Appl Microbiol Biotechnol (2010) 86:525–533
DOI 10.1007/s00253-009-2231-1

BIOTECHNOLOGICALLY RELEVANT ENZYMES AND PROTEINS

The noncellulosomal family 48 cellobiohydrolase from *Clostridium phytofermentans* ISDg: heterologous expression, characterization, and processivity

Xiao-Zhou Zhang · Zuoming Zhang · Zhiguang Zhu · Noppadon Sathitsuksanoh · Yunfeng Yang · Y.-H. Percival Zhang

Received: 21 July 2009 / Revised: 29 August 2009 / Accepted: 31 August 2009 / Published online: 15 October 2009

Abstract Family 48 glycoside hydrolases (cellobiohydrolases) are among the most important cellulase components for crystalline cellulose hydrolysis mediated by cellulolytic bacteria. Open reading frame (Cphy_3368) of *Clostridium phytofermentans* ISDg encodes a putative family 48 glycoside hydrolase (CpCel48) with a family 3 cellulose-binding module. CpCel48 was successfully expressed as two soluble intracellular forms with or without a C-terminal His-tag in *Escherichia coli* and as a secretory active form in *Bacillus subtilis*. It was found that calcium ion enhanced activity and thermostability of the enzyme. CpCel48 had high activities of 15.1 U μmol^{-1} on Avicel and 35.9 U μmol^{-1} on regenerated amorphous cellulose (RAC) with cellobiose as a main product and cellotriose and cellotetraose as by-products. By contrast, it had very weak activities on soluble cellulose derivatives (e.g., carboxymethyl cellulose (CMC)) and did not significantly decrease the viscosity of the CMC solution. Cellotetraose was the smallest oligosaccharide substrate for CpCel48. Since processivity is a key characteristic for cellobiohydrolases, the new initial false/right attack model was developed for estimation of processivity by considering the enzyme's substrate specificity, the crystalline structure of homologous Cel48 enzymes, and the configuration of cellulose chains. The processivities of CpCel48 on Avicel and RAC were estimated to be ~3.5 and 6.0, respectively. Heterologous expression of secretory active cellobiohydrolase in *B. subtilis* is an important step for developing recombinant cellulolytic *B. subtilis* strains for low-cost production of advanced biofuels from cellulosic materials in a single step.

Keywords *Clostridium phytofermentans* · Cellobiohydrolase · Family 48 exoglucanase · *Bacillus subtilis* · Characterization · Processivity

X.-Z. Zhang · Z. Zhang · Z. Zhu · N. Sathitsuksanoh · Y.-H. P. Zhang
Department of Biological Systems Engineering,
Virginia Polytechnic Institute and State University,
210-A Seitz Hall,
Blacksburg, VA 24061, USA

Z. Zhang
Key Laboratory for Molecular Enzymology and Engineering of the Ministry of Education, Jilin University,
Changchun 130023, China

N. Sathitsuksanoh · Y.-H. P. Zhang
Institute for Critical Technology and Applied Science (ICTAS),
Virginia Polytechnic Institute and State University,
Blacksburg, VA 24061, USA

Y. Yang
Biosciences Division, Oak Ridge National Laboratory,
Oak Ridge, TN 37831, USA

Y. Yang · Y.-H. P. Zhang (✉)
DOE BioEnergy Science Center (BESC),
Oak Ridge, TN 37831, USA
e-mail: biofuels@vt.edu

Introduction

The production of cellulosic ethanol from nonfood lignocellulosic biomass would bring benefits, such as promoting rural economy, decreasing greenhouse gas emissions, and enhancing national energy security (Lynd et al. 2008; Zhang 2008). In order to decrease cellulase use, increase volumetric productivity, and reduce capital investment, consolidated bioprocessing (CBP) has been proposed by integrating cellulase production, cellulose

hydrolysis, and ethanol fermentation in a single step (Lynd et al. 2002, 2008). *Clostridium phytofermentans* ISDg is an obligately anaerobic, mesophilic Gram-positive bacterium (Warnick et al. 2002). This bacterium can utilize a broad range of carbon sources, including polysaccharides (e.g., cellulose, xylan, pectin, polyglacturonic acid, starch), oligosaccharides (e.g., cellobiose, gentibiose, lactose, maltose), and monomeric sugars (e.g., glucose, fructose, mannose, galactose, xylose, and ribose) for generation of ethanol, acetate, CO_2, and hydrogen (Warnick et al. 2002).

Clostridium thermocellum and *C. phytofermentans* have been regarded as model microorganisms for CBP (Lu et al. 2006; Warnick et al. 2002; Zhang and Lynd 2005a). *C. thermocellum* produces a complexed cellulase system, called cellulosome (Bayer et al. 2004; Doi and Kosugi 2004). The association of cellulosome on the surface of the microbe acting on cellulose has been demonstrated to enhance microbial cellulose hydrolysis rates by approximately two- to fivefold as compared to cell-free enzymatic hydrolysis (Lu et al. 2006). In contrast to the complexed cellulase systems of *C. thermocellum*, *C. phytofermentans* appears not to produce complexed cellulases due to the lack of the scaffolding, cohesin, and dockerin sequences at http://genome.jgi-psf.org/finished_microbes/clopi/clopi.home.html. This observation raised a question—how *C. phytofermentans* hydrolyzes cellulosic materials efficiently without cellulosomes. Based on the carbohydrate-active enzymes (CAZy) database (Cantarel et al. 2009), the *C. phytofermentans* genome contains more than 100 open reading frames (ORFs) encoding putative glycoside hydrolases. However, none of them have been cloned, purified, and characterized.

The glycoside hydrolase family 48 exoglucanases are widely believed to play important roles in crystalline cellulose hydrolysis mediated by bacterial cellulase systems. Their role is believed to be somewhat similar to that of the *Trichoderma* CBH1 (Cel7A) (Teeri 1997; Zhang et al. 2006b). Family 48 exoglucanases work as dominant catalytic components of cellulosomes, such as *Clostridium cellulolyticum* CelF (Reverbel-Leroy et al. 1997), *Clostridium cellulovorans* ExgS (Liu and Doi 1998), *Clostridium jusui* CelD (Kakiuchi et al. 1998), and *C. thermocellum* CelS (Kruus et al. 1995; Wang et al. 1994), or as key noncomplexed cellulase components, such as *Cellulomonas fimi* CbhB (Shen et al. 1995), *Clostridium stercorarium* Avicelase II (Bronnenmeier et al. 1991), *Thermobifida fusca* Cel48 (Irwin et al. 2000), *Paenibacillus barcinonensis* BP-23 Cel48C (Sánchez et al. 2003), and *Ruminococcus albus* 8 Cel48A (Devillard et al. 2004). For example, *C. thermocellum* CelS accounts for ~30% of the weight of the cellulosome isolated from the Avicel-grown culture but its level decreases to ~10% of the weight of that isolated from cellobiose-grown culture, implying that CelS plays a key role for crystalline cellulose hydrolysis (Bayer et al. 1985; Zhang and Lynd 2005c).

