Title: Mini-scaffoldin enhanced mini-cellulosome hydrolysis performance on low-accessibility cellulose (Avicel) more than on high-accessibility amorphous cellulose

Authors: Chun You, Xiao-Zhou Zhang, Y.-H. Percival Zhang

PII: S1369-703X(12)00029-0
DOI: doi:10.1016/j.bej.2012.01.011
Reference: BEJ 5446

To appear in: *Biochemical Engineering Journal*

Received date: 13-10-2011
Revised date: 26-1-2012
Accepted date: 30-1-2012

Please cite this article as: C. You, X.-Z. Zhang, Y.-H.P. Zhang, Mini-scaffoldin enhanced mini-cellulosome hydrolysis performance on low-accessibility cellulose (Avicel) more than on high-accessibility amorphous cellulose, *Biochemical Engineering Journal* (2010), doi:10.1016/j.bej.2012.01.011

Mini-scaffoldin enhanced mini-cellulosome hydrolysis performance on low-accessibility cellulose (Avicel) more than on high-accessibility amorphous cellulose

Chun You, [1] Xiao-Zhou Zhang, [1,2] Y.-H. Percival Zhang [1,2,3,4 *]

[1] Biological Systems Engineering Department, Virginia Tech, 210-A Seitz Hall, Blacksburg, Virginia 24061, USA

[2] Gate Fuels Inc., 3107 Alice Drive, Blacksburg, Virginia 24060, USA

[3] Institute for Critical Technology and Applied Science (ICTAS), Virginia Tech, Blacksburg, Virginia 24061, USA

[4] DOE BioEnergy Science Center (BESC), Oak Ridge, Tennessee 37831, USA

[*]Corresponding author: YPZ, Email: ypzhang@vt.edu,

Tel: (540) 231-7414 [O], Fax: (540) 231-3199

Running title: Cellulose hydrolysis by synthetic cellulosomes

Abstract

A glycoside hydrolase family 5 *Bacillus subtilis* endoglucanase, a family 9 *Clostridium thermocellum* processive endoglucanase, and a family 48 *Clostridium phytofermentans* ISDg cellobiohydrolase were assembled together by the high-affinity interaction between three cohesins in a mini-scaffoldin (mini-CipA) and dockerins in three cellulases, forming the mini-cellulosome. This mini-cellulosome exhibited enhanced hydrolytic activity on low-accessibility cellulose (microcrystalline cellulose, Avicel) and high accessibility cellulose (regenerated amorphous cellulose, RAC) as compared to the non-complexed cellulase mixture at the same enzyme amount. The stimulation factors (SF, i.e., activity ratios of the mini-cellulosome to the non-complexed cellulase mixture) were larger on Avicel than on RAC regardless of substrate/enzyme ratios. Also, SF increased when substrate/enzyme ratio increased. The different hydrolysis patterns of the mini-cellulosome and cellulase mixture on Avicel and RAC suggested that the construction of synthetic cellulosomes would be an efficient way to significantly enhance cellulose hydrolysis rate and digestibility, especially in the case of low-accessibility recalcitrant cellulose at low enzyme usage.

Keywords: Biofuels; Cellulase engineering; Enzymatic cellulose hydrolysis; Synthetic cellulosome; Synergy

1. Introduction

Lignocellulosic biomass is the most abundant renewable bioresource [1]. The utilization of a small fraction of collectable low-cost cellulosic materials, including crop residues (e.g., corn stover) as well as dedicated bioenergy grass and wood, would produce a significant fraction of sustainable transportation biofuel so that it would decrease reliance on crude oil, enhance energy security, and decrease net greenhouse gas emissions [2]. Cost-effective sugar release from recalcitrant lignocellulose, however, remains challenging [3, 4]. During biomass saccharification, low mass-specific activity cellulase results in a large use of cellulase [5]. The weight ratio of substrate to enzyme (i.e., [S]/[E]) for cellulose hydrolysis is at least one order of magnitude higher than that for starch hydrolysis, resulting in higher enzyme cost [5]. Therefore, increasing specific activity of cellulase would reduce enzyme usage so to enhance overall economy of biomass-based biorefineries. Mass-specific activity of cellulase can be enhanced by several approaches: improvement in individual components by directed evolution [6, 7] and rational design [8], reconstitution of non-complexed cellulase cocktails [1, 9, 10], and construction of complexed cellulases (called synthetic cellulosomes) [11-18].

Enzymatic hydrolysis of cellulose requires synergetic action among endoglucanase (EG, EC 3.2.1.4), cellobiohydrolase (CBH) (EC 3.2.1.91), and beta-glucosidase (EC 3.2.1.21). In the *Trichoderma* fungal cellulase system, the dominant components are EG I (Cel7B) & III (Cel5A), cellobiohydrolase I (Cel7A) and cellobiohydrolase II (Cel6A), suggesting that these components play a central role in hydrolyzing cellulose. EG cuts accessible β-1,4-glucosidic bond of cellulose chains randomly. CBH I and CBH II act on reducing end and non-reducing end, respectively. It is thought that the respective cellulase component in bacterial cellulase system are glycoside hydrolase family 5 endoglucanase Cel5, family 48 cellobiohydrolase (Cel48)

acting on reducing end, and family 9 processive endoglucanase (Cel9) acting on non-reducing end [10].

Anaerobic bacteria and fungi often produce complexed cellulases -- cellulosomes, whose catalytic units are linked by non-hydrolytic scaffoldins [19-22]. Inspired by natural cellulosomes, Bayer and his coworkers have proposed to construct designer cellulosomes with tailored subunit components through the high-affinity interaction between cohesins and dockerins [23]. A number of building blocks, including cohesins, dockerins, catalytic modules (e.g., cellulases and hemicellulases), carbohydrate binding modules (CBMs), and linkers, have been reassembled into various designer cellulosomes *in vitro* [11-15]. However, most of these designer cellulosome studies focused on the assembly of different cellulosomes and demonstrated enhanced synergy due to scaffoldins, few studies attempted to investigate the ratios of substrate to enzyme on the stimulation effects of cellulosomes on different substrates in that decreasing cellulase usage (i.e., mg cellulase per g of cellulose) is vital to cost-effective sugar release from pretreated cellulosic materials [5]. Only Fierobe et al. (2002) reported that the stimulation effects of mini-cellulosomes on Avicel and bacterial cellulose, and found that stimulation effects were inversely related to substrate concentration at a fixed cellulase concentration [15]. It was vital and interesting to study the in-depth relationship between different action mode cellulase components linked by scaffoldins or not, and cellulose characteristics.

Substrate accessibility to cellulase may be the most important substrate characteristic impacting enzymatic cellulose hydrolysis, as compared to degree of polymerization (DP), crystallinity, pore volume, particle size, and so on [4, 10, 24]. Microcrystalline cellulose (Avicel) is made from wood pulp by acid hydrolysis that can remove most amorphous cellulose and all hemicellulose, but it still contains a significant fraction of amorphous cellulose [10, 25].

Regenerated amorphous cellulose (RAC) is prepared from Avicel through a series of steps: cellulose slurrying in water, cellulose dissolution in concentrated phosphoric acid, and regeneration in water [26]. As a result, RAC, a completely disordered insoluble substrate, has approximately 20 times cellulose accessibility of that of Avicel [10, 27], while it has the same DP when ice-cold concentrated phosphoric acid is used [28]. Avicel and RAC represent two extreme model cellulosic materials featuring very low and very high substrate accessibility, respectively [10]. Most pretreated cellulosic materials produced by dilute acid pretreatment, steam explosion, hot water, cellulose solvent-based lignocellulose fractionation, soaking in aqueous ammonia have substrate accessibility falling between those of Avicel and RAC [4, 10]. Therefore, the study of cellulosome and non-complexed cellulase hydrolysis on these two model cellulosic materials at different [S]/[E] ratios would help to develop advanced enzyme systems so as to decrease cellulase usage.

In this study, we assembled the designer mini-cellulosome containing three bacterial cellulases -- an endoglucanase Cel5, a processive endoglucanase Cel9, and a cellobiohydrolase Cel48. Family 5 endoglucanase (BsCel5), family 9 processive endoglucanase (CtCel9) and family 48 cellobiohydrolase (CpCel48) were obtained from *B. subtilis* [7], *C. thermocellum* [10] and *C. phytofermentans* ISDg [29], respectively. A mini-scaffoldin (mini-CipA) containing three type I cohesins and one family 3 CBM was truncated from *C. thermocellum* CipA [19, 30, 31]. The stimulation effects of this designer mini-cellulosome were investigated on two model substrates at different [S]/[E] ratios compared to the non-complexed mixture.

2. Materials and Methods

2.1. Chemicals

All chemicals were reagent grade or higher, purchased from Sigma (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA), unless otherwise noted. Microcrystalline cellulose –Avicel PH105 (20 μm) – was purchased from FMC (Philadelphia, PA). RAC was prepared from Avicel as previously described [26]. The oligonucleotides were synthesized by Integrated DNA Technologies (Coraville, IA).

2.2. Strains and medium

Escherichia coli JM109 was used as a host cell for DNA manipulation, and *E. coli* BL21 Star (DE3) (Invitrogen, Carlsbad, CA) was used as a host cell for recombinant protein expression. The Luria-Bertani (LB) medium was used for *E. coli* cell culture and recombinant protein expression. Ampicillin (100 μg/mL) was added in the LB medium.

