

**Construction of an efficient degradation system
for cellulosic biomass**

Jungu BAE

2015

Contents

Introduction	1
Chapter I	13
Exoproteome profiles of <i>Clostridium cellulovorans</i> on various carbon sources	
Chapter II	39
Proximity effect among cellulose-degrading enzymes displayed on the <i>Saccharomyces cerevisiae</i> cell surface	
Conclusions	62
Acknowledgements	63
Publications	64

Abbreviations

BG	β -glucosidase
BSA	Bovine serum albumin
CAZy	Carbohydrate-active enzymes
CBP	Consolidated bioprocessing
CBH	Cellobiohydrolase
CBM	Carbohydrate binding module
CbpA	Carbohydrate binding protein A
CE	Carbohydrate esterase
DNS	3,5-Dinitrosalicylic acid
EG	Endoglucanase
EGFP	Enhanced green fluorescent protein
ESI	Electrospray ionization
FDR	False discovery rate
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GH	Glycosyl hydrolases
GPI	Glycosylphosphatidylinositol
HCA	Hierarchical cluster analysis
HCD	Higher-energy c-trap dissociation
LC	Liquid chromatography
MES	2-Morpholinoethanesulfonic acid
MS	Mass spectrometry
OD	Optical density
PASC	Phosphoric acid swollen cellulose
PBS	Phosphate-buffered saline
PCA	Principal component analysis
PCR	Polymerase chain reaction
PL	Polysaccharolytic lyases
PNPG	p-nitrophenyl glucopyranoside

RFU	Relative fluorescence units
SDS	Sodium dodecyl sulfate
SLH	Surface layer homology
SpC	Spectral count
SSP	Substrate specific proteins
TMT	Tandem mass tag
YPD	Yeast extract peptone dextrose

Introduction

Petroleum is the world's main energy source. As society is completely dependent on petroleum, many problems including global warming, the limitations in the supply of fossil fuels, and energy security have arisen. To resolve those problems, the development of renewable alternative energy and construction of a sustainable society are required. Much effort has been made to develop alternative energies such as solar energy, wind power, geothermal power, wave power, biomass, etc. Among them, biomass has several unique characteristics; it can produce liquid transportation fuel such as biodiesel or bioethanol, which can be used directly in our current infrastructure. Moreover, not only energy but also value-added chemicals can be made, which suggests that biomass is a candidate resource for the complete substitution of petroleum. Because of these characteristics, biomass has been attracting great attention as a promising alternative energy.

Biorefinery

A biorefinery is a facility that integrates biomass conversion processes and equipments to produce fuels, power, and chemicals from biomass (1). As biomass is renewable and carbon-neutral unlike petroleum, it is expected that a sustainable society can be constructed through biorefineries. In a sugar platform concept, once polymeric forms of biomass have degraded to monomeric sugars, these sugars can be fermented to various kinds of biofuel or chemical precursors using metabolically engineered microorganisms. This sugar platform concept could completely replace the current petroleum platform.

In the USA and Brazil, corn and sugarcane have been used as biomass feedstock to produce sugars and convert them to bioethanol. As they consist of starches and the extraction efficiency of glucose is high enough, bioethanol is now produced from glucose at an industrial level and used in transportation fuel. However, these grain biomasses have limitations because there are increasing global problems in population and not enough arable land; thus, cellulosic biomass that includes remaining products after the extraction of glucose is required.

Cellulosic biomass

Cellulosic biomass mainly consists of polysaccharides such as cellulose and hemicellulose,

and it includes waste paper, grasses, leaves, and agricultural residues such as rice straw. There are several advantages of using cellulosic biomass over grain biomass for the production of biofuels and value-added chemical products. First, the waste materials can be utilized by microorganisms. Second, not only glucose, but also other C5 monomers including xylose are produced during the saccharification process, and they can eventually be fermented to precursors for broader kinds of chemicals. Third, some cellulosic crops grow on marginal lands that are currently unsuitable for arable farming, which leads to the effective use of unused lands and further reduction in greenhouse-gas emissions and the cost of biofuels (2).

However, the cost of biofuel production from cellulosic biomass is currently too high. First, the recalcitrancy of the crystalline cellulose; it is resistant to degradation as it is insoluble in water and composed of hydrogen-bonded crystalline fibers (3). Learning from nature how cellulose is degraded by microorganisms and the application of a synergistic reaction between cellulases will be critical for improving the efficiency of cellulose degradation. Second, the biofuel production process consists of multiple steps including enzyme production, saccharification, and fermentation, which require a complicated infrastructure and enhance contamination risks. Consolidated bioprocessing (CBP), which integrates the whole biofuel-production process, is expected to contribute to the reduction of biofuel-production costs (4, 5) (Fig. 1).

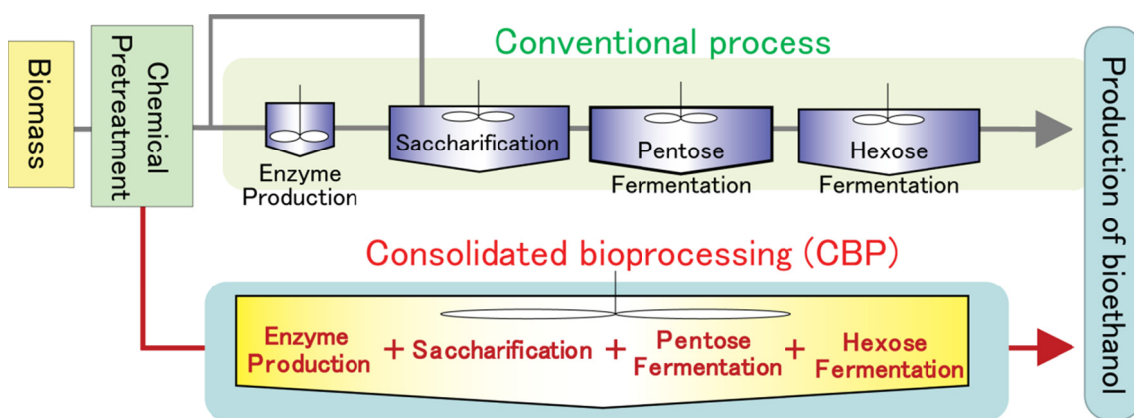


Fig. 1 Consolidated bioprocessing (CBP)

Cellulose degradation system

Several microorganisms have the ability to degrade cellulose efficiently. There are two types of cellulase degradation systems: free enzyme systems as secreted enzymes in

non-cellulosome-producing aerobic fungi and cellulosome systems as cell surface displayed enzymes in cellulosome-producing anaerobic bacteria.

Aerobic fungi, including *Trichoderma reesei* and *Neurospora crassa*, which do not produce cellulosomes, have adopted the free enzyme system that involves the secretion of various polysaccharide-degradation enzymes such as cellulases, including endo- and exoglucanases, which destroy the cellulose crystal structure and degrade cellulose into small compounds. Exoglucanases and endoglucanases are known to work synergistically (6, 7). The synergistic action has been shown in the sequential reaction of endoglucanase and exoglucanase. The suggested model for the mechanism of the synergistic action is that endoglucanases target the amorphous regions of cellulose to create new ends, from which exoglucanases can begin substrate hydrolysis into the crystalline regions of the polysaccharide (8).

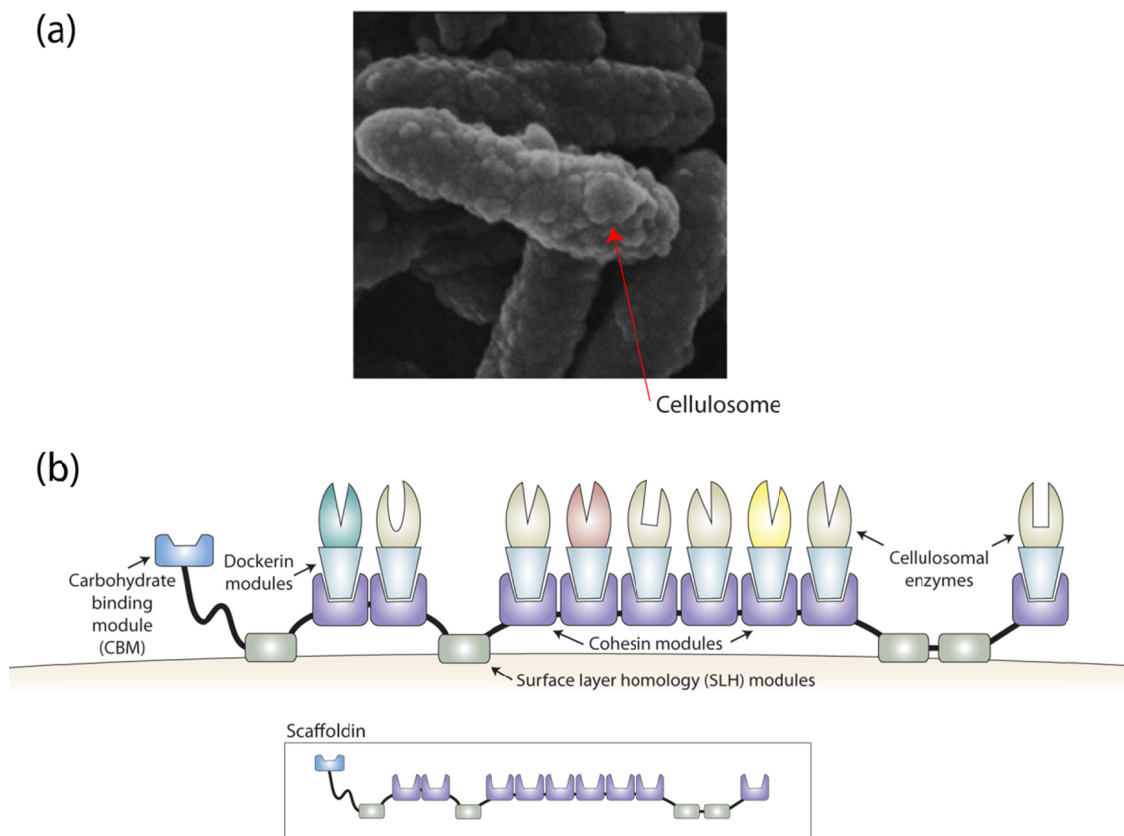


Fig.2 SEM image of *C. cellulovorans* (a) (9) and an illustration of cellulosome (b)

Some anaerobic bacteria including *Clostridium thermocellum* and *Clostridium cellulovorans* are known to produce an extracellular enzyme complex called “cellulosome” (Fig. 2a) (10). A cellulosome consists of a scaffoldin with several bound enzymes (11, 12), and this system has been considered the most efficient system for crystalline cellulose degradation (Fig. 2b). The scaffoldin is mainly composed of cohesin modules, a carbohydrate-binding module (CBM) (13), and surface layer homology (SLH) modules (14). The complex is constructed through interaction between a cohesin module and a dockerin module that is attached to each cellulosomal enzyme (15, 16). The cellulose degradation activity of the cellulosome is dramatically decreased when it lacks scaffoldin, which suggests that the cellulases react more synergistically when they are close, i.e. there is a proximity effect (17). The efficient cellulose degradation activity of the cellulosome mainly comes from the proximity effect. This proximity effect between cellulases should be applied to the biomass-degradation system for CBP to enhance the saccharification efficiency and lower the biofuel production cost.

Proteome analysis

Although there have been extensive studies on cellulosome-producing microorganisms and the efficiency of the cellulosomal system has been suggested, many of the strategies to degrade cellulose efficiently are poorly understood. By analyzing microorganisms comprehensively, we can learn from nature and apply their attractive strategies to the biomass-degradation system.

Proteome analysis that directly reflects the abundance of proteins is considered an effective approach for the comprehensive analysis of the biomass-degrading strategy of cellulosome-producing microorganisms. Extensive proteome analyses of *C. cellulovorans* and *C. thermocellum* have revealed that their cellulosomal gene expressions greatly differ depending on their growth substrates such as cellobiose, cellulose, xylose, and pectin (18, 19). For example, cellulase gene expression occurs in the presence of all substrates but increases in cellulose, and the expression of the genes encoding xylanase or pectate lyase increases when grown on xylan or pectin, respectively (20, 21).

Recent progress in analytical equipment and methods has enabled quantitative proteomic analyses using liquid chromatography/mass spectrometry systems (LC-MS/MS). The variation in cellulosomal enzyme-encoding gene expression between organisms grown on cellobiose and

cellulose has been investigated using this technology (22). Furthermore, a comprehensive comparison of the changes in the cellulosomal composition in *C. thermocellum* as a response to different carbon sources has confirmed the substrate-specific regulation at the protein level (23). However, these proteomic analyses have been focused on cellulosome samples purified with cellulose, not on non-cellulosomal enzymes.

Table 1 The number of cellulosomal and non-cellulosomal GH^a- and PL^b-encoding genes in the major cellulosome-producing clostridia.

Organism	Cellulosomal	Non-cellulosomal	Cellulosomal GHs + PLs /
	GHs ^a + PLs ^b	GHs + PLs	Non-cellulosomal GHs + PLs
<i>C. cellulovorans</i>	29	63	0.46
<i>C. cellulolyticum</i>	47	42	1.1
<i>C. thermocellum</i>	53	14	3.8

^a GHs: Glycosyl hydrolases, ^b PLs: Polysaccharolytic lyases

The potential importance of non-cellulosomal enzymes in cellulosome-producing microorganisms such as *C. cellulovorans* has been suggested based on whole genome sequence analyses, which have revealed that the number of non-cellulosomal enzyme-coding genes was relatively higher than those in other clostridia (24, 25). The ratio of cellulosomal glycosyl hydrolases (GHs) and polysaccharolytic lyases (PLs) to non-cellulosomal GHs and PLs is 0.46 (29/63) (Table 1). However, if the protein samples are not purified with avicel using CBM-avicel interaction for the analysis of non-cellulosomal enzymes, the number of identified proteins decreases due to the suppression of ionization in the mass spectrometer, which results from low separation performance (26). To overcome ionization suppression without pre-separation, significant improvements in separation performance are necessary. The monolithic silica column is a novel separation medium used in liquid chromatography that exhibits a lower column backpressure because of its high permeability. This property enables the preparation of a long column, leading to a better performance compared to conventional columns (27-29). According to a previous study (19), a higher number of proteins were identified using the monolithic silica column prepared in a long-fused silica capillary (300 cm) compared with when using a conventional particle-packed

column. There is an urgent need to conduct comprehensive proteome analysis of cellulosomal and non-cellulosomal enzymes using an LC-MS/MS system with a long monolithic silica column.

Yeast surface display system for proximity effect and consolidated bioprocessing (CBP)

CBP, which integrates all of the steps needed for biofuel production, has the potential to reduce the biofuel production costs compared to conventional processes that consist of multiple separated steps (4, 5). For CBP, microorganisms that degrade biomass substrates and produce desired products at high yields and titers are necessary. However, as no such microorganism has been found in nature, two strategies, one native strategy and one recombinant strategy, have been suggested. In the native strategy, ethanol fermentation ability is endowed to biomass-degrading microorganisms through genetic engineering. For example, *C. thermocellum* has been engineered to produce ethanol, but the titer and yield of ethanol were not high because of the low tolerance for ethanol (30). Additionally, it cannot metabolize degraded C5 products. In recombinant strategy, ethanol-fermentable microorganisms are genetically engineered to produce biomass-degrading enzymes to enable saccharification. *Saccharomyces cerevisiae* is a promising candidate for the recombinant strategy, as it can easily be genetically manipulated, is tolerant to various stresses that it might experience in an industrial environment, and is already able to produce ethanol from both glucose and xylose (31). There have been considerable efforts made to enhance the cellulose-degrading abilities of the engineered yeasts (32-35).

A yeast surface display system (36, 37) has been developed and applied as one of the recombinant strategies to construct yeast that can degrade cellulose and produce ethanol directly (Fig. 3). In the yeast surface display system, the gene for the target enzyme is fused with the secretion signal sequence, 3'-half of the α -agglutinin gene, and the GPI anchor attachment signal, which enable the immobilization of the target enzyme on the yeast cell wall. An arming yeast (38), that displays endoglucanase, cellobiohydrolase (exoglucanase), and β -glucosidase, has been constructed and successfully produced ethanol directly from cellulose as substrate (39). In addition, as one of the characteristics of arming yeast, the proximity effect between displayed cellulases has been proposed since a number of cellulases are densely concentrated on the surface of the yeast. These characteristics support the idea that the yeast surface display system could be effective for cellulosic biofuel production, which enables CBP and exploits the proximity effect shown in cellulosome.

However, it has not been confirmed whether the arming yeast shows proximity effect. A synergistic reaction such as a proximity effect is an important factor for enhancing the saccharification efficiency, and it is essential to demonstrate whether the yeast surface display system shows the proximity effect to reveal the potential of the system.

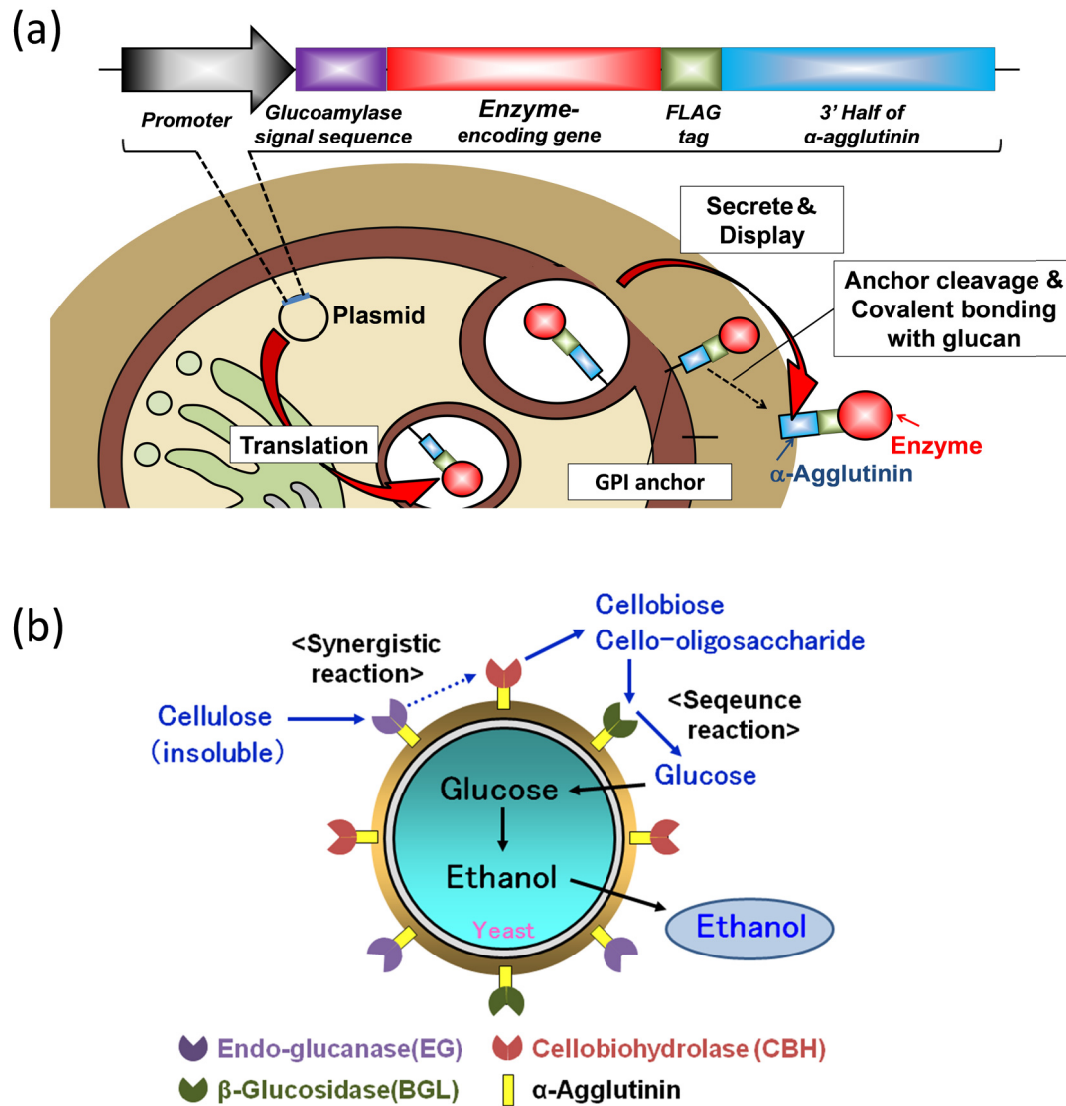


Fig. 3 Yeast surface display system (a) and an arming yeast displaying cellulases (b)