The processivity of the exoenzyme is defined as the number of catalytic events during the association of the enzyme with the polymeric substrate before its dissociation (Breyer and Matthews 2001; Von Hippel et al. 1994). The extent of processivity of enzymes (e.g., exoglucanases, λ-exonucleases, 20S proteasome, T7 DNA polymerases, yeast RNA polymerase II) ranges from ~20 to more than 20,000 (Breyer and Matthews 2001). Cellobiohydrolases are thought to guide detached single-polymer chains through long and deep active-site clefts, leading to processive (stepwise) degradation of the cellulose (Teeri et al. 1998). Many studies of the processivity of cellobiohydrolase have been conducted based on fungal *Trichoderma reesei* Cel7A (Kipper et al. 2005; Medve et al. 1998; von Ossowski et al. 2003). The model of false/right attacks estimated that the processivity values for *T. reesei* Cel7A on Avicel and bacterial microcrystalline cellulose were estimated to be ca. 5–10 (Medve et al. 1998) and 23 (von Ossowski et al. 2003), respectively. Another model based on the burst kinetics implied that the processivity extents of *T. reesei* Cel7A were 88, 42, and 34 on labeled bacterial cellulose, bacterial microcrystalline cellulose, and endoglucanase-pretreated bacterial cellulose, respectively (Kipper et al. 2005). However, no study has yet been made to estimate the processivity of bacterial glycoside hydrolase family 48 enzymes.

Here, we cloned the *cpcel48* gene, expressed CpCel48 in *E. coli* or *B. subtilis*, purified, and characterized them. In addition, we developed a new false/right initial attack model for estimating the processivity of CpCel48 on Avicel and regenerated amorphous cellulose (RAC).

Materials and methods

Chemicals

All chemicals were reagent grade or higher, purchased from Sigma (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA), unless otherwise noted. RAC was prepared from Avicel, as previously described (Zhang et al. 2006a). Cellodextrins were prepared by hydrolysis of a mixture of concentrated HCl/H_2SO_4 and separated by chromatography (Zhang and Lynd 2003). The oligonucleotides were synthesized by Integrated DNA Technologies (Coraville, IA).

Strains and media

Escherichia coli DH5α was used as a host cell for DNA manipulation, and *E. coli* BL21 Star (DE3) (Invitrogen, Carlsbad, CA) or *B. subtilis* WB800 (Wu et al. 2002) were

used as the host for recombinant protein expression. *B. subtilis* transformation was performed by the competent cell method (Anagnostopoulos and Spizizen 1961). The Luria–Bertani (LB) medium was used for *E. coli* cell growth and recombinant protein expression. The LB/glucose (LBG) and M9/glucose (M9G) media denotes the LB or M9 media supplemented with 0.4% (*w*/*v*) glucose. Ampicillin (100 μg ml^{-1}) and kanamycin (20 μg ml^{-1}) were added in the *E. coli* and *B. subtilis* media, respectively.

Construction of plasmids

The DNA sequence encoding the mature CpCel48 (34-919 amino acids; GenBank accession number ABX43721) were amplified from the genomic DNA of *C. phytofermentans* ISDg (ATCC accession number 700394) by a primer pair of P1 (5′-GGAAT TCCAT ATGGG TGAAA CTGAG CAAGC-3′, *Nde*I site underlined) and P2 (5′-GCCGC TCGAG TGGTT CGATA CCCCA-3′, *Xho*I site underlined). The polymerase chain reaction (PCR) product was digested with *Nde*I/*Xho*I and ligated into the corresponding sites of the vector pET20b (Novagen) to generate pET20b-Cpcel48 for expression of CpCel48 with a C-terminal His-tag. For expression of CpCel48′ without the His-tag, another primer pair of P1 and P3 (5′-CGACT CTCGA GTTTA TGGTT CGATA CCCCA ATTAA GT-3′, *Xho*I site underlined) was used to amplify the *cpcel48* gene, and the PCR product after double digestion was cloned into the *Nde*I/*Xho*I sites of pET20b to generate pET20b-Cpcel48′. For expression of secretory CpCel48′ in *B. subtilis*, the *cpcel48* gene was amplified with a primer pair P4 (5′-CTAGC TGCAG GTGAA ACTGA GCAAG C-3′, *Pst*I site underlined) and P5 (5′-GATAC CCGGG TTTTA TGGTT CGATA CCCCA-3′, *Xma*I site underlined). The PCR product was digested with *Pst*I/*Xma*I and cloned into the corresponding sites of the *E. coli–B. subtilis* shuttle vector pP43NMK (Zhang et al. 2005) to generate pP43N-Cpcel48′, where the *cpcel48* gene without its signal peptide-encoding fragment was fused with a NprB signal peptide-encoding sequence.

