博士学位論文

Engineering of a Cellulolytic Bacterium and a Lignocellulose-degrading Enzyme for Utilization of Cellulosic Biomass

セルロース系バイオマス利用に向けたセルロース系バイオマス分解細菌とその分解酵素の改変

2016 年 3 月

市川　俊輔

Table of contents

Chapter 1. Introduction

We heavily depend on products derived from petroleum, generating concerns about global warming and the depletion of fossil resources [1)]. The price of petroleum is highly instable over a decade [2)]. One prediction showed that petroleum will deplete 50 years later without developing new petroleum fields which includes political and technological problems [3)]. For sustainable development of human societies, we should establish strategies to generate the products from renewable and environmentally-friendly materials instead of petroleum.

In this concept, we should produce the products from inexpensive materials, because of the cheaper price of the products from petroleum. One of the candidate materials is biomass, because of the renewability and the abundant amount on the earth [4)]. It is generated by photosynthesis using solar energy and carbon dioxide. The utilization of cellulosic biomass does not contribute the increase of carbon oxide concentration in the atmosphere, does not promote global warming. The conversion of biomass into the products, such as fuels, plastics and other chemicals, is an attractive and feasible strategy, which is called as biorefinery.

We have consumed a large amount of petroleum as liquid fuels [5)]. Liquid fuels can be produced from biomass, called biofuel. Ethanol, which can be mixed with gasoline, has been produced from corn flour in the United States or cane sugar in Brazil. Diesel fuel has been mainly produced from vegetable oil in Europe. Producing a high volume of these fuels from food has created a link between food and fuel prices. Instead of using food, fuel production could come from more abundant and underused resources such as cellulosic biomass [70)].

Large number of companies has tried to produce liquid fuels from cellulosic biomass. But only a few economic biofuel productions have been achieved by some companies, e.g. POET-DSM Advanced Biofuels, E. I. du Pont de Nemours and Company, Abengoa Bioenergy, and Beta Renewables [6-9)]. It seems difficult to add value to biofuel: these companies have to make the fullest possible use of "economies of scale", founding huge biofuel production plants and collecting a large amount of cellulosic biomass such as corn stover from extensive cultivation fields to achieve economic biofuel production.

Japanese government had showed the plan to develop economic biofuel production in Japan. The government said that we should decrease the cost of biofuel production more than three times to establish economic biofuel production [10)]. To achieve the cost reduction, Research Association of Innovative Bioethanol Technology was organized and tried to develop the way to decrease the cost of biofuel production.

However, it could not achieve the cost reduction, and it is disbanded in 2014. Strategic Energy Plan 2014 showed that the government positioned biofuel production from cellulosic biomass as less important in Japan [11]. In contrast to petroleum, the energy density of cellulosic biomass is quite low, the long-distance transportation of cellulosic biomass is not economically favorable. In the project of Research Association of Innovative Bioethanol Technology, a large amount of a cellulosic biomass, nepier grass, was cultivated in Indonesia, and it was transported to Japan. To achieve economic biofuel production, we should collect cellulosic biomass from regional area, and should develop the efficient conversion process of the cellulosic biomass into biofuel.

For the efficient conversion of cellulosic biomass, we must first overcome the problems caused by biomass recalcitrance [54]. Because of that, conventional biofuel production processes should include expensive and complicated steps. Cellulosic biomass is milled, chemically treated and enzymatically treated to produce fermentable sugars. The sugars are converted to ethanol by yeast. The ethanol is distillated to purify it. Japanese government estimated that the price of biofuel should be less than 40 Japanese yen to replace petroleum oils such as gasoline. For efficient biomass degradation, the chemical pretreatment step is necessary, however it costs ca. 40 yen [12]. The price of lignocellulose-degrading enzyme solution has decreased dramatically, however it is ca. 50 Japanese yen in 2012 [10]. Distillation step also significantly increase the cost of biofuel production [13].

For the significant cost reduction, here I propose the simplified biofuel production process in which cellulosic biomass is degraded and converted to liquid fuel simultaneously (Fig. 1 in chapter 4). In this strategy, the soil bacteria, clostridia, which are strong cellulosic biomass degraders, are applied. Because the clostridia do not require the expensive chemical pretreatment and the enzymatic solutions, it is expected that the utilization of clostridia will decrease the cost of biofuel production [14].

For further efficient degradation of cellulosic biomass, the characterization of clostridia seems important. The clostridia secrete multienzyme machines called cellulosome for the degradation of cellulosic biomass [15]. Cellulosomes are composed of numerous kinds of cellulases and related enzyme subunits, which are assembled into the complex by virtue of a unique type of scaffolding subunit (Fig. 1 in chapter 2). The scaffolding subunit is essential: the scaffolding unit deletion mutant of a clostridia can not form the cellulosome, resulting the defect of the degradation ability of cellulosic biomass [16]. The scaffolding subunit contains a single carbohydrate-binding module (CBM), which exhibits recognition of, and strong binding to, crystalline cellulose. It is reported that the CBM largely contributes to the cellulose degradation ability of

cellulosome [17].

Several investigators have shown that the binding of lignocellulose-degrading enzyme to the insoluble substrates such as cellulose is a key step in cellulosic biomass degradation [18-20]. Thus, any effort to develop a kinetic framework for the enzymatic degradation of cellulose must identify and characterize the binding mechanisms that lead to the formation of enzyme–substrate complexes on insoluble cellulose. Key processes, such as adsorption and desorption, reversibility, exchange, and competition, can be determined from binding isotherms and other measures of enzyme–substrate formation [21]. To understand the degradation mechanisms of cellulosome, the binding feature of cellulosomal CBM is an important factor for resolving the binding behavior observed with intact cellulosome. To date, the association constants have been determined for the cellulosomal CBM from the cellulosomes of *Clostridium thermocellum*, *Clostridium cellulolyticum*, *Clostridium cellulovorans*, and *Clostridium acetobutylicum* [22-25]. These studies were performed with purified crystalline cellulose, although cellulosic biomass includes various polysaccharides and morphologically heterogeneous celluloses [26,27]. *Clostridium josui* also produces a cellulosome, but characterization of the corresponding CBM has not been reported [28,29]. To understand the degradation mechanisms of cellulosome, the detail the binding features of the cellulosomal CBM from a clostridia, *C. josui*, to various components of cellulosic biomass was characterized in chapter 2.

To improve the enzymatic activity of lignocellulose-degrading enzymes, random mutagenesis or site-directed mutagenesis have been taken so far. But few researchers have reported examples of mutants showing significantly higher cellulase activity toward insoluble substrates [30]. Most of lignocellulose-degrading enzymes from clostridia are generally composed of one or more catalytic modules and one or more non-catalytic accessory modules, e.g. CBM. CBMs could enhance activity of the enzymes by promoting binding of the catalytic module on the surface of target insoluble substrates [31].

Cellulosic biomass is a recalcitrant material because of its complicated structure containing cellulose, xylan and other polysaccharides [27]. CBMs are highly diverse modules that can be grouped into 71 families on the basis of amino acid sequence similarities, and can discriminate the fine structure of polysaccharides in cellulosic biomass [121]. Thus, it is speculated that replacement of the original CBM in a cellulase with the other CBM should switch its specificity for various polysaccharides. A family 11 CBM from *C. thermocellum* Cel5E preferentially binding to ß-1,4- and ß-1,3-1,4-mixed linkage glucans [32]. A cellulosomal family 3 CBM from *C. josui* mainly binds to microcrystalline cellulose [33]. A family 6 CBM from *Clostridium stercorarium* xylanase 11A binds to xylan [34]. To provide new insights into the engineering of lignocellulose-degrading

enzyme, a lignocellulose-degrading enzyme Cel5E from *C. thermocellum* was engineered with these different CBMs and their activities are characterized in chapter 3. I propose the possibility that the replacement of CBMs will be one of the ways to improve the degradation ability of lignocellulose-degrading enzymes for cellulosic biomass (Fig. 6 in chapter 3).

The clostridia produces slight amount of ethanol, which is the most popular compounds as biofuel. Because of their cellulosic biomass degradation and biofuel production ability, the utilization of the clostridia open the possibility of the simplified biofuel production process, called consolidated bioprocessing. It is suggested the consolidated bioprocessing could reduce the cost of biofuel production significantly [14]. In fact, ethanol has not attractive feature as fuel, high energy consumption of distillation step to purify it, for example [13]. We can collect insoluble fuel compounds, e.g. hydrocarbon, higher alcohol, fatty acid ethyl ester, without distillation. The bacterial production pathway of such water-insoluble fuel compounds was reported, enabling the challenge to engineer cellulolytic bacteria to produce the water-insoluble fuel compounds [35]. In chapter 4, I report the metabolic engineering of the clostridia, *C. thermocellum*, to produce higher alcohol from cellulose in which the costly processes of the conventional biofuel production, the supplementation of lignocellulose-degrading enzymes and distillation of fuel compounds, are eliminated (Fig. 1 in chapter 4).

Chapter 2. Characterization of a lignocellulose-degrading enzyme: binding characters of cellulosome from *Clostridium josui*

2.1. Abstract

To understand the lignocellulose degradation activity of the lignocellulose-degrading enzyme complex "cellulosome" from a cellulolytic bacterium *Clostridium josui*, a carbohydrate-binding module of the scaffoldin (*Cj*CBM3) was characterized. *Cj*CBM3 shows the binding to crystalline cellulose, non-crystalline cellulose and soluble polysaccharides. The binding isotherm of *Cj*CBM3 to acid-swollen cellulose is best fitted by the Langmuir two-site model, suggesting that there are two *Cj*CBM3 binding sites on acid-swollen cellulose with different affinities. The second site shows lower affinity and larger binding capacity, suggesting that the cellulosome is directly targeted to the cellulose surface with high affinity, where larger amounts of the cellulosome bind to cellulose with low affinity.

2.2. Introduction

Cellulosic biomass is essentially recalcitrant and insoluble. In biodegradation of the biomass, enzymes hydrolyzing celluloses and other plant cell wall polysaccharides bind to the surface of these insoluble materials [18,19]. Therefore, these enzymes often harbor carbohydrate-binding modules (CBMs) to facilitate attachment to the surface of their substrates [31]. Currently, based on amino acid sequence similarities, 71 families of CBMs, including two missing families, are described in the latest update of the CAZy database [121].

Cellulosomes are multicellulolytic enzyme complexes that largely contribute to lignocellulose degradation ability of cellulolytic bacteria, e.g. Clostridia [36-38]. Cellulosomes contain a non-catalytic scaffolding protein, or scaffoldin, consisting of multiple cohesin domains and at least one CBM. The scaffoldin CBM belongs to family 3a and plays an important role in targeting the cellulosome to its substrate (Fig. 1). To understand fully the functions of the cellulosome, characterization of this family 3a CBM is needed. To date, the association constants have been determined for the scaffoldin CBM3 from the cellulosomes of *Clostridium thermocellum*, *Clostridium cellulolyticum*, *Clostridium cellulovorans*, and *Clostridium acetobutylicum* [22-25]. These studies were performed with purified crystalline cellulose, although cellulosic biomass includes various polysaccharides and morphologically heterogeneous celluloses [26,27]. *Clostridium josui* also produces a cellulosome, but characterization of the corresponding CBM3 has not been reported [28,29]. In this study, we report that a scaffoldin CBM3 from *C. josui* (*Cj*CBM3) bind to wide range of substrates, i.e. crystalline celluloses, a non-crystalline cellulose and soluble polysaccharides. Furthermore, we show that *Cj*CBM3 recognizes two different structures on acid-swollen cellulose, which is a morphologically heterogeneous cellulose [39].

2.3. Materials and Methods

2.3.1. Plasmid construction

The DNA fragments coding for *Cj*CBM3 and for a family 1 CBM from Cel7A of *Trichoderma reesei* NBRC 31329 (*Tr*CBM1) were amplified from genomic DNA of the respective organisms using the primer pair listed in Table 1. The resulting products from polymerase chain reaction were ligated into pQE30 (Qiagen, Hilden, Germany) or pRSET/CFP (Invitrogen, CA) [28], respectively. Mutants of *Cj*CBM3, *Cj*CBM3-W119A, and *Cj*CBM3-Y66A/W119A, were constructed by site-directed mutagenesis [40].

2.3.2. Protein purification

The constructs were expressed in *Escherichia coli* JM109 or BL21(DE3) (TOYOBO, Kyoto, Japan). Transformed *E. coli* cells were grown overnight at 37 °C in Luria–Bertani broth supplemented with 50 μg/mL ampicillin. Once the cells reached an optical density of about 0.5 at 600 nm, isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 1 mM. Cells then were collected by centrifugation at 6000 × g and disrupted by sonication. The expressed proteins in cell-free extracts were purified using Ni-NTA agarose resin (Qiagen) according to the manufacturer's instructions. Purified proteins were dialyzed with 20 mM potassium phosphate buffer, pH 7.0, overnight at 4 °C. Dialyzed proteins then were centrifuged at 18,000 × g for 2 min; supernatants were aliquoted and stored at -80°C. The final protein preparations were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), and protein concentration was determined by measuring absorbance at 280 nm. The molar extinction coefficients for *Cj*CBM3, W119A, Y66A/W119A, *Tr*CBM1-cyan fluorescence protein (CFP), and CFP (34410, 28860, 27520, 30760, and 24800 $M^{-1}cm^{-1}$, respectively) were predicted from the tryptophan, tyrosine, and cysteine content of the proteins [41].

2.3.3. Substrate preparation

Carboxymethyl cellulose (Sigma-Aldrich, St. Louis, MO, USA), xyloglucan (Megazyme, Wicklow, Ireland), birch wood xylan (Fluka, Buchs, Switzerland), oat-spelt xylan (Fluka), chitin (Wako Pure Chemical, Tokyo, Japan), methyl cellulose (Nacalai tesque, Kyoto, Japan), hydroxypropyl cellulose (Hercules Incorporated, Wilmington, DE, USA), soluble starch (Merck KGaA, Darmstadt, Germany), β-glucan from barley (Sigma-Aldrich), guar (Sigma-Aldrich), lichenan (Sigma-Aldrich), laminarin (Sigma-Aldrich), pectin

(Nacalai tesque), and highly crystalline cellulose from *Halocynthia* (Synaptech, Yamanashi, Japan) were used as commercially available substrates. Mannan was kindly provided by Dr. Araki of Mie University, Japan. Bacterial microcrystalline cellulose was prepared from cultures of *Gluconacetobacter xylinus* subsp. *xylinus* (NBRC 13772) as described previously [42]. Ball-milled cellulose was prepared by ball mill processing of the KC flock (Nippon Paper Chemicals, Tokyo, Japan) with distilled water for 72 h at 4 °C. Acid-swollen cellulose was prepared by phosphoric acid treatment of Avicel (Funakoshi, Tokyo, Japan) [43]. Prepared substrates were stored at 4 °C.

2.3.4. Macroarray assay

The CBM binding of various polysaccharides was detected by macroarray assay [44,45]. Polysaccharides were applied as 1 μL aliquots to untreated mixed nitrocellulose sheets (MF-Millipore membrane filters) (Millipore, Billerica, MA, USA) in a 5-fold dilution series, yielding spots containing 0.32-1000 ng polysaccharides on the membranes. Sheets were left to dry at room temperature for 30 min before blocking for 1 h with 5% (w/v) skim milk protein in 5 × phosphate-buffered saline prepared from a 10 × stock solution containing 1.36 M NaCl, 380 mM KCl, 80 mM Na_2HPO_4, and 15 mM KH_2PO_4, pH 7.2. The nitrocellulose sheets were incubated with 1.25 μg/mL of the appropriate protein in 5 × phosphate-buffered saline at room temperature for 1 h, extensively washed with 5 × phosphate buffered saline, and then incubated with a 5000-fold dilution of anti-His horseradish peroxidase conjugate (Qiagen) in 5 × phosphate-buffered saline. After washing with distilled water, sheets were incubated in freshly prepared horseradish peroxidase substrate (0.5% 3,3'-diaminobenzidine tetrahydrochloride, 0.005% H_2O_2) to detect CBM binding. The reaction was stopped by washing the nitrocellulose sheets with distilled water.

2.3.5. Affinity gel electrophoresis

The affinity of *Cj*CBM3 for soluble polysaccharides was determined by affinity gel electrophoresis, as described by Araki *et al*. [46], using xyloglucan, carboxymethyl cellulose, methyl cellulose, mannan, and guar at a concentration of 0.1% (w/v) in a polyacrylamide gel. The electrophoresis was carried out for 30 min at room temperature in native polyacrylamide gels containing 12% (w/v) acrylamide. Bovine serum albumin was used as a nonbinding control protein.

2.3.6. Adsorption isotherm measurement

The binding characteristics of *Cj*CBM3 to insoluble cellulose were determined by adsorption isotherm measurements. All measurements were carried out at room temperature in 1.5 mL Sumilon Proteo-save SS tubes (Sumitomo Bakelite, Tokyo, Japan) containing 1.2-12 μM *Cj*CBM3 mixed with 17 μg acid-swollen cellulose in 20 mM potassium phosphate buffer (pH 7.0) to a final aqueous volume of 0.3 mL. Each solution was gently mixed and incubated at room temperature for 30 min to allow the adsorption system to equilibrate; during the incubation, the solution was mixed after 15 min. The samples then were centrifuged at 25 °C and 18,000 × g for 2 min to remove protein-covered substrates. The samples then were centrifuged at 18,000 × g for 1 min, and the clear supernatant was collected and used to determine the unbound protein concentration at an absorbance of 280 nm. Each measurement was done in triplicate. Bound protein concentrations were calculated using unbound protein concentration and total protein concentration. Langmuir adsorption isotherms were drawn for each value and were evaluated using ORIGIN software (MicroCal, Northampton, MA, USA) [47].

Equation (1) is the Langmuir one-site model.

$$[PC] = \frac{[PC]_{max} Ka[P]}{1 + Ka[P]} \qquad (1)$$

where [PC] is the concentration of bound protein (moles · g cellulose^{-1}), [P] is the concentration of unbound protein (molar), $[PC]_{max}$ is the total concentration of available binding sites (moles · g cellulose^{-1}), and *K*a is the equilibrium association constant (liters · mol^{-1}). Equation (2) is the Langmuir two-site model that assumes heterogeneity, such that two independent classes of binding sites are present.

$$[PC] = \frac{[PC_1]_{max} Ka_1[P]}{1 + Ka_1[P]} + \frac{[PC_2]_{max} Ka_2[P]}{1 + Ka_2[P]} \qquad (2)$$

where $[PC_1]_{max}$ is the total concentration of high-affinity binding sites characterized by the equilibrium association constant Ka_1, and $[PC_2]_{max}$ is the total concentration of low-affinity binding sites characterized by the equilibrium association constant Ka_2.

2.4. Results

2.4.1. Preparation of recombinant protein solutions

*Cj*CBM3, mutant proteins of *Cj*CBM3, *Ti*CBM1-CFP, and CFP were prepared as described in the Materials and Methods section. As *Ti*CBM1 was not retained in the soluble fraction, *Ti*CBM1 was produced as a fusion to cyan fluorescence protein (CFP), permitting successful purification. *Ti*CBM1-CFP bound to Avicel with an association constant ($1.4 \cdot 10^5$ M^{-1}) similar to values previously reported for the parent protein [48,49], indicating that *Ti*CBM1-CFP maintains the binding characteristics of *Ti*CBM1. The final preparations showed a single band on SDS-PAGE (Fig. 2). Circular dichroism spectra showed that the structure of *Cj*CBM3 was not affected by the amino acid substitutions (data not shown).

2.4.2. Affinity of *Cj*CBM3 for various polysaccharides

The macroarray assay detected binding of various polysaccharides by *Ti*CBM1-CFP and *Cj*CBM3 [44,45] (Fig. 3). Both CBMs are classified as type-A CBMs, meaning that these CBMs can bind to the surface of crystalline cellulose [31]. In fact, *Ti*CBM1-CFP exhibited binding only to bacterial microcrystalline cellulose and ball-milled cellulose, polymers that have partly crystalline regions [42]. The fusion protein failed to show binding to acid-swollen cellulose or soluble polysaccharides, suggesting that these polymer types do not include crystalline regions permitting interaction with type-A CBMs. In control reactions, CFP did not exhibit detectable binding to any of the tested polysaccharides (data not shown). In contrast, *Cj*CBM3 showed binding not only to bacterial microcrystalline cellulose and ball-milled cellulose, but also to acid-swollen cellulose, methyl cellulose, xyloglucan, chitin, mannan, and guar. The bindings of *Cj*CBM3 for soluble polysaccharides were determine again by affinity gel electrophoresis. *Cj*CBM3 binding to xyloglucan and methyl cellulose, but not to carboxymethyl cellulose, also was detected in aqueous conditions [46] (Fig. 4).

By analogy to other CBM3s, the aromatic amino acids tyrosine 66 (Y66) and tryptophan 119 (W119) of *Cj*CBM3 are predicted to have a central role in binding to crystalline cellulose surfaces [50] (Fig. 5 and 6). Mutation of W119 alone or together with Y66 decreased the binding activities against crystalline cellulose and soluble substrates, suggesting that the crystalline cellulose binding site was equal to the soluble substrate binding site (Fig. 3).

2.4.3. Adsorption isotherm analysis of *Cj*CBM3 to cellulose

Because the morphology of cellulose is heterogeneous, the binding of *Cj*CBM3 to cellulose may include more than one type of binding event. To identify the distinct binding events, adsorption isotherm analysis was performed. The adsorption isotherm of *Cj*CBM3 to acid-swollen cellulose was fitted better by the Langmuir two-binding-site model compared to a single-binding-site model, indicating that there are two *Cj*CBM3 binding sites on acid-swollen cellulose with different affinities (Fig. 7). At the first binding site, the association constant was $4.1 \cdot 10^6$ M^{-1}, a value comparable to those of other scaffoldin CBMs [23,24] (Table 2). An association constant could not be determined for the second binding site, indicating that the value was too small to resolve by this model. This low-affinity binding is not negligible, because the total concentration of low-affinity binding sites, $[PC_2]_{max}$, was large (Table 2, Fig. 7). Given that higher ion intensity conditions are known to inhibit hydrophilic interactions, we repeated the binding assay in the presence of 1 M sodium chloride. Under high-salt conditions, the total concentration of high-affinity binding sites, $[PC_1]_{max}$, was not altered. On the other hand, the total concentration of low-affinity binding sites, $[PC_2]_{max}$, was drastically decreased in the presence of 1 M NaCl, suggesting that hydrophilic interactions contribute to binding in the second binding event (Table 2).

2.5. Discussion

*Cj*CBM3 and *Ti*CBM1 are classified as type-A CBMs, which can bind to surface of crystalline cellulose. Consistent with this classification, *Ti*CBM1-CFP exhibited binding to bacterial microcrystalline cellulose and ball-milled cellulose, polymers known to include crystalline regions. However, *Cj*CBM3 exhibited binding not only to bacterial microcrystalline cellulose and ball-milled cellulose but also to non-crystalline acid-swollen cellulose and soluble polysaccharides (Fig. 3). Three-dimensional structures of *C. thermocellum* CipA CBM3 and *C. cellulolyticum* CipC CBM3, proteins that have amino acid sequence similarity with *Cj*CBM3, reveal the presence of a flat surface involved in binding to crystalline cellulose surfaces, and a groove located opposite to the flat surface [50,51] (Fig. 6). To determine whether this groove can be a non-crystalline polysaccharide binding site in *Cj*CBM3, W119A and Y66A/W119A mutants were constructed [19] (Fig. 5). Tyrosine 66 and tryptophan 119 were in the flat surface. If the groove were the binding site of soluble substrates, these mutations would have no effect on affinity. However, both mutants exhibited decreased binding activities against crystalline cellulose and soluble substrates, suggesting that *Cj*CBM3 binds to the soluble substrates at the crystalline cellulose binding site (Fig. 3).

The binding of *Cj*CBM3 to acid-swollen cellulose includes a low-affinity interaction with a large number of binding sites (Table 2). Nonspecific binding between proteins and polysaccharides is not consistent with the low-affinity binding, as demonstrated by the observation that bovine serum albumin does not show significant binding (Fig. 7). On the other hand, the adsorption isotherm of *Cj*CBM3 to highly crystalline cellulose from *Halocynthia* was clearly saturated, suggesting that the low affinity binding is the interaction between *Cj*CBM3 and heterologous structure of acid-swollen cellulose [52](Fig. 7). Sugimoto *et al*. reported the occurrence of a steric exclusion effect in Cel7A upon CBM1 binding of cellulose [53]. These authors also indicated that Hill's equation provided a better fit for their assay's binding isotherms [53]. In the present study, the isotherm of *Cj*CBM3 was better fitted by the Langmuir two-site model than by Hill's equation. A similar two-site binding phenomenon was reported for CBM2 from *Cellulomonas fimi*; however, the nature of this low-affinity binding site on cellulose remains unclear [47]. The low-affinity binding of *Cj*CBM3 was not observed in the presence of 1 M sodium chloride, suggesting that hydrophilic interactions contribute to the binding in the second binding event of *Cj*CBM3 (Table 2).

