Journal of Bioscience and Bioengineering
VOL. 128 No. 4, 398–404, 2019

ELSEVIER

www.elsevier.com/locate/jbiosc

Enzymatic and molecular characterization of an endoglucanase E from *Clostridium cellulovorans* 743B

Ryo Kozaki and Hideo Miyake*

Graduate School of Bioresources, Mie University, 1577 Kurimamachiya-cho, Tsu, Mie 514-8507, Japan

Received 25 January 2019; accepted 15 March 2019
Available online 10 April 2019

Endoglucanase E (EngE) is a cellulosomal enzyme of the glycoside hydrolase family 5 generated by the cellulosome-producing bacterium *Clostridium cellulovorans* 743B. Although its basic activities and properties have been characterized, its substrate specificity, product range, and steady-state kinetics remain unclear. The current study prepared recombinant EngE (rEngE) and analyzed its substrate specificity and product range using thin layer chromatography. When carboxymethyl cellulose (CMC) or phosphoric acid swollen cellulose was used as a substrate, disaccharides and trisaccharides were the main products. However, no product was detected with microcrystalline cellulose as the substrate. This indicated that rEngE is a cellulase that hydrolyzes low-crystallinity cellulose. Furthermore, products were detected when glucomannan, lichenan, or β-glucan was used, but no product was obtained with xylan. These results suggested that rEngE hydrolyzes the β-1,4 glycosidic bond between glucose residues of the substrate. In the kinetic analysis, at CMC concentrations of ≥3 mg/mL, the reaction rate decreased. Application of the above data to three substrate inhibition models generated a better fit to a model that generates products not only from the enzyme–substrate complex but also from enzyme–substrate–substrate (ESS) complexes, in which two substrates are bound to the enzymes. In addition, it was found that a carbohydrate-binding module (CBM) contained in EngE binds to cellulose. Therefore, substrate inhibition likely occurred because the binding site of CBM may correspond to one of the substrate-binding sites in the ESS complex.

[**Key words:** *Clostridium cellulovorans*; Endoglucanase; Substrate inhibition; Carbohydrate-binding module; Cellulase]

Currently, 150 glycoside hydrolase (GH) families are registered in the carbohydrate-active enzymes (CAZy) database (http://www.cazy.org), 52 of which are involved in plant cell wall degradation (1). GH5, which includes enzymes that hydrolyze cellulose, xyloglucan, mannan and xylan and others as substrates, is one of these families. Many of their substrate specificities have been determined showing EC numbers close to 20 (2). GH5, which is one of the largest GH families, is further classified into 56 subfamilies according to the CAZy database. Members of GH5 show various substrate specificities.

Clostridium cellulovorans 743B is an anaerobic and mesophilic bacterium that forms multi-enzyme complexes referred to as cellulosomes on its surface (3–5). Cellulosomes are assembled by the interaction of a dockerin domain contained in some carbohydrate-related enzymes (i.e., cellulosomal enzymes) and certain cohesin domains contained in scaffold proteins. Whole-genome analysis of *C. cellulovorans* revealed 80 genes encoding GHs, including 9 genes encoding cellulosomal enzymes belonging to the GH5 family (6,7). Among them, only mannanase A (ManA) and endoglucanase B (EngB) have been characterized so far. They have been shown to function as cellulosomal enzymes and their substrate specificity has been determined (8,9).

Recently, profiles of some native cellulosomal enzymes of *C. cellulovorans* strains adapted to various carbon sources were analyzed revealing that endoglucanase E (EngE), belonging to the GH5 family, and cellulose binding protein A (CbpA), a scaffold protein, are co-expressed (10). Scaffold proteins serve as bases for constructing cellulosomes. In addition, the *engE*-coding protein, without a signal peptide (amino acids 1–31), is a relatively large enzyme with a molecular weight of 112,000 (11), and represents a unique enzyme consisting of 3 surface-layer homology (SLH) domains that self-assemble on the surface of microbial cells, a catalytic domain belonging to the GH5 family, a carbohydrate-binding module (CBM) belonging to the CBM65 family, and a dockerin domain from the N-terminal. Although researchers such as Tamaru and Doi have already evaluated EngE activity on polysaccharides including the effects of temperature and pH, its specific substrate(s), range of products, and enzyme kinetics remain unknown (11). The objective of the current study was to further characterize EngE. It is felt that these results may help broaden the repertoire of the GH5 family of enzymes and promote the practical application of EngE.

MATERIALS AND METHODS

Cloning of engE, engEΔCBM65-Doc and cbm65 from *C. cellulovorans* The gene (NCBI Reference Sequence: WP_010073437.1) encoding EngE without the signal peptide, a nucleotide sequence encoding amino acid residues 32 to 1067, was

* Corresponding author. Tel./fax: +81 59 231 9611.
E-mail address: miyake@bio.mie-u.ac.jp (H. Miyake).

1389-1723/$ – see front matter
https://doi.org/10.1016/j.jbiosc.2019.03.013

amplified via polymerase chain reaction (PCR) using *C. cellulovorans* genomic DNA and the following primers as described previously (11,12): forward primer, 5′-ATACCATGGCAGAAGCTAACTACACAACAA-3′; reverse primer, 5′-CGGCTCGAGTTATATTGCTTTTTTTAAGAA-3′. The reaction product encoding EngE was ligated into a *Nco*I/*Xho*I-digested pET-41a(+) glutathione-*S*-transferase (GST) fusion vector (Merck, Darmstadt, Germany) to create a plasmid. The gene encoding EngE without the signal peptide, CBM65 and the dockerin domain (*engEΔCBM65-Doc*), a nucleotide sequence encoding amino acid residues 32 to 843, was amplified using PCR from *C. cellulovorans* genomic DNA and the following primers; forward primer, 5′-AAACCCGGGGCAGAAGCTAACTACACAACAAAAG-3′; reverse primer, 5′−AAAACTCGAGACCAAAATTATTATTATCCCACCACAAAC-3′. The reaction product encoding EngEΔCBM65-Doc was ligated into a *Sma*I/*Xho*I-digested pET-50b(+) NusA (Nus tag) fusion vector (Merck) to create a plasmid. The gene encoding CBM65, a nucleotide sequence encoding amino acid residues 844 to 1013, was amplified by PCR from the GST-EngE expression plasmid, described above, using the following primers; forward primer, 5′-AAACCATGGGCACAGGTGAGAGACTGGG-3′; reverse primer, 5′-AAAACTCGAGAAGATTTACAGGTTGTGCATTCTTGTAAAC-3′. A plasmid which expressed the amplified gene fragment as a fusion protein with GST was constructed via the same method used to construct GST-EngE. The sequences were confirmed using an Applied Biosystems 3100 DNA sequencer (Thermo Fisher Scientific).