Enzyme expression and purification

The strains *E. coli* BL21(pET20b-Cpcel48) or BL21 (pET20b-Cpcel48′) were cultivated in the LB medium supplemented with 1.2% glycerol at 37 °C for expression of CpCel48 or CpCel48′, respectively. When A_{600} reached ~2.0, 50 μmol l^{-1} isopropyl-beta-D-thiogalactopyranoside was added, and the temperature was decreased to 18 °C for ~16 h. For purification of CpCel48 by the nickel–resin column, the cell pellet was resuspended in a 50 mmol l^{-1} phosphate buffer (pH 8.0) containing 5 mmol l^{-1} imidazole and 300 mmol l^{-1} NaCl. The cells were lysed by ultrasonication. The supernatant of the crude cell extract was applied to the nickel column, and then the column was washed in a 50-mM phosphate buffer (pH 8.0) with 5 mM imidazole and 300 mmol l^{-1} NaCl. The bound CpCel48 was eluted by a 50 mmol l^{-1} phosphate buffer (pH 8.0) with 100 mmol l^{-1} imidazole and 300 mmol l^{-1} NaCl. The enzyme was desalted by a PD-10 desalting column (Amersham) and then stored at 4 °C in a 50 mM phosphate buffer (pH 8.0) containing 50 mmol l^{-1} NaCl and 5% glycerol. For purification of CpCel48′ on RAC, the cell pellets were resuspended with a 50-mmol l^{-1} Tris-HCl buffer (pH 8.0). After ultrasonication and centrifugation, RAC was added for adsorbing the target protein at room temperature. After washing once with deionized water, the CpCel48′ was desorbed by using 80% ethylene glycol (Hong et al. 2008). For expression of secretory CpCel48′ in *B. subtilis*, the strain WB800 (pP43N-Cpcel48′) was cultivated in the LBG medium at 37 °C. The secretory CpCel48 in the culture was precipitated by trichloroacetic acid (TCA) for qualitative analysis by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) or purified by affinity adsorption on RAC.

Enzyme assays

Enzyme activity was assayed at 50 °C by using initial reaction rates within 20 min rather than a fixed conversion within a fixed time (Zhang et al. 2009). The reaction mixtures (0.5 ml) contained 1% (*w*/*v*) of the substrate (e.g., Avicel, RAC, carboxymethyl cellulose (CMC)) in a 50-$mmol^{-1}$ MES buffer (pH 6.0) containing 1 mmol l^{-1} $CaCl_2$. Enzyme concentration in the reactions was 10 μg ml^{-1} (~0.1 μmol l^{-1}), unless otherwise noted. The reactions were terminated by boiling for 5 min. After centrifugation, aliquots of the supernatants were assayed for the release of the reducing sugars. Concentration of reducing sugars was determined by the 2,2′-bicinchoninate method (Zhang and Lynd 2005b). One unit of enzyme was defined as 1 μmol of reducing ends generated per minute

Viscosity determination

The reactions containing the purified enzyme (0.01 μg CpCel5B ml^{-1} or 10 μg CpCel48 ml^{-1}), and 1% low-viscosity CMC were incubated at 50 °C. The samples were withdrawn at different time points, boiled for 5 min, and then diluted tenfold by deionized water. The viscosity of the CMC solutions was determined based on the time of outflow at 23 °C by using an Ubbelohde viscometer tube (Fisher Scientific).

Chromatographic assays

The enzymatic samples were separated on a Whatman 60 Å silica gel plate with ethyl acetate–H_2O–methanol (8:3:4, *v*/*v*)

(Zhang and Lynd 2006). The soluble cellodextrins released from cellulosic materials were measured by using a Beckman high-performance liquid chromatography (HPLC) equipped with a Bio-Rad HPX-42A column (Richmond, CA) and a refractive index detector at a flow rate of 0.4 mL/min (Zhang and Lynd 2003).

Other assays

Protein concentration was measured by the Bradford assay (Bradford 1976). Protein samples were separated by 10% SDS-PAGE. Using bovine serum albumin as the standard, the amount of CpCel48′ in the SDS-PAGE gel was estimated with the densitometry analysis software Quantity One (version 4.4.0, Bio-Rad, Hercules, CA).

Results

Sequence analysis

The CAZy database predicts that the ORF Cphy_3368 (GenBank accession number ABX43721) encodes a putative glycoside hydrolase family 48 enzyme (CpCel48), which is clustered with a downstream gene encoding a putative family 9 endoglucanase. This enzyme is the only family 48 glycoside hydrolase and is likely to be the sole exoglucanase in the chromosome of *C. phytofermentans* ISDg. The *cpcel48* gene encodes a protein of 919 amino acids containing a signal peptide (residue 1-33) with a typical cleavage site (ALA-G) as predicted by SignalP 3.0 (Bendtsen et al. 2004). The BLAST analysis suggested that the amino acid sequences from 41-666, 675-758, and 771-919 correspond to a glycoside hydrolase family 48 catalytic domain, an immunoglobulin-like domain (Mosbah et al. 2000), and a family 3 cellulose-binding module, respectively. The catalytic domain of CpCel48 has similarities to that of *C. stercorarium* Avicelase II (83%) (Bronnenmeier et al. 1991), *Clostridium cellulyticum* CelF (81%) (Reverbel-Leroy et al. 1997), and other Cel48 cellulases (57% to 80%).

Overexpression and purification

Three different forms of CpCel48 were heterologously expressed in *E. coli* and *B. subtilis*. They were intracellular CpCel48 with a His-tag by *E. coli*, intracellular CpCel48′ without a His-tag by *E. coli*, and secretory CpCel48′ by *B. subtilis*. *E. coli* BL21 (DE3) (pET20b-Cpcel48) expressed ca. 50 mg soluble CpCel48 per liter of culture (lane 2, Fig. 1a). CpCel48 was purified, and a yield of ~39.5% was obtained by using a nickel resin (lane 4, Fig. 1a). After an affinity adsorption on RAC, followed by desorption by ethylene glycol, CpCel48′ was purified, and a yield of ~10.4% was obtained (lane 5, Fig. 1a).

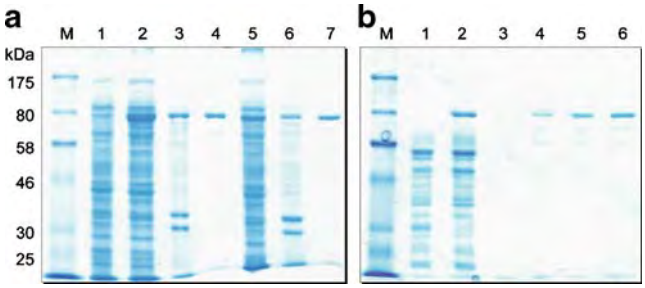

Fig. 1 SDS-PAGE analysis of expression of CpCel48 in *E. coli* and *B. subtilis*. **a** The expression and purification of CpCel48 in *E. coli*. *Lanes M* protein marker; *1* BL21(pET20b); *2* BL21(pET20b-Cpcel48), soluble fraction; *3* BL21(pET20b-Cpcel48), insoluble fraction; *4* the purified CpCel48; *5* BL21(pET20b-Cpcel48′), soluble fraction; *6* BL21(pET20b-Cpcel48′), insoluble fraction; *7* the purified CpCel48′. **b** CpCel48 expression and secretion in *B. subtilis*. *Lanes 1* WB800(pP43NMK), 48 h; *2* WB800(pP43N-Cpcel48′), 48 h; *3* WB800(pP43NMK), 48 h; *4* WB800(pP43N-Cpcel48′), 24 h; *5* WB800(pP43N-Cpcel48′), 48 h; *6* WB800(pP43N-Cpcel48′), 72 h. Samples loaded in all lanes were from 100-μL cell culture. *Lanes 1* and *2* were proteins precipitated by TCA from the culture supernatant. *Lanes 3–6* were partially purified CpCel48′ after RAC adsorption and desorption