2.3. Construction of plasmids

The sequences of all PCR primers used are listed in Table 1. Plasmids and recombinant proteins are summarized in Fig. 1. All plasmid sequences were verified by DNA sequencing. The DNA sequence encoding the truncated mini-CipA (26-723 amino acids, GenBank Accession number: L08665) was amplified based on the genomic DNA of *C. thermocellum* ATCC 27405 by a primer pair of mini-CipA_For and mini-CipA_Rev. The PCR product was digested with NdeI/XhoI and ligated into NdeI/XhoI-digested pET20b (Novagen, Madison, WI), yielding pET20b-mini-CipA. The dockerin module was added to the C-terminus of Cpcel48 by overlap-extension PCR. The DNA encoding the catalytic domains of CpCel48 (GenBank Accession number: ABX43721) with or without CBM were amplified from pP43N-Cpcel48 [29] by two primer pairs of CpCel48_For and CpCel48_Rev as well as CpCel48_For and CpCel48'_Rev,

respectively. The DNA fragment encoding a dockerin domain (DocS , 673 -741 amino acids) from CtCelS (GenBank Accession number: L06942) was amplified from the genomic DNA of *C. thermocellum* by a primer pair of DocS_For and DocS_Rev. The two resultant overlapping fragments were mixed as the template for the next round PCR by a primer pair of CpCel48_For/DocS_Rev. These resultant fragments were cloned into NdeI/XhoI-digested pET20b, thereby generating pET20b-Cpcel48 and pET20b-Cpcel48'. pET20b-Ctcel9 and pET20b-Ctcel9' were also obtained using PCR amplification and overlap-extension polymerase chain reaction method. The DNA encoding the mature CtCel9 (28-739 amino acids, GenBank Accession number: CAA43035) was amplified from the genomic DNA of *C. thermocellum* by a primer pair of CtCelF_For and CtCelF_Rev. The PCR product was digested with NdeI/XhoI and ligated into NdeI/XhoI-digested pET20b, generating pET20b-Ctcel9. The DNA encoding the CtCel9' (CtCel9 without CBM3c) was also amplified from the genomic DNA of *C. thermocellum* by a primer pair of CtCelF_For and CtCelF'_Rev, the DNA encoding a dockerin domain (DocF from CtCelF, 669-739 amino acids) was amplified from the genomic DNA of *C. thermocellum* by a primer pair of DocF_For and CtCelF_Rev. The two resultant overlapping fragments were mix as the PCR template by a primer pair of CtCelF_For and CtCelF_Rev. The resultant fragment was digested with NdeI/XhoI and cloned into NdeI/XhoI-digested pET20b, thereby generating pET20b-Ctcel9'. Plasmid of pET20b-Bscel5 and pET20b-Bscel5' were constructed in the same way as pET20b-Cpcel48 and pET20b-Cpcel48'. The DNA sequence encoding mature BsCel5 (30–499 amino acids, GenBank Accession number: CAA82317) was amplified from the genomic DNA of *B. subtilis* 168. The dockerin of BsCel5 and BsCel5' was DocK (821-895 amino acids) from CtCelK (NCBI Reference Sequence: YP_001036843) of *C. thermocellum*.

151

152 *2.4. Protein expression and purification*

153 *E. coli* BL21 Star (DE3) cell containing the plasmid was cultivated in 150 mL of LB
154 medium supplemented with 1.2% glycerol at 37 °C in 500 mL flask at a rotary rate of 220 rpm.
155 When A_{600} reached approximately 0.75, 100 μM isopropyl-beta-D-thiogalactopyranoside (IPTG,
156 a final concentration) was added and the cultivation temperature was decreased to 16 °C for ~16
157 h. After centrifugation at 5,000 rpm for 10 min, the cell pellets were resuspended in 40 mL of 50
158 mM HEPES buffer (pH 7.5) containing 5 mM imidazole and 50 mM NaCl. The cells were lysed
159 by ultrasonication. After centrifugation at 12,000 rpm for 15 min, 10 μl of the supernatant was
160 applied to SDS-PAGE for checking protein expression levels. The supernatant of the crude cell
161 lysate was applied to a column containing 2 mL of nickel-NTA resin (Promega, Madison, WI)
162 and then the resin was washed with 10 mL of 50 mM HEPE buffer (pH 7.5) containing 20 mM
163 imidazole and 50 mM NaCl. The bound proteins were eluted by 5 mL of 50 mM HEPES buffer
164 (pH 7.5) containing 250 mM imidazole and 50 mM NaCl. The enzyme was desalted by a PD-10
165 desalting column (Amersham, Piscataway, NJ) and then stored at -20 °C in a 50 mM HEPES
166 buffer (pH 7.5) containing 50 mM NaCl and 5% glycerol.

167

168 *2.5. RAC affinity pull-down analysis*

169 For determination of the composition of the adsorbed cellulosome, 5 μg of mini-CipA
170 was mixed with CtCel9'at various molar ratios of 1:1, 1:3, 1:6, and 0:3 in 50 mM HEPES (pH

171 7.5) containing 50 mM NaCl and 10 mM $CaCl_2$ for 5 min. The mini-cellulosome was adsorbed
172 by 100 μg of RAC at 0°C for 10 min. After centrifugation at 5,000 rpm for 10 min, the protein in
173 the supernatant was precipitated by 10% trichloroacetic acid (TCA). After washing in ice-cold
174 acetone once, the protein pellets were dissolved in 40 μl of 1 x SDS-loading buffer, and 10 μl of
175 the sample was loaded to SDS-PAGE gel. The RAC pellets were washed with 1 mL of 50 mM
176 HEPES (pH 7.5) containing 300 mM NaCl and 10mM $CaCl_2$ two times, followed by washing
177 with 1 mL of ice-cold water two times. After centrifugation at 5,000 rpm for 10 min, the RAC-
178 containing pellets containing the adsorbed mini-cellulosome were resuspended in 1 × SDS-
179 loading buffer in final volume of 40 μL. After boiling for 5 min and centrifugation at 5,000 rpm
180 for 10 min, 10 μL of the supernatant was loaded to SDS-PAGE gel. RAC affinity pull-down was
181 also conducted to check the composition of the four component mini-cellulosomes.
182
183 *2.6. Native PAGE*
184 The native gel included the separating gel and stacking gel. The 12%
185 acrylamide/bisacrylamide separating gel contained 2 mL of H_2O, 1.2 mL of 40%
186 acrylamide/bisacrylamide, 0.75 mL of 2 M Tris-HCl (pH 8.8), 40 μL of 10% ammonium
187 persulfate, and 3 μL of tetramethylethylenediamine (TEMED). The 5%
188 acrylamide/bisacrylamide stacking gel contained 1.1 mL of H_2O, 190 μL of 40%
189 acrylamide/bisacrylamide, 94 μL of 2M Tris-HCl (pH8.8), 15 μL of 10% APS, and 1.5 μL of
190 TEMED. The protein samples were mixed at the following ratio at room temperature in a 50
191 mM HEPES buffer (pH 7.5) containing 50 mM NaCl and 10 mM $CaCl_2$. After 5 min, the protein
192 samples were loaded to a non-denaturing PAGE gel. For the formation of the four-component

mini-cellulosome, three cellulase components of equimolar were mixed together, and then equimolar mini-CipA was added into the cellulase mixture.

2.7. Enzyme activity assays

Ten nM of BsCel5 and BsCel5', 30 nM of CtCel9 and CtCel9', 60 nM of CpCel48 and CpCel48' were used to determine the specific activity on Avicel and RAC in the presence of mini-CipA, where mini-CipA to cellulase molar ratio was 1:3, and the absence of mini-CipA in 0.5 mL of 50 mM HEPES buffer (pH 7.5) containing 10 mM $CaCl_2$ and 50 mM NaCl at 37^{o}C for 20 min. One unit (U) of enzyme activity was defined as the amount of enzyme producing one μmol of reducing sugar per min. For the formation of the four-component mini-cellulosome, equimolar BsCel5', CtCel9 and CpCel48 were mixed in 50 mM HEPES buffer (pH 7.5) containing 10 mM $CaCl_2$ and 50mM NaCl, and then equimolar mini-CipA was added. Hydrolysis of Avicel and RAC by the mini-cellulosome was investigated at a cellulase loading from 0.005, 0.01, 0.02, 0.04, to 0.08 g/L (not including mini-CipA) in a 50 mM HEPES (pH 7.5) containing 50 mM NaCl, 10 mM $CaCl_2$ and 4 g/L Avicel or RAC at 37 °C. Sixty unit of β-glucosidase (Novozyme 188 purchased from Sigma-Aldrich) per gram of cellulose was supplemented in all hydrolysis experiments for eliminating hydrolysis product inhibition. The hydrolysis of Avicel and RAC by the same amount of β-glucosidase without cellulase was performed as a control. The reactions were stopped by boiling for 5 min. After centrifugation at 10,000 rpm for 1 min, reducing sugar in the supernatants was determined by the modified 2,2'-bicinchoninate method at 75 ^{o}C for 30 min with glucose as a reference [28].

2.8. Other assays

Protein mass concentration was measured by the Bio-Rad Bradford protein dye reagent method (Bio-Rad, Hercules, CA) with bovine serum albumin as a reference. The protein mass based on the Bradford method was calibrated by their absorbance (280 nm) in 6 M guanidine hydrochloride [15]. The adsorption efficiency of cellulases and the mini-cellulosome to cellulosic substrates was determined as described previously [27]. The purity of protein samples was examined by 12% SDS-PAGE. The native PAGE and SDS-PAGE were stained by the Bio-Rad Bio-Safe Colloidal Coomassie Blue G-250. The intensity of the band in the gel was analyzed with Quantity One software (Bio-Rad, Version 4.6.7). The particle sizes of individual protein and the mini-cellulosome in the aqueous solution were directly estimated by the Malvern Zetasizer Nano ZS system (Worcestershire, UK).

3. Results

3.1. Individual cellulase components and mini-CipA

Two of the three cellulases used in this study were the same as those enzymes used in our previous work except CpCel9 [10]. These cellulases were (i) a non-cellulosomal *B. subtilis* family 5 endoglucanase (BsCel5), (ii) a cellulosomal *C. thermocellum* family 9 processive endoglucanase (CtCel9), and (iii) a non-cellulosomal *C. phytofermentans ISDg* family 48 cellobiohydrolase (CpCel48) [10] (Fig. 1). In this study, family 9 cellulase (CtCel9) from *C. thermocellum* was used instead of CpCel9 due to: (i) the expression level of CtCel9 was higher than that of CpCel9, accompanied with a higher protein purification yield, (ii) they had comparable activity at the experimental temperature, and (iii) CtCel9 contained its own type I dockerin. In order to change wild-type Cel5 and Cel48 to cellulosomal cellulase, another two type I dockerin modules from *C. thermocellum* were added to the C-terminal of BsCel5 and

CpCel48, respectively. Here we chose three sequence-different type I dockerins for three cellulases because these three cellulases could be co-expressed in a consolidated bioprocessing *B. subtilis* strain.