Production and purification of recombinant proteins Each plasmid DNA was used for transformation in *Escherichia coli* BL21 (DE3). A transformant was grown in 2× YT medium [1.6% Hipolypepton (Wako Pure Chemical Industries, Ltd., Osaka, Japan), 1.0% yeast extract (Nacalai Tesque, Inc., Kyoto, Japan), and 0.5% NaCl] at 37°C to an optical density of 0.5−0.7 at 600 nm, induced with isopropyl β-D-thio-galactopyranoside (0.1 mM final concentration), and harvested by centrifugation after 24 h at 20°C.

Recombinant EngE (rEngE) was purified by affinity chromatography using Glutathione Sepharose 4B (GE Healthcare, Chicago, IL, USA). To remove the GST tag added to the recombinant, 1 U of enterokinase (Merck) per mg of recombinant was added. Subsequently, the recombinant enterokinase and the cleaved GST tag were removed via GST affinity chromatography. The through fractions were dialyzed against 50 mM Tris−HCl buffer (pH 7.5) and subjected to anionic exchange chromatography using a HiTrap DEAE FF (GE Healthcare). Eluted fractions were concentrated using centrifugal ultrafiltration devices, and the concentrated solution was subjected to gel filtration chromatography via HiLood 16/600 Superdex 200 pg (GE Healthcare). Mutant EngE, lacking CBM65 and the dockerin domain (rEngEΔCBM65-Doc), was purified through affinity chromatography using Ni Sepharose 6 Fast Flow (GE Healthcare). In order to remove the Nus tag added to the recombinant, 1 U of Turbo3C Protease (Accelagen, San Diego, CA, USA) per mg of recombinant was added. Subsequently, Turbo3C Protease and the cleaved Nus tag were removed using Ni affinity chromatography. GST-CBM65 was purified via affinity chromatography using Glutathione Sepharose 4B (GE Healthcare). The amount of each recombinant protein obtained at each purification step was monitored by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Substrate specificity and product range determination Low-viscosity carboxymethyl cellulose (low-viscosity CMC; Product Number: C5678, Sigma−Aldrich, St. Louis, MO, USA), microcrystalline cellulose (Merck), phosphoric acid swollen cellulose (PASC), lichenan (Megazyme, Wicklow, Ireland), glucomannan (Megazyme), β-glucan (Megazyme), birchwood xylan (Sigma−Aldrich), and beechwood xylan (Sigma−Aldrich) were used as the substrates. PASC was prepared using microcrystalline cellulose (13). To determine substrate specificity of EngE, 5.6 μg of purified rEngE and 8.0 mg of each substrate were incubated in 1 mL of McIlvaine buffer (pH 5) containing 25 mM NaCl at 25°C for 24 h. McIlvaine buffer was adjusted to specified pH by mixing 0.2 M disodium hydrogenphosphate and 0.1 M citric acid. The products were analyzed by thin-layer chromatography (TLC) (glass TLC plate, Silica gel 60, Merck). Sample separation was performed on the TLC plate using a mixture of butanol, acetic acid, and water in a 2:1:1 volume ratio. Following development, spots of sugars were visualized by heating the plate after spraying with concentrated H_2SO_4. For substrates on which no product was observed by TLC analysis, the amount of reducing sugar was quantified before and after rEngE addition using a modified version of the Somogyi−Nelson method (14), and relative activity was detected by comparing with that of CMC.

Measurement of cellulase activity EngE-catalyzed reactions were carried out using 50 nM rEngE or 1 μM rEngEΔCBM65-Doc at 25°C and pH 5.0 (in McIlvaine buffer containing 50 mM NaCl). Low-viscosity CMC was used as the substrate. Hydrolyzes of CMC were followed by a modification of the Somogyi−Nelson method (14). Reaction rates were measured at each substrate concentration (0.5−35 mg/mL), and plotted against substrate concentrations. The reaction rate was calculated using released reducing sugar concentrations (glucose equivalent per min) under above stated assay conditions.

Cellulose binding assay of CBM Binding GST-CBM65 to insoluble cellulose was carried out using the method described by Yoda et al (15). Five milligrams of crystalline cellulose was placed in illustra MicroSpin columns (GE Healthcare), dissolved in 50 mM Tris−HCl buffer (pH 7.5) containing 50 mM NaCl and washed several times by centrifugation. Subsequently, 200 μL of 2.5 μM GST-CBM65 replaced with the same buffer was added to the column, stirred, and incubated on ice for 1 h. Following centrifugation, the pellet was washed several times with the same buffer using centrifugation. Proteins bound to microcrystalline cellulose were eluted with 5 % sodium dodecyl sulfate by centrifugation. Both the washed fraction as well as the eluted fraction were subjected to SDS-PAGE. Bovine serum albumin (BSA) (Wako) was used as the negative control.

Binding of soluble cellulose to GST-CBM65 was observed via isothermal titration calorimetry, VP-ITC (MicroCal, Northampton, MA, USA), at pH 7.5 and 25°C. Twenty-eight, 10 μL portions of 50 μM GST-CBM65 were injected into 56 μM low-viscosity CMC solution. Dilution heat resulting from the injection was negligible. The data were analyzed using the software (Origin 7.0) supplied with the calorimeter.

