Biocatalysis, Protein Engineering, and Nanobiotechnology **Biotechnology and Bioengineering**
DOI 10.1002/bit.26242

Engineering the *N*-terminal End of CelA Results in Improved Performance and Growth of *Caldicellulosiruptor bescii* on Crystalline Cellulose†

Sun-Ki Kim[1,3], Daehwan Chung[2,3], Michael E. Himmel[2,3], Yannick J. Bomble[2,3], and Janet Westpheling[1,3*]

Sun-Ki Kim
Email: skkim@uga.edu
Daehwan Chung
Email: Daehwan.Chung@nrel.gov
Michael E. Himmel
Email: Mike.Himmel@nrel.gov
Yannick J. Bomble
Email: yannick.bomble@nrel.gov
Janet Westpheling
Email: janwest@uga.edu

[1]Department of Genetics, University of Georgia, Athens, Georgia 30602, USA

[2]Biosciences Center, National Renewable Energy Laboratory, Golden, Colorado

[3]The BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 37831, USA

* Corresponding author: Email: janwest@uga.edu

Phone: +1-706-542-1436 FAX: +1-706-542-3910

†This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/bit.26242]

Additional Supporting Information may be found in the online version of this article.

Received October 3, 2016; Revision Received November 16, 2016; Accepted December 18, 2016

Abstract

CelA is the most abundant enzyme secreted by *Caldicellulosiruptor bescii* and has been shown to outperform mixtures of commercially available exo- and endoglucanases *in vitro.* CelA contains both a glycoside hydrolase family 9 endoglucanase and a glycoside hydrolase family 48 exoglucanase known to be synergistic in their activity, connected by three cellulose-binding domains via linker peptides. Here, repeated aspartate residues were introduced into the N-terminal ends of CelA GH9 and GH48 domains to improve secretion efficiency and/or catalytic efficiency of CelA. Among several constructs, the highest activity on carboxymethylcellulose (CMC), 0.81 ± 0.03 mg/mL was observed for the *C. bescii* strain containing CelA with 5-aspartate tag at the *N*-terminal end of GH9 domain – an 82% increase over wild type CelA. In addition, Expression of CelA with *N*-terminal repeated aspartate residues in *C. bescii* results in a dramatic increase in its ability to grow on Avicel.

Keywords: biomass deconstruction; CelA; repeated aspartate residues; *Caldicellulosiruptior*

Introduction

Organisms from the hyperthermophilic genus *Caldicellulosiruptor* have the unusual ability to grow on a variety of lignocellulosic biomass substrates without conventional pretreatment (Yang et al., 2009). However, to be industrially relevant, where acidic thermal chemical pretreatments are likely to be used to enhance digestibility, we chose the industrial cellulose standard, Avicel, to be used as substrate for these studies. The crystalline cellulose content in Avicel is relatively high and thought to be similar to cellulose in pretreated biomass, as most acidic pretreatments remove hemicellulose and amorphous cellulose (Ishizawa et al., 2009; Park et al., 2010). Avicel is also well suited to studies of the effects of modifying cellulase signal peptides because enzyme loss to lignins and hemicelluloses is less likely, providing more accurate digestibility analysis. Unlike most cellulolytic species of the *Clostridium* genus that rely on complex protein structures, called cellulosomes, to attach to and solubilize plant cell walls, *Caldicellulosiruptor* species secrete primarily free multifunctional enzymes into the exoproteome. One such cellulase, CelA, is the most abundant enzyme secreted by *C. bescii* (Lochner et al., 2011; Yang et al., 2009), and we recently reported that a deletion of *celA* in *C. bescii* resulted in a dramatic reduction in its ability to utilize complex biomass (Young et al., 2014).

Secreted bacterial proteins are typically synthesized with *N*-terminal signal peptides that target them for either protein secretion (Sec, protein secretion) or the twin-arginine translocation (Tat) export pathway (Fu et al., 2007; Mergulhao et al., 2005). Whereas the Sec pathway is universally conserved, essential, and typically the primary route for protein export, the Tat pathway is found in some, but not all, bacteria (Palmer and Berks, 2012). The *C. bescii* genome includes genes necessary for the Sec pathway, but does not include genes related to the Tat pathway (Kataeva et al., 2009). The signal peptide of CelA is a typical Sec-type signal peptide composed of a positively charged *N*-terminus (N-region), a central hydrophobic region (H-region), and a polar *C*-

terminal region (C-region) (Fu et al., 2007) making it likely that CelA is exported by the Sec pathway.