In addition, the *cpcel48* gene was cloned into the *E. coli—B. subtilis* shuttle expression vector pP43NMK (Zhang et al. 2005). The plasmid contained an in-frame NprB signal peptide-encoding sequence, allowing the expression of secretory CpCel48′ in *B. subtilis*. *B. subtilis* WB800 (pP43N-Cpcel48′) produced soluble CpCel48′ as the dominant protein in the extracellular proteome of WB800 (lane 2, Fig. 1b). The CpCel48′ concentration was estimated to be ~15.8 mg l^{-1} of the culture at 48 h. The secretory CpCel48′ levels increased from 24 to 72 h (lanes 4–6, Fig. 1b), and its activity was confirmed by the assay on cellulose (data not shown).

Enzyme characterization

Figure 2a shows that CpCel48 exhibited the optimum pH between pH 5 and 6. At 50 °C, the enzyme activity in the presence of 1 mM Ca^{2+} was 49.3% higher than that without Ca^{2+} addition. Ca^{2+} also increased the optimum temperature for CpCel48 in the 20 min reaction from 55 to 60 °C, as shown in Fig. 2b. The addition of Ca^{2+} also increased its thermostability at high temperatures of 45–55 °C, as shown in Fig. 2c. The enzyme lost its activity quickly when the temperature was higher than 50 °C. An addition of 5 mM DTT did not affect its activity (data not shown).

Table 1 presents the specific activities of CpCel48 on different substrates. CpCel48 had activities of 15.1 and 35.9 U μmol^{-1} on insoluble cellulosic materials—Avicel and RAC, respectively. By contrast, this enzyme had much

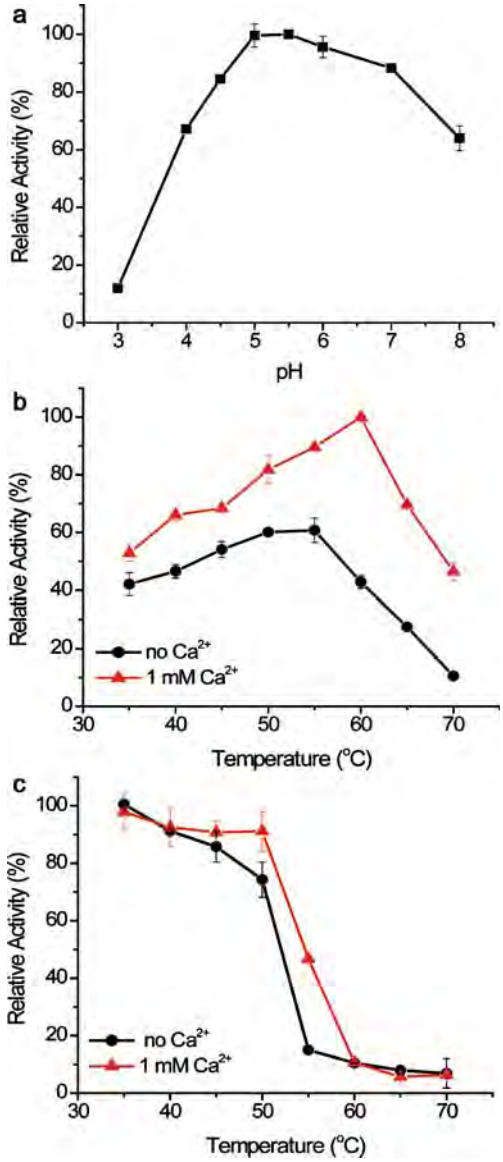

Fig. 2 Activity as function of pH (**a**) and temperature (**b**) of CpCel48 on Avicel in a 50-mM MES buffer. **c** Thermostability of CpCel48. The CpCel48 (20 μg mL^{-1}) solutions were incubated at different temperature for 30 min, and its residual activities were determined at 40 °C on Avicel

lower activities of 2.7, 3.8, and 4.6 U μmol^{-1} on low-viscosity CMC, medium-viscosity CMC, and hydroxyethylcellulose, respectively. It was highly possible that the observed weak activity of CpCel48 on substituted substrates was attributed to its action on unmodified sugar chains of soluble cellulose derivatives and its processivity was blocked by the substituted groups. CpCel48 also had a weak promiscuous activity on xylan (1.2 U μmol^{-1}). No activity was detected on three chromogenic substrates: (4-nitrophenyl-β-D-cellobioside (*p*NPC), 4-nitrophenyl-β-D-lactopyranoside (*p*NPL), and 4-methylumbelliferyl-β-D-cellobioside (*p*MUC). Also, CpCel48′ expressed in *E. coli* and secretory CpCel48′ expressed in *B. subtilis* had similar activities to CpCel48 (data not shown).

Figure 3 shows the viscosity profiles of the CMC solutions that were catalyzed by CpCel48 and *C. phytofermentans* endoglucanase Cel5B (GenBank accession number ABX43556). Endoglucanase Cel5B drastically decreased the viscosity of the CMC solution, while CpCel48 slightly decreased the viscosity which was similar to a typical pattern of exoglucanases (Kruus et al. 1995; Liu and Doi 1998; Morag et al. 1991; Watson et al. 2009). The relationship between the changes in viscosities and reducing ends also suggested that CpCel48 was a typical exoglucanase (data not shown).

