Bioresource Technology 102 (2011) 5988–5994

Contents lists available at ScienceDirect

Bioresource Technology

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

Enzymatic hydrolysis of cellulose by the cellobiohydrolase domain of CelB from the hyperthermophilic bacterium *Caldicellulosiruptor saccharolyticus*

Joshua I. Park [a,b], Michael S. Kent [a,d], Supratim Datta [a,b], Bradley M. Holmes [a,b], Zhaohua Huang [e], Blake A. Simmons [a,b], Kenneth L. Sale [a,c], Rajat Sapra [a,b,*]

[a] *Deconstruction Division, Joint BioEnergy Institute, Emeryville, CA 94608, USA*
[b] *Department of Biomass Science and Conversion Technology, Sandia National Laboratories, Livermore, CA 94551, USA*
[c] *Biosystems R&D Department, Sandia National Laboratories, Livermore, CA 94551, USA*
[d] *Sandia National Laboratories, Albuquerque, NM 87123, USA*
[e] *Departments of Bioengineering, Therapeutic Sciences, and Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94143, USA*

ARTICLE INFO

Article history:
Received 27 September 2010
Received in revised form 5 February 2011
Accepted 8 February 2011
Available online 13 February 2011

Keywords:
Caldicellulosiruptor saccharolyticus
Cellulase
Cellobiohydrolase
Carbohydrate binding module
Hyperthermophile

ABSTRACT

The *celB* gene of *Caldicellulosiruptor saccharolyticus* was cloned and expressed in *Escherichia coli* to create a recombinant biocatalyst for hydrolyzing lignocellulosic biomass at high temperature. The GH5 domain of CelB hydrolyzed 4-nitrophenyl-β-D-cellobioside and carboxymethyl cellulose with optimum activity at pH 4.7–5.5 and 80 °C. The recombinant GH5 and CBM3-GH5 constructs were both stable at 80 °C with half-lives of 23 h and 39 h, respectively, and retained >94% activity after 48 h at 70 °C. Enzymatic hydrolysis of corn stover and cellulose pretreated with the ionic liquid 1-ethyl-3-methylimidazolium acetate showed that GH5 and CBM3-GH5 primarily produce cellobiose, with product yields for CBM3-GH5 being 1.2- to 2-fold higher than those for GH5. Confocal microscopy of bound protein on cellulose confirmed tighter binding of CBM3-GH5 to cellulose than GH5, indicating that the enhancement of enzymatic activity on solid substrates may be due to the substrate binding activity of CBM3 domain.

Published by Elsevier Ltd.

1. Introduction

Recently there has been renewed interest liberating fermentable sugars from pretreated lignocellulosic biomass by utilizing mixtures of enzymes that hydrolyze biomass for ethanol and other biofuels production (Banerjee et al., 2010; Gao et al., 2010). Employing active biomass hydrolyzing enzymes at extremes of temperature and pH may be advantageous for industrial scale production of fermentable sugars from lignocellulosic biomass, because these conditions facilitate overcoming biomass recalcitrance and prevent the growth of contaminating microorganisms (Blumer-Schuette et al., 2008). Cellulases, which include endoglucanases (EC 3.2.1.4), cellobiohydrolases (EC 3.2.1.91), and β-glucosidases (3.2.1.21), are members of the glycoside hydrolase (GH) families of enzymes that catalyze the cleavage of β-1,4-glycosidic bonds of cellulose to glucose (Henrissat and Davies, 1997; Lynd et al., 2002). In nature, these enzymes are found in microorganisms that thrive in extreme environments (Sunna et al., 1997; Bergquist et al., 1999; Blumer-Schuette et al., 2008); therefore, cellulases from extremophiles are attractive as potential biocatalysts for industrial depolymerization of cellulose.

Caldicellulosiruptor saccharolyticus (*C. sac*) is a thermophilic anaerobic bacterium first isolated from thermal springs in New Zealand (Sissons et al., 1987). *C. sac* is capable of growing on cellulose as its sole carbon source at an optimum growth temperature of 70 °C, presumably by secreting the cellulolytic enzymes required to produce glucose for survival and growth (Rainey et al., 1994). Recently, it was reported that *C. sac* also grows on insoluble switchgrass at 70 °C without the need for any other carbon sources (Yang et al., 2009). Genomic analysis of *C. sac* revealed the presence of several gene clusters encoding multi-domain hydrolases, including cellulases and xylanases (Gibbs et al., 2000; van de Werken et al., 2008). To date, both *celA* and *celB* from *C. sac* have been confirmed as bi-functional cellulolytic genes (Saul et al., 1990; Te'o et al., 1995). However, detailed biochemical studies have not been reported for either recombinant endoglucanases or cellobiohydrolases from *C. sac*, presumably due to the difficulty of producing sufficient quantities of recombinant proteins. Previously, Saul and