Since CBMs usually facilitate catalytic module of cellulase to bind to the substrate, most of cellulase components in non-complexed cellulase systems have their own CBMs. In contrast, most cellulosomal cellulases do not contain a CBM3 possibly because non-hydrolytic scaffoldins contain a CBM3 module [19, 30]. Whether CBM is an elemental component for individual cellulase or cellulosomal cellulase, however, remained unclear because the addition of CBM did not always enhance CBM-free cellulase activity [6]. Therefore, we constructed six different cellulase components with or without CBM (Fig. 1). BsCel5 containing a family 5 cellulase catalytic module, a CBM3b, and a dockerin from *C. thermocellum* CelK, while BsCel5' contained a family 5 cellulase catalytic module and a dockerin domain. So did CpCel48, while its dockerin was from *C. thermocellum* CelS. CtCel9 contained a family 9 cellulase catalytic module, a CBM3c domain and its own dockerin, while CtCel9' contained a family 9 cellulase catalytic module and a dockerin domain. Mini-CipA, a truncated mini-CipA from of the *C. thermocellum* CipA, contained the first three type I cohesin modules and one CBM3b module, which was located between the second and the third cohesin module (Fig. 1). All of the recombinant proteins were produced in *E. coli* BL21(DE3). The cellulase components and mini-CipA were purified to homogeneity, examined by SDS-PAGE (Fig. 2). Their molecular weights were the same as expected (Fig. 1).

3.2. Analysis of the mini-cellulosome assembly

The interaction between dockerin-containing cellulases and the cohesin-containing mini-CipA was examined by RAC affinity pull-down experiment (Fig. 3). The adsorption of six cellulases on RAC was examined first. Nearly all of two CBM3b-containing cellulases (BsCel5 and CpCel48, 0.08 g/L) were adsorbed by 4 g/L RAC while about 75% of CBM3c-containing CtCel9 was adsorbed, suggesting that CBM3b was a stronger substrate-binding module than CBM3c. When CBMs were removed from cellulase components, their adsorption abilities were decreased greatly. For example, ~30% BsCel5', ~80% CpCel48' and ~5% CtCel9' were adsorbed by RAC. When mini-CipA was mixed with any cellulase component with a molar ratio of 1:3 followed by RAC adsorption, there were little detectable proteins in the supernatant, suggesting that CBM-free cellulases can be totally adsorbed by RAC through the help of mini-CipA.

Since the catalytic modules of BsCel5 and CpCel48 still had some weak binding ability to cellulose, while the catalytic module of CtCel9 rarely bound to the cellulose. Thus, CtCel9' was chosen to determine the molar ratio of cohesin and dockerin in the mini-cellulosome. When CtCel9' was alone, nearly all of CtCel9' remained in the supernatant (Fig. 3A). When the molar ratios of mini-CipA and CtCel9' were 1:1 and 1:3, no protein bands were detected in the supernatant, suggesting that both mini-CipA and CtCel9' were totally bound to RAC through the CBM3b module in mini-CipA. When the ratio was 1:6, a half of CtCel9' was bound by RAC and the other half of CtCel9' was remained in the supernatant. These results suggested that the binding ratio of mini-CipA and dockerin-containing cellulase was approximately 1:3, in good agreement with previous reports pertaining to a molar ratio of cohesin to dockerin of 1:1 [19, 30].

The mini-cellulosome made from a mixture of equimolar BsCel5', CtCel9', CpCel48' and mini-CipA were mixed with RAC. After the pellets were washed, the mini-cellulosome bound on the surface of RAC was examined by SDS-PAGE. There were four protein bands (Fig.

3B), representing four different components. The band densities of four bands suggested that the molar ratio of BsCel5':CtCel9': CpCel48': mini-CipA was approximately 1:1:1:1, as expected. Another mini-cellulosome was prepared by mixing equimolar BsCel5', CtCel9, CpCel48 and mini-CipA. After RAC pull-down, this mini-cellulosome exhibited three bands in SDS-PAGE (Fig. 3B, Lane 5) rather than four bands because molecular masses of CtCel9 and mini-CipA are very close so that these two bands overlapped together in SDS-PAGE. The above experiments suggested that three-cohesion-containing mini-CipA can bind a mixture containing equal molar three dockerin-containing cellulases at a ratio of 1:1:1:1.

The *in vitro* assembly of two-component and four-component mini-cellulosomes was examined in native gel (Fig. 4). Lanes 1, 3, 5 and 7 presented a single band for BsCel5', CtCel9, CpCel48, and mini-CipA, respectively. When mini-CipA and individual cellulases (i.e., BsCel5', CtCel9, and CpCel48) at a molar ratio of 1:3 were mixed, a single band was observed in Lanes 2, 4, and 6, respectively, suggested that the dockerin-containing individual cellulases formed static enzyme complexes with the cohesin-containing mini-CipA. When equimolar BsCel5', CtCel9, CpCel48 and mini-CipA were mixed *in vitro*, a large band, along with some nearby minor bands, was observed (Fig. 4, Lane 8), suggesting the formation of the four-component mini-cellulosome. There were no small-size bands representing free dockerin-containing cellulases (Fig. 4, Lane 8), suggesting that all of the dockerin-containing cellulases were bound with mini-CipA, forming the mini-cellulosome. Since each dockerin-containing cellulase can randomly bind with any one of cohesins of mini-CipA, the resulting mini-cellulosome was not homogeneous, while the whole group had equimolar four components, similar to natural cellulosomes and designer mini-cellulosomes [11-15].

The particle sizes of the mini-cellulosome and individual components were examined by the Malvern Zetasizer Nano ZS system (Fig. 5). The diameters of BsCel5', CtCel9, CpCel48, and mini-CipA were approximately 5.6 nm, 9.3 nm, 10.64 nm, and 8.3 nm, respectively. In contrast, the diameter of the mini-cellulosome made from a mixture containing equimolar BsCel5', CtCel9, CpCel48 and mini-CipA, was larger, being approximately 16.5 nm. Also, no small size peaks were observed for this mixture, suggesting that all the dockerin-containing cellulases were integrated into the mini-CipA. All of the above evidences suggested that mixing equimolar of three dockerin-containing cellulases and a mini-CipA containing three dockerin modules resulted in the mini-cellulosome with an aggregated molar ratio of 1:1:1:1.

3.3. Individual cellulase activities in the absence and presence of mini-CipA

The specific activities of the six cellulases were examined on two substrates -- Avicel and RAC (Table 2). All the cellulases showed much higher activities on RAC than on Avicel. For example, the activity of BsCel5' on RAC was 1.29 ± 0.09 μmol/mg/min, 8.6 times the activity on Avicel (0.15 ± 0.02 μmol/mg/min). However, the specific activity of BsCel5 was lower than that of BsCel5', suggesting this CBM3b may not be important for BsCel5. The activities of CtCel9 were about 3-4 fold higher than those of CtCel9' on both substrates, suggesting that the CBM3c was important for maintaining the activity of this family 9 cellulase, as reported previously [32]. The specific activities of CpCel48 were somewhat higher than those of CpCel48' on both substrates.

When mini-CipA was mixed with individual dockerin-containing cellulase at a molar ratio of 1:3, all cellulases in the presence of mini-CipA exhibited higher activity than themselves alone (Table 2). Mini-CipA did not show any activity against Avicel and RAC (Table 2). The

329 stimulation factor (SF) was defined as the ratio of soluble sugar released by the mini-cellulosome
330 to soluble sugar released by the non-complexed cellulase mixture at the same cellulase amount.
331 All the CBM-free cellulases showed higher SF values than their CBM-containing counterparts.
332 Also, SF values obtained on Avicel were higher than those on RAC (Table 2). It was noted that
333 SF values ranged greatly. For example, SF values of CtCel9 and CtCel9' on both substrates were
334 more than two, much higher than those of BsCel5 and CpCel48, possibly because CtCel9 was a
335 native cellulase component from *C. thermocellum* cellulosome but BsCel5 and CpCel48 were
336 not. In support to this, the similar observation has been reported elsewhere [15].
337 Because the activities of BsCel5', CtCel9 and CpCel48 in the presence of mini-CipA
338 were higher than those of their counterparts, they were chosen for the formation of the four-
339 component mini-cellulosome in the following hydrolysis experiments.
340
341 *3.4. Hydrolysis of Avicel and RAC by the four-component mini-cellulosome*
342 Hydrolysis performances of Avicel and RAC by the four-component mini-cellulosome
343 containing equimolar BsCel5', CtCel9, and CpCel48 were investigated at different cellulase
344 loadings of 0.005 and 0.08 g/L (not including mini-CipA) (Fig. 6). The cellulose hydrolysis
345 experiments at the cellulase loading amount of 0.01, 0.02, and 0.04 g/L were also carried out
346 (data not shown). On Avicel, when cellulase concentration was 0.08 g/L, approximately 1000
347 μM glucose was released by the mini-cellulosome after 24 h, i.e., 5% of the substrate was
348 hydrolyzed (Fig. 6A). In contrast, approximately 450 μM glucose was released by the same
349 amount of the non-complexed cellulase mixture after 24 h (Fig. 6A). At this enzyme
350 concentration, the SF value of the mini-cellulosome was about 2.12-fold on Avicel. When the
351 cellulase concentration was decreased to 0.005 g/L, the mini-cellulosome released *ca.* 250 μM

soluble sugar after 24 h on Avicel and exhibited approximately three times of activity higher than the non-complexed cellulase mixture (Fig. 6B). On RAC, individual cellulases, their mixture, and the mini-cellulosome always showed much faster hydrolysis rates than those on Avicel. When cellulase concentration was 0.08 g/L, approximately 6500 μM and 5000 μM glucose were released by the mini-cellulosome and the non-complexed cellulase mixture after 24 h, respectively (Fig. 6C) and the SF value of tri-functional mini-cellulosome was *ca.* 1.2 on RAC, much lower than that on Avicel (SF = 2.12). When cellulase concentration was decreased to 0.005 g/L, the mini-cellulosome released approximately 1300 μM glucose after 24 h on RAC and exhibited approximately 2.3-fold of activity of the non-complexed cellulase mixture (Fig. 6D). This SF value was also lower than on Avicel (i.e., ~ 3.0) at the same enzyme and substrate loadings.

