Enzyme and Microbial Technology 99 (2017) 32–37

Contents lists available at ScienceDirect

Enzyme and Microbial Technology

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

Characterization and crystal structure of a thermostable glycoside hydrolase family 45 1,4-β-endoglucanase from *Thielavia terrestris*

Jian Gao[a,1], Jian-Wen Huang[b,1], Qian Li[a], Weidong Liu[a], Tzu-Ping Ko[c], Yingying Zheng[a], Xiansha Xiao[a], Chih-Jung Kuo[d], Chun-Chi Chen[a,*], Rey-Ting Guo[a,*]

[a] *Industrial Enzymes National Engineering Laboratory, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, China*
[b] *AsiaPac Biotechnology Co., Ltd, Dongguan, 523808, China*
[c] *Institute of Biological Chemistry, Academia Sinica, Taipei, 11529, Taiwan*
[d] *Department of Veterinary Medicine, National Chung Hsing University, Taichung, City 402, Taiwan*

ARTICLE INFO

Article history:
Received 29 August 2016
Received in revised form 1 December 2016
Accepted 15 January 2017
Available online 17 January 2017

Keywords:
Thielavia terrestris
Endoglucanase
Thermostable
Pichia pastoris
Crystal structure

ABSTRACT

1,4-β-Endoglucanase is one of the most important biocatalysts in modern industries. Here, a glycoside hydrolase (GH) family 45 endoglucanase from thermophilic fungus *Theilavia terrestris* (TtCel45A) was expressed in *Pichia pastoris*. The recombinant protein shows optimal activity at 60 °C, pH 4–5. The enzyme exhibits extraordinary thermostability that more than 80% activity was detected after heating at 80 °C for 2.5 h. The high resolution crystal structures of apo-form enzyme and that in complex with cellobiose and cellotetraose were solved to 1.36–1.58 Å. The protein folds into two overall regions: one is a six-stranded β-barrel, and the other one consists of several extended loops. Between the two regions lies the substrate-binding channel, which is an open cleft spanning across the protein surface. A continuous substrate-binding cleft from subsite −4 to +3 were clearly identified in the complex structures. Notably, the flexible V–VI loop (^{113}Gly-^{114}Gly-^{115}Asp-^{116}Leu-^{117}Gly-^{118}Ser) is found to open in the presence of −1 sugar, with D115 and L116 swung away to yield a space to accommodate the catalytic acid D122 and the $^{2,5}B$ boat conformation of −1 sugar during transition state. Collectively, we characterized the enzyme properties of *P. pastoris*-expressed TtCel45A and solved high-resolution crystal structures of the enzyme. These results are of great interests in industrial applications and provide new insights into the fundamental understanding of enzyme catalytic mechanism of GH45 endoglucanases.

1. Introduction

Cellulose, which represents the main structure of plant cell wall, is the most abundant renewable resource on earth. Cellulose is a high degree polymer, whose main chain is composed of 1,4-β-glycosidic linked D-glucoses. Cellulose decomposition has been found very useful in a wide variety of industrial applications, including biofuel production, food and feed manufacture, laundry, and textile processing [1]. Using cellulases to achieve cellulose decomposition is more and more popular owing to good efficiency, eco-friendliness, and good specificity [2]. At least three groups of enzymes are included in cellulase family: *endo*-1,4-β-*endo*glucanse (EC 3.2.1.4, endoglucanase), cellobiohydrolase (EC 3.2.1.91), and β-glucosidase (EC 3.2.1.21). Among them, endoglucanse, which catalyzes random hydrolysis of 1,4-β-glycosidic bonds to cleave cellulose main chain into smaller fragments, is considered a key biocatalyst in cellulose degradation and is the most widely developed cellulase product.

