RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Fusion enzyme consisting of bacterial expansin and endoglucanase for degradation of highly crystalline cellulose

Kazunori Nakashima, Koji Endo, Naomi Shibasaki-Kitakawa and Toshikuni Yonemoto*

Department of Chemical Engineering, Tohoku University, Aoba-yama 6-6-07, Aoba-ku, Sendai 980-8579, Japan

*Author for correspondence
Toshikuni Yonemoto
Tel/Fax: +81-22-795-7255
E-mail: toshiy@rpel.che.tohoku.ac.jp

Abstract

We have prepared novel fusion enzymes consisting of *Bacillus subtilis* expansin EXLX1 and *Clostridium thermocellum* endoglucanase CelD. These two components are directly fused, or fused using flexible glycine-serine peptide linkers (GGGGS, GS linker) with different lengths: a triplicate $(GGGGS)_3$ linker (GS3) and a sextuple $(GGGGS)_6$ linker (GS6), resulting in fusion enzymes EXLX1-CelD, EXLX1-GS3-CelD and EXLX1-GS6-CelD. The binding ability and digestibility of these fusion enzymes towards a series of cellulose substrates with different crystallinity index (CrI) was examined. Fused with EXLX1, CelD acquired high binding ability to various kinds of cellulose and the linker length had no impact on function. In the degradation of cellulose, EXLX1-GS3-CelD exhibited the highest degradation activity among the fusion enzymes examined, suggesting that linker length between the two proteins has a significant impact on the activity of the fusion enzyme. EXLX1-GS3-CelD was found to function more effectively towards higher crystalline celluloses.

Keywords: crystallinity index, expansin, endoglucanase, fusion protein, pretreatment, saccharification

Introduction

Biorefineries are promising processes for the production of biofuels and bio-based chemicals from renewable natural resources that can replace our dependency on limited fossil fuel resources.[1–3] Lignocellulosic biomass, composed of cellulose, hemicellulose and lignin, is an abundant renewable resource that has significant potential for use in biorefinery processes.[4] Biological degradation of cellulose using the cellulolytic enzyme cellulase is the most feasible strategy for the production of fermentable glucose from cellulose.[5] However, a challenging problem in cellulose degradation is the low digestibility of cellulose by cellulase. As cellulose is a highly crystalline polymer, where polysaccharide chains are tightly packed with each other by multivalent inter- and intra-molecular hydrogen bonding,[6] crystalline cellulose is recalcitrant to biodegradation by cellulase. Therefore, pretreatment of cellulose is necessary to enhance its digestibility by disrupting the highly ordered and rigid structure of cellulose.[7,8] The pretreatment process often involves chemicals such as acids and/or alkaline that should be removed before enzymatic hydrolysis. Moreover, high-energy consumption is also required because of treatment at high temperature and pressure. Thus, the pretreatment process is complicated and costly. If cellulases acquire the ability to disrupt the high crystallinity of cellulose, the pretreatment and the saccharification process can be simultaneously conducted in a single vessel, where neither chemical reagent nor high energy would be required. This system represents a potential novel and cost-effective unit process that leads to the construction of innovative biorefinery processes.

Expansin is a plant protein capable of loosening the plant cell wall and disrupting the cellulose crystal.[9,10] Not only plants but also filamentous fungi and yeasts such as *Trichoderma reesei*,[11,12] *Aspergillus fumigatus*[13] and *Kluyveromyces lactis*[14] are known to possess an expansin-like protein, named swollenin. Swollenin exhibits crystal-disrupting

activity for cellulosic materials. More recently, expansin-like proteins have been also found in bacteria, such as *Bacillus subtilis* and this protein is named EXLX1.[15–17] These expansin families were shown to accelerate cellulose degradation when added to the reaction mixture.[13,14,16]

In the present study, we have prepared fusion enzymes consisting of expansin and endoglucanase (cellulase). *Bacillus subtilis* expansin EXLX1 and *Clostridium thermocellum* endoglucanase CelD were connected by peptide linkers that varied in length. By connecting the two components, the synergistic effect on cellulose degradation was expected to be significantly enhanced, because the fusion enzymes could potentially disrupt the crystalline structure of cellulose by EXLX1 and simultaneously be degraded to fermentable sugars by CelD. The binding ability and digestibility of the fusion enzymes to cellulose were evaluated using a series of cellulosic substrates with different crystallinity.

