Bioresource Technology 112 (2012) 10–17

Contents lists available at SciVerse ScienceDirect

Bioresource Technology

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

Construction and characterization of chimeric cellulases with enhanced catalytic activity towards insoluble cellulosic substrates

Amar A. Telke [a], Sunil S. Ghatge [a], Seo-Hee Kang [a], Sundarapandian Thangapandian [a], Keun-Woo Lee [a], Hyun-Dong Shin [b], Youngsoon Um [c], Seon-Won Kim [a,*]

[a] Division of Applied Life Sciences (BK21), PMBBRC, Gyeongsang National University, 900 Gajwa-dong, Jinju, Gyeongnam 660-701, Republic of Korea
[b] School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
[c] Center for Environmental Technology Research, KIST, 39-1 Hawolgok-Dong, Seoul 136-791, Republic of Korea

ARTICLE INFO

Article history:
Received 10 January 2012
Received in revised form 13 February 2012
Accepted 15 February 2012
Available online 22 February 2012

Keywords:
Alicyclobacillus acidocaldrious
Endoglucanase
Cellulose binding module
Chimeric enzymes
Computational modeling

ABSTRACT

The chimeric proteins viz. CBM3–Cel9A, CBM4–Cel9A and CBM30–Cel9A, are constructed by fusion of family 3, 4, and 30 cellulose binding modules (CBMs) to N-terminus of family 9 endoglucanase (Cel9A) from *Alicyclobacillus acidocaldrious*. The chimeric enzymes were successfully expressed in *Escherichia coli* and purified to homogeneity. The chimeric enzymes showed significant increase in Avicel (8–12 folds) and filter paper (7–10 folds) degradation activities compared to Cel9A endoglucanase. Computational protein modeling and simulation on the chimeric enzymes were applied to analyze the fused CBMs effect on the increased insoluble cellulosic substrates degradation activity. Thin layer chromatography analysis of the enzymatic hydrolysis products and distribution of reducing sugars between soluble and insoluble fractions indicated processive cleavage of insoluble cellulosic substrates by the chimeras. The fused CBMs played a critical accessory role for the Cel9A catalytic domain and changed its character to facilitate the processive cleavage of insoluble cellulosic substrates.

1. Introduction

Cellulosic substrates derived from tailored crops will be used for production of biofuel such as ethanol which is called the next generation biofuel. Cellulose is the most abundant carbon source in nature, but it is very difficult to degrade, because of its insolubility and quasi-crystalline structure, and its presence in plant cell walls in a matrix with other polymers that limit access to the cellulose surface. An efficient breakdown of cellulose is a major limiting step for the next generation biofuel production and the cellulases acting on insoluble cellulose are key enzymes in this process.

A number of cellulases isolated from various microorganisms have been classified in glycosyl hydrolase family 9 (GH9). The GH9 cellulases show a broad range of substrate specificity and are categorized into four different subgroups of A, B, C, and D on basis of protein molecular architectures. The subgroup A includes enzymes consisting of only catalytic module without a cellulose binding module (CBM). The subgroup B enzymes possess the family 3 CBM in an immediate downstream of catalytic module. The subgroup C enzymes contain N-terminal immunoglobulin (Ig)-like domain followed by single catalytic module. The enzymes belongs to subgroup D contain not only the Ig-like domain but also at least one additional CBM. Although the GH9 cellulases that are belongs to subgroup A, B, C, and D have different molecular architecture but they contain similar $(\alpha–\alpha)_6$-barrel fold in catalytic module. In general, the catalytic module and a CBM in a cellulase molecule can function independently to each other. The glycoside hydrolases that utilize insoluble cellulosic substrates are modular proteins being comprised of catalytic module appended to one or more noncatalytic CBMs. It is well known that CBMs play a critical role in enzymatic hydrolysis of plant structural and storage polysaccharides (Cosgrove, 2000). Many CBMs have now been identified experimentally and several hundred putative CBMs can be further elucidated on a basis of amino acid sequence homology (Boraston et al., 2004). There are currently 43 families of CBMs and these CBMs display substantial variation in ligand specificity (Shoseyov et al., 2006). The CBMs are composed of 30–200 amino acids and present next to either C-terminal or N-terminal of cellulase catalytic module in single, double or triple domains. The contribution of CBMs to the activity of cellulases against insoluble cellulose was shown in several cases by removal of CBM from insoluble cellulose degrading cellulase or by grafting CBM onto cellulase consisting of only catalytic module. The family 3, 4, and 30 CBMs are commonly found with GH9 catalytic module and their removal from the catalytic module

* Corresponding author. Address: Division of Applied Life Sciences (BK21), Gyeongsang National University, 900 Gajwa-dong, Jinju, Gyeongnam 660-701, Republic of Korea. Tel.: +82 55 772 1362.
E-mail address: swkim@gnu.ac.kr (S.-W. Kim).

0960-8524/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.biortech.2012.02.066

results into significant loss in enzyme activity. The CBMs have accessory role for retaining the hydrolysis property of GH9 catalytic module. For example, the removal of family 30 CBM from *Clostridium thermocellum* endoglucanase CelJ led to 99% and 90% losses of catalytic activity towards CMC and Avicel, respectively (Arai et al., 2003). *Clostridium cellulolyticum* Cel9E is inactivated by deletion of its family 4 CBM (Gaudin et al., 2000).

The CBMs have been known to increase enzyme–substrate proximity and also enhance accessibility of the substrate by modifying surface structure of insoluble cellulose. If cellulase is engineered to bind to insoluble cellulose in high affinity through fusion of an effective CBM, the enzyme concentration on the insoluble substrate surface can be increased significantly. Subsequently, the catalytic efficiency of the engineered cellulases to insoluble cellulose can be improved by the CBM. In the present study, chimeric proteins were constructed by fusion of family 3, 4, and 30 CBMs to N-terminus of GH9 endoglucanase (Cel9A) from *Alicyclobacillus acidocaldarious* ATCC27009. The Cel9A endoglucanase was chosen for this study because of its hyper-thermophillic nature and broad substrate specificity. The absence of CBM in the Cel9A endoglucanase might be the reason for its low hydrolytic activity towards insoluble cellulosic substrates. The family 3, 4, and 30 CBMs from *C. thermocellum* 35319 were adopted for the fusion due to their carbohydrate chain specificity and thermal stability.

2. Methods

2.1. Chemicals

Avicel PH101, carboxymethyl cellulose (CMC), barley β-glucan, *p*-nitrophenyl-β-D-glucoside (*p*-NPG) and *p*-nitrophenyl-β-D-cellobioside (*p*-NPC) were obtained from Sigma Chemicals (St. Louis, MO, USA). TLC plates were obtained from Merck Laboratory (Germany).