References

1. **Kamm B, Kamm M.** 2004. Principles of biorefineries. *Appl. Microbiol. Biotechnol.* **64**:137-145.
2. **Gelfand I, Sahajpal R, Zhang X, Izaurrealde RC, Gross KL, Robertson GP.** 2013. Sustainable bioenergy production from marginal lands in the US Midwest. *Nature* **493**:514-517.
3. **Mansfield SD, Mooney C, Saddler JN.** 1999. Substrate and enzyme characteristics that limit cellulose hydrolysis. *Biotechnol. Prog.* **15**:804–816.
4. **Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS.** 2002. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **66**:506–577.
5. **Olson DG, McBride JE, Shaw AJ, Lynd LR.** 2012. Recent progress in consolidated bioprocessing. *Curr. Opin. Biotechnol.* **23**:396-405.
6. **Murashima K, Kosugi A, Doi RH.** 2002. Synergistic effects on crystalline cellulose degradation between cellulosomal cellulases from *Clostridium cellulovorans*. *J. Bacteriol.* **184**:5088–5095.
7. **Berger E, Zhang D, Zverlov VV, Schwarz WH.** 2007. Two noncellulosomal cellulases of *Clostridium thermocellum*, Cel9I and Cel48Y, hydrolyse crystalline cellulose synergistically. *FEMS Microbiol. Lett.* **268**:194–201.
8. **Gilbert HJ.** 2010. The biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiol.* **153**:444–455.
9. **Tamaru Y, Miyake H, Kuroda K, Ueda M, Doi RH.** 2010. Comparative genomics of the mesophilic cellulosome-producing *Clostridium cellulovorans* and its application to biofuel production via consolidated bioprocessing. *Environ. Technol.* **31**:889-903.
10. **Doi RH, Park JS, Liu CC, Malburg LM, Tamaru Y, Ichiishi A, Ibrahim A.** 1998. Cellulosome and noncellulosomal cellulases of *Clostridium cellulovorans*. *Extremophiles* **2**:53–60.
11. **Shoseyov O, Takagi M, Goldstein MA, Doi RH.** 1992. Primary sequence analysis of *Clostridium cellulovorans* cellulose binding protein A. *Proc. Natl. Acad. Sci. USA* **89**:3483–3487.
12. **Shoseyov O, Doi RH.** 1990. Essential 170-kDa subunit for degradation of crystalline cellulose

- by *Clostridium cellulovorans* cellulase. Proc. Natl. Acad. Sci. USA **87**:2192–2195.
13. **Goldstein MA, Takagi M, Hashida S, Shoseyov O, Doi RH, Segel IH.** 1993. Characterization of the cellulose-binding domain of the *Clostridium cellulovorans* cellulose-binding protein A. J. Bacteriol. **175**:5762–5768.
 14. **Kosugi A, Amano Y, Murashima K, Doi RH.** 2004. Hydrophilic domains of scaffolding protein CbpA promote glycosyl hydrolase activity and localization of cellulosomes to the cell surface of *Clostridium cellulovorans*. J. Bacteriol. **186**:6351–6359.
 15. **Lytle B, Wu JH.** 1998. Involvement of both dockerin subdomains in assembly of the *Clostridium thermocellum* cellulosome. J. Bacteriol. **180**:6581–6585.
 16. **Bayer EA, Morag E, Lamed R.** 1994. The cellulosome – a treasure-trove for biotechnology. Trends Biotechnol. **12**:379–386.
 17. **Fierobe HP, Bayer EA, Tardif C, Czjzek M, Mechaly A, Belaich A, Lamed R, Shoham Y, Belaich JP.** 2002. Degradation of cellulose substrates by cellulosome chimeras. Substrate targeting versus proximity of enzyme components. J. Biol. Chem. **277**:49621–49630.
 18. **Cho W, Jeon SD, Shim HJ, Doi RH, Han SO.** 2010. Cellulosomic profiling produced by *Clostridium cellulovorans* during growth on different carbon sources explored by the cohesin marker. J. Biotechnol. **145**:233–239.
 19. **Morisaka H, Matsui K, Tatsukami Y, Kuroda K, Miyake H, Tamaru Y, Ueda M.** 2012. Profile of native cellulosomal proteins of *Clostridium cellulovorans* adapted to various carbon sources. AMB Express **2**:37.
 20. **Han SO, Cho HY, Yukawa H, Inui M, Doi RH.** 2004. Regulation of expression of cellulosomes and noncellulosomal (hemi)cellulolytic enzymes in *Clostridium cellulovorans* during growth on different carbon sources. J. Bacteriol. **186**:4218–4227.
 21. **Han SO, Yukawa H, Inui M, Doi RH.** 2005. Effect of carbon source on the cellulosomal subpopulations of *Clostridium cellulovorans*. Microbiology **151**:1491–1497.
 22. **Gold ND, Martin VJ.** 2007. Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. J. Bacteriol. **189**:6787–6795.
 23. **Raman B, Pan C, Hurst GB, Rodriguez M, Jr., McKeown CK, Lankford PK, Samatova NF, Mielenz JR.** 2009. Impact of pretreated Switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: a quantitative proteomic analysis. PLoS

ONE **4**:e5271.

24. **Tamaru Y, Miyake H, Kuroda K, Nakanishi A, Kawade Y, Yamamoto K, Uemura M, Fujita Y, Doi RH, Ueda M.** 2010. Genome sequence of the cellulosome-producing mesophilic organism *Clostridium cellulovorans* 743B. *J. Bacteriol.* **192**:901–902.
25. **Tamaru Y, Miyake H, Kuroda K, Nakanishi A, Matsushima C, Doi RH, Ueda M.** 2011. Comparison of the mesophilic cellulosome-producing *Clostridium cellulovorans* genome with other cellulosome-related clostridial genomes. *Microb. Biotechnol.* **4**:64–73.
26. **Buhrman DL, Price PI, Rudewiczcor PJ.** 1996. Quantitation of SR 27417 in human plasma using electrospray liquid chromatography-tandem mass spectrometry: A study of ion suppression. *J. Am. Soc. Mass Spectrom.* **7**:1099-1105.
27. **Miyamoto K, Hara T, Kobayashi H, Morisaka H, Tokuda D, Horie K, Koduki K, Makino S, Nunez O, Yang C, Kawabe T, Ikegami T, Takubo H, Ishihama Y, Tanaka N.** 2008. High-efficiency liquid chromatographic separation utilizing long monolithic silica capillary columns. *Anal. Chem.* **80**:8741–8750.
28. **Morisaka H, Kobayashi K, Kirino A, Furuno M, Minakuchi H, Nakanishi K, Ueda M.** 2009. Performance of wide-pore monolithic silica column in protein separation. *J. Sep. Sci.* **32**:2747–2751.
29. **Minakuchi H, Nakanishi K, Soga N, Ishizuka N, Tanaka N.** 1996. Octadecylsilylated porous silica rods as separation media for reversed-phase liquid chromatography. *Anal. Chem.* **68**:3498–3501.
30. **Brown SD, Guss AM, Karpinets TV, Parks JM, Smolin N, Yang S, Land ML, Klingeman DM, Bhandiwad A, Rodriguez M, Jr., Raman B, Shao X, Mielenz JR, Smith JC, Keller M, Lynd LR.** 2011. Mutant alcohol dehydrogenase leads to improved ethanol tolerance in *Clostridium thermocellum*. *Proc. Natl. Acad. Sci. USA* **108**:13752-13757.
31. **Ota M, Sakuragi H, Morisaka H, Kuroda K, Miyake H, Tamaru Y, Ueda M.** 2013. Display of *Clostridium cellulovorans* xylose isomerase on the cell surface of *Saccharomyces cerevisiae* and its direct application to xylose fermentation. *Biotechnol. Prog.* **29**:346-351.
32. **Yamada R, Taniguchi N, Tanaka T, Ogino C, Fukuda H, Kondo A.** 2010. Cocktail delta-integration: a novel method to construct cellulolytic enzyme expression ratio-optimized yeast strains. *Microb. Cell Fact.* **9**:32.

33. **Kuroda K, Matsui K, Higuchi S, Kotaka A, Sahara H, Hata Y, Ueda M.** 2009. Enhancement of display efficiency in yeast display system by vector engineering and gene disruption. *Appl. Microbiol. Biotechnol.* **82**:713-719.
34. **Inokuma K, Hasunuma T, Kondo A.** 2014. Efficient yeast cell-surface display of exo- and endo-cellulase using the SED1 anchoring region and its original promoter. *Biotechnol. Biofuels* **7**:8.
35. **Fan LH, Zhang ZJ, Yu XY, Xue YX, Tan TW.** 2012. Self-surface assembly of cellulosomes with two miniscaffoldins on *Saccharomyces cerevisiae* for cellulosic ethanol production. *Proc. Natl. Acad. Sci. USA* **109**:13260-13265.
36. **Kondo A, Ueda M.** 2004. Yeast cell-surface display – applications of molecular display. *Appl. Microbiol. Biotechnol.* **64**:28-40.
37. **Kuroda K, Ueda M.** 2011. Cell surface engineering of yeast for applications in white biotechnology. *Biotechnol. Lett.* **33**:1-9.
38. **Anonymous.** 1997. Arming yeast with cell-surface catalysts. *Chem. Eng. News.* **75**:32.
39. **Fujita Y, Ito J, Ueda M, Fukuda H, Kondo A.** 2004. Synergistic saccharification, and direct fermentation to ethanol, of amorphous cellulose by use of an engineered yeast strain codisplaying three types of cellulolytic enzyme. *Appl. Environ. Microbiol.* **70**:1207-1212.

Chapter I Exoproteome profiles of *Clostridium cellulovorans* on various carbon sources

The cellulosome, a protein complex produced by many cellulolytic Gram-positive anaerobic bacteria, such as *Clostridium*, efficiently degrades plant cell wall polysaccharides (1-3). The cellulosome was first found in *C. thermocellum* (4-6), and much progress has been made to understand the characteristic of the highly cellulolytic protein complex. However, the molecular mechanism of the formation of the cellulosome has not been characterized (7). Previously, we sequenced the entire genome of *C. cellulovorans* and identified all genes, including those that encode proteins of known and unknown functions, related to cellulosomal composition (8). Genome analysis of *C. cellulovorans* indicated the presence of 57 cellulosomal protein-encoding genes and 168 non-cellulosomal (hemi)cellulolytic protein-encoding genes. The non-cellulosomal (hemi)cellulolytic proteins (hereafter called non-cellulosomal proteins) have a N-terminal signal peptide but do not have dockerin domains. The 57 cellulosomal protein-encoding genes included 4 scaffold protein-encoding genes and 53 cellulosomal protein-encoding genes with dockerin domains. The major scaffold protein, CbpA, is comprised of 9 cohesin domains that bind to various cellulosomal proteins via cohesin-dockerin interactions (9). Interestingly, even though most of the cellulosomal proteins were glycoside hydrolases (GHs), proteins such as proteases, protease inhibitors, and unknown proteins were also included. These proteins seemingly not related to saccharification may be important in the degradation of various resources. However, we only have general information regarding the proteins that actively degrade biomass, although genome analysis has provided many interesting insights into the characteristics of *C. cellulovorans*. Additionally, *C. cellulovorans* has a larger number of non-cellulosomal proteins (10) that have polysaccharide degradation ability but not dockerin domains in their amino acid sequences compared to other cellulosome-producing clostridia.

In almost all previous studies of the cellulosome, cellulosomal proteins have been the focus (11, 12). However, it is important to note that *C. cellulovorans* must utilize non-cellulosomal proteins for polysaccharide degradation in the native condition (13). There have been several reports about non-cellulosomal proteins from either *C. thermocellum* or *C. cellulovorans* using gel-based

methods, but they were only limited to small fraction of the non-cellulosomal proteins (14, 15). Without understanding the contribution of non-cellulosomal proteins comprehensively, it has been impossible to determine the strategy by which cellulosome-producing bacteria degrade many kinds of substrates.

In this chapter, we performed quantitative proteomic analysis using tandem mass tag (TMT)-labeled samples on a mass spectrometer in order to characterize the quantitative changes in cellulosomal and non-cellulosomal proteins secreted by *C. cellulovorans* (16). Using a unique, long-monolithic silica capillary column (480 cm)-based liquid chromatography mass spectrometry (LC-MS/MS) approach, we successfully identified and quantified the exoproteome of *C. cellulovorans*, including both cellulosomal and non-cellulosomal proteins, without cellulosome isolation and pre-fractionation. Through this approach, we attempted to clarify how *C. cellulovorans* optimizes cellulosomal and non-cellulosomal proteins to degrade different substrates.

Materials and Methods

Growth substrates

Cellobiose (Sigma, MO, USA), xylan (Sigma), and pectin (Sigma) were used in the growth experiments. Phosphoric acid swollen cellulose (PASC) was prepared from a form of microcrystalline cellulose (Merck, Darmstadt, Germany) and was used in the growth experiments.

Cell culture and medium

C. cellulovorans 743B (ATCC 35296) was grown anaerobically as described previously (17), except for the carbon sources, which were 0.3% (w/v) cellobiose, 0.3% (w/v) xylan, 0.3% (w/v) pectin, or 0.3% (w/v) PASC (18).

Estimating the growth of the anaerobic bacteria

The growth curves of anaerobic *C. cellulovorans* on cellobiose, xylan, pectin, and PASC were determined by bacterial protein estimation, as described by Bensadoun and Weinstein (19). Five hundred microliters of cell culture was centrifuged for 10 min at $13,000 \times g$. The cell pellets were washed with 500 μ L of phosphate buffered saline (pH 7.4) and incubated with 400 μ L of

sodium deoxycholate (2%) for 20 min at 37°C. One hundred microliters of trichloroacetic acid (24%) was added to the suspension, which was centrifuged at $13,000 \times g$ for 10 min. Fifty microliters of resolubilization solution (5% SDS, 2N NaOH) was added to the suspension and vigorously mixed. The protein concentration was measured using a Protein Assay Bicinchoninate kit (Nacalai Tesque, Kyoto, Japan), with bovine serum albumin as the standard.

Preparation of extracellular proteins for quantitative proteome analysis

Proteome samples were prepared from *C. cellulovorans* culture media. The culture (50 mL) was centrifuged ($6,000 \times g$, 25°C), and the supernatant was subjected to ultrafiltration using Amicon Ultra YM-10 (Millipore, Bedford, MA) to obtain the extracellular proteins. The concentrated samples from each culture were dissolved in 100 μ L of triethylammonium hydrogen carbonate buffer (200 mM). Five microliters of tris(2-carboxyethyl)phosphine (200 mM) was added. The reaction was performed for 60 min at 55°C. Five microliters of iodoacetamide (375 mM) was then added, and the mixtures were reacted for 30 min, protected from light, at room temperature. Two microliters of freshly prepared sequence-grade modified trypsin (1 μ g/ μ L; Promega) was added, and the proteins were digested overnight at 37°C. The 4 proteome samples (cellobiose, xylan, pectin, and PASC) were labeled using a TMT (tandem mass tag) 6-plex labeling kit (Thermo Fisher Scientific, MA, USA) with reporters at $m/z = 126, 127, 128, \text{ and } 129$, respectively, in 41 μ L of CH_3CN . After 60 min of reaction at room temperature, 8 μ L of 5% (w/v) hydroxylamine was added to each tube and mixed for 15 min. In addition, a mixture of tryptic fragments from all substrates was reacted with TMT-131 as an internal standard for quantification. The aliquots were then combined, and the pooled sample was evaporated under vacuum. The samples were then dissolved in 100 μ L of trifluoroacetic acid (0.1%) before LC-MS/MS analysis.

LC-MS/MS analysis

Proteome analyses were performed using a liquid chromatography mass spectrometry system equipped with a long monolithic column, as described in a previous study (13). A monolithic silica capillary column was prepared from a mixture of tetramethoxysilane and methyltrimethoxysilane (20). Tryptic digests were separated by reversed-phase chromatography using a monolithic silica capillary column (480 cm long, 0.1 mm ID) at a flow rate of 500 nL/min.

The gradient was provided by changing the mixing ratio of the 2 eluents: A, 0.1% (v/v) formic acid and B, 80% acetonitrile containing 0.1% (v/v) formic acid. The gradient was started with 5% B, increased to 45% B for 600 min, further increased to 95% B to wash the column, returned to the initial condition, and held for re-equilibration. The separated analytes were detected on a mass spectrometer with a full scan range of 350–1500 m/z (resolution 60000), followed by 10 data-dependent HCD MS/MS scans acquired for TMT reporter ions. A normalized collision energy of 80% in HCD (higher-energy c-trap dissociation) with a 0.1 ms activation time was used. An electrospray ionization (ESI) voltage of 2.4 kV was applied directly to the LC buffer distal to the chromatography column using a microtee. The ion transfer tube temperature on the LTQ Velos ion trap was set to 300°C. Duplicate analyses were performed for each sample in 3 independent experiments. Blank runs were inserted between different samples.

Data analysis

The mass spectrometry data files were analyzed for protein identification and quantification using a WF_LTQ_Orbitrap_Mascot_HCD_ReporterQuantitation template in Protein Discoverer software (Thermo Fisher Scientific). Protein identification was performed using the Mascot algorithm against the *C. cellulovorans* protein database (4254 sequences) from NCBI (<http://www.ncbi.nlm.nih.gov/>) with a precursor mass tolerance of 20 ppm and a fragment ion mass tolerance of 20 mmu. Carbamidomethylation of cysteine and the TMT 6-plex at the N-terminus were set as fixed modifications. Protein quantification was performed using the Reporter Ions Quantifier with the TMT 6-plex method. The data were then filtered with a cut-off criteria of a q-value ≤ 0.05 , corresponding to a 5% FDR (false discovery rate) on a spectral level. The detected spectral counts of the identified peptides were used to estimate the abundance of cellulosomal and non-cellulosomal proteins. Proteins with no missing values in 3 replicates were accepted in the protein quantitation analysis. Global median normalization was carried out to normalize the amount of tryptic digests injected into the mass spectrometer. The heat map was constructed using Cluster 3.0 (21), which can perform hierarchical cluster analysis. Euclidean distance was used to measure the similarities of the protein profile patterns within the clustering analysis. To visualize the clustering results from Cluster 3.0, Java TreeView (22) software was used.

Results

In order to elucidate that quantitative changes in cellulosomal and non-cellulosomal proteins in the exoproteome of *C. cellulovorans* on different growth substrates, *C. cellulovorans* was cultured under anaerobic conditions with cellobiose, xylan, pectin, or PASC as the sole carbon source. Proteome samples were prepared from the supernatants without cellulosome purification and pre-fractionation. Using a quantitative proteomic approach based on an isobaric tag method, we obtained the quantitative proteomic profiles of 4 kinds of *C. cellulovorans* exoproteomes, as illustrated in Fig. 1.

