Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 336 (2005) 299–308
BBRC
www.elsevier.com/locate/ybbrc

Cloning, sequence analysis, and characterization of a novel β-glucosidase-like activity from *Pichia etchellsii*

Pranita Roy, Saroj Mishra *, Tapan K. Chaudhuri

Department of Biochemical Engineering and Biotechnology, Indian Institute of Technology, Delhi, Hauz-Khas, New Delhi 110016, India

Received 1 July 2005
Available online 19 August 2005

Abstract

Genomic DNA fragment encoding a novel β-glucosidase-like activity of the yeast *Pichia etchellsii* was cloned and expressed in *Escherichia coli*. An open-reading frame of 1515 bp, termed *mug*A, coding for a protein of predicted molecular mass of approximately 54 kDa was confirmed for this activity. The sequence of the deduced protein did not show homology with the generic β-glucosidases but a high degree of identity was seen with several Ser-Asp (SD)-rich cell-surface-associated proteins. The secondary structure prediction program 3D-PSSM indicated the protein to be composed of largely helical and coiled structures, which was confirmed by circular dichroism spectroscopy. The encoded protein, MUGA, was purified by about 53-fold and characterized as a monomer of 52.1 kDa by SDS–PAGE and MALDI-TOF. The protein displayed high hydrolytic activity on methylumbelliferyl β-D-glucoside but relatively very little hydrolysis of *p*-nitrophenyl β-D-glucoside and gentiobiose, characteristic substrates for β-glucosidases. The binding experiments performed between *P. etchellsii* cells and the purified *E. coli* expressed MUGA indicated binding with the cell surface, which was monitored by fluorescence microscopy. In competition experiments with the SD dipeptide, less protein was shown to bind to the cell surface, in a concentration-dependent manner.

Keywords: β-Glucosidase-like protein; MUG hydrolysis; CD spectroscopy; Cell-surface-associated protein; Fluorescence microscopy

β-Glucosidases (β-D-glucoside glucohydrolase, EC 3.2.1.21) are able to cleave the β-glucosidic linkages in di- and oligoglucosaccharides, and several other glycoconjugates. These enzymes are widely distributed in the living world and play pivotal roles in many biological processes. The physiological roles associated with this enzyme are diverse and depend on the location of the enzyme and the biological system in which these occur. In cellulolytic microorganisms, β-glucosidase is involved in cellulase induction and cellulose hydrolysis [1,2]. In plants, this enzyme is involved in β-glucan synthesis during cell wall development, pigment metabolism, fruit ripening, and defense mechanisms [3,4]. In humans and other mammals, the enzyme is involved in the hydrolysis of glucosyl ceramides [5] and the deficiency of the enzyme leads to Gaucher's disease. Like many hydrolases, these enzymes can be used for synthesizing a variety of glycoconjugates such as alkyl glucosides, aminoglycosides, and special disaccharide fragments of phytoalexin—elicitor oligosaccharides which are involved in plant and other microbial defense mechanisms (For review, see [6]).

A number of bacterial, yeast, and fungal β-glucosidase genes have been cloned. At the sequence level, they are placed in family 1 and family 3 of glycoside hydrolase families [7], which are 99 (last update, May 19, 2005) in number. The sequence based classification is continuously updated (URL: http://afmb.cnrs-mrs.fr/~pedro/CAZY) and is useful in characterizing the enzymes from the structural point of view but the substrate specificity with respect to the aglycone moiety still serves a primary, or, in some cases, the only lead in isolating and characterizing unknown or structurally undefined glucosidases. In general,

* Corresponding author. Fax: +91 11 2658 2282.
E-mail addresses: saroj@dbeb.iitd.ernet.in, saroj98@hotmail.com (S. Mishra).

0006-291X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2005.08.067

β-glucosidases hydrolyze *p*-nitrophenyl β-D-glucoside (*p*NPG) and gentiobiose, the former being the substrate of choice for routine measurement of enzyme activity. It has also been used as a substrate in the chromogenic screening of recombinant colonies expressing this enzyme. While many β-glucosidases are reported to be bound to the cell wall or involved in cell-surface-associated phenomena, the sequenced genes have not indicated any structural features normally associated with the cell-surface-bound proteins. A number of enzymes have been crystallized, particularly from family 1, and the structures are typical $(\beta/\alpha)_8$ barrel structures. The structure of a barley β-D-glucan exohydrolase (Exo I) of *Hordeum vulgare* [8], an enzyme related to family 3 enzymes based on sequence similarity, is known and also contains characteristic $(\beta/\alpha)_8$ barrel in one of the two domains. The other domain is an α/β sandwich structure.

In our search for enzymes that are thermo-tolerant and have glycosyltransferase activities that can be used for the enzymatic synthesis of various glycoconjugates, we have reported on the cloning and expression of two β-glucosidases, *Bgl*I [9] and *Bgl*II [10], from *Pichia etchellsii* into *Escherichia coli*. Both the enzymes have been well characterized [11,12] and *Bgl*II has been extensively used for synthesis of oligosaccharides [13], alkyl and terpene glucosides [14], and asparagine mimetics [15]. Two additional enzymes, BGLI and BGLII [16], have been purified from the cell wall of the yeast and their detailed properties described. All the four enzymes efficiently hydrolyzed methylumbelliferyl β-D-glucoside (MUG) and *p*NPG. In this paper, we describe, for the first time, isolation and properties of a novel enzymatic activity that resembled β-glucosidases in terms of its ability to hydrolyze MUG but displayed relatively very little hydrolytic activity on *p*NPG and had no sequence similarity with the generic β-glucosidases. Instead, sequence identity was seen with several members of the Ser-Asp repeat protein family which are cell-surface-associated. The purified protein was predicted to contain helical/coiled regions, which was confirmed by circular dichroism spectroscopy. Binding experiments were performed which indicated strong binding between the yeast cells and the purified protein. Investigations of this protein might further explain the role of β-glucosidases in cell-surface-associated phenomena.