We conclude that *Cj*CBM3 shows binding not only to crystalline cellulose but also to non-crystalline cellulose and soluble polysaccharides through the crystalline cellulose binding site. The binding isotherm of

*Cj*CBM3 to acid-swollen cellulose was best fitted by the Langmuir two-site model. Hydrophilic interactions appear to contribute to the binding events at the second site. Under physiological conditions, the second site showed reduced affinity and increased binding capacity. These data indicate that the cellulosome is directly targeted to the cellulose surface with high affinity, where larger amounts of the cellulosome bind to cellulose with low affinity.

2.6. Figures and tables

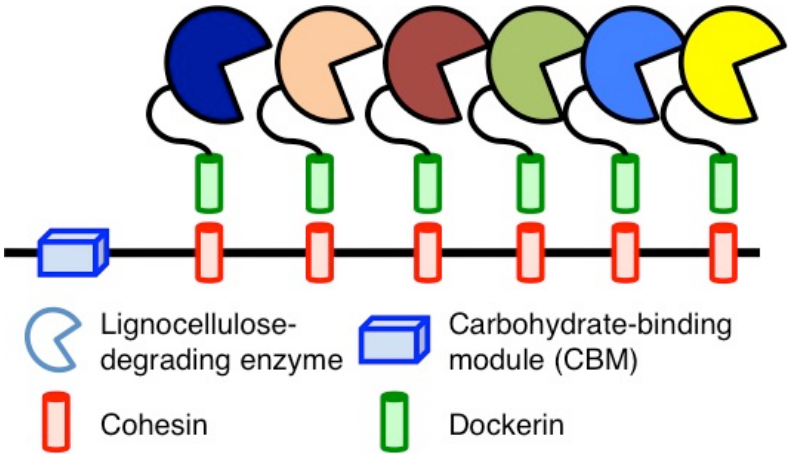


Figure 1. A model of cellulosome structure from *Clostridium josui*

Cellulosome largely contribute to lignocellulose degradation ability of Clostridia. The bacteria produce lignocellulose-degrading enzymes appended to dockerin modules, and a scaffold protein named scaffoldin which includes carbohydrate-binding module and cohesin modules. Strong interactions between the cohesin modules and the dockerin modules form multicellulolytic enzyme complex, cellulosome. The family 3 carbohydrate-binding module plays an important role in targeting the cellulosome to its substrate.

Table 1. Primers used for the DNA constructs

Construct	Primer	Direction
*Cj*CBM3	5’-AAGGATCCGCAGCTGATACTGGCG-3’	Forward
	5’-AAGCTTTCAACCATTAGGTGTTGAACCA-3’	Reverse
*Ti*CBM1-CFP	5'-GGATCCCCTACCCAGTCTCACTACGG-3’	Forward
	5'-GAATTCGGGGGAGGTCAGGCACTGAGAGTAGTAAGG-3’	Reverse
*Cj*CBM3W119A	5'-AACGCTATGGATGTTACTTCAAAGGT-3'	Forward
	5'-TGAAGTAACATCCATAGCGTTGTT-3'	Reverse
*Cj*CBM3Y66A/W119A	5'-GACGCTTCTAATTTTGATCAATCAAA-3'	Forward
	5'-TTGATCAAAATTAGAAGCGTCGTT-3'	Reverse

Incorporated *Bam*HI, *Hin*dIII, and *Eco*RI restriction sites are underlined.

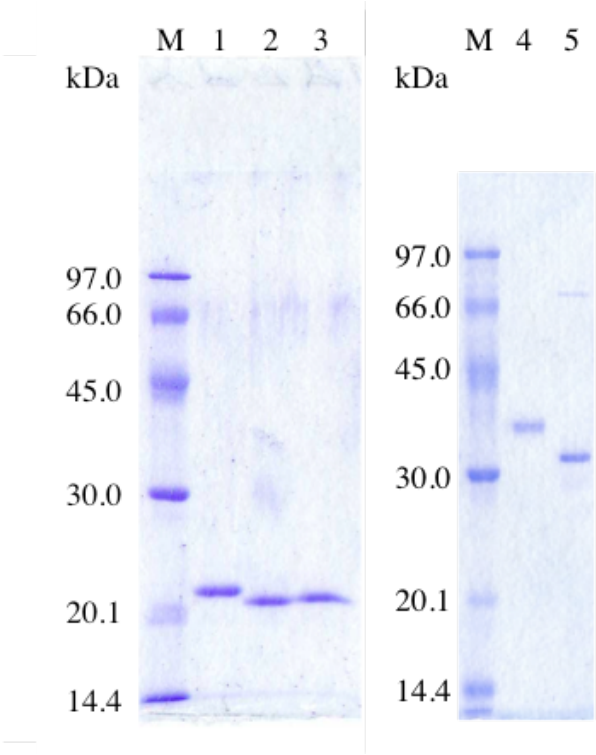


Figure 2. SDS-PAGE of *Cj*CBM3, *Cj*CBM3 mutants, *Ti*CBM1-CFP, and CFP.

The gel was stained with Rapid Stain CBB Kit (Nakalai tesque). M, molecular weight marker (Amersham™ LMW Calibration Kit; GE Healthcare, Buckinghamshire, UK). Lane 1, *Cj*CBM3; lane 2, *Cj*CBM3W119A; lane 3, *Cj*CBM3Y66A/W119A; lane 4, *Ti*CBM1-CFP; lane 5, CFP.

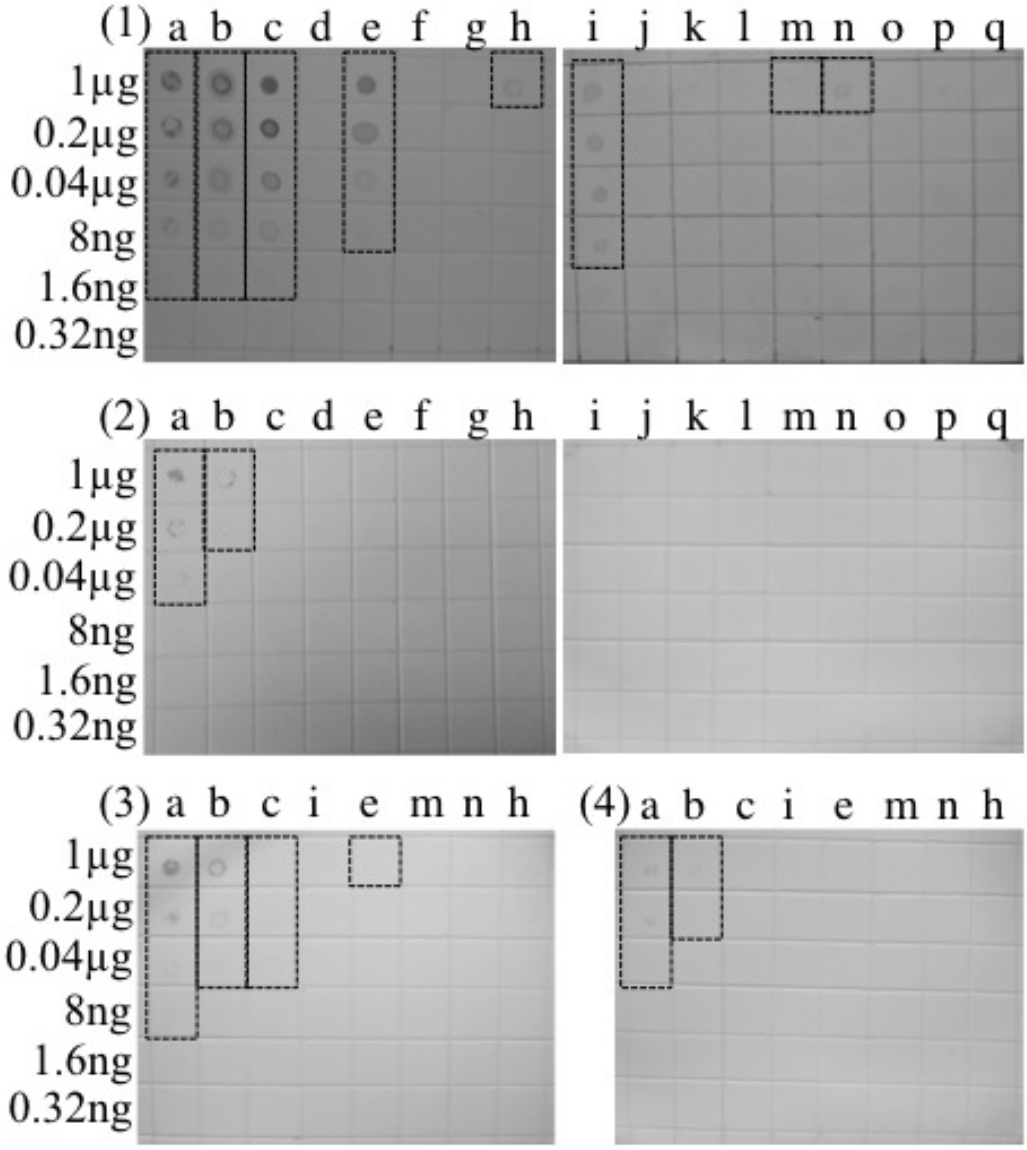


Figure 3. CBM macroarray assays.

CBM macroarray assay of the binding of (1) *Cj*CBM3, (2) *Ti*CBM1-CFP, (3) *Cj*CBM3 W119A, and (4) *Cj*CBM3 Y66A/W119A to dilution series of bacterial microcrystalline cellulose (lane a), ball-milled cellulose (lane b), acid-swollen cellulose (lane c), carboxymethyl cellulose (lane d), xyloglucan (lane e), birchwood xylan (lane f), oatspelt xylan (lane g), chitin (lane h), methyl cellulose (lane i), hydroxypropyl cellulose (lane j), soluble starch (lane k), β-glucan (lane l), mannan (lane m), guar (lane n), lichenan (lane o), laminarin (lane p), and pectin (lane q). Samples were applied to nitrocellulose in a dilution series yielding amounts as indicated, then probed with 1.2 μg/mL of the respective protein. CFP did not exhibit binding to any of the tested polysaccharides (data not shown).

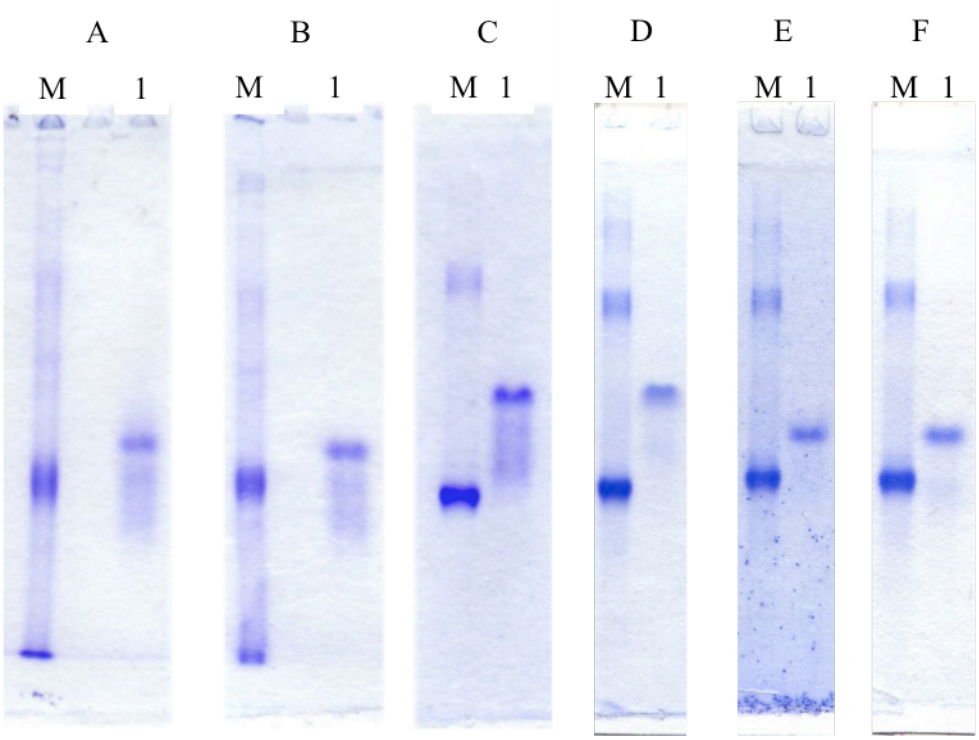


Figure 4. Affinity gel electrophoresis of *Cj*CBM3.

The affinity of *Cj*CBM3 for soluble cellulose was analyzed by affinity gel electrophoresis in a gel containing carboxymethyl cellulose (B), xyloglucan (C), methyl cellulose (D), guar (E), mannan (F), or no added polysaccharide (A). M, bovine serum albumin; lane 1, *Cj*CBM3.

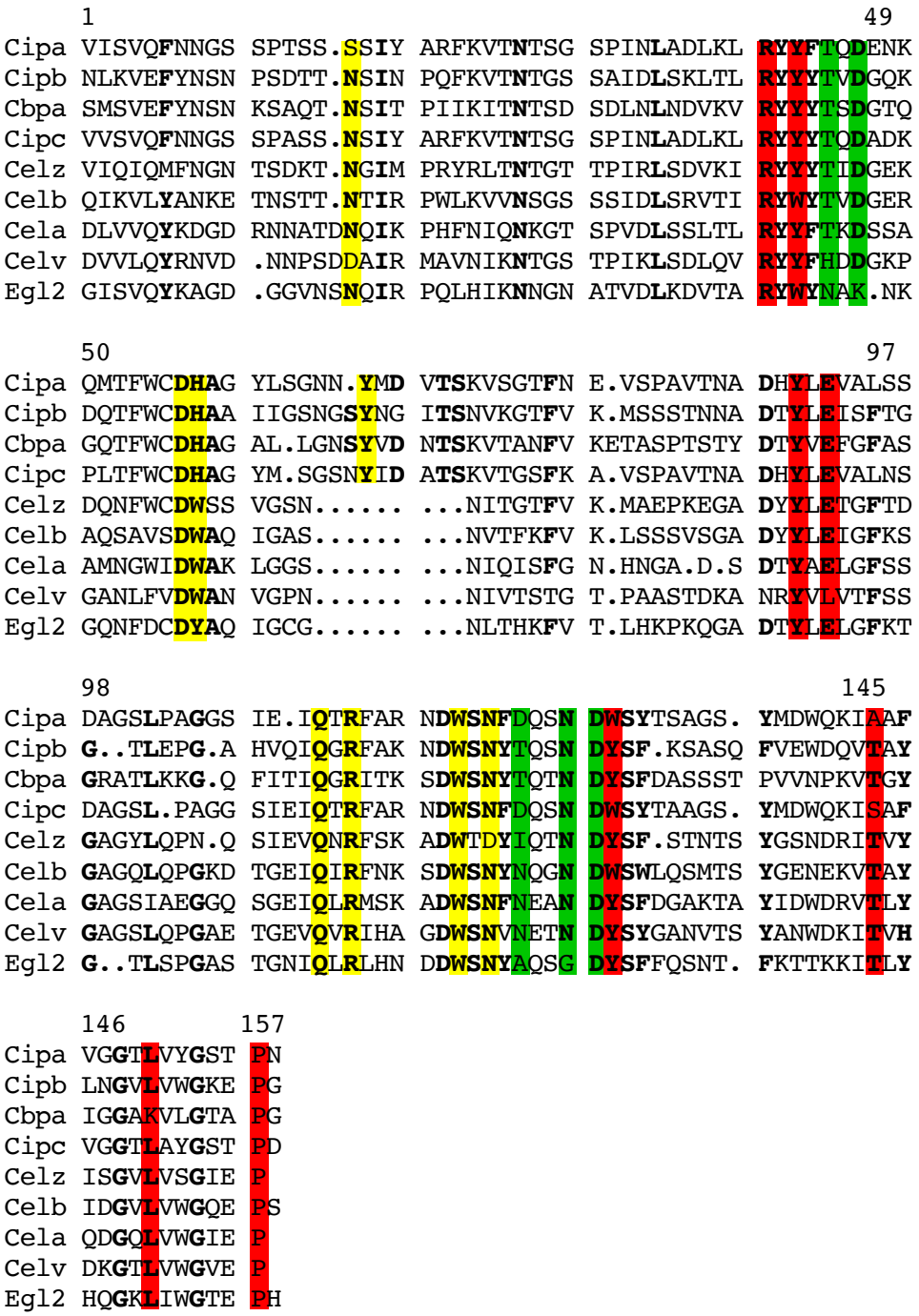


Figure 5. Structure-based sequence alignment of selected CBMs from family 3 [50].

Regions of secondary structure are marked and labeled. Conserved residues are shown in bold type. Conserved surface-exposed residues, which are not involved in the stabilization of loops of secondary structure elements, cluster into different discrete regions, and have been color-coded as follows: yellow, conserved surface residues which form a planar strip at the bottom of the CBM molecule; red, conserved surface residues which form a shallow groove on the top of the CBM molecule; green, residues which

interact with the calcium ion, where the non-conserved residue at position 123 interacts with the calcium via its backbone carbonyl group. The CBMs shown in the figure are from the following proteins: CipA, the scaffoldin subunit from *C. josui* (Swiss-Prot accession number O82830); CipB, the scaffoldin subunit from *C. thermocellum* (Swiss-Prot accession number Q06851); CipC, the scaffoldin subunit of *C. cellulolyticum* (GenBank accession number U40345); CbpA, the scaffoldin subunit from *C. cellulovorans* (Swiss-Prot accession number P38058); CelZ, cellulase CelZ (Avicelase I) from *C. stercorarium* (Swiss-Prot accession number P23659); CelB, cellobiohydrolase/endocellulase from *Caldocellum saccharolyticum* (Swiss-Prot accession number P10474); CelA, endo-glucanase (cellulaseA) from *Bacillus lautus* (Swiss-Prot accession number P29719); CelV, endo-glucanaseV from *Erwinia carotovora* (Swiss-Prot accession number S39962); Egl2, endo-glucanase (cellulase Egl2) from *B. subtilis* (Swiss-Prot accession number A27198).

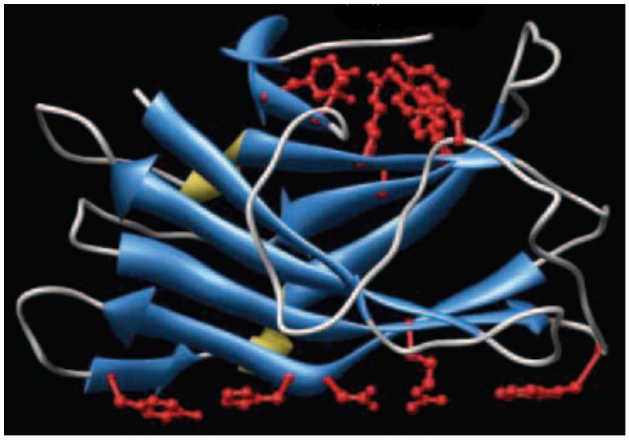


Figure 6. The three dimensional structure of *Ct*CBM3 [50)]

β-sheet fragments appear as blue strips. The amino acids at the flat (bottom) or the groove (top) binding sites are indicated in red. Tyrosine 66 and tryptophan 119 were in the flat surface in CBM3 from *C. josui*.

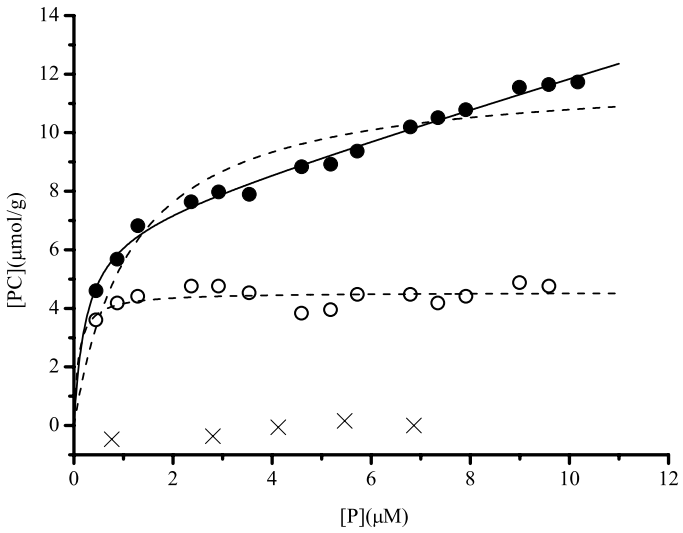


Figure 7. Adsorption isotherms of *Cj*CBM3 to celluloses.

The binding of *Cj*CBM3 to acid-swollen cellulose (●), of *Cj*CBM3 to crystalline cellulose from *Halocynthia* (○), and of bovine serum albumin to acid-swollen cellulose (×) are presented. The vertical axis is the concentration of bound CBM and the horizontal axis is the concentration of unbound CBM. The dashed lines show curve fitting using the Langmuir one-site model; the solid line shows curve fitting using the Langmuir two-sites model.

Table 2. Binding parameters of *Cj*CBM3 with acid-swollen cellulose

buffer	$[PC_1]_{max}$ (μmol/g cellulose)	Ka_1 (M^{-1})	$[PC_2]_{max}$ (μmol/g cellulose)	Ka_2 (M^{-1})
20mM KPB	6.88 (0.70)	4.1×10^6 (1.4×10^6)	5.3×10^4 (1.4×10^8)	1.0×10^1 (2.5×10^4)
20mM KPB + 1M NaCl	6.93 (0.41)	2.8×10^7 (8.1×10^6)	3.1×10^0 (0.3×10^0)	3.8×10^5 (1.6×10^5)

$[PC]_{max}$ is the total concentration of available binding sites, and *K*a is the equilibrium association constant. Standard errors are shown in parentheses. The parameters of the low-affinity binding, $[PC_2]_{max}$ and Ka_2, could not be determined because the standard errors were too large. Note that the low-affinity binding (demonstrated by the small magnitude of the association constant in 20 mM KPB) is not observed upon the addition of 1 M NaCl.

Chapter 3. Engineering of a lignocellulose-degrading enzyme: the replacement of carbohydrate-binding modules switches the specificity of the activity of an endoglucanase Cel5E from *Clostridium thermocellum*

3.1. Abstract

Most cellulases contain carbohydrate-binding modules (CBMs) that largely contribute to their activity for insoluble substrates. *Clostridium thermocellum* Cel5E is an endoglucanase having xylanolytic activity. The Cel5E originally has a family 11 CBM preferentially binding to ß-1,4- and ß-1,3-1,4-mixed linkage glucans. In this study, we replaced the CBM with a different type of CBM, either a family 3 microcrystalline cellulose-directed CBM from *Clostridium josui* scaffoldin, or a family 6 xylan-directed CBM from *Clostridium stercorarium* xylanase 11A. Chimeric endoglucanases showed enhanced activity that was affected by CBM binding specificity. These chimeric enzymes could efficiently degrade milled cellulosic biomasses, such as corn hull, because of heterologous components in the plant cell wall, indicating that diverse CBMs play roles in degradation of lignocellulosic materials.

3.2. Introduction

To ensure a reliable future source of energy and raw materials, the utilization of sustainable biomass has considerable advantages over petroleum-based energy sources. Cellulosic biomass, obtained as agricultural byproducts and industrial residues, is an abundant, inexpensive, and renewable source of sugars, and is a desirable feedstock for the sustainable production of fuels and chemical products. Cellulose and hemicellulose in cellulosic biomass can be degraded by microbial enzymes into oligosaccharides or monosaccharides, which are fermentable to biofuels. However, this degradation step is a bottleneck in the process that is preventing establishment of biofuel technology [54]. To improve the efficiency of the degradation step, steps to increase the enzymatic activity, for example by random mutagenesis or site-directed mutagenesis, have been taken. But few researchers have reported examples of mutants showing significantly higher cellulase activity toward insoluble substrates [30].

Lignocellulose-degrading enzymes are generally composed of one or more catalytic modules and one or more non-catalytic accessory modules, *e.g*. a carbohydrate-binding module (CBM). CBMs enhance activity of the enzymes by promoting binding of the catalytic module on the surface of target insoluble substrates [31]. Wang *et al*. reported that artificial *Fervidobacterium nodosum* Rt17-B1 Cel5As fused with CBMs showed increased cellulase activity for microcrystalline cellulose [55].

Lignocellulose is a recalcitrant material because of its complicated structure containing cellulose, xylan and other polysaccharides [27]. CBMs are highly diverse modules that can be grouped into 71 families on the basis of amino acid sequence similarities, and can discriminate the fine structure of polysaccharides in plant cell walls. Thus, it is speculated that replacement of the original CBM in a cellulase with another CBM should switch its specificity for various polysaccharides.

An endoglucanase, Cel5E, from *Clostridium thermocellum* CelH (*Ct*Cel5E) shows endoglucanase and endoxylanase activities. A family 11 CBM (*Ct*CBM11) is appended to the native *Ct*Cel5E [32]. Here, we show that replacement of *Ct*CBM11 by *Clostridium josui* CipA carbohydrate-binding module 3 or *Clostridium stercorarium* xylanase 11A carbohydrate-binding module 6 changes the specificity of *Ct*Cel5E activity.

3.3. Materials and Methods

3.3.1. Preparation of purified *Ct*Cel5E

The genes encoding *Ct*Cel5E and *Ct*Cel5E-*Ct*CBM11 were amplified from genomic DNA of *Clostridium thermocellum* stain 132 and inserted into vector pQE30 (Qiagen, Hilden, Germany), yielding p*Ct*Cel5E and p*Ct*Cel5E-*Ct*CBM11, respectively [32]. Genes for a family 3 CBM from *Clostridium josui* CipA (*Cj*CBM3), and a family 6 CBM from *Clostridium stercorarium* xylanase 11A (*Cs*CBM6) were also amplified and connected downstream of the *Ct*Cel5E gene to generate *Ct*Cel5E-*Cj*CBM3 and *Ct*Cel5E-*Cs*CBM6 sequences [33,34] (Fig. 1). These fusion genes were inserted in pQE30, yielding p*Ct*Cel5E-*Cj*CBM3 and p*Ct*Cel5E-*Cs*CBM6, respectively. The primers used in this study are shown in Table 1. These plasmids were expressed in *Escherichia coli* JM109 and the expressed proteins were purified using a Ni-NTA agarose resin according to the manufacturer's instructions (Qiagen). Purified proteins were dialyzed against 50 mM potassium phosphate buffer (pH 7.0).

3.3.2. Enzymatic activity measurement

*Ct*Cel5Es (1.3 ㎛ M) was mixed with either 1.5 % carboxymethyl cellulose (Sigma-Aldrich, St. Louis, MO, USA), 0.5 % ball-milled cellulose, 1 % Avicel (Funakoshi, Tokyo, Japan), 0.5 % oat spelt xylan (Fluka, Buchs, Switzerland) or 1% milled corn hulls. The reactions were carried out in 300 ㎛ L of 50 mM potassium phosphate buffer (pH 7.0) at 60 ˚C. The amounts of released reducing sugars were determined by the dinitrosalicylic acid method [56]. One unit of activity was defined as moles of glucose equivalents per ㎛ mole of enzyme per minute.

3.3.3. Binding analysis

*Ct*Cel5Es (1.3 ㎛ M) was mixed with either 1.5 % carboxymethyl cellulose, 0.5 % ball-milled cellulose, 1 % Avicel, 0.5 % oat spelt xylan. The mixtures were incubated in 300 ㎛ L of 50 mM potassium phosphate buffer (pH 7.0) at 60 ˚C for 30 minutes. After the incubation, the mixtures were centrifuged and the supernatants were applied to SDS-PAGE. The band sizes of *Ct*Cel5Es were quantified with ImageJ software to determine the amount of *Ct*Cel5Es binding to the substrates after the incubation.

3.3.4. Substrate preparations

Carboxymethyl cellulose (Sigma-Aldrich, St. Louis, MO, USA), microcrystalline cellulose (Funakoshi, Tokyo, Japan) and oat-spelt xylan (Fluka, Buchs, Switzerland) were used as commercially available substrates. Ball-milled cellulose was prepared by ball mill processing of 3 g of the KC flock (Nippon Paper Chemicals, Tokyo, Japan) with 1 L of distilled water for 72 h at 4 °C. Corn hull was kindly provided by Dr. Khanok Ratanakhanokchai in King Mongkut's university of Technology Thonburi, Thailand. The corn hull was milled to a final particle size of 1 mm by knife milling.

3.4. Results

3.4.1. Preparation of *Ct*Cel5Es

The genes encoding *Ct*Cel5E, *Ct*Cel5E-*Ct*CBM11, *Ct*Cel5E-*Cj*CBM3 and *Ct*Cel5E-*Cs*CBM6 were expressed in *E. coli* JM109, and the resultant proteins were purified with Ni-NTA agarose resin. The purified recombinant protein solutions appeared as a single band following sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Fig. 2).

3.4.2. The enzymatic activities of *Ct*Cel5E for celluloses and a xylan

The relative activities of *Ct*Cel5E-CBMs against substrates are shown in Figure 3. The activities were compared with that of *Ct*Cel5E. There was no significant difference in activity among *Ct*Cel5E-*Ct*CBM11, *Ct*Cel5E-*Cs*CBM6, *Ct*Cel5E catalytic module for carboxymethyl cellulose, indicating that these CBMs do not affect enzymatic activity for this soluble substrate [31]. However, *Ct*Cel5E-*Cj*CBM3 showed somewhat lower activity, indicating microcrystalline cellulose-directed CBMs, such as family 1 CBM, affect catalytic activity toward soluble substrates as described by Nozaki *et al* [57].

For oat spelt xylan, *Ct*Cel5E-*Ct*CBM11 and *Ct*Cel5E-*Cs*CBM6 showed 1.5 and 2.3 times higher activities than *Ct*Cel5E. For ball-milled cellulose, the relative activities of *Ct*Cel5E-*Ct*CBM11, *Ct*Cel5E-*Cj*CBM3 and *Ct*Cel5E-*Cs*CBM6 were 9.9, 2.9 and 5.3 times higher than that of *Ct*Cel5E. The relative activities of *Ct*Cel5E-*Ct*CBM11, *Ct*Cel5E-*Cj*CBM3 and *Ct*Cel5E-*Cs*CBM6 were 2.3, 3.1 and 1.3 times higher than that of *Ct*Cel5E for Avicel cellulose.

3.4.3. The binging amount of *Ct*Cel5Es to celluloses and a xylan

The binding amounts of *Ct*Cel5E-*Ct*CBM11, *Ct*Cel5E-*Cj*CBM3 and *Ct*Cel5E-*Cs*CBM6 to ball-milled cellulose, Avicel cellulose or oat spelt xylan were analyzed with SDS-PAGE. The band of *Ct*Cel5E-*Cj*CBM3 was not observed after the incubation with ball-milled cellulose, indicating that all of *Ct*Cel5E-*Cj*CBM3 bound to ball-milled cellulose in the reaction condition (Fig. 4 and 5). *Ct*Cel5E-*Cs*CBM6 bound to oat spelt xylan, but *Ct*Cel5E-*Ct*CBM11 and *Ct*Cel5E-*Cj*CBM3 did not. Only 7 % and 16 % of *Ct*Cel5E-*Ct*CBM11 bound to Avicel cellulose and ball-milled cellulose, respectively. The binding amount of these *Ct*Cel5Es also did not correlate with their catalytic activity for Avicel cellulose and ball-milled cellulose (Fig. 5).

3.4.4. The enzymatic activities of *Ct*Cel5E for a cellulosic biomass, milled corn hull

Sixty-three μg of reducing sugars was released from 3 mg of milled corn hulls by the original enzyme, *Ct*Cel5E-*Ct*CBM11 after the reaction for 24 hours. In the case of the CBM-replaced enzymes *Ct*Cel5E-*Cj*CBM3 and *Ct*Cel5E-*Cs*CBM6, 82 and 91 μg of reducing sugars were released from the milled corn hull for 24 hours, indicating that CBM replacements could enhance catalytic activity of the endoglucanase for lignocellulosic materials (Fig. 6).

3.5. Discussion

CBMs are highly diverse modules and can discriminate the fine structure of polysaccharides in lignocellulose. Thus, it is speculated that replacement of the original CBM in a cellulase with another CBM should switch its specificity for various polysaccharides. In this study, we replaced the CBM11 appended to *Ct*Cel5E with CBM3 from *Clostridium josui* scaffoldin, or CBM6 from *Clostridium stercorarium* xylanase 11A.

There was no significant difference in activity among these *Ct*Cel5Es for carboxymethyl cellulose, in contrast, *Ct*Cel5E-*Ct*CBM11, *Ct*Cel5E-*Cj*CBM3 and *Ct*Cel5E-*Cs*CBM6 showed higher activities against insoluble substrates (ball-milled cellulose, Avicel, and oat spelt xylan, respectively) than the catalytic domain *Ct*Cel5E without any CBM (Fig. 3). *Ct*Cel5E-*Ct*CBM11 showed the highest activity for ball-milled cellulose. *Ct*Cel5E-*Cj*CBM3 and *Ct*Cel5E-*Cs*CBM6 demonstrated higher activities than *Ct*Cel5E-*Ct*CBM11 for Avicel and oat spelt xylan, respectively, indicating that replacement of the CBM switches the specificity of the catalytic module *Ct*Cel5E for insoluble substrates.

CBMs bind to insoluble polysaccharides and enhance the catalytic activity of their enzymes for insoluble substrates [31]. It has been reported that *Cj*CBM3 is not able to bind oat spelt xylan, whereas *Ct*CBM11 and *Cs*CBM6 do [50,58-60]. The affinity of *Ct*CBM11 and *Cs*CBM6 to oat spelt xylan is 1.7×10^{1} M^{-1} and 2.3×10^{6} M^{-1}, respectively [50,59,60]. This suggests that the quite low affinity of *Ct*CBM11, which should cause low binding of *Ct*Cel5E-*Ct*CBM11 to oat spelt xylan, confers a small increase in *Ct*Cel5E activity (Fig. 5).

The *Ct*Cel5E-*Ct*CBM11, *Ct*Cel5E-*Cj*CBM3 and *Ct*Cel5E-*Cs*CBM6 enzymes showed improvement of *Ct*Cel5E activity for ball-milled cellulose and Avicel cellulose (Fig. 3). The *Ct*Cel5E-*Ct*CBM11 and *Ct*Cel5E-*Cs*CBM6 showed higher activities than *Ct*Cel5E-*Cj*CBM3 for ball-milled cellulose. It has also been reported that *Ct*CBM11, *Cj*CBM3 and *Cs*CBM6 bind to insoluble cellulose substrates [34,58,60]. Structural studies have reported that each CBM shows a specific binding character due to its unique binding site structure: family 3 CBM has a flat binding site that preferentially interacts with the surface of crystalline cellulose [50]. On the other hand, *Ct*CBM11 has a cleft structure that binds to cellulose, xylan, ß-1,3-1,4-glucan, lichenan and glucomannan [60]. *Cs*CBM6 also has a cleft structure that binds to cellulose and xylan [61]. These cleft structures can capture a cellulose chain in amorphous region of cellulose microfibrils. Ball milling of cellulose leads to a considerable reduction in cellulose crystallinity, thus it converts crystalline cellulose into

amorphous cellulose [62]. Therefore, we hypothesized that larger binding amounts of *Ct*Cel5E-*Ct*CBM11 and *Ct*Cel5E-*Cs*CBM6 to ball-milled cellulose lead to the higher activities. Unexpectedly, the binding amounts of *Ct*Cel5E-*Ct*CBM11 and *Ct*Cel5E-*Cs*CBM6 were lower than that of *Ct*Cel5E-*Cj*CBM3 (Fig. 4 and 5). The binding amount of these *Ct*Cel5Es also did not correlate with their catalytic activity for Avicel cellulose (Fig. 5). The structural differences of CBMs should confer different recognition characteristics for the surface structure of insoluble polysaccharides [63]. *Ct*CBM11, *Cj*CBM3 and *Cs*CBM6 might carry *Ct*Cel5E to different sites in ball-milled cellulose and Avicel cellulose; thus, fusion enzymes may result in different ratios of improved *Ct*Cel5E activity (Fig. 3).

Lignocellulose is a recalcitrant material because of its complicated structure including cellulose, xylan and other polysaccharides [27]. Because each CBM targets different structural sites in plant cell walls, the CBM replacements may enhance the catalytic activity of *Ct*Cel5E in complex lignocellulosic materials, as was the case in the improvement of activity toward Avicel cellulose and xylan. Inedible parts of corn are one of the candidates as lignocellulose materials for biofuel production [64]. The amount of reducing sugars released from milled corn hulls by the CBM-replaced enzymes *Ct*Cel5E-*Cj*CBM3 and *Ct*Cel5E-*Cs*CBM6 were both greater than the amount released by the original enzyme, *Ct*Cel5E-*Ct*CBM11 (Fig. 6). This result suggests that CBM replacements could enhance catalytic activity of the endoglucanase for lignocellulosic materials including various types of celluloses [65].

To improve the efficiency of lignocellulose degradation, attempts have been made to increase the enzymatic activity, but only a few successful results have been reported. We showed that replacement of *Ct*CBM11 with *Cj*CBM3 or *Cs*CBM6 changed the specificity of *Ct*Cel5E activity. These CBM-replaced chimeric Cel5Es showed increased activity for Avicel cellulose, xylan and a type of cellulosic biomass, milled corn hulls. CBMs having different substrate specificities focus the catalytic module on different sites of cellulosic biomass and improve their catalytic activity toward polysaccharides in cellulosic biomass.

3.6. Figures and tables

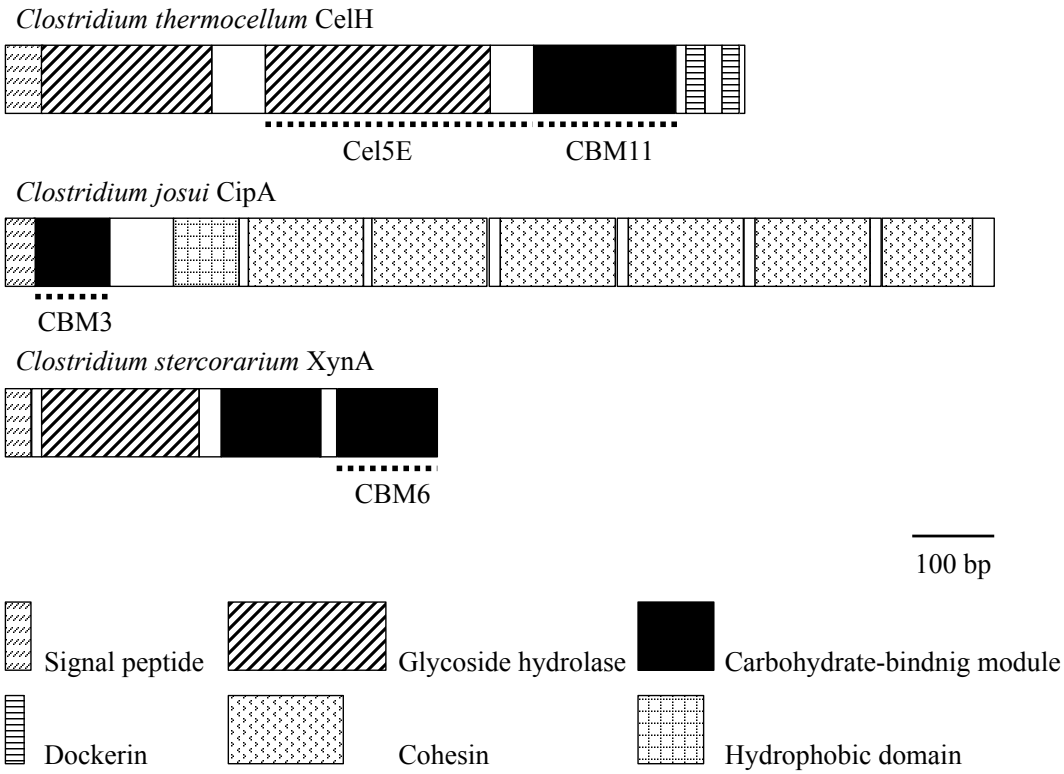


Figure 1. Isolated cellulase and carbohydrate-binding module genes

Cel5E from *C. thermocellum* CelH was used as a model enzyme. CBM11 from *C. thermocellum* CelH, CBM3 from *C. josui* CipA and CBM6 from *C. stercorarium* XynA were also isolated and fused to the C-terminal end of *Ct*Cel5E [32-34]. Dotted lines show isolated region of the genes.

name	sequence (5' → 3')
*Ct*CelHF	GGATCCTCACCAAAGGCTGTCGACCC
*Ct*CelHR1	AAGCTTTTAATATGTTCCGGGATTAAATAGTGC
*Ct*CelHR2	AAGCTTTTAAGTAGCACCAATCAGCTTGATATT
*Ct*CelHR3	GCATGCATATGTTCCGGGATTAAATAGTGC
*Cj*CBM3F	GCATGCGCAGCTGATACTGGCG
*Cj*CBM3R	AAGCTTTTATCAACCATTAGGTGTTGAACCATATACA
*Cs*CBM6F	GCATGCAGTTCACCAGTGCCTGCACC
*Cs*CBM6R	GGTACCTTACTTAAGTTCCTGATTT

Table 1. Primers used for the DNA constructs

Incorporated *Bam*HI, *Sph*I, *Kpn*I and *Hin*dIII restriction sites are underlined.

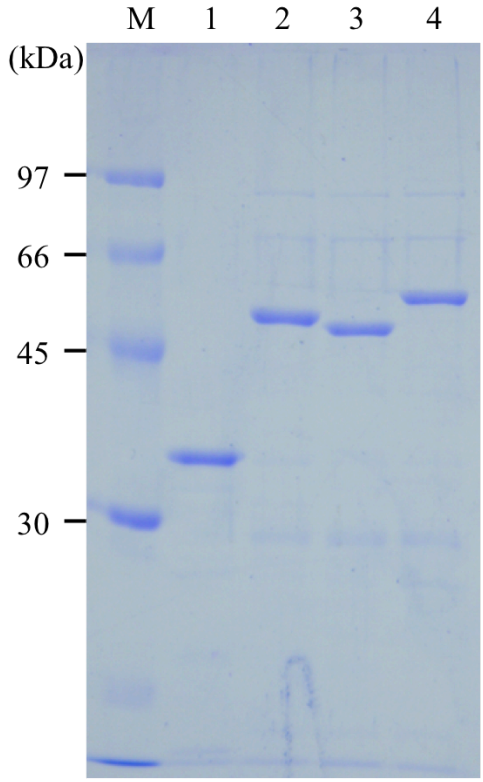


Figure 2. Purities of recombinant enzymes

M, molecular weight marker (LMW Electrophoresis Calibration Kit; GE Healthcare Life Sciences, Buckinghamshire, UK); lane 1, *Ct*Cel5E; lane 2, *Ct*Cel5E-*Cj*CBM3; lane 3, *Ct*Cel5E-*Cs*CBM6; lane 4, *Ct*Cel5E-*Ct*CBM11. The gel was stained with Rapid Stain CBB Kit (Nakalai Tesque, Kyoto, Japan).

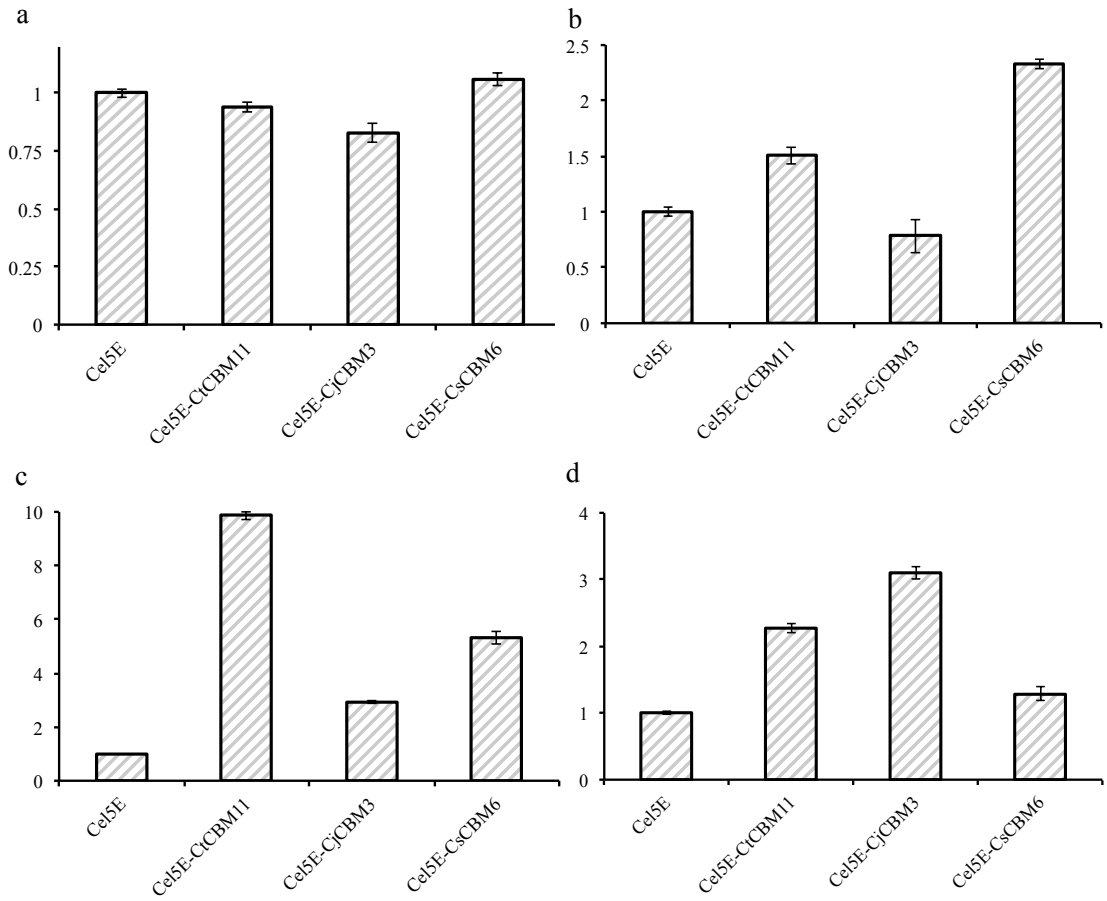


Figure 3. Relative activities of *Ct*Cel5E with and without CBMs

Enzymatic activities of *Ct*Cel5E, *Ct*Cel5E-*Ct*CBM11, *Ct*Cel5E-*Cj*CBM3 and *Ct*Cel5E-*Cs*CBM6 were analyzed with (a) 1.5 % carboxymethyl cellulose, (b) 0.5 % oat spelt xylan, (c) 0.5 % ball-milled cellulose and (d) 1 % Avicel cellulose at 60 ˚C. The activities were compared with that of *Ct*Cel5E. Bars show standard error. Each measurement was done in triplicate.

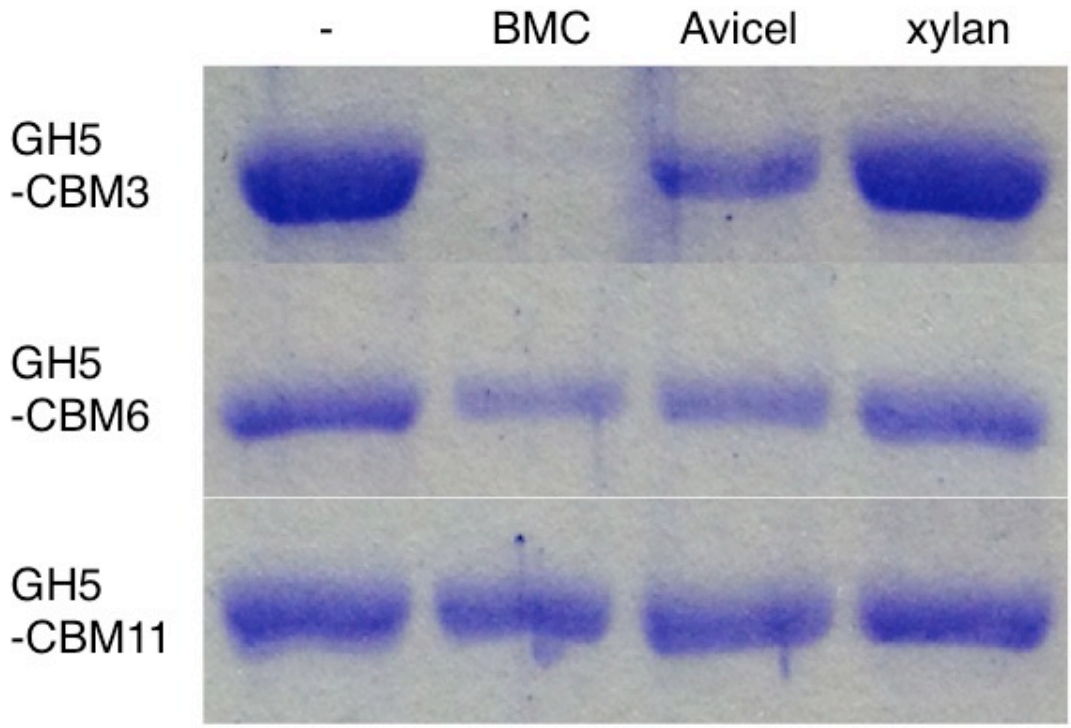


Figure 4. The binding of *Ct*Cel5Es to the celluloses and xylans.

*Ct*Cel5E-*Cj*CBM3, *Ct*Cel5E-*Cs*CBM6 and *Ct*Cel5E-*Ct*CBM11 were incubated with ball-milled cellulose (BMC), Avicel cellulose, oat spelt xylan or without substrate. After centrifugation of the incubated mixtures, supernatants were applied to SDS-PAGE. The binding amounts of the *Ct*Cel5Es were determined by the band sizes which were quantified with ImageJ software.

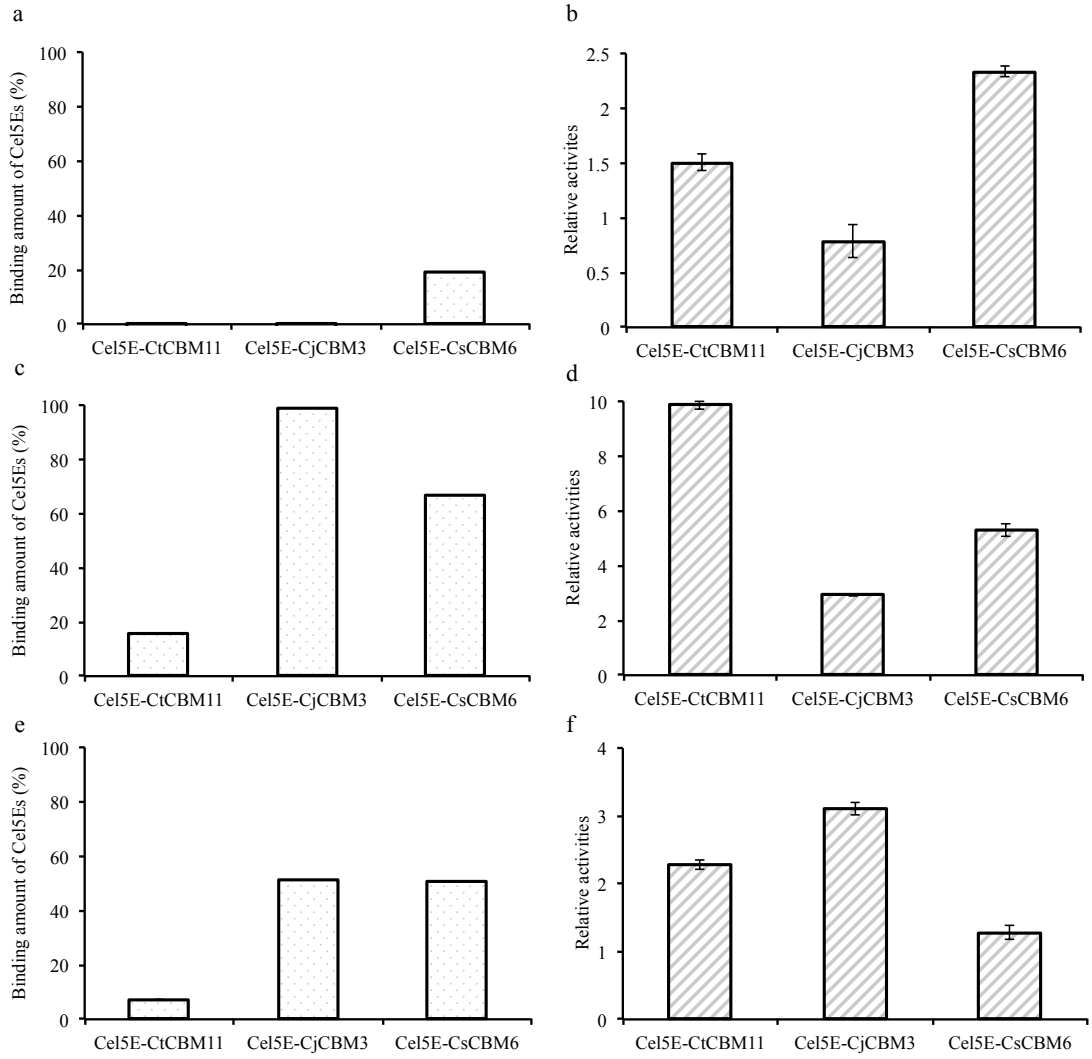


Figure 5. Comparison of relative activities and binding ratio of *Ct*Cel5Es for insoluble polysaccharides.

The amounts of *Ct*Cel5E-*Ct*CBM11, *Ct*Cel5E-*Cj*CBM3 and *Ct*Cel5E-*Cs*CBM6 bound during enzymatic activity measurements were quantified (Fig. 4). The amount bound and the relative activities for oat spelt xylan (a, b), ball-milled cellulose (c, d) and Avicel cellulose (e, f) are represented. The activities were compared with that of *Ct*Cel5E. Bars show standard error. Notice that the amount bound did not correlate with activity.

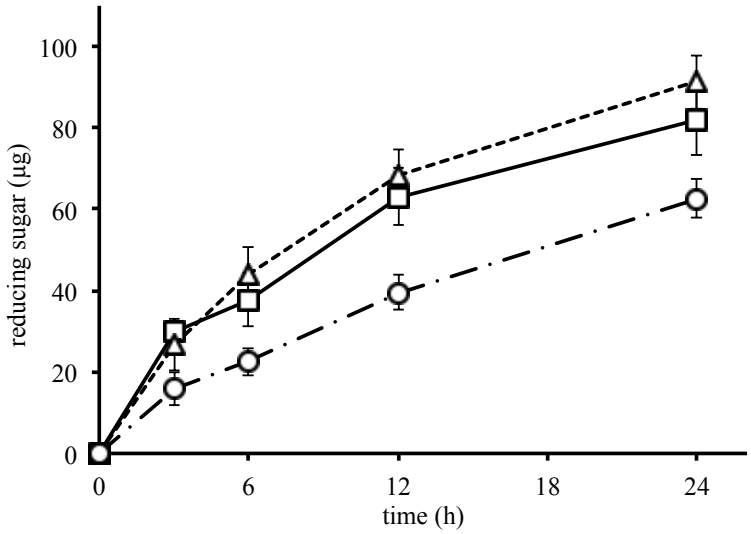


Figure 6. Catalytic activities of *Ct*Cel5Es for milled corn hulls

Reducing sugar release of *Ct*Cel5E-*Ct*CBM11 (○), *Ct*Cel5E-*Cj*CBM3 (□) and *Ct*Cel5E-*Cs*CBM6 (▸) from 3 mg milled corn hulls were quantitated. Bars show standard error.

Chapter 4. Engineering of a cellulolytic bacterium: higher alcohols production from cellulose by the engineered *Clostridium thermocellum*

4.1. Abstract

Achieving economic biofuel production from cellulosic biomass will require significant cost reductions. Enzymatic degradation of cellulosic biomass and distillation of water-soluble fuel compounds substantially increase the cost of biofuel production. Consolidated bioprocessing is a strategy to circumvent expensive biofuel production steps. *Clostridium thermocellum* is a promising bacterium for consolidated bioprocessing because it does not require the supplementation of lignocellulose-degrading enzymes. To produce water-insoluble fuel compounds, *C. thermocellum* was engineered to express a fatty acyl-ACP reductase and an aldehyde-deformylating oxygenase from *Synechococcus elongatus* PCC 7942. Expression of the aldehyde-deformylating oxygenase gene was clearly detected, whereas only slight expression of the fatty acyl-ACP reductase gene was detected. Cells expressing the fatty acyl-ACP reductase and the aldehyde-deformylating oxygenase accumulated fatty aldehydes (higher alcohol precursors). After cultivation with cellulose, the higher alcohols, decanol and dodecanol, were detected in the organic solvent phase of the culture broth, indicating that the strain secreted the higher alcohols. These results suggest that the engineered *C. thermocellum* strain, expressing fatty acyl-ACP reductase and aldehyde-deformylating oxygenase genes, directly produces and secretes higher alcohols from cellulose without the supplementation of cellulases. The higher alcohols can be collected by phase separation.

4.2. Introduction

Heavy dependence on products derived from petroleum has generated concerns about global warming and the depletion of fossil resources [66]. Biomass is a sustainable resource, and its utilization has been described as carbon neutral because it does not contribute to global warming. Cellulosic biomass is a particularly attractive candidate sustainable resource material because of its abundance [67].

Numerous studies have examined the use of cellulosic biomass in biorefinery processes for the production of petroleum products, including fuel compounds. The development of economic processes for biofuel production from cellulosic biomass will require significant cost reductions, however [68]. Some synthetic biology approaches for biofuel production have utilized microorganisms such as *Escherichia coli* and *Saccharomyces cerevisiae* to produce fuel compounds from sugars [69,70]. Lignocellulose-degrading enzymes are used to convert cellulosic biomass into fermentable sugars; however, the efficiency of the enzymatic degradation process is low [54] (Fig. 1).

Simplified "consolidated bioprocessing" strategies have been reported, in which microorganisms secrete enzymes for producing fermentable sugars from cellulosic biomass and simultaneously ferment them into fuel compounds [14,71-75]. The consolidated bioprocessing approach is promising because it eliminates the need to add lignocellulose-degrading enzymes, which significantly increases the cost of biofuel production. Lignocellulose-degrading enzymes such as cellulases can be expressed in engineered strains of *E. coli* or *S. cerevisiae*, but their lignocellulose-degradation activity is low. The efficient production of lignocellulose-degrading enzymes by non-cellulolytic microorganisms remains a challenging task [73,76].

The clostridia are anaerobic soil bacteria, and some species can degrade cellulosic biomass without chemical pretreatment or the addition of lignocellulose-degrading enzymes [77-79]. These microorganisms are promising for use in consolidated bioprocessing applications because of their ability to degrade cellulosic biomass and produce ethanol [80,81]. Despite their high lignocellulose-degrading ability, only a few studies have examined biofuel production using genetically engineered cellulolytic clostridia. Heterologous expression of *Zymomonas mobilis* pyruvate decarboxylase and alcohol dehydrogenase in *Clostridium cellulolyticum* was shown to result in increased ethanol production [82]. Another reports described a genetically engineered *C. cellulolyticum* and *Clostridium thermocellum* strain that produces isobutanol from cellulose [83,84]. Metabolic engineering studies indicated that co-cultivation of *C. thermocellum* with engineered *Thermoanaerobacterium saccharolyticum* leads to increased ethanol production [85].

To collect water-soluble fuel compounds such as ethanol and butanol from culture broth, a distillation step is required, which also significantly increases the cost of biofuel production [13,86] (Fig. 1). Furthermore, ethanol and butanol are toxic and repress growth of the microorganisms. To produce higher concentrations of ethanol, researchers have attempted to enhance the ethanol tolerance of *C. thermocellum* [87-89]. In contrast to these alcohols, water-insoluble compounds such as aliphatic hydrocarbons are less toxic to *C. thermocellum* [90]. Additionally, due to their insolubility in water, it is expected that these compounds could be more easily collected from the culture broth by phase separation [91].

It has been reported that the expression of just two enzyme genes, fatty acyl-acyl carrier protein (ACP) reductase (ACR) and aldehyde-deformylating oxygenase (ADO), are sufficient to engineer *E. coli* strain to produce higher alcohols and hydrocarbons [35]. It is expected that introducing the genes encoding ACR and ADO into *C. thermocellum* would enable the organism to produce higher alcohols and hydrocarbons.

Because wild type *C. thermocellum* can utilize cellulosic biomass, the engineered *C. thermocellum* might be able to convert it into higher alcohols and hydrocarbons directly, without the supplementation of cellulases [14]. Additionally, the higher alcohols and hydrocarbons could potentially be collected without distillation because these compounds are water-insoluble [35,91] (Fig. 1). To verify this hypothesis, in this study, *C. thermocellum* was engineered with ACR and ADO genes to produce water-insoluble fuel compounds from cellulose.

4.3. Materials and Methods

4.3.1. Strains and growth conditions

Clostridium thermocellum DSM 1313 was anaerobically cultured in CTFUD medium containing 3 g/L of sodium citrate tribasic dehydrate, 1.3 g/L of $(NH_4)_2SO_4$, 1.5 g/L of KH_2PO_4, 130 mg/L of $CaCl_2$ $2H_2O$, 500 mg/L of $_L$-cysteine-HCl, 11.56 g/L of 3-morpholinopropanesulfonic acid, 2.6 g/L of $MgCl_2$ $6H_2O$, 1 mg/L of $FeSO_4$ $7H_2O$, 4.5 g/L of yeast extract, and 1 mg/L of resazurin (pH 7.0), with cellobiose or cellulose at 55 °C [92)]. Transformed *C. thermocellum* cells were cultured in CTFUD medium supplemented with 24 μg/mL of thiamphenicol. *Escherichia coli* JM109 and BL21(DE3) (TOYOBO) were cultured in Luria broth supplemented with 50 μg/mL of ampicillin or 20 μg/mL of chloramphenicol at 37 °C.

4.3.2. Plasmid construction

The plasmid pNW33N was obtained from the Bacillus Genetic Stock Center. Primers used in this study are listed in Table 1. To remove the repeat region in pNW33N, the DNA fragment was amplified by PCR using primers 62 and 63 and then ligated. The resultant plasmid, designated pMU102, contained an origin of replication that functions in *C. thermocellum* [92,93)]. The ACR genes from *Synechococcus elongatus* PCC 7492 (WP_011242364.1), *C. acetobutilycum* ATCC 824 (WP_010965585.1), *C. cellulovorans* ATCC 35296 (WP_010074197.1), and *Thermosynechococcus elongatus* BP-1 (NP_682102.1) and the ADO gene from *S. elongatus* PCC 7492 (WP_011378104.1) were amplified by PCR using primers 35 and 36, 152 and 153, 156 and 157, 169 and 170, and 37 and 38, respectively [35,94)]. The genomic DNA of *T. elongatus* BP-1 was kindly provided by Dr. Tabata of Kazusa DNA Research Institute, Japan. The resulting products were then ligated into pMU102 (Fig. 2). The promoter region of the cellobiose phosphorylase gene and the terminator region of the hypoxanthine phosphoribosyl transferase gene from *C. thermocellum* were amplified by PCR using primers 102 and 103 and 104 and 105, respectively [85)]. The resulting products were inserted upstream of the ACR gene or downstream of the ADO gene, respectively. A ribosome-binding site and a histidine-tag sequence from pQE30 (QIAGE) were amplified by PCR using primers 39 and 40 and 41 and 42, respectively. The resulting products were inserted upstream of the ACR and ADO genes, respectively.

4.3.3. Transformation of *C. thermocellum*

The plasmids for *C. thermocellum* transformation were extracted and purified from *E. coli* BL21 (DE3) [95]. An aliquot (2 mL) of a *C. thermocellum* DSM 1313 stock culture was inoculated into 100 mL of CTFUD medium with 0.5% cellobiose and cultivated at 55 °C without stirring until reaching an optical density at 600 nm (OD_{600}) of 0.8. A total of 50 mL of the culture broth was centrifuged in a 50-mL plastic centrifuge tube at 6,500 × *g* for 12 min at 4 °C. The supernatant was discarded by decantation, with care taken not to disturb the pellet. The pellet was washed three times with autoclaved pure water. The washed cells were suspended in 100 μL of autoclaved pure water. A total of 20 μL of cell suspension and 1 μg of plasmid were mixed in a 1-mm electroporation cuvette. A square electrical pulse was applied using a Gene Pulser Xcell Electroporation System (Bio-Rad). The voltage and pulse length were 1500 V and 1.5 ms, respectively. The pulsed cells were suspended in 5 mL of CTFUD medium with 0.5% cellobiose. The cells were incubated at 51 °C in a dry bath incubator for 18 h. A total of 100 μL of the recovered cell culture was plated onto CTFUD solid medium containing 1% agarose and 6 μg/mL thiamphenicol. The cells were incubated at 55 °C for 3-5 days [92]. The transformed cells were isolated and cultured in CTUFD medium with 0.5% cellobiose and 6 μg/mL of thiamphenicol. The culture broth was analyzed using PCR to confirm that the isolated transformants harbored the appropriate plasmids.

4.3.4. Gene expression analysis

Expression of the ACR and ADO genes in transformed *C. thermocellum* strains was analyzed by Western blotting. The transformed *C. thermocellum* strains were cultured in CTFUD with 24 μg/mL of thiamphenicol at 55 °C [95]. The cells were lysed and the proteins were resolved by SDS-PAGE and electrotransferred onto PVDF membranes using a CompactBLOT apparatus (ATTO). ACR and ADO were detected on the membranes using an anti–His-tag antibody conjugated with horseradish peroxidase (QIAGEN) and ECL Western Blotting Detection Regents (GE Healthcare Life Sciences).

4.3.5. Fatty aldehyde composition analysis.

Transformed *C. thermocellum* was cultured in 25 mL of CTFUD medium. The cells were harvested and washed three times with 1 mL of purified water. The cell pellets were placed on ice and then vacuum-dried for 24 h at 25 °C, after which 1 mL of 3 M NaOH was added to the dried cells. The cell suspension was heated at 90 °C for 1 h, cooled to room temperature, and then 2 mL of 3 M HCl was added

and the suspension was reheated at 90 °C for 10 min. After cooling to room temperature, aliphatic compounds were extracted three times with 1 mL of 1:1 (v/v) hexane/diethyl ether. The organic phase was dried at room temperature, and the pellet was dissolved in methanol and analyzed using gas chromatography-mass spectrometry (GC-MS) (Shimadzu) [89]. The parameters for GC-MS were as follows: run time, 19.0 min; column, DB5-MS (Agilent Technologies); injection volume, 1 μL; inlet, 300 °C splitless; carrier gas, helium; flow, 1.3 mL/min; oven temp, hold at 100 °C, for 3 min, to 320 at 20 °C/min, hold at 320 °C for 5 min; detector, GCMS-QP2010; detector temp, 200 °C; scan range, 50-550 *m/z*. The resulting peak data were searched against the NIST electron impact fragmentation pattern library [35].

4.3.6. Preparation of purified *Se*ADO

The ADO gene from *S. elongatus* PCC 7942 (WP_011378104.1) (*Se*ADO) was amplified by PCR using primers 98 and 99 and then ligated into pQE30 [35]. The *Se*ADO gene was then expressed in *E. coli* JM109. Transformed *E. coli* cells were grown in Luria broth supplemented with 50 μg/mL of ampicillin. Once the cells reached an OD_{600} of approximately 0.5, isopropyl-1-thio-ß-D-galactopyranoside was added to a final concentration of 1 mM. Cells then were collected by centrifugation at 6,000 · *g* and disrupted by sonication. The proteins present in the cell-free extract were purified using Ni-NTA agarose resin (Qiagen) according to the manufacturer's instructions. Purified proteins were dissolved in 100 mM HEPES buffer (pH 7.2) containing 0.1 M KCl and 10% glycerol. The purified solution was aliquotted and stored at −80 °C.

4.3.7. ADO assay

Extracts containing fatty aldehydes from *C. thermocellum* were mixed with 130 μM purified *Se*ADO, 80 μM $Fe(NH_4)_2(SO_4)_2$, 10 μM spinach ferredoxin (Sigma-Aldrich), 1.8 μM spinach ferredoxin reductase (Sigma-Aldrich), and 2 mM NADPH in 100 μL of 100 mM HEPES buffer (pH 7.2) containing 0.1 M KCl and 10% glycerol. The reactions were incubated at 25 °C for 3 h, quenched by the addition of 100 μL of ethyl acetate, and mixed using a vortex mixer to extract the hydrocarbon product and unreacted substrate. A 1-μL aliquot of the ethyl acetate layer was injected into the GC-MS instrument [96]. The parameters for GC-MS were as described above [35].

4.3.8. Production of fuel compounds by engineered *C. thermocellum* strains

Clostridium thermocellum strains harboring ACR and ADO genes were cultured anaerobically in 100 mL of CTFUD medium with 10 g of Avicel cellulose and 24 μg/mL of thiamphenicol at 55 °C for 3 days. During cultivation, 200 μL of rapeseed oil was overlaid onto the medium. After cultivation, the oil was collected and mixed with 100 μL of 2 M methanolic KOH at room temperature for 1 min [98]. The aliphatic compounds in the organic phase were analyzed by GC-MS. The parameters for GC-MS were as described above [35]. Three independent experiments were conducted.

4.4. Results

4.4.1. Expression of ACR and ADO genes

The ACR and the ADO genes were amplified by PCR and ligated into pMU102 [92,93]. Ribosome-binding site sequences and histidine-tag sequences were inserted upstream of the genes. The genes were expressed using the promoter region of the cellobiose phosphorylase gene and the terminator region of the hypoxanthine phosphoribosyl transferase gene from *C. thermocellum* [85] (Fig. 2). The resulting plasmids were used to transform *C. thermocellum* [92]. Successful isolation of the transformants was confirmed by PCR (Fig. 3). Each transformant was cultured in CTFUD medium with 24 μg/mL of thiamphenicol, and the levels of ACR and ADO expression were analyzed by Western blotting. No specific band was detected in the strain harboring pMU102. In the isolated transformants harboring *Se*ACR and *Se*ADO genes, *Se*ADO expression was clearly detected. In contrast, only slight expression of *Se*ACR was detected (Fig. 4 and 5). ACRs from *C. acetobutilycum* ATCC 824, *C. cellulovorans* ATCC 35296, and *T. elongatus* BP-1 were also expressed instead of *Se*ACR; however, the expression levels of these ACRs were also low (Fig. 5). Some strains did not express ACR or ADO, even if they harbored the appropriate plasmid (Fig. 3 and 5).

4.4.2. Production of fatty aldehydes in *C. thermocellum* cells

Clostridium thermocellum cells expressing *Se*ACR and *Se*ADO were collected, and aliphatic compounds within the cells were extracted using a previously described procedure [89]. The peaks exhibiting ≥90% similarity to fatty aldehydes were detected (Fig. 6). Authentic standards of the detected fatty aldehydes were not commercially available; therefore, the presence of fatty aldehydes in the cell extract was confirmed by reacting the samples with purified *Se*ADO, which converts fatty aldehydes to hydrocarbons [97]. After reacting *Se*ADO with the extract of *C. thermocellum* cells expressing *Se*ACR and *Se*ADO, the intensity of peaks exhibiting ≥90% similarity to fatty aldehydes decreased, and the hydrocarbon pentadecan was detected in the reaction solution (Fig. 7). The presence of pentadecan in the reaction solution was also confirmed based on the retention time of an authentic standard. These results indicate that *C. thermocellum* cells expressing *Se*ACR and *Se*ADO contain fatty aldehydes.

4.4.3. Production and secretion of water-insoluble fuel compounds from cellulose

Clostridium thermocellum expressing *Se*ACR and *Se*ADO was cultivated in 100 mL of CTFUD medium with 100 g/L of Avicel cellulose. Rapeseed oil was overlaid onto the medium to capture water-insoluble fuel compounds produced by the bacteria. After cultivation, the oil was collected and methanolyzed, and aliphatic compounds in the organic solvent were analyzed by GC-MS [98]. We did observe peaks exhibiting ≥90% similarity to higher alcohols (Fig. 8). Two higher alcohols, decanol and dodecanol, were identified based on the retention time of authentic standards (Fig. 8). The amounts of decanol and dodecanol detected in the organic solvent were 3.1 and 19.5 μg, respectively (Fig. 9).

4.5. Discussion

The recalcitrance of cellulosic biomass is a significant barrier to biofuel production. Lignocellulose-degrading enzymes must therefore be used to convert cellulosic biomass into fermentable sugars; however, the efficiency of this process is low [54]. *Clostridium thermocellum* is a promising microorganism for biofuel production because it can degrade cellulosic biomass without chemical pretreatment or the addition of lignocellulose-degrading enzymes [77,78] (Fig. 1). It was reported that the expression of *S. elongatus* PCC 7492 ACR and ADO in *E. coli* is sufficient to confer the ability to synthesize higher alcohols and hydrocarbons [35]. In this study, the genes encoding ACR and ADO were expressed in *C. thermocellum* to produce water-insoluble fuel compounds from cellulose (Fig. 1).

*Se*ADO was clearly expressed, whereas the level of *Se*ACR expression was quite low (Fig. 4). To achieve higher levels of ACR expression in *C. thermocellum*, other ACRs were examined. No gene homologous to *Se*ACR was identified in *C. thermocellum*. The ACR from *C. acetobutilycum* ATCC 824 was also active in *E. coli* [94]. ACR genes from *C. cellulovorans* ATCC 35296 and *T. elongatus* BP-1 are homologues of the *Ca*ACR and *Se*ACR genes, respectively. We have speculated that genes from other clostridia could be expressed in *C. thermocellum*. Because *T. elongatus* is a thermophile, it is possible that its proteins could be expressed and would remain stable in *C. thermocellum*, which is cultivated under thermophilic conditions [99]. These ACRs, *Ca*ACR, *Cc*ACR and *Te*ACR, were also expressed in *C. thermocellum* instead of *Se*ACR; however, the expression levels were low (Fig. 5). Achieving higher levels of heterologous expression of ACR might be difficult, as it has been reported that *E. coli* also expresses ACRs at low levels [100].

No higher alcohols and hydrocarbons were detected in the extracts of 5-mL cultures of any of the transformants after the addition of an equal amount of ethyl acetate, nor were any higher alcohols and hydrocarbons detected in the methanol extracts of disrupted transformants [35] (data not shown).

Subsequently, in order to determine the production of higher alcohols and hydrocarbons, cells of the strain expressing *Se*ACR and *Se*ADO were cultured in 50 mL of medium, collected, and then the aliphatic compounds were extracted [89]. No higher alcohols and hydrocarbons were detected in the extract; however, peaks exhibiting ≥90% similarity to fatty aldehydes were detected. After the extract was mixed with purified *Se*ADO, the hydrocarbon pentadecan was detected, indicating that cells expressing *Se*ACR and *Se*ADO contained fatty aldehydes (Fig. 7). Fatty aldehydes are intermediates of the higher alcohol and hydrocarbon

synthesis pathway [76] (Fig. 10). Thus, a strain expressing *Se*ACR and *Se*ADO would be expected to produce higher alcohols and hydrocarbons; however, these compounds were not detected in these experiments.

Consequently, the strain was cultured with 10% Avicel cellulose, and the presence of aliphatic compounds contained in a small amount of oil overlaid onto the medium was assessed using GC-MS. In this experiment, the higher alcohols decanol and dodecanol were successfully detected in the overlay oil, while no hydrocarbons were detected (Fig. 9). As ADO reportedly requires molecular oxygen to function, the enzyme may not have been active in *C. thermocellum*, which is cultivated under anaerobic conditions [101,102].

Higher alcohols are commodity chemicals used in fuels, detergents, emollients, and food additives [103,104]. The ACR from *S. elongatus* catalyzes the production of fatty aldehydes rather than higher alcohols [100]. It was also reported that the combined deletion of 13 aldehyde reductases (ALRs) in *E. coli* (AdhE, YqhD, FucO, YiaY, EutG, GldA, AdhP, YjgB, YahK, YghA, YbbO, DkgA, and BetA) results in a >90% reduction in the production of higher alcohols and a significant increase in fatty aldehyde production, suggesting that these ALRs are involved in higher alcohol synthesis [105] (Fig. 10). *Clostridium thermocellum* harbors genes homologous to *E. coli* ALRs. An NCBI protein Blast similarity search revealed that a number of *C. thermocellum* genes (Clo1313_2130, Clo1313_1827, Clo1313_0166, and Clo1313_1798) are homologous to AdhE, YqhD, FucO, YiaY, GldA, and EutG (Table 2). The other genes, (Clo1313_1833 and Clo1313_0076), (Clo1313_1283, Clo1313_0815, Clo1313_2272, Clo1313_0720), and (Clo1313_1992, Clo1313_2081, Clo1313_1946) are similar to ALRs in *E. coli* (AdhP, Yjgb, YahK), (YghA, YbbO), and DkgA, respectively. These results suggest that the ALRs might convert fatty aldehydes to higher alcohols in *C. thermocellum* (Fig. 10).

The other significant barrier to biofuel production is the cost of distillation of water-soluble fuel compounds [13,86] (Fig. 1). To avoid this expensive step, *C. thermocellum* was engineered to produce water-insoluble fuel compounds. However, as some oleaginous microorganisms accumulate lipids in the cell, harvesting the cells and extracting the products also increases the cost of biofuel production [106,107]. Thus, product secretion is an important factor, as secretion of water-insoluble fuel compounds allows for phase separation to occur, facilitating recovery from the fermentation broth and alleviating potential toxicity issues [91]. Additionally, secretion of fuel compounds has the potential to relieve product-feedback inhibition, thus increasing final concentrations [108]. In this study, the higher alcohols decanol and dodecanol were produced and detected in the oil overlaid onto the culture broth, indicating that the engineered *C. thermocellum* strain secreted water-insoluble fuel compounds (Fig. 9).

In this study, *C. thermocellum* was engineered to produce water-insoluble fuel compounds from cellulose. ACR and ADO genes were successfully expressed in the engineered strain. Fatty aldehydes, which are higher alcohol precursors, were detected in cells of the *C. thermocellum* strain expressing *Se*ACR and *Se*ADO. The *C. thermocellum* strain expressing *Se*ACR and *Se*ADO produced and secreted higher alcohols from cellulose, without the addition of cellulases, suggesting that the strain directly converts cellulose into water-insoluble fuel compounds that can be easily collected by phase separation (Fig. 1).

4.6. Figures and tables

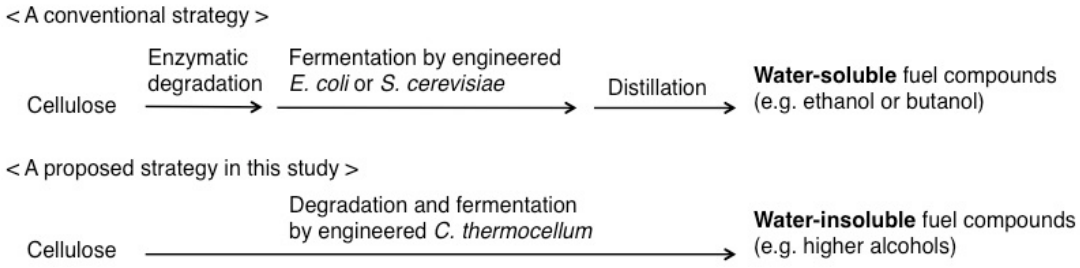


Figure 1. Conventional and proposed strategies for biofuel production.

Conventional strategies include expensive steps, such as enzymatic degradation of cellulosic biomass and distillation of water-soluble fuel compounds. In this study, a cellulolytic bacterium, *Clostridium thermocellum*, was engineered to produce and secrete water-insoluble fuel compounds in order to circumvent the costly steps in the conventional process.

Primer	Sequence (5' to 3')
35	ATTCCCGGGATTACCGGTATGTTCGGTCTTATCGGTCATCTC
36	ATTGGTACCATTGCGGCCGCTCAAATTGCCAATGCCAAGG
37	ATTGCGGCCGCATTAGATCTATGCCGCAGCTTGAAGC
38	ATTGGTACCATTGTTAACTCAAACGGCCGCAAGG
39	ATTCCCGGGATTGGGCCCAATTATAATAGATTCAATTGTGAGCGGA
40	ATTACCGGTGTGATGGTGATGGTGATGCG
41	ATTGCGGCCGCATTATGCATAATTATAATAGATTCAATTGTGAGCGGA
42	ATTAGATCTGTGATGGTGATGGTGATGCG
62	GAATTCGAGCTCGGTACCCG
64	TAATGAAGAAAGCAGACAAGTAAGCCT
98	ATTGGATCCATGCCGCAGCTTGAAGC
99	ATTAAGCTTTCAAACGGCCGCAAGG
102	TAAGCATGCGAGTCGTGACTAAGAACGTCAAAGTAA
103	TAAGGGCCCCTTTCGTCCTCCTTAAAATTTTCG
104	TAAGTTAACAAAATATACAAAGGTTTCTTGTGTTTTTAATAC
105	TAAGTCGACTTCAACTAGTTCCCTCCTCATTCTT
120	GGGGGCTACCCCAAAAACTT
152	ATTACCGGTATGAAGCTTGAATTTACATGTGGC
153	ATTGCGGCCGCTTAGATTTGTCTACTAACCAAAAATGTAAACTC
156	ATTACCGGTATGAAGAGCGAAACTGTTTGTGA
157	ATTGCGGCCGCTTATAGTTGTCTACTAACTACGTAAGTCATTTCC
169	ATTACCGGTATGTTTGGATTAATTGGTCATCTGAC
170	ATTGCGGCCGCTTACTGAGGATTCAACAGTAGTGGC
171	TGTTCGCCTTCAATCACGATC

Table 1. Primers used in this study.

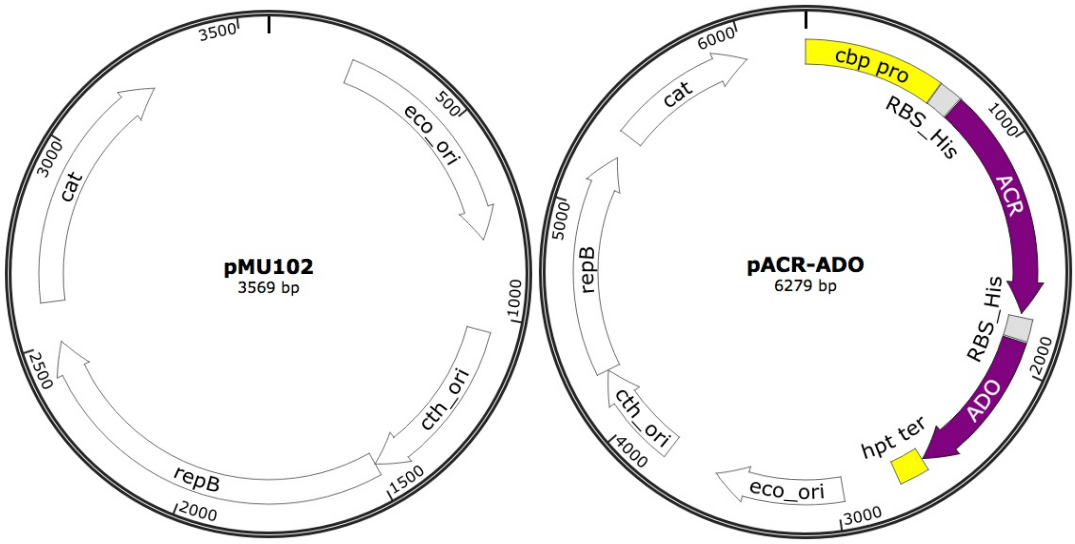
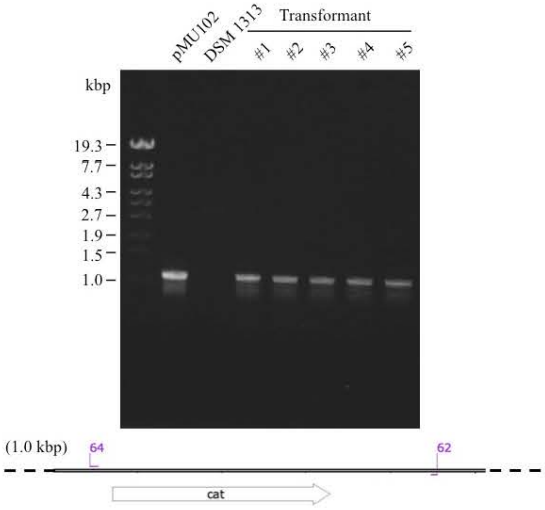
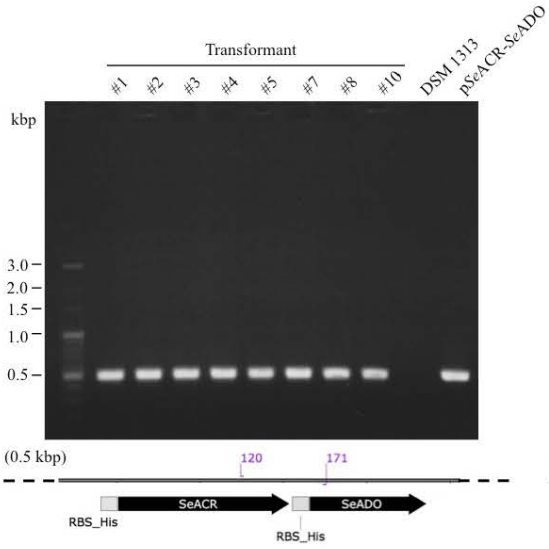
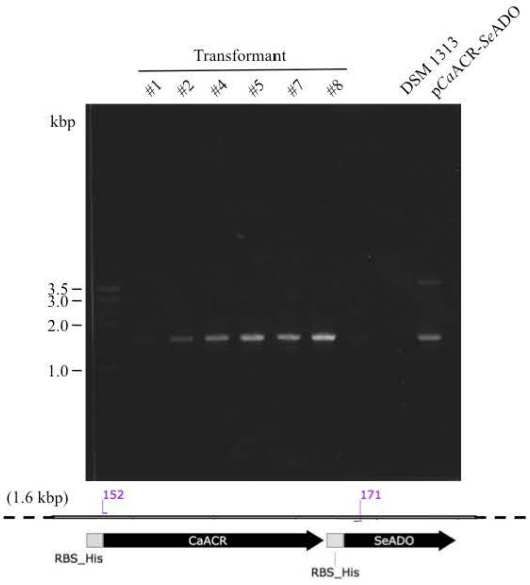


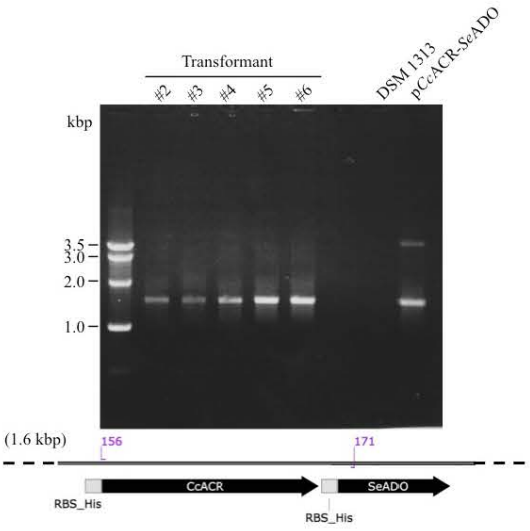
Figure 2. Construction of the plasmid used for higher alcohol and hydrocarbon production.

The plasmid pMU102 was used as a vector (left). ACR and ADO genes were inserted into pMU102 (right). cbp pro: a promoter from the cellobiose phosphorylase gene of *C. thermocellum*. RBS_His: ribosome-binding sites and histidine-tag sequences from pQE30. ACR: fatty acyl-ACP reductase gene. ADO: aldehyde-deformylating oxygenase gene. hpt ter: terminator from the hypoxanthine phosphoribosyl transferase gene of *C. thermocellum*. eco_ori: replication origin function in *E. coli*. cth_ori: replication origin function in *C. thermocellum*. repB: replication protein for the pMU102 origin of replication. cat: chloramphenicol acetyltransferase (provides resistance to chloramphenicol and thiamphenicol).


pMU102
DSM 1313
Transformant
#1
#2
#3
#4
#5
kbp
19.3
7.7
4.3
2.7
1.9
1.5
1.0
(1.0 kbp)
64
62
cat


Transformant
#1
#2
#3
#4
#5
#7
#8
#10
DSM 1313
pSeACR-SeADO
kbp
3.0
2.0
1.5
1.0
0.5
(0.5 kbp)
120
171
SeACR
SeADO
RBS_His
RBS_His


Transformant
#1
#2
#4
#5
#7
#8
DSM 1313
pCaACR-SeADO
kbp
3.5
3.0
2.0
1.0
(1.6 kbp)
152
171
CaACR
SeADO
RBS_His
RBS_His


Transformant
#2
#3
#4
#5
#6
DSM 1313
pCcACR-SeADO
kbp
3.5
3.0
2.0
1.0
(1.6 kbp)
156
171
CcACR
SeADO
RBS_His
RBS_His

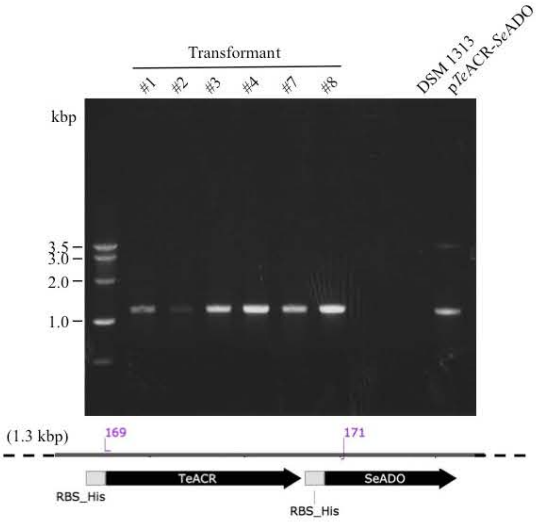

Transformant
#1
#2
#3
#4
#7
#8
DSM 1313
pTeACR-SeADO
kbp
3.5
3.0
2.0
1.0
(1.3 kbp)
169
171
TeACR
SeADO
RBS_His
RBS_His

Figure 3. Isolation of *C. thermocellum* transformed to express ACRs and ADO.

Isolated transformants were cultivated in CTFUD with cellobiose and thiamphenicol. The cultured cells were used as samples for PCR. Portions of the plasmids and primers used in the experiments are illustrated in each panel. The predicted sizes of amplified products are shown in parentheses. Wild-type strain DSM 1313 was used as a negative control. The plasmids used for transformation served as positive controls.

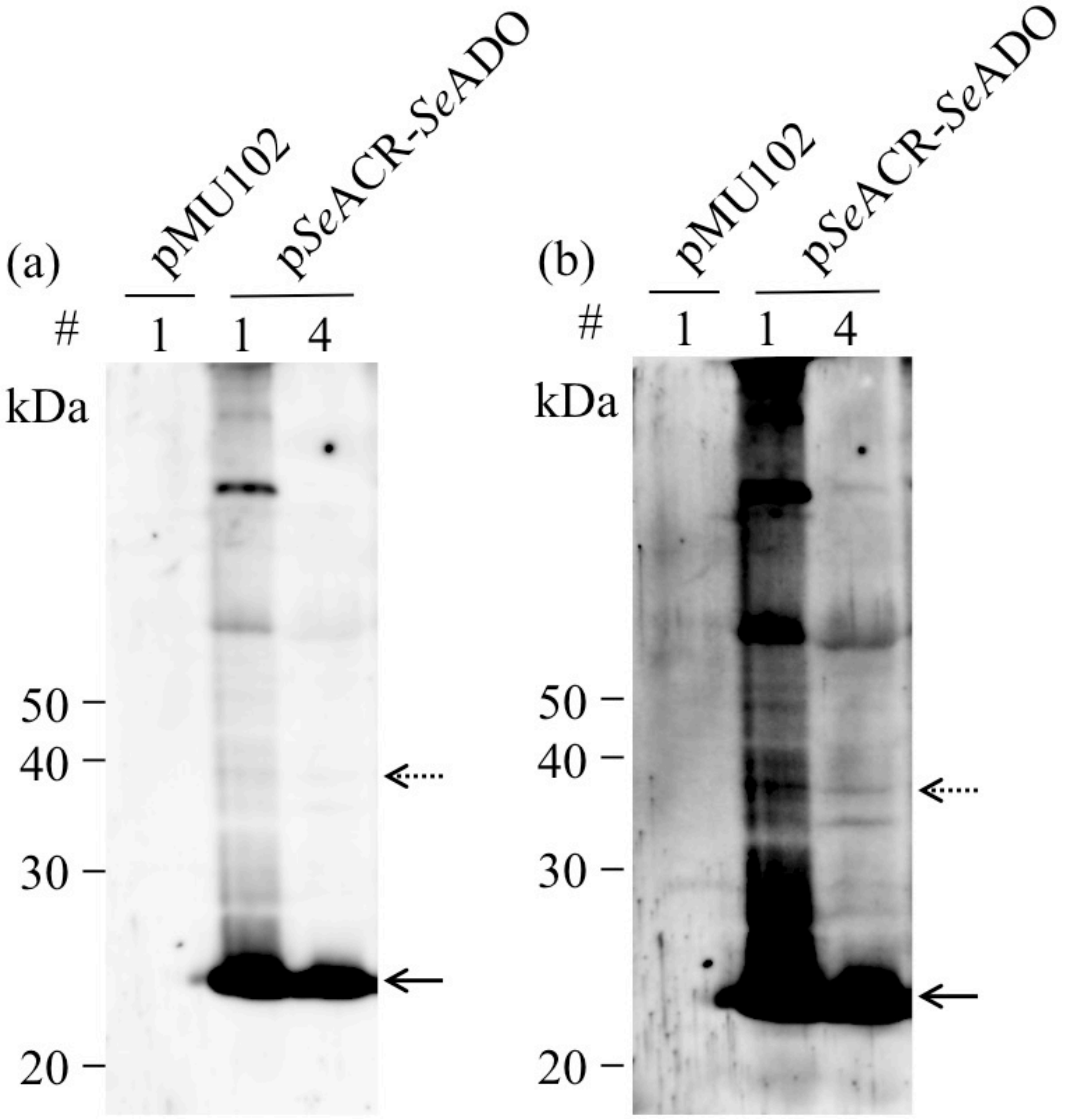


Figure 4. Expression of *Se*ACR and *Se*ADO in *C. thermocellum*.

Cell lysates of each isolated transformant were analyzed by Western blotting to evaluate the expression of *Se*ACR and *Se*ADO (a, b). The exposure time was longer in (b). Dotted and solid arrows indicate *Se*ACR and *Se*ADO, respectively. The molecular weights of *Se*ACR and *Se*ADO are 38.9 and 27.9 kDa, respectively. Numerals above the lanes indicate the strains isolated in this study (Fig. 3).

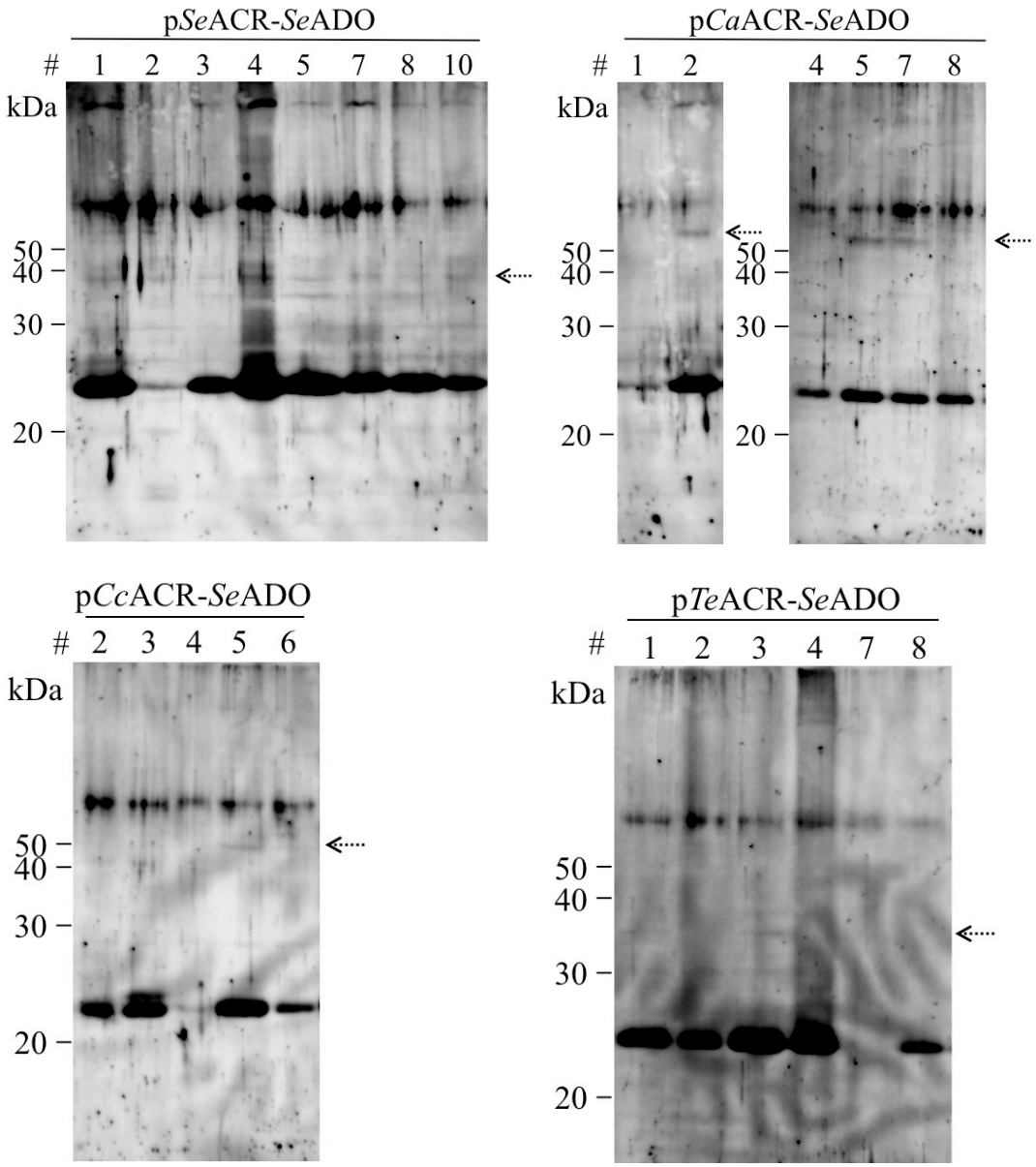
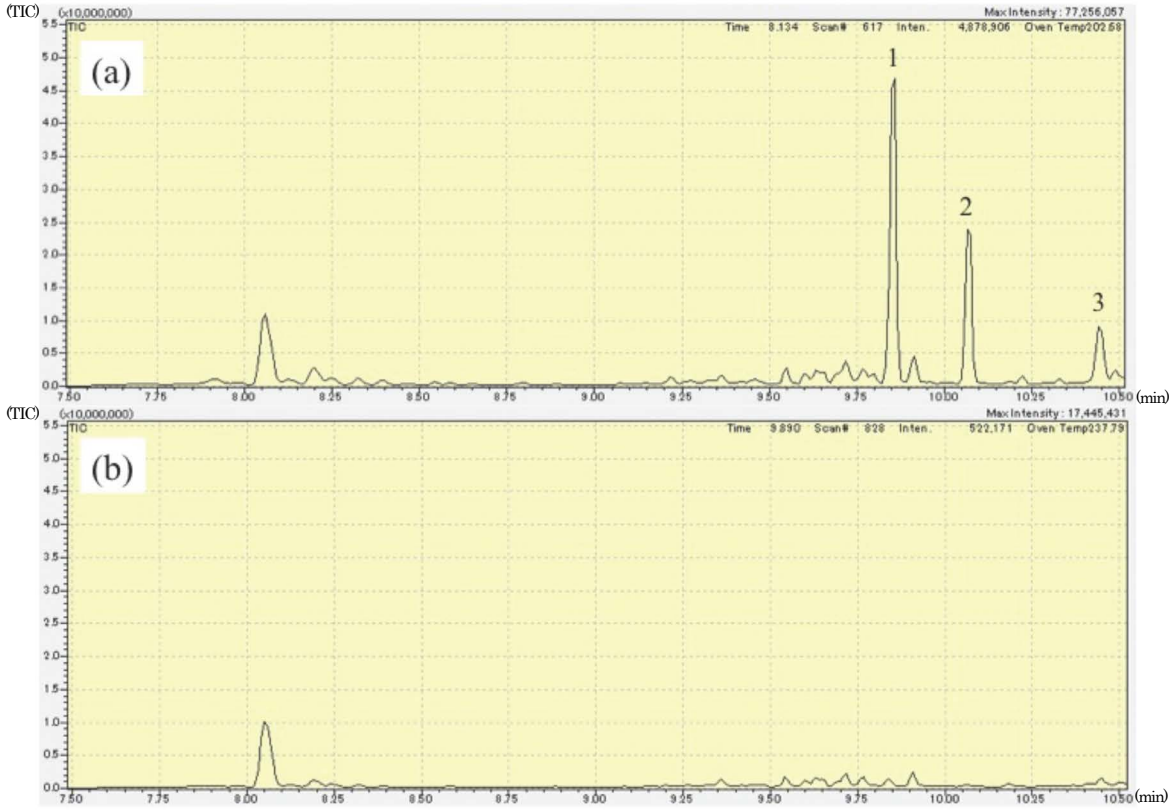
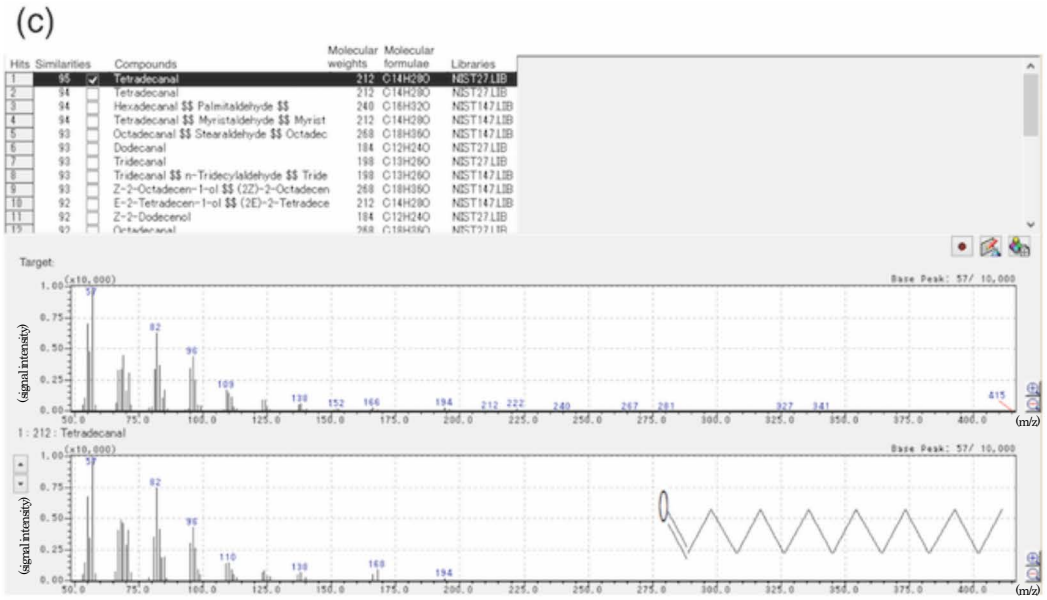


Figure 5. Expression of *Se*ACR, *Ca*ACR, *Cc*ACR, and *Te*ACR and *Se*ADO in *C. thermocellum*.

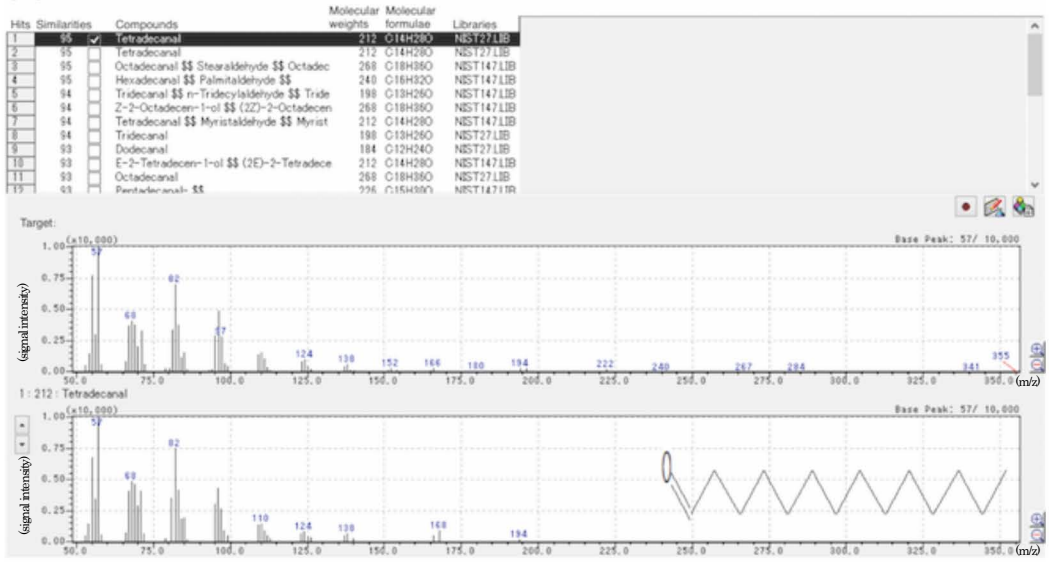
Cell lysates of each transformant were analyzed by Western blotting to assess the expression of ACRs and ADO. The expression of *Se*ADO and *Se*ACR, *Ca*ACR, *Cc*ACR, and *Te*ACR is shown. Numerals above the lanes indicate the strains isolated in this study (Fig. 3). Dotted arrows indicate ACR expression. The molecular weights of the expressed proteins are as follows: *Se*ACR, 38.9 kDa; *Ca*ACR, 53.7 kDa; *Cc*ACR, 53.0 kDa; *Te*ACR, 40.0 kDa; *Se*ADO, 27.9 kDa.


(TIC)
(x10,000,000)
TIC
Max Intensity : 77,256,057
Time 8.134 Scan# 617 Inten. 4,878,906 Oven Temp202.58
(a)
1
2
3
(min)
(TIC)
(x10,000,000)
TIC
Max Intensity : 17,445,431
Time 9.890 Scan# 828 Inten. 522,171 Oven Temp237.79
(b)
(min)

(c)


Hits Similarities Compounds Molecular weights Molecular formulae Libraries
1 95 Tetradecanal 212 C14H28O NIST27.LIB
2 94 Tetradecanal 212 C14H28O NIST27.LIB
3 94 Hexadecanal $$ Palmitaldehyde $$ 240 C16H32O NIST147.LIB
4 94 Tetradecanal $$ Myristaldehyde $$ Myrist 212 C14H28O NIST147.LIB
5 93 Octadecanal $$ Stearaldehyde $$ Octadec 268 C18H36O NIST147.LIB
6 93 Dodecanal 184 C12H24O NIST27.LIB
7 93 Tridecanal 198 C13H26O NIST27.LIB
8 93 Tridecanal $$ n-Tridecylaldehyde $$ Tride 198 C13H26O NIST147.LIB
9 93 Z-2-Octadecen-1-ol $$ (2Z)-2-Octadecen 268 C18H36O NIST147.LIB
10 92 E-2-Tetradecen-1-ol $$ (2E)-2-Tetradece 212 C14H28O NIST147.LIB
11 92 Z-2-Dodecenol 184 C12H24O NIST27.LIB
Target:
Base Peak: 57/ 10,000
(signal intensity)
(m/z)
1 : 212 : Tetradecanal
Base Peak: 57/ 10,000
(signal intensity)
(m/z)

(d)


Hits Similarities Compounds Molecular weights Molecular formulae Libraries
1 95 Tetradecanal 212 C14H28O NIST27.LIB
2 95 Tetradecanal 212 C14H28O NIST27.LIB
3 95 Octadecanal $$ Stearaldehyde $$ Octadec 268 C18H36O NIST147.LIB
4 95 Hexadecanal $$ Palmitaldehyde $$ 240 C16H32O NIST147.LIB
5 94 Tridecanal $$ n-Tridecylaldehyde $$ Tride 198 C13H26O NIST147.LIB
6 94 Z-2-Octadecen-1-ol $$ (2Z)-2-Octadecen 268 C18H36O NIST147.LIB
7 94 Tetradecanal $$ Myristaldehyde $$ Myrist 212 C14H28O NIST147.LIB
8 94 Tridecanal 198 C13H26O NIST27.LIB
9 93 Dodecanal 184 C12H24O NIST27.LIB
10 93 E-2-Tetradecen-1-ol $$ (2E)-2-Tetradece 212 C14H28O NIST147.LIB
11 93 Octadecanal 268 C18H36O NIST27.LIB
Target:
Base Peak: 57/ 10,000
(signal intensity)
(m/z)
1 : 212 : Tetradecanal
Base Peak: 57/ 10,000
(signal intensity)
(m/z)

(e)

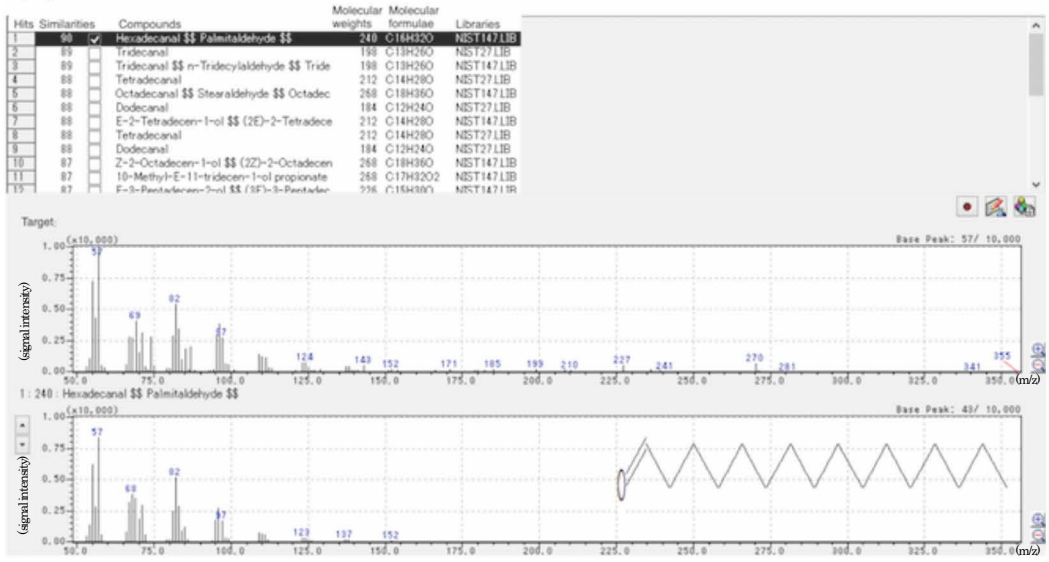

Hits Similarities Compounds Molecular weights Molecular formulae Libraries
1 90 Hexadecanal $$ Palmitaldehyde $$ 240 C16H32O NIST147.LIB
2 89 Tridecanal 198 C13H26O NIST27.LIB
3 89 Tridecanal $$ n-Tridecylaldehyde $$ Tride 198 C13H26O NIST147.LIB
4 88 Tetradecanal 212 C14H28O NIST27.LIB
5 88 Octadecanal $$ Stearaldehyde $$ Octadec 268 C18H36O NIST147.LIB
6 88 Dodecanal 184 C12H24O NIST27.LIB
7 88 E-2-Tetradecen-1-ol $$ (2E)-2-Tetradece 212 C14H28O NIST147.LIB
8 88 Tetradecanal 212 C14H28O NIST27.LIB
9 88 Dodecanal 184 C12H24O NIST27.LIB
10 87 Z-2-Octadecen-1-ol $$ (2Z)-2-Octadecen 268 C18H36O NIST147.LIB
11 87 10-Methyl-E-11-tridecen-1-ol propionate 268 C17H32O2 NIST147.LIB
Target:
Base Peak: 57/ 10,000
(signal intensity)
(m/z)
1 : 240 : Hexadecanal $$ Palmitaldehyde $$
Base Peak: 43/ 10,000
(signal intensity)
(m/z)

Figure 6. GC-MS analysis of the cell extract of the strain expressing *Se*ACR and *Se*ADO.

The cell extract of the strain expressing *Se*ACR and *Se*ADO was analyzed by GC-MS (a). Analysis of culture medium control without *C. thermocellum* is shown in (b). Fragmentation patterns and results of similarity searches for compounds represented by peaks 1, 2, and 3 are shown in panels (c), (d), and (e), respectively.

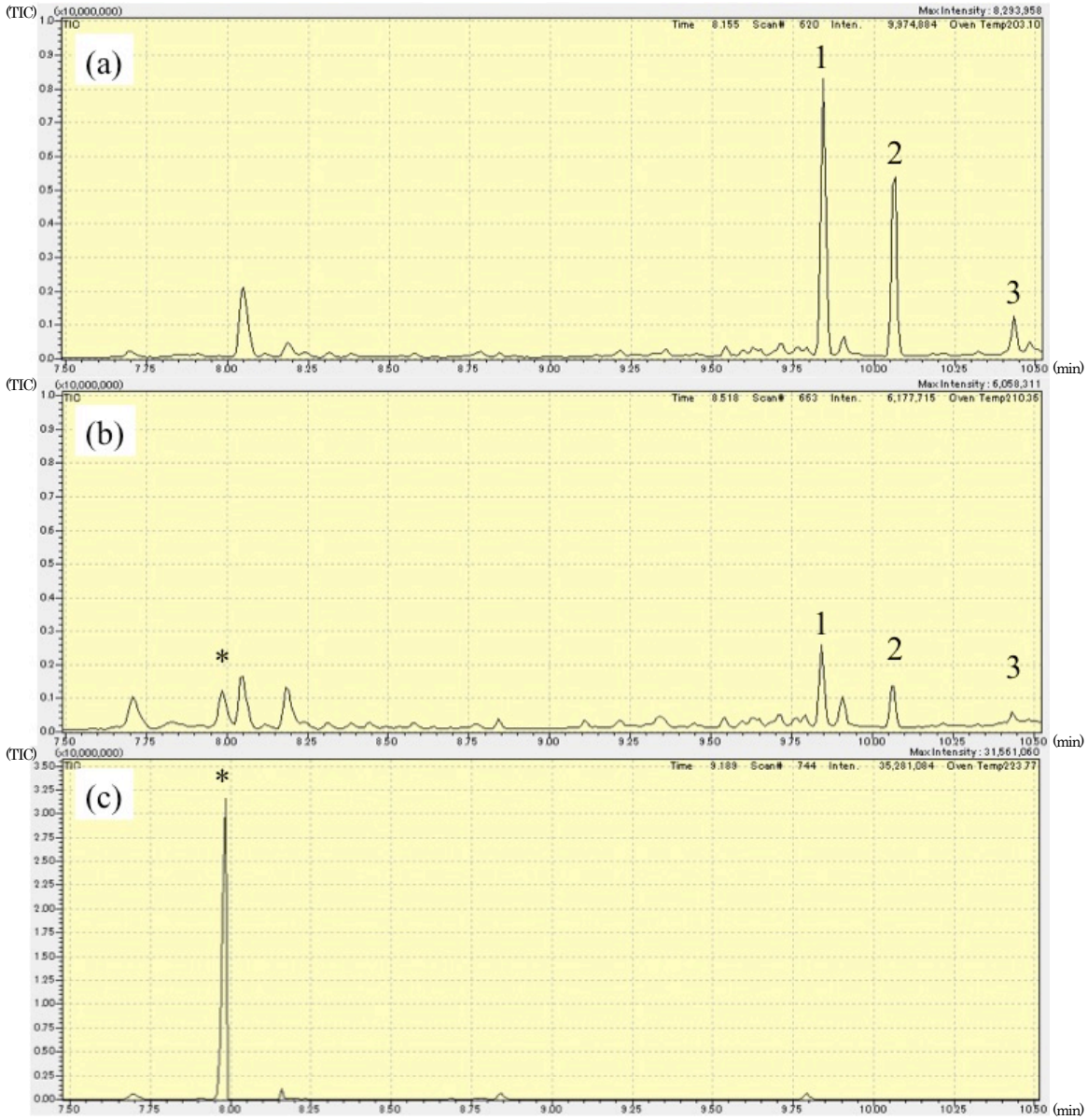
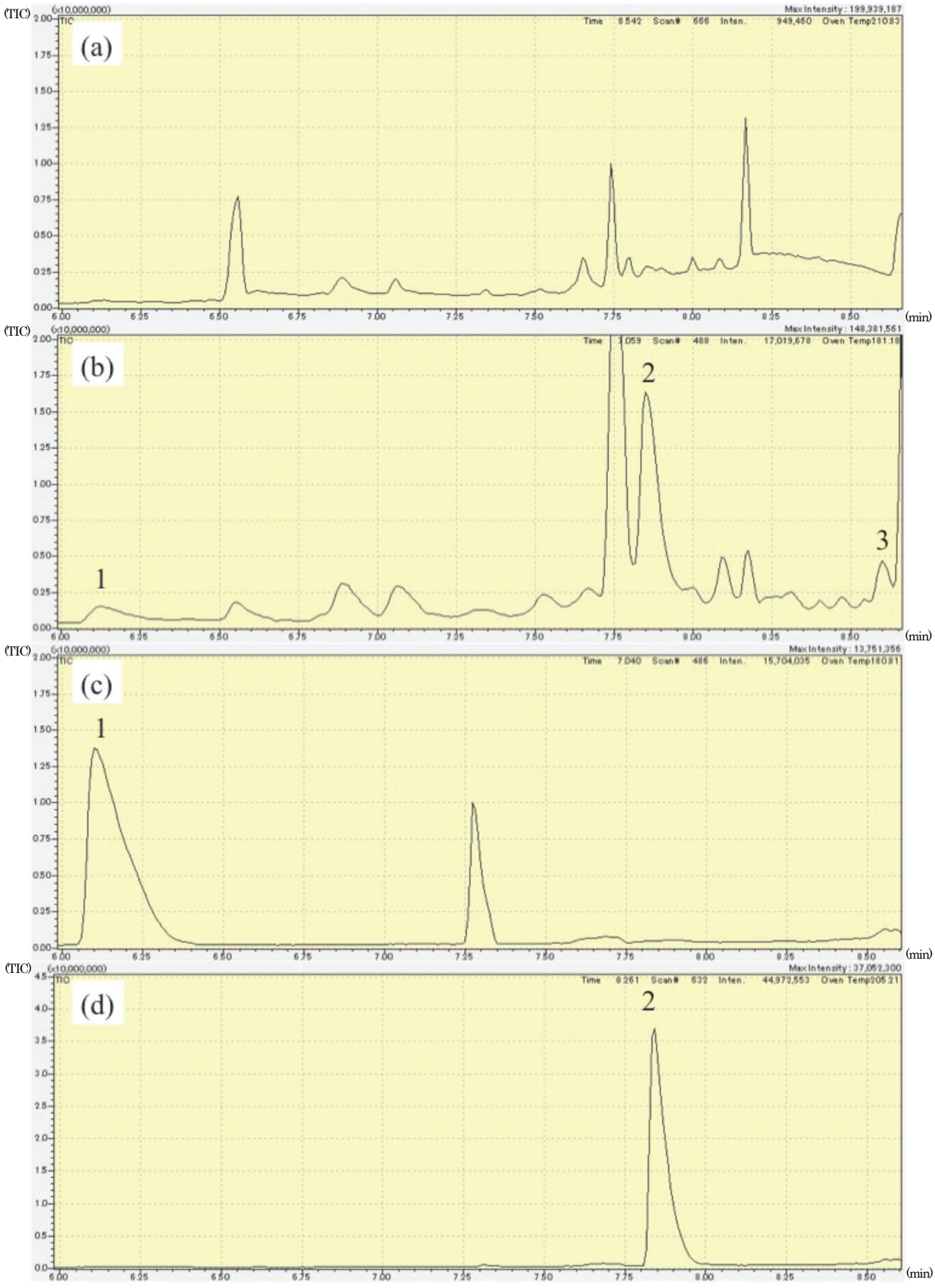
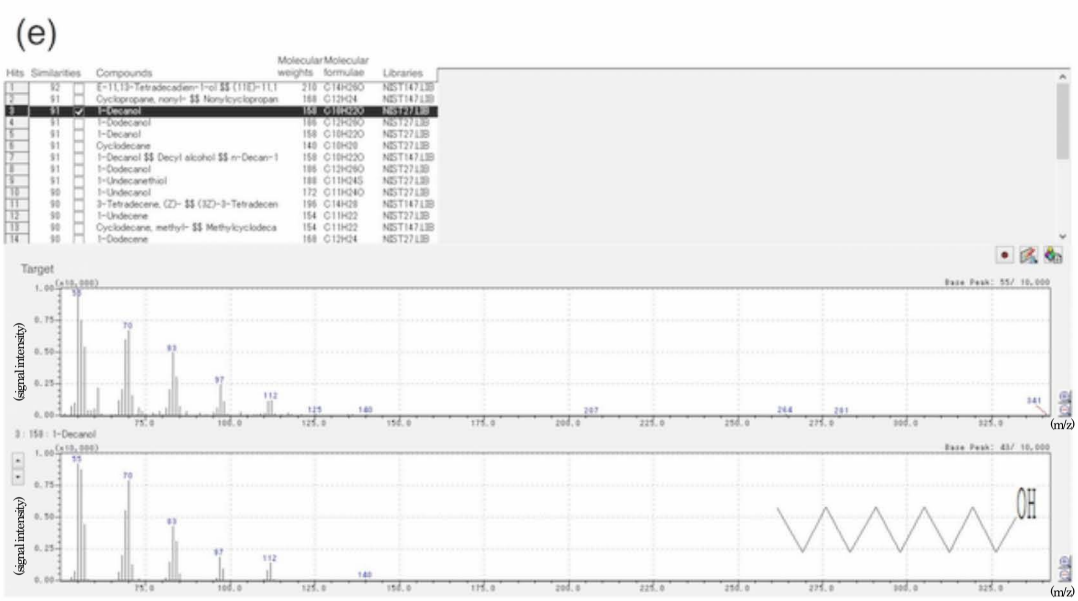


Figure 7. Accumulation of fatty aldehydes in cells of the strain expressing *Se*ACR and *Se*ADO.

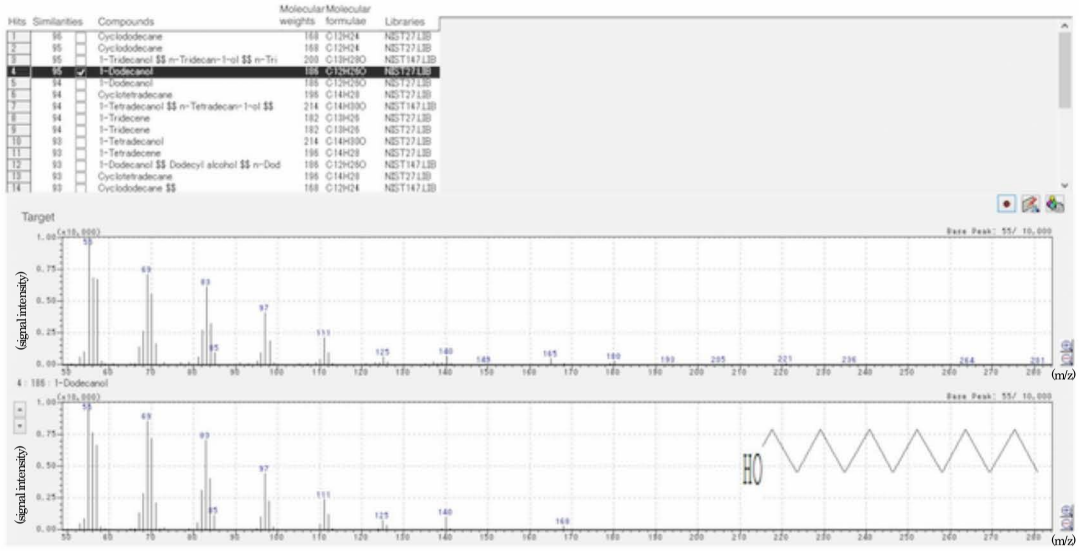
(a) Analysis of the cell extract of the strain expressing *Se*ACR and *Se*ADO without purified *Se*ADO. (b) Analysis of the cell extract of the strain expressing *Se*ACR and *Se*ADO with added purified *Se*ADO. The reaction solutions were analyzed by GC-MS. Analysis of an authentic pentadecan standard is shown in (c). Peaks 1, 2, and 3 exhibited ≥90% similarity to fatty aldehydes (Fig. 6). Asterisk indicates pentadecan peak.


(TIC)
(a)
(b)
(c)
(d)
1
2
3
(min)

(e)


Hits Similarities Compounds Molecular weights Molecular formulae Libraries
1 92 E-(1,13)-Tetradecadien-1-ol $$ (11E)-11,1 210 C14H26O NIST147.LIB
2 91 Cyclopropane, nonyl- $$ Nonylcyclopropan 168 C12H24 NIST147.LIB
3 91 1-Decanol 158 C10H22O NIST27.LIB
4 91 1-Dodecanol 186 C12H26O NIST27.LIB
5 91 1-Decanol 158 C10H22O NIST27.LIB
6 91 Cyclododecane 168 C12H24 NIST27.LIB
7 91 1-Decanol $$ Decyl alcohol $$ n-Decan-1 158 C10H22O NIST147.LIB
8 91 1-Dodecanol 186 C12H26O NIST27.LIB
9 91 1-Undecanethiol 188 C11H24S NIST27.LIB
10 90 1-Undecanol 172 C11H24O NIST27.LIB
11 90 3-Tetradecene, (Z)- $$ (3Z)-3-Tetradecen 196 C14H28 NIST147.LIB
12 90 1-Undecene 154 C11H22 NIST27.LIB
13 90 Cyclodecane, methyl- $$ Methylcyclodeca 154 C11H22 NIST147.LIB
14 90 1-Dodecene 168 C12H24 NIST27.LIB
Target
Base Peak: 55/ 10,000
(signal intensity)
(m/z)
3 : 158 : 1-Decanol
Base Peak: 43/ 10,000
OH

(f)


Hits Similarities Compounds Molecular weights Molecular formulae Libraries
1 96 Cyclododecane 168 C12H24 NIST27.LIB
2 95 Cyclododecane 168 C12H24 NIST27.LIB
3 95 1-Tridecanol $$ n-Tridecan-1-ol $$ n-Tri 200 C13H28O NIST147.LIB
4 95 1-Dodecanol 186 C12H26O NIST27.LIB
5 94 1-Dodecanol 186 C12H26O NIST27.LIB
6 94 Cyclotetradecane 196 C14H28 NIST27.LIB
7 94 1-Tetradecanol $$ n-Tetradecan-1-ol $$ 214 C14H30O NIST147.LIB
8 94 1-Tridecene 182 C13H26 NIST27.LIB
9 94 1-Tridecene 182 C13H26 NIST27.LIB
10 93 1-Tetradecanol 214 C14H30O NIST27.LIB
11 93 1-Tetradecene 196 C14H28 NIST27.LIB
12 93 1-Dodecanol $$ Dodecyl alcohol $$ n-Dod 186 C12H26O NIST147.LIB
13 93 Cyclotetradecane 196 C14H28 NIST27.LIB
14 93 Cyclododecane $$ 168 C12H24 NIST147.LIB
Target
Base Peak: 55/ 10,000
(signal intensity)
(m/z)
4 : 186 : 1-Dodecanol
Base Peak: 55/ 10,000
HO

(g)

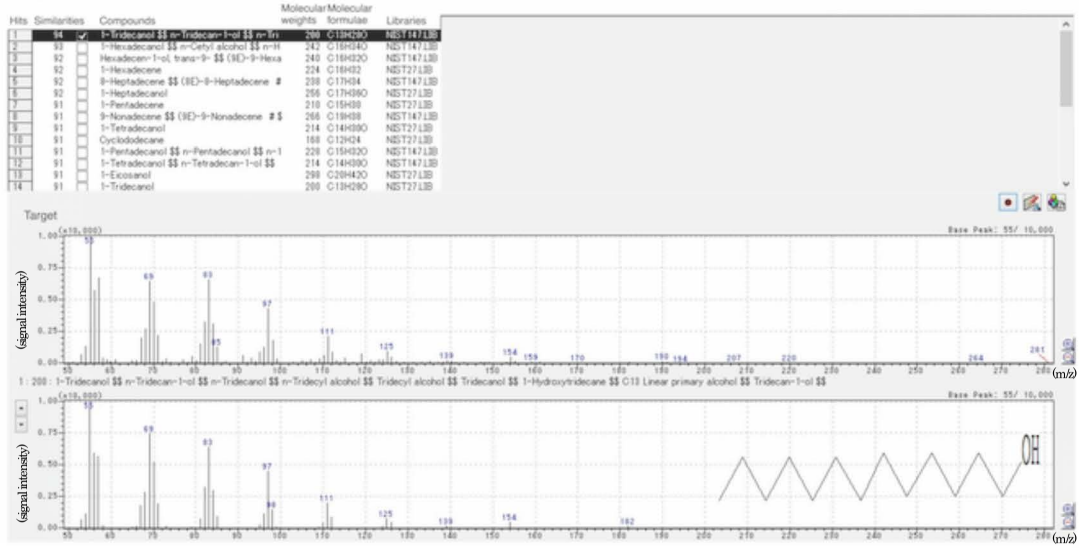

Hits Similarities Compounds Molecular weights Molecular formulae Libraries
1 94 1-Tridecanol $$ n-Tridecan-1-ol $$ n-Tri 200 C13H28O NIST147.LIB
2 93 1-Hexadecanol $$ n-Cetyl alcohol $$ n-H 242 C16H34O NIST147.LIB
3 92 Hexadecen-1-ol, trans-9- $$ (9E)-9-Hexa 240 C16H32O NIST147.LIB
4 92 1-Hexadecene 224 C16H32 NIST27.LIB
5 92 9-Heptadecene $$ (9E)-9-Heptadecene # 238 C17H34 NIST147.LIB
6 92 1-Heptadecanol 256 C17H36O NIST27.LIB
7 91 1-Pentadecene 210 C15H30 NIST27.LIB
8 91 9-Nonadecene $$ (9E)-9-Nonadecene # $ 266 C19H38 NIST147.LIB
9 91 1-Tetradecanol 214 C14H30O NIST27.LIB
10 91 Cyclododecane 168 C12H24 NIST27.LIB
11 91 1-Pentadecanol $$ n-Pentadecanol $$ n-1 228 C15H32O NIST147.LIB
12 91 1-Tetradecanol $$ n-Tetradecan-1-ol $$ 214 C14H30O NIST147.LIB
13 91 1-Eicosanol 298 C20H42O NIST27.LIB
14 91 1-Tridecanol 200 C13H28O NIST27.LIB
Target
Base Peak: 55/ 10,000
(signal intensity)
(m/z)
1 : 200 : 1-Tridecanol $$ n-Tridecan-1-ol $$ n-Tridecanol $$ n-Tridecyl alcohol $$ Tridecyl alcohol $$ Tridecanol $$ 1-Hydroxytridecane $$ C13 Linear primary alcohol $$ Tridecan-1-ol $$
Base Peak: 55/ 10,000
OH

Figure 8. GC-MS analysis of compounds produced from Avicel cellulose by the strain expressing *Se*ACR and *Se*ADO.

The strain expressing *Se*ACR and *Se*ADO was cultivated with Avicel cellulose. Aliphatic compounds in the organic solvent phase of the culture broth were analyzed by GC-MS. Gas chromatograms of aliphatic compounds of the strain harboring pMU102 and p*Se*ACR-*Se*ADO and authentic decanol and dodecanol standards are shown in (a), (b), (c), and (d), respectively. Fragmentation patterns and the results of similarity searches for the compounds represented by peaks 1, 2, and 3 are shown in panels (e), (f), and (g), respectively.

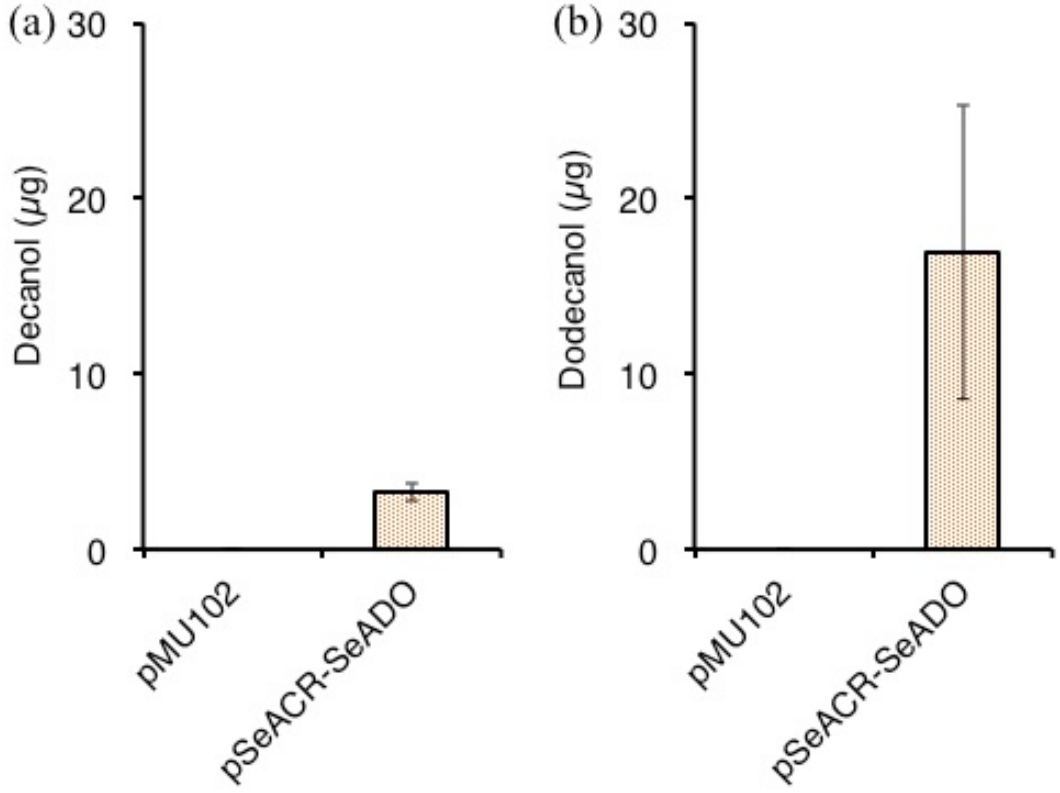


Figure 9. Production of higher alcohols from cellulose.

The strain expressing *Se*ACR and *Se*ADO was cultivated with Avicel cellulose. The higher alcohols decanol (a) and dodecanol (b) contained in the organic solvent phase of the culture broth were quantified by GC-MS. Three independent experiments were conducted. Bars show standard error.

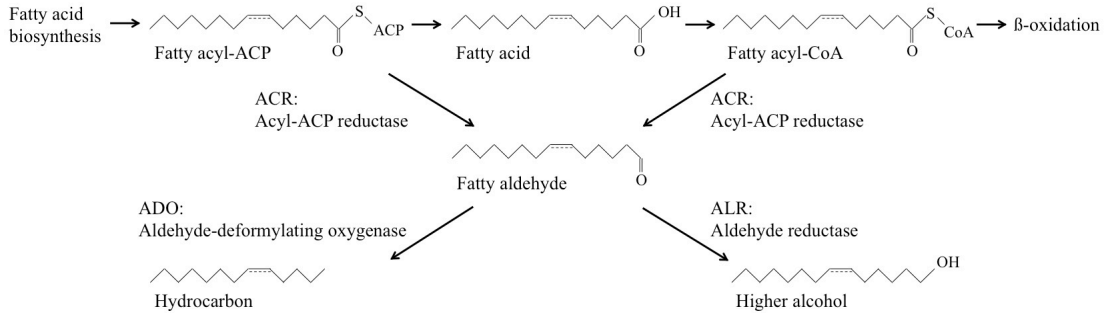


Figure 10. Higher alcohol and hydrocarbon production pathway.

ACR converts fatty acyl-ACP and fatty acyl-CoA into fatty aldehyde. ADO converts fatty aldehyde into hydrocarbon. ALR converts fatty aldehyde into higher alcohol. The *C. thermocellum* strain engineered in this study to express *Se*ACR and *Se*ADO produced higher alcohols. Our data suggest that ALRs in *C. thermocellum* might convert fatty aldehydes into higher alcohols.

ALRs in *E. coli*	Putative ALRs in *C. thermocellum*	Max score	Total score	Query cover	E value	Identity	Description
AdhE	Clo1313_1798	1145	1145	96%	0	63%	acetaldehyde dehydrogenase
	Clo1313_1827	139	139	42%	3.0E-36	30%	alcohol dehydrogenase
	Clo1313_0166	132	132	31%	6.0E-34	30%	alcohol dehydrogenase
	Clo1313_2130	82.4	82.4	43%	1.0E-17	25%	NADH-dependent butanol dehydrogenase
YqhD	Clo1313_2130	276	276	95%	9.0E-90	38%	NADH-dependent butanol dehydrogenase
	Clo1313_1827	118	118	91%	1.0E-30	28%	alcohol dehydrogenase
	Clo1313_0166	86.3	86.3	72%	9.0E-20	24%	alcohol dehydrogenase
	Clo1313_1798	61.2	61.2	92%	4.0E-11	23%	acetaldehyde dehydrogenase
FucO	Clo1313_1798	232	232	97%	4.0E-69	36%	acetaldehyde dehydrogenase
	Clo1313_0166	171	171	100%	1.0E-49	31%	alcohol dehydrogenase
	Clo1313_1827	164	164	96%	5.0E-47	32%	alcohol dehydrogenase
	Clo1313_2130	100	100	95%	1.0E-24	24%	NADH-dependent butanol dehydrogenase
YiaY	Clo1313_1798	227	227	94%	2.0E-67	38%	acetaldehyde dehydrogenase
	Clo1313_1827	191	191	98%	3.0E-57	32%	alcohol dehydrogenase
	Clo1313_0166	149	149	97%	1.0E-41	28%	alcohol dehydrogenase
	Clo1313_2130	108	108	89%	2.0E-27	25%	NADH-dependent butanol dehydrogenase
EutG	Clo1313_1798	221	221	87%	4.0E-65	36%	acetaldehyde dehydrogenase
	Clo1313_1827	172	172	93%	6.0E-50	31%	alcohol dehydrogenase
	Clo1313_0166	155	155	93%	7.0E-44	29%	alcohol dehydrogenase
	Clo1313_2130	80.5	80.5	82%	2.0E-17	22%	NADH-dependent butanol dehydrogenase
GldA	Clo1313_1827	52.4	52.4	83%	2.0E-09	25%	alcohol dehydrogenase
AdhP	Clo1313_1833	108	108	90%	9.0E-28	29%	alcohol dehydrogenase
	Clo1313_0076	95.1	95.1	80%	3.0E-23	28%	alcohol dehydrogenase GroES domain protein
YjgB	Clo1313_1833	82	82	61%	1.0E-18	29%	alcohol dehydrogenase
	Clo1313_0076	54.3	54.3	92%	3.0E-09	24%	alcohol dehydrogenase GroES domain protein
YahK	Clo1313_1833	70.5	70.5	64%	1.0E-14	28%	alcohol dehydrogenase
	Clo1313_0076	60.8	60.8	76%	3.0E-11	27%	alcohol dehydrogenase GroES domain protein
YghA	Clo1313_0815	98.2	98.2	82%	7.0E-25	33%	short-chain dehydrogenase
	Clo1313_1283	97.8	97.8	83%	9.0E-25	29%	3-ketoacyl-ACP reductase
	Clo1313_0720	58.9	58.9	76%	4.0E-11	28%	enoyl-ACP reductase
	Clo1313_2272	58.5	58.5	82%	5.0E-11	24%	short-chain dehydrogenase
YbbO	Clo1313_1283	64.3	64.3	68%	3.0E-13	26%	3-ketoacyl-ACP reductase
	Clo1313_0815	55.5	55.5	69%	6.0E-10	25%	short-chain dehydrogenase
DkgA	Clo1313_1992	67.4	67.4	91%	8.0E-14	30%	aldo/keto reductase
	Clo1313_2081	58.2	58.2	60%	8.0E-11	25%	aldo/keto reductase
	Clo1313_1946	47.8	47.8	92%	2.0E-07	23%	voltage-gated potassium channel
BetA	-						

Table 2. Putative aldehyde reductases in *C. thermocellum*.

Homologous *E. coli* aldehyde reductases genes in *C. thermocellum* were identified by NCBI protein BLAST search. Genes showing an E value <1.0E-06 are listed. No gene homologous to BetA was identified.

Chapter 5. Discussion

Biorefinery, the conversion of cellulosic biomass into the products, such as fuels, plastics and other chemicals, is an attractive and feasible strategy to circumvent the energy and environmental concerns of petroleum refinery. Biofuel production is one of the most important concepts in biorefinery, because we have consumed a large amount of petroleum as liquid fuels [5]. Only a few economic biofuel productions from cellulosic biomass have been achieved so far. We need to innovate the way to reduce the cost of biofuel production for further prevalence of biofuel.

Conventional biofuel production processes include expensive and complicated steps. Cellulosic biomass is milled, chemically treated and enzymatically treated to produce fermentable sugars. The sugars are converted to ethanol by yeast. The ethanol is distillated to purify it. Especially, the chemical pretreatment step, the enzymatic degradation step, and distillation step increase the cost of biofuel production [13].

Cellulosic biomass is a highly recalcitrant material, the degradation of cellulosic biomass chemical pretreatment and large amount of expensive lignocellulose-degrading enzyme solution. Clostridia are anaerobic soil bacteria, and some species originally produce a set of lignocellulose-degrading enzymes, and they can degrade cellulosic biomass without chemical pretreatment or the addition of lignocellulose-degrading enzymes [14]. Thus, utilization of clostridia opens the possibility for the significant cost reduction of biofuel production with eliminating the chemical pretreatment and the supplementation of lignocellulose-degrading enzymes.

The feasible way for the biofuel production with clostridia is the consolidated bioprocessing, in which a cellulolytic bacterium degrades cellulosic biomass and produces fuel compounds, simultaneously. Because the clostridia originally produces slight amount of ethanol, they are candidate bacteria for the consolidated bioprocessing. It is suggested the consolidated bioprocessing could reduce the cost of biofuel production significantly [13]. Ethanol is not attractive feature as fuel, high energy consumption of distillation step to purify it, for example. We can collect insoluble fuel compounds, e.g. hydrocarbon, higher alcohol, fatty acid ethyl ester, without distillation. There are some oleaginous microorganisms accumulate such insoluble fuel compounds in the cell, harvesting the cells and extracting the products also increase the cost of biofuel production [106,107], suggesting that product secretion is an important factor for the cost reduction. I reported in chapter 3, *C. thermocellum* expressing an acyl-ACP reductase and an aldehyde-deformylating oxgenase produced and secreted higher alcohols. This result suggesting the possibility that the engineered *C.*

thermocellum can secrete these water-insoluble fuel compounds from cellulosic biomass without costly process, the supplementation of lignocellulose-degrading enzyme solution and the distillation of fuel compounds, however, the titer is quite low so far, 3.1 and 19.5 μg of decanol and dodecanol, respectively.

These compounds were synthesized via fatty acid synthesis pathway (Fig. 10 in chapter 3). In some papers, the increased titers of the fatty acid derived fuel compounds in *E. coli* have been reported. Deletion of the second gene of the ß-oxidation pathway and overexpression of thioesterases produced 1.2 g/L of fatty acids, which were directly converted into fatty alcohols (0.6 g/L) by overexpression of fatty acyl-CoA reductases [109]. A dynamic sensor-regulator system has been used to balance fatty acid pathway intermediates, producing 1.5 g/L of fatty acid ethyl esters [110]. Eliminating fatty acid breakdown via ß-oxidation, while overexpressing the rate-limiting acetyl-CoA carboxylase, the engineered *E. coli* produced 4.5 g/L of fatty acids [111,112]. The engineering with reversing ß-oxidation allowed for production of 7 g/L of extracellular fatty acids [113]. These strategies might be applied for the engineering of the clostridia to increase the titers of fatty acid derived fuel compounds such as higher alcohols.

As I mentioned above, the engineered strain of *C. thermocellum* successfully produced higher alcohols, however the titers were quite low. Some investigators reported the examples of biofuel production from the engineered clostridia. *C. cellulyticum* expressing pyruvate decarboxylase and alcohol dehydrogenase from *Z. mobilis* produced 0.8 g/L of ethanol [82]. The other engineered strain of *C. cellulolyticum* produced 0.7 g/l of isobutanol [83]. The similar engineering generated *C. thermocellum* strain producing 5.4 g/L of isobutanol. In the other study, a mutant strain of *C. thermocellum* produced 3.4 g/L of ethanol by removing side product formation [114].

Recent progresses in metabolic engineering and synthetic biology generate the powerful *Escherichia coli* strains for biorefinery. The engineered *E. coli* strains produced elevated titers of biofuel, e.g. 30 g/L n-butanol, 51 g/L isobutanol, and 143 g/L isopropanol, from the fermentable sugars [115-117]. The results indicate that the engineering for fuel compound productions in the cellulolytic bacteria needs the breakthroughs to increase the titers.

Improvement of the efficiency of bacterial cellulosic biomass degradation is a subject of great interest. The clostridia secrete cellulosome, which are multienzyme machines for the degradation of cellulosic biomass. Scaffolding subunit of cellulosome is essential for the cellulosic biomass degradation ability, but overexpression of the scaffolding subunit did not increase the cellulose degradation efficiency of *Clostridium thermocellum* [37]. *Clostridium cellulolyticum* engineered with the heterologous expression of Cel9A from

Clostridium phytofermentans degraded pure cellulose 1.5 times faster than wild type starin [118]. The decrease of pyruvate accumulation by heterologous expression of pyruvate decarboxylase and alcohol dehydrogenase from *Zymomonas mobilis* also leaded to 150% increase of cellulose consumption in *C. cellulolyticum* [82].

Cellulosome could degrade cellulosic biomass and mainly produce cellobiose. Cellobiose inhibits the activity of cellulosome [119]. Prawitwong *et al*. reported that supplementation of only one cellobiose-degrading enzyme, ß-glucosidase, increased the cellulose degradation ability of *C. thermocellum*, resulting complete degradation of 10 % pure cellulose, suggesting that the cellobiose accumulation is one of the limiting factor in cellulosic biomass degradation by clostridia. In this system, the most part of glucans in a cellulosic biomass, alkaline-pretreated rice straw, was degraded, and approx. 8 % of fermentable sugars was produced. The fermentable sugars can be converted to ethanol by yeast [120]. This report suggested that the cultivation of *C. thermocellum* supplemented with ß-glucosidase could be applied for the fermentable sugar production for biofuel production. In this thesis, I reported successful expression of foreign genes in *C. thermocellum*. It is expected that further engineering of *C. thermocellum*, e.g. the expression and the secretion of ß-glucosidase, will decrease cost of cellulosic biomass degradation and fermentable sugar production. As I mentioned above, the engineered *E. coli* strains produce higher titers of n-butanol, isobutanol, and isopropanol from fermentable sugars. These engineering of bacterial production for fermentable sugar and fuel compound production will be a feasible way to achieve economic biofuel production.

Summary

For sustainable development of human societies, we should produce the products such as biofuel from renewable and environmentally-friendly materials instead of petroleum. Cellulosic biomass is a candidate material for the biofuel production. To achieve commercial biofuel production from cellulosic biomass, the significant cost reduction is required. Here I proposed the biofuel production strategy in which a cellulolytic bacterium degrades cellulosic biomass and produces fuel compounds, simultaneously.

Enzymatic degradation of cellulosic biomass is the costly process in the biofuel production. To understand the detail feature of the enzymatic degradation of cellulosic biomass, one of the most powerful lignocellulose-degrading enzyme produced from a cellulolytic bacterium was characterized. The lignocellulose-degrading enzyme harbor a carbohydrate-binding module, which is used to determine the binding character of the enzyme in this study. I report the binding characters at the molecular level: the carbohydrate-binding module binds to various components of cellulosic biomass such as crystalline and non-crystalline form of cellulose, and xyloglucan.

Few researchers have reported the examples of engineering of lignocellulose-degrading enzymes for improvement of the degradation ability of cellulosic biomass so far. Here, I show that the engineered lignocellulose-degrading enzyme with the replacement of carbohydrate-binding modules changed the specificity of the activity. Additionally, I also showed that the engineered enzyme could efficiently degrade a cellulosic biomass, milled corn hull, than the original enzyme, suggesting the possibility that the replacement of carbohydrate-binding modules will be one of the ways to improve the degradation ability of lignocellulose-degrading enzymes for cellulosic biomass.

Some cellulolytic bacteria produce water-soluble fuel compounds such as ethanol. To collect water-soluble fuel compounds from the culture broth, a distillation step is required, which also significantly increases the cost of biofuel production. In contrast to the water-soluble fuel compounds, it is expected that water-insoluble compounds such as higher alcohols are more easily collected from the culture broth by phase separation due to their insolubility in water. Here I report that the engineered cellulolytic bacterium produced higher alcohols from cellulose in which the costly processes, the supplementation of lignocellulose-degrading enzymes and distillation of fuel compounds, were eliminated.

These results in the characterization and the engineering of lignocellulose-degrading enzyme and cellulolytic bacteria in this study will be basic information of process developments for the further cost

reduction of the biofuel production, resulting the establishment of economic biofuel production from cellulosic biomass.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr. Shuichi Karita for the continuous support of my Ph. D study. I would like to thank Dr. Issei Kobayashi and Dr. Takeo Yano for the support to prepare the thesis. I appreciate to Dr. Kazuo Sakka, Mie University, for kindly sharing the gene pulser. We thank Drs. Lee R. Lynd and Daniel G. Olson, Dartmouth College, for their suggestions regarding *C. thermocellum* transformation. We thank Drs. Takashi Mishima and Naoto Isono, Mie University, for their instruction in GC-MS analysis.

References

1. Climate Change 2014 Synthesis Report Summary for Policymakers. International Panel on Climate Change.
2. White paper of energy. 2004. Agency for Natural Resources and Energy. Japan.
3. BP Statistical Review of World Energy. 2005. BP public limited company.
4. The asian biomass handbook - a guide for biomass production and utilization -. 2008. The Japan Institute of Energy.
5. http://oil-info.ieej.or.jp/index.html. The Oil Information Center.
6. http://poet-dsm.com. POET-DSM Advanced Biofuels
7. http://www.dupont.com. E. I. du Pont de Nemours and Company.
8. http://www.abengoabioenergy.com/web/en/. Abengoa Bioenergy.
9. http://www.betarenewables.com. Beta Renewables.
10. http://www.nedo.go.jp/content/100513589.pdf. 2012. Ministry of Economy, Trade and Industry. Japan.
11. Strategic Energy Plan 2014. Ministry of Economy, Trade and Industry. Japan.
12. **Chundawat, SPS, Beckham GT, Himmel ME, Dale BE.** 2011. Deconstruction of lignocellulosic biomass to fuels and chemicals. Annu. Rev. Chem. Biomol. Eng. **2**:121–145.
13. **McAloon A, Taylor F, Yee W, Ibsen K, Wooley R.** 2000. Determining the cost of producing ethanol from corn starch and lignocellulosic feedstocks. National Renewable Energy Laboratory Technical Report NREL/TP-580-28893.
14. **Lynd LR, van Zyl WH, McBride JE, Laser M.** 2005. Consolidated bioprocessing of cellulosic biomass: an update. Curr. Opin. Biotechnol. **16**:577-583.
15. **Doi RH, Kosugi A.** 2004. Cellulosomes: plant-cell-wall-degrading enzyme complexes. Nat. Rev. Microbiol. **2**:541-551.
16. **Olson DG, Giannone RJ, Hettich RL, Lynd LR.** 2013. Role of the cipA scaffoldin protein in cellulose solubilization, as determined by targeted gene deletion and complementation in *Clostridium thermocellum*. J. Bacteriol. **195**:733-739.
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. **Arantes V, Saddler JN.** 2010. Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. Biotechnol. Biofuels **3**:4-14.

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

20. **Maglione G, Russell JB, Wilson DB.** 1997. Kinetics of cellulose digestion by *Fibrobacter succinogenes* S85. Appl. Environ. Microbiol. **63**:665-669.

21. **Jung H, Wilson DB, Walker LP.** 2002. Binding mechanisms for *Thermobifida fusca* Cel5A, Cel6B, and Cel48A cellulose-binding modules on bacterial microcrystalline cellulose. Biotechnol. Bioeng. **80**:380-392.

22. **Gal L, Pages S, Gaudin C, Belaich A, Reverbel-Leroy C, Tardif C, Belaich JP.** 1997. Characterization of the cellulolytic complex (cellulosome) produced by *Clostridium cellulolyticum*. Appl. Environ. Microbiol. **63**:903-909.

23. **Goldstein MA, Takagi M, Hashida S, Shoseyo 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.

24. **Morag E, Lapidot A, Govorko D, Lamed R, Wilchek M, Bayer EA, Shoham Y.** 1995. Expression, purification, and characterization of the cellulose-binding domain of the scaffoldin subunit from the cellulosome of *Clostridium thermocellum*. Appl. Environ. Microbiol. **61**:1980-1986.

25. **Sabathé F, Soucaille P.** 2003. Characterization of the cipA scaffolding protein and in vivo production of a minicellulosome in *Clostridium acetobutylicum*. J. Bacteriol. **185**:1092-1096.

26. **Bansal P, Hall M, Realff MJ, Lee JH, Bommarius AS.** 2009. Modeling cellulase kinetics on lignocellulosic substrates. Biotechnol. Adv. **27**:833-848.

27. **Burton RA, Gidley MJ, Fincher GB.** 2010. Heterogeneity in the chemistry, structure and function of plant cell walls. Nat. Chem. Biol. **10**:724-732.

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

29. **Sukhumavasi J, Ohmiya K, Shimizu S, Ueno K.** 1988. *Clostridiurn josui* sp. nov., a cellulolytic,

moderate thermophilic species from Thai compost. Int. J. Syst. Bacteriol. **38**:179-182.

30. **Zhang YHP, Himmel ME, Mielenz JR.** 2006. Outlook for cellulase improvement: Screening and selection strategies. Biotechnol. Adv. **24**:452–481.
31. **Boraston AB, Bolam DN, Gilbert HJ, Davies GJ.** 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. **382**:769–781.
32. **Yagüe E, Béguin P, Aubert JP.** 1990. Nucleotide sequence and deletion analysis of the cellulase-encoding gene *celH* of *Clostridium thermocellum*. Gene **89**:61–67.
33. **Kakiuchi M, Isui A, Suzuki K, Fujino T, Fujino E, Kimura T, Karita S, Sakka K, Ohmiya K.** 1998. Cloning and DNA sequencing of the genes encoding *Clostridium josui* scaffolding protein CipA and cellulase CelD and identification of their gene products as major components of the cellulosome. J. Bacteriol. **180**:4303-4308.
34. **Sakka K, Kojima Y, Kondo T, Karita S, Ohmiya K, Shimada K.** 1993. Nucleotide sequence of the *Clostridium stercorarium xynA* gene encoding xylanase A: identification of catalytic and cellulose binding domains. Biosci. Biotechnol. Biochem. **57**:273-277.
35. **Schirmer A, Rude MA, Li X, Popova E, del Cardayre SB.** 2010. Microbial biosynthesis of alkanes. Science. **329**:559-562.
36. **Hong W, Zhang J, Feng Y, Mohr G, Lambowitz AM, Cui GZ, Liu YJ, Cui Q.** 2014. The contribution of cellulosomal scaffoldins to cellulose hydrolysis by *Clostridium thermocellum* analyzed by using thermotargetrons. Biotechnol. Biofuels. **7**:80.
37. **Olson DG, Giannone RJ, Hettich RL, Lynd LR.** 2013. Role of the cipA scaffoldin protein in cellulose solubilization, as determined by targeted gene deletion and complementation in *Clostridium thermocellum*. J. Bacteriol. **195**:733-739.
38. **Zverlov VV, Klupp M, Krauss J, Schwarz WH.** 2008. Mutations in the scaffoldin gene, cipA, of *Clostridium thermocellum* with impaired cellulosome formation and cellulose hydrolysis: insertions of a new transposable element, IS 1447, and implications for cellulase synergism on crystalline cellulose. J. Bacteriol. **190**:4321-4327.
39. **McLean BW, Boraston AB, Brouwer D, Sanaie N, Fyfe CA, Warren RAJ, Kilburn DG, Haynes CA.** 2002. Carbohydrate-binding modules recognize fine substructures of cellulose. J. Biol. Chem. **277**:50245-50254.
40. **Jindou S, Xu Q, Kenig R, Shulman M, Shoham Y, Bayer EA, Lamed R.** 2006. Novel architecture

of family-9 glycoside hydrolases identified in cellulosomal enzymes of *Acetivibrio cellulolyticus* and *Clostridium thermocellum*. FEMS Microbiol. Lett. **254**:308-314.

41. **Gill SC, Hippel PH.** 1989. Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. **182**:319-326.
42. **Gilkes NR, Jervis E, Henrissat B, Tekant B, Miller Jr RC, Warren RA, Kilburn DG.** 1992. The adsorption of a bacterial cellulase and its two isolated domains to crystalline cellulose. J. Biol. Chem. **267**:6743-6749.
43. **Wood TM.** 1988. Preparation of crystalline, amorphous, and dyed cellulase substrates. Methods Enzymol. **160**:19-25.
44. **Araki Y, Karita S, Tsuchiya T, Kondo M, Goto M.** 2010. Family 17 and 28 carbohydrate-binding modules discriminated different cell-wall sites in sweet potato roots. Biosci. Biotechnol. Biochem. **74**:802-805.
45. **McCartney L, Gilbert HJ, Bolam DN, Boraston AB, Knox JP.** 2004. Glycoside hydrolase carbohydrate-binding modules as molecular probes for the analysis of plant cell wall polymers. Anal. Biochem. **326**:49-54.
46. **Araki Y, Karita S, Tanaka A. Kondo M, Goto M.** 2009. Characterization of family 17 and family 28 carbohydrate-binding modules from *Clostridium josui* Cel5A. Biosci. Biotechnol. Biochem. **73**:1028-1032.
47. **Creagh AL, Ong E, Jervis E, Kilburn DG, Haynes CA.** 1996. Binding of the cellulose-binding domain of exoglucanase Cex from *Cellulomonas fimi* to insoluble microcrystalline cellulose is entropically driven. Proc. Natl. Acad. Sci. U. S. A. **93**:12229-12234.
48. **Guo J, Catchmark JM.** 2013. Binding specificity and thermodynamics of cellulose-binding modules from *Trichodema reesei* Cel7A and Cel6A. Biomacromol. **14**:1268-1277.
49. **Ståhlberg J, Johansson G, Pettersson G.** 1991. A new model for enzymatic hydrolysis of cellulose based on the two-domain structure of cellobiohydrolase I. Bio/technol. **9**:286-290.
50. **Tormo J, Lamed R, Chirino AJ, Morag E, Bayer EA, Shoham Y, Steitz TA.** 1996. Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. EMBO J. **15**:5739-5751.
51. **Shimon LJ, Pagés S, Belaich A, Belaich JP, Bayer EA, Lamed R, Shoham Y, Frolow F.** 2000. Structure of a family IIIa scaffoldin CBD from the cellulosome of *Clostridium cellulolyticum* at 2.2 Å

resolution. Acta Crystallogr. Sect. D. **56**:1560-1568.

52. **Igarashi K, Wada M, Hori R, Samejima M.** 2006. Surface density of cellobiohydrolase on crystalline celluloses. FEBS Journal. **273**:2869-2878.

53. **Sugimoto N, Igarashi K, Wada M, Samejima M.** 2012. Adsorption characteristics of fungal family 1 cellulose-binding domain from *Trichoderma reesei* cellobiohydrolase 1 on crystalline cellulose: negative cooperative adsorption via a steric exclusion effect. Langmuir **28**:14323-14329.

54. **Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD.** 2007. Biomass recalcitrance: engineering plants and enzymes for biofuels production. Science **315**:804-807.

55. **Wang Y, Tang R, Tao J, Wang X, Zheng B, Feng Y.** 2012. Chimeric cellulase matrix for investigating intramolecular synergism between non-hydrolytic disruptive functions of carbohydrate-binding modules and catalytic hydrolysis. J. Biol. Chem. **287**:29568-29578.

56. **Miller GL.** 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal. Biochem. **31**:426-428.

57. **Nozaki K, Nishijima H, Arai T, Mizuno M, Sato N, Amano Y.** 2011. Regulation of adsorption behavior of carbohydrate-binding module family 1 and endo-ß-1,4-glucanase onto crystalline cellulose. J. Appl. Glycosci. **58**:133-138.

58. **Ichikawa S, Karita S, Kondo M, Goto M.** 2014. Cellulosomal carbohydrate-binding module from *Clostridium josui* binds to crystalline and non-crystalline cellulose, and soluble polysaccharides. FEBS Lett. **588**:3886-3890.

59. **Bharali S, Purama RK, Majumder A, Fontes CMGA, Goyal A.** 2007. Functional characterization and mutation analysis of family 11, carbohydrate-binding module (*Ct*CBM11) of cellulosomal bifunctional cellulase from *Clostridium thermocellum*. Indian J. Microbiol. **47**:109-118.

60. **Carvalho AL, Goyal A, Prates JAM, Bolam DN, Gilbert HJ, Pires VMR, Ferreira LMA, Planas A, Romão MJ, Fontes CMGA.** 2004. The family 11 carbohydrate-binding module of *Clostridium thermocellum* Lic26A-Cel5E accommodates ß-1,4- and ß-1,3−1,4-mixed linked glucans at a single binding site. J. Biol. Chem. **279**:34785-34793.

61. **Boraston AB, McLean BW, Chen G, Li A, Warren RAJ, Kilburn DG.** 2002. Co-operative binding of triplicate carbohydrate-binding modules from a thermophilic xylanase. Mol. Microbiol. **43**:187–194.

62. **Yabushita M, Kobayashi H, Hara K, Fukuoka A.** 2014. Quantitative evaluation of ball-milling effects on the hydrolysis of cellulose catalysed by activated carbon. Catal. Sci. Technol. **4**:2312-2317.

63. **Gilbert HJ.** 2003. How carbohydrate binding modules overcome ligand complexity. Structure **11**:609-610.

64. **Youngs H, Somerville C.** 2014. Plant science. Best practices for biofuels. Science **344**:1095-1096.

65. **Hromádková Z, Ebringerová A.** 1995. Isolation and characterization of hemicelluloses of corn hulls. Chem. Papers **49**:97-101.

66. **Tett SFB, Stott PA, Allen MR, Ingram WJ, Mitchell JFB.** 1999. Causes of twentieth-century temperature change near the Earth's surface. Nature **399**:569-572.

67. **Tilman D, Hill J, Lehman C.** 2006. Carbon-negative biofuels from low-input high-diversity grassland biomass. Science **314**:1598-1600.

68. **Regalbuto JR.** 2009. Cellulosic biofuels - got gasoline? Science **325**:822-824.

69. **Stephanopoulos G.** 2007. Challenges in engineering microbes for biofuels production. Science **315**:801-804.

70. **Peralta-Yahya PP, Zhang F, del Cardayre SB, Keasling JD.** 2012. Microbial engineering for the production of advanced biofuels. Nature **488**:320-328.

71. **Chung D, Cha M, Guss AM, Westpheling J.** 2014. Direct conversion of plant biomass to ethanol by engineered *Caldicellulosiruptor bescii*. Proc. Natl. Acad. Sci. U. S. A. **111**:8931-8936.

72. **Olson DG, McBride JE, Shaw AJ, Lynd LR.** 2012. Recent progress in consolidated bioprocessing. Curr Opin Biotechnol. **23**:396-405.

73. **Mazzoli R, Lamberti C, Pessione E.** 2012. Engineering new metabolic capabilities in bacteria: lessons from recombinant cellulolytic strategies. Trends Biotechnol. **30**:111-119.

74. **Yamada R, Hasunuma T, Kondo A.** 2013. Endowing non-cellulolytic microorganisms with cellulolytic activity aiming for consolidated bioprocessing. Biotechnol. Adv. **31**:754-63.

75. **Hasunuma T, Okazaki F, Okai N, Hara KY, Ishii J, Kondo A.** 2013. A review of enzymes and microbes for lignocellulosic biorefinery and the possibility of their application to consolidated bioprocessing technology. Bioresour. Technol. **135**:513-522.

76. **Gronenberg LS, Marcheschi RJ, Liao JC.** 2013. Next generation biofuel engineering in prokaryotes. Curr. Opin. Chem. Biol. **17**:462-471.

77. **Izquierdo JA, Pattathil S, Guseva A, Hahn MG, Lynd LR.** 2014. Comparative analysis of the ability of *Clostridium clariflavum* strains and *Clostridium thermocellum* to utilize hemicellulose and unpretreated plant material. Biotechnol. Biofuels. **7**:136.

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

79. **Bayer EA, Lamed R, White BA, Flint HJ.** 2008. From cellulosomes to cellulosomics. Chem. Rec. **8**:364-377.

80. **Akinosho H, Yee K, Close D, Ragauskas A.** 2014. The emergence of *Clostridium thermocellum* as a high utility candidate for consolidated bioprocessing applications. Front Chem. **2**:66.

81. **Blumer-Schuette SE, Brown SD, Sander KB, Bayer EA, Kataeva I, Zurawski JV, Conway JM, Adams MW, Kelly RM.** 2014. Thermophilic lignocellulose deconstruction. FEMS Microbiol. Rev. **38**:393-448.

82. **Guedon E, Desvaux M, Petitdemange H.** 2002. Improvement of cellulolytic properties of *Clostridium cellulolyticum* by metabolic engineering. Appl. Environ. Microbiol. **68**:53-58.

83. **Higashide W, Li Y, Yang Y, Liao JC.** 2011. Metabolic engineering of *Clostridium cellulolyticum* for production of isobutanol from cellulose. Appl. Environ. Microbiol. **77**:2727-2733.

84. **Lin PP, Mi L, Morioka AH, Yoshino KM, Konishi S, Xu SC, Papanek BA, Riley LA, Guss AM, Liao JC.** 2015. Consolidated bioprocessing of cellulose to isobutanol using *Clostridium thermocellum*. Metab. Eng. **31**:44-52.

85. **Argyros DA, Tripathi SA, Barrett TF, Rogers SR, Feinberg LF, Olson DG, Foden JM, Miller BB, Lynd LR, Hogsett DA, Caiazza NC.** 2011. High ethanol titers from cellulose by using metabolically engineered thermophilic, anaerobic microbes. Appl. Environ. Microbiol. **77**:8288-8294.

86. **Xue C, Zhao JB, Chen LJ, Bai FW, Yang ST, Sun JX.** 2014. Integrated butanol recovery for an advanced biofuel: current state and prospects. Appl. Microbiol. Biotechnol. **98**:3463-3474.

87. **Williams TI, Combs JC, Lynn BC, Strobel HJ.** 2007. Proteomic profile changes in membranes of ethanol-tolerant *Clostridium thermocellum*. Appl. Microbiol. Biotechnol. **74**:422-432.

88. **Yang S, Giannone RJ, Dice L, Yang ZK, Engle NL, Tschaplinski TJ, Hettich RL, Brown SD.** 2012. *Clostridium thermocellum* ATCC27405 transcriptomic, metabolomic and proteomic profiles after ethanol stress. BMC Genomics. **13**:336.

89. **Timmons MD, Knutson BL, Nokes SE, Strobel HJ, Lynn BC.** 2009. Analysis of composition and structure of *Clostridium thermocellum* membranes from wild-type and ethanol-adapted strains. Appl. Microbiol. Biotechnol. **82**:929-939.

90. **Berberich JA, Knutson BL, Strobel HJ, Tarhan S, Nokes SE, Dawson KA.** 2000. Toxicity effects of

compressed and supercritical solvents on thermophilic microbial metabolism. Biotechnol. Bioeng. **70**:491-497.

91. **Doshi R, Nguyen T, Chang G.** 2013. Transporter-mediated biofuel secretion. Proc. Natl. Acad. Sci. U. S. A. **110**:7642-7647.
92. **Olson DG, Lynd LR.** 2012. Transformation of *Clostridium thermocellum* by electroporation. Methods Enzymol. **510**:317-330.
93. **Tripathi SA, Olson DG, Argyros DA, Miller BB, Barrett TF, Murphy DM, McCool JD, Warner AK, Rajgarhia VB, Lynd LR, Hogsett DA, Caiazza NC.** 2010. Development of pyrF-based genetic system for targeted gene deletion in Clostridium thermocellum and creation of a pta mutant. Appl. Environ. Microbiol. **76**:6591-6599.
94. **Choi YJ, Lee SY.** Microbial production of short-chain alkanes. 2013. Nature **502**:571-574.
95. **Guss AM, Olson DG, Caiazza NC, Lynd LR.** 2012. Dcm methylation is detrimental to plasmid transformation in *Clostridium thermocellum*. Biotechnol. Biofuels. **5**:30.
96. **Olson DG, Lynd LR.** 2012. Computational design and characterization of a temperature-sensitive plasmid replicon for gram positive thermophiles. J. Biol. Eng. **6**:5.
97. **Zhang J, Lu X, Li JJ.** 2013. Conversion of fatty aldehydes into alk(a/e)nes by *in vitro* reconstituted cyanobacterial aldehyde-deformylating oxygenase with the cognate electron transfer system. Biotechnol. Biofuels **6**:86.
98. **Ichihara K, Shibahara A, Yamamoto K, Nakayama T.** 1996. An improved method for rapid analysis of the fatty acids of glycerolipids. Lipids **31**:535-539.
99. **Nakamura Y, Kaneko T, Sato S, Ikeuchi M, Katoh H, Sasamoto S, Watanabe A, Iriguchi M, Kawashima K, Kimura T, Kishida Y, Kiyokawa C, Kohara M, Matsumoto M, Matsuno A, Nakazaki N, Shimpo S, Sugimoto M, Takeuchi C, Yamada M, Tabata S.** 2002. Complete genome structure of the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1. DNA Res. **9**:123-130.
100. **Lin F, Das D, Lin XN, Marsh EN.** 2013. Aldehyde-forming fatty acyl-CoA reductase from cyanobacteria: expression, purification and characterization of the recombinant enzyme. FEBS J. **280**:4773-4781.
101. **Li N, Chang WC, Warui DM, Booker SJ, Krebs C, Bollinger JM Jr.** 2012. Evidence for only oxygenative cleavage of aldehydes to alk(a/e)nes and formate by cyanobacterial aldehyde

decarbonylases. Biochemistry **51**:7908-7916.

102. **Andre C, Kim SW, Yu XH, Shanklin J.** 2013. Fusing catalase to an alkane-producing enzyme maintains enzymatic activity by converting the inhibitory byproduct H_2O_2 to the cosubstrate O_2. Proc. Natl. Acad. Sci. U. S. A. **110**:3191-3196.

103. **Lennen RM, Pfleger BF.** 2013. Microbial production of fatty acid-derived fuels and chemicals. Curr. Opin. Biotechnol. **24**:1044-1053.

104. **Akhtar MK, Turner NJ, Jones PR.** 2013. Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities. Proc. Natl. Acad. Sci. U. S. A. **110**:87-92.

105. **Rodriguez GM, Atsumi S.** 2014. Toward aldehyde and alkane production by removing aldehyde reductase activity in *Escherichia coli*. Metab. Eng. **25**:227-237.

106. **Alper H, Stephanopoulos G.** 2009. Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? Nat. Rev. Microbiol. **7**:715-723.

107. **Huffer S, Roche CM, Blanch HW, Clark DS.** 2012. *Escherichia coli* for biofuel production: bridging the gap from promise to practice. Trends Biotechnol. **30**:538-545.

108. **Kung Y, Runguphan W, Keasling JD**. 2012. From fields to fuels: recent advances in the microbial production of biofuels. ACS Synth. Biol. 2012 **1**:498-513.

109. **Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, Del Cardayre SB, Keasling JD.** 2010. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. Nature **463**:559-562.

110. **Zhang F, Carothers JM, Keasling JD.** 2012. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. Nat. Biotechnol. **30**:354-359.

111. **Liu T, Vora H, Khosla C.** 2010. Quantitative analysis and engineering of fatty acid biosynthesis in *E. coli*. Metab. Eng. **12**:378-386.

112. **Liu H, Yu C, Feng D, Cheng T, Meng X, Liu W, Zou H, Xian M.** 2012. Production of extracellular fatty acid using engineered *Escherichia coli*. Microb. Cell Fact. **11**:41.

113. **Dellomonaco C, Clomburg JM, Miller EN, Gonzalez R.** 2011. Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals. Nature **476**:355-359.

114. **Papanek B, Biswas R, Rydzak T, Guss AM.** 2015. Elimination of metabolic pathways to all traditional fermentation products increases ethanol yields in *Clostridium thermocellum*. Metab. Eng. **32**:49-54.

115. **Shen CR, Lan EI, Dekishima Y, Baez A, Cho KM, Liao JC.** 2011. Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. Appl. Environ. Microbiol. **77**:2905-2915.

116. **Inokuma K, Liao JC, Okamoto M, Hanai T.** 2010. Improvement of isopropanol production by metabolically engineered *Escherichia coli* using gas stripping. J. Biosci. Bioeng. **110**:696-701.

117. **Baez A, Cho KM, Liao JC.** 2011. High-flux isobutanol production using engineered *Escherichia coli*: a bioreactor study with *in situ* product removal. Appl. Microbiol. Biotechnol. **90**:1681-1690.

118. **Ravachol J, Borne R, Meynial-Salles I, Soucaille P, Pagès S, Tardif C, Fierobe HP.** 2015. Combining free and aggregated cellulolytic systems in the cellulosome-producing bacterium *Ruminiclostridium cellulolyticum*. Biotechnol. Biofuels **8**:114.

119. **Gefen G, Anbar M, Morag E, Lamed R, Bayer EA.** 2012. Enhanced cellulose degradation by targeted integration of a cohesin-fused β-glucosidase into the *Clostridium thermocellum* cellulosome. Proc. Natl. Acad. Sci. U. S. A. **109**:10298-10303.

120. **Prawitwong P, Waeonukul R, Tachaapaikoon C, Pason P, Ratanakhanokchai K, Deng L, Sermsathanaswadi J, Septiningrum K, Mori Y, Kosugi A.** 2013. Direct glucose production from lignocellulose using *Clostridium thermocellum* cultures supplemented with a thermostable β-glucosidase. Biotechnol. Biofuels **6**:184.

121 Carbohydrate-Active enZymes Database: CAZy, http://www.cazy.org