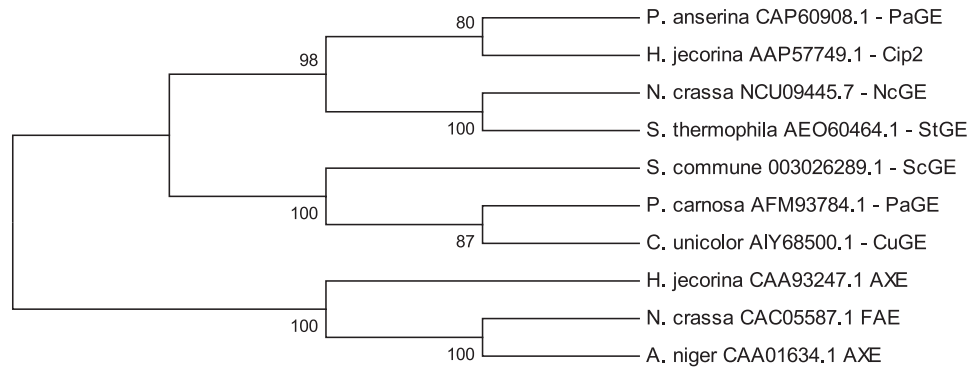
A**B**

Fig. 7. Amino acid sequence alignment (A) and phylogenetic tree analysis (B) of characterized glucuronoyl esterases and other carbohydrate esterases.

The alignment of characterized GEs including NcGE is shown. The residues forming the catalytic triad are indicated as follows: nucleophile serine (solid arrow); glutamic acid (arrowhead) and histidine (open arrow). The consensus sequences of GEs proposed in this study including the one previously identified (GCSRXXG; Topakas et al., 2010) are underlined. GEs from *Cerrena unicolor* (CuGE), *Neurospora crassa* (NcGE), *Phanerochaete carnosae* (PcGE), *Podospora anserina* (PaGE), *Schizophyllum commune* (ScGE), *Sporotrichum thermophile* (StGE2), and *Hypocrea jecorina* (Cip2), together with two acetyl xylan esterases from *H. jecorina* and *Aspergillus niger*, and one feruloic acid esterase (FAE) from *N. crassa*, were analyzed.

and hemicellulose from the plant cell walls (Znameroski et al., 2012). In this report, we showed that the putative protein NCU09445.7 in *N. crassa* is actually a functional glucuronoyl esterase. CBM involved in the interaction of cellulolytic enzymes with the insoluble substrates was not found in the NcGE sequence. This is similar to ScGE from *S. commune* (Wong et al., 2012), PcGE2 from *P. chrysosporium* (Duranová et al., 2009b), and StGE from *S. thermophile* (Vafiadi et al., 2009). In contrast, Cip2, PcGE1 from *P. chrysosporium* (Duranová et al., 2009b), and CuGE from *C. unicolor* (d'Errico et al., 2015) contain type 1 CBM at the N-terminus of the catalytic domain. CBMs are classified into subfamilies based on the amino acid sequence similarity and most of the fungal CBMs in cellulases belong to the family 1, in which three aromatic residues create a planar structure to bind to the crystallized surfaces of chitin/cellulose (Mello and Polikarpov, 2014).

When expressed in *P. pastoris*, a protein of 44 kDa which

is nearly equal to the calculated size of myc- and hexahistidine-tagged NcGE was produced. This agrees well with the absence of potential *N*-glycosylation site in NcGE. Although 15 *O*-glycosylation sites were predicted by NetOGlyc 4.0 server, it is unlikely that NcGE is highly *O*-glycosylated.

Toward the substrate 3-(4-methoxyphenyl) propyl methyl 4-*O*-methyl- α -D-glucopyranosiduronate (Fig. 1), NcGE displayed much higher K_m value (15 mM) compared with ScGE (1.78 mM). It thus seems that NcGE does not prefer this substrate, although it clearly cleaved the ester linkage as a typical GE. The reason may be the synthetic substrate simulates but could not exactly represent the natural bonds of target substrate (Wong et al., 2012), and thus the natural substrate and physiological role of NcGE are yet to be elucidated.