Growth confirmation

To determine the growth of *C. cellulovorans* with the 4 different substrates, we performed bacterial protein estimation (23) (Fig. 2). *C. cellulovorans* can grow on pectin and xylan as a sole carbon source, whereas other cellulosome-producing bacteria, such as *Clostridium thermocellum*, cannot (17). PASC was used as a cellulose substrate instead of Avicel for faster growth of *C. cellulovorans* and higher amount of produced proteins, comparable to when the other substrates were used, i.e. cellobiose, xylan, and pectin, which leads to more convenient and accurate quantitative analysis. Moreover, even though it is not crystalline cellulose, we can see the difference between when the polymer and the disaccharide were used as the substrates. Although there was an initial delay of growth in pectin and PASC cultures compared with cellobiose and xylan cultures, 12-day supernatants from cellobiose, xylan, pectin, and PASC cultures had comparable amounts of protein. Hence, we determined that *C. cellulovorans* cultured for 12 days was in stationary phase and that we could obtain the appropriate amounts of protein for proteome analysis at this time point. We performed quantitative proteome analysis of *C. cellulovorans* supernatant cultured with 4 kinds of substrates for 12 days.

Quantitative proteome analysis of C. cellulovorans culture supernatants

We used TMT (tandem mass tag, 6-plex) isobaric tags and nano LC-MS/MS with a long monolithic silica capillary column (480 cm). Each TMT-labeled peptide sample was mixed at a specific ratio such that the mixed samples were estimated to contain the same amount of

TMT-labeled peptides from each sample. The peptide concentrations of each sample were measured using the bicinchoninic acid assay (BCA). Principal component analysis (PCA) was performed to examine the similarity of the protein production profiles between 3 biological replicates of culture supernatant for each substrate. A score plot of PCA showed high similarity between each of the biological replicates, and the plots of each substrate formed individual groups, indicating that the protein profiles of the exoproteomes reflected the difference in substrate (Fig. 3).

C. cellulovorans has 4254 protein-coding genes in its genome. The mass spectrometry data were used for protein identification with the protein database built from the genome analysis of *C. cellulovorans*. In total, with the cut-off criteria, we identified 639 proteins in each sample from 3 biological replicates. All identified proteins are summarized in Table S1 in Supplemental file 1 of the paper “Exoproteome profiles of *Clostridium cellulovorans* on various carbon sources” (<http://aem.asm.org/content/suppl/2013/10/01/AEM.02137-13.DCSupplemental/zam999104808so1.pdf>). The quantitative data were combined and standardized using a median approach to normalize the amount of tryptic digests injected into the mass spectrometer. We then assessed the variance between all biological replicates. Scatter plots of all comparisons were created using the data from the combination of the biological replicates (Fig. 4).

Quantitative exoproteome analysis focused on saccharification

The work flow of the data analysis and a summary of the data are shown in Fig. 5. From the 639 identified proteins, cellulosomal and non-cellulosomal proteins (Table 1) involved in saccharification for further study were selected and the number of the proteins were 79. Of these, 35 were cellulosomal proteins and 44 were non-cellulosomal proteins. A previous genomic study identified 57 and 168 genes encoding cellulosomal and non-cellulosomal proteins, respectively. Thus, our results indicate that 61.4% (35 of 57) of cellulosomal proteins and 26.1% (44 of 168) of non-cellulosomal proteins were produced under the conditions analyzed in this study.

Using the standardized relative quantitative data of 79 cellulosomal and non-cellulosomal proteins, hierarchical cluster analysis was performed to generate a heatmap and to observe the trends of extracellular proteins (Fig. 6). In the hierarchical cluster analysis, the 3 biological replicates clearly grouped into small clusters. Within each cluster, generally, many xylanases from the xylan samples, many pectate lyases from the pectin samples, and many kinds of glucanases from the

cellobiose and PASC samples were most commonly observed. These results indicate that *C. cellulovorans* recognized the extracellular environment (in this case, substrate) and secreted the optimal enzymes necessary for each substrate (24).

Table 1 The list of selected 79 proteins including 35 cellulosomal proteins and 44 non-cellulosomal proteins

	Locus	Name ^a	Modular content ^b	ΣSpC ^c
cellulosomal proteins	Clocel_2741	EngZ	GH9,CBM3	96
	Clocel_1432		GH5	19
	Clocel_2600		GH5,CBM32	84
	Clocel_3650		GH44	12
	Clocel_0930		GH9,CBM3	49
	Clocel_2576		GH9,CBM3	96
	Clocel_0983		GH5	17
	Clocel_3111		GH5	202
	Clocel_0619		GH5	146
	Clocel_1150	EngB	GH5	7
	Clocel_3359	EngE	GH5	541
	Clocel_2822	EngH	GH9,CBM3	251
	Clocel_2821	EngK	CBM4,GH9	854
	Clocel_2819	EngL	GH9	180
	Clocel_2816	EngM	CBM4,GH9	157
	Clocel_1624	EngY	CBM30,GH9	114
	Clocel_2818	ManA	NA	268
	Clocel_4119		CBM35,GH26,CBM35	294
	Clocel_2607	ManB	CBM35,GH26	242
	Clocel_2575		CBM35,GH26	59
	Clocel_2295	XynA	GH11,CE4	94
	Clocel_2900	XynB	CBM22,GH10	420
	Clocel_2823	ExgS	GH48	827

	Clocel_1623	PelA	PL1,PL9	259
	Clocel_0148		NA	26
	Clocel_1010		PL11	22
	Clocel_0147		NA	120
	Clocel_0145		NA	83
	Clocel_0146		NA	244
	Clocel_0982		NA	60
	Clocel_3193		NA	153
	Clocel_3112		NA	186
	Clocel_2240		NA	75
	Clocel_2824	CbpA	CBM3	1593
	Clocel_2820	HbpA	NA	87
non-cellulosomal proteins	Clocel_3242	EngD	GH5,CBM2	51
	Clocel_3662	EngF	GH5,CBM17	95
	Clocel_2815	EngN	GH9	30
	Clocel_1478	EngO	CBM4,GH9	163
	Clocel_2020		GH51	93
	Clocel_0912		GH5	169
	Clocel_3197		GH130	77
	Clocel_3196		GH130	142
	Clocel_1011		PL1,CBM13	9
	Clocel_1430		GH31	33
	Clocel_1455		GH53,CBM61	40
	Clocel_4124		GH26,CBM59	9
	Clocel_1134		GH26,CBM23	164
	Clocel_2535		GH43	55
	Clocel_0034		GH31	65
	Clocel_3657		GH43	54
	Clocel_3489		GH2	13
	Clocel_2507		GH42	34

Clocel_3191	GH53	20
Clocel_2595	GH43	30
Clocel_1420	CBM30,GH9	66
Clocel_2536	GH51,GH43	11
Clocel_1243	NA	54
Clocel_1254	NA	318
Clocel_0733	PL9	36
Clocel_3834	PL10	165
Clocel_0873	PL9	69
Clocel_3380	PL9	36
Clocel_3113	PL10	22
Clocel_0724	GT1	156
Clocel_4071	NA	124
Clocel_2256	GH105	50
Clocel_0211	CE8	15
Clocel_0032	GH94	188
Clocel_0041	GH42	60
Clocel_0391	GH94	266
Clocel_0560	GH36	14
Clocel_1172	PL1	76
Clocel_1342	CBM4	18
Clocel_1369	GH23	26
Clocel_2606	GH5,CBM46	488
Clocel_2882	GH1	268
Clocel_4053	CBM27	7
Clocel_4083	GH2	45

^a Name: Names of only the reported proteins were shown.

^b Modular content: They are determined based on CAZy database.

GH: glycoside hydrolases, PL: polysaccharide lyase, CBM: carbohydrate binding module, NA: not annotated (not included in CAZy database)

^c Σ SpC: sum of spectral count using data analysis for each protein

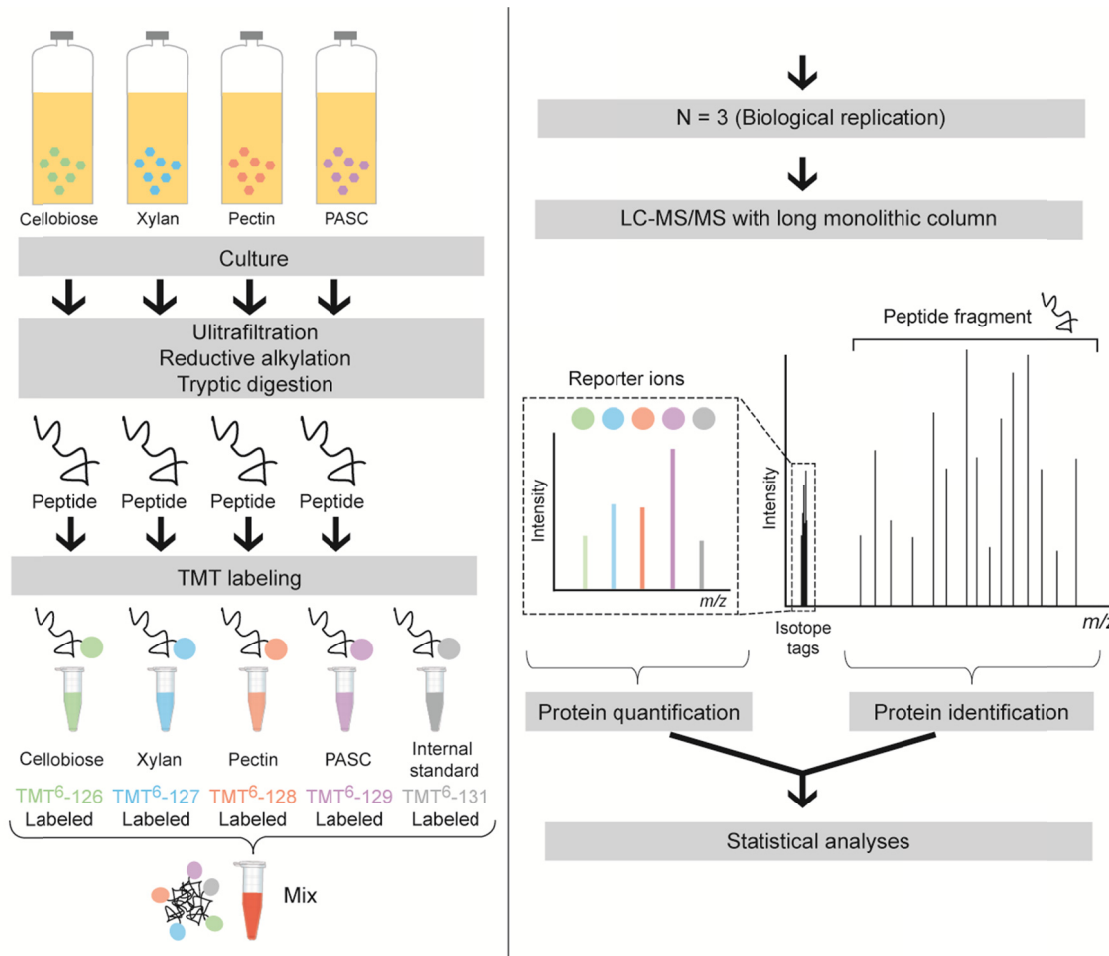


Fig. 1 Schematic representation of the workflow to quantify proteins in the exoproteome

The proteins in culture supernatants of *C. cellulovorans* grown with cellobiose, xylan, pectin, or PASC as a sole carbon source were individually digested with trypsin, and each peptide sample was labeled with Tandem Mass Tags (TMT) with various reporters at $m/z = 126$, 127 , 128 , and 129 , respectively. The mixture of labeled samples was applied to the LC-MS/MS with a long monolithic column to analyze the proteins secreted on each growth substrate. Statistical analyses of the proteome data were performed to elucidate the quantitative changes in the exoproteome profile on 4 different growth substrates.

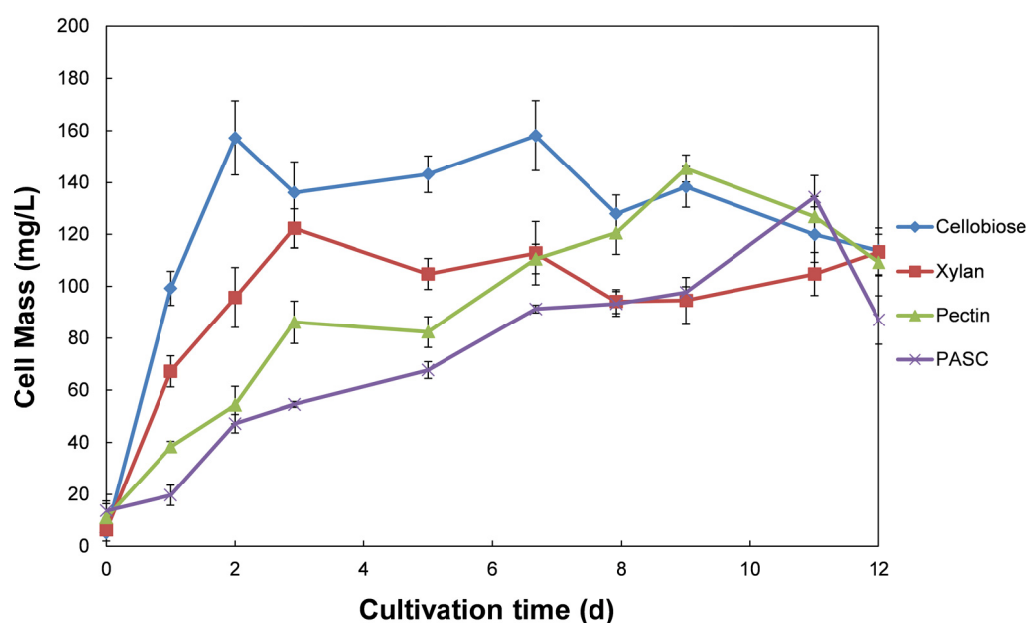


Fig. 2 Growth confirmation of *C. cellulovorans* cultured with 4 different growth substrates

The growth curve was determined by measuring the amount of protein inside the cells. Cells were in stationary phase in all samples, and cells cultured for 12 days were used for the proteome analysis.

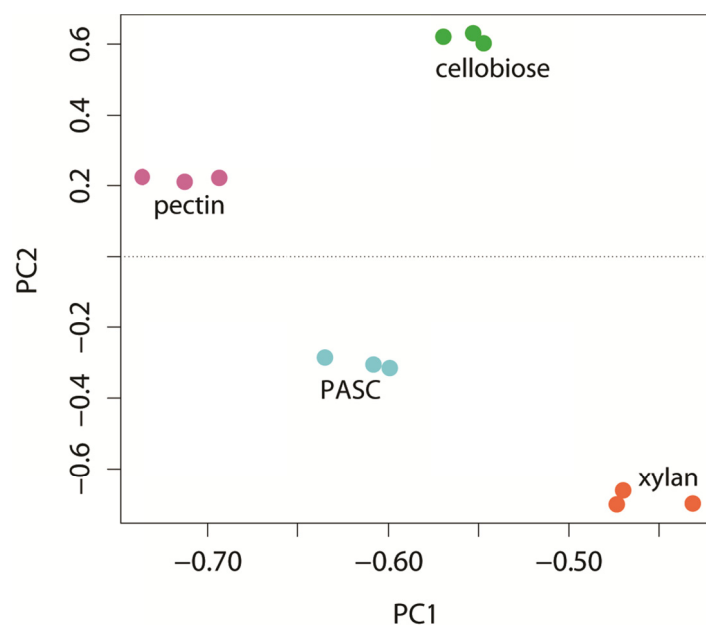


Fig. 3 Principal component analysis of the data from the 3 biological replicates used in this study

Proteome data for each substrate clustered in close proximity. The cumulative contribution rate for principal component PC1 to PC2 was 60%. PASC: phosphoric acid-swollen cellulose.

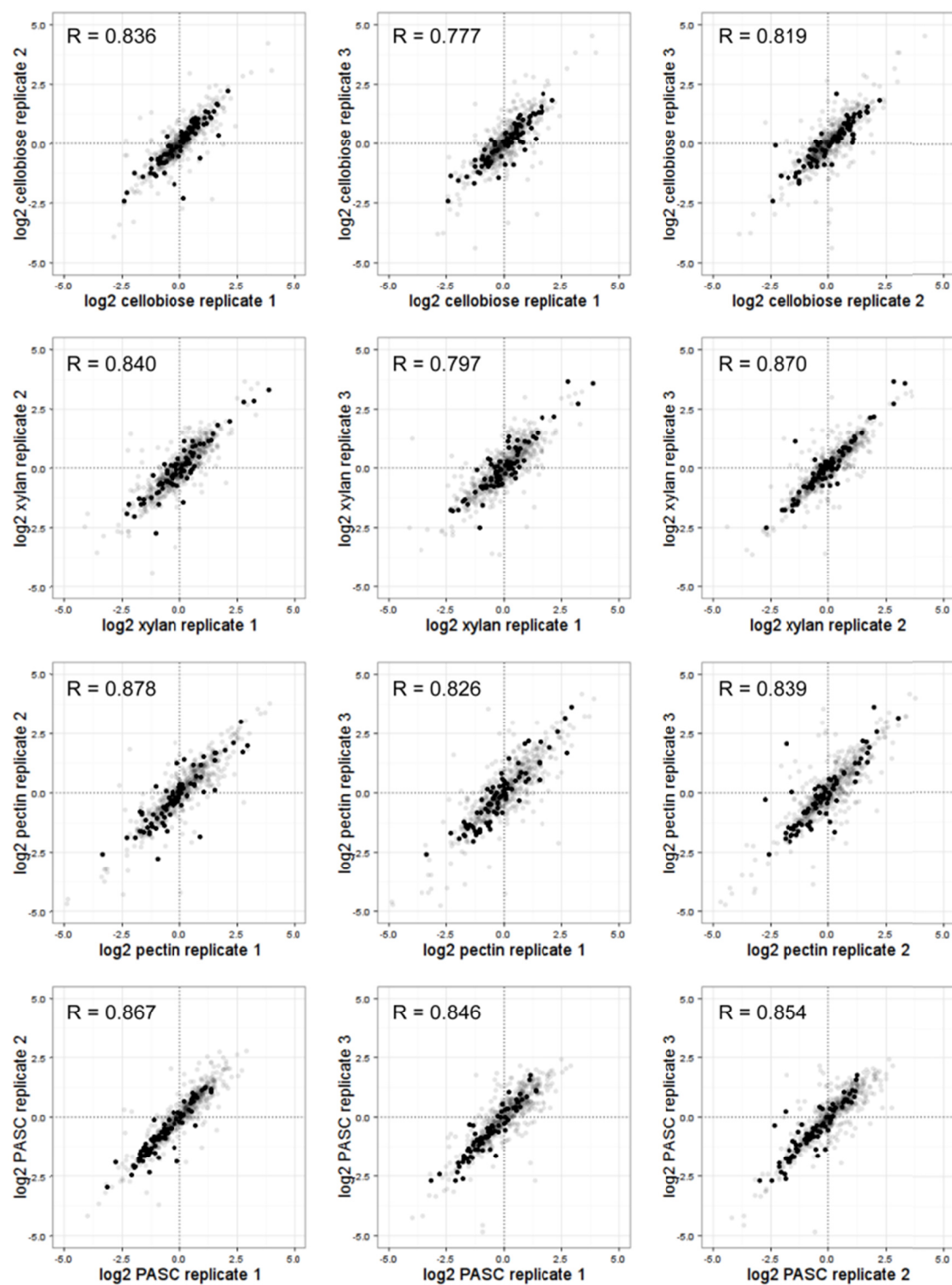


Fig. 4 Scatter plots of the 3 biological replicates for each substrate

Fold change values between the biological replicates of each substrate were compared.