The effects of substrate/enzyme ([S]/[E]) mass ratios from 50 to 800 on stimulation factor are shown in Fig. 7. On both substrates, the SF values remained nearly constant at high [S]/[E] ratios and decreased as [S]/[E] decreased. Clearly, the mini-cellulosome exhibited higher SF on low accessibility recalcitrant Avicel than on RAC. This result suggested that the construction of the mini-cellulosomes would be efficient for enhancing cellulose conversion, especially in the case of low-accessibility recalcitrant cellulose at low enzyme usage.

Furthermore, the adsorbed efficiencies of the non-complexed cellulase mixture and the mini-cellulosome were examined on Avicel and RAC (Fig. 8). Regardless of substrate type, more cellulosome were adsorbed than the non-complexed cellulase mixture mainly due to scaffoldin's CBM. For low-accessibility Avicel, adsorbed cellulase efficiency greatly increased with an increase in [S]/[E] (Fig. 8A). When [S]/[E] was 50 (i.e., a typical industrial cellulase loading), only 18% or 26% of cellulases and cellulosomes were adsorbed by Avicel, suggesting

that Avicel had low substrate accessibility. By contrast, most of the non-complexed cellulases and the cellulosome were adsorbed by high-accessibility RAC regardless of substrate concentration (Fig. 8B).

4. Discussion

The four-component mini-cellulosome containing the *B. subtilis* family 5 endoglucanase, *C. thermocellum* family 9 processive endoglucanase, and *C. phytofermentans* ISDg family 48 cellobiohydrolase were assembled *in vitro* and its hydrolysis performance was investigated on Avicel and RAC. The mini-cellulosome exhibited faster hydrolysis rate than the non-complexed cellulase mixture on Avicel and RAC (Fig. 6).The four-component mini-cellulosomes had much higher SF values on Avicel than on RAC, especially when enzyme concentrations were low (Fig. 7). This finding was in partial agreement with previous works [15, 33]. Furthermore, we presented in-depth understanding of different hydrolysis mechanisms of mini-cellulosome and non-complexed enzyme mixture on Avicel and RAC in terms of [S]/[E] ratio.

Avicel is a heterogeneous substrate; its glucan chains are aligned in the same direction; and highly ordered hydrogen bonds among adjacent sugar chains result in low surface accessibility to cellulase [4, 27]. Avicel hydrolysis by the cellulase mixture was a peeling or layer-by-layer hydrolysis process. The ends of β-glucosidic bond on the surface of Avicel generated by adsorbed endoglucanase cannot by hydrolyzed by exoglucanase until endoglucanase moved to elsewhere and exoglucanase moved to the right site [28]. For low-accessibility Avicel, the ends of cellulose chains were limited to exoglucanase [34-36]. In contrary to Avicel, RAC is a homogeneous amorphous cellulose, whose highly ordered hydrogen bonds in the cellulose chains are disrupted [25], its surface area is at least 20 times higher than

that of Avicel based on the adsorption of a cellulase-size molecule [10]. High-accessibility RAC allowed endoglucanase to efficiently and rapidly hydrolyze β-glucan bonds of all cellulose chains, resulting in a rapid reduction in DP within a short time [28]. As a result, the reducing and non-reducing ends of RAC were fast more than exoglucanase. Therefore, each cellulase component in the non-complexed cellulase mixture worked independently so that the synergy between exoglucanase and endoglucanase were not vital to complete hydrolysis of RAC [10].

CBM3b in mini-CipA enabled to enrich more cellulases to be adsorbed on the surface of cellulose, called substrate targeting effect, resulting in enhanced hydrolysis rate (Table 1). On Avicel, the new ends on the surface of Avicel generated by endoglucanase can be easily accessed by the proximate exoglucanases linked by a scaffoldin, called enzyme proximity effect [15] or substrate channeling [37], resulting in higher hydrolysis rate than the non-complexed cellulase mixture. In contrast, all new ends of RAC generated by endoglucanase were not limited, which can be hydrolyzed by exoglucanase, resulting in weaker enzyme proximate effects on RAC than on Avicel. Therefore, less SF effect was observed on RAC than on Avicel (Fig. 7).

Additionally, SF increased when [S]/[E] increased for both substrates (Fig. 7). This trend may be attributed to a dynamic result from two causes. The first cause was substrate inhibition effect for non-complexes cellulase [38]. When [S]/[E] ratio increased, the distance between adsorbed individual cellulase components in the non-complexed cellulase mixture increased, resulting in lower apparent activity (i.e., higher SF); while the distance between the scaffoldin-linked cellulases was nearly constant independent of substrate concentration. The second cause was that the ratio of the adsorbed mini-cellulosome to adsorbed non-complexed cellulase mixture decreased with an increase in [S]/[E] (Fig.8). It meant that more cellulosome was adsorbed on the substrate, exhibiting higher apparent activity. It was worth pointing out that in cases of high-

accessibility RAC at regardless of [S/E] ratio (Fig. 8B) and low-accessibility Avicel at high [S]/[E] ratios (Fig. 8A), the effect of the second cause could be minimal because most of cellulosome and cellulases were adsorbed.

In addition to substrate accessibility and [S]/[E] ratios investigated here, numerous factors could influence enzymatic hydrolysis performance and SF, such as different enzyme components used [39, 40], CBM location [6], CBM type [41], substrate properties (e.g., presence of hemicellulose and lignin) [40]. It was noted that not all the enzymes from different families would result in synergies of mini-cellulosomes (Fierobe, personal communication).

Most pretreatments, such as dilute acid, steam explosion, hot water, ammonia-based technology, generate pretreated biomass with the substrate accessibility larger than Avicel and less than RAC [4]. Their seemingly two-phase enzymatic hydrolysis profiles implied that the pretreated biomass may have at least two fractions: a fraction of amorphous cellulose with high accessibility and a fraction of recalcitrance cellulose with low accessibility [42]. Therefore, the construction of cellulosomes would be efficient to increase cellulose digestibility of the pretreated biomass that contained low-accessibility cellulose core to more than 90% at a low enzyme usage.

In a word, the mini-cellulosome and the non-complexed cellulase mixture had different hydrolysis mechanisms on low-accessibility recalcitrant cellulose and high-accessibility amorphous cellulose. The stimulation factor from scaffoldin was stronger on recalcitrant Avicel than on amorphous cellulose, suggesting that the construction of cellulosome may be an effective way to increase mass-specific activity of cellulase, especially for low substrate accessibility recalcitrant cellulosic substrate at an increased substrate/enzyme ratio.

Acknowledgement

This work was supported mainly by the DOE BioEnergy Science Center (BESC). BESC 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 College of Agriculture and Life Sciences Bioprocessing and Biodesign Research Center at Virginia Tech.

References

[1] Y.-H.P. Zhang, M.E. Himmel, J.R. Mielenz, Outlook for cellulase improvement: screening and selection strategies, Biotechnol. Adv., 24 (2006) 452-481.

[2] L.R. Lynd, P.J. Weimer, W.H. van Zyl, I.S. Pretorius, Microbial cellulose utilization: fundamentals and biotechnology, Microbiol. Mol. Biol. Rev., 66 (2002) 506-577.

[3] L.R. Lynd, M.S. Laser, D. Bransby, B.E. Dale, B. Davison, R. Hamilton, M. Himmel, M. Keller, J.D. McMillan, J. Sheehan, C.E. Wyman, How biotech can transform biofuels, Nat. Biotechnol., 26 (2008) 169-172.

[4] J.A. Rollin, Z. Zhu, N. Sathisuksanoh, Y.-H.P. Zhang, Increasing cellulose accessibility is more important than removing lignin: A comparison of cellulose solvent-based lignocellulose fractionation and soaking in aqueous ammonia, Biotechnol. Bioeng., 108 (2011) 22-30.

[5] Y.-H.P. Zhang, Production of biocommodities and bioelectricity by cell-free synthetic enzymatic pathway biotransformations: Challenges and opportunities, Biotechnol. Bioeng., 105 (2010) 663-677.

[6] W. Liu, X.-Z. Zhang, Z.-M. Zhang, Y.-H.P. Zhang, Engineering of *Clostridium phytofermentans* endoglucanase Cel5A for improved thermostability Appl. Environ. Microbiol., 76 (2010) 4914-4917.

[7] X.-Z. Zhang, Y.-H.P. Zhang, Simple, fast and high-efficiency transformation system for directed evolution of cellulase in *Bacillus subtilis*, Microb. Biotechnol., 4 (2011) 98-105.

[8] Y. Li, D.C. Irwin, D.B. Wilson, 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 (2010) 2582-2588.

[9] L. Rosgaard, S. Pedersen, J. Langston, D. Akerhielm, J.R. Cherry, A.S. Meyer, Evaluation of minimal *Trichoderma reesei* cellulase mixtures on differently pretreated barley straw substrates, Biotechnol. Prog., 23 (2007) 1270-1276.

[10] H.H. Liao, X.Z. Zhang, J.A. Rollin, Y.-H.P. Zhang, A minimal set of bacterial cellulases for consolidated bioprocessing of lignocellulose, Biotechnol. J., 6 (2011) 1409-1418.

[11] S. Morais, Y. Barak, J. Caspi, Y. Hadar, R. Lamed, Y. Shoham, D.B. Wilson, E.A. Bayer, Contribution of a xylan-binding module to the degradation of a complex cellulosic substrate by designer cellulosomes, Appl. Environ. Microbiol., 76 (2010) 3787-3796.