The amino acid sequences of NcGE and 6 other GEs with experimentally-proven GE activity were aligned (Fig. 7). It is evident that the motif GCSRXXG containing the

nucleophile serine is strongly conserved. This consensus sequence was identified by aligning the sequences of three characterized GEs from *H. jecorina* (Cip2), *S. thermophile* (StGE2), and *P. chrysosporium* with other putative GE sequences showing more than 40% amino acid sequence identity with StGE (Topakas et al., 2010). In this paper, using the sequences of seven characterized GEs, we propose that this motif could be expanded to VTGCSRXXGKGA (the original motif is underlined). Besides, based on the crystal structures of Cip2 (Pokkuluri et al., 2011) and StGE2 (Charavgi et al., 2013), the catalytic triad of GEs was suggested to be composed of two more amino acids, glutamic acid and histidine, located C-terminal to the nucleophile serine. The molecular docking study of StGE2 with the model substrate (Katsimpouras et al., 2014) also supported this idea, which led to the hypothesis that Ser-Glu-His is the standard catalytic triad for GEs. Taking these into consideration, we propose two more consensus sequences constituting the catalytic triad of GEs: PQESG and HC containing the glutamic acid and histidine residues (underlined; position 328 and 439 in NcGE, respectively). These sequences are strongly conserved in all GE sequences, suggesting again the importance of these residues as the components of catalytic triad. In support of our proposal, the crystal structures of Cip2 and StGE (Charavgi et al., 2013; Pokkuluri et al., 2011) showed that the disulfide bond is formed between the cysteine residues equivalent to Cys304 and Cys440 in NcGE that links the consensus sequences VTGCSRXXGKGA and HC, hence bringing serine and histidine residues close together in the catalytic site (Charavgi et al., 2013; Katsimpouras et al., 2014). Cys336 involved in another disulfide bond with Cys412 is near E328 and both of these residues are also highly conserved, possibly assisting in introducing glutamic acid to create the catalytic triad. Therefore, these two disulfide bonds may play an important role to form the catalytic triad in CE15. Although our proposal needs further verification by additional experiments, it might be informative in identifying novel GE sequences, elucidation of their catalytic mechanism, and in achieving the ultimate goal-complete degradation of lignocellulose.