Analysis of substrate inhibition As rEngE is a glycoside hydrolase, we examined 3 potential models of substrate inhibition (Fig. 1). One was a general substrate inhibition model (Eq. 1) in which inhibition was achieved by forming an enzyme−substrate−substrate (ESS) complex, wherein 2 substrates were bound to an enzyme in addition to an enzyme−substrate (ES) complex (Fig. 1A) (16).

$$v = \frac{k_{cat}[E]_0[S]}{|S|^2/K_i + [S] + K_m} \quad [1]$$

The $k_{cat}[E]_0$ of Eq. 1 was defined as the maximum reaction rate, V_{max1}. The other substrate inhibition model involved products obtained from both ES and ESS complexes (Fig. 1B, C). Fig. 1B shows a random-ordered substrate inhibition model in which one substrate binds to the enzyme in a productive or a non-productive binding manner, following which another substrate binds in the opposite binding manner to form an ES_pS_n complex.

K_{m1}, K_{m2}, K_{i1}, and K_{i2} correspond to Eqs. 2, 3, 4 and 5, respectively.

$$K_{m1} = [E][S]/[ES_p] \quad (2)$$

$$K_{m2} = [ES_n][S]/[ES_pS_n] \quad (3)$$

$$K_{i1} = [ES_p][S]/[ES_pS_n] \quad (4)$$

$$K_{i2} = [E][S]/[ES_n] \quad (5)$$

Further, the total amount of the enzyme, $[E]_0$, is expressed as

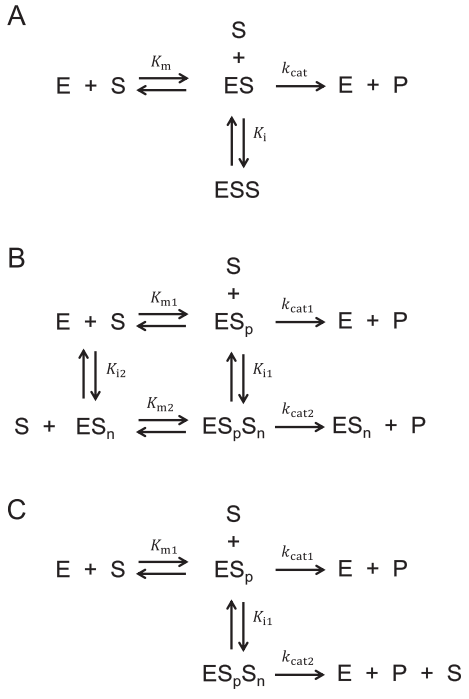

FIG. 1. Schematic of 3 kinetic models for substrate inhibition. (A) General substrate inhibition model. Substrate inhibition is caused by the dead-end ESS complex. (B) Random-ordered substrate inhibition model. One substrate binds to the enzyme in a productive or a non-productive manner, and another substrate binds in the opposite binding manner to form an ES_pS_n complex. (C) Sequential substrate inhibition model. One substrate binds to the enzyme in a productive binding manner, and another substrate binds in a non-productive binding manner to form an ES_pS_n complex. In panels B and C, the product is generated from the ES_p complex as well as the ES_pS_n complex.

$$[E]_0 = [E] + [ES_p] + [ES_n] + [ES_pS_n] \quad (6)$$

and the reaction rate v is the sum of $k_{cat1}[ES_p]$ and $k_{cat2}[ES_pS_n]$ as follows:

$$v = k_{cat1}[ES_p] + k_{cat2}[ES_pS_n] \quad (7)$$

From Eqs. 2–7, the reaction rate is expressed as follows:

$$v = \frac{k_{cat1}[E]_0[S] + k_{cat2}[E]_0[S]^2/K_{i1}}{[S]^2/K_{i1} + [S]K_{m1}/K_{i2} + [S] + K_{m1}} \quad (8)$$

$$K_{m1}K_{i1} = K_{m2}K_{i2} \quad (9)$$

The $k_{cat1}[E]_0$ and $k_{cat2}[E]_0$ of Eq. 8 was defined as the maximum reaction rate V_{max1} and V_{max2}, respectively.

The manner in which one substrate binds to the enzyme in a productive binding manner, and another substrate binds in a non-productive binding manner to form an ES_pS_n complex in a sequential substrate inhibition model is shown (Fig. 1C). It corresponds to a model sans K_{i2} and K_{m2} paths of the random-ordered substrate inhibition model.

K_{m1} and K_{i1} correspond to Eqs. 10 and 11, respectively.

$$K_{m1} = [E][S]/[ES_p] \quad (10)$$

$$K_{i1} = [ES_p][S]/[ES_pS_n] \quad (11)$$

Further, the total amount of the enzyme, $[E]_0$, is expressed as

$$[E]_0 = [E] + [ES_p] + [ES_pS_n] \quad (12)$$

and the reaction rate v is the sum of $k_{cat1}[ES_p]$ and $k_{cat2}[ES_pS_n]$ as follows:

$$v = k_{cat1}[ES_p] + k_{cat2}[ES_pS_n] \quad (13)$$

From Eq. 10–13, the reaction rate is expressed as follows:

$$v = \frac{k_{cat1}[E]_0[S] + k_{cat2}[E]_0[S]^2/K_{i1}}{[S]^2/K_{i1} + [S] + K_{m1}} \quad (14)$$

The $k_{cat1}[E]_0$ and $k_{cat2}[E]_0$ of Eq. 14 were defined as the maximum reaction rate V_{max1} and V_{max2}, respectively.

The obtained data were then applied to Eq. 1, Eq. 8 or Eq. 14 using Origin software (OriginLab Co., Northampton, MA, USA) to calculate the relevant kinetic and statistical parameters, as residual standard deviation ($Sy.x$), adjusted R-square (R_f^2) and Akaike's information criterion (AIC) for each substrate inhibition model. The substrate inhibition model which best explains the observed inhibitory activity of rEngE was subsequently verified.

Analysis of product inhibition The reaction rate was determined via a reaction mixture containing 50 nM rEngE, 7 mg/mL low-viscosity CMC, and 0 mM or 2.0 mM cellobiose at 25°C and a pH of 5.0 (in McIlvaine buffer containing 50 mM NaCl). The extent of CMC hydrolysis was determined by the same method described above, and inhibition of cellobiose was investigated by comparing the reaction rates in the presence or absence of cellobiose.