2.2. Bacterial strains, plasmids and growth conditions

The codon optimized synthetic gene encoding Cel9A endoglucanase in plasmid of pUCel9A was obtained from GenScript (NJ, USA). The genomic DNA of *C. thermocellum* ATCC 35319 was used as source of gene for family 3, 4 and 30 CBMs. The genomic DNA of *Saccharophagus degradans* strain 2–40 ATCC43961 was used as source for processive endoglucanase Cel5H. The expression vector pET-28a(+) was purchased from Novagen (Madison, WI, USA). *Escherichia coli* DH5a and *E. coli* BL21 (DE3) were used as host for cloning and expression, respectively. Luria–Bertani (LB) medium supplemented with kanamycin (50 μg/ml) was used for the cultivation of *E. coli* cells.

2.3. Construction, cloning and sequencing of chimeric genes

The chimeric genes were constructed using a three-step overlapping polymerase chain reaction (PCR) (Zhang et al., 2010). The construction procedure is shown in Fig. 1. The pUCel9A plasmid was used as template for amplification of full length Cel9A endoglucanse. The families of CBM3 from the cellulosome-integrating protein CipA, CBM4 from cellobiohydrolase A (CbhA) and CBM 30 from the processive endoglucanase CelJ were amplified using PCR from genomic DNA of *C. thermocellum*. The full length gene of processive endoglucanase Cel5H was amplified from genomic DNA of *S. degradans* strain 2–40. The CbhA (GH9-$Fn_{3,2}$-CBDIII) gene was cloned by using PCR from genomic DNA of *C. thermocellum*. The sequences of the primers used in this study are given in Table S1. The restriction sites were introduced at the 5′ and 3′ ends of the primers, respectively. The primers were obtained from Bioneer Co., Ltd. (South Korea).

The PCR was performed using 30 successive cycles as follows: denaturation at 98 °C for 0.5 min, annealing at 56 °C for 0.5 min, and extension at 72 °C for the time dependent on the PCR products length. The Phusion high-fidelity DNA polymerase (Finnzymes, USA) was utilized during the process. PCR products were separated by 1% agarose gel electrophoresis and extracted from the gel using gel extraction kit. The extracted DNA fragments were digested with restriction enzymes and ligated into the vector pET-28a(+). The ligation mixture was used to transform chemically competent *E. coli* DH5a. The plasmid isolated from these transformants was verified by restriction analysis, and the gene sequence was confirmed by DNA sequencing. The plasmids with the correct sequences for Cel5H, CbhA, Cel9A, CBM3–Cel9A, CBM4–Cel9A and CBM30–Cel9A were named pECel5H, pECbhA, pECel9A, pECBM3–Cel9A, pECBM4–Cel9A and pECBM30–Cel9A, respectively.

The enzymes used for DNA manipulations were purchased from New England Biolabs (UK). DNA sequencing was performed by SolGent (South Korea). Genomic DNA extraction was performed according to the method described as earlier (Sambrook and Russell, 2001). Plasmid DNA was extracted using Qiagen Spin Column Plasmid Mini-Preps kit (USA). Restricted plasmids and PCR products were recovered from agarose gel using Qiagen gel extraction kit (USA).

2.4. Expression and purification of the chimeric proteins

E. coli BL21(DE3) cells harboring pECel5H, pECbhA, pECel9A, pECBM3–Cel9A, pECBM4–Cel9A and pECBM30–Cel9A were grown in LB medium containing kanamycin (50 μg/ml) at 37 °C. When the culture reached an A_{600} of 0.7, IPTG (isopropyl-β-D-thiogalactopyranoside) was added to a final concentration of 0.2 mM. The culture was then incubated overnight at 22 °C under shaking condition (180 rpm). The culture after overnight induction was centrifuged (10,000*g*, 20 min, 4 °C), and the cell pellet was collected. The proteins were extracted using the CelLyticB system (Sigma) according to the manufacturer's instructions. The chimeric proteins were purified on a precharged nickel Sepharose column under native conditions (HisTrap, GE Healthcare/Amersham, Piscataway, NJ).

2.5. Enzyme assays

Most assays for chimeric enzymes were performed at pH 5.0 and 70 °C with 0.1–1.0 nmol purified enzymes in reaction buffer containing 50 mM sodium acetate (pH 5.0), 200 mM NaCl, and 10 mM $CaCl_2$. The one unit of enzyme activity was defined as the amount of enzyme releasing 1 μmol of reducing sugar per min. The effect of pH and temperature on catalytic activity was examined in the pH range 4.0–8.0 and 10–90 °C, respectively. Sodium acetate (4.0–5.6) and piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (6.0–8.0) buffers were used for maintaining the pH at a final concentration of 50 mM.

The carboxymethyl cellulose (CMC) and barley β-glucan assays were performed with 1% (w/v) substrate in a total volume of 0.5 ml for 10 min reaction time. Phosphoric acid-swollen cellulose (PASC) was prepared according to method described in previous reports (Wood, 1971; Zhang et al., 2006). The PASC (0.1%, w/v) degradation assay was performed in a total volume of 0.5 ml for 10 min. The dinitrosalicylic acid (DNS) assay was used to detect product formation in these reactions (Ghose, 1987). Digestion of Avicel (1%, w/v) or Whatman No. 1 filter paper (3 mg) was carried out in total volume of 0.25 ml for 2 h at 60 °C. The reaction was stopped by incubation for 15 min at 95 °C and the residual substrate was separated by centrifugation at 10,000g. The products of Avicel or filter paper digestion were then digested with 50 U of β-glucosidase (Sigma, USA) in a total volume of 0.35 ml at 37 °C for 1 h. The β-glucosidase was inactivated by incubation for 15 min at 95 °C. The glucose con-