Experimental

Preparation of proteins and fusion enzymes

Genomic DNA of *C. thermocellum* NBRC 103400 and *B. subtilis* subsp. NBRC 13719 were obtained from the NBRC, NITE (Kisarazu, Japan). *Escherichia coli* DH5α strain and *E. coli* BL21(DE3) strain (NIPPON GENE, Toyama, Japan) were used as a host strain for plasmid construction and protein expression, respectively. PCR was performed using PrimeSTAR HS DNA Polymerase (Takara Bio Inc., Otsu, Japan) and the oligonucleotide primers (Operon Biotechnologies, Tokyo, Japan) listed in Table 1. Expression vector pET-22b(+), In-Fusion HD Cloning Kit (Takara Bio Inc.) and restriction enzymes (New England Biolabs Japan, Tokyo, Japan) were used for plasmid construction. Other chemicals used in this study were of reagent grade.

We have constructed three different types of fusion enzymes consisting of EXLX1 and CelD connected with a flexible glycine-serine peptide linker (GGGGS, GS linker) with different lengths, that is, (i) no linker, (ii) the triplicate $(GGGGS)_3$ (GS3 linker), and (iii) the sextuple $(GGGGS)_6$ (GS6 linker). The schematic representations of the three different types of fusion enzymes, which are directly-fused (EXLX1-CelD), connected with GS3 linker (EXLX1-GS3-CelD), and connected with GS6 linker (EXLX1-GS6-CelD), are illustrated in Fig. 1.

The gene encoding EXLX1 (deposited as YoaJ protein, GenBank: AAB84448.1) was amplified by PCR from *B. subtilis* genomic DNA with the primer set, p1-F/p2-R. These primers were designed to remove a deduced signal sequence at the *N*-terminus of the protein.[16] The amplified EXLX1 fragment was inserted into the expression vector pET-22b(+) digested by *Nde*I and *Xho*I using the In-Fusion cloning method to obtain the EXLX1 expression vector. The CelD gene (GenBank: X04584.1) of *C. thermocellum* originally contains a non-catalytic "dockerin domain" at the *C*-terminus, which is known to cause the formation of insoluble inclusion bodies.[18] Therefore, a truncated CelD gene without the *C*-terminal dockerin domain and the *N*-terminal signal sequence was cloned from *C. thermocellum* genomic DNA using the primer set, p3-F/p4-R. The PCR product was inserted into pET-22b(+) to obtain the CelD expression vector. To construct the EXLX1-CelD expression vector, the CelD gene was amplified with the primer set p5-F/p4-R and the PCR product inserted into the *Xho*I site of the EXLX1 expression vector to give the EXLX1-CelD expression vector. In a similar manner, the EXLX1-GS3-CelD expression vector was constructed using the primer p6-F as the forward primer rather than the p5-F primer. A EXLX1 fragment amplified with the primer set p1-F/p7-R and a CelD fragment amplified with the primer set p8-F/p4-R were simultaneously inserted into pET-22b(+) digested with

*Nde*I/*Xho*I, providing the EXLX1-GS3-CelD expression vector possessing an *EcoR*I site between EXLX1 and the GS linker region. This vector was digested with *Nde*I/*EcoR*I and then ligated with the EXLX1 fragment amplified with the primer set p1-F/p9-R, resulting in the EXLX1-GS6-CelD expression vector.

The five types of proteins and enzymes shown in Fig. 1 were overexpressed in *E. coli* BL21(DE3) as the host strain. *E. coli* transformed with the corresponding expression vector were grown at 37 °C in Luria–Bertani (LB) medium with 100 mg/L of ampicillin. When the absorbance of the media at 600 nm reached 0.4, the expression of the protein was induced by adding to a final concentration 1 mmol/L of isopropyl *β*-D-1-thiogalactopyranoside (IPTG) to the medium and then the cells were allowed to continue growing at 37 °C for 12 h. *E. coli* cells expressing the proteins were collected by centrifugation (3,000 rpm, 10 min) and lysed by sonication (Branson sonifier 250, Branson Ultrasonics, Danbury, CT, USA). After centrifugation of the lysed solution at 12,000 rpm for 10 min, the supernatant was recovered and concentrated by ultrafiltration (MWCO: 10,000 Da). The concentrated protein solution was diluted with binding buffer (20 mmol/L Tris-HCl, 500 mmol/L NaCl, 10 mmol/L imidazole, pH 8.0), followed by His-tag purification by a Ni-nitrilotriacetic acid affinity column (Bio-Scale Mini Profinity eXact cartridge, Bio-Rad Laboratories. Inc., Hercules, FL, USA). The expressed proteins were analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE).