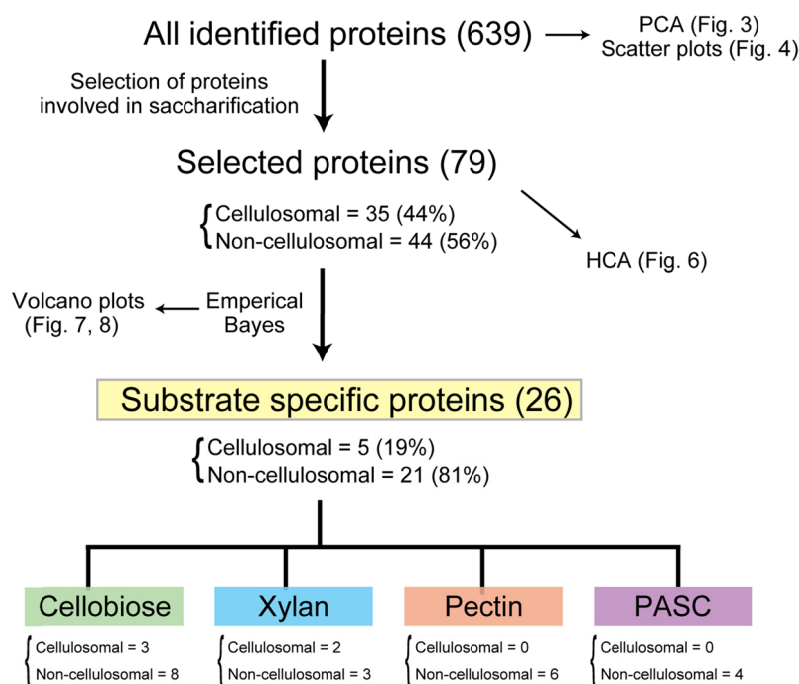


Fig. 5 Work flow and summary of the data analysis of the *C. cellulovorans* exoproteome

The total number of proteins identified and the number of selected proteins involved in saccharification are shown. The numbers of cellulosomal and non-cellulosomal proteins in the total data set and among the substrate-specific extracellular proteins are also shown. The number of 'substrate-specific extracellular proteins' was 26 and 81% of them were non-cellulosomal proteins. Among the 57 cellulosomal genes that has been reported (10), 35 cellulosomal proteins were produced in all of the samples in this study. The numbers of 'pectin-' and 'PASC'-specific cellulosomal proteins were 0, which means that 35 cellulosomal proteins were produced but substrate specific cellulosomal proteins with significant fold change among them were not observed.

Table 2 Protein abundance index of exoproteome

	# protein ^a	# SSP ^b	ΣSpC ^c	ΣSpC / # protein
cellulosomal	35	5	7937	227
non-cellulosomal	44	21	3924	89

^a # protein: the number of selected proteins involved in saccharification among 639 identified proteins

^b # SSP: the number of substrate specific proteins

^c ΣSpC: sum of spectral count using data analysis for each protein

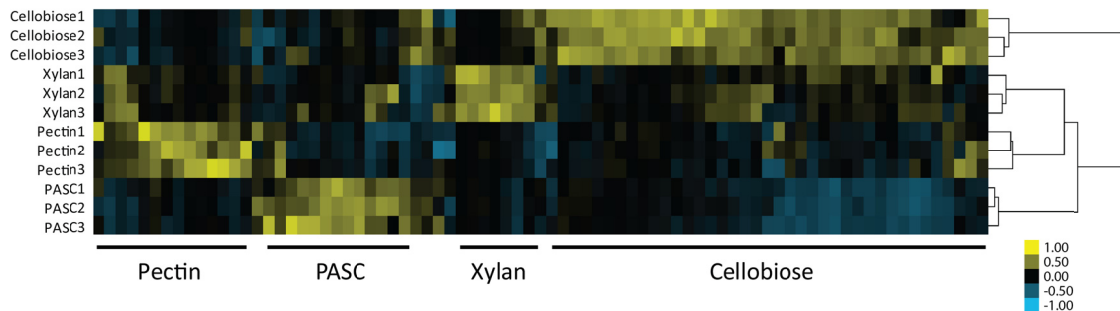


Fig. 6 Exoproteome profiles of *C. cellulovorans* cultivated with 4 different growth substrates shown in a heat map after hierarchical cluster analysis

The quantitative proteome data from 3 biological replicates of each substrate sample, 12 samples in total, were used for the hierarchical cluster analysis. The tree structure represents the similarity between the data of each of the 12 samples and shows that the biological replicates grouped together, indicating that their exoproteome profiles were similar and the profiles were reproducible. Distinctive differences in the exoproteome profiles depending on the growth substrate were confirmed, and most of the proteins produced in response to a given growth substrate were proteins able to degrade that substrate.

Abundance of the cellulosomal and non-cellulosomal proteins

Among the selected 79 cellulosomal and non-cellulosomal proteins, the spectral counts of each protein was used to compare the abundance of the cellulosomal and non-cellulosomal proteins. The average of the spectral counts of the cellulosomal proteins was much larger than that of the non-cellulosomal proteins (Table 2). In addition, a scaffolding protein of the cellulosome (Clocel_2824) showed the largest counts, and basic cellulosomal proteins (Clocel_2823 and Clocel_3359), reported to be constitutively produced (13), and another cellulosomal protein produced in xylan culture (Clocel_2821) also showed high counts (third and fourth, and second respectively) (Table 1). The results indicate that the cellulosomal proteins were more highly produced than the non-cellulosomal proteins.

Substrate-specific extracellular proteins

To determine the proteins that showed a significant fold change between different growth

substrates, an empirical Bayes moderated t-test was performed. *P*-values were adjusted with the Benjamini-Hochberg method to avoid the problem of multiple testing. Volcano plots were generated to visualize extracellular proteins present in different amount. The criteria that we adopted for the extracellular proteins present in different amount were an FDR-adjusted *P*-value < 0.01 and a fold change of protein ratio > 2. Extracellular proteins defined as present in different amount were identified in all comparisons, and they are shown as blue dots in the volcano plots (Fig. 7).

Pectin/cellobiose and xylan/cellobiose volcano plots are shown in Fig. 8, as examples of the analysis. We identified proteins differentially produced in pectin media compared to cellobiose media. Proteins classified as polysaccharide lyases (PLs) by CAZy (<http://www.cazy.org/>) are putative pectin degrading proteins. The PLs among the proteins differentially produced in pectin media are indicated in pink font with an underline. In total, 9 PLs were identified in the group of 79 selected proteins (Table 1), and 7 of these PLs were differentially produced in pectin media compared to cellobiose media (Fig. 8a). Among these proteins is Cloce1_1623 (PelA), a well-known cellulosomal pectate lyase that was classified as a member of PL1 and PL9 families; it is reported to degrade polygalacturonic acid into digalacturonic acid or trigalacturonic acid (25). Moreover, 4 of the 7 PLs were also highly produced in xylan media (Fig. 8b), possibly because xylan and pectin have similar structures, including pentose sugar moieties.

We defined the proteins that were present in different amount on one specific substrate compared to all of the 3 other substrates as "substrate-specific extracellular proteins" (i.e., cellobiose-specific proteins, xylan-specific proteins, pectin-specific proteins, and PASC-specific proteins). The list of these substrate-specific extracellular proteins is shown in Table 3, and the details are discussed in the following sections. The number of substrate-specific extracellular proteins was 26 (Fig. 5), and other 53 proteins were basal proteins that did not show significant fold change on different growth substrates.

Substrate-specific extracellular proteins for cellobiose

Eleven cellobiose-specific proteins were identified; 3 were cellulosomal proteins and 8 were non-cellulosomal proteins (Table 3). Five of the 11 cellobiose-specific proteins were endoglucanases: Cloce1_1624 (EngY) (26), Cloce1_2815 (EngN) (27), Cloce1_1478 (EngO) (28), Cloce1_3242 (EngD) (29), and Cloce1_3662 (EngF) (30). These proteins have been previously

reported; however, they are not enzymes known to be involved in cellobiose degradation. Therefore, cellobiose might be a trigger for the detection of extracellular cellulose. Among the cellobiose-specific proteins, Clocel_3662 (EngF) showed the highest fold change (Fig. 8 and Table 3).

Table 3 Substrate specific extracellular proteins

		Locus	Name	CAZy family ^a	vs cellobiose		vs xylan		vs pectin		vs PASC	
					log2-fold change	FDR-adjusted P-value ^b	log2-fold change	FDR-adjusted P-value ^b	log2-fold change	FDR-adjusted P-value ^b	log2-fold change	FDR-adjusted P-value ^b
Cellobiose specific	cellulosomal proteins	Clocel_1624	EngY	CBM30, GH9			1.82	2.54E-07	1.65	4.94E-07	2.28	3.57E-08
		Clocel_2575		CBM35, GH26			1.13	7.86E-05	2.09	4.94E-07	2.11	3.72E-07
		Clocel_3112		NA			2.1	1.14E-08	2.25	1.11E-08	2.71	3.30E-09
	non-cellulosomal proteins	Clocel_1011		PL1, CBM13			1.72	5.38E-06	1.27	4.54E-05	1.53	9.41E-06
		Clocel_1420		CBM30, GH9			1.46	2.91E-04	2.07	1.46E-05	2.23	6.88E-06
		Clocel_1478	EngO	CBM4, GH9			1.31	1.86E-05	2.11	4.32E-07	1.75	1.43E-06
		Clocel_2815	EngN	NA			2.56	7.13E-08	1.57	3.65E-06	2.09	3.17E-07
		Clocel_3196		GH130			2.24	3.59E-06	2.03	6.96E-06	1.38	9.36E-05
		Clocel_3197		GH130			1.68	6.86E-07	1.46	2.01E-06	1.03	1.93E-05
		Clocel_3242	EngD	GH5, CBM2			1.52	5.90E-07	1.39	1.04E-06	1.82	9.82E-08
		Clocel_3662	EngF	GH5, CBM17			4.13	2.54E-06	2.47	1.04E-04	3.6	5.15E-06
Xylan specific	cellulosomal proteins	Clocel_1432		GH5	2.72	3.94E-05			2.7	6.34E-05	2.99	1.47E-05
		Clocel_2295	XynA	GH11, CE4	4.5	5.49E-08			4.38	1.20E-07	4.51	5.41E-08
	non-cellulosomal proteins	Clocel_0912		GH5	4.16	2.10E-09			3.77	1.16E-08	2.89	5.02E-08
		Clocel_1430		GH31	1.3	7.07E-06			1.98	4.23E-07	2.44	5.02E-08
		Clocel_2595		GH43	5.5	4.91E-09			5.08	1.67E-08	4.62	4.41E-08
Pectin specific	non-cellulosomal proteins	Clocel_0873		PL9	2.32	3.64E-07	1.49	1.92E-05			3.91	6.07E-09
		Clocel_1172		PL1	2.93	3.65E-06	2.01	1.08E-04			3.79	4.21E-07
		Clocel_1243		NA	2.13	1.14E-07	1.64	1.05E-06			1.03	3.79E-05
		Clocel_2256		GH105	3.23	1.46E-06	3.47	1.24E-06			3.92	3.22E-07
		Clocel_2507		GH42	3.19	1.75E-05	2.6	1.43E-04			2.64	1.23E-04
		Clocel_3380		PL9	3.24	1.14E-07	2.76	4.23E-07			3.62	3.44E-08
PASC specific	non-cellulosomal proteins	Clocel_0032		GH94	1.02	1.65E-05	1.15	8.02E-06	1.48	1.22E-06		
		Clocel_2606		GH5, CBM46	2.11	1.68E-08	1.55	1.47E-07	2.33	6.07E-09		
		Clocel_3657		GH43	1.92	1.69E-05	1.03	2.08E-03	1.67	1.13E-04		
		Clocel_4124		GH26, CBM59	1.56	6.88E-06	2.36	2.57E-07	2.54	2.39E-07		

^a Refer to <http://www.cazy.org/>

^b *P*-values were adjusted for multiple testing with the Benjamini-Hochberg method.

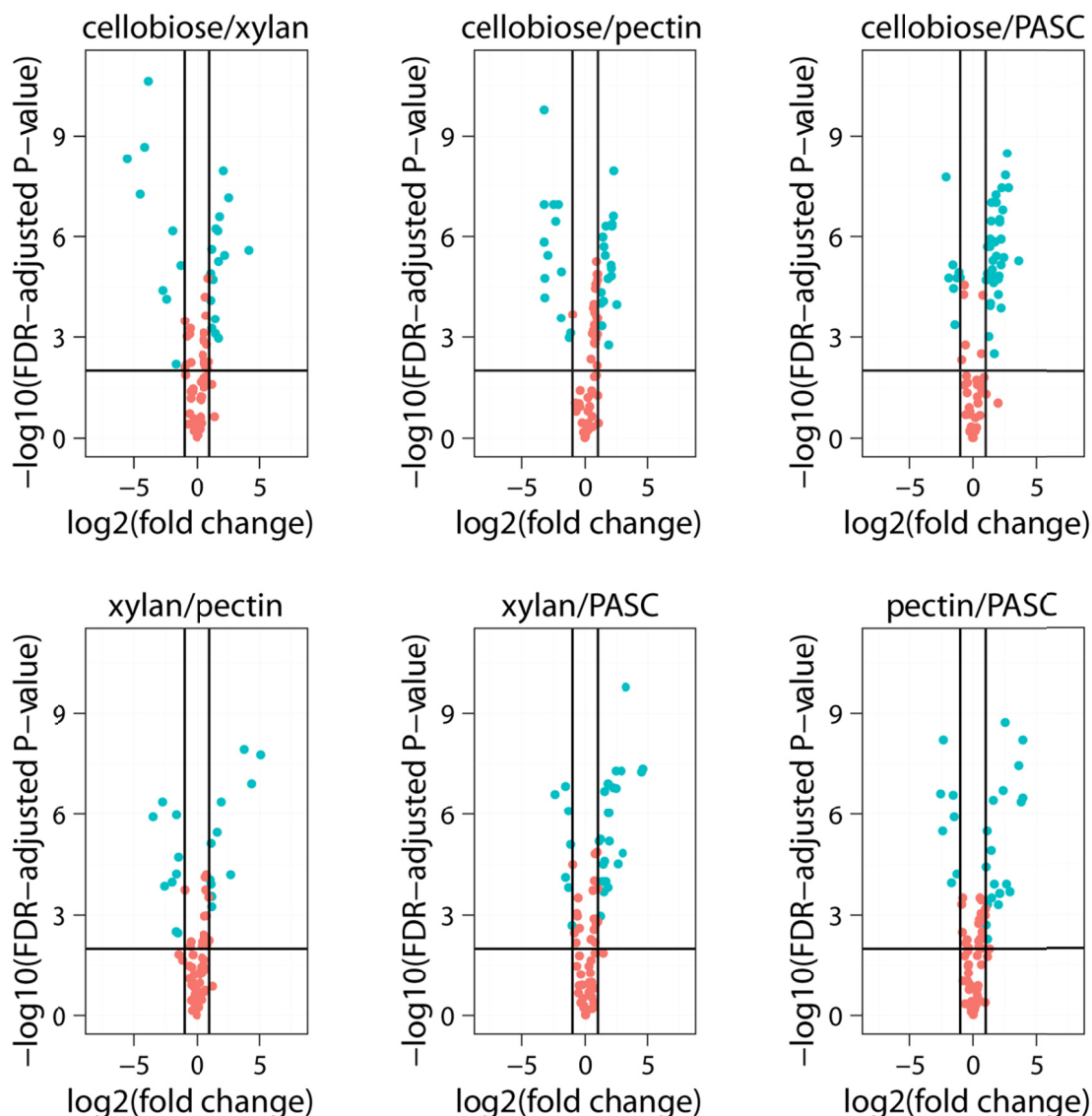


Fig. 7 Volcano plots show proteins differentially secreted on two different substrates

The differentially secreted proteins were defined as the proteins secreted on a specific substrate that showed an FDR adjusted P-value < 0.01 and a fold change of protein ratio > 2 when compared to secretion on a different substrate; they are shown as blue dots.

Substrate-specific extracellular proteins for xylan

Xylan-specific proteins comprised 2 cellosomal proteins and 3 non-cellosomal proteins (Table 3). Amongst the identified cellosomal proteins, Clocel_1432 was classified as a member of the GH5 family by CAZy. Generally, GH5 enzymes have varying substrate specificities. Because endo- β -1,4-xylanase activity has been reported for GH5 enzymes, this protein was produced in xylan

could have endo- β -1,4-xylanase activity. In the previous study, Clocel_2295 (XynA) was classified as a member of the GH11 and carbohydrate esterase (CE)4 families, and endo-xylanase activity and deacetylase activity were confirmed (31). The primary degradation products from xylan by this enzyme were xylobiose and xylotriose. Among the non-cellulosomal proteins, Clocel_0912, classified as a GH5 enzyme, is thought to have endo- β -1,4-xylanase activity. Clocel_2595, classified as a GH43 enzyme, is thought to have β -xylosidase activity and produce xylose by degrading xylan. Clocel_1430, classified as a GH31 enzyme, is thought to have α -xylosidase activity. In short, *C. cellulovorans* can first degrade xylan into xylobiose or xylotriose through the activity of cellulosomal XynA or other endo- β -1,4-xylanases and then degrade xylobiose to xylose through the activity of Clocel_2595 and Clocel_1430. The supernatant from cultures with xylan as a substrate contained xylose, but xylobiose or xylotriose were not detected (data not shown), consistent with their conversion to xylose.

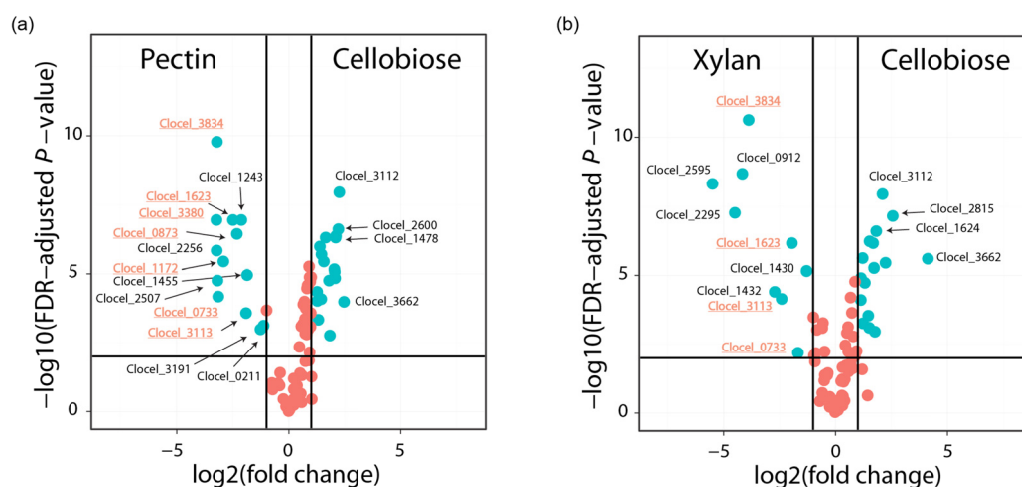


Fig. 8 Two representative volcano plots, (a) pectin/cellobiose and (b) xylan/cellobiose, show extracellular proteins present in different amount

The extracellular proteins present in different amount were defined as the proteins secreted on a specific substrate that showed an FDR adjusted P -value < 0.01 and a fold change of protein ratio > 2 when compared to secretion on a different substrate; they are shown as blue dots. The names in pink font with underlines are known or putative pectin degrading enzymes which are categorized into polysaccharide lyase (PL) family by CAZy.