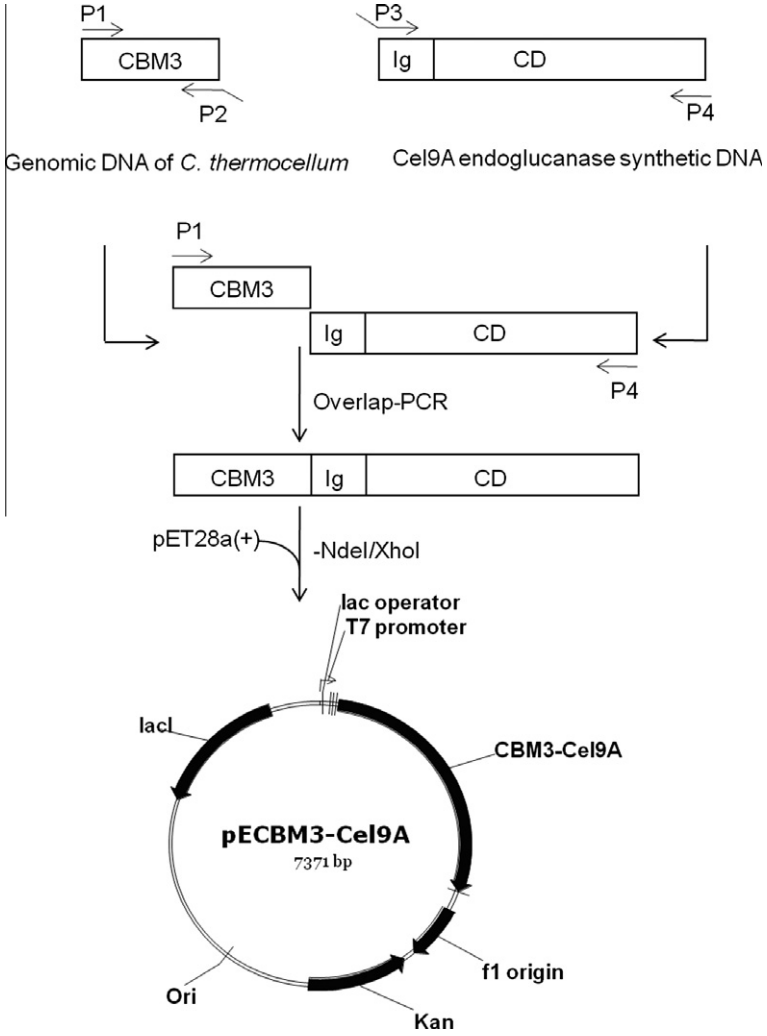

Fig. 1. Schematic diagram for construction of chimeric enzymes. The CBM3 from *C. thermocellum* is fused to Ig-like domain of Cel9A endoglucanase from *A. acidocaldarious*. The fusions of CBM4 and CBM30 from *C. thermocellum* with Cel9A endoglucanase follow the same strategy.

centration in final reaction mixture was measured using glucose oxidase kit (GAGO-20 from Sigma, USA), according to manufacturer's instructions. Release of cellobiose was calculated as 50% of the rate of glucose accumulation (Watson et al., 2009). The hydrolysis of *p*-nitrophenyl-β-D-glycoside (10 mM) and *p*-nitrophenyl-β-D-cellobioside (10 mM) was assayed by monitoring the released *p*-nitrophenol concentration at 410 nm after addition of NaOH at final concentration of 0.1 M (Eckert et al., 2002).

2.6. Protein quantification and SDS–PAGE analysis

Protein concentrations were determined by using the Pierce BCA protein assay kit (Thermoscientific, USA) with purified bovine serum albumin as standard. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) was carried out in a vertical polyacrylamide slab gel. Electrophoresis was performed on 5% staking and 12% polyacrylamide gel under denaturing condition. The buffer solutions for the staking and separating gel were 1 M Tris–HCl (pH 6.8) and 1.5 M Tris–HCl (pH 8.8), respectively.

2.7. Cellulose binding assay

Avicel binding property of the chimeric enzymes was studied as described previously with minor modifications (Li et al., 2007). The binding assay was carried out in 2.0 ml tubes. The 1 mg of Avicel was mixed with 0.5% bovine serum albumin in sodium acetate buffer (20 mM, pH 5). The mixture was incubated at 25 °C for 30 min to avoid nonspecific binding. The same amounts (2 nmol) of native or chimeric enzymes were added to this solution. Then, the tubes were placed in intelli-mixer RM-2 (Rose Scientific Ltd., Canada) for 1 h with a rotation speed of 20 rpm. The reaction mixtures were centrifuged at 10,000*g* for 5 min, and the amount of unbound enzyme was estimated from the residual cellulase activity in the supernatant. The amount of Avicel bound enzyme was calculated from the difference between the initial enzyme activity and the unbound enzyme activity. Control reactions were performed under the same condition, except in the absence of Avicel. Also, the Avicel bound proteins were qualitatively analyzed by using SDS–PAGE.

2.8. Cellulase processivity assay and synergistic interaction

Cellulase processivity was evaluated by measuring the ratio of soluble to insoluble reducing sugars according to the procedure of Watson et al. (2009). The processivity ratio is obtained like as μmol of soluble reducing sugar (cellobiose standard) divided by μmol insoluble sugar (glucose standard). Filter paper was used as substrate for cellulase processivity test. After the filter paper hydrolysis and centrifugation, the soluble reducing sugar in supernatant fraction was measured using the glucose oxidase method as described above. The insoluble reducing sugar in the precipitate

fraction was determined using a modified 2,2-bicinchoninate assay (Doner and Irwin, 1992). The precipitated filter paper was washed with 6 M guanidine hydrochloride to remove bound cellulase, and then washed again four times with assay buffer and water. The reducing sugar of the treated filter paper was measured using Pierce microBCA reagent (USA) with glucose as a standard. The processivity was investigated using 1 nmol of CBM3–Cel9A chimeric enzyme and *S. degradans* Cel5H endoglucanase. Synergistic interaction between CBM3–Cel9A and CbhA towards filter paper was investigated. The reaction mixture was composed of 7 mg of filter paper disc, and equimolar mixture (1:1) of CBM3–Cel9A and CbhA in the reaction buffer containing 50 mM sodium acetate (pH 5.0), 200 mM NaCl, and 10 mM $CaCl_2$. The reaction was carried out at 60 °C for 10 h and the reducing sugar products were measured by DNS method. As a control, the same reaction was performed with Cel9A instead of CBM3–Cel9A. The degree of synergistic effect (DSE) was defined as ratio of the observed activity of combined enzymes to the sum of observed individual activities. All enzyme activity assays were run in triplicate, and data were presented as means ± standard deviation.

2.9. Thin-layer chromatography

One microliter of samples was spotted on a silica 60 TLC plate (Merck, USA), and air dried. Chromatograms were developed using nitromethane, 1-propanol, and water (2:5:1.5) (v/v/v) (Watson et al., 2007). TLC plate was dipped in a mixture of 0.3% (w/v) α-naphthol and 5% (v/v) sulfuric acid in methanol and heated to 110 °C for 10 min to visualize resolved products.