RESULTS AND DISCUSSION

Expression and purification of recombinant proteins The molecular architecture of each recombinant protein is presented (Fig. 2). GST-EngE, rEngE⊿CBM65-Doc and GST-CBM65 were successfully expressed by the *E. coli* expression system. The GST tag in GST-EngE was cleaved using protease via various chromatography techniques to ultimately produce purified rEngE. In order to investigate the influence of CBM65 and the dockerin domain on the catalyzed reaction, rEngE⊿CBM65-Doc, a mutant lacking these domains, was prepared. REngE⊿CBM65-Doc was purified via Ni affinity chromatography. GST-CBM65 was prepared to investigate its binding ability to cellulose. GST-CBM65 was purified without cleaving the GST tag in order to develop GST as an alternative structure to the catalytic domain of EngE. Each purified recombinant protein produced a major band upon SDS-PAGE analysis, with a molecular size which agreed well with the molecular weight deduced from the each nucleotide sequence (rEngE; molecular weight of 112,590, rEngE⊿CBM65-Doc; molecular weight of 88,468, and GST-CBM65; molecular weight of 51,498) (Fig. 3).

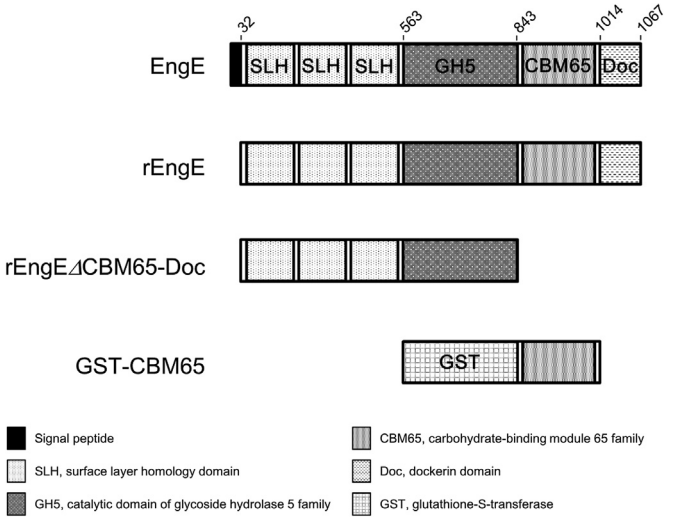

FIG. 2. Molecular architecture of recombinant proteins.

Identification of the specific substrate and product range of rEngE GH5 is one of the largest CAZy GH families, and its constituent GHs are known to hydrolyze various substrates. Thus, we aimed to identify specific substrate(s) as well as products of rEngE by TLC, through analyzing reaction products using several celluloses, with different degrees of crystallinity, and some hemicelluloses as substrates. Specifically, each reaction solution contained rEngE and CMC, PASC, microcrystalline cellulose, lichenan, glucomannan, β-glucan, birchwood xylan, or beechwood xylan as test substrates, which were incubated for 24 h and applied to TLC plates (Fig. 4). With respect to celluloses, the main products detected were di- and tri-saccharides as soluble cellulose for CMC and amorphous cellulose for PASC, whereas no microcrystalline cellulose products were detected (Fig. 4A). By contrast, hydrolysis of glucomannan mainly produced trisaccharides and longer oligosaccharides, and hydrolysis of lichenan and β-glucan mainly produced di- and tri-oligosaccharides and longer oligosaccharides (Fig. 4B). However, no products were detected when beechwood or birchwood xylan was used as the substrate (Fig. 4C). Relative activities of beechwood xylan and birchwood xylan on CMC

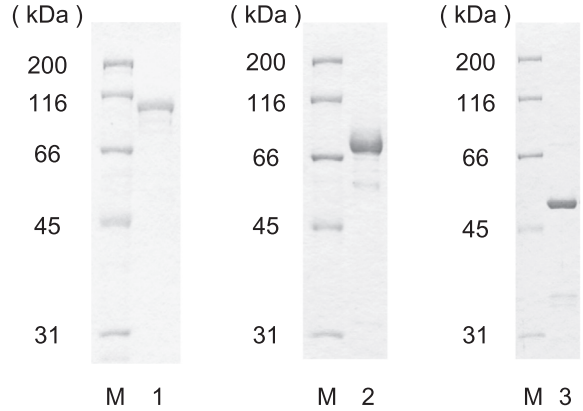

FIG. 3. SDS-PAGE of purified recombinant proteins. Gels were stained with Coomassie brilliant blue and the samples were loaded in the following order: M, protein marker (molecular masses shown on the left); 1, purified rEngE; 2, purified rEngE⊿CBM65-Doc; 3, purified GST-CBM65.

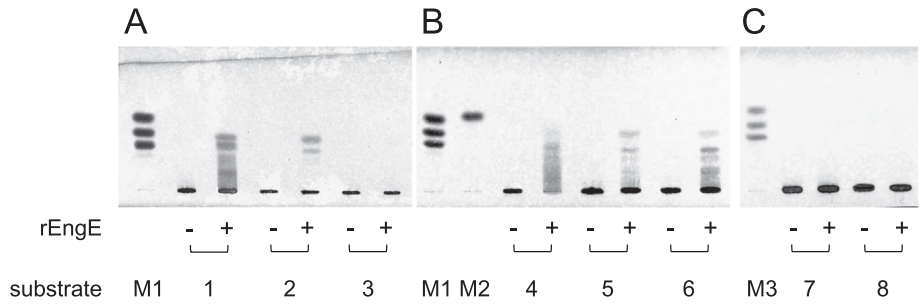

FIG. 4. TLC analysis of the degradation products of EngE. Low-viscosity CMC, PASC, microcrystalline cellulose, glucomannan, lichenan, β-glucan, beechwood xylan, and birchwood xylan were incubated with (+) or without (−) rEngE at 25°C for 24 h. (A) M1, cellooligosaccharides (glucose, cellobiose, cellotriose); 1, low-viscosity CMC; 2, PASC; 3, microcrystalline cellulose. (B) M2, mannose; 4, glucomannan; 5, lichenan; 6, β-glucan. (C) M3, xylooligosaccharides (xylose, xylobiose, xylotriose); 7, beechwood xylan; 8, birchwood xylan.