The hydrolysis products of CpCel48 on soluble cellodextrins with degree of polymerization (DP) from three to six were analyzed by thin layer chromatography, as shown in Fig. 4a. CpCel48 showed no detectable activity on cellotriose (G3) and cleaved cellotetraose (G4) to cellobiose (G2), indicating that the minimum sugar chain length for CpCel48 was four. This provides an explanation for the observation that *p*NPC, *p*NPL, and *p*MUC cannot be hydrolyzed by this enzyme (Table 1). Cellopentaose (G5) was cleaved to G2 and G3, and cellohextraose (G6) to G2, G3, and G4. The hydrolysis products on insoluble cellulosic materials Avicel and RAC were analyzed by HPLC at a low ratio of enzyme/substrate in Fig. 4b, c. G2 was the dominant product, and G3 and G4 were minor products, suggesting that CpCel48 was a cellobiohydrolase. On Avicel, the molar ratios of G2, G3, and G4 were 77.0–79.6%, 13.3–16.4%, and 6.1–7.9%, respectively. On RAC, the molar ratio of G2 was much higher (86.6–88.1%) while G3 and G4 accounted for 8.1–8.8% and 3.6–4.6%, respectively.

Processivity

To estimate the processivity (P) of CpCel48, we established a new false/right initial attack model. Natural cellulose chains have two types of cellulose configuration (50% for each, Fig. 5; Wood et al. 1989; Zhang and Lynd 2004). Considering the catalytic mechanism and crystal structure of Cel48 attacking from reducing ends of cellulose chains (Guimaraes et al. 2002; Parsiegla et al. 1998), Cel48 enzymes are suggested to cut every two glucosidic bonds along the cellulose chain. As shown in Fig. 5, the first cuttings of type A chains by CpCel48 result in right products (αG2) and false products (1-α)G4, where α represents a fraction of G2 to (G2+G4); while the first cutting of type B chain results in false products (G3) only. After initial false or right attacks, CpCel48 hydrolyzes along the cellulose chain, resulting in (P−1) molecules of cellobiose released, until the processive actions of CpCel48 were stopped. Therefore, the product compositions of

Table 1 Specific activities of CpCel48

Substrate	Specific activity (U/μmole)
Microcrystalline cellulose (Avicel)	15.1±0.6
Regenerated amorphous cellulose	35.9±0.3
Carboxymethyl cellulose (low viscosity)	2.7±0.7
Carboxymethyl cellulose (medium viscosity)	3.8±0.2
Hydroxyethylcellulose	4.6±0.5
Oat spelts xylan	1.2±0.9
4-nitrophenyl-β-D-cellobioside	ND
4-nitrophenyl-β-D-lactopyranoside	ND
4-methylumbelliferyl-β-D-cellobioside	ND

ND not detected

CpCel48 for the first round of adsorption, hydrolysis, and dissociation were rewritten as

$$\text{Type A} \quad 50\alpha\% \times [\text{G4} + (\text{P} - 1)\text{G2}] \tag{1}$$

$$\text{Type A} \quad 50(1 - \alpha)\% \times [\text{P} \cdot \text{G2}] \tag{2}$$

$$\text{Type B} \quad 50\% \times [\text{G3} + (\text{P} - 1)\text{G2}] \tag{3}$$

In sum, the enzymatic hydrolysate contained G2—50% [2(P-1)+(1-α)], G3—50%, and G4—50α%. As a result, the values of α and P can be rewritten as

$$\frac{[G3]}{[G4]} = \frac{1}{\alpha} \tag{4}$$

$$\frac{[G2]}{[G3] + [G4]} = \frac{2(P - 1) + (1 - a)}{1 + a} \tag{5}$$

The above *P* estimation required two conditions: (1) the ratio of the substrate to the enzyme is far in excess so that all enzymes are bound on the cellulose (i.e., no soluble G4 can be further hydrolyzed by free CpCel48; otherwise it would result in an overestimation of *P* value); and (2) cellulose digestibility is very low (i.e., most of substrate hydrolysis occurs at the first round of adsorption, hydrolysis, and dissociation; otherwise it would result in an overestimation of *P* value). It was found that enzymatic cellulose hydrolysis mediated by higher enzyme loadings or longer hydrolysis time resulted in higher ratios of $\frac{[G2]}{[G3]+[G4]}$, i.e., overestimated *P* values (data not shown). The results showed that the ratios of G2:G3:G4 by CpCel48 were nearly constant for 6 h, as shown in Fig. 4a. These results were obtained under carefully chosen conditions (i.e., very high ratio of substrate to enzyme, short hydrolysis time, and limited cellulose hydrolysis). For the hydrolysis of Avicel and RAC, the ratios of G3 to G4 were nearly constant, being ~2, suggesting that α equaled 50% (i.e., the initial false attack of 50% of type A chains resulted in one G4; the initial right attack of the other 50% of type A chains resulted in one G2). Based on Eq. 5, the *P* values on Avicel and RAC were 3.50 and 6.04, respectively; i.e., CpCel48 conducted 3.5 processive attacks on Avicel and 6.0 attacks on RAC before its dissociation from the substrates.

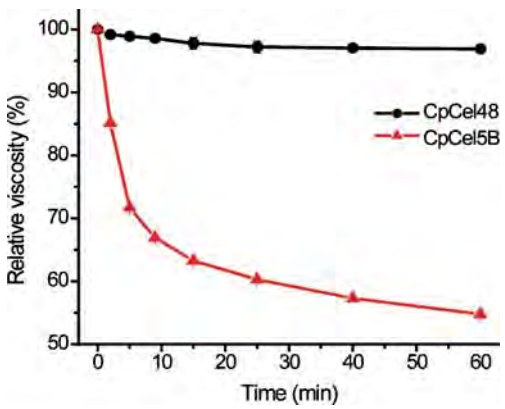

Fig. 3 Viscosity profiles of the CMC solutions hydrolyzed by CpCel48 and CpCel5B

Discussion

In this study, the first cellulase from *C. phytofermentans* ISDg was cloned, purified, and characterized. This enzyme belongs to the family 48 glycoside hydrolases. Several pieces of evidence suggested that CpCel48 is a cellobiohydrolase: (1) CpCel48 exhibited high activities of 15. U μmol^{-1} on Avicel and 35.9 U μmol^{-1} on RAC, but very low activities of 2.7–4.6 U μmol^{-1} on soluble chemically substituted celluloses; (2) CpCel48 was not able to significantly decrease the CMC solution viscosity; and (3) CpCel48 produced a major hydrolysis product—cellobiose.