2.10. Macromolecular docking

In order to prepare the individual fused models of Cel9A catalytic domain (CD) with CBM3 and CBM4, macromolecular docking simulations were carried out using the DOT v 2.0 program (Ten Eyck et al., 1995). During the DOT docking calculations, CBMs were assigned as the moving molecules whereas CD was kept as stationary molecule. From protein data bank, the X-ray structures of CD of Cel9A of *A. acidocaldarius* (PDB code: 3H3K), CBM3 (PDB code: 1NBC) and CBM4 (PDB code: 3P6B) of *C. thermocellum* were obtained to be used in this study (Alahuhta et al., 2010; Tormo et al., 1996). The Cel9A is a monomer of 537 amino acid residues consisting of two domains: the Ig-like domain from residues 1 to 85 followed by the catalytic domain from residues 86 to 537. The structure of Cel9A in complex with cello-oligosaccharide was used for study. A cubical grid of 192 Å in size was introduced to accommodate both the moving and stationary molecules. This size was fixed by calculating the longest diameter of the moving and stationary molecules and following the criteria that the grid must be at least 2M + S, where M and S represent the diameters of the moving and stationary molecules. The DOT performs a systematic, rigid-body search of moving molecule translated and rotated about a second stationary molecule. The intermolecular energy from the sum of electrostatic and van der Waals terms was calculated for every configuration of the two molecules. The size of the grid and the number of rotational orientations of the moving molecule used in the docking calculation determines the calculation time but not the number of atoms in the molecules. Finally, the program ranks the docked poses based on the energy terms and top 20 poses are provided.

2.11. Molecular dynamics (MD) simulations

The MD simulations can provide significant insights in understanding molecular mechanism and play critical role in predicting structure and structural motions. It was also proved to be a powerful method to investigate structural and dynamical information of macromolecular structure in atomic details (Braun et al., 2008; Cordomi et al., 2008; Yoon et al., 2008). Initial coordinates for the protein atoms were taken from the CD with no fused CBM (CD-X) and fused structures of CD-CBM3 and CD-CBM4. All of these fused and non-fused structures contain an unaffected pentose moiety at the active site present in CD. The protonation states of all ionizable residues were set to their normal states at pH 7. Three MD simulations were performed for the systems including CD-X, CD-CBM3, and CD-CBM4 structures. All MD simulations were performed with GROMOS96 43a1 force field using GROMACS 4.5.3 package running on a high performance Linux cluster computer (Hess et al., 2008; Van Der Spoel et al., 2005). During the MD simulations, all the protein atoms including metal ions (Zn^{2+}, Ca^{2+}, and Mg^{2+}) were surrounded by a cubic water box of SPC3 water molecules that extended 10 Å from the protein and periodic boundary conditions were applied in all directions. The systems were neutralized with Na^+ or Cl^- counter ions replacing the water molecules and energy minimization was performed using steepest descent algorithm for 10,000 steps. A 100 ps position restrained MD simulations was performed for every system followed by 5 ns production runs with a time step of 2 fs at constant pressure (1 atm), temperature (300 K). The electrostatic interactions were calculated by the PME algorithm and all bonds were constrained using LINCS algorithm. A twin range cutoff was used for long-range interactions including 0.9 nm for van der Waals and 1.4 nm for electrostatic interactions. The snapshots were collected at every 1 ps and stored for further analyses. The system stability and catalytically important structural changes in every system was analyzed by GROMACS, Accelrys Discovery Studio, and PyMol programs.

3. Results and discussion

3.1. Construction and purification of chimeric enzymes

Knowledge of the elementary steps in cellulase action is essential for building enhanced models of cellulose deconstruction that will guide for development of enhanced cellulase systems. Previously, site directed mutagenesis of amino acids surrounding the active site has been performed in order to obtain higher catalytic efficiency of GH9 endoglucanase towards insoluble cellulosic substrates (Li et al., 2010; Ni et al., 2010). In the present study, a fusion protein approach in which CBMs were fused with Cel9A endoglucanase for improved affinity to insoluble cellulosic substrates was developed. The Cel9A endoglucanase has considerable activity towards CMC, lichenan, and cello-oligosaccharides (Eckert et al., 2002). The absence of CBM might be the reason for very low activity towards insoluble cellulosic substrates. In order to generate the efficient chimeric enzymes, the choice of suitable CBM is critical. To date, characteristics of CBMs have been explored extensively (Shoseyov et al., 2006; Tormo et al., 1996). The CBMs from *C. thermocellum* were used for fusion with Cel9A endoglucanase of *A. acidocaldrious*. The reason is that *C. thermocellum* is a gram-positive thermophile like as *A. acidocaldrious* and its cellulolytic enzymes and CBMs are well elucidated. The family 3, 4 and 30 CBMs are commonly found in GH9 cellulases from *C. thermocellum*. The CBM3 of the cellulosome-integrating protein CipA from *C. thermocellum* was bound to crystalline cellulose in a reversible manner (Carrard et al., 2000). The CBM4 in *C. thermocellum* CbhA and CelK was predicted to play an important role in the degradation of plant cell wall (Zverlov et al., 1998). The CBM30 of endoglucanase CelJ has affinity towards both soluble and insoluble cellulosic substrates. The family 3, 4, and 30 CBMs, presenting critical roles in degradation of insoluble cellulose, seem to have different ligand specificity. Thus, the family 3, 4 and 30 CBMs from *C. thermocellum*

were fused separately to N-terminal domain of Cel9A endoglucanase through overlap PCR technique as described in Fig. 1. The fusion location of the CBMs to Cel9A was referred from the protein motif analysis of family 9 cellulases i.e. *C. thermocellum* CbhA, CelK, and CelJ (Fig. S1). The fused genes were subsequently inserted into expression vector pET-28a(+), that encodes the N-terminal His_6 tag. The resulting construct were used for protein expression in *E. coli* BL21(DE3). The expressed proteins were extracted using CelLyticB reagent. The soluble expressions of proteins were checked by SDS–PAGE (Fig. 2). These proteins were purified by using precharged nickel Sepharose column under native conditions. The binding was achieved in presence of 20 mM imidazole and then washed with 1× sodium phosphate buffer (pH 7.5–8.0) containing 40 mM imidazole. Finally, the elution was carried out with 250 mM imidazole. The fractions containing the desired protein were pooled and dialyzed overnight at 4 °C against 50 mM Tris–HCl buffer (pH 7.5). Finally enzymes were stored in 50 mM Tris–HCl buffer (pH 7.5) containing 10% glycerol. Each of the purified preparations gave single band on SDS–PAGE and their molecular sizes were good agreement with those calculated from the nucleotide sequences (Fig. 2).

3.2. Properties of chimeric enzymes

The optimal temperature and pH of the chimeric enzymes were determined at temperature range of 10–90 °C and pH range of 4.0–8.0 (Fig. 3). For comparative purposes, the Cel9A endoglucanase was also subjected to a similar analysis. The results showed that all chimeric enzymes exhibited an optimal pH of 5.0, whereas the native Cel9A endoglucanase displayed an optimal pH of 5.5. Not surprisingly, all native and chimeric enzymes exhibited same optimum temperature of 70 °C. The chimeric enzymes were stable at pH range of 5.0–8.0, when stored overnight at 4 °C (data not shown here). While, there was a little loss of enzyme activity at 60 °C after 1 h. The half life of native Cel9A endoglucanase was 30 min at 75 °C, while those of chimeric enzymes were 20 min at the same temperature.