Preparation of cellulose with different crystallinities

Cellulose shows extremely high crystallinity when compared with other natural polysaccharides such as amylose and hemicellulose. This is attributed to the multivalent hydrogen bond network formed in cellulose chains, resulting in the low digestibility of

cellulose by cellulase. Phosphoric acid is well known to dissolve cellulose and disrupt the crystalline structure, leading to the acceleration of enzymatic degradation.[19,20] To make a detailed investigation of the cellulose-binding and crystal-disrupting function of EXLX1 and fusion enzymes, cellulose with different crystallinities were prepared by using different concentrations of phosphoric acid (76.9 wt%, 77.7 wt% and 82.1 wt%), according to a previous method.[20]

Crystalline cellulose Avicel (PH-101) was used as a starting material for the preparation of cellulose with different crystallinities. Avicel is basically composed of only cellulose with crystalline structure, differing from native biomass composed of cellulose, xylose, and lignin, with complicated heterostructure. For these reasons, Avicel is commonly utilized in the investigation of cellulase function,[5,13-15] and was also employed as a substrate to investigate the properties of fusion enzymes in the present study. Phosphoric acid at specified concentrations (53 ml) was slowly added to Avicel (0.1 g) with constant stirring at 120 rpm and 4 °C for 40 min. Two hundred mL of water was added to the solution with vigorous stirring, resulting in a white precipitate of the regenerated cellulose. The cellulose was collected by centrifugation at 5,000 rpm for 15 min, and then washed with deionized water until the pH of the solution was ~5. The suspension of cellulose was stored at 4 °C before use.

The crystallinity index (CrI) of the prepared cellulose was examined by X-ray powder diffraction (XRD).[21] Cellulose samples after lyophilization were scanned by an X-ray diffractometer (RINT-2200VHF+/PC, Rigaku, Tokyo, Japan) from 2θ = 5–30° with a scan speed of 1°/min and a step size of 0.02°. The CrI was calculated by eqn. (1).

$$\mathrm{CrI} = \frac{I_{002} - I_{\mathrm{AM}}}{I_{002}} \times 100 \quad (1)$$

Here, I_{002} is the intensity of the peak at 2θ = 22.5° and I_{am} is the intensity of the peak at 2θ =

158 18°.

159

160 **Assay of the binding ability of the protein to cellulose**

161 The binding abilities of EXLX1, CelD and the fusion enzymes to various kinds of cellulose
162 with different crystallinities were studied. A five mL cellulose suspension (5 g/L) with
163 different CrI in acetate buffer (50 mmol/L, pH5) was pre-heated at 40 °C for 30 min. The
164 proteins were added to the suspension at a concentration of 1.0 μmol/L, followed by gently
165 shaking at 100 rpm and 40 °C. The experimental time for the binding ability was relatively
166 short (~2 h), such that the amount of cellulose degradation was not significant. Samples (600
167 μL) were collected at scheduled times and centrifuged to separate the precipitant and the
168 supernatant. The concentration of free protein unbound to cellulose in the supernatant was
169 measured by the Bradford assay (Bio-Rad Protein Assay #500-0006, Bio-Rad Laboratories.
170 Inc.), and the ratio of bound proteins to cellulose was determined by equation. (2).

171 $$\text{Ratio of bound protein } [\%] = \left(1 - \frac{\text{protein in supernatant [mol]}}{\text{initial protein [mol]}}\right) \times 100 \qquad (2)$$

172

173 **Assay of the digestibility of enzymes towards the cellulose substrate**

174 The digestibility of enzymes towards the cellulose substrate with various CrI values was
175 examined under long incubation periods, using the same conditions as described in the
176 binding assay. During degradation, samples (300 μL) were collected at scheduled times and
177 centrifuged to separate the supernatant containing the generated sugars from the unreacted
178 solid cellulose. The concentrations of glucose (G1) and cellobiose (G2) in the supernatant
179 were measured by HPLC (ACQUITY UPLC H-Class, Waters, Milford, MA, USA) equipped
180 with a BEH amide column (130 Å, 1.7 μm, 2.1 × 50 mm) and an evaporative light scattering
181 (ELS) detector, where both acetonitrile and ultrapure water containing 0.2 vol% triethylamine

were used as the mobile phase. Digestibility of the cellulose was calculated by eqn. (3).

$$\text{Digestibility [\%]} = \left(\frac{\text{produced G1 [mol]} + \text{produced G2 [mol]} \times 2}{\text{initial G1 residue in cellulose [mol]}} \right) \times 100 \qquad (3)$$

Results and discussion

Verification of proteins and fusion enzymes

The nucleotide sequences of target genes (EXLX1, CelD, EXLX1-CelD, EXLX1-GS3-CelD, and EXLX1-GS6-CelD) in each expression vector were verified by DNA sequencing. Figure 2 shows the result of SDS-PAGE analysis for each protein that is expressed in *E. coli*, followed by His-tag purification. EXLX1 (23.1 kDa) was successfully expressed as reported previously.[16] Similarly, truncated CelD (62.0 kDa) and the fusion proteins, EXLX1-CelD (85.1 kDa), EXLX1-GS3-CelD (86.1 kDa) and EXLX1-GS6-CelD (87.0 kDa), were highly expressed in *E. coli*. All of the proteins were successfully purified by Ni-NTA affinity chromatography, because no bands of endogenous proteins from *E. coli* were observed in SDS-PAGE analysis.

Characterization of cellulose with different crystallinities

Figure 3 shows the photographs (a) and XRD pattern (b) of cellulose pretreated with different concentrations of phosphoric acid, along with the results of untreated crystalline cellulose. Cellulose became more swollen as the concentration of phosphoric acid was increased (Fig. 3a). Similarly, the XRD peak intensity at 2θ = 22.5° assigned to the crystalline cellulose decreased with increasing concentration of phosphoric acid (Fig. 3b), indicating that the disruption of cellulose crystal. For a characteristic XRD pattern of the cellulose treated with

82.1% phosphoric acid, the peak intensity at 2θ = 22.5° decreased and at the same time the peak around 2θ = 18~20° indicating amorphous structure remarkably increased. Similar changes in peak shape during pretreatment are also found in other reports.[20,22] The CrI values of the pretreated cellulose based on the XRD analysis were calculated using equation (1) and are listed in Table 2. The CrI of the cellulose drastically decreased with increasing phosphoric acid concentration. These results indicate that changing the phosphoric acid concentration was useful for the preparation of cellulose with different CrI values. The prepared cellulosic substrates were used for the following experiments.

Evaluation of the binding ability of the protein to cellulose

Cellulose binding abilities of EXLX1, CelD, and fusion enzymes were studied using a series of cellulose substrates with different CrI. Figure 4 shows the time courses of the ratio of bound protein to the cellulosic substrate. For the lowest crystalline cellulose with a CrI of 13.7, EXLX1 reached equilibrium with the ratio of bound protein > 95%. The binding ability of CelD was quite low because of the absence of the carbohydrate binding module (CBM) in the protein molecule.[18] Fusion enzymes (*i.e.*, EXLX1-CelD, EXLX1-GS3-CelD, EXLX1-GS6-CelD) bound in rather high abundance to the cellulose, resulting in a bound protein ratio of ~90%. EXLX1 was reported to bind to cellulose *via* hydrophobic interactions between aromatic amino acids such as Trp in a protein molecule and pyranose rings of cellulose chain,[23] and would work in the same way as the CBM.[24,25] A similar tendency was also observed in other cellulose substrates with CrI values of 57.6 and 74.0. For the highly crystalline cellulose with a CrI of 80.3, however, the binding velocity of each protein was relatively low with a lower ratio of bound protein compared with the others. Fusion of EXLX1 was considered to enhance the cellulose binding ability of CelD, regardless of the

linker length between the two proteins.

Evaluation of the digestibility of the enzymes for the cellulose substrate

In a preliminary examination, fusion enzymes connected with single $(GGGGS)_1$ (GS1 linker) and duplicate $(GGGGS)_2$ (GS2 linker) showed almost same digestibility as directly-fused enzyme with no linker. Therefore, the fusion enzymes with GS1 and GS2 linker were not considered for the further investigation. Then, we focused on the three fusion enzymes whose linker lengths were drastically changed, that is, EXLX1-CelD, EXLX1-GS3-CelD, and EXLX1-GS6-CelD. Figure 5 shows the degradation profiles of the pretreated cellulose with the lowest crystallinity (CrI = 13.7). During the initial stages of the reaction (~24 h), EXLX1-GS3-CelD and EXLX1-GS6-CelD afforded higher digestibility when compared with CelD, whereas the digestibility of the mixed system (EXLX1 + CelD) and that