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

[13] J. Caspi, D. Irwin, R. Lamed, Y. Li, H.P. Fierobe, D.B. Wilson, E.A. Bayer, Conversion of *Thermobifida fusca* free exoglucanases into cellulosomal components: comparative impact on cellulose-degrading activity, J. Biotechnol., 135 (2008) 351-357.

[14] F. Mingardon, A. Chanal, A.M. Lopez-Contreras, C. Dray, E.A. Bayer, H.P. Fierobe, Incorporation of fungal cellulases in bacterial minicellulosomes yields viable, synergistically acting cellulolytic complexes, Appl. Environ. Microbiol., 73 (2007) 3822-3832.

[15] H.P. Fierobe, E.A. Bayer, C. Tardif, M. Czjzek, A. Mechaly, A. Belaich, R. Lamed, Y. Shoham, J.P. Belaich, Degradation of cellulose substrates by cellulosome chimeras. Substrate targeting versus proximity of enzyme components, J. Biol. Chem., 277 (2002) 49621-49630.

[16] S.L. Tsai, J. Oh, S. Singh, R. Chen, W. Chen, Functional assembly of minicellulosomes on the *Saccharomyces cerevisiae* cell surface for cellulose hydrolysis and ethanol production, Appl. Environ. Microbiol., 75 (2009) 6087-6093.

[17] S.L. Tsai, G. Goyal, W. Chen, Surface Display of a Functional Minicellulosome by Intracellular Complementation Using a Synthetic Yeast Consortium and Its Application to Cellulose Hydrolysis and Ethanol Production, Appl. Environ. Microbiol., 76 (2010) 7514-7520.
[18] C. You, X.Z. Zhang, N. Sathitsuksanoh, L.R. Lynd, Y.-H.P. Zhang, Enhanced microbial cellulose utilization of recalcitrant cellulose by an ex vivo cellulosome-microbe complex, Appl. Environ. Microbiol., (2011).
[19] E.A. Bayer, J.P. Belaich, Y. Shoham, R. Lamed, The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides, Annu. Rev. Microbiol., 58 (2004) 521-554.
[20] P.J.M. Steenbakkers, X.-L. Li, E.A. Ximenes, J.G. Arts, H. Chen, L.G. Ljungdahl, H.J.M. Op den Camp, Noncatalytic Docking Domains of Cellulosomes of Anaerobic Fungi, J. Bacteriol., 183 (2001) 5325-5333.
[21] A.C.J. Freelove, D.N. Bolam, P. White, G.P. Hazlewood, H.J. Gilbert, A Novel Carbohydrate-binding Protein Is a Component of the Plant Cell Wall-degrading Complex of Piromyces equi, J. Biol. Chem., 276 (2001) 43010-43017.
[22] B.R.S. Ali, L. Zhou, F.M. Graves, R.B. Freedman, G.W. Black, H.J. Gilbert, G.P. Hazlewood, Cellulases and hemicellulases of the anaerobic fungus Piromyces constitute a multiprotein cellulose-binding complex and are encoded by multigene families, FEMS Microbiol. Let., 125 (1995) 15-22.
[23] E.A. Bayer, E. Morag, R. Lamed, The cellulosome--a treasure-trove for biotechnology, Trends Biotechnol., 12 (1994) 379-386.
[24] R. Kumar, C.E. Wyman, Access of cellulase to cellulose and lignin for poplar solids produced by leading pretreatment technologies, Biotechnol. Prog., 25 (2009) 807-819.
[25] N. Sathitsuksanoh, Z.G. Zhu, S. Wi, Y.-H.P. Zhang, Cellulose solvent-based biomass pretreatment breaks highly ordered hydrogen bonds in cellulose fibers of switchgrass, Biotechnol. Bioeng., 108 (2011) 521–529.
[26] Y.-H.P. Zhang, J. Cui, L.R. Lynd, L.R. Kuang, A transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure, Biomacromolecules, 7 (2006) 644-648.
[27] J. Hong, X. Ye, Y.-H.P. Zhang, Quantitative determination of cellulose accessibility to cellulase based on adsorption of a nonhydrolytic fusion protein containing CBM and GFP with its applications, Langmuir, 23 (2007) 12535-12540.
[28] Y.-H.P. Zhang, L.R. Lynd, Determination of the number-average degree of polymerization of cellodextrins and cellulose with application to enzymatic hydrolysis, Biomacromolecules, 6 (2005) 1510-1515.
[29] X.Z. Zhang, Z. Zhang, Z. Zhu, N. Sathitsuksanoh, Y. Yang, Y.-H.P. Zhang, The noncellulosomal family 48 cellobiohydrolase from *Clostridium phytofermentans* ISDg: heterologous expression, characterization, and processivity, Appl. Microbiol. Biotechnol., 86 (2010) 525-533.
[30] R.H. Doi, A. Kosugi, Cellulosomes: plant-cell-wall-degrading enzyme complexes, Nat. Rev. Microbiol., 2 (2004) 541-551.
[31] U.T. Gerngross, M.P. Romaniec, T. Kobayashi, N.S. Huskisson, A.L. Demain, Sequencing of a *Clostridium thermocellum* gene (cipA) encoding the cellulosomal SL-protein reveals an unusual degree of internal homology, Mol. Microbiol., 8 (1993) 325-334.
[32] T. Arai, A. Kosugi, H. Chan, R. Koukiekolo, H. Yukawa, M. Inui, R.H. Doi, Properties of cellulosomal family 9 cellulases from *Clostridium cellulovorans*, Appl. Microbiol. Biotechnol., 71 (2006) 654-660.
[33] H.P. Fierobe, F. Mingardon, A. Mechaly, A. Belaich, M.T. Rincon, S. Pages, R. Lamed, C. Tardif, J.P. Belaich, E.A. Bayer, Action of designer cellulosomes on homogeneous versus complex substrates: controlled incorporation of three distinct enzymes into a defined trifunctional scaffoldin, J. Biol. Chem., 280 (2005) 16325-16334.

543 [34] M. Kurašin, P. Väljamäe, Processivity of cellobiohydrolases is limited by the substrate, J. Biol. Chem.,
544 286 (2011) 169-177.
545 [35] Y.-H.P. Zhang, L.R. Lynd, A functionally-based model for hydrolysis of cellulose by fungal cellulase,
546 Biotechnol. Bioeng., 94 (2006) 888-898.
547 [36] S.E. Levine, J.M. Fox, H.W. Blanch, D.S. Clark, A mechanistic model of the enzymatic hydrolysis of
548 cellulose, Biotechnol. Bioeng., 107 (2010) 37-51.
549 [37] Y.-H.P. Zhang, Substrate channeling and enzyme complexes for biotechnological applications,
550 Biotechnol. Adv., 29 (2011) 715-725.
551 [38] X. Huang, M.H. Penner, Apparent substrate inhibition of the *Trichoderma reesei* cellulase system., J.
552 Agric. Food. Chem., 39 (1991) 2096-3000.
553 [39] S. Moraïs, A. Heyman, Y. Barak, J. Caspi, D.B. Wilson, R. Lamed, O. Shoseyov, E.A. Bayer, Enhanced
554 cellulose degradation by nano-complexed enzymes: Synergism between a scaffold-linked
555 exoglucanase and a free endoglucanase, J. Biotechnol., 147 (2010) 205-211.
556 [40] S. Moraïs, Y. Barak, J. Caspi, Y. Hadar, R. Lamed, Y. Shoham, D.B. Wilson, E.A. Bayer, Cellulase-
557 xylanase synergy in designer cellulosomes for enhanced degradation of a complex cellulosic
558 substrate, mBio, 1 (2010) e00285-00210.
559 [41] X. Ye, Z. Zhu, C. Zhang, Y.-H.P. Zhang, Fusion of a family 9 cellulose-binding module improves
560 catalytic potential of *Clostridium thermocellum* cellodextrin phosphorylase on insoluble cellulose,
561 Appl. Microbiol. Biotechnol., 92 (2011) 551-560.
562 [42] S. Wald, C.R. Wilke, H.W. Blanch, Kinetics of the enzymatic hydrolysis of cellulose, Biotechnol.
563 Bioeng., 26 (1984) 221-230.

564

565

565 **Figure legends**

566 **Fig. 1.** Schematic representation of the recombinant proteins used in this study.

567

568 **Fig. 2.** SDS-PAGE of the purified recombinant proteins produced in *E. coli* for in vitro assembly
569 of the mini-cellulosomes. M, protein marker.

570

571 **Fig. 3.** SDS-PAGE of RAC affinity pull-down for the determination the composition of the
572 mini-cellulosome containing mini-CipA and CtCel9' at different molar ratios (A): P, RAC-
573 containing pellets; S, supernatant; and M, protein marker; for the determination the composition
574 of the mini-cellulosome containing BsCel5', CtCel9', CpCel48' and mini-CipA (B); for the
575 determination the composition of the mini-cellulosome containing BsCel5', CtCel9, CpCel48
576 and mini-CipA (C).

577

578 **Fig. 4.** (A) Native gel of individual cellulosomal components and two-component and four-
579 component cellulosomes. The protein loadings in native gel were 1.77 μg BsCel5' (Lane 1), 1.77
580 μg BsCel5' + 1 μg mini-CipA (Lane 2), 3.16 μg CtCel9 (Lane 3), 3.16 μg CtCel9 + 1 μg mini-
581 CipA (Lane 4), 4.23 μg CpCel48 (Lane 5), 4.23 μg CpCel48 + 1 μg mini-CipA (Lane 6) ,1 μg
582 mini-CipA (Lane 7), and 0.59 μg BsCel5'/1.05 μg CtCel9/ 1.41 μg CpCel48/1 μg mini-CipA
583 (Lane 8).

584

585 **Fig.5.** The particle sizes of the four-component mini-cellulosome and individual protein
586 components determined by the Malvern Zetasizer Nano ZS system.