The chimeric enzymes had broad substrate specificity and efficiently hydrolyzed the soluble cellulosic substrates as well as insoluble ones (Table 1). The catalytic activities of CBM4–Cel9A and CBM30–Cel9A chimeric enzymes towards CMC increased by 1.7 and 1.8 folds compared to that of the native Cel9A endoglucanase.

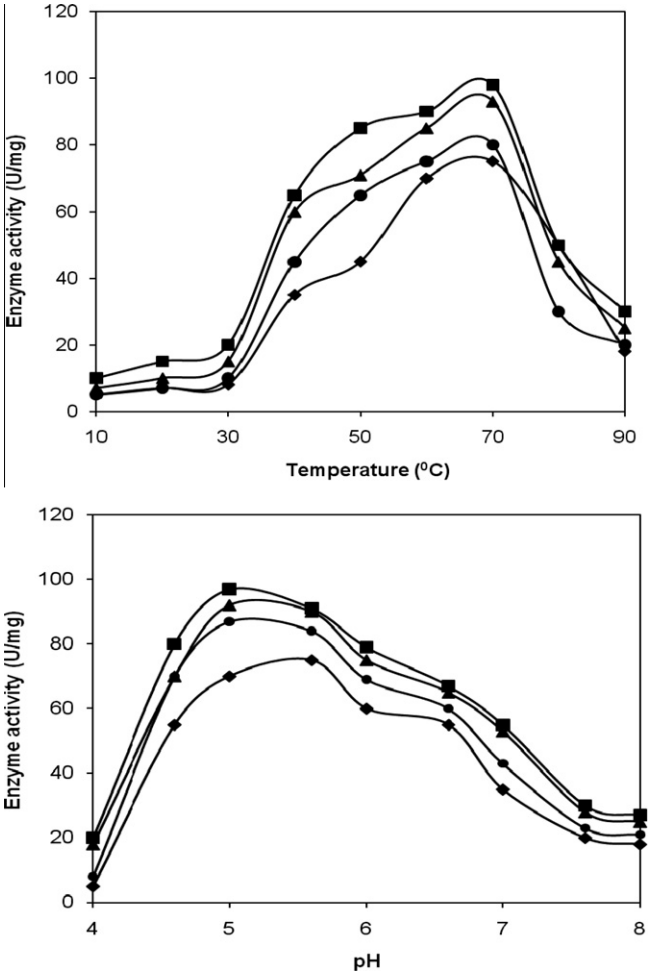

Fig. 3. Optimum pH and temperature for chimeric enzymes. Symbols are indicated as following: Cel9A (squares), CBM3–Cel9A (circles), CBM4–Cel9A (triangles), and CBM30–Cel9A (diamond). Standard deviation errors are within 10%.

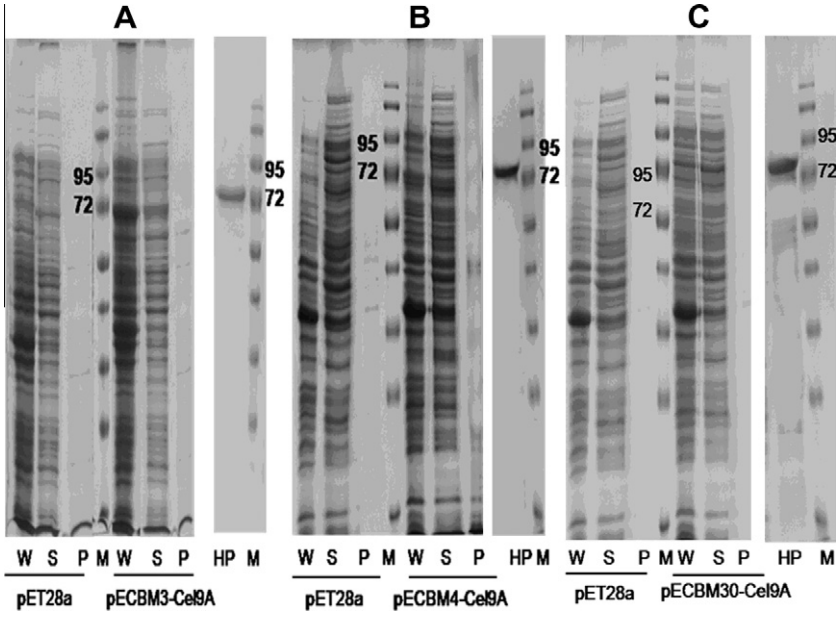

Fig. 2. SDS–PAGE analysis for expression and purification of chimeric enzymes: (A) CBM3–Cel9A, (B) CBM4–Cel9A, and (C) CBM30–Cel9A. Lanes are indicated as follow: (W) whole cell proteins, (S) soluble proteins, (P) cell pellet proteins, (HP) His-tag purified protein, and (M) molecular weight markers.

It was also found that 2 folds increase in catalytic activities of the chimeric enzymes towards PASC compared to the native enzyme. The catalytic activities of the chimeric enzymes towards insoluble polysaccharides (Avicel and filter paper) were significantly enhanced by 7.7–12.0 folds compared to those of the native enzyme. However, there was no significant increase in the chimeric enzymes activities towards *p*-NPC. No detectable activity towards *p*-NPG was observed in all enzymes. Among all chimeric enzymes, CBM3–Cel9A showed the highest catalytic activity towards insoluble polysaccharides. Avicel binding affinity of the chimeric enzymes was determined to define its correlation with enhanced catalytic activity towards insoluble polysaccharides. It was found that the chimeric enzymes showed higher binding affinity towards Avicel as compared to the native Cel9A endoglucanase. The binding ratios of Cel9A, CBM3–Cel9A, CBM4–Cel9A and CBM30–Cel9A were 5%, 50%, 35% and 37%, respectively (data not shown). It has been claimed that the efficiency of cellulases is directly related to their affinity for the substrates (Carrard et al., 2000; Klyosov, 1990). In according to it, CBM3–Cel9A presented the highest binding affinity and hydrolysis efficiency towards Avicel among the chimeric enzymes. The CBMs of chimeric enzymes was suspected to determine the binding affinity. The CBM3 contains a flat cellulose binding surface and the aromatic residues present in the flat surface are involved in the interaction with hydrophobic region of crystalline cellulosic substrates (Tormo et al., 1996). The CBM4 contains cellulose binding cleft, in which the aromatic residues of Trp68 and Trp118 contribute to the binding to amorphous region of crystalline cellulose (Alahuhta et al., 2010). The CBM30 binds to insoluble cellulose but the cellulose binding mechanism is still unknown (Arai et al., 2003). The target binding sites of the family 3, 4 and 30 CBMs are different, which might be responsible for different enzyme activity pattern of the chimeric enzymes to cellulosic substrates.