activity were 1.1% and 0.25%, respectively, whereas EngE only showed slight xylanase activity. These results demonstrated that EngE may preferably hydrolyze celluloses with low crystallinity. Since the main chain of hydrolyzed substrates is composed of glucose residues, EngE was revealed to be an endo-β-1,4-glucanase that hydrolyzes β-1,4 glycosidic bonds between glucose residues. Although several carbohydrate-related enzymes of *C. cellulovorans* have been reported to date, small oligosaccharides such as di- and tri-saccharides have been rarely detected as main products (17,18). *C. cellulovorans* can efficiently degrade cellulosic biomass by constructing cellulosomes on the cell surface (3–5). Since EngE retains the dockerin domain and is expressed with a scaffolding protein, CbpA, it is considered to act as a constituent of the cellulosome (10). Alternatively, even when it is not assembled as a cellulosome, EngE would still be able to bind to the cell surface since it has 3 SLH domains similar to CbpA (19). Therefore, although EngE is a secreted enzyme, it is expected to localize on the cell surface. *C. cellulovorans* incorporates and metabolizes small sugars such as cellobiose (20). Since the main products of EngE are di- and tri-saccharides, EngE might contribute to the efficient production of small sugars in the process of degrading cellulosic biomass. *C. cellulovorans* would rapidly incorporate these sugars by localizing EngE on the cell surface.

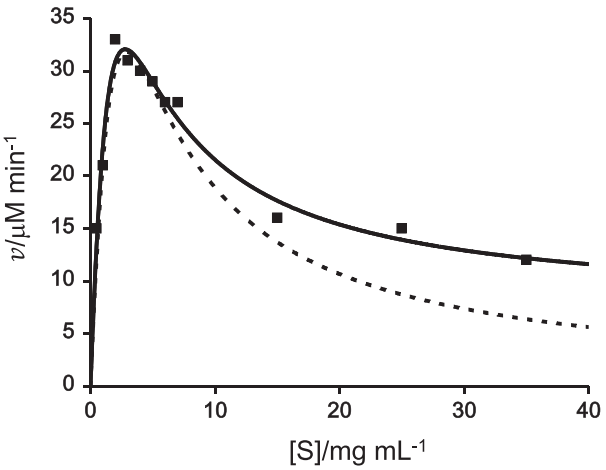

FIG. 5. Effect of substrate concentration on the reaction rate of the enzymatic hydrolysis of CMC. Reaction rates (v) measured at each substrate concentration [S], were plotted against the substrate concentration. Solid line shows the data fitted to the random-ordered model of substrate inhibition (Eq. 8) and sequential model of substrate inhibition (Eq. 14), respectively. The dashed line shows the data fitted to the general substrate inhibition model (Eq. 1).

Enzyme kinetics Steady-state kinetics provides a simple and rapid means for assessing enzyme activity. We performed a detailed analysis of the steady-state kinetics of rEngE using CMC, which was confirmed as a rEngE substrate. The reaction rate decreased gradually at substrate concentrations of 3 mg/mL and above (Fig. 5). This decrease might be attributed to product or substrate inhibition. The reaction rate was determined with a reaction mixture containing 7 mg/mL CMC, and 0 mM or 2.0 mM cellobiose, respectively. A sufficient decrease in the reaction rate was observed at this substrate concentration, where the cellobiose concentration corresponded to the amount of disaccharide produced when 10% of added CMC was degraded. The reaction rate was 30 μM/min in the presence of cellobiose, whereas it was 27 μM/min in the absence of cellobiose. Thus, the observed reaction rate was almost similar, thereby eliminating product inhibition as a cause. In addition, it is conceivable that reaction rate may decrease due to the viscosity of CMC. However, some reports have indicated that even when CMC concentrations over 10 mg/mL are used as a substrate, the reaction rate corresponds to the Michaelis–Menten type (21–23). In the case of EngE, the decrease in reaction rate after CMC concentration exceeded 3 mg/mL, was considered to be due to inhibition by the substrate. To examine the possibility of substrate inhibition, the data in the direct plot (Fig. 5) were analyzed using general substrate inhibition (Eq. 1), random-ordered model of substrate inhibition (Eq. 8), and sequential model of substrate inhibition (Eq. 14) to evaluate the 3 substrate inhibition models (Fig. 1). The curves generated from random-ordered and sequential models of substrate inhibition were almost similar and were highly consistent with the data. However, the curve obtained from general substrate inhibition was not consistent with the data. The values of the kinetics and statistical parameters are shown (Tables 1 and 2). The kinetic parameters of random-ordered and sequential substrate inhibition models were almost similar for all

TABLE 1. Kinetic parameters of substrate inhibition by rEngE.

Model	K_{m1}/mg mL^{-1}	K_{i1}/mg mL^{-1}	V_{max1}/μM min^{-1}	K_{m2}/mg mL^{-1}	K_{i2}/mg mL^{-1}	V_{max2}/μM min^{-1}
General substrate inhibition model	1.4 ± 0.4	7.4 ± 2.1	58 ± 8	NA	NA	NA
Sequential substrate inhibition model	3.7 ± 2.7	1.6 ± 1.4	117 ± 68	NA	NA	7.4 ± 2.1
Random-ordered substrate inhibition model	3.7 ± 3.0	1.6 ± 1.5	117 ± 74	2.3×10^{-25}	2.6×10^{25}	7.4 ± 2.3

Standard error of K_{m2} and K_{i2} in the random-ordered substrate inhibition model could not be calculated. NA, not applicable.

TABLE 2. Statistical parameters of kinetic analyses for the three models of substrate inhibition by rEngE.