Heterologous expression of soluble Cel48 enzymes in *E. coli* has been reported to be difficult; for example, *C. thermocellum* CelS (Kruus et al. 1995; Wang et al. 1994) and *C. Cellulovorans* ExgS (Liu and Doi 1998). Therefore, it was encouraging to produce active CpCel48 in *E. coli*. *B.*

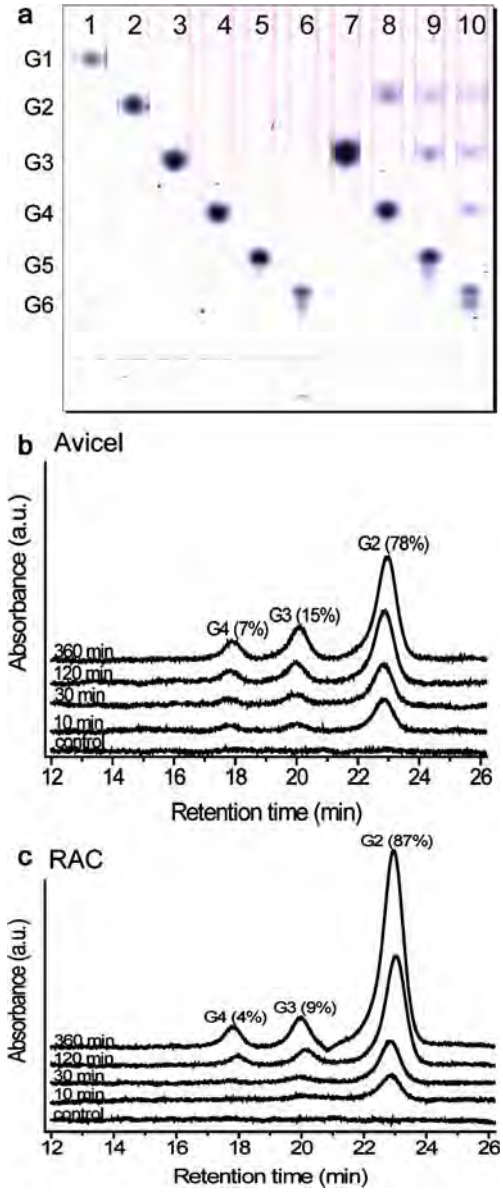

Fig. 4 **a** Thin layer chromatography analysis of hydrolysis products. G1–G6 (*lanes 1–6*) refer to the positions of standards (4 μg of each): glucose (*G1*), cellobiose (*G2*), cellotriose (*G3*), cellotetraose (*G4*), cellopentaose (*G5*), and cellohextaose (*G6*), respectively. Reactions contained G3 (0.4%), G4 (0.4%), G5 (0.4%), and G6 (0.4%) were incubated at 37 °C for 48 h with 50 μg mL^{-1} CpCel48. Samples loaded in *lanes 7–10* were the hydrolysis products of G3 (1 μL), G4 (1 μL), G5 (1 μL), and G6 (2 μL), respectively. HPLC analysis of hydrolysis products on Avicel (**b**) and on RAC (**c**). The reactions (1 mL) containing 3% (*w*/*v*) Avicel or 1% RAC were incubated at 50 °C with 10 μg ml^{-1} CpCel48

subtilis strains, industrially safe microorganisms, have been widely used for producing enzymes and polymers. Efficient secretion of active cellulases is important in the development of recombinant cellulolytic *B. subtilis* that can produce cellulase, hydrolyze cellulose, and ferment the soluble sugars to biocommodities in a single step (Arai et

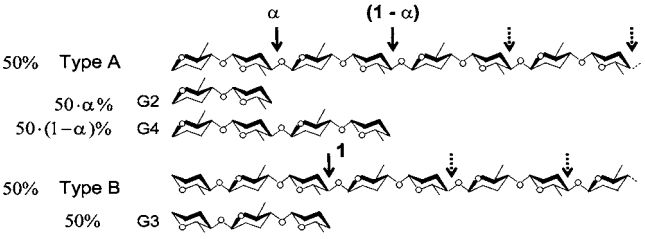

Fig. 5 The scheme of initial false/right attack model for CpCel48. Three possible cleavage sites for the initial attack are indicated by *solid arrows*

al. 2007; Cho et al. 2004; Doi 2008). This study is the first report of the expression of a secretory family 48 glycoside hydrolase in *B. subtilis*, where cellobiohydrolases are vital to hydrolysis of crystalline cellulose (Teeri 1997; Zhang et al. 2006b).

Distinctive from the previous two models for the *T. reesei* CBH1 (Cel7A)—$P = [G2]/([G1] + [G3])$ (von Ossowski et al. 2003) and $P = [G2]/[G1] + 0.5$ (Medve et al. 1998), this initial false/right attack model (Eqs. 4 and 5) for glycoside hydrolase family 48 enzymes was proposed based on the substrate specificity of the enzyme, the crystal structure of homologous Cel48 enzymes, and the configuration of natural cellulose. Since the experimental conditions—such as enzyme loading, substrate to enzyme ratio, and reaction time—were known to influence the *P* values (positive bias), the experiments were designed carefully to ensure the reliability of the initial false/right attack model. The comparison of *P* value of 3.5 on Avicel (approximately seven to nine anhydroglucose units) and nine sugar-binding sites of family 48 cellulase (Guimaraes et al. 2002) implied that CpCel48, which was bound on the single cellulose chain that was detached from Avicel, does not processively hydrolyze the cellulose chain. This poor processivity may be attributed to a highly cooperative interchain and intrachain hydrogen bond interaction for crystalline cellulose (Qian 2008; Zhang and Lynd 2004). The *P* value was increased to 6.0 (i.e., 12–14 anhydroglucose units) on RAC, suggesting that CpCel48 hydrolyzes the sugar chain more processively on RAC, which does not have orderly hydrogen bonds than on microcrystalline cellulose. Interestingly, the ratio of processivity on Avicel relative to RAC (3.5 vs 6) was approximately equal to the ratio of specific activities on the respective substrates (15.1 vs 35.9 U μmol^{-1}), implying an intrinsic correlation between enzyme processivity and enzyme activity. Since it was suggested that processivity of exoenzyme came at a large cost in terms of enzyme hydrolysis rate (Horn et al. 2006; von Ossowski et al. 2003), it was possible that the relatively weaker processivity of CpCel48 might be correlated with its higher specific activity. Considering the implicated relationship among enzyme processivity, mass-specific activity, and substrate accessibility, it might

be better to increase substrate accessibility for the actions of nonprocessive endoenzymes having higher turnover numbers by using cellulose solvent-based biomass pretreatment (Sathitsuksanoh et al. 2009; Zhang et al. 2007; Zhu et al. 2009) rather than to improve the properties of processive enzymes on low accessibility crystalline cellulose (Hong et al. 2007).