3.3. Analysis of hydrolysis products and processivity

The hydrolysis products of CMC, PASC, filter paper, Avicel, cellotetraose and *p*-NPC by CBM3–Cel9A were qualitatively analyzed by TLC (Fig. 4). Cellobiose and cellotetraose were detected after hydrolysis of CMC, PASC, filter paper, and Avicel. The chimeric enzyme was suspected to produce cellobiose and cellotetraose simultaneously as cleavage products. In the time course reaction of PASC hydrolysis (Fig. 4B), the intensity of cellobiose spots on TLC increased as the increase of reaction time from 30 to 150 min, whereas there was no significant change of the cellotetraose spots intensity. It indicated cellotetraose was further degraded into cellobiose by the chimeric enzyme. When cellotetraose and *p*-NPC were used as substrates (Fig. 4D and E), the chimeric enzyme efficiently hydrolyzed the substrates with concomitant release of cellobiose. Previously, it was claimed that the type of cleavage products is determined by the active site configuration of Cel9A and the strong binding affinity of the −2 and −1 subsites towards two glucose moieties is responsible for the production of cellobiose as a major cleavage product (Eckert et al., 2009). The report also suggested the production possibility of cellotetraose as a cleavage product, based on the crystal structure of Cel9A endoglucanase bound to cellotetraose at the binding subsites of −4 and −1. Therefore, the analytical results on hydrolysis products of cellulosic substrates by CBM3–Cel9A are well corresponded to the previous report. It suggests that chimeric enzyme tends to hydrolyze cellulose to cellobiose, which is different to a typical character of endoglucanase generating cellodextrin as a main hydrolysis product. The chimeric enzyme CBM3–Cel9A has both endo and exo types of cellulase activity and shows unique mechanism for cellulose substrate hydrolysis which will give new insight for cellulase engineering. The production of cellobiose could be an indication of processivity. To evaluate the processivity of CBM3–Cel9A chimeric enzyme, the concentration of soluble and insoluble sugar were measured after hydrolysis of filter paper and compared with the known processive endoglucanase of Cel5H from *S. degradans*. The CBM3–Cel9A chimeric enzyme produced 81% soluble reducing sugar and 19% insoluble reducing sugar after 2 h hydrolysis of filter paper. The Cel5H reference enzyme produced 80% and 20% of soluble and insoluble reducing sugars, respectively. Typical endoglucanases have been known to produce 60–70% of soluble reducing sugars from filter paper substrate, whereas the engineered CBM3–Cel9A endoglucanase produced 81% of soluble reducing sugar with the processive ratio of 4.26. Processive endoglucanases have been reported to have processive values of 3.1–7.0 (Li et al., 2007). Thus, the fused CBM3 was suspected to prevent desorption of Cel9A endoglucanase from cellulose, and to force the enzyme to progress continuously along the cellulose surface and to proceed with high processivity. Cellulases that can produce high percentage of soluble fermentable sugar are required for cost effective fermentation of cellulose carbon source and the chimeric cellulase CBM3–Cel9A is a promising cellulase for the goal.

3.4. Synergistic interaction between CbhA and CBM3–Cel9A chimeric enzyme

Processive endoglucanase has been known to have synergetic effect with cellobiohydrolase on cellulose degradation (Vazana et al., 2010). Cellobiohydrolase (CbhA) from thermophilic bacterium, *C. thermocellum* was used to see the synergetic effect with CBM3–Cel9A and native Cel9A (Fig. S2). The combined activity of the coupled enzymes was greater than the sum of the individual enzymes activities. The CBM3–Cel9A chimeric enzyme showed 2.15° of synergistic effect (DSE) with CbhA on degradation of filter paper, whereas the native Cel9A enzyme had 1.5° of synergistic effect (DSE) with CbhA. It suggests that the CbhA presents the higher synergistic effect with CBM3–Cel9A chimeric enzyme as compared to the native Cel9A. Complete cellulose hydrolysis has been performed by combined action of multiple enzymes, viz. cellobiohydrolase, endoglucanase, cellodextrinase and β-glucosidase. The synergistic action among the multiple enzymes is desired for

Table 1
Substrate specific activity of native and chimeric enzymes towards soluble and insoluble cellulosic substrates.

Substrates	Specific activity (U/μmol)			
	Cel9A (59 kDa)	CBM3–Cel9A (76 kDa)	CBM4–Cel9A (80 kDa)	CBM30–Cel9A (80 kDa)
CMC	4325 ± 70	6600 ± 90	7320 ± 80	7720 ± 90
Barley β-glucan	11,800 ± 200	12,800 ± 190	15,000 ± 170	15,000 ± 180
PASC	649 ± 50	1000 ± 60	1260 ± 70	1280 ± 70
Avicel	0.50 ± 0.09	6.0 ± 0.9	4.0 ± 0.7	5.0 ± 0.8
Filter paper	0.70 ± 0.08	7.0 ± 1.0	5.0 ± 1.0	6.0 ± 0.9
p-NPG	ND[a]	ND[a]	ND[a]	ND[a]
p-NPC	118 ± 30	120 ± 30	130 ± 20	135 ± 20

[a] No activity was detected.