A great number of endoglucanase have been identified, most of which are from bacteria and fungi. Based on the Carbohydrate-Active Enzyme database (CAZy), endoglucanases are categorized into 13 glycoside hydrolase (GH) families based on protein sequence similarity and structure of the catalytic domain [3] (http://www.cazy.org). GH5, GH7, GH12, GH44, and GH51 enzymes catalyze via a retaining mechanism, while those of GH6, GH8, GH9, GH45, GH48, GH74, GH124, and GH131 via an inverting mechanism. Unlike the enzymes in most GH families, where similar protein fold is shared by various enzymes with distinct substrate specificities, GH45 and GH48 members are exclusively endoglucanases. Moreover, GH45 endoglucanases have been found to exhibit higher activity on amorphous cellulose and possess a great potential in commercial applications [4]. Therefore, functional and

* Corresponding authors.
E-mail addresses: ccckate0722@gmail.com (C.-C. Chen), guo_rt@tib.cas.cn (R.-T. Guo).
[1] JG and JWH contribute equally.

http://dx.doi.org/10.1016/j.enzmictec.2017.01.005
0141-0229/

structural studies of GH45 enzymes are of interest to both academic and industrial areas.

GH45 enzymes fold into a six-stranded β-barrel and harbor an open substrate-binding cleft spanning across the protein surface. Based on the relative position of the β-barrel and substrate-binding cleft, currently solved GH45 structures can be subdivided into two groups. Group I contains enzymes from the *Humicola* species and *Melanocarpus albomyces*, and group II contains those from the blue mussel *Mytilus edulis* and *Phanerochaete chrysosporium* (Fig. S1). In group I enzymes, the substrate-binding cleft is formed between the β-barrel and a region consisting of several long loops [4,5], while that in group II enzymes is formed mainly by the β-barrel [6]. In group I, two catalytic aspartic acids located near the center of the substrate-binding channel serve as catalytic acid and base. Notably, previous structural analyses indicate that a flexible loop between β-strand V and VI of group I enzymes show variable conformations in apo- and substrate-bound structure [4,5,7]. The V–VI loop is considered to serve as a lid that flips to close the enzyme active center to maintain an environment suitable for catalytic reaction.

Structural stability and catalytic activity at elevated temperatures are amongst the most appreciated properties of an industrial enzyme. Thermophilic organisms are notable sources of heat stable and thermophilic enzymes. *Thielavia terrestris* is a soil-born thermophilic ascomycete fungus, which secretes an array of biomass-degrading enzymes [8,9]. Although a limited number of cellulolytic enzymes from *T. terretris* have been characterized, these enzymes are highly efficient, thermostable, and of great application potentials [10,11]. Here, the gene encoding the catalytic domain of TtCel45A is expressed in *Pichia pastoris* and the enzyme property of the recombinant protein is examined. Furthermore, the crystal structure of TtCel45A and complex structure with oligosaccharides were solved at high resolution to reveal the protein folding and substrate-binding pattern. These results provide an important basis for further applications of TtCel45A and shall expand our understanding in molecular mechanism of GH45 endoglucanases.

2. Materials and methods

2.1. Cloning and expression of TtCel45A in P. pastoris

The gene encoding catalytic domain (residue 22–222) of TtCel45A (GenBank access No. XP_003651003.1) was chemically synthesized and then cloned into pPICZαA vector (Invitrogen). The pPICZαA/TtCel45A plasmid DNA was amplified in *Escherichia coli* cells and linearized by restriction enzyme *Pme*I prior to transformation into *P. pastoris* via electroporation. The transformants were selected on YPD plates (1% yeast extract, 2% peptone, 2% dextrose, 2% agar) containing 100 μg/ml zeocin (Invitrogen) and incubated at 30 °C for two days. Single colonies were inoculated in 5 ml YPD medium and then amplified in 50 ml BMGY (1% yeast extract, 2% peptone, 100 mM potassium phosphate pH 6.0, 1.34% yeast nitrogen base with ammonium sulfate without amino acid, 4×10^{-5}% biotin and 1% glycerol) at 30 °C for 24 h. Cells were harvested by centrifugation and resuspended in 50 ml BMMY (1% yeast extract, 2% peptone, 100 mM potassium phosphate pH 6.0, 1.34% yeast nitrogen base with ammonium sulfate without amino acid, 4×10^{-5}% biotin, and 0.5% methanol) and incubated at 30 °C for 24 h. Afterward, 1% methanol is supplemented every 24 h for protein induction for 4-5 days.