Substrate-specific extracellular proteins for pectin

Pectin-specific proteins comprised 6 non-cellulosomal proteins (Table 3). Clocel_2507, classified as a GH42 enzyme, is thought to have β -galactosidase activity. Clocel_0873 and Clocel_3380 were classified as PL9 enzymes that shows activity to degrade pectate polymer into disaccharide. Clocel_2256, classified as a GH105 enzyme, is thought to have unsaturated rhamnogalacturonyl hydrolase activity. Clocel_1172 was classified as a PL1. In summary, *C. cellulovorans* can first degrade pectate into pectate disaccharide through the activity of pectate lyases and then degrade pectate disaccharide into monomers through the activity of Clocel_2507, which is similar to xylan degradation.

Substrate-specific extracellular proteins for PASC

PASC-specific proteins comprised 4 non-cellulosomal proteins (Table 3). Clocel_4124 was especially enriched, and a domain of Clocel_4124 was classified as a CBM59 domain, which was reported to have the ability to bind to cellulose. Clocel_2606, classified as a GH5 and a CBM46 enzyme, is thought to have cellulose-binding ability and be able to degrade cellulose. Clocel_0032, classified as a GH94, is expected to have cellodextrin phosphorylase activity.

Discussion

In this study, we described a strategy by which *C. cellulovorans* degrades several different substrates via the optimization of cellulosomal and non-cellulosomal proteins in the line of cellulosomic analysis, and we identified the importance of non-cellulosomal proteins for the optimal degradation of various biomasses.

We focused on 79 cellulosomal and non-cellulosomal proteins, out of the 639 extracellular proteins identified. The 79 proteins could be divided into two types: basal proteins that were produced rather abundantly and consistently irrespective of the type of growth substrate, and substrate-specific extracellular proteins that were differentially produced proteins in one specific substrate. First, we considered the abundance of cellulosomal and non-cellulosomal proteins in the exoproteome. Although it is necessary to perform an exact quantification using antibody and a synthetic compound to precisely compare the amounts of each protein, the spectral counts, shown in

Table 2 and S2, correlate with the abundance (32). The comparison of spectral counts of cellulosomal and non-cellulosomal proteins (Table 2) showed that the cellulosomal proteins were more abundant and, therefore, the highly abundant cellulosomal proteins might play a major role in the degradation of biomass.

The analysis of substrate-specific extracellular proteins showed that 81% (21 proteins) of 26 substrate-specific extracellular proteins were non-cellulosomal proteins (Fig. 5 and Table 3). *C. cellulovorans* has a larger number of non-cellulosomal genes than the other cellulosome-forming clostridia (10), judging from the presence of dockerin domains and N-terminal signal peptides. We confirmed that non-cellulosomal proteins were not only encoded in the genome but were actually produced at the protein level. In previous reports, quantitative changes in cellulosomal proteins were generally thought to be important, and the cellulosome was purified in most other proteomic studies of the cellulosome (23, 33). However, our results indicate the possibility that most of the cellulosomal proteins including Clocel_2824, Clocel_2823, and Clocel_3359 were rather basal proteins and non-cellulosomal proteins were produced with drastic change for the optimized degradation of various biomasses as substrate-specific extracellular proteins, which provides new insight into the importance of non-cellulosomal proteins. Interestingly, in recent study of *C. cellulolyticum* (34), transcriptomic analysis revealed that an "accessory" set of CAZymes specifically expressed for each of the non-cellulose substrates consisted mostly of non-cellulosomal protein-encoding genes, which supports the results in this study.

There has been a study that demonstrated the synergistic activity of the cellulosomes and non-cellulosomal proteins (14) by comparing the sum of each activity of cellulosomes and non-cellulosomal proteins and the measured activity of mixture of cellulosomes and non-cellulosomal proteins. Furthermore, Ding *et al.* (35) compared the mechanisms by which non-complexed fungal secreted cellulases and multienzyme complexes (cellulosomes) deconstruct plant cell walls. Whereas cellulosomes peeled off individual cellulose microfibrils from the cell wall surface, the fungal secreted cellulases penetrated inside the cellulose microfibril network and dissolved the entire wall in a uniform manner, which resulted in faster degradation of cellulose compared to that by cellulosomes. Not only the proximity effect of cellulosome but also the penetrating effect of non-complexed secreted cellulases was important, which supports the significance of non-cellulosomal proteins. Non-cellulosomal proteins are non-complexed proteins

like fungal secreted cellulases (Fig. 9). The penetrating effects of non-cellulosomal proteins could explain why non-cellulosomal proteins accounted for such a large fraction of the substrate-specific extracellular proteins and how cellulosomes and non-cellulosomal proteins show synergistic activity.

In conclusion, this study has quantitatively characterized the exoproteome profiles of *C. cellulovorans* using LC-MS/MS with a long monolithic column. This study reports the whole exoproteome of *C. cellulovorans*, including cellulosomal and non-cellulosomal proteins, from cell culture supernatant without a purification and pre-fractionation process. We selected 79 cellulosomal and non-cellulosomal proteins out of 639 identified proteins and found that the abundance of cellulosomal proteins was higher than that of non-cellulosomal proteins, based on spectral counts of the proteins. Subsequently, we further identified 26 substrate-specific extracellular proteins (Fig. 5). The substrate-specific extracellular proteins for each of the 4 substrates consisted of proteins expected to degrade the particular substrate, e.g., xylanases for xylan and pectin lyase for pectin. Surprisingly, among the substrate-specific extracellular proteins, the number of non-cellulosomal proteins identified was almost 4 times the number of cellulosomal proteins. The results indicate that cellulosomal proteins, produced in large amounts, play a major role in substrate degradation, and that non-cellulosomal proteins, rather specifically produced, optimize the extracellular complement of the secretome for the efficient degradation of various components of a biomass. We propose that non-cellulosomal proteins penetrate the surface of a biomass and dissolve the entire cell wall in a uniform manner and that cellulosomes vigorously degrade a biomass from its surface; non-cellulosomal proteins and cellulosomal proteins could thus work synergistically (Fig. 9).

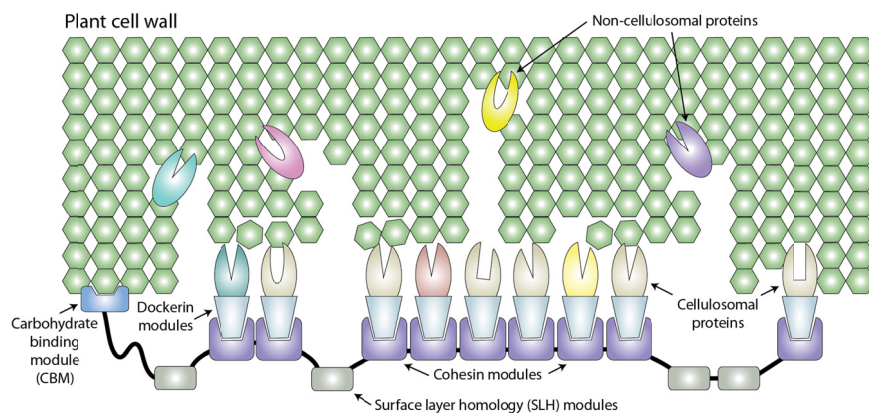


Fig. 9 A schematic model of the biomass degradation process by cellulosomal and non-cellulosomal proteins secreted by *C. cellulovorans*

Summary

The cellulosome is a complex of cellulosomal proteins bound to scaffolding proteins. This complex is considered the most efficient system for cellulose degradation. *C. cellulovorans*, known to produce cellulosomes, changes the composition of its cellulosomes depending on the growth substrate. However, studies have only investigated cellulosomal proteins; changes in non-cellulosomal proteins have rarely been examined. In this chapter, we performed a quantitative proteome analysis of the whole exoproteome of *C. cellulovorans*, including cellulosomal and non-cellulosomal proteins, to illustrate how various substrates are efficiently degraded. *C. cellulovorans* was cultured with cellobiose, xylan, pectin, or phosphoric acid swollen cellulose (PASC) as the sole carbon source. PASC was used as a cellulose substrate for more accurate quantitative analysis. Using an isobaric tag method and a liquid chromatography mass spectrometer equipped with a long monolithic silica column, 639 proteins were identified and quantified in all 4 samples. Among these, 79 proteins were involved in saccharification, including 35 cellulosomal and 44 non-cellulosomal proteins. We compared protein abundance by spectral count and found that cellulosomal proteins were more abundant than non-cellulosomal proteins. Next, we focused on the fold change of the proteins depending on the growth substrate. Drastic changes were mainly observed among the non-cellulosomal proteins. These results indicate that cellulosomal proteins were primarily produced to efficiently degrade any substrate, and that non-cellulosomal proteins were specifically produced to optimize the degradation of a particular substrate. This chapter highlights the importance of non-cellulosomal proteins as well as cellulosomes for the efficient degradation of various substrates.

References

1. **Bayer EA, Belaich JP, Shoham Y, Lamed R.** 2004. The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu. Rev. Microbiol.* **58**:521-554.
2. **Doi RH, Kosugi A.** 2004. Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat. Rev. Microbiol.* **2**:541-551.
3. **Desvaux M.** 2005. The cellulosome of *Clostridium cellulolyticum*. *Enzyme. Microb. Technol.*

37:373-385.

4. **Bayer EA, Kenig R, Lamed R.** 1983. Adherence of *Clostridium thermocellum* to cellulose. J. Bacteriol. **156**:818-827.
5. **Lamed R, Setter E, Bayer EA.** 1983. Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. J. Bacteriol. **156**:828-836.
6. **Lamed R, Setter E, Kenig R, Bayer EA.** 1983. The cellulosome: a discrete cell surface organelle of *Clostridium thermocellum* which exhibits separate antigenic, cellulose-binding and various cellulolytic activities. Biotechnol. Bioeng. Symp. **13**:163-181.
7. **Desvaux M, Khan A, Scott-Tucker A, Chaudhuri RR, Pallen MJ, Henderson IR.** 2005. Genomic analysis of the protein secretion systems in *Clostridium acetobutylicum* ATCC 824. Biochim. Biophys. Acta **1745**:223-253.
8. **Tamaru Y, Miyake H, Kuroda K, Nakanishi A, Kawade Y, Yamamoto K, Uemura M, Fujita Y, Doi RH, Ueda M.** 2010. Genome sequence of the cellulosome-producing mesophilic organism *Clostridium cellulovorans* 743B. J. Bacteriol. **192**:901–902.
9. **Doi RH, Tamaru Y.** 2001. The *Clostridium cellulovorans* cellulosome: an enzyme complex with plant cell wall degrading activity. Chem. Rec. **1**:24–32.
10. **Tamaru Y, Miyake H, Kuroda K, Nakanishi A, Matsushima C, Doi RH, Ueda M.** 2011. Comparison of the mesophilic cellulosome-producing *Clostridium cellulovorans* genome with other cellulosome-related clostridial genomes. Microb. Biotechnol. **4**:64–73.
11. **Shoseyov O, Doi RH.** 1990. Essential 170-kDa subunit for degradation of crystalline cellulose by *Clostridium cellulovorans* cellulase. Proc. Natl. Acad. Sci. USA **87**:2192–2195.
12. **Cho W, Jeon SD, Shim HJ, Doi RH, Han SO.** 2010. Cellulosomic profiling produced by *Clostridium cellulovorans* during growth on different carbon sources explored by the cohesin marker. J. Biotechnol. **145**:233–239.
13. **Morisaka H, Matsui K, Tatsukami Y, Kuroda K, Miyake H, Tamaru Y, Ueda M.** 2012. Profile of native cellulosomal proteins of *Clostridium cellulovorans* adapted to various carbon sources. AMB Express **2**:37.
14. **Han SO, Cho HY, Yukawa H, Inui M, Doi RH.** 2004. Regulation of expression of cellulosomes and noncellulosomal (hemi)cellulolytic enzymes in *Clostridium cellulovorans* during growth on different carbon sources. J. Bacteriol. **186**:4218-4227.

15. **Morag E, Bayer EA, Lamed R.** 1990. Relationship of cellulosomal and noncellulosomal xylanases of *Clostridium thermocellum* to cellulose-degrading enzymes. J. Bacteriol. **172**:6098-6105.
16. **Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Johnstone R, Mohammed AK, Hamon C.** 2003. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal. Chem. **75**:1895–1904.
17. **Sleat R, Mah RA, Robinson R.** 1984. Isolation and characterization of an anaerobic, cellulolytic bacterium, *Clostridium cellulovorans* sp. nov. Appl. Environ. Microbiol. **48**:88–93.
18. **Zhang YH, Cui J, Lynd LR, Kuang LR.** 2006. A transition from cellulose swelling to cellulose dissolution by *o*-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. Biomacromolecules **7**:644–648.
19. **Bensadoun A, Weinstein D.** 1976. Assay of proteins in the presence of interfering materials. Anal. Biochem. **70**:241–250.
20. **Motokawa M, Kobayashi H, Ishizuka N, Minakuchi H, Nakanishi K, Jinnai H, Hosoya K, Ikegami T, Tanaka N.** 2002. Monolithic silica columns with various skeleton sizes and through-pore sizes for capillary liquid chromatography. J. Chromatogr. A **961**:53–63.
21. **de Hoon MJ, Imoto S, Nolan J, Miyano S.** 2004. Open source clustering software. Bioinformatics **20**:1453–1454.
22. **Saldanha AJ.** 2004. Java Treeview—extensible visualization of microarray data. Bioinformatics **20**:3246–3248.
23. **Raman B, Pan C, Hurst GB, Rodriguez M, Jr., McKeown CK, Lankford PK, Samatova NF, Mielenz JR.** 2009. Impact of pretreated Switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: a quantitative proteomic analysis. PLoS One **4**:e5271.
24. **Nataf Y, Bahari L, Kahel-Raifer H, Borovok I, Lamed R, Bayer EA, Sonenshein AL, Shoham Y.** 2010. *Clostridium thermocellum* cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factors. Proc. Natl. Acad. Sci. USA **107**:18646-18651.
25. **Tamaru Y, Doi RH.** 2001. Pectate lyase A, an enzymatic subunit of the *Clostridium cellulovorans* cellulosome. Proc. Natl. Acad. Sci. USA **98**:4125–4129.

26. **Arai T, Kosugi A, Chan H, Koukiekolo R, Yukawa H, Inui M, Doi RH.** 2006. Properties of cellulosomal family 9 cellulases from *Clostridium cellulovorans*. Appl. Microbiol. Biotechnol. **71**:654–660.
27. **Tamaru Y, Karita S, Ibrahim A, Chan H, Doi RH.** 2000. A large gene cluster for the *Clostridium cellulovorans* cellulosome. J. Bacteriol. **182**:5906–5910.
28. **Han SO, Yukawa H, Inui M, Doi RH.** 2005. Molecular cloning and transcriptional and expression analysis of *engO*, encoding a new noncellulosomal family 9 enzyme, from *Clostridium cellulovorans*. J. Bacteriol. **187**:4884–4889.
29. **Foong FC, Doi RH.** 1992. Characterization and comparison of *Clostridium cellulovorans* endoglucanases-xylanases EngB and EngD hyperexpressed in *Escherichia coli*. J. Bacteriol. **174**:1403–1409.
30. **Ishi A, Sheweita S, Doi RH.** 1998. Characterization of EngF from *Clostridium cellulovorans* and identification of a novel cellulose binding domain. Appl. Environ. Microbiol. **64**:1086–1090.
31. **Kosugi A, Murashima K, Doi RH.** 2002. Xylanase and acetyl xylan esterase activities of XynA, a key subunit of the *Clostridium cellulovorans* cellulosome for xylan degradation. Appl. Environ. Microbiol. **68**:6399–6402.
32. **Choi H, Fermin D, Nesvizhskii AI.** 2008. Significance analysis of spectral count data in label-free shotgun proteomics. Mol. Cell. Proteomics **7**:2373–2385.
33. **Gold ND, Martin VJ.** 2007. Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. J. Bacteriol. **189**:6787–6795.
34. **Xu C, Huang R, Teng L, Wang D, Hemme CL, Borovok I, He Q, Lamed R, Bayer EA, Zhou J, Xu J.** 2013. Structure and regulation of the cellulose degradome in *Clostridium cellulolyticum*. Biotechnol. Biofuels **6**:73.
35. **Ding SY, Liu YS, Zeng Y, Himmel ME, Baker JO, Bayer EA.** 2012. How does plant cell wall nanoscale architecture correlate with enzymatic digestibility? Science **338**:1055–1060.

Chapter II Proximity effect between cellulose-degrading enzymes displayed on the *Saccharomyces cerevisiae* cell surface

Cellulose is the most abundant organic polymer on earth (1). Lignocellulosic biomass has increasingly attracted attention as a promising alternative feedstock for biofuel (2). However, producing biofuels from lignocellulosic biomass by using the current technology is exceedingly expensive because of 2 possible reasons: (a) the recalcitrancy that hinders the deconstruction and use of the feedstock; and (b) the involvement of multiple steps, because of which a complicated infrastructure is required and the risk of contamination is enhanced. Overcoming these obstacles requires both synergistic reaction between cellulases and consolidated bioprocessing (CBP), which integrates the whole biofuel-production process (3, 4).

Certain naturally occurring microorganisms can degrade cellulose. Aerobic fungi such as *T. reesei* secrete several kinds of cellulases, mainly endoglucanase (EG), cellobiohydrolase (CBH), and β -glucosidase (BG), to completely degrade cellulose, and the free cellulases are recognized to degrade cellulose synergistically (5). Conversely, anaerobic bacteria such as *C. thermocellum* and *C. cellulovorans* produce a complex of cellulases, the cellulosome, and degrade lignocellulosic biomass efficiently (6-8). When the cellulases are located near each other within the cellulosome, they exhibit higher cellulose-degrading activity than they do when not present in the cellulosome; this is referred to as the proximity effect (9, 10).

In a previous study, to exploit the proximity effect and achieve CBP, we constructed an arming yeast that simultaneously displays, on the surface of yeast cells, the 3 kinds of cellulases that are necessary for completely degrading cellulose into glucose (11). The yeast strain could directly degrade and ferment cellulose to ethanol. Because the cellulases were concentrated on the cell surface and were nearer each other than they are when they exist freely, the cellulases were expected to degrade cellulose with a greater degree of synergy (i.e., to exhibit the proximity effect) as they do when present in the cellulosome. However, the proximity effect has not been demonstrated in the previous studies because of the challenges associated with quantifying the amounts of the displayed cellulases.

In this chapter, we used genome integration to construct a yeast strain that simultaneously displays EG II, CBH II from *T. reesei*, and BG I from *Aspergillus aculeatus*, and we tagged each of

the cellulases with a polypeptide or a fluorescent protein to enable the quantification of the displayed enzymes. The mixture of single-cellulase-displaying yeasts was prepared by quantifying and adjusting the amount of the displayed cellulases. Subsequently, we measured the proximity effect between the displayed cellulases by comparing phosphoric-acid-swollen cellulose (PASC)- and crystalline cellulose (Avicel)-degrading activities of the 3-cellulase-displaying-yeast strain with those of the mixture of single-cellulase-displaying yeasts. Furthermore, in this report, we discuss the relationship between the proximity effect and the distance between cellulases.

Materials and Methods

Plasmid construction

All the primers used for constructing plasmids are listed in Table 1. PCR was performed using KOD-plus-DNA polymerase (Toyobo, Osaka, Japan). The plasmids used for cellulase display were constructed as follows: Before constructing the plasmids used for genome integration, we constructed the multicopy plasmids used for displaying the cellulases tagged with a fluorescent protein or polypeptide. The *T. reesei* EG II gene present in pEG (12) was PCR-amplified using the BglII-FLAG-linker-EG-F and EG-SphI-extra-R primers to attach a FLAG tag to the N-terminus of EG. The amplified EG DNA fragment was inserted into the pRS425display (12) that was cut using BglII and SphI. The resulting plasmid included 2 FLAG tags, that is, on the N- and C-termini of EG, because a FLAG-tag-encoding sequence was originally included in the pRS425display; the plasmid constructed was named pRS425DF-EG.