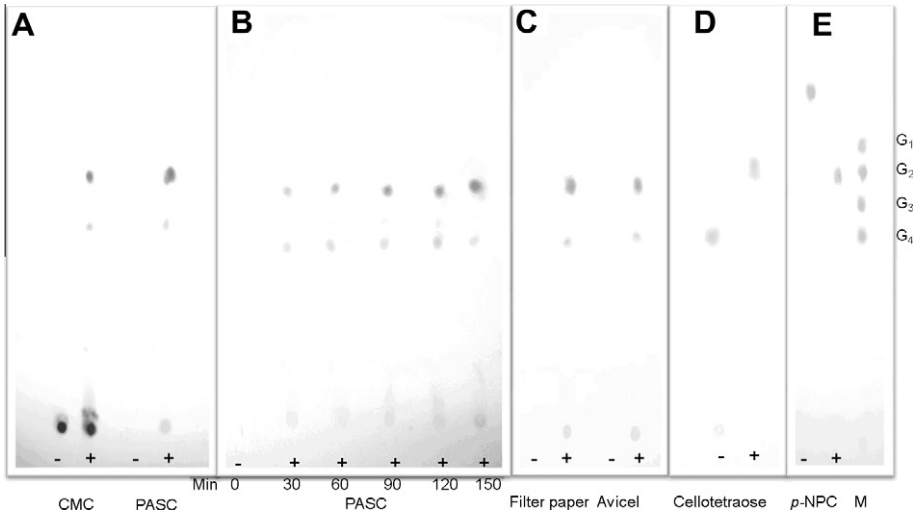

Fig. 4. TLC analyses of hydrolysis products for PASC, CMC, filter paper, Avicel, cellotetraose and *p*-NPC. (A) Extensive (5 h) hydrolysis products of CMC and PASC. (B) Time course of hydrolysis products of PASC from 0 to 150 min. (C) Extensive (16 h) hydrolysis products of filter paper and Avicel. (D) Extensive (2 h) hydrolysis product of cellotetraose. (E) Extensive (2 h) hydrolysis product of *p*-NPC and M: Standard marker, whereas G1 to G4 represent glucose, cellobiose, cellotriose, and cellotetraose, respectively. The PASC, CMC, cellotetraose and *p*-NPC were treated with 0.1 nmol of CBM3–Cel9A at 60 °C. The same reaction was performed with Avicel and filter paper with 1.0 nmol of CBM3–Cel9A. The reactions were performed in absence (−) and presence (+) of the enzyme.

an effective cellulose hydrolysis. Therefore, the processive chimeric enzyme can be a potential candidate for development of a free cellulases mixture (cocktail) for complete hydrolysis of cellulosic substrates to fermentable sugar as compared to the wild type. In order to minimize the cost of cellulases cocktail, it is necessary to minimize the number of components in cellulases cocktail that are employed for cellulose hydrolysis. An approach is use of multifunction cellulases harboring both exo and endo activities of cellulose hydrolysis and the chimeric cellulase CBM3–Cel9A can be a potential multifunction cellulase for it. We are investigating the synergistic effect of CBM3–Cel9A with other cellulases to design an efficient cellulases cocktail of three components (i.e. cellobiohydrolase, CBM3–Cel9A, and β-glucosidase) for cost effective conversion of cellulose substrates to fermentable sugar.

3.5. Macromolecular docking and selection of binding poses

Driven by the wealth of new experimental data on cellulases, computer simulations are beginning to play an increasingly significant role in understanding the structure–function relationships of enzymatic cellulose deconstruction. Previous computational analysis work suggested that a critical function of the Ig-like domain was to ensure the proper folding of catalytic domain to assist the proper orientation of key active site and catalytic residues (Liu et al., 2010). The wet-lab experimental results suggest that the fused CBMs were enhanced the catalytic activity of Cel9A. In order to investigate the effect of the fused CBMs on the active site of Cel9A and their interaction with Ig-like domain, the full structures of CBM3–Cel9A and CBM4–Cel9A chimeric enzymes were constructed by macromolecular docking and computational modeling. The macromolecular docking procedure has ranked top 20 energetically favorable poses based on the electrostatic and van der Waals energy terms (Table S2). These 20 poses were filtered from 2000 configurations that were predicted initially. The CBMs were fused to CD through its N-terminal end and thus the docked poses were selected in such a way that the fusion would be possible. The 19th docked pose from the 20 poses sorted by its final energy terms was selected as the best pose since the C-terminal of the CBM3 and N-terminal of CD were closely located. The final sum of energy of the selected pose of CD-CBM3 docked model was −9.59 kcal/mol. No other docked pose within the top 20 poses has shown the fusing terminals close to each other, which was a prerequisite for the fusion. In terms of CD-CBM4 fusion, the second top pose from the top 20 docked poses ranked based on the energy terms was selected as the best pose to be used in the fusion and further study (Table S3). The final energy of this selected pose is −17.74 kcal/mol, very close to that of the top ranked pose (−21.28 kcal/mol). Finally, the selected poses of CD-CBM3 and CD-CBM4 structures were used in preparing the fused models by joining the N-terminal of CD and C-terminal of CBMs (Fig. S3). These fused models were subjected to MD simulations in order to refine the binding poses and to observe the structural changes due to the fused CBM parts.

3.6. Molecular dynamics simulations and basic MD analyses

The basic check for system stability can be normally achieved by the root mean square deviation (RMSD) plotting over the time. In this study, the RMSD values of Cα atoms of all the systems were calculated to investigate the changes upon fusion of two different CBMs to the CD. The CD of the enzyme includes an Ig-like domain in its N-terminal region and the fused CBMs interact through this domain. The comparison of RMSD values between the CDs of CD-X, CD-CBM3 and CD-CBM4 systems revealed the extent of structural deviation occurred in CDs of fused models (Fig. S4). The observed difference in RMSD values between CBM parts of CD-CBM3 and CD-CBM4 fused models has shown that these modules are indeed highly deviant to each other. Further, the root mean square fluctuations (RMSF) were calculated and compared with each other to observe the highly flexible regions of the proteins (Fig. S5). The high fluctuation regions were observed in the CBM part along with slight deviations in CD regions. The active site key residues D143, D146, and E515, which are important for glycoside hydrolysis, have not shown much fluctuation during the simulation in all three systems and showed similar RMSF values. The key catalytic residues, D143, D146, and E515 are involved in the hydrolysis of sugar along with two tryptophan residues, namely, W343 and W401 in the active site. The W343 is located perpendicular and the W401 is located parallel to the bound sugar moiety. The representative structures taken from the last 2 ns of the simulations were superimposed to observe the structural changes specially focused in the active site. It was observed that the conformation of D143, D146 and E515 residues were similar to

each other in all three systems (Fig. S6). The number of hydrogen bonds between the Ig-like domain and rest of the CD was calculated along with the number of hydrogen bonds between the fused CBM and Ig-like domain to observe difference between the CBM3 and CBM4 bindings (Fig. S7). The number of hydrogen bond between Ig-like domain and rest of the CD has been maintained throughout the simulation time (Fig. S7). The computational results indicate that interaction between the Ig-like domain and the rest of Cel9A is not changed by fusion of CBMs in the computational protein models and the fused CBMs do not have a negative effect on the active site configuration of Cel9A, but bring the positive stimulatory effect on hydrolysis of cellulose substrate.

4. Conclusion

Chimeric cellulases were successfully constructed and expressed in *E. coli*. Chimeric cellulases displayed enhanced binding affinity and hydrolytic activity towards Avicel and filter paper. The fused CBMs play a critical accessory role for catalytic module and change its character to facilitate processive cleavage of cellulosic substrates. Chimeric cellulase of CBM3-Cel9A presented a synergistic interaction with cellobiohydrolase for cellulose hydrolysis and could be a potential component of effective cellulase cocktail. *In silico* experimental results suggest that fused CBMs do not have a negative effect on the active site configuration of Cel9A, but bring positive stimulatory effect on hydrolysis of cellulose substrate.