587

588 **Fig. 6.** Enzyme hydrolysis profiles on Avicel and RAC by the non-complexed enzyme mixture --
589 BsCel5', CtCel9 and CpCel48 and the mini-cellulosome containing equimolar BsCel5', CtCel9,
590 CpCel48 and mini-CipA at the same enzyme amount. Cellulase loadings were 0.08 (A&C) and
591 0.005 g/L (B&D); substrates were Avicel (A&B) and RAC (C&D).
592
593 **Fig. 7.** The relationship of stimulation factors in terms of the ratio of substrate to enzyme on
594 Avicel and RAC after 24 h.
595
596 **Fig. 8.** The adsorption efficiency of the non-complexed enzyme mixture containing BsCel5',
597 CtCel9 and CpCel48 and the mini-cellulosomes containing BsCel5', CtCel9, CpCel48, and mini-
598 CipA by Avicel (A) and RAC (B) in terms of the ratio of substrate to enzyme.
599

599 **Table 1** The primers used to amplify gene fragments.

Gene	Template	Primer name	Sequence	RE*
Mini-CipA	Genome of *Clostridium thermocellum*	Mini-CipA_For	GTAGTA**CATATG**GTATCGGCGGCCACAATGACAG	NdeI
		Mini-CipA_Rev	GCAGTA**CTCGAG**ATTCGAATCATCTGTCGGTGTTG	XhoI
CpCel48 and CpCel48'	Genome of *Clostridium phytofermentans*	Cpcel48_For	CCTCTG**CATATG**GGTGAAACTGAGCAAGC	NdeI
		Cpcel48_Rev	GTAGAGGACCCACCTCCTCCAGATCCTGGTTCGATACCCCAATTAAGTTTTCC	
		CpCel48'_Rev	GTAGAGGACCCACCTCCTCCAGATCCAACCTTAACATCTCCTACTACCACAAC	
	Genome of *Clostridium thermocellum*	DocS_For	GGATCTGGAGGAGGTGGGTCCTCTACTAAATTATACGGCGACGTC	
		DocS_Rev	GCATTA**CTCGAG**GTTCTTGTACGGCAATGTATC	XhoI
CtCel9 and CtCel9'	Genome of *Clostridium thermocellum*	CtCelF_For	GCTTCA**CATATG**GCGGATTTCAACTATGGTGAGGCAC	NdeI
		CtCelF_Rev	GGACCAT **CTCGAG** CTGTTCAGCCGGGAATTTTTCAATAAG	XhoI
		DocF_For	CACCGGGGGAAGAATTTGGAGATGTGAATTTTGACGGAAG	
		CtCelF'_Rev	CACATCTCCAAATTCTTCCCCCGGTGTTTCAAAAG	
BsCel5 and BsCel5'	Genome of *Bacillus subtilis*	BsCel5_For	CCTCAG**CATATG**GCAGGGACAAAAACGCC	NdeI
		BsCel5_Rev	CTCCGGTTCTTCTGGGTCTACTCCTCCATTTGGTTCTGTTCCCCAAATCAGTTTTC	
		BsCel5'_Rev	CTCCGGTTCTTCTGGGTCTACTCCTCCAGAAATACCATTTTCCTGTGTGGGTTTATC	
	Genome of *Clostridium thermocellum*	DocK_For	GGAGGAGTAGACCCAGAAGAACCGGAGGTTATTTATG	
		DocK_Rev	GCCGCC**CTCGAG**TTTATGTGGCAATACATCTATC	XhoI

600 *RE, restriction enzyme site. Restriction enzyme sites included in primer sequences for cloning purposes

601 are indicated in bold, and the overlapping sequences are indicated by underlined text

602

602 **Table 2.** The specific activity of cellulases on Avicel and RAC in the presence and absence of mini-CipA.

	Specific activity (μmol/mg/min)		Specific activity in the presence of mini-CipA (μmol/mg/min)		Stimulation factor (SF)	
	Avicel	RAC	Avicel	RAC	Avicel	RAC
BsCel5	0.12 ± 0.01	0.94 ± 0.04	0.19 ± 0.02	1.06 ± 0.05	1.50	1.13
BsCel5'	0.15 ± 0.02	1.29 ± 0.09	0.25 ± 0.08	1.54 ± 0.12	1.70	1.19
CtCel9	0.032 ± 0.002	0.26 ± 0.01	0.094 ± 0.008	0.65 ± 0.03	2.93	2.48
CtCel9'	0.010 ± 0.000	0.059 ± 0.003	0.045 ± 0.003	0.19 ± 0.01	4.48	3.26
CpCel48	0.021 ± 0.001	0.047 ± 0.002	0.025 ± 0.003	0.052 ± 0.001	1.19	1.10
CpCel48'	0.015 ± 0.003	0.035 ± 0.001	0.019 ± 0.003	0.042 ± 0.002	1.27	1.20
Mini-CipA	0 ± 0	0 ± 0				