Materials and methods

Strains, plasmids, and culture conditions. Pichia etchellsii (Deutsche Sammelung für Mikroorganismen, Germany) was used as the donor of β-glucosidase gene. *E. coli* DH5α cells were used as the host and pUC19 as cloning vector. *P. etchellsii* was maintained on YPD (0.5% yeast extract, 1% bacto-peptone, and 2% D-glucose) medium. *E. coli* was grown in LB at 37 °C according to standard protocols [17]. For *E. coli* transformants, ampicillin (amp) was added at 50 μg/ml to LB medium.

Cloning of β-glucosidase gene. Chromosomal DNA from *P. etchellsii* was prepared using zymolyase according to Pandey and Mishra [9]. The pUC19 plasmid was made by standard alkaline lysis method [17]. Partially *Sau*3AI digested chromosomal DNA of 5–20 kbp size was ligated to dephosphorylated pUC19 at the *Bam*HI site. The recombinant plasmids were transformed into competent *E. coli* DH5α cells and the cells were plated on LB containing amp, X-Gal, and IPTG. The white colonies were picked and further plated on LB + amp plates. Screening for β-glucosidase activity was done by overlaying the overnight grown colonies with 1 mM MUG (Sigma) and incubating the plates at 37 °C. Hydrolysis of MUG by the cloned β-glucosidase resulted in the release of 4-methylumbelliferrone (or methylumbelliferone) making the positive clones fluorescent under UV. MUG positive clones were tested further on *p*NPG plates. For this, the positive clones were streaked on LB + amp containing filter-sterilized *p*NPG at 4 mM. The colonies were incubated at 37 °C for 16–18 h and checked for the appearance of a yellow zone (*p*-nitrophenol liberation) around the clones. A positive and a negative control were included in all plate assays.

Sequence analysis and delineation of the correct ORF. Sequencing reactions on the plasmid pMG8 were carried out on Applied Biosystems automatic sequencing unit at Microsynth, Switzerland. The second strand sequencing was carried out at the DNA sequencing facility of the Department of Biochemistry, University of Delhi, South Campus. ORF search and sequence analysis were done using DNASTAR program. The correct reading frame was delineated by sub-cloning experiments. Based on identification of an overly expressed new protein of approximate 50 kDa molecular mass in the pMG8:DH5α clone, the possible initiating codon was identified, which was flanked by a unique *Pst*I site (shown in bold in Fig. 1) at the 5′end. The plasmid pMG8 was digested with *Pst*I and the ~2.85 kbp DNA fragment (containing the putative reading frame), of the three fragments generated by *Pst*I digest, was gel-eluted and ligated to *Pst*I digested/CIP treated pUC19 vector. The ligation mixture was used for transforming competent DH5α cells following the standard protocol [17]. The transformants were screened for MUG hydrolyzing activity as described above. The reading frame was further confirmed by digesting the plasmid pMG8 with *Eco*RI, which cut the putative ORF intragenically at a single site. The fragment containing the entire plasmid backbone and the 3′end (~50%) of the gene was self-ligated and transformed into competent DH5α cells. The plasmids were analyzed for the correct DNA changes and screened for MUG hydrolyzing activity.

Confirmation of origin of mugA by the PCR method. The origin of *mugA* insert in pMG8:DH5α was confirmed by PCR experiment. The chromosomal DNA of *P. etchellsii* was used as the template (1 μg) in PCR (100 μl) and the two, forward (5′-CCT CGG AAT CAG ACA CGG-3′) and reverse (5′-CCG TTT GGG CAT TTC CGC-3′) primers (underlined by arrows in Fig. 1), were used. The amplification conditions were as follows: initial denaturation at 94 °C/2 min, followed by 28 cycles of denaturation at 94 °C/30 s, annealing at 45 °C /30 s, and extension at 72 °C/3 min. A step of extension at 72 °C/10 min was carried out at the end. About 10 μl of the sample was run on 0.7% agarose gel along with the λ (*Eco*RI–*Hin*dIII cut) markers.

Sequence analysis and prediction of secondary structure. The putative sequence of the delineated protein, termed MUGA, was subjected to homology search using the National Centre for Biotechnological Information (NCBI) on-line program BLAST against protein (BlastX) and nucleotide (BlastN) sequences stored in GenBank. β-Glucosidase sequences of family 1 and family 3 were downloaded (http://afmb.cnrs-mrs.fr/~pedro/CAZY) and a sequence comparison was performed using pairwise and multiple sequence alignment program CLUSTAL W [18]. The sequence was analyzed for secondary structural elements by the 3D-PSSM program [19] available online. The sequence was also analyzed by the 'DAS' TM-segment prediction program for prediction of trans-membrane proteins (URL: http://www.sbc.su.se/~miklos/DAS/).

Purification of MUGA protein. For purification of intracellularly localized MUGA, recombinant E. coli cells were grown overnight at 37 °C in 1-L LB containing 50 μg/ml amp with vigorous shaking. Harvested cells were suspended in 1/10th of the original culture volume in Tris–Cl buffer, pH 8.3, and subjected to sonication (Soniprep) for 30 min at maximum output. The clear cell-free extract obtained after centrifugation (10,000*g*, 20 min) was subjected to differential ammonium sulfate precipitation between 30% and 80%, and the precipitate was dissolved in minimal volume of 20 mM Tris–Cl buffer, pH 8.3. Dialyzed and concentrated enzyme was loaded on a column (1 × 14 cm) containing DEAE–Sepharose equilibrated