2.2. Protein purification

Culture supernatant of *P. pastoris* transformant containing 1.0 M ammonium sulfate was applied to a phenyl sepharose column (GE Healthcare) which was equilibrated by buffer containing 25 mM Tris, 1.0 M ammonium sulfate, pH 7.5. The column was washed by a gradient forming by 1.0–0 M ammonium sulfate. Fractions containing target protein were collected, and then dialyzed twice against 5 l of buffer containing 25 mM Tris, pH 7.5. The protein was further purified by diethylaminoethyl (DEAE) column by using a gradient formed by 25 mM Tris, pH 7.5 and 0–400 mM NaCl. Target protein was collected, dialyzed against Tris buffer as mentioned above, and concentrated for crystallization trials. The protein purity was verified by SDS-PAGE analysis.

2.3. Enzyme activity measurement

The cellulase activity was determined by dinitrosalicylic acid (DNS) method as previous described with minor modifications [12]. In this study, equal amounts of enzyme solution (50 mM sodium citrate buffer, pH 4.5) and 1% (w/v) carboxy-methyl cellulose (CMC) were mixed and incubated in a water bath at 50 °C for 15 min. The reaction was then mixed with DNS and incubated in boiling water for 5 min to remove residual enzyme activity. After cooling in cold water bath for 5 min, the absorbance of 540 nm was measured for calculation of the enzyme activity. One unit of activity is defined as the amount of enzyme that releases 1 μ mole product per minute under the assay conditions.

The optimal temperature and pH of TtCel45A were determined at different temperatures (from 50 to 80 °C) or in different buffers (pH 3–8) for 15 min. The buffer systems were 0.1 M sodium citrate (pH 3–6) and 0.1 M potassium phosphate (pH 6–8). Enzyme activity was determined as described above. For thermostability analysis, normalized TtCel45A proteins were incubated at 80 °C or 90 °C for different time periods prior to activity measurements.

2.4. Crystallization and data collection

Purified recombinant TtCel45A protein was crystallized by using sitting-drop vapor diffusion method and Crystal Screen kits (Hampton Research). Crystals were grown in a reservoir solution containing 0.2 M $MgCl_2$, 0.1 M Tris pH 7.5, and w/v 30% PEG4000 at room temperature for five months. The TtCel45A-CBI crystals were obtained by co-crystallization of TtCel45A and 10 mM cellobiose (Megazyme), and the TtCel45A-CTT crystal were obtained by soaking the TtCel45A-CBI crystals in mother liquid containing 10 mM cellotetraose (Megazyme) at room temperature for 1 h. X-ray diffraction data sets were collected on beam lines BL15A1 and BL13B1 of National Synchrotron Radiation Research Center (NSRRC) in Hsinchu, Taiwan. The diffraction images were processed by using HKL2000 [13], and 5% randomly selected reflections were set aside for calculating R_{free} as a monitor prior to structure refinements.

2.5. Structure determination and refinement

The crystal structure of apo-form TtCel45A was solved by molecular replacement method with Phaser MR program [14] using a hypothetical model generated by Phyre2 web portal [15] using *M. albomyces* GH45 endoglucanase structure as a template (PDB ID, 1L8F). Structure refinement was carried out by using Coot [16] and Refmac5 [17], and water molecules were modeled according to 1.0 σ map level. The complex structures TtCel45A-CBI and TtCel45A-CTT were solved by using molecular replacement method using TtCel45A apo-form structure as a searching model. Subsequent refinements were performed as described above. All data collection and refinement statistics are summarized in Table 1, and all structure figures were prepared by using PyMOL (http://pymol.sourcefore.net/).