In addition to the signal peptide, the *N*-terminal portion of the exported protein also plays an important role in its translocation. Changes in the *N*-terminus outside the signal sequence of exported proteins have a dramatic effect on protein transport (Le Loir et al., 1998; Li et al., 1988). Highly secreted proteins typically possess a negatively charged residue in the first five amino acids (Macintyre et al., 1990). Previous studies showed that fusion tag systems consisting of the *pelB* signal sequence and repeated aspartate tags improved both expression and secretion of lipase B (CalB) from *Candida antarctica* and asparaginase isozyme II (AsnB) from *E. coli* (Kim et al., 2015a; Kim et al., 2015b). Most studies used *E. coli* as a model host because the molecular mechanisms and elements for Gram positive and Gram negative bacterial Sec pathways are almost identical (Mori and Ito, 2001). In addition, the attachment of repeated aspartates enhanced catalytic efficiency of CalB and α-1,2-fucosyltransferase (FucT2) from *Helicobacter pylori* (Chin et al., 2015; Kim et al., 2015b). These findings suggest that alteration of the *N*-terminal terminus of CelA GH9 and CelA GH48 domains might improve secretion efficiency and/or catalytic efficiency of CelA as it did for recombinant CalB, AnsB, and FucT2. Based on this hypothesis, repeating aspartate tags were introduced into the N-terminal ends of CelA GH9 and GH48 domains to examine effects of length and location on protein activity and secretion. Here, we show that these modifications resulted in increases in both the activity of the exoproteome acting on CMC and Avicel; as well as a dramatic effect on the ability of cells containing these modified N-terminal ends of CelA GH9 and GH48 domains to grow on Avicel.

Materials and Methods

Strains, media, and culture conditions. *E. coli* and *C. bescii* strains used in this study are listed in Table 1. All *C. bescii* strains were cultured in anaerobic conditions at 65°C on solid or in liquid low

osmolarity defined (LOD) medium (Farkas et al., 2013), as previously described, with 5 g/L maltose or cellobiose as sole carbon source for routine growth and transformation experiments (Farkas et al., 2013). For growth of uracil auxotrophs, the defined medium with 40 μM uracil was used. *E. coli* DH5α strain was used as a host for plasmid DNA construction and preparation using standard techniques. *E. coli* cells were cultivated in LB medium containing apramycin (50 μg/mL). Plasmid DNA was isolated using a Qiagen Miniprep Kit (Qiagen, Valencia, CA, USA). Chromosomal DNA from *C. bescii* strains was extracted using the Quick-gDNA MiniPrep (Zymo, Irving, CA) as previously described (Chung et al., 2011).

Construction and transformation of vectors with modified. N-terminal ends of CelA GH9 and GH48 domains**.** Q5 High-Fidelity DNA polymerase (New England BioLabs, Ipswich, MA, USA) was used for all plasmid construction, restriction enzymes (New England BioLabs, Ipwich, MA, USA) and the Fast-link DNA ligase kit (Epicentre Biotechnologies, Madison, WI, USA) were used for plasmid construction according to the manufacturer's instructions. Plasmids pSKW06, pSKW07, pSKW08, and pSKW09 are identical to pDCW173 (an expression vector previously used to examine extracellular CelA, Chung et al., 2015a) except that these vectors contain repeated aspartate tags at the *N*-terminal end of CelA mature sequence (Fig. S1). To make these changes, A 13.3 kb DNA fragment with different lengths of repeated aspartate tags was amplified with the forward primer, SK031 (with XmaI site), and the reverse primers, SK032, SK033, SK034, or SK035 (with an XmaI site) using pDCW173 as template. For example, SK031 (the forward primer) and SK032 (the reverse primer) were used to amplify a CelA gene with 3-aspartate tag at the *N*-terminal end of GH9 domain. After gene amplification, all the PCR products were cut with XmaI and ligated to construct pSKW06, pSKW07, pSKW08, and pSKW09. Plasmids pSKW14, 15, and 16 were designed to attach repeated aspartate tags at the *N*-terminal end of CelA GH48 domain (Fig. S2). Using the pDCW173 as template and the forward primer, SK046 (with XmaI site), and the reverse primers, SK047, SK048, or SK049 (with XmaI site), CelA genes with repeated aspartate tags at the *N*-terminal end of GH48 domain were obtained. These linear DNA fragments were

digested with XmaI and ligated to construct pSKW14, pSKW15, and pSKW16. Plasmids were introduced into *E. coli* DH5α cells by electroporation in a 1-mm-gap cuvette at 1.8 kV and transformants were selected for apramycin resistance. All plasmids were sequenced by Automatic Sequencing (Genewiz, South Plainfield, NJ, USA). Electrotransformation of *C. bescii* cells was performed as previously described (Groom et al., 2014). After being electro-pulsed with plasmid DNA (~0.5 μg), the cultures were recovered in low osmolarity complex (LOC) medium (Farkas et al., 2013) at 65°C. Recovery cultures were transferred to liquid LOD medium (Farkas et al., 2013) without uracil to allow selection of uracil prototrophs. Cultures were plated on solid LOD media to obtain isolated colonies that were picked into 10 mL of LOD medium. Total DNA was extracted. Taq polymerase (Sigma, St. Louis, MO, USA) was used for PCR reactions. PCR amplification was done with primers (DC460 and DC228) outside of the gene cassette on the plasmid to confirm the presence of the plasmid containing the CelA gene. Primers used for plasmid construction and confirmation are listed in Table S1.