It would be important and interesting to compare different enzymes of family 48 glycoside hydrolases in one laboratory in the future, because they play an important role in crystalline cellulose hydrolysis.

Acknowledgments This work was supported mainly by the DOE BioEnergy Science Center (BESC). The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. This work was also partially supported by the USDA Bioprocessing and Biodesign Center and DuPont Young Professor Award. Zuoming Zhang was partially supported by the China Scholar Council. Noppadon Sathitsuksanoh was partially supported by the ICTAS scholar program.

References

Anagnostopoulos C, Spizizen J (1961) Requirements for Transformation in *Bacillus subtilis*. J Bacteriol 81(5):741–746

Arai T, Matsuoka S, Cho HY, Yukawa H, Inui M, Wong SL, Doi RH (2007) Synthesis of *Clostridium cellulovorans* minicellulosomes by intercellular complementation. Proc Natl Acad Sci U S A 104 (5):1456–1460

Bayer EA, Setter E, Lamed R (1985) Organization and distribution of the cellulosome in *Clostridium thermocellum*. J Bacteriol 163:552–559

Bayer EA, Belaich JP, Shoham Y, Lamed R (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. Annu Rev Microbiol 58:521–554

Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340 (4):783–795

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254

Breyer WA, Matthews BW (2001) A structural basis for processivity. Protein Sci 10(9):1699–1711

Bronnenmeier K, Rucknagel KP, Staudenbauer WL (1991) Purification and properties of a novel type of exo-1, 4-beta-glucanase (Avicelase II) from the cellulolytic thermophile *Clostridium stercorarium*. Eur J Biochem 200(2):379–385

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucl Acids Res 37(suppl_1):D233–D238

Cho HY, Yukawa H, Inui M, Doi RH, Wong SL (2004) Production of minicellulosomes from *Clostridium cellulovorans* in *Bacillus subtilis* WB800. Appl Environ Microbiol 70(9):5704–5707

Devillard E, Goodheart DB, Karnati SKR, Bayer EA, Lamed R, Miron J, Nelson KE, Morrison M (2004) *Ruminococcus albus* 8 mutants defective in cellulose degradation are deficient in two processive endocellulases, Cel48A and Cel9B, both of which possess a novel modular architecture. J Bacteriol 186(1):136–145

Doi RH (2008) Cellulases of mesophilic microorganisms: cellulosome and noncellulosome producers. Ann N Y Acad Sci 1125:267–279

Doi RH, Kosugi A (2004) Cellulosomes: plant-cell-wall-degrading enzyme complexes. Nat Rev Microbiol 2:541–551

Guimaraes BG, Souchon H, Lytle BL, David Wu JH, Alzari PM (2002) The crystal structure and catalytic mechanism of cellobiohydrolase CelS, the major enzymatic component of the *Clostridium thermocellum* cellulosome. J Mol Biol 320:587–596

Hong J, Ye X, Zhang YHP (2007) Quantitative determination of cellulose accessibility to cellulase based on adsorption of a nonhydrolytic fusion protein containing CBM and GFP with its applications. Langmuir 23(25):12535–12540

Hong J, Ye X, Wang Y, Zhang Y-HP (2008) Bioseparation of recombinant cellulose binding module-protein by affinity adsorption on an ultra-high-capacity cellulosic adsorbent. Anal Chem Acta 621:193–199

Horn SJ, Sikorski P, Cederkvist JB, Vaaje-Kolstad G, Sorlie M, Synstad B, Vriend G, Varum KM, Eijsink VGH (2006) Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. Proc Nat Acad Sci USA 103 (48):18089–18094

Irwin DC, Zhang S, Wilson DB (2000) Cloning, expression and characterization of a family 48 exocellulase, Cel48A, from *Thermobifida fusca*. Eur J Biochem 267(16):4988–4997

Kakiuchi M, Isui A, Suzuki K, Fujino T, Fujino E, Kimura T, Karita S, Sakka K, Ohmiya K (1998) Cloning and DNA sequencing of the genes encoding *Clostridium josui* scaffolding protein CipA and cellulase CelD and identification of their gene products as major components of the cellulosome. J Bacteriol 180:4303–4308

Kipper K, Valjamae P, Johnansson G (2005) Processive action of cellobiohydrolase Cel7A from *Trichoderma reesei* is revealed as 'burst' kinetics on fluorescent polymeric model substrates. Biochem J 385:527–535

Kruus K, Wang WK, Ching J, Wu JH (1995) Exoglucanase activities of the recombinant *Clostridium thermocellum* CelS, a major cellulosome component. J Bacteriol 177:1641–1644

Liu CC, Doi RH (1998) Properties of *exgS*, a gene for a major subunit of the *Clostridium cellulovorans* cellulosome. Gene 211 (1):39–47

Lu Y, Zhang Y-HP, Lynd LR (2006) Enzyme-microbe synergy during cellulose hydrolysis by *Clostridium thermocellum*. Proc Natl Acad Sci USA 103(44):16165–16169

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66:506–577

Lynd LR, Laser MS, Bransby D, Dale BE, Davison B, Hamilton R, Himmel M, Keller M, McMillan JD, Sheehan J et al (2008) How biotech can transform biofuels. Nat Biotechnol 26(2):169–172

Medve J, Karlsson J, Lee D, Tjerneld F (1998) Hydrolysis of microcrystalline cellulose by cellobiohydrolases I and endoglucanase II from *Trichoderma reesei*: adsorption, sugar production pattern, and synergism of the enzymes. Biotechnol Bioeng 59:621–634

Morag E, Halevy I, Bayer EA, Lamed R (1991) Isolation and properties of a major cellobiohydrolase from the cellulosome of *Clostridium thermocellum*. J Bacteriol 173(13):4155–4162

Mosbah A, Belaich A, Bornet O, Belaich JP, Henrissat B, Darbon H (2000) Solution structure of the module X2_1 of unknown function of the cellulosomal scaffolding protein CipC of *Clostridium cellulolyticum*. J Mol Biol 304(2):201–217