References

- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. et al. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.*, **37**, D233–D238, doi:10.1093/nar/gkn663.
- Charavgi, M. D., Dimarogona, M., Topakas, E., Christakopoulos, P., and Chrysina, E. D. (2013) The structure of a novel glucuronoyl esterase from *Myceliophthora thermophila* gives new insights into its role as a potential biocatalyst. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **69**, 63–73, doi:10.1107/S0907444912042400.
- d'Errico, C., Jørgensen, J. O., Krogh, K. B. R. M., Spødsberg, N., Madsen, R. et al. (2015) Enzymatic degradation of lignin-carbohydrate complexes (LCCs): Model studies using a fungal glucuronoyl esterase from *Cerrena unicolor*. *Biotechnol. Bioeng.*, **112**, 914–922, doi:10.1002/bit.25508.
- Đuranová, M., Hirsch, J., Kolenová, K., and Biely, P. (2009a) Fungal glucuronoyl esterases and substrate uronic acid recognition. *Biosci. Biotechnol. Biochem.*, **73**, 2483–2487, doi:10.1271/bbb.90486.
- Đuranová, M., Špáníková, S., Wösten, H. aB., Biely, P., and De Vries, R. P. (2009b) Two glucuronoyl esterases of *Phanerochaete chrysosporium*. *Arch. Microbiol.*, **191**, 133–140, doi:10.1007/s00203-008-0434-y.
- Gasteiger, E., Hoogland, C., a Gattiker, Duvaud, S., Wilkins, M. et al. (2005) Protein identification and analysis tools on the ExPASy server. *Proteomics Protoc. Handb.*, 571–607, doi:10.1385/1-59259-890-0:571.
- Katsimpouras, C., Bénarouche, A., Navarro, D., Karpusas, M., Dimarogona, M. et al. (2014) Enzymatic synthesis of model substrates recognized by glucuronoyl esterases from *Podospora anserina* and *Myceliophthora thermophila*. *Appl. Microbiol. Biotechnol.*, **98**, 5507–5516, doi:10.1007/s00253-014-5542-9.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A. et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948, doi:10.1093/bioinformatics/btm404.
- Li, X.-L., Špáníková, S., de Vries, R. P., and Biely, P. (2007) Identification of genes encoding microbial glucuronoyl esterases. *FEBS Lett.*, **581**, 4029–4035, doi:10.1016/j.febslet.2007.07.041.
- Mello, B. L. and Polikarpov, I. (2014) Family 1 carbohydrate binding-modules enhance saccharification rates. *AMB Express.*, **4**, 36, doi:10.1186/s13568-014-0036-9.
- Pokkuluri, P. R., Duke, N. E. C., Wood, S. J., Cotta, M. a., Li, X. L. et al. (2011) Structure of the catalytic domain of glucuronoyl esterase Cip2 from *Hypocrea jecorina*. *Proteins Struct. Funct. Bioinforma.*, **79**, 2588–2592, doi:10.1002/prot.23088.
- Sasagawa, T., Matsui, M., Kobayashi, Y., Otagiri, M., and Moriya, S. (2011) Plasmid high-throughput recombinant gene expression systems in *Pichia pastors* using newly developed plasmid vectors. *Plasmid*, **65**, 65–69, doi:10.1016/j.plasmid.2010.08.004.
- Špáníková, S. and Biely, P. (2006) Glucuronoyl esterase—Novel carbohydrate esterase produced by *Schizophyllum commune*. *FEBS Lett.*, **580**, 4597–4601, doi:10.1016/j.febslet.2006.07.033.
- Špáníková, S., Poláková, M., Joniak, D., Hirsch, J., and Biely, P. (2007) Synthetic esters recognized by glucuronoyl esterase from *Schizophyllum commune*. *Arch. Microbiol.*, **188**, 185–189, doi:10.1007/s00203-007-0241-x.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A., and Kumar, S. (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.*, **30**, 2725–2729, doi:10.1093/molbev/mst197.
- Topakas, E., Moukoulis, M., Dimarogona, M., Vafiadi, C., and Christakopoulos, P. (2010) Functional expression of a thermophilic glucuronoyl esterase from *Sporotrichum thermophile*: identification of the nucleophilic serine. *Appl. Microbiol. Biotechnol.*, **87**, 1765–1772, doi:10.1007/s00253-010-2655-7.
- Vafiadi, C., Topakas, E., Biely, P., and Christakopoulos, P. (2009) Purification, characterization and mass spectrometric sequencing of a thermophilic glucuronoyl esterase from *Sporotrichum thermophile*. *FEMS Microbiol. Lett.*, **296**, 178–184, doi:10.1111/j.1574-6968.2009.01631.x.
- Wong, D. W. S., Chan, V. J., McCormack, A. a., Hirsch, J., and Biely, P. (2012) Functional cloning and expression of the *Schizophyllum commune* glucuronoyl esterase gene and characterization of the recombinant enzyme. *Biotechnol. Res. Int.*, **2012**, 951267, doi:10.1155/2012/951267.
- Znameroski, E. a., Coradetti, S. T., Roche, C. M., Tsai, J. C., Iavarone, a. T. et al. (2012) Induction of lignocellulose-degrading enzymes in *Neurospora crassa* by cellodextrins. *Proc. Natl. Acad. Sci.*, **109**, 6012–6017, doi:10.1073/pnas.1118440109.