Enzyme activity assays. Cellulolytic activity was measured using 10 g/L of either CMC or Avicel in MES reaction buffer (pH 5.5) as previously described (Kanafusa-Shinkai et al., 2013). Cells were grown in a 2 L volume of LOD medium with 40 mM MOPS containing 5 g/L cellobiose as carbon source. Culture broth was centrifuged (6,000 × g at 4°C for 15 min), filtered (glass fiber, 0.7 μm), to separate out cells. The 2 L of extracellular protein (ECP) was loaded to a hollow fiber cartridge with 3 kDa molecular weight cut off (GE healthcare, Buckinghamshire, UK) and eluted with 50 mL buffer 20 mM MES/2 mM β–mercaptoethanol (pH 5.5). Concentration of the 50 mL ECP was increased further (~25 times) with Vivaspin column (10 kDa molecular weight cut off, Sartorius, Goettingen, Germany). Protein concentrations were determined using the Bio-Rad protein assay kit with bovine serum albumin (BSA) as the standard. Twenty five μg/mL of the extracellular protein fraction was added to each reaction and incubated at 75°C (1 h for CMC and 24 h for Avicel). Reducing sugars in the supernatant were measured using dinitrosalicylic acid (DNS). Samples and

standards (glucose) were mixed 1:1 with DNS reaction solution, boiled for two min and absorbance was measured at OD_{575}. Activity was reported as mg/mL of glucose released.

Growth of *C. bescii* strains on cellobiose and Avicel. To measure growth on cellobiose, cells were sub-cultured twice in LOD medium with 5 g/L maltose as sole carbon source and this culture was used to inoculate media with 5 g/L cellobiose (1% total volume for all experiments) as sole carbon source, 50 mL LOD medium with 40 mM MOPS, and incubated at 65°C with shaking at 150 rpm. Cell growth on cellobiose was measured by optical density (OD) at 680 nm using a Jenway Genova spectrophotometer. To measure growth on Avicel, cells were subcultured in LOD medium with 5 g/L maltose first and then finally with 5 g/L Avicel. Colony-forming units (CFU) were measured by plating cells on LOC medium. We previously reported plating efficiencies for LOD medium (Chung et al., 2011; Farkas et al., 2013) and recently showed that the plating efficiencies for LOD and LOC medium are the same (Kim et al., 2016).

Results and Discussion

Introduction of aspartate residues into the N-terminal ends of CelA GH9 and GH48 domains**.** CelA is the most abundant extracellular enzyme produced by *C. bescii* (Lochner et al., 2011; Yang et al., 2009) and has cellulase activity almost equal to the entire purified *C. bescii* exoproteome (Brunecky et al., 2013). The signal peptide of CelA has been used successfully to secrete other proteins from *C. bescii* and is currently the most effective signal peptide so far described for members of this genus (Chung et al., 2015b). The introduction of repeated aspartate residues has been shown to increase the secretion; as well as the specific activity of the lipase B (CalB) from *Candida antarctica* and the asparaginase isozyme II (AsnB) from *E. coli* (Kim et al., 2015a; Kim et al., 2015b). To investigate whether or not repeated aspartate residues would improve the secretion and activity of CelA in *C. bescii*, aspartate residues were introduced into the N-terminal ends of CelA GH9 and GH48 domains as shown in Fig. 1 and Fig. S3. Residues were placed in various

locations within the signal sequence similar to the engineering of the lipase signal peptide (Kim et al., 2015b). These CelA variants were named as indicated in Fig. 1 and details of their construction are shown in Fig. S1,S2, and S3. For the D3-GH48 and D5-GH48 variants, one of the two identical CBM3b domains was always recombined out of the signal peptide during replication in *E. coli* DH5α (Fig. 1) and this is why these constructs only have one CBM3b. Plasmids containing the various changes; as well as the wild type version of CelA, were transformed into JWCB29, a strain of *C. bescii* containing a deletion of *celA* (Young et al., 2014) to eliminate background activity from the wild type protein. Transformants were selected for uracil prototrophy as previously described (Groom et al., 2014). All of the constructed variants readily transformed JWCB29 with the exception of pSKW08 (D7-GH9) – in spite of repeated attempts. Transformants containing the constructs were grown at 65°C to accommodate the expression of *C. thermocellum pyrF* gene used for complementation and plasmid selection. PCR analysis was used to verify the presence of the plasmid in the transformants using primers (DC460 and DC228) that amplify that portion of the plasmid containing the open reading frame of CelA (Fig. S4 and S5). Total DNA from *C. bescii* strains containing the plasmids was used to back-transform *E. coli* DH5α strain to verify autonomous replication of the plasmids and two different restriction endonuclease digests performed on plasmid DNA purified from two independent back-transformants resulted in identical digestion patterns to the original plasmid (Fig. S4 and S5). These results indicated that the plasmids were successfully transformed into *C. bescii* and structurally stable during transformation and replication in *C. bescii* and back-transformation into *E. coli.*