Parsiegla G, Juy M, Reverbel-Leroy C, Tardif C, Belaich JP, Driguez H, Haser R (1998) The crystal structure of the processive endocellulase CelF of *Clostridium cellulolyticum* in complex with a thiooligosaccharide inhibitor at 2.0 Å resolution. EMBO J 17:5551–5562

Qian X (2008) The effect of cooperativity on hydrogen bonding interactions in native cellulose Iβ from ab initio molecular dynamics simulations. Mol Simulation 34(2):183–191

Reverbel-Leroy C, Pages S, Belaich A, Belaich J-P, Tardif C (1997) The processive endoglucanase CelF, a major component of the *Clostridium cellulolyticum* cellulosome: purification and characterization of the recombinant form. J Bacteriol 179:46–52

Sánchez MM, Pastor FIJ, Diaz P (2003) Exo-mode of action of cellobiohydrolase Cel48C from *Paenibacillus* sp. BP-23. Eur J Biochem 270(13):2913–2919

Sathitsuksanoh N, Zhu Z, Templeton N, Rollin J, Harvey S, Zhang Y-HP (2009) Saccharification of a potential bioenergy crop, *Phragmites australis* (common reed), by lignocellulose fractionation followed by enzymatic hydrolysis at decreased cellulase loadings. Ind Eng Chem Res 48:6441–6447

Shen H, Gilkes NR, Kilburn DG, Miller RCJ, Warren RA (1995) Cellobiohydrolase B, a second exo-cellobiohydrolase from the cellulolytic bacterium *Cellulomonas fimi*. Biochem J 311:67–74

Teeri TT (1997) Crystalline cellulose degradation: new insights into the function of cellobiohydrolases. Trends Biotechnol 15:160–167

Teeri TT, Koivula A, Linder M, Wohlfahrt G, Divne C, Jones TA (1998) *Trichoderma reesei* cellobiohydrolases: why so efficient on crystalline cellulose? Biochem Soc Trans 26:173–178

Von Hippel PH, Fairfield FR, Dolejsi MK (1994) On the processivity of polymerases. Ann NY Acad Sci 726:118–131

von Ossowski I, Stahlberg J, Koivula A, Piens K, Becker D, Boer H, Harle R, Harris M, Divne C, Mahdi S et al (2003) Engineering the Exo-loop of *Trichoderma reesei* Cellobiohydrolase, Cel7A. A comparison with *Phanerochaete chrysosporium* Cel7D. J Mol Biol 333(4):817–829

Wang WK, Kruus K, Wu JH (1994) Cloning and expression of the *Clostridium thermocellum celS* gene in *Escherichia coli*. Appl Microbiol Biotechnol 42:346–352

Warnick TA, Methe BA, Leschine SB (2002) *Clostridium phytofermentans* sp. nov., a cellulolytic mesophile from forest soil. Int J Syst Evol Microbiol 52(4):1155–1160

Watson BJ, Zhang H, Longmire AG, Moon YH, Hutcheson SW (2009) Processive endoglucanases mediate degradation of cellulose by the *Saccharophagus degradans* cellulolytic system. J Bacteriol

Wood TM, McCrae SI, Bhat KM (1989) The mechanism of fungal cellulase action: synergism between enzyme components of *Penicillium pinophilum* cellulase in solubilizing hydrogen bond-ordered cellulose. Biochem J 260:37–43

Wu SC, Yeung JC, Duan Y, Ye R, Szarka SJ, Habibi HR, Wong SL (2002) Functional production and characterization of a fibrin-specific single-chain antibody fragment from *Bacillus subtilis*: effects of molecular chaperones and a wall-bound protease on antibody fragment production. Appl Environ Microbiol 68 (7):3261–3269

Zhang Y-HP (2008) Reviving the carbohydrate economy via multi-product biorefineries. J Ind Microbiol Biotechnol 35(5):367–375

Zhang Y-HP, Lynd LR (2003) Cellodextrin preparation by mixed-acid hydrolysis and chromatographic separation. Anal Biochem 322:225–232

Zhang Y-HP, Lynd LR (2004) Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. Biotechnol Bioeng 88:797–824

Zhang Y-HP, Lynd LR (2005a) Cellulose utilization by *Clostridium thermocellum*: bioenergetics and hydrolysis product assimilation. Proc Natl Acad Sci USA 102:7321–7325

Zhang Y-HP, Lynd LR (2005b) Determination of the number-average degree of polymerization of cellodextrins and cellulose with application to enzymatic hydrolysis. Biomacromolecules 6:1510–1515

Zhang Y-HP, Lynd LR (2005c) Regulation of cellulase synthesis in batch and continuous cultures of *Clostridium thermocellum*. J Bacteriol 187:99–106

Zhang Y-HP, Lynd LR (2006) Biosynthesis of radio-labeled cellodextrins by the *Clostridium thermocellum* cellobiose and cellodextrin phosphorylases for measurement of intracellular sugars. Appl Microbiol Biotechnol 70(1):123–129

Zhang XZ, Cui ZL, Hong Q, Li SP (2005) High-level expression and secretion of methyl parathion hydrolase in *Bacillus subtilis* WB800. Appl Environ Microbiol 71(7):4101–4103

Zhang Y-HP, Cui J-B, Lynd LR, Kuang LR (2006a) A transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: evidences from enzymatic hydrolysis and supramolecular structure. Biomacromolecules 7(2):644–648

Zhang Y-HP, Himmel M, Mielenz JR (2006b) Outlook for cellulase improvement: screening and selection strategies. Biotechnol Adv 24(5):452–481

Zhang Y-HP, Ding S-Y, Mielenz JR, Elander R, Laser M, Himmel M, McMillan JD, Lynd LR (2007) Fractionating recalcitrant lignocellulose at modest reaction conditions. Biotechnol Bioeng 97 (2):214–223

Zhang Y-HP, Hong J, Ye X (2009) Cellulase assays. Meth Mol Biol 581:213–231

Zhu Z, Sathitsuksanoh N, Vinzant T, Schell DJ, McMillan JD, Zhang Y-HP (2009) Comparative study of corn stover pretreated by dilute acid and cellulose solvent-based lignocellulose fractionation: enzymatic hydrolysis, supramolecular structure, and substrate accessibility. Biotechnol Bioeng 103:715–724