Abbreviations: [C2mim][OAc], 1-ethyl-3-methylimidazolium acetate; *C. sac*, *Caldicellulosiruptor saccharolyticus*; CBM, carbohydrate binding module; CMC, carboxymethyl cellulose; CV, column volume; GH, glycoside hydrolase; *p*NPC, 4-nitrophenyl-β-D-cellobioside; *p*NPG, 4-nitrophenyl-β-D-glucopyranoside.

* Corresponding author. Address: 5885 Hollis Street 4th Floor, Emeryville, California 94608, USA. Tel.: +1 510 508 9956; fax: +1 510 486 4252.
E-mail addresses: rajatsapra@gmail.com, rsapra@lbl.gov (R. Sapra).

0960-8524/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.biortech.2011.02.036

coworkers reported that *C. sac* has a multi-domain gene, *celB*, containing a carbohydrate binding module (CBM3) flanked by two putative GH10 and GH5 domains in a 3120 bp-long open reading frame (Saul et al., 1990). Based on the enzyme activities of truncated *celB* constructs on soluble model substrates, CelB was annotated as a bi-functional cellulase that has both cellobiohydrolase (GH10) and endoglucanase (GH5) activities (Saul et al., 1990). The *celB* gene also contains an N-terminal putative signal peptide sequence (amino acid residues 1–36) without a putative transmembrane domain, which suggests that CelB is an extracellular enzyme excreted to hydrolyze cellulosic materials outside of the cell membrane. Recent proteomic analysis of the *C. sac* secretome identified CelB as one of the two extracellular cellulases (Muddiman et al., 2010); therefore, CelB may be a key extracellular enzyme for degradation of polysaccharides for survival and growth of *C. sac*.

Since the earlier study of CelB showed enzymatic activity at high temperature (Saul et al., 1990), expression and purification of the functional domain of CelB were pursued with the codon-optimized gene to create a recombinant biocatalyst. A total of five recombinant CelB constructs consisting of combinations of the three functional domains (GH10-CBM3-GH5, GH10-CBM3, CBM3-GH5, GH10, and GH5) were expressed in *Escherichia coli*, but only CBM3-GH5 and GH5 expressed as soluble proteins. Therefore, this study was aimed at investigating the influence of the CBM3 domain on the thermostability and the enzymatic activity of the GH5 domain on CMC, microcrystalline cellulose (Avicel), Avicel pretreated with 1-ethyl-3-methylimidazolium acetate ([C2mim][OAc]), corn stover, and corn stover pretreated with [C2mim][OAc].

2. Methods

2.1. Cloning truncated celB constructs for protein expression in E. coli

The amino acid sequence of CelB from *C. sac* (UniProt ID: A4XIF7) containing GH10, CBM3, and GH5 domains was obtained from the UniProt database (www.uniprot.org). Based on the amino acid sequence of CelB, the codon-optimized gene was synthesized for protein expression in *E. coli* (GenScript, Piscataway, NJ; Supplementary Fig. 1). The domain boundaries in the *celB* gene were defined by Pfam annotations (http://pfam.sanger.ac.uk) to design PCR primers for creating truncated constructs of *celB* (GH10-CBM3-GH5, GH10-CBM3, CBM3-GH5, GH10, and GH5). All PCR primers included the attB1 and attB2 recombination sites for Gateway cloning (Invitrogen, Carlsbad, CA). The amplicons were inserted into the pDONR221 plasmid to create entry clones (Supplementary Fig. 2). The entry clones in the pDONR221 vector were then recombined with the pET DEST42 vector to create constructs for protein expression in *E. coli*. All constructs in the pET DEST42 vector contained the nucleotide sequence for the c-terminal V5 epitope and His (×6) tags. The nucleotide sequence of each construct was verified by DNA sequencing (Quintara, Berkeley, CA).