Model	*Sy.x*	R_f^2	AIC
General substrate inhibition model	3.6	0.78	39
Sequential substrate inhibition model	1.6	0.96	27
Random-order substrate inhibition model	1.7	0.95	38

Sy.x, residual standard deviation; R_f^2, adjusted R-squared; AIC, Akaike's information criterion.

parameters except K_{m2} and K_{i2} of the random-ordered substrate inhibition model, where K_{m2} and K_{i2} are extremely numerical values (Table 1). These results indicated that the random-ordered substrate inhibition model has not the paths of K_{m2} and K_{i2}. Therefore, it was strongly suggested that EngE adheres to a sequential substrate inhibition model. In order to evaluate this result from another viewpoint, *Sy.x*, a calculated standard deviation value, R_f^2, a coefficient of determination, and AIC for were used for evaluating the optimal variables of the statistical model. The parameters of sequential substrate inhibition model yielded good values for *Sy.x*, R_f^2, and AIC. The values of *Sy.x* and R_f^2 in the random-ordered substrate inhibition model were nearly equivalent to those of the sequential substrate inhibition model, but the AIC values were significantly different (Table 2). The AIC value of the random-ordered substrate inhibition model was higher than that of the sequential substrate inhibition model. As the model with the smallest AIC value is considered to be optimal, the random-ordered substrate inhibition model did not qualify as the optimal model. This indicated that the optimal substrate inhibition model for rEngE was the sequential substrate inhibition model (Fig. 1C), and that rEngE generate products from the ESS complex.

Other GHs have been shown to exhibit substrate inhibition, including GH5 xyloglucanase (XG5 from *Paenibacillus pabuli*), GH12 xyloglucanase (XG12 from *Bacillus licheniformis*), GH5 β-mannanase (ManF from *Bacillus stearothermophilus*), and GH27 α-galactosidase (AglB from *Aspergillus niger*) (24–27). Substrate inhibition has also been shown to occur in Cellulysin (Merck), a crude extract of cellulase from *Trichoderma viride* (28). However, reports of substrate inhibition are relatively rare, overall. In particular, ours is the first study to show that ESS complexes yield products. A comparison of kinetic parameters for CMC among GH5 enzymes, including EngE is shown (Table 3) (29,30). Although the reaction temperature was different, rEngE showed the highest k_{cat}/K_m values among GH5 enzymes, which were more than twice that of EngD and more than 7 times that of FmEG (Table 3). Thus, EngE appears to have the greatest catalytic efficiency among the known GH5 enzymes.

Binding of CBM65 to cellulose and its effect on the catalytic reaction Substrate inhibition of rEngE was found to adhere to the sequential substrate inhibition model through kinetic analysis, as it was revealed that rEngE forms an ESS complex by binding productively with the substrate and then binding non-productively with another substrate. These substrate binding sites were assumed to be the active site and the substrate binding site of CBM65. As the substrate productively binds to the active site of the catalytic domain, the binding affinity of cellulose as substrate was analyzed using CBM65 fused to a protein with an equivalent molecular weight instead of the catalytic domain of EngE. Proteins bound to crystalline celluloses were detected by SDS-PAGE. BSA was not detected in the elution fraction, but GST-CBM65 was slightly detected in the eluted fraction (Fig. 6A). Therefore, it was suggested that the CBM65 contained in EngE binds weakly to microcrystalline cellulose. Binding to soluble cellulose was observed using VP-ITC. The isothermal for the binding of GST-CBM65 to CMC is shown (Fig. 6B). It revealed that GST-CBM65 binds specifically to soluble cellulose such as CMC. Since the average molecular weight of CMC is approximately 90,000, the thermodynamic parameters were; $n = 0.023$, $K_d = 1.2 \times 10^{-6}$ M, and $\Delta H = -1.9 \times 10^5$ J/mol. These results indicated that CBM65 binds to cellulose regardless of crystallinity. Reportedly, the CBM65 contained in Ra2535 from *Rumicoccus albus* 8 binds to crystalline cellulose and CMC (31). Also, Cel5A from *Eubacterium cellulosolvens* has two CBM65s, which bind to CMC and acid-swollen cellulose, but cannot bind to crystalline cellulose (15,32). Therefore, the CBM65 of EngE seems to have similar binding properties to the CBM65 of Ra2535. Further, the value of K_d for CMC of GST-CBM65 was 1.0–1.4 μM, and the value of K_{i1} of the sequential substrate inhibition model calculated from the kinetic analysis was 2.2–33 μM in terms of molar concentration. It was found that both values obtained from different analyses were very close, and that another substrate binding site could likely be a substrate binding site of CBM65. Thus, it may be surmised that substrate inhibition occurs because the binding site of CBM65 corresponds to one of the substrate-binding sites in the ESS complex. On the other hand, since it was assumed that CBM65 was involved in the catalytic reaction by binding with CMC, rEngEΔCBM65-Doc lacking a CBM65 and a dockerin domain was prepared and its enzyme activity examined. Purified rEngEΔCBM65-Doc showed little activity. It was speculated that CBM65 plays a critical role in the structural stability of the catalytic domain.

Among the GH5 enzymes, EngE showed the greatest efficiency for catalyzing CMC, even though it caused substrate inhibition at higher substrate concentrations. Also, since its main products are disaccharides and trisaccharides, it may contribute to the efficient production of small sugars during the process of degrading cellulosic biomass. It binds to cellulose via its CBM65, and *C. cellulovorans* efficiently and rapidly takes up these sugars by localizing it on the cell surface (19). By mimicking such a system in *vivo*, by expressing EngE outside the cells of microorganisms which can take up small sugars, such as butanol-producing bacteria, it may localize to the cell surface layer like a cellulosome. Furthermore, the enzyme can be recycled while microbial cells grow, to directly produce alcohol from cellulose or cellulosic biomass.

TABLE 3. Comparison of kinetic parameters for CMC among GH5 enzymes.

Enzyme	Strain	Model	K_{m1}/mg mL^{-1}	K_{i1}/mg mL^{-1}	k_{cat1}/s^{-1}	k_{cat2}/s^{-1}	$k_{cat1}\ K_{m1}^{-1}$/mL mg^{-1} s^{-1}
EngE	*C. cellulovorans*	Sequential substrate inhibition model	3.7	1.6	39	2.5	11
EngD (29)	*C. cellulovorans*	Michaelis–Menten model	6.5	NA	30	NA	4.6
FmEG (30)	*F. mediterranea* MF3/22	Michaelis–Menten model	3.0	NA	4.5	NA	1.5

Kinetic parameters for EngE, EngD, and FmEG were determined at 25°C, 37°C, and 50°C, respectively. NA, not applicable.