The *T. reesei* CBH II gene in pCBH (12) was PCR-amplified using the BglII-CBH-F and CBH-SphI-R primers. The amplified CBH DNA fragment was inserted into the pRS426display (12), which was cut using BglII and SphI, and the resulting plasmid was named pRS426D-CBH. The EGFP gene present in pULSG1 (13) was PCR-amplified using the SphI-EGFP-F and EGFP-SphI-R primers, and the amplified EGFP DNA fragment was inserted into the pRS426D-CBH plasmid that was cut using SphI. The resulting plasmid was named pRS426DE-CBH.

The *A. aculeatus* BG I gene in pBG (12) was PCR-amplified using the NotI-BG-F and BG-SphI-R primers. The amplified BG I gene fragment was inserted in a plasmid that was constructed as follows: The sequence encoding the *GAPDH* promoter, the secretion signal of

Table 1 Plasmids and primers used in this study

Plasmids and primers	Feature or sequence
<i>plasmids</i>	
pRS425DF-EG	Cell surface display of endoglucanase II with FLAG tag, <i>LEU2</i>
pRS426DE-CBH	Cell surface display of cellobiohydrolase II with EGFP, <i>URA3</i>
pRS423Dm-BG	Cell surface display of β -glucosidase I with mCherry, <i>HIS3</i>
pRS402	<i>ADE2</i>
pRS403	<i>HIS3</i>
pRS405	<i>LEU2</i>
pRS406	<i>URA3</i>
pRS403DF-EG	Cell surface display of endoglucanase II with FLAG tag, <i>HIS3</i>
pRS405DF-EG	Cell surface display of endoglucanase II with FLAG tag, <i>LEU2</i>
pRS406DF-EG	Cell surface display of endoglucanase II with FLAG tag, <i>URA3</i>
pRS403DE-CBH	Cell surface display of cellobiohydrolase II with EGFP, <i>HIS3</i>
pRS405DE-CBH	Cell surface display of cellobiohydrolase II with EGFP, <i>LEU2</i>
pRS406DE-CBH	Cell surface display of cellobiohydrolase II with EGFP, <i>URA3</i>
pRS403Dm-BG	Cell surface display of β -glucosidase I with mCherry, <i>HIS3</i>
pRS405Dm-BG	Cell surface display of β -glucosidase I with mCherry, <i>LEU2</i>
pRS406Dm-BG	Cell surface display of β -glucosidase I with mCherry, <i>URA3</i>
<i>primers</i>	
BglII-FLAG-linker-EG-F	5'-ATGC <u>AGATCT</u> IGATTACAAGGATGACGATGACAAGGGTGGATCTACTGTCTGGGG CCAGTGTG-3'
EG-SphI-extra-R	5'-TGCAGTCGGACGATGCGCATGCCTTTCTTGCGAGACACGAGCTG-3'
BglII-CBH-F	5'-ATGC <u>AGATCT</u> CAAGCTTGCTCAAGCGTCTGG-3'
CBH-SphI-R	5'-ATGCGCATGCCAGGAACGATGGGTTTGCGTTTG-3'
SphI-EGFP-F	5'-ATGCGCATGCCGGTGGATCTGGTGGCGTGA-3'
EGFP-SphI-R	5'-ATGCGCATGCCTTGTACAGCTCGTCCATGCC-3'
NotI-BG-F	5'-ATGCGCGGCCGCGATGAACTGGCGTTCTCTCCTC-3'
BG-SphI-R	5'-ATGCGCATGCTTGACACCTTCGGGAGCGC-3'
SphI-5linker-mCherry-F	5'-ATGCGCATGCCGGTGGATCTGGTGGCGTGAGCAAGGGCGAGGAG-3'
mCherry-SphI-R	5'-ATGCGCATGCCTTGTACAGCTCGTCCATGCC-3'
infusion GAPDH promoter F	5'-CGGGCCCCCCTCGAGACCAGTTCTCACACGGAACAC-3'
infusion GAPDH terminator R	5'-ACCGCGGTGGCGCCGCTTTGATTATGTTCTTTCTATTTGAATGAGATATGA-3'

Underlined sequences: Restriction enzyme sites

glucoamylase from *Rhizopus oryzae*, the multicloning site, the strep tag, and the C-terminal region of α -agglutinin that are present in the pRS426display were inserted into the Sall–SacII section of pRS423 (American Type Culture Collection, VA, USA) after digestion with Sall and SacII. The resulting plasmid was named pRS423display. The amplified BG I gene fragment was inserted into the pRS423display, and this plasmid was named pRS423D-BG. The mCherry gene present in pmCherry-N1 (Clontech, CA, USA) was PCR-amplified using the SphI-5linker-mCherry-F and mCherry-SphI-R primers, and the fragment was inserted into pRS423D-BG that was cut using SphI. The resulting plasmid was named pRS423Dm-BG.

The gene cassettes used for the cell-surface display of FLAG-EG, CBH-EGFP, and BG-mCherry (*GAPDH* promoter, secretion signal of glucoamylase, the cellulase containing a polypeptide tag or a fluorescence protein, and the C-terminal region of α -agglutinin) were amplified with the infusion *GAPDH* promoter F and infusion *GAPDH* terminator R primers (Table 1). Each of the amplified fragments was inserted into the each of the 3 plasmids, pRS403, pRS405, and pRS406 (American Type Culture Collection), that were cut using XhoI and NotI; thus, we constructed the 9 plasmids that were used for genome integration (Table 1).

Table 2 Constructed yeast strains

Name	Introduced plasmids
Cont-yeast ^a	pRS402, pRS403, pRS405, pRS406
EG-yeast ^b	pRS402, pRS403DF-EG, pRS405DF-EG, pRS406DF-EG
CBH-yeast ^c	pRS402, pRS403DE-CBH, pRS405DE-CBH, pRS406DE-CBH
BG-yeast ^d	pRS402, pRS403Dm-BG, pRS405Dm-BG, pRS406Dm-BG
ALL-yeast ^e	pRS402, pRS403Dm-BG, pRS405DF-EG, pRS406DE-CBH

^aCont-yeast: a yeast strain that displays no cellulase

^bEG-yeast: a yeast strain that displays only endoglucanase (EG)

^cCBH-yeast: a yeast strain that displays only cellobiohydrolase (CBH)

^dBG-yeast: a yeast strain that displays only β -glucosidase (BG)

^eALL-yeast: a yeast strain that simultaneously displays EG, CBH, and BG

MIX-yeast was prepared by mixing EG-yeast, CBH-yeast, and BG-yeast.

Strains and media

S. cerevisiae wild-type strain W303-1A (*MATa*, *leu2-3/112*, *ura3-1*, *trp1-1*, *his3-11/15*, *ade2-1*, *can1-100*) was used for the cell-surface display. The cellulase-displaying yeast strains constructed in this study are listed in Table 2. Because W303-1A is colored red because of the accumulation of the adenine precursor, the strain was transformed with pRS402; this eliminated the red color and allowed mCherry fluorescence to be observed. Yeast transformants were aerobically cultivated in yeast extract peptone dextrose (YPD) medium (1% (w/v) yeast extract, 2% glucose, 2% peptone) that was buffered at pH 6.0 with 50 mM 2-morpholinoethanesulfonic acid (MES). *Escherichia coli* DH5 α (F^- , ϕ 80dlacZ Δ M15, Δ (*lacZYA-argF*)U169, *deoR*, *recA1*, *endA1*, *hsdR17*(r_K^- , m_K^+), *phoA*, *supE44*, λ^- , *thi-1*, *gyrA96*, *relA1*) was used as a host for manipulating recombinant DNA and this strain was grown in Luria–Bertani medium (1% tryptone, 0.5% yeast extract, 0.5% sodium chloride) containing 100 μ g/mL ampicillin.

Immunostaining and microscopic examination of the 5 yeast strains

All cellulase-displaying yeast strains were precultivated for 24 h at 30°C in YPD medium to inoculate the same number of yeast cells for main-cultivation, and cultivated for 72 h at 30°C in YPD medium in all the following experiments.

After the cultivation, OD_{0.5} yeast cells (equal to the yeast cell number in 1 mL of yeast suspension at OD₆₀₀ = 0.5) of 5 yeast strains listed in Table 2 were harvested, washed once with phosphate-buffered saline, pH 7.4 (PBS), and then used for quantifying the displayed cellulases. The cell pellets were resuspended using 50 μ L of PBS containing 1% bovine serum albumin, and then the yeast suspensions were incubated on a rotary shaker; all incubations in this experiment were at room temperature. After the suspensions were incubated for 30 min, 1 μ L of an anti-FLAG mouse monoclonal primary antibody (Sigma-Aldrich, MO, USA) was added to them. Subsequently, they were incubated for 1.5 h, following which they were washed twice with PBS and the precipitates were resuspended using 50 μ L of PBS. Next, we added 8 μ L of CFTM405S-conjugated anti-mouse goat monoclonal secondary antibodies (Biotium, CA, USA) and incubated the cell suspensions for another 1.5 h. After the incubation, the suspensions were washed twice with PBS and the precipitates were resuspended using 40 μ L of PBS. These immunostained yeast strains were examined under an inverted microscope (IX71; Olympus, Tokyo, Japan). The green fluorescence of EGFP was detected

through a U-MNIBA2 mirror unit containing a BP470-490 excitation filter, a DM505 dichroic mirror, and a BA510-550 emission filter (Olympus). The red fluorescence of mCherry was detected through a U-MWIG2 mirror unit containing a BP520-550 excitation filter, a DM565 dichroic mirror, and a BA580IF emission filter (Olympus). The blue fluorescence of the CFTM405S fluorophore was detected through a U-MNUA2 mirror unit containing a BP360-370 excitation filter, a DM400 dichroic mirror, and a BA420-460 emission filter (Olympus).

Fluorometric analysis

After immunofluorescence labeling, the yeast cells were suspended in 200 μ L of PBS in 96-well black plates (353945, BD Falcon) and the fluorescence of the cells was measured using a Fluoroskan Ascent fluorometer (Labsystems OY, Helsinki, Finland). Filter pairs of 355/460 nm, 485/510 nm, and 584/612 nm were used for detecting the fluorescence of CFTM405S, EGFP, and mCherry, respectively.

Cellulase-activity assay

BG activity was measured using 20 mM of *p*-nitrophenyl glucopyranoside (PNPG) (30.125 mg/5 mL) as a substrate. OD0.05 yeast cells were collected by centrifuging them for 1 min at 12,000 $\times g$, and then washed once with a 50 mM citrate buffer. The cell pellet was resuspended using 750 μ L of the 50 mM citrate buffer and incubated in a heat block at 50°C for 30 min. The PNPG solution was also incubated at 50°C for 30 min, and 250 μ L of this solution was added to the yeast suspension to start the reaction. During the reaction, samples were withdrawn at 5, 10, and 15 min and mixed with 0.2M Na₂CO₃ in a 1:1 ratio, and 200 μ L of the mixed solution was used for measuring Abs₄₀₀ using the Fluoroskan Ascent fluorometer and the 96-well black plates mentioned in the preceding section.

EG and CBH activities and cellulase proximity effects were measured using 1% PASC (14) as a substrate. OD10 or OD0.1 yeast cells were collected and washed once with the 50 mM citrate buffer. The cell pellet was resuspended using 500 μ L of a 100 mM citrate buffer and incubated in a heat block at 50°C for 1 h, after which 500 μ L of 2% PASC was added to start the reaction (which was performed at 50°C). Reaction samples were withdrawn at 6, 12, 24, and 48 h of incubation in the case of OD10 yeast cells and at 24, 48, 84, and 120 h of incubation in the case of OD0.1 yeast

cells, and then the supernatants were mixed with DNS solution (NaOH, 16 g/L; potassium sodium tartrate, 300 g/L; 3,5-dinitrosalicylic acid [DNS], 5 g/L) in a 1:2 ratio to quantify the reduced ends of degraded products. The mixed solutions were incubated at 100°C for 5 min, and after cooling them to room temperature, 200 μ L of each solution was used for measuring Abs₅₃₀ using the Fluoroskan Ascent fluorometer and 96-well black plates.

Results

Construction of a yeast strain that simultaneously displays 3 kinds of cellulose-degrading enzymes on the cell surface

To demonstrate that the 3 kinds of cellulases displayed on yeast cell surface, EG, CBH, and BG, degrade cellulose synergistically (Fig. 1), we constructed these 5 yeast strains: Cont-yeast (yeast displaying no cellulases), EG-yeast (yeast displaying FLAG-tagged EG), CBH-yeast (yeast displaying CBH fused with EGFP), BG-yeast (yeast displaying BG fused with mCherry), and ALL-yeast (yeast simultaneously displaying EG-FLAG, CBH-EGFP, and BG-mCherry) (Table 2). We constructed 13 plasmids that were used for the display (Fig. 2), and these plasmids were transformed into the yeast *S. cerevisiae* W303-1A strain to obtain the aforementioned 5 yeast strains (Tables 1 and 2). To readily visualize and quantify the displayed cellulases, EGFP and mCherry were fused with CBH and BG, respectively; in the case of EG, the attached FLAG tag was detected by immunostaining.

To confirm that the 5 yeast strains displayed each of the cellulases properly, we immunostained the yeast cells by using an anti-FLAG primary antibody and a CFTM405S-conjugated anti-mouse secondary antibody. The blue fluorescence of CFTM405S, the green color of EGFP, and the red color of mCherry were observed under a fluorescence microscope (Fig. 3). This microscopic analysis revealed that, whereas Cont-yeast exhibited no fluorescence, EG-yeast fluoresced blue, CBH-yeast fluoresced green, BG-yeast fluoresced red, and ALL-yeast exhibited fluorescence in all 3 colors. Notably, almost 100% of EG-, CBH-, BG-, and ALL-yeast cells displayed cellulases. These results indicate that ALL-yeast simultaneously displayed the 3 kinds of cellulases and that all 5 yeast strains were successfully constructed.

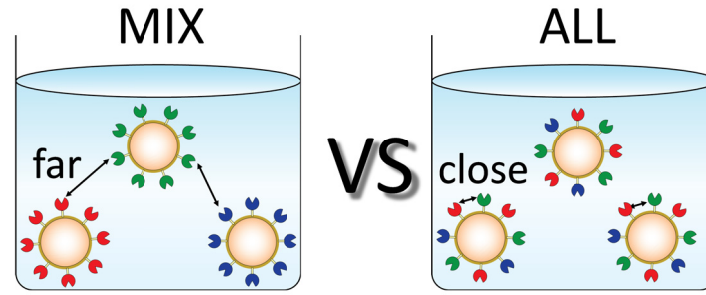


Fig. 1 Schematic illustration of experiments conducted to demonstrate the proximity effect between cellulases displayed on the surface of yeast cells

MIX-yeast is a mixture of 3 single-cellulase-displaying yeasts, EG-yeast (for example, blue), CBH-yeast (green), and BG-yeast (red), which are listed in Table 2. MIX-yeast was prepared by mixing the 3 single-cellulase-displaying yeasts in a ratio corresponding to the display ratio of ALL-yeast. The proximity effect between the cellulases displayed on ALL-yeast, which simultaneously displays EG, CBH, and BG, was determined by comparing the PASC-degrading activity of ALL-yeast and MIX-yeast.

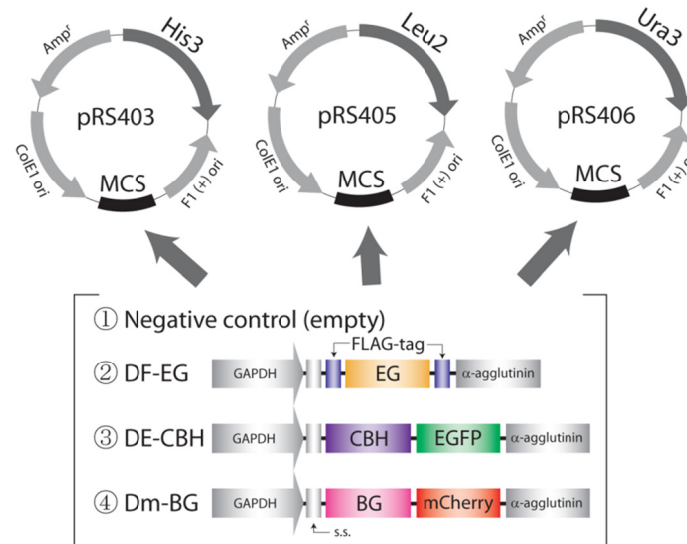


Fig. 2 Maps of the plasmids used for generating the 5 cellulase-displaying yeast strains

The 3 gene cassettes DF-EG, DE-CBH, and Dm-BG were constructed and inserted into pRS403, pRS405, and pRS406. D: Display, F: FLAG tag, E: EGFP, m: mCherry. All of the plasmids were introduced into yeast as listed in Table 2. Genes encoding FLAG tag, EGFP, and mCherry were fused with the cellulase genes to visualize and quantify the displayed cellulases.

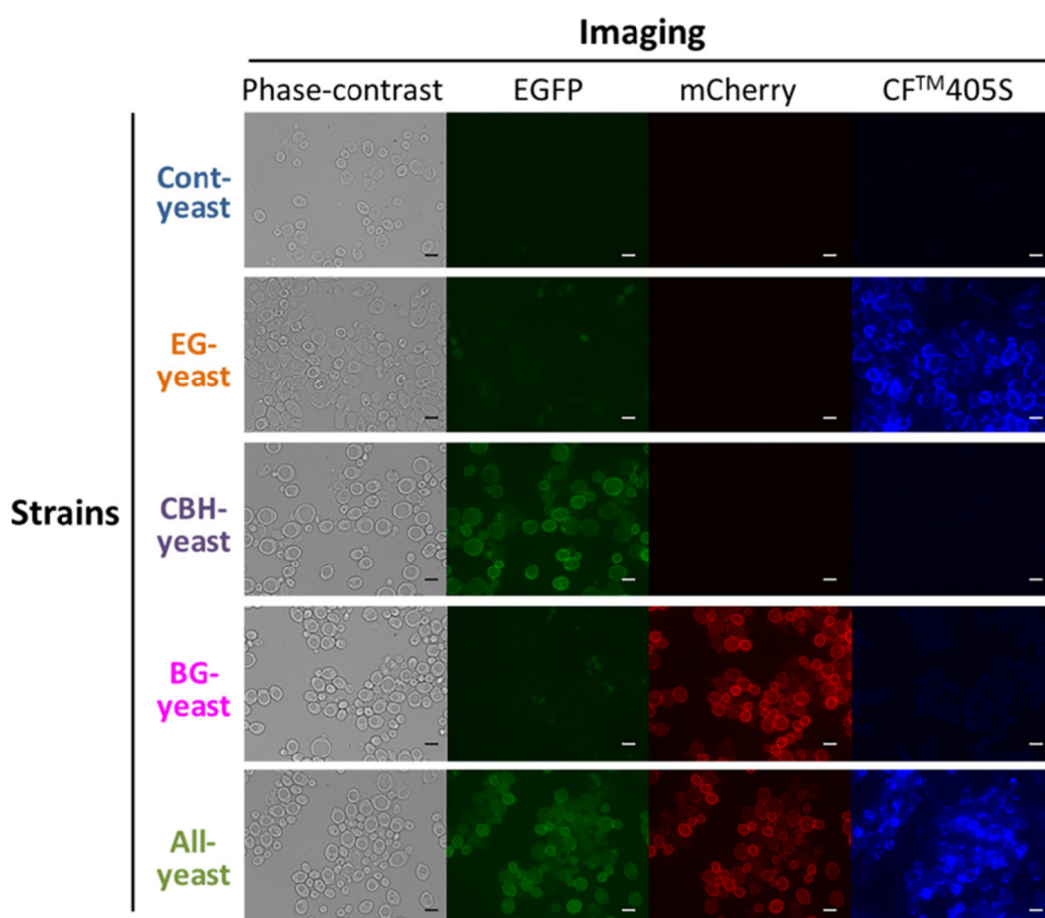


Fig. 3 Confirmation of cellulase display by immunostaining and fluorescence microscopy

Cells of 5 yeast strains were immunostained using an anti-FLAG mouse monoclonal primary antibody and CFTM405S-conjugated anti-mouse goat monoclonal secondary antibodies. Fluorescence microscopy was used to examine the fluorescence of EGFP on CBH-yeast and ALL-yeast, CFTM405S on EG-yeast and ALL-yeast, and mCherry on BG-yeast and ALL-yeast (abbreviations are listed in Table 2). Scale bar = 5 μ m.