2.2. Protein expression and purification

Small scale protein expression of truncated *celB* in *E. coli* and cellulase activity screenings on the soluble substrate carboxymethyl-cellulose (CMC) were performed as described previously (Allgaier et al., 2010). For larger scale protein expression of recombinant CBM3-GH5 or GH5, 10 mL of overnight culture was inoculated into 1 L of fresh LB medium containing 50 μg/mL of carbenicillin. The cells were grown at 37 °C with constant shaking at 200 RPM. Protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG), to a final concentration of 0.5 mM, to the cell culture at an OD_{600nm} between 0.6 and 0.9, and then the cells were grown at 30 °C for overnight with constant shaking at 200 RPM. The cells containing the expressed proteins were harvested by centrifugation at 6,000 g for 15 min at 4 °C. The cell pellet was resuspended in 20 mL of lysis buffer (300 mM NaCl, 20 mM Tris–Cl, 20 mM imidazole, pH 8, 0.2% Triton X-100, and 1 mM phenylmethylsulfonylfluoride) per L culture, and then stored at −80 °C until the frozen cell pellet was ready for protein purification.

To purify the recombinant proteins, the frozen cell suspension was thawed at room temperature, and lysozyme (1 mg/mL) and a protease inhibitor cocktail (Calbiochem, Gibbstown, NJ) were added to lyse the cells. The cell suspension was then incubated on ice for 30 min. Then the cell lysate was sonicated with three cycles of 30 s-pulse. The nucleic acids in the cell lysate suspension were digested by adding Benzonase (25 U/ml; Novagen, Gibbstown, NJ) and incubating at room temperature for 30 min, and the cell lysates were heated at 70 °C for 30 min to precipitate the native heat-labile *E. coli* proteins. The soluble extract containing the recombinant CBM3-GH5 or GH5 was obtained by centrifugation at 12,000 g for 30 min at 4 °C. The soluble extract was collected, and 10 mM β-mercaptoethanol was added to the extract. The resulting soluble extract was filtered through 0.45 μm membrane and the filtrate was used for purification of the recombinant proteins. The His(×6)-tagged recombinant CBM3-GH5 or GH5 was purified using an ÄKTA Explorer (GE Healthcare Life Sciences, Piscataway, NJ). A HisTrap FF Ni^{2+}-NTA affinity column (GE Healthcare) was pre-equilibrated with a wash buffer containing 300 mM NaCl, 20 mM Tris–Cl, and 20 mM imidazole, pH 8.0 prior to loading the soluble extract onto the column. The flow-rate was held at constant rate of 1 mL/min for all subsequent steps. The column was washed with 10 column volumes (CV) of the wash buffer. A linear gradient was applied to the column with elution buffer containing 300 mM NaCl, 20 mM Tris–Cl, 500 mM imidazole, pH 8.0 over 20 CV. The eluates containing the recombinant protein were pooled, and the buffer was exchanged to 20 mM Tris–Cl, pH 8.0 using an Econo-Pac DG-10 column (Bio-Rad, Hercules, CA). The buffer exchanged protein was loaded onto a HiTrap Q HP anion-exchange column (GE Healthcare) that was pre-equilibrated with 20 mM Tris–Cl, pH 8.0. Then a linear gradient was applied to the column with 1 M NaCl, 20 mM Tris–Cl, pH 8.0 over 20 CV to elute the bound protein. The collected fractions were analyzed by SDS-PAGE. The protein concentrations were determined by Bradford assay (Bio-Rad).

2.3. Cellulase activity screening and enzyme assays

Cellulase activity of the *celB* constructs was tested by Congo red assay on CMC-agar plates (Teather and Wood, 1982). The soluble extracts from the heat-treated cell lysates, which had been incubated at 75 °C for 30 min as described above, containing expressed recombinant proteins were spotted on CMC-agar plate (0.2% CMC and 1.5% agar), and then the plate was incubated at 37 °C for 4 h. After incubation, the plate was rinsed with 70% of ethanol, and was stained by the Congo red dye solution (0.1% in water) for 1 h. Unbound dye was washed away with 1 M NaCl until the cleared-zones were visible. CMC was also used to determine the optimum pH (pH_{opt}) and temperature (T_{opt}) of CBM3-GH5 and GH5 activities. A buffer cocktail containing sodium acetate (NaOAc), 2-(*N*-morpholino) ethanesulfonic acid (MES), and 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) was prepared for pH-dependent enzyme activity measurement as described previously (Allgaier et al., 2010). The enzyme activity of CBM3-GH5 or GH5 was measured by mixing 1 μg of enzyme and 1% CMC in 50 μL of total volume of reaction containing 100 mM NaOAc, 50 mM MES, and 50 mM HEPES at various pH. After enzyme reaction at T_{opt}, the 3,5-dinitrosalicylic acid (DNS)