Expression of CelA with *N*-terminally repeated aspartic acid residues in *C. bescii* resulted in an increase in the activity of the exoproteome on CMC and Avicel. To investigate effect of the aspartate residues in the N-terminal ends of CelA GH9 and GH48 domains on CelA activity, *C. bescii* cells were grown at 65°C and proteins present in the extracellular (ECP) fraction were assayed at 75°C. Carboxymethylcellulose (CMC) and Avicel are traditionally used to assay for endoglucanase and exoglucanase activity, respectively, by sugar release (Adney et al., 1994). All of

the *C. bescii* strains containing the engineered N-terminal ends of CelA GH9 and GH48 domains, D3-GH9, D5-GH9, D9-GH9, D3-GH48, D5-GH48, and D7-GH48 showed a significant increase in activity on CMC, compared with that of wild type CelA. For Avicel hydrolysis; however, only the *C. bescii* strain containing D5-GH48 showed a significant improvement in activity (Fig. 2). Among them, the highest activity on CMC, 0.81 ± 0.03 mg/mL was observed for the *C. bescii* strain containing D5-GH9 – an 82% increase over wild type CelA. This result is consistent with previous studies showing that among a series of variants containing repeated aspartate residues of various lengths, five aspartates are optimum for the improved activity of lipase B (CalB) from *Candida antarctica* and the asparaginase isozyme II (AsnB) from *E. coli* (Kim et al., 2015a; Kim et al., 2015b). For CalB, the improved activity of the enzyme could be ascribed to a conformational change in CalB structure caused by the negative charge of the five aspartate residues (Kim et al., 2015b). We speculate that the repeated aspartate residues in the N-terminal ends of CelA GH9 and GH48 domains might play the same role in improving the activity of CelA by changing its structure. As shown in Fig. 2, the effect of the repeated aspartate residues on the endoglucanase activity of CelA was more dramatic than this effect on the exoglucanase activity. This result might also be explained by a change in the structure of the CelA GH9, as that domain of the protein is known to be responsible for the endoglucanase activity of CelA (Yi et al., 2013). CelA has also been shown to be glycosylated and although the predicted sites of glycosylation are not predicted to affect the GH9 domain of the protein, changes in the structure of the protein might affect the pattern or extent of glycosylation resulting in an effect on protein activity and/or thermostability (Chung et al., 2015a; Jeoh et al., 2008).

Expression of CelA with *N*-terminal repeated aspartic acid residues in *C. bescii* results in a dramatic increase in its ability to grow on Avicel. Whereas the activity of the exoproteome of most of the *C. bescii* strains expressing different versions of the N-terminal ends of CelA GH9 and GH48 domains we tested did not affect activity on Avicel, there was a surprising and dramatic effect on the ability of these *C. bescii* strains to grow on Avicel. As shown in Fig. 3A and 3B,while there

was a slight difference in the lag phase for some of the strains, the final cell densities were almost identical, indicating that the addition of the aspartic acid residues did not result in a general growth defect for any of the strains. Expression of the D3-GH9, D5-GH9 or D9-GH9 CelA in *C. bescii*, resulted in a substantial increase in the ability of these strains to grow on Avicel (Fig. 3C). Among them, the highest cell densities were observed for cells containing the CelA variants with DH3-GH9, D5-GH9 and D9-GH9 – all changes that would be expected to affect the endoglucanase activity of the enzyme. For the D3-GH9 variant, there was a 3.3 and 5.5 fold increase at 24 and 48 h, respectively ($4.4 \pm 0.1 \times 10^6$ CFU/mL at 24 h) and (7.8×10^5 CFU/mL at 48 h), over that for wild type CelA. For the D5-GH9 variant, there was a 4.6 and 24.5 fold increase observed at 24 and 48 h, respectively ($6.2 \pm 0.3 \times 10^6$ CFU/mL at 24 h) and ($3.5 \pm 0.2 \times 10^6$ CFU/mL at 48 h), over that for wild type CelA. For the D9-GH9 variant there was a 3.8 and 6.0 fold increase at 24 and 48 h, respectively ($5.2 \pm 0.6 \times 10^6$ CFU/mL at 24 h) and ($8.6 \pm 2.9 \times 10^5$ CFU/mL at 48 h), over that for wild type CelA. The other variants were not significantly different from wild type CelA. These results suggest that the endoglucanase activity may be rate limiting for the growth of *C. bescii* on crystalline cellulose substrate. This conclusion is consistent with our previous study showing that the addition of the E1 endoglucanase from *Acidothermus cellulolyticus* improved both the *in vitro* and *in vivo* cellulolytic activity of the exoproteome, likely by synergistic interaction with CelA (Chung et al., 2015b).

Conclusions

Following the hypothesis that insertion of multiple aspartic acid residues into the N-terminal ends of either the GH9A or GH48 ORFs might affect enzyme activity proved to be true in the case of *C. bescii* CelA. This modification to the wild type sequence resulted in increased performance of the endoglucanase domain (D3- and D5-GH9), and to a lesser extent an increase in the activity of the exoglucanase domain (D5-GH48) was also observed. We attribute these changes in activity to

potential small changes in the protein packing for both the GH9 and GH48. Importantly, most efficient cellulase systems contain highly active exocellulase enzymes capable of decrystallizing cellulose, so the observed increase in activity on Avicel bears directly upon the ability of CelA to degrade realistic biomass feedstocks destined for biofuels production.

Acknowledgements

This work was supported by the BioEnergy Science Center, US DOE Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the US DOE under contract DE-AC05-00OR22725. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Adney, W. S., Ehrman, C. I., Baker, J. O., Thomas, S. R., Himmel, M. E. 1994. Cellulase assays - methods from empirical mathematical-models. Acs Sym Ser 566:218-235.

Brunecky, R., Alahuhta, M., Xu, Q., Donohoe, B. S., Crowley, M. F., Kataeva, I. A., Yang, S. J., Resch, M. G., Adams, M. W. W., Lunin, V. V., Himmel, M. E., Bomble, Y. J. 2013. Revealing nature's cellulase diversity: the digestion mechanism of *Caldicellulosiruptor bescii* CelA. Science 342**:**1513-1516.