603

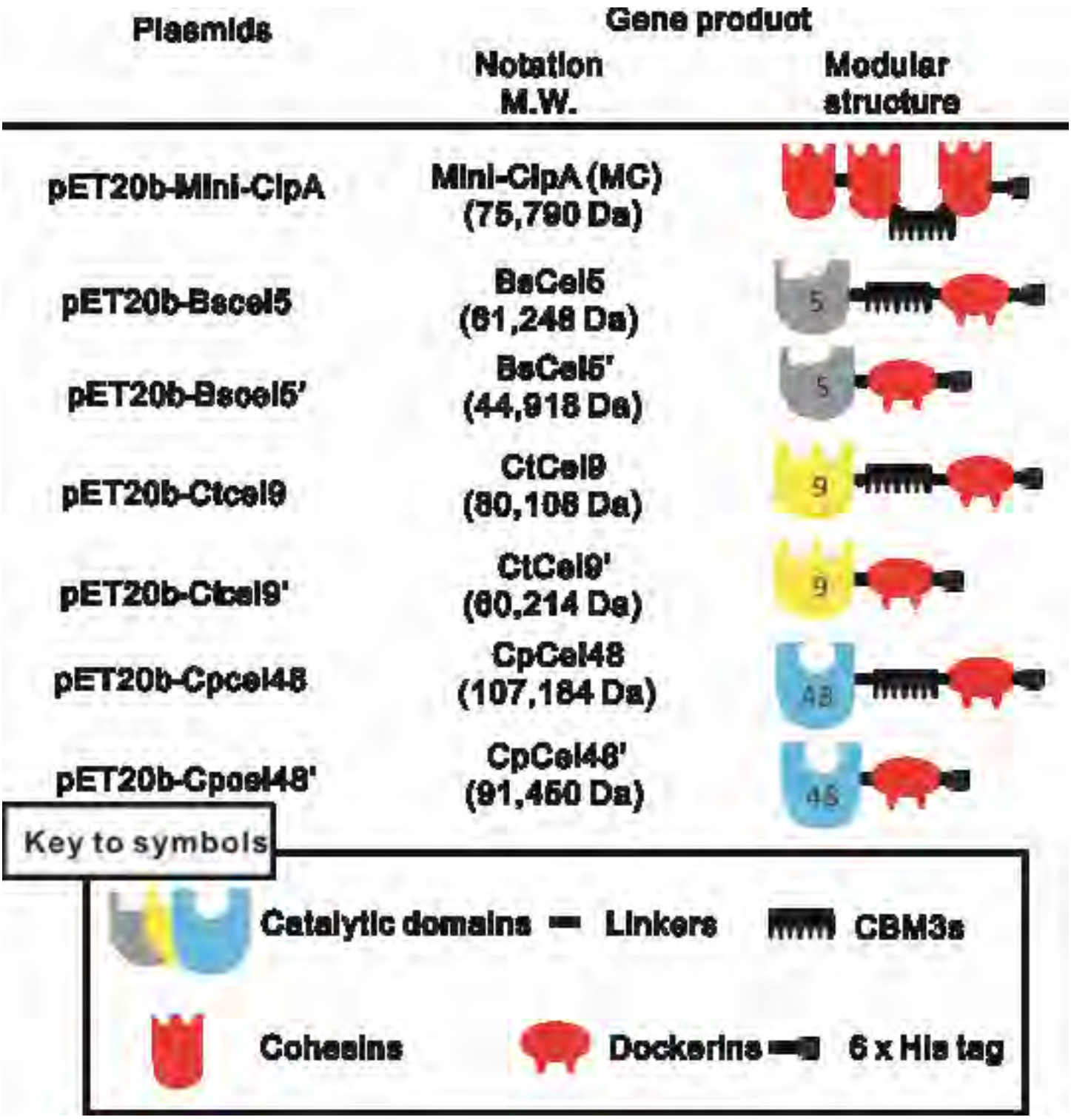
604

Key highlight

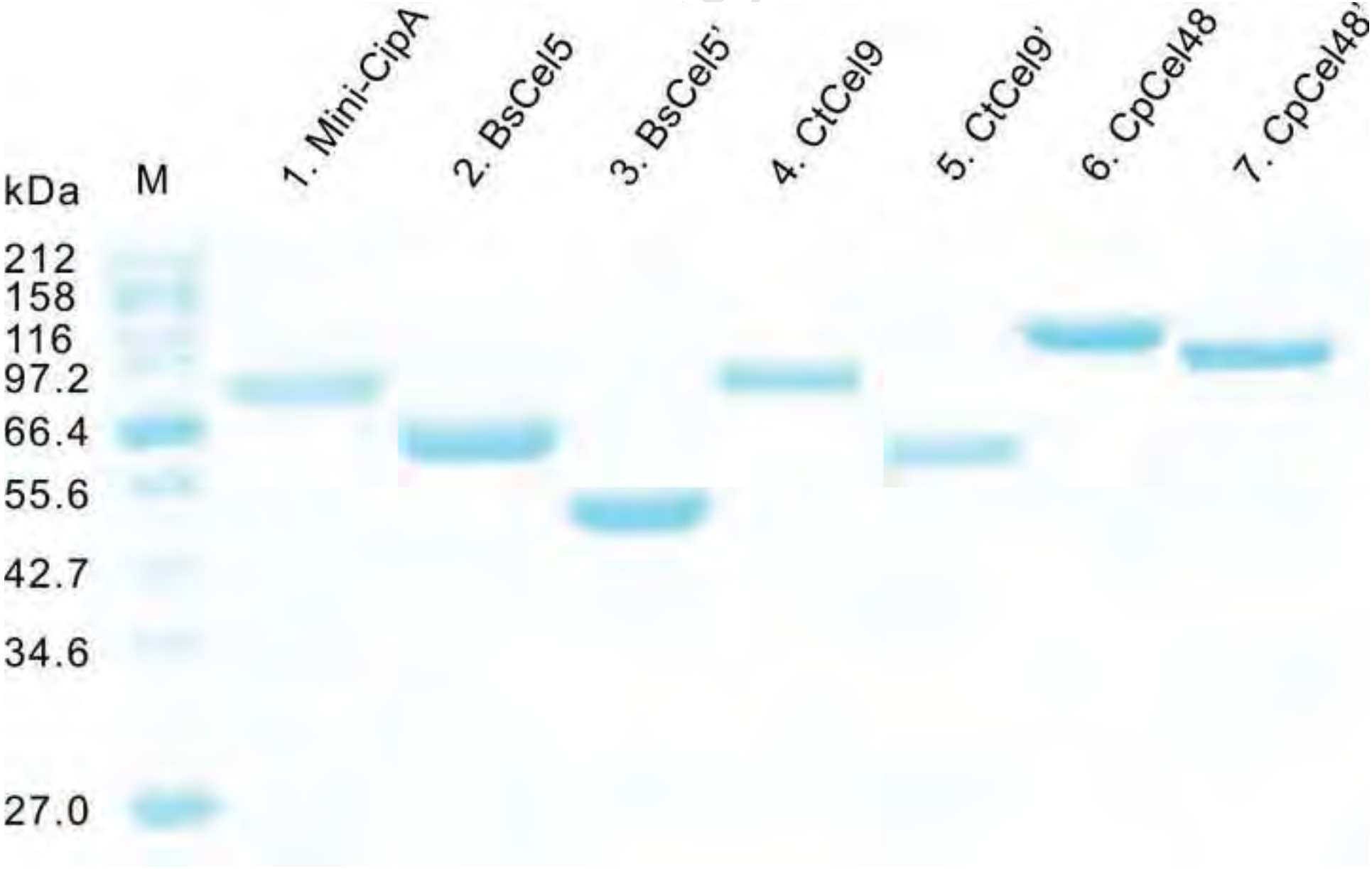
1. The synergistic effect of trifunctional cellulosome compared to non-complex cellulase mixture was more significant on Avicel than on RAC.

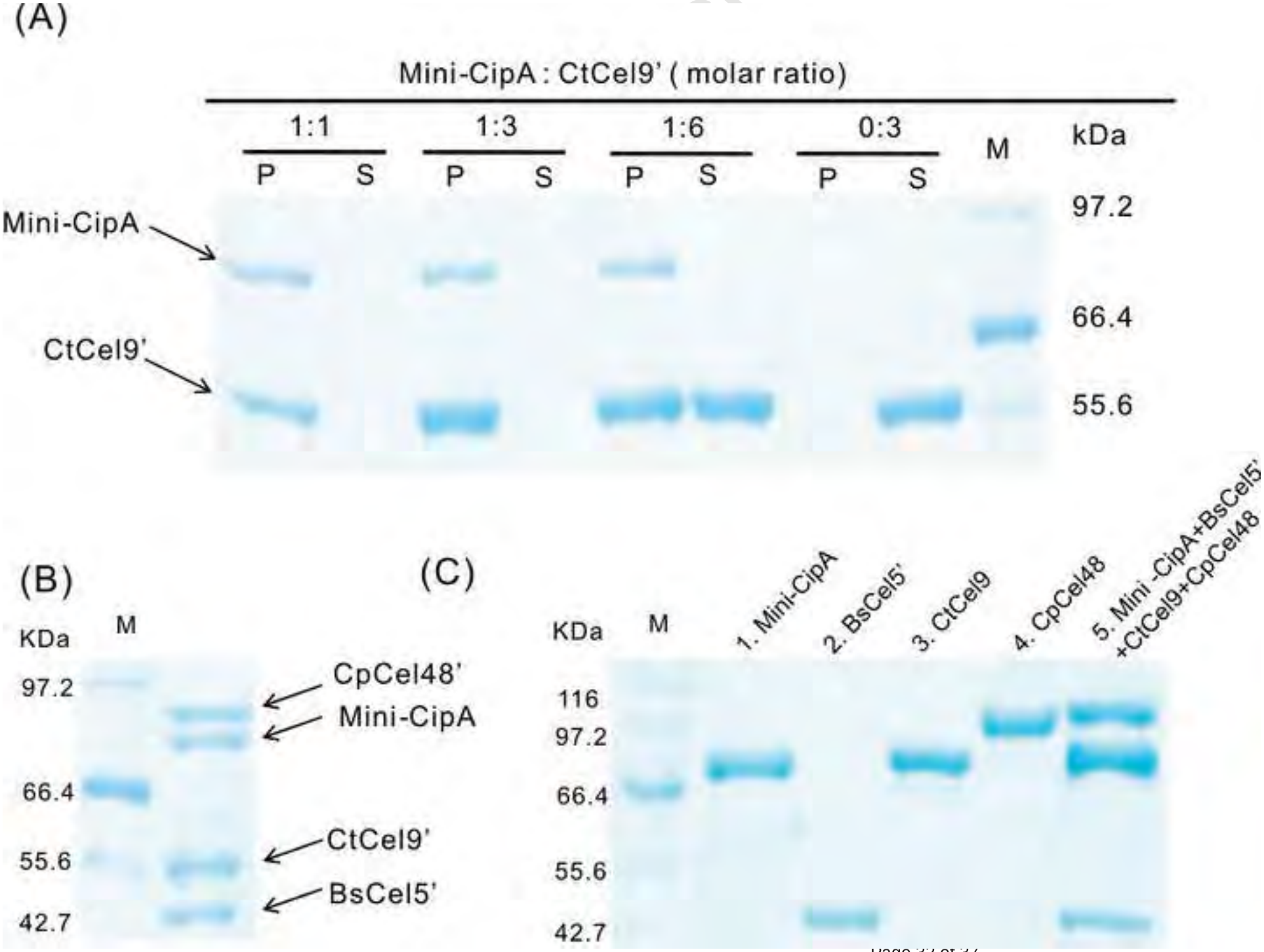
2. The synergistic effect of trifunctional cellulosome compared to non-complex cellulase mixture increased when substrate/enzyme ratio increased, i.e., enzyme usage decreased.

3. Construction of cellulosome should be an efficient way to increase cellulose conversion rate, especially in the cases of hydrolyzing recalcitrant low-accessibility crystalline cellulose and low enzyme usage.



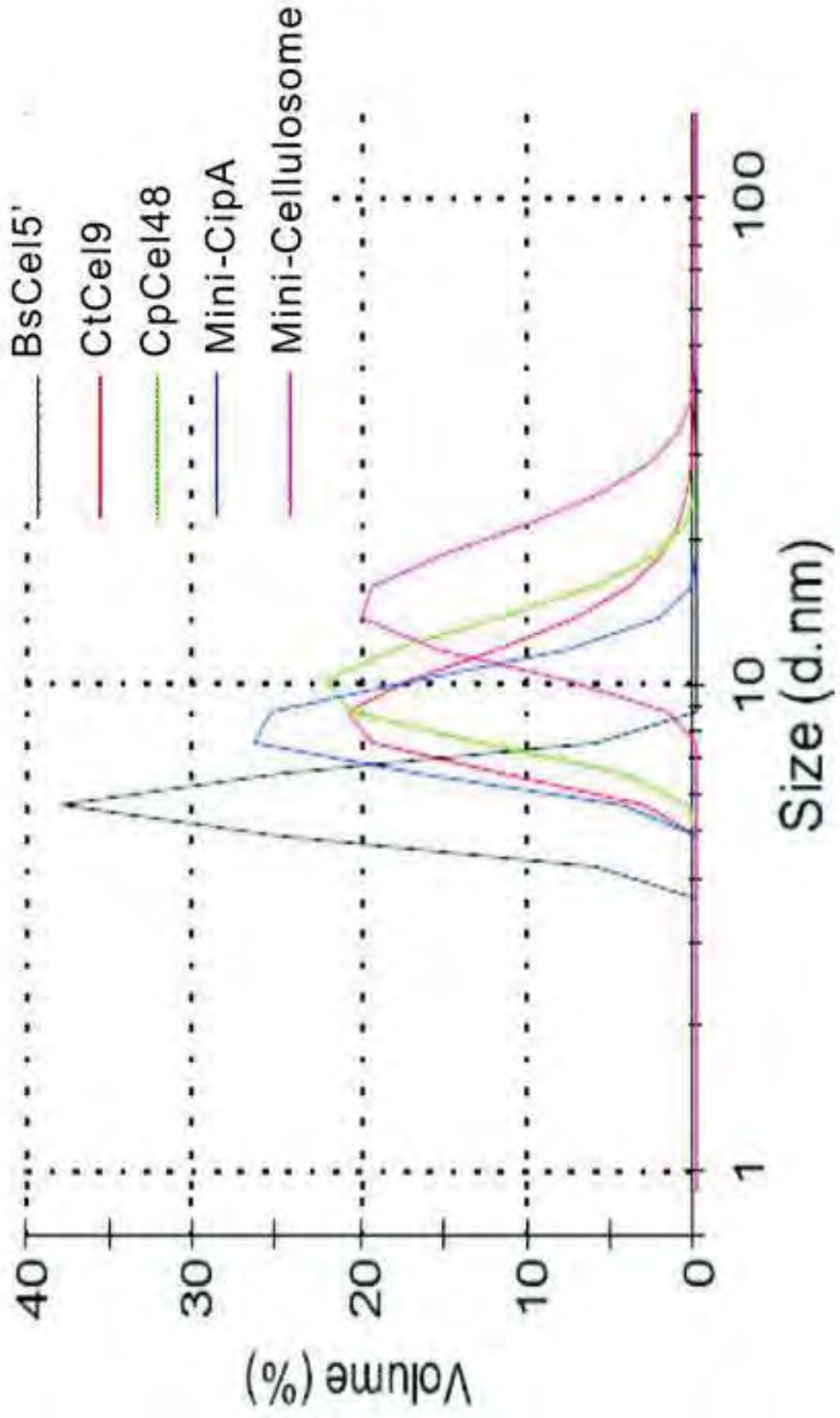
Figure(s)


kDa
M
1. Mini-CipA
2. BsCel5
3. BsCel5'
4. CtCel9
5. CtCel9'
6. CpCel48
7. CpCel48'
212
158
116
97.2
66.4
55.6
42.7
34.6
27.0