Confirmation of cellulase activities of the 5 yeast strains constructed

After constructing the 5 yeast strains, we examined their cellulase activities to measure the PASC-degrading activities of EG-yeast, CBH-yeast, and ALL-yeast, or PNPG to measure the BG activities of BG-yeast and ALL-yeast. All 5 yeast strains were incubated at 50°C for 1 h to kill and prevent ALL-yeast from ingesting the produced glucose before the reaction. Incubation at 50°C does not affect the activity of the displayed cellulases because their optimal temperature is 50°C. After this inactivation step, we confirmed that glucose was not consumed by ALL-yeast during the

reaction (data not shown). The results presented in Figure 4 show that all of the constructed yeast strains exhibited their specific activities. Notably, the PASC-degrading activity of EG-yeast decreased slightly over time whereas that of ALL-yeast was exactly proportional to the time course (Fig. 4a); this result suggests that, whereas EG displayed on EG-yeast was inhibited by the degradation products generated from PASC (e.g., cellobiose), the glucose-tolerant BG that was displayed on ALL-yeast degraded oligosaccharides to glucose, and thus the EG activity of ALL-yeast was not inhibited.

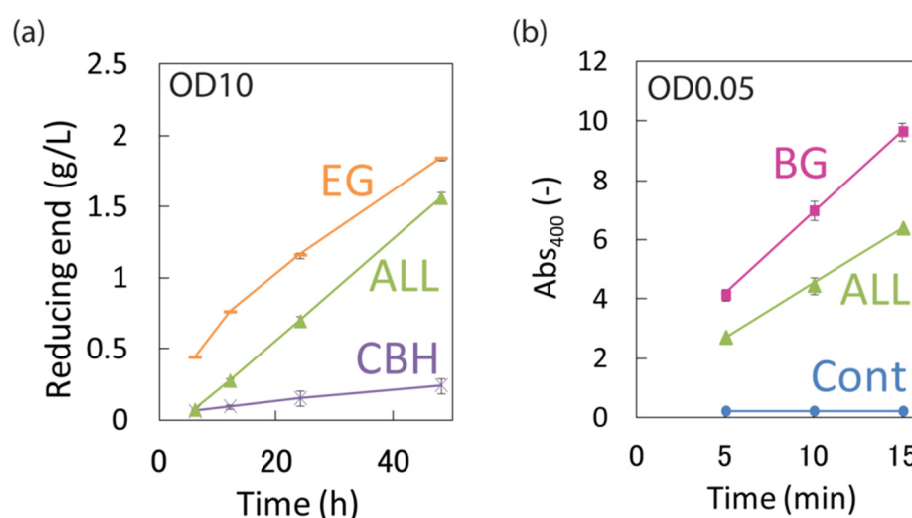


Fig. 4 Confirmation of the PASC-degrading activity and BG activity of the 5 cellulase-displaying yeast strains

(a) PASC-degrading activities (reducing end) of EG and CBH displayed on EG-yeast, CBH-yeast, and ALL-yeast were measured. Yeast cells were used at a concentration of OD10 in the reaction. The PASC-degrading activity of Cont-yeast was subtracted from the activities measured for EG-yeast, CBH-yeast, and ALL-yeast and the values obtained are plotted here. (b) BG activities (Abs₄₀₀) of BG displayed on BG-yeast and ALL-yeast were measured using PNPG as a substrate. Yeast cells were used at a concentration of OD0.05 in the reaction. The reactions were performed at 50°C, and the yeasts were incubated at 50°C for 1 h before the reaction started to prevent ALL-yeast from taking up the produced glucose. The data represent reducing ends (g/L) generated by the yeast strains in the PASC-degrading reaction, and the average ($\pm$ standard deviation) of 3 independent experiments is shown.

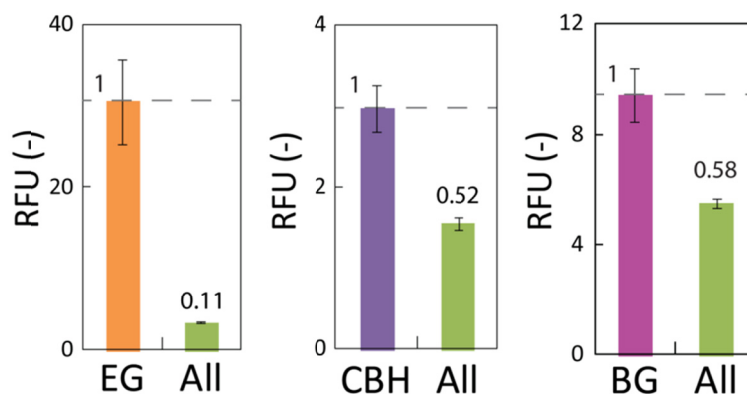


Fig. 5 Comparison of the amounts of cellulases displayed on ALL-yeast and EG-, CBH-, or BG-yeast

The amounts of the displayed cellulases were determined by measuring the fluorescence after immunostaining. Filter pairs of 355/460 nm, 485/510 nm, and 584/612 nm were used to quantify the fluorescence of CFTM405S, EGFP, and mCherry, respectively. RFU: relative fluorescence units. The data represent the RFU values of the yeast strains, and the average ($\pm$ standard deviation) of 3 independent experiments is shown.

Determination of the amounts of the displayed cellulases and their proximity effect

To compare the cellulase activity of ALL-yeast with that of MIX-yeast, which was a mixture of single-cellulase-displaying yeasts (Fig. 1), we had to quantify and adjust the amount of the cellulases used in the reaction. First, the amount of EG displayed on EG-yeast and ALL-yeast was quantified by measuring the fluorescence intensity after immunostaining and it was 1:0.11 (Fig. 5). Next, the fluorescence intensity of EGFP was measured to quantify the CBH displayed on CBH-yeast and ALL-yeast, and the ratio was 1:0.52. Finally, the fluorescence intensity of mCherry was measured to quantify the BG displayed on BG-yeast and ALL-yeast, and the ratio was 1:0.58.

The BG activity of BG-yeast was approximately 1.6 times higher than that of ALL-yeast (Fig. 4b). Because EG-yeast and CBH-yeast did not exhibit any activity toward PNPG, this result suggests that the ratio of the activity corresponds to the ratio of the amount of the enzymes displayed on BG-yeast and ALL-yeast. As the display ratio calculated for BG on BG-yeast:ALL-yeast on the basis of the fluorescence intensity was almost the same as the ratio determined from the PNPG assay, it indicates that the display ratio calculated from the fluorescence intensity is reliable.

MIX-yeast was prepared by mixing the 3 single-cellulase-displaying yeasts in the aforementioned ratios: MIX-yeast was composed of OD 1.1 of EG-yeast + OD 5.2 of CBH-yeast + OD 5.8 of BG-yeast; thus, the total amount of the used cellulases derived from MIX-yeast and ALL-yeast was the same. Because the total amount of MIX-yeast added up to OD12.1, we added OD2.1 of Cont-yeast to OD10 of ALL-yeast to adjust the total amount of yeast cells. ALL-yeast or MIX-yeast was mixed with PASC, and their PASC-degrading activities were compared. The activity of ALL-yeast was 1.25 fold higher than that of MIX-yeast, which suggests that the cellulases simultaneously displayed on ALL-yeast synergistically degraded PASC (Fig. 6a). Based on this result, we surmised that the distance between the different kinds of cellulases (i.e., the distance between yeast cells in the case of MIX-yeast versus the distance between cellulases on single yeast cells in the case of ALL-yeast) was critical for the synergistic reaction between the cellulases; this is because the distance between the cellulases in ALL-yeast is substantially shorter than that in MIX-yeast (Fig. 1).

We conducted an additional experiment to support the proximity effect (i.e., to demonstrate that increased PASC-degrading activity is related to the distance between cellulases). The amount of the yeast cells used in the PASC-degrading reaction was lowered to OD0.1 for the purpose of further increasing the distance between cellulases in MIX-yeast while maintaining the same distance between the cellulases in ALL-yeast. At OD0.1, relative to MIX-yeast, ALL-yeast exhibited even higher (2.22 fold higher) PASC-degrading activity than it did when the activities of these yeasts were measured at OD10 (Fig. 6b). The difference in PASC-degrading activity between ALL-yeast and MIX-yeast at the low concentration was approximately 5-fold greater than that when the high concentration of yeast was used. This result suggests that the proximity effect functioned in enhancing PASC degradation by the 3 kinds of cellulases displayed on ALL-yeast.

Discussion

In this study, we successfully constructed a yeast strain that simultaneously displays 3 kinds of cellulases and demonstrated that the simultaneous display promoted synergistic reaction between the different kinds of cellulases (the proximity effect). More specifically, simultaneous display and the proximity effect have been demonstrated after adjusting the amounts of the displayed enzymes

by using arming yeast. Previous reports have indicated that a yeast strain that simultaneously displays several kinds of enzymes exhibits higher activity than a mixture of single-enzyme-displaying yeast strains; however, the amounts of the displayed enzymes were not quantified, and thus the amount of the enzymes used was not adjusted (11, 15).

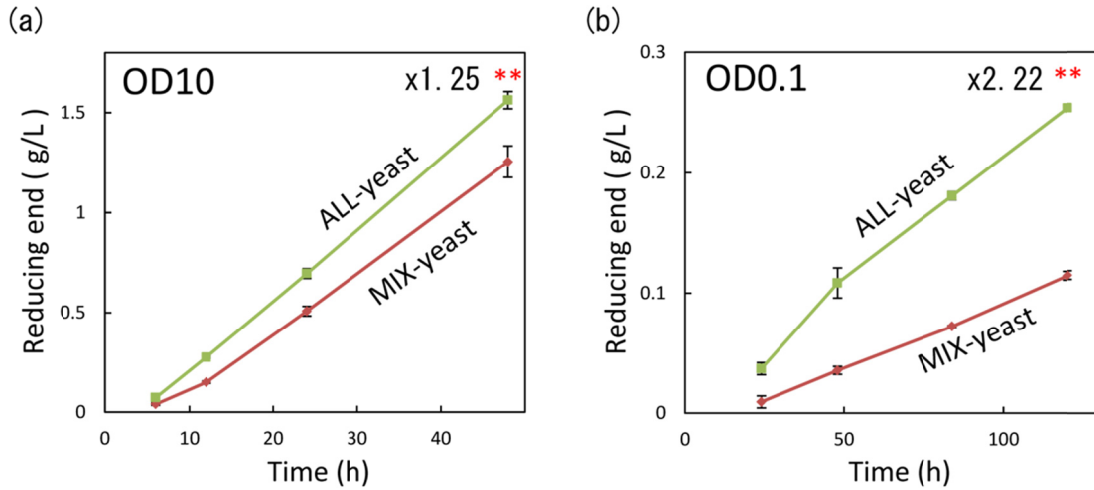


Fig. 6 Proximity effect between cellulases displayed on ALL-yeast

MIX-yeast was prepared based on the ratio of the amount of each cellulase displayed on ALL-yeast to that on the single-yeast strains. (a) Yeast cells were used at OD10 in the reaction. MIX-yeast: OD1.1 of EG-yeast + OD5.2 of CBH-yeast + OD5.8 of BG-yeast. ALL-yeast: OD2.1 of Cont-yeast + OD10 of ALL-yeast. (b) Yeast cells were used at OD0.1 in the reaction. At this lower yeast concentration, the distance between yeast cells was greater than that when yeast cells were used at OD10, and, consequently, the distance between the different kinds of cellulases in MIX-yeast was greater at OD0.1 than at OD10. To obtain the OD0.1 yeast strains, OD10 MIX-yeast and OD10 ALL-yeast were diluted 100 fold. The yeast strains were used in reactions containing 1% PASC, and the reduced ends of the degraded products were quantified using the DNS assay. The data represent the reduced ends (g/L) generated by the yeast strains in the PASC-degrading reaction; the average ($\pm$ standard deviation) of 3 independent experiments is shown. $**P < 0.01$, calculated using the data at end-point and t tests. ALL-yeast showed 1.25 fold (a) or 2.22 fold (b) higher PASC-degrading activity. The activity of Cont-yeast was subtracted from the activities of ALL-yeast and MIX-yeast, and they were plotted.

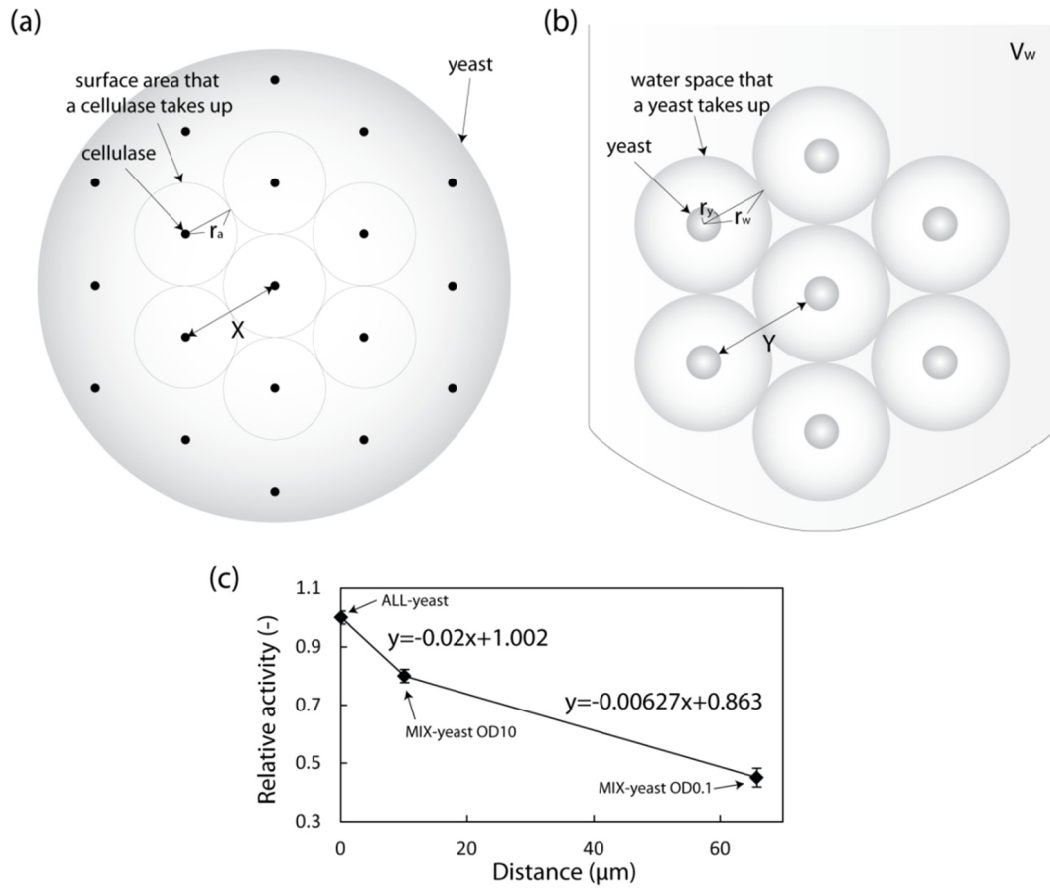


Fig. 7 Approximate calculation of the distance between cellulases in ALL-yeast and MIX-yeast and the relationship between distance and the proximity effect

(a) The distance between the different kinds of cellulases displayed on ALL-yeast (X) was calculated. The area occupied by a displayed cellulase was set to be a circle to simplify the calculation. r_a , radius of the area that a displayed cellulase occupies. (b) The distance between the different kinds of cellulases displayed on MIX-yeast (Y) was calculated. The volume occupied by a yeast cell was set to be a sphere to simplify the calculation. r_w , radius of the water space that a yeast cell occupies; $V_w = 1 \text{ mL } (10^{-6} \text{ m}^3)$, total volume of water used in the reaction; $r_y = 2.5 \text{ } \mu\text{m}$, radius of a yeast cell. (c) Relative activities of ALL-yeast and MIX-yeast in PASC degradation are plotted against the distance between cellulases.

The distances between different kinds of cellulases in ALL-yeast and MIX-yeast were calculated approximately to analyze the relationship between the distance and the proximity effect. We hypothesized that the cellulases are evenly distributed on the cell surface of yeast, and we set the

area or the volume occupied by a displayed cellulase or a yeast cell to be a circle or a sphere, respectively, to simplify the calculations (Fig. 7). Formulae [1] and [2] provided below were used for calculating the distance between the 3 cellulases in ALL-yeast, and formulae [3], [4], and [5] were used for calculating the distance between the cellulases in MIX-yeast.

$$\pi r_a^2 = 4\pi r_y^2 / n_d \quad [1]$$

$$X = 2r_a \quad [2]$$

$$4/3\pi r_w^3 = V_w / n_y \quad [3]$$

$$r_w - r_y = r_c \quad [4]$$

$$Y = 2r_c \quad [5]$$

The variables are the following: r_a , putative radius of the area that a displayed cellulase occupies; $r_y = 2.5 \mu\text{m}$, putative radius of a yeast cell; $n_d = 10^4$, total number of displayed cellulases on the surface of a yeast cell (16); X , putative distance between the 3 kinds of cellulases displayed on ALL-yeast; r_w , putative radius of the surrounding space that a yeast cell occupies; $V_w = 1 \text{ mL}$ (10^{-6} m^3), total volume of water used in the reaction; $n_y = 5.55 \times 10^8$ (OD10) or 5.55×10^6 (OD0.1), the number of yeast cells used in the reaction; Y , putative distance between the different kinds of cellulases displayed on MIX-yeast.

Table 3 Relationship between distance and the proximity effect^a

	Distance ^b (μm)	Relative activity ^c
ALL-yeast	0.100	1.00 ± 0.02
MIX-yeast OD10	10.1	0.80 ± 0.02
MIX-yeast OD0.1	65.9	0.45 ± 0.03

^a Reducing-end values at 48 and 120 h were used in the calculation for the OD10 and OD0.1 samples, respectively.

^b The distance between the different kinds of cellulases. To calculate the distance approximately, we set the variables of the radius of yeasts, the amount of total displayed cellulases, and the number of cells in an OD1 yeast-cell culture to be $2.5 \mu\text{m}$ (putative), 10^4 cellulases/yeast cell (a previous study), and 5.55×10^7 (counted), respectively.

^c Relative activity is (reducing-end value at endpoint of MIX-yeast or ALL-yeast)/(reducing-end value at endpoint of ALL-yeast)

The distance between the 3 kinds of cellulases in ALL-yeast was 0.1 μm , and the distance in MIX-yeast, which is equal to the distance between yeast cells, was 10.1 μm in case of OD10 and 65.1 μm in case of OD0.1 (Table 3). Relative activity was calculated by dividing the reducing end value of ALL-yeast or MIX-yeast at the endpoint (48 h for OD10 and 120 h for OD0.1) by the value of ALL-yeast. Our results showed that the proximity effect increased when the distance was shortened (Fig. 7c). Furthermore, the absolute value of the gradient between 0.1 and 10.1 μm (0.02) was >3-fold higher than that between 10.1 and 65.1 μm (0.00627), which implies that the proximity effect changes more drastically when the distance range shrinks.