Chin, Y. W., Kim, J. Y., Lee, W. H., Seo, J. H. 2015. Enhanced production of 2'-fucosyllactose in engineered *Escherichia coli* BL21star(DE3) by modulation of lactose metabolism and fucosyltransferase. J Biotechnol 210**:**107-115.

Chung, D., Young, J., Bomble, Y. J., Vander Wall, T. A., Groom, J., Himmel, M. E., Westpheling, J. 2015a. Homologous expression of the *Caldicellulosiruptor bescii* CelA reveals that the extracellular protein is glycosylated. Plos One 10:3.

Chung, D., Young, J., Cha, M., Brunecky, R., Bomble, Y. J., Himmel, M. E., Westpheling, J. 2015b. Expression of the *Acidothermus cellulolyticus* E1 endoglucanase in *Caldicellulosiruptor bescii* enhances its ability to deconstruct crystalline cellulose. Biotechnol Biofuels 8:113.

Chung, D., Huddleston, J. R., Farkas, J., Westpheling, J. 2011. Identification and characterization of CbeI, a novel thermostable restriction enzyme from *Caldicellulosiruptor bescii* DSM 6725 and a member of a new subfamily of HaeIII-like enzymes. J Ind Microbiol Biotechnol 38**:**1867-1877.

Farkas, J., Chung, D., Cha, M., Copeland, J., Grayeski, P., Westpheling, J. 2013. Improved growth media and culture techniques for genetic analysis and assessment of biomass utilization by *Caldicellulosiruptor bescii*. J Ind Microbiol Biotechnol 40**:**41-49.

Fu, L. L., Xu, Z. R., Li, W. F., Shuai, J. B., Lu, P., Hu, C. X. 2007. Protein secretion pathways in *Bacillus subtilis*: Implication for optimization of heterologous protein secretion. Biotechnol Adv 25**:**1-12.

Groom, J., Chung, D., Young, J., Westpheling, J. 2014. Heterologous complementation of a *pyrF* deletion in *Caldicellulosiruptor hydrothermalis* generates a new host for the analysis of biomass deconstruction. Biotechnol Biofuels 7:132.

Ishizawa, C. I., Jeoh, T., Adney, W. S., Himmel, M. E., Johnson, D. K., Davis, M. F. 2009. Can delignification decrease cellulose digestibility in acid pretreated corn stover? Cellulose 16**:**677-686.

Jeoh, T., Michener, W., Himmel, M. E., Decker, S. R., Adney, W. S. 2008. Implications of cellobiohydrolase glycosylation for use in biomass conversion. Biotechnol Biofuels 1:10.

Kanafusa-Shinkai, S., Wakayama, J., Tsukamoto, K., Hayashi, N., Miyazaki, Y., Ohmori, H., Tajima, K., Yokoyama, H.

2013. Degradation of microcrystalline cellulose and non-pretreated plant biomass by a cell-free extracellular cellulase/hemicellulase system from the extreme thermophilic bacterium *Caldicellulosiruptor bescii*. J Biosci Bioeng 115:64-70.

Kataeva, I. A., Yang, S. J., Dam, P., Poole, F. L., Yin, Y., Zhou, F. F., Chou, W. C., Xu, Y., Goodwin, L., Sims, D. R., Detter, J. C., Hauser, L. J., Westpheling, J., Adams, M. W. W. 2009. Genome sequence of the anaerobic, thermophilic, and cellulolytic bacterium "*Anaerocellum thermophilum*" DSM 6725. J Bacteriol 191:3760-3761.

Kim, S. K., Chung, D., Himmel, M. E., Bomble, Y. J., Westpheling, J. 2016. Heterologous expression of family 10 xylanases from *Acidothermus cellulolyticus* enhances the exoproteome of *Caldicellulosiruptor bescii* and growth on xylan substrates. Biotechnol Biofuels 9:176.

Kim, S. K., Min, W. K., Park, Y. C., Seo, J. H. 2015a. Application of repeated aspartate tags to improving extracellular production of *Escherichia coli* L-asparaginase isozyme II. Enzyme Microb Technol 79-80:49-54.

Kim, S. K., Park, Y. C., Lee, H. H., Jeon, S. T., Min, W. K., Seo, J. H. 2015b. Simple amino acid tags improve both expression and secretion of *Candida antarctica* lipase B in recombinant *Escherichia coli*. Biotechnol Bioeng 112:346-355.

Le Loir, Y., Gruss, A., Ehrlich, S. D., Langella, P. 1998. A nine-residue synthetic propeptide enhances secretion efficiency of heterologous proteins in *Lactococcus lactis*. J Bacteriol 180:1895-1903.

Li, P., Beckwith, J., Inouye, H. 1988. Alteration of the amino terminus of the mature sequence of a periplasmic protein can severely affect protein export in *Escherichia coli*. Proc Natl Acad Sci USA 85:7685-7689.

Lochner, A., Giannone, R. J., Rodriguez, M., Jr., Shah, M. B., Mielenz, J. R., Keller, M., Antranikian, G., Graham, D. E., Hettich, R. L. 2011. Use of label-free quantitative proteomics to distinguish the secreted cellulolytic systems of *Caldicellulosiruptor bescii* and *Caldicellulosiruptor obsidiansis*. Appl Environ Microbiol 77:4042-4054.

Macintyre, S., Eschbach, M. L., Mutschler, B. 1990. Export incompatibility of N-terminal basic residues in a mature polypeptide of *Escherichia coli* can be alleviated by optimizing the signal peptide. Mol Gen Genet 221:466-474.

Mergulhao, F. J. M., Summers, D. K., Monteiro, G. A. 2005. Recombinant protein secretion in *Escherichia coli*. Biotechnol Adv 23:177-202.

Mori, H., Ito, K. 2001. The Sec protein-translocation pathway. Trends Microbiol 9:494-500.

Palmer, T., Berks, B. C. 2012. The twin-arginine translocation (Tat) protein export pathway. Nat Rev Microbiol 10:483-496.

Park, S., Baker, J. O., Himmel, M. E., Parilla, P. A., Johnson, D. K. 2010. Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. Biotechnol Biofuels 3:10.

Yang, S. J., Kataeva, I., Hamilton-Brehm, S. D., Engle, N. L., Tschaplinski, T. J., Doeppke, C., Davis, M., Westpheling, J., Adams, M. W. 2009. Efficient degradation of lignocellulosic plant biomass, without pretreatment, by the thermophilic anaerobe "*Anaerocellum thermophilum*" DSM 6725. Appl Environ Microbiol 75**:**4762-4769.

Yi, Z., Su, X., Revindran, V., Mackie, R. I., Cann, I. 2013. Molecular and biochemical analyses of CbCel9A/Cel48A, a highly secreted multi-modular cellulase by *Caldicellulosiruptor bescii* during growth on crystalline cellulose. Plos One 8**:**12.

Young, J., Chung, D., Bomble, Y. J., Himmel, M. E., Westpheling, J. 2014. Deletion of *Caldicellulosiruptor bescii* CelA reveals its crucial role in the deconstruction of lignocellulosic biomass. Biotechnol Biofuels 7:142.

Figure Captions

Figure 1. Schematic diagram of the structures of recombinant CelA expression cassettes. Genes were expressed under the transcriptional control of the *C. bescii* S-layer promoter. SP, CelA signal peptide; CBM, carbohydrate-binding module; GH, glycoside hydrolase; D3, 3 aspartates; D5, 5 aspartates; D7, 7 aspartates; D9, 9 aspartates.

Figure 2. Relative enzymatic activity of the extracellular fraction of *C. bescii* strains on CMC and Avicel. Activity of extracellular protein (25 μg/mL concentrated protein) of *C. bescii* strains overexpressing CelA, D3-GH9, D5-GH9, D9-GH9, D3-GH48, D5-GH48, and D7-GH48 was measured in triplicate.

Figure 3. Growth of *C. bescii* strains on cellobiose (A, B) and Avicel (C). Growth of *C. bescii* strains overexpressing CelA, D3-GH9, D5-GH9, D9-GH9, D3-GH48, D5-GH48, and D7-GH48 was compared.

Table 1. Strains and plasmids used in this study

Name	Description	Reference
E. coli		
JW516	DH5α containing pSKW06 (Apramycin^R)	This study
JW517	DH5α containing pSKW07 (Apramycin^R)	This study
JW518	DH5α containing pSKW08 (Apramycin^R)	This study
JW519	DH5α containing pSKW09 (Apramycin^R)	This study
JW524	DH5α containing pSKW14 (Apramycin^R)	This study
JW525	DH5α containing pSKW15 (Apramycin^R)	This study
JW526	DH5α containing pSKW16 (Apramycin^R)	This study
C. bescii		
JWCB29	Δ*pyrFA ldh::ISCbe4* Δ*cbe1* Δ*celA* (*ura*$^-$/5-FOA^R)	(Young et al., 2014)
JWCB46	JWCB29 containing pDCW173 (*ura*$^+$/5-FOA^S)	(Chung et al., 2015a)
JWCB69	JWCB29 containing pSKW06 (*ura*$^+$/5-FOA^S)	This study
JWCB70	JWCB29 containing pSKW07 (*ura*$^+$/5-FOA^S)	This study
JWCB71	JWCB29 containing pSKW09 (*ura*$^+$/5-FOA^S)	This study
JWCB79	JWCB29 containing pSKW14 (*ura*$^+$/5-FOA^S)	This study
JWCB80	JWCB29 containing pSKW15 (*ura*$^+$/5-FOA^S)	This study
JWCB81	JWCB29 containing pSKW16 (*ura*$^+$/5-FOA^S)	This study
Plasmids		
pDCW173	Expression vector for extracellular CelA and source of CelA open reading frame (Apramycin^R)	(Chung et al., 2015a)
pSKW06	Expression vector for CelA with 3 aspartate tag at the N-terminus of GH9 domain (Apramycin^R)	This study
pSKW07	Expression vector for CelA with 5 aspartate tag at the N-terminus of GH9 domain (Apramycin^R)	This study
pSKW08	Expression vector for CelA with 7 aspartate tag at the N-terminus of GH9 domain (Apramycin^R)	This study
pSKW09	Expression vector for CelA with 9 aspartate tag at the N-terminus of GH9 domain (Apramycin^R)	This study
pSKW14	Expression vector for CelA with 3 aspartate tag at the N-terminus of GH48 domain (Apramycin^R)	This study
pSKW15	Expression vector for CelA with 5 aspartate tag at the N-terminus of GH48 domain (Apramycin^R)	This study

pSKW16	Expression vector for CelA with 7 aspartate tag at the N-terminus of GH48 domain (Apramycin^{R})	This study

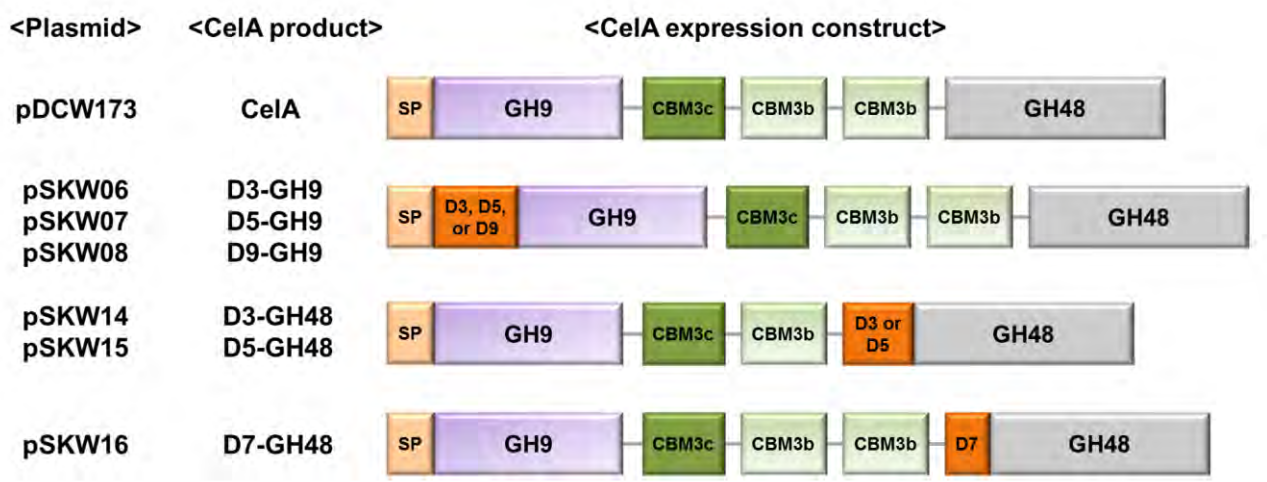

Figure 1

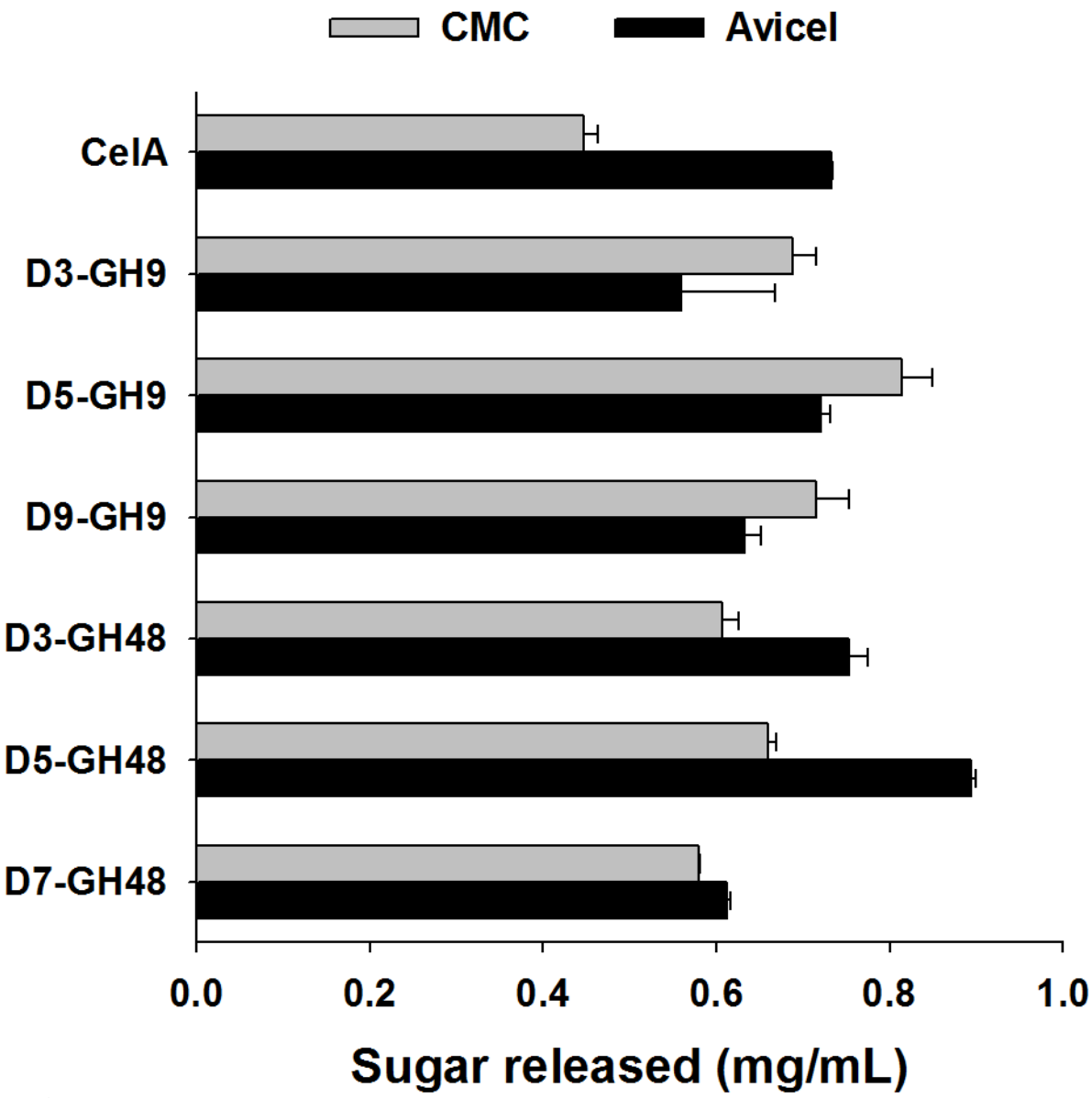

Figure 2

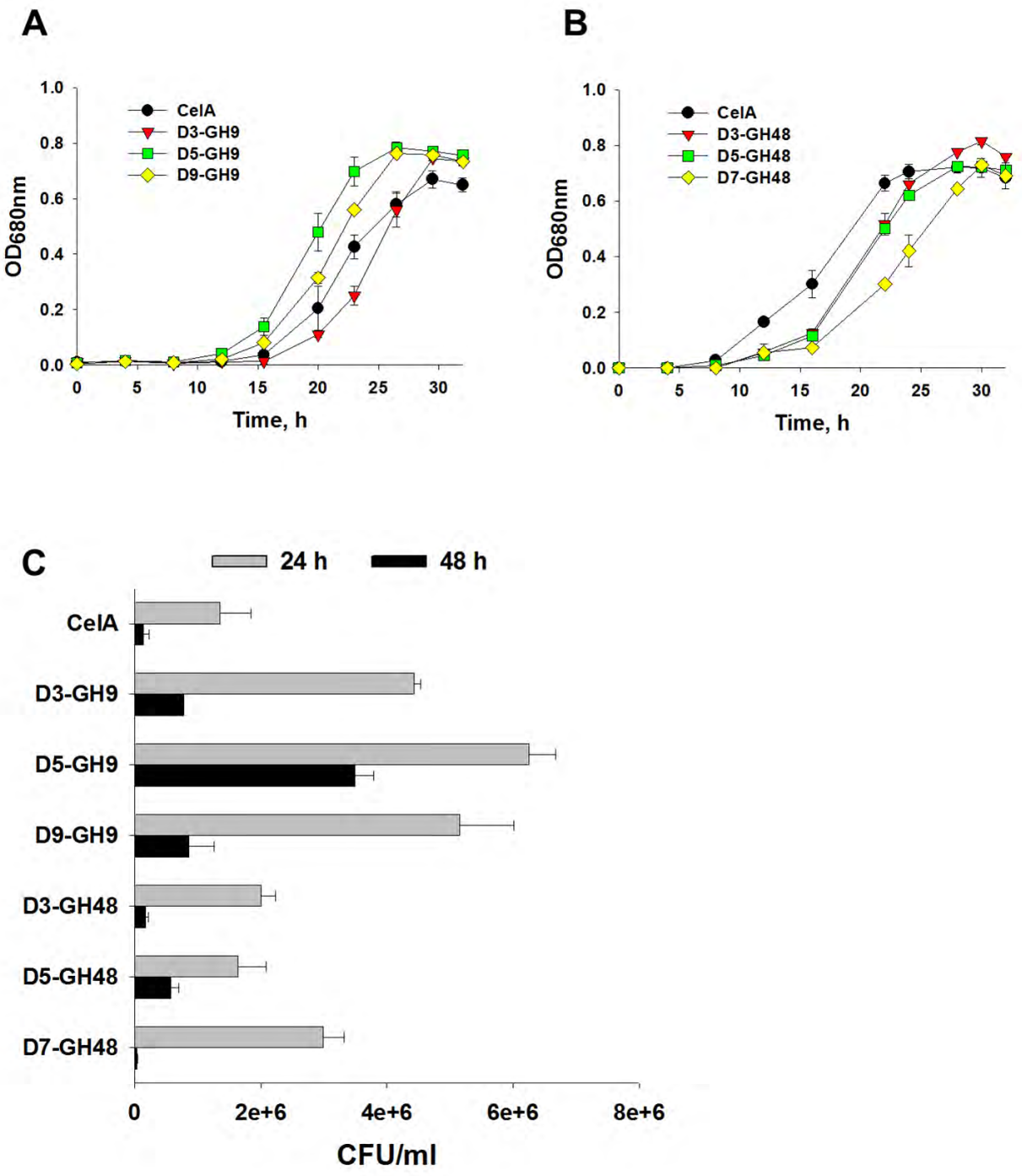

Figure 3