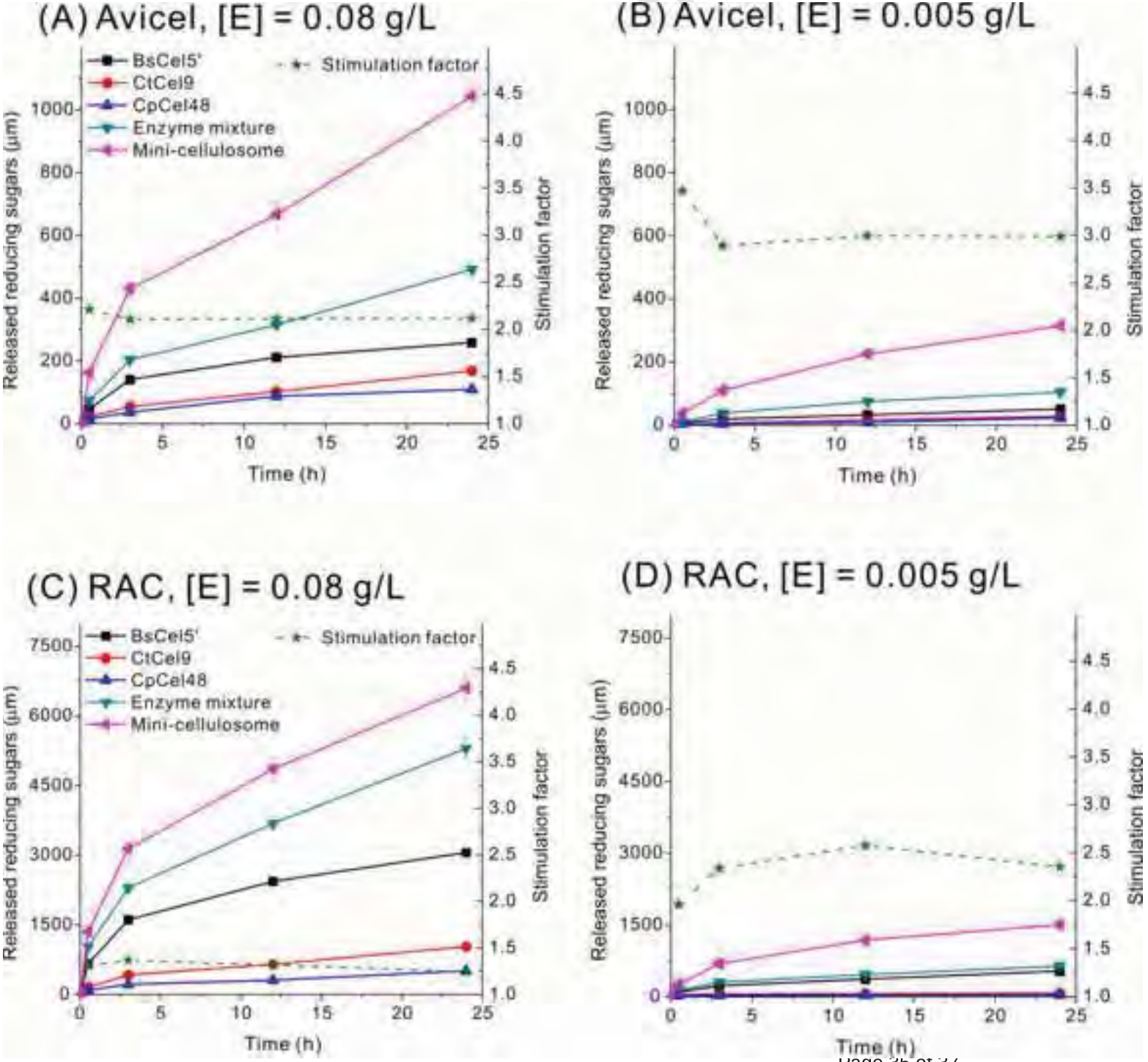

(A)
Mini-CipA : CtCel9' (molar ratio)
1:1
1:3
1:6
0:3
P
S
M
kDa
Mini-CipA
CtCel9'
97.2
66.4
55.6
(B)
KDa
M
97.2
66.4
55.6
42.7
CpCel48'
Mini-CipA
CtCel9'
BsCel5'
(C)
KDa
M
1. Mini-CipA
2. BsCel5'
3. CtCel9
4. CpCel48
5. Mini -CipA+BsCel5'
+CtCel9+CpCel48
116
97.2
66.4
55.6
42.7

1. BsCel5'
2. Mini-CipA+BsCel5'
3. CtCel9
4. Mini-CipA+CtCel9
5. CpCel48
6. Mini-CipA+CpCel48
7. Mini-CipA
8. Mini-CipA+BsCel5'+CtCel9+CpCel48

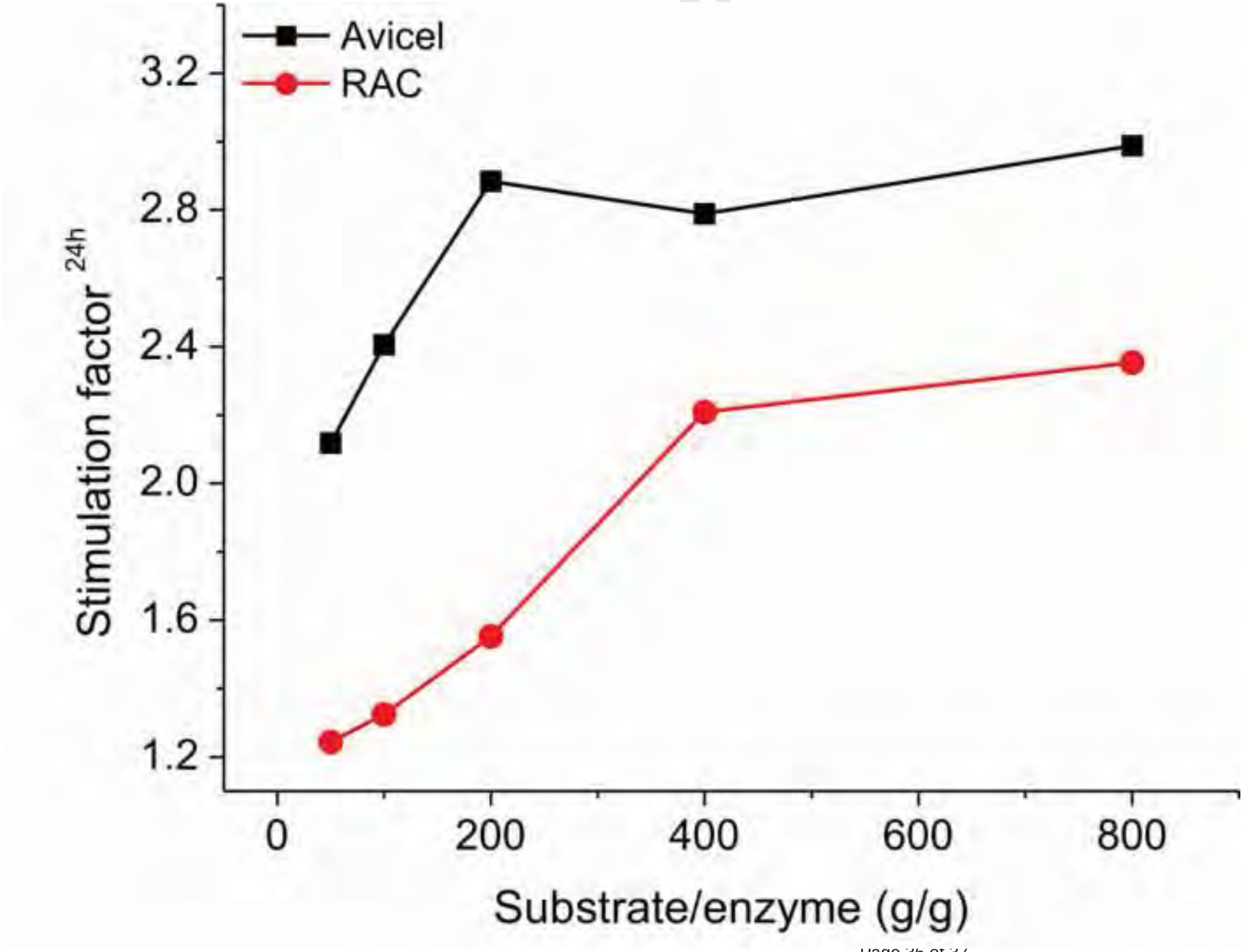
Figure(s)


BsCel5'
CtCel9
CpCel48
Mini-CipA
Mini-Cellulosome
Volume (%)
40
30
20
10
0
1
10
100
Size (d.nm)

Figure(s)


(A) Avicel, [E] = 0.08 g/L
(B) Avicel, [E] = 0.005 g/L
(C) RAC, [E] = 0.08 g/L
(D) RAC, [E] = 0.005 g/L
BsCel5'
CtCel9
CpCel48
Enzyme mixture
Mini-cellulosome
Stimulation factor
Released reducing sugars (μm)
Time (h)

Figure(s)


Avicel
RAC
3.2
2.8
2.4
2.0
1.6
1.2
Stimulation factor 24h
0
200
400
600
800
Substrate/enzyme (g/g)