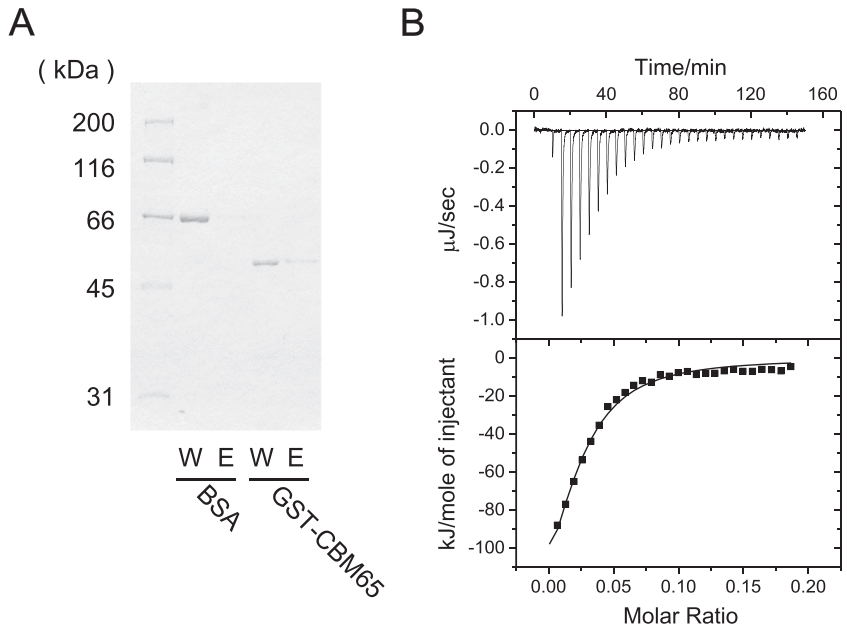

FIG. 6. Cellulose binding assay of CBM65. (A) BSA and GST-CBM65 bound and/or unbound to microcrystalline cellulose were detected by SDS-PAGE. BSA was used as the negative control. W and E indicate the washed fraction and the eluted fraction, respectively. (B) Isothermal titration calorimetry of the binding of GST-CBM65 to CMC at pH 7.5 and 25°C. The thermogram (top) and binding isotherm (bottom) are shown. In total, 10 μL of GST-CBM65 solution (50 μM) was injected 28 times into the low-viscosity CMC solution (56 μM). In the binding isotherm, solid squares show experimental values and the solid line shows a theoretical curve fitted to the one sets of sites model.

ACKNOWLEDGMENTS

The authors are very grateful to Professor Akiyoshi Tanaka of Mie University for stimulating discussions about substrate inhibition models. This work was funded by the Program for Promotion of Basic and Applied Researches for Innovations in Bio-oriented Industry, the National Agriculture and Food Research Organization of Japan. A part of this research was funded by the Science and Technology Research Promotion Program for Agriculture, Forestry, Fisheries, and Food Industry, the Ministry of Agriculture, Forestry and Fisheries of Japan.

References

1. **Rodrigues Mota, T., Matias de Oliveira, D., Marchiosi, R., Ferrarese-Filho, O., and Dantas dos Santos, W.:** Plant cell wall composition and enzymatic deconstruction, AIMS Bioeng., **5**, 63–77 (2018).
2. **Aspeborg, H., Coutinho, P. M., Wang, Y., Brumer, H., and Henrissat, B.:** Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5), BMC Evol. Biol., **12**, 186–202 (2012).
3. **Bayer, E. A., Shimon, L. J., Shoham, Y., and Lamed, R.:** Cellulosomes-structure and ultrastructure, J. Struct. Biol., **124**, 221–234 (1998).
4. **Matano, Y., Park, J. S., Goldstein, M. A., and Doi, R. H.:** Cellulose promotes extracellular assembly of *Clostridium cellulovorans* cellulosomes, J. Bacteriol., **176**, 6952–6956 (1994).
5. **Doi, R. H. and Tamaru, Y.:** The *Clostridium cellulovorans* cellulosome: an enzyme complex with plant cell wall degrading activity, Chem. Rec., **1**, 24–32 (2001).
6. **Tamaru, Y., Miyake, H., Kuroda, K., Nakanishi, A., Kawade, Y., Yamamoto, K., Uemura, M., Fujita, Y., Doi, R. H., and Ueda, M.:** Genome sequence of the cellulosome-producing mesophilic organism *Clostridium cellulovorans* 743B, J. Bacteriol., **192**, 901–902 (2010).
7. **Tamaru, Y., Miyake, H., Kuroda, K., Nakanishi, A., Matsushima, C., Doi, R. H., and Ueda, M.:** Comparison of the mesophilic cellulosome-producing *Clostridium cellulovorans* genome with other cellulosome-related clostridial genomes, Microb. Biotechnol., **4**, 64–73 (2011).
8. **Tamaru, Y. and Doi, R. H.:** The *engL* gene cluster of *Clostridium cellulovorans* contains a gene for cellulosomal ManA, J. Bacteriol., **182**, 244–247 (2000).
9. **Foong, F., Hamamoto, T., Shoseyov, O., and Doi, R. H.:** Nucleotide sequence and characteristics of endoglucanase gene *engB* from *Clostridium cellulovorans*, J. Gen. Microbiol., **137**, 1729–1736 (1991).
10. **Morisaka, H., Matsui, K., Tatsukami, Y., Kuroda, K., Miyake, H., Tamaru, Y., and Ueda, M.:** Profile of native cellulosomal proteins of *Clostridium cellulovorans* adapted to various carbon sources, AMB Express, **2**, 37 (2012).
11. **Tamaru, Y. and Doi, R. H.:** Three surface layer homology domains at the N terminus of the *Clostridium cellulovorans* major cellulosomal subunit EngE, J. Bacteriol., **181**, 3270–3276 (1999).
12. **von Heijne, G.:** Signal sequences. The limits of variation, J. Mol. Biol., **184**, 99–105 (1985).
13. **Zhang, Y. H., Cui, J., Lynd, L. R., and Kuang, L. R.:** A transition from cellulose swelling to cellulose dissolution by *o*-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure, Biomacromolecules, **7**, 644–648 (2006).
14. **Hiromi, K., Takasaki, Y., and Ono, S.:** Kinetics of hydrolytic reaction catalyzed by crystalline bacterial α-amylase. III. The influence of temperature, Bull. Chem. Soc. Jpn., **36**, 563–569 (1963).
15. **Yoda, K., Toyoda, A., Mukoyama, Y., Nakamura, Y., and Minato, H.:** Cloning, sequencing, and expression of a *Eubacterium cellulosolvens* 5 gene encoding an endoglucanase (Cel5A) with novel carbohydrate-binding modules, and properties of Cel5A, Appl. Environ. Microbiol., **71**, 5787–5793 (2005).
16. **Armstrong, E. F.:** Enzymes. By J.B.S. Haldane, M.A, pp. 919–920, in: Plimmer, R. H. A. and Hopkins, F. G. (Eds.), Monographs on biochemistry, vol. vii+235. Longmans, Green & Co., London (1930).
17. **Han, S. O., Yukawa, H., Inui, M., and Doi, R. H.:** Isolation and expression of the *xynB* gene and its product, XynB, a consistent component of the *Clostridium cellulovorans* cellulosome, J. Bacteriol., **186**, 8347–8355 (2004).
18. **Han, S. O., Yukawa, H., Inui, M., and Doi, R. H.:** Molecular cloning and transcriptional and expression analysis of *engO*, encoding a new noncellulosomal family 9 enzyme, from *Clostridium cellulovorans*, J. Bacteriol., **187**, 4884–4889 (2005).
19. **Kosugi, A., Murashima, K., Tamaru, Y., and Doi, R. H.:** Cell-surface-anchoring role of N-terminal surface layer homology domains of *Clostridium cellulovorans* EngE, J. Bacteriol., **184**, 884–888 (2002).
20. **Sleat, R., Mah, R. A., and Robinson, R.:** Isolation and characterization of an anaerobic, cellulolytic bacterium, *Clostridium cellulovorans* sp. nov. Appl. Environ. Microbiol., **48**, 88–93 (1984).
21. **Balasubramanian, N. and Simoes, N.:** *Bacillus pumilus* S124A carboxymethyl cellulase; a thermo stable enzyme with a wide substrate spectrum utility, Int. J. Biol. Macromol., **67**, 132–139 (2014).
22. **Zhang, F., Chen, J. J., Ren, W. Z., Nie, G. X., Ming, H., Tang, S. K., and Li, W. J.:** Cloning, expression and characterization of an alkaline thermostable GH9 endoglucanase from *Thermobifida halotolerans* YIM 90462 T, Bioresour. Technol., **102**, 10143–10146 (2011).
23. **Blanchette, C., Lacayo, C. I., Fischer, N. O., Hwang, M., and Thelen, M. P.:** Enhanced cellulose degradation using cellulase-nanosphere complexes, PLoS One, **7**, e42116 (2012).

24. **Gloster, T. M., Ibatullin, F. M., Macauley, K., Eklof, J. M., Roberts, S., Turkenburg, J. P., Bjornvad, M. E., Jorgensen, P. L., Danielsen, S., Johansen, K. S., and other 4 authors:** Characterization and three-dimensional structures of two distinct bacterial xyloglucanases from families GH5 and GH12, J. Biol. Chem., **282**, 19177–19189 (2007).
25. **Talbot, G. and Sygusch, J.:** Purification and characterization of thermostable β-mannanase and α-galactosidase from *Bacillus stearothermophilus*, Appl. Environ. Microbiol., **56**, 3505–3510 (1990).
26. **Ethier, N., Talbot, G., and Sygusch, J.:** Gene cloning, DNA sequencing, and expression of thermostable β-mannanase from *Bacillus stearothermophilus*, Appl. Environ. Microbiol., **64**, 4428–4432 (1998).
27. **Liao, J., Okuyama, M., Ishihara, K., Yamori, Y., Iki, S., Tagami, T., Mori, H., Chiba, S., and Kimura, A.:** Kinetic properties and substrate inhibition of α-galactosidase from *Aspergillus niger*, Biosci. Biotechnol. Biochem., **80**, 1747–1752 (2016).
28. **Liaw, E. T. and Penner, M. H.:** Substrate-velocity relationships for the *Trichoderma viride* cellulase-catalyzed hydrolysis of cellulose, Appl. Environ. Microbiol., **56**, 2311–2318 (1990).
29. **Lee, C. Y., Yu, K. O., Kim, S. W., and Han, S. O.:** Enhancement of the thermostability and activity of mesophilic *Clostridium cellulovorans* EngD by in vitro DNA recombination with *Clostridium thermocellum* CelE, J. Biosci. Bioeng., **109**, 331–336 (2010).
30. **Pan, R., Hu, Y., Long, L., Wang, J., and Ding, S.:** Extra carbohydrate binding module contributes to the processivity and catalytic activity of a non-modular hydrolase family 5 endoglucanase from *Fomitiporia mediterranea* MF3/22, Enzyme Microb. Technol., **91**, 42–51 (2016).
31. **Iakiviak, M., Devendran, S., Skorupski, A., Moon, Y. H., Mackie, R. I., and Cann, I.:** Functional and modular analyses of diverse endoglucanases from *Ruminococcus albus* 8, a specialist plant cell wall degrading bacterium, Sci. Rep., **6**, 29979 (2016).
32. **Luis, A. S., Venditto, I., Temple, M. J., Rogowski, A., Basle, A., Xue, J., Knox, J. P., Prates, J. A., Ferreira, L. M., Fontes, C. M., Najmudin, S., and Gilbert, H. J.:** Understanding how noncatalytic carbohydrate binding modules can display specificity for xyloglucan, J. Biol. Chem., **288**, 4799–4809 (2013).