Acknowledgements

This work was supported by a grant from the International Collaborative R&D Program of Knowledge Economy Technology Innovation Program, MKE, and a Grant (NRF-2010-C1AAA001-0029084) from the National Research Foundation, MEST, and the financial support from the KIST through Future Key Technology Program, Korea. One of the authors (Dr. Amar A. Telke) is supported by postdoctoral fellowship from the BK21 Program, MEST, Korea.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.biortech.2012.02.066.

References

Alahuhta, M., Xu, Q., Bomblea, Y., Brunecky, R., Adneya, W., Dinga, S., Himmela, M., Lunin, V., 2010. The unique binding mode of cellulosomal CBM4 from *Clostridium thermocellum* cellobiohydrolase A. J. Mol. Biol. 402, 374–387.

Arai, T., Araki, R., Tanaka, A., Karita, S., Kimura, T., Sakka, K., Ohmiya, K., 2003. Characterization of a cellulase containing a family 30 carbohydrate-binding module (CBM) derived from *Clostridium thermocellum* CelJ: importance of the CBM to cellulose hydrolysis. J. Bacteriol. 185, 504–512.

Boraston, A.B., Bolam, D.N., Gilbert, H.J., Davies, G.J., 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. 382, 769–781.

Braun, G.H., Jorge, D.M., Ramos, H.P., Alves, R.M., da Silva, V.B., Giuliatti, S., Sampaio, S.V., Taft, C.A., Silva, C.H., 2008. Molecular dynamics, flexible docking, virtual screening, ADMET predictions, and molecular interaction field studies to design novel potential MAO-B inhibitors. J. Biomol. Struct. Dyn. 25, 347–355.

Carrard, G., Koivula, A., Soderlund, H., Beguin, P., 2000. Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. Proc. Natl. Acad. Sci. USA 97, 10342–10347.

Cordomi, A., Ramon, E., Garriga, P., Perez, J.J., 2008. Molecular dynamics simulations of rhodopsin point mutants at the cytoplasmic side of helices 3 and 6. J. Biomol. Struct. Dyn. 25, 573–587.

Cosgrove, D.J., 2000. Loosening of plant cell walls by expansions. Nature 407, 321–326.

Doner, L.W., Irwin, P.L., 1992. Assay of reducing end-groups in oligosaccharide homologues with 2,2′-bicinchoninate. Anal. Biochem. 202, 50–53.

Eckert, K., Zielinski, F., Lo Leggio, L., Schneider, E., 2002. Gene cloning, sequencing, and characterization of a family 9 endoglucanase (CelA) with an unusual pattern of activity from the thermoacidophile *Alicyclobacillus acidocaldarius* ATCC27009. Appl. Microbiol. Biotechnol. 60, 428–436.

Gaudin, C., Belaich, A., Champ, S., Belaich, J.P., 2000. CelE, a multidomain cellulase from *Clostridium cellulolyticum*: a key enzyme in the cellulosome? J. Bacteriol. 182, 1910–1915.

Ghose, T., 1987. Measurement of cellulase activities. Pure Appl. Chem. 59, 257–268.

Hess, B., Kutzner, C., van der Spoel, D., Lindahl, E., 2008. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J. Comput. Chem. 4, 435–447.

Klyosov, A.A., 1990. Trends in biochemistry and enzymology of cellulose degradation. Biochemistry 29, 10577–10585.

Li, Y., Irwin, D.C., Wilson, D.B., 2007. Processivity, substrate binding, and mechanism of cellulose hydrolysis by *Thermobifida fusca* Cel9A. Appl. Environ. Microbiol. 73, 3165–3172.

Li, Y., Irwin, D.C., Wilson, D.B., 2010. Increased crystalline cellulose activity via combinations of amino acid changes in the family 9 catalytic domain and family 3c cellulose binding module of *Thermobifida fusca* Cel9A. Appl. Environ. Microbiol. 76, 2582–2588.

Liu, H., Pereira, J.H., Adams, P.D., Sapra, R., Simmons, B.A., Sale, K.L., 2010. Molecular simulations provide new insights into the role of the accessory immunoglobulin-like domain of Cel9A. FEBS Lett. 584, 3431–3435.

Ni, J., Takehara, M., Watanabe, H., 2010. Identification of activity related amino acid mutations of a GH9 termite cellulase. Bioresour. Technol. 101, 6438–6443.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual, third ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Shoseyov, O., Shani, Z., Levy, I., 2006. Carbohydrate binding modules: biochemical properties and novel applications. Microbiol. Mol. Biol. Rev. 70, 283–295.

Ten Eyck, L.F., Mandell, J., Roberts, V.A., Pique, M.E., 1995. Surveying molecular interactions with DOT. In: Proceedings of the 1995 ACM/IEEE Supercomputing Conference, San Diego, IEEE Computer Society Press, Los Alamitos, California, USA.

Tormo, J., Lamed, R., Chirino, A., Morag, E., Bayer, A., Shoham, Y., Steitz, T., 1996. Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. EMBO J. 15, 5739–5751.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., Berendsen, H.J., 2005. GROMACS: fast, flexible, and free. J. Comput. Chem. 26, 1701–1718.

Vazana, Y., Morais, S., Barak, Y., Lamed, R., Bayer, E.A., 2010. Interplay between *Clostridium thermocellum* family 48 and family 9 cellulases in cellulosomal versus noncellulosomal states. Appl. Environ. Microbiol. 76, 3236–3243.

Watson, B.J., Zhang, H., Longmire, A.G., Moon, Y.H., Hutcheson, S.W., 2009. Processive endoglucanases mediate degradation of cellulose by *Saccharophagus degradans*. J. Bacteriol. 191, 5697–5705.

Wood, T.M., 1971. The cellulase of *Fusarium solani*. Purification and specificity of the -(1–4)-glucanase and the -D-glucosidase components. Biochem. J. 121, 353–362.

Yoon, J., Park, J., Jang, S., Lee, K., Shin, S., 2008. Conformational characteristics of unstructured peptides: alpha-synuclein. J. Biomol. Struct. Dyn. 25, 505–515.

Zhang, Y., Chen, S., Xu, M., Cavaco-Paulo, A., Wu, J., Chen, J., 2010. Characterization of *Thermobifida fusca* cutinase-carbohydrate-binding module fusion proteins and their potential application in bioscouring. Appl. Environ. Microbiol. 76, 6870–6876.

Zhang, Y.H., Cui, J., Lynd, L.R., Kuang, L.R., 2006. A transition from cellulose swelling to cellulose dissolution by *o*-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. Biomacromolecules 7, 644–648.

Zverlov, V.V., Velikodvorskaya, G.V., Schwarz, W.H., Bronnenmeier, K., Kellermann, J., Staudenbauer, W.L., 1998. Multidomain structure and cellulosomal localization of the *Clostridium thermocellum* cellobiohydrolase CbhA. J. Bacteriol. 180, 3091–3099.