We have found that the distance between cellulases is critical for the occurrence of the synergistic reaction: the strength of the proximity effect increased with decrease in the distance between the different kinds of cellulases, and the proximity effect increased more sharply at shorter distances than at longer distances (Fig. 7). These results suggest that enhancing the display efficiency would lead to increased proximity effect because the distance between displayed cellulases would be decreased. In the development of arming technology, substantial effort has been devoted to improving display efficiency by using mainly 3 strategies. First, display can be enhanced by integrating increased numbers of genes; this can be achieved by inserting 2 or 3 cassettes of genes into a plasmid or by means of δ -integration (17). Yamada *et al.* (18) reported a cocktail δ -integration method that they used to successfully insert multiple genes and readily optimize the ratio of displayed cellulases, which resulted in an increase in cellulose-degrading activity. Second, disruption of cell-wall proteins or proteins related to cell-wall formation can increase display efficiency. SED1 is a major structural cell-wall protein expressed in the stationary phase (19), and it is considered to compete with α -agglutinin for cell-surface display. A *SED1*-disrupted yeast strain was constructed previously, which showed increased display efficiency, particularly during the stationary-growth phase (20, 21). Moreover, Matsuoka *et al.* (22) reported that the disruption of the MNN2 gene that encodes α -1,2-mannosyltransferase increased overall display efficiency, but did so particularly on the outer surface where high-molecular-weight substrates can bind. Third, the gene promoter or enzyme-anchoring domain used can be changed to enhance display (23). In the study, we used the GAPDH promoter and the α -agglutinin C-terminal-anchor domain. In a recent study, the GAPDH promoter and the α -agglutinin anchor domain were replaced with the SED1 promoter and the SED1 anchor domain respectively; the promoter replacement increased the overall display efficiency and

the anchor-domain replacement increased the display efficiency on the outer surface. Using these strategies to enhance display efficiency is expected to lead not only to an increase in the total amounts of displayed cellulases but also to an increase in the proximity effect.

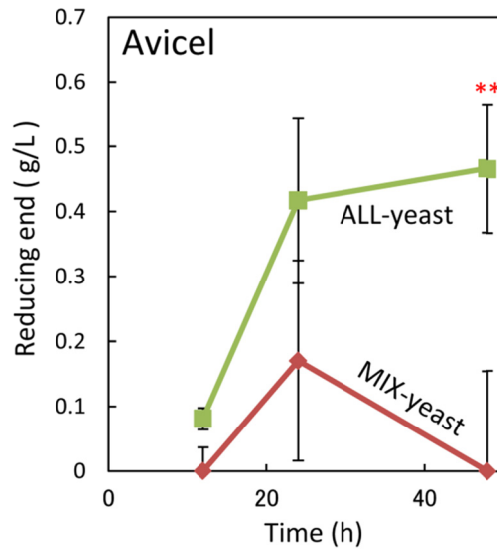


Fig. 8 Schematic illustration of experiments conducted to demonstrate the proximity effect between cellulases displayed on the surface of yeast cells

MIX-yeast is a mixture of 3 single-cellulase-displaying yeasts, EG-yeast (for example, blue), CBH-yeast (green), and BG-yeast (red), which are listed in Table 2. MIX-yeast was prepared by mixing the 3 single-cellulase-displaying yeasts in a ratio corresponding to the display ratio of ALL-yeast. The proximity effect between the cellulases displayed on ALL-yeast, which simultaneously displays EG, CBH, and BG, was determined by comparing the PASC-degrading activity of ALL-yeast and MIX-yeast.

We have demonstrated the proximity effect of arming yeast on PASC, but whether it shows the proximity effect on Avicel is also an interesting issue. Krauss *et al.* (24) reported that the addition of scaffoldin to the scaffoldin-disrupted *C. thermocellum* culture supernatant increased the cellulase activity only when Avicel was used as a substrate, whereas it did not differ greatly when PASC or soluble cellulose was used, which suggests exclusive proximity effect on Avicel. From this report, we expected that the arming yeast would show even greater proximity effect on Avicel, and we

measured the Avicel-degrading activity of the MIX-yeast and ALL-yeast (Fig. 8). Although the Avicel-degrading activity of MIX-yeast was not high enough to be detected, ALL-yeast showed significantly higher Avicel-degrading activity than MIX-yeast, which suggests that the cellulases displayed on ALL-yeast degraded Avicel synergistically and the proximity effect could be even bigger than on PASC.

Arming yeasts offer the advantageous feature of serving as a biocatalyst of cellulose degradation, and, additionally, these cellulase-displaying yeasts can be used to further study the proximity effect. Studies on cellulosomes have yielded unexpected results showing that decreasing the distance between cellulases in the cellulosome does not necessarily lead to increased cellulose-degrading activity (25, 26); this was unexpected because of the lack of information on the natural fine structure of the cellulosome. In the previous study in which a designed minicellulosome was used, proximity effects of cellulases at varying distances were evaluated (25). The lengths of the linkers between cohesins (domains in the scaffoldin structure of the cellulosome that cellulases can bind to) were varied using, for example, 0, 5, or approximately 30 amino acids, and the designed minicellulosome featuring the shortest linker exhibited the lowest cellulose-degrading activity because of the conformational constraint and steric hindrance between the cellulases. Moreover, a similar cellulose-degrading activity was reported when a linker length of up to 128 amino acids was used, even though the distance between the cellulases increased (26).

A 30 amino-acid linker is estimated to be around 10 nm long. In our study, the distance between the cellulases in ALL-yeast was calculated to be 100–10 nm when the display efficiency was 10^4 – 10^5 per cell, which is comparable to the distance between cellulases in the minicellulosome. Further investigation of the proximity effect at this close range conducted by using arming yeast could provide valuable insights into the proximity effect and conformational constraints when combined with the results of cellulosome analysis (27).

In conclusion, we have successfully constructed an ALL-yeast strain that simultaneously displays 3 kinds of cellulases and have determined that ALL-yeast exhibited higher PASC-degrading activity than MIX-yeast did when the same amount of cellulases was used, which demonstrates the proximity effect. Because the proximity effect increased more drastically when the distances between the cellulases decreased, further increasing the cellulase-display efficiency could enhance the proximity effect and thus cellulose-degrading activity. We expect that combining

proximity-effect studies conducted in various contexts, such as by using arming yeast and cellulosomes, will provide key insights that could facilitate the development of novel strategies for efficiently degrading cellulose and for enabling synergistic and continuous reactions in several other fields.

Summary

Proximity effect is a form of synergistic effects exhibited when cellulases work within a short distance from each other, and this effect can be a key factor in enhancing saccharification efficiency. In this chapter, we evaluated the proximity effect between 3 cellulose-degrading enzymes displayed on the yeast-cell surface, that is, endoglucanase, cellobiohydrolase, and β -glucosidase. We constructed 2 kinds of arming yeasts through genome integration: ALL-yeast, which simultaneously displayed the 3 cellulases (thus, the different cellulases were near each other); and MIX-yeast, a mixture of 3 kinds of single-cellulase-displaying yeasts (the cellulases were far apart). The cellulases were tagged with a fluorescence protein or polypeptide to visualize and quantify their display. To evaluate the proximity effect, we compared the activities of ALL-yeast and MIX-yeast with respect to degrading phosphoric-acid-swollen cellulose after adjusting for the cellulase amounts. ALL-yeast exhibited 1.25 fold or 2.22 fold higher activity than MIX-yeast did at a yeast concentration of OD10 or OD0.1. At OD0.1, the distance between the 3 cellulases was greater than that at OD10 in MIX-yeast, but the distance remained the same in ALL-yeast; thus, the difference between the cellulose-degrading activities of ALL-yeast and MIX-yeast increased (to 2.22 fold) at OD0.1, which strongly supports the proximity effect between the displayed cellulases. Proximity effect was also observed for crystalline cellulose (Avicel). We expect the proximity effect to further increase when enzyme-display efficiency is enhanced, which would further increase cellulose-degrading activity. This arming-yeast technology can also be applied to examine proximity effects in other diverse fields.

References

1. **Klemm D, Heublein B, Fink HP, Bohn A.** 2005. Cellulose: fascinating biopolymer and

- sustainable raw material. *Angew. Chem. Int. Ed. Engl.* **44**:3358–3393.
2. **Carroll A, Somerville C.** 2009. Cellulosic biofuels. *Annu. Rev. Plant. Biol.* **60**:165–182.
 3. **Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS.** 2002. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **66**:506–577.
 4. **Lynd LR, Laser MS, Bransby D, Dale BE, Davison B, Hamilton R, Himmel M, Keller M, McMillan JD, Sheehan J, Wyman CE.** 2008. How biotech can transform biofuels. *Nat. Biotechnol.* **26**:169–172.
 5. **Jalak J, Kurasin M, Teugjas H, Valjamae P.** 2012. Endo-exo synergism in cellulose hydrolysis revisited. *J. Biol. Chem.* **287**:28802–28815.
 6. **Tamaru Y, Karita S, Ibrahim A, Chan H, Doi RH.** 2000. A large gene cluster for the *Clostridium cellulovorans* cellulosome. *J. Bacteriol.* **182**:5906–5910.
 7. **Lamed R, Setter E, Bayer EA.** 1983. Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. *J. Bacteriol.* **156**:828–836.
 8. **Johnson EA, Sakajoh M, Halliwell G, Madia A, Demain AL.** 1982. Saccharification of complex cellulosic substrates by the cellulase system from *Clostridium thermocellum*. *Appl. Environ. Microbiol.* **43**:1125–1132.
 9. **Fierobe HP, Bayer EA, Tardif C, Czjzek M, Mechaly A, Belaich A, Lamed R, Shoham Y, Belaich JP.** 2002. Degradation of cellulose substrates by cellulosome chimeras. Substrate targeting versus proximity of enzyme components. *J. Biol. Chem.* **277**:49621–49630.
 10. **Morais S, Barak Y, Caspi J, Hadar Y, Lamed R, Shoham Y, Wilson DB, Bayer EA.** 2010. Cellulase-xylanase synergy in designer cellulosomes for enhanced degradation of a complex cellulosic substrate. *mBio* **1**:e00285-10.
 11. **Fujita Y, Ito J, Ueda M, Fukuda H, Kondo A.** 2004. Synergistic saccharification, and direct fermentation to ethanol, of amorphous cellulose by use of an engineered yeast strain codisplaying three types of cellulolytic enzyme. *Appl. Environ. Microbiol.* **70**:1207–1212.
 12. **Nakanishi A, Kuroda K, Ueda M.** 2012. Direct fermentation of newspaper after laccase-treatment using yeast codisplaying endoglucanase, cellobiohydrolase, and β -glucosidase. *Renew. Energ.* **44**:199–205.
 13. **Matsui K, Kuroda K, Ueda M.** 2009. Creation of a novel peptide endowing yeasts with acid tolerance using yeast cell-surface engineering. *Appl. Microbiol. Biotechnol.* **82**:105–113.

14. **Zhang YH, Cui J, Lynd LR, Kuang LR.** 2006. A transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. *Biomacromolecules* **7**:644–648.
15. **Nakatani Y, Yamada R, Ogino C, Kondo A.** 2013. Synergetic effect of yeast cell-surface expression of cellulase and expansin-like protein on direct ethanol production from cellulose. *Microb. Cell. Fact.* **12**:66.
16. **Shibasaki S, Ueda M, Iizuka T, Hirayama M, Ikeda Y, Kamasawa N, Osumi M, Tanaka A.** 2001. Quantitative evaluation of the enhanced green fluorescent protein displayed on the cell surface of *Saccharomyces cerevisiae* by fluorometric and confocal laser scanning microscopic analyses. *Appl. Microbiol. Biotechnol.* **55**:471–475.
17. **Sakai A, Shimizu Y, Hishinuma F.** 1990. Integration of heterologous genes into the chromosome of *Saccharomyces cerevisiae* using a delta sequence of yeast retrotransposon Ty. *Appl. Microbiol. Biotechnol.* **33**:302–306.
18. **Yamada R, Taniguchi N, Tanaka T, Ogino C, Fukuda H, Kondo A.** 2010. Cocktail delta-integration: a novel method to construct cellulolytic enzyme expression ratio-optimized yeast strains. *Microb. Cell. Fact.* **9**:32.
19. **Shimoi H, Kitagaki H, Ohmori H, Iimura Y, Ito K.** 1998. Sed1p is a major cell wall protein of *Saccharomyces cerevisiae* in the stationary phase and is involved in lytic enzyme resistance. *J. Bacteriol.* **180**:3381–3387.
20. **Kuroda K, Matsui K, Higuchi S, Kotaka A, Sahara H, Hata Y, Ueda M.** 2009. Enhancement of display efficiency in yeast display system by vector engineering and gene disruption. *Appl. Microbiol. Biotechnol.* **82**:713–719.
21. **Kotaka A, Sahara H, Kuroda K, Kondo A, Ueda M, Hata Y.** 2010. Enhancement of beta-glucosidase activity on the cell-surface of sake yeast by disruption of *SED1*. *J. Biosci. Bioeng.* **109**:442–446.
22. **Matsuoka H, Hashimoto K, Saijo A, Takada Y, Kondo A, Ueda M, Ooshima H, Tachibana T, Azuma M.** 2014. Cell wall structure suitable for surface display of proteins in *Saccharomyces cerevisiae*. *Yeast* **31**:67–76.
23. **Inokuma K, Hasunuma T, Kondo A.** 2014. Efficient yeast cell-surface display of exo- and endo-cellulase using the SED1 anchoring region and its original promoter. *Biotechnol. Biofuels.*

- 7:8.
24. **Krauss J, Zverlov VV, Schwarz WH.** 2012. In vitro reconstitution of the complete *Clostridium thermocellum* cellulosome and synergistic activity on crystalline cellulose. Appl. Environ. Microbiol. **78**:4301-4307.
 25. **Vazana Y, Barak Y, Unger T, Peleg Y, Shamshoum M, Ben-Yehezkel T, Mazor Y, Shapiro E, Lamed R, Bayer EA.** 2013. A synthetic biology approach for evaluating the functional contribution of designer cellulosome components to deconstruction of cellulosic substrates. Biotechnol. Biofuels. **6**:182.
 26. **Molinier AL, Nouailler M, Valette O, Tardif C, Receveur-Brechot V, Fierobe HP.** 2011. Synergy, structure and conformational flexibility of hybrid cellulosomes displaying various inter-cohesins linkers. J. Mol. Biol. **405**:143–157.
 27. **Bae J, Morisaka H, Kuroda K, Ueda M.** 2013. Cellulosome complexes: natural biocatalysts as arming microcompartments of enzymes. J. Mol. Microbiol. Biotechnol. **23**:370–378.

Conclusions

The present study has been carried out to construct an efficient cellulosic biomass degradation system.

In chapter I, we carried out a quantitative proteome analysis of the whole exoproteome of *Clostridium cellulovorans* grown on various carbon sources to clarify the strategy to efficiently degrade cellulosic biomass. We identified 639 proteins in all samples, and of them, 79 proteins were involved in saccharification. While cellulosomal proteins were more abundant than non-cellulosomal proteins, the fold change analysis of the proteins depending on the growth substrates revealed that the amounts of non-cellulosomal proteins changed more drastically. This finding highlights the importance of non-cellulosomal proteins as well as cellulosomal proteins for the efficient degradation of various substrates.

In chapter II, we evaluated the proximity effect between cellulases displayed on *Saccharomyces cerevisiae*. Engineered yeast that simultaneously displays endoglucanase (EG), cellobiohydrolase (CBH), and β -glucosidase (BG) (*i.e.* arming yeast) was constructed to mimic the cellulosomal system of *C. cellulovorans*, and the displayed cellulases were demonstrated to degrade cellulose with a high degree of synergy because of the close distance between them (the proximity effect). The results indicate the effectiveness of the arming yeast and that the increase in the amount of displayed enzymes will further enhance the proximity effect and cellulose-degradation activity.

Learning from the biomass degradation strategy of *C. cellulovorans*, the cellulosomal proteins studied in the exoproteome analysis can be displayed on the yeast cell surface for establishment of CBP and demonstration of the proximity effect. Furthermore, secreting substrate specific non-cellulosomal proteins could also lead to more efficient cellulose degradation system. The methodology throughout this study can be applied to determine and reconstruct the enzyme combinations optimized for the degradation of other biomass substrates.

Acknowledgements

This thesis is submitted by the author to Kyoto University for the Doctor degree of Agriculture. The studies presented here have been carried out under the direction of Professor Mitsuyoshi Ueda in the Laboratory of Biomacromolecular Chemistry, Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, during 2010-2015.

I would like to express my sincere gratitude to Professor Mitsuyoshi Ueda for his continuous guidance, warm patience, and full supports throughout the course of this study.

I deeply thank Associate Professor Kouichi Kuroda for his guidance and critical comments throughout my study. I am also deeply grateful to Assistant Professor Hironobu Morisaka for his experimental suggestions and technical supports. I express my appreciation to Secretary Fukuko Suzuki for her support of everything in my laboratory life. My deep appreciation goes to all the members in this laboratory for their constant encouragement and support throughout this work.

I am grateful to Kashiyaama Scholarship Foundation, JT Asia Foundation, and Iwatani Foundation for financial supports.

Lastly, I would like to sincerely thank my parents and sister for their loves, deep understanding, continuous encouragements, and supports.

Publications

Chapter I

Matsui, K.*, Bae, J.*, Esaka, K., Morisaka, H., Kuroda, K. & Ueda, M. (*:equal contribution) Exoproteome profiles of *Clostridium cellulovorans* on various carbon sources. Appl. Environ. Microbiol. 79(21), 6576-6584 (2013).

Chapter II

Bae, J., Kuroda, K. & Ueda, M. Proximity effect between cellulose-degrading enzymes displayed on yeast cell surface. Appl. Environ. Microbiol. 81(1), 59-66 (2015).

Other publications

Nakanishi, A., Bae, J.G., Fukai, K., Tokumoto, N., Kuroda, K., Ogawa, J., Nakatani, M., Shimizu, S. & Ueda M. Effect of pretreatment of hydrothermally processed rice straw with laccase-displaying yeast on ethanol fermentation. Appl. Microbiol. Biotechnol. **94**(4), 939-948 (2012).

Nakanishi, A., Bae, J., Kuroda, K. & Ueda, M. Construction of novel selection system of optimized combination of CBM (Carbohydrate Binding Module) for biomass using yeast cell surface engineering. AMB Express **2**:56 (2012).

Bae, J., Morisaka, H., Kuroda, K. & Ueda, M. Cellulosome complexes - natural biocatalysts as arming micro-compartments of enzymes. J. Mol. Microbiol. Biotechnol. 23(4-5):370-378 (2013).

Japanese publications

中西昭仁、Bae Jungu、黒田浩一 「バイオリファイナリーからのフェノールプラットフォーム - フェノール化合物への変換」、シーエムシー出版「エコバイオリファイナリー - 脱石油社会へ移行するための環境ものづくり戦略 - 」第8章 229-236 (2010)

中西昭仁、Bae Jungu、黒田浩一、植田充美 「細胞表層工学を利用した最適なセルラーゼカクテルの構築」、シーエムシー出版「バイオマス分解酵素研究の最前線」第 18 章 202-207 (2012)

Bae Jungu、中西昭仁、黒田浩一、植田充美 「リグニン分解酵素表層提示酵母を用いたバイオマス変換」、シーエムシー出版「バイオマス分解酵素研究の最前線」第 29 章 272-277 (2012)

Bae Jungu、中西昭仁、黒田浩一 「効率的バイオリファイナリーに向けたバイオ前処理技術と耐性育種戦略」、生物工学会誌特集、90 (7)、382-385 (2012)

Bae Jungu、黒田浩一、植田充美 「アーミング酵母によるリグニン変換技術とその利用」、シーエムシー出版「リグニン利用の最新動向」第 6 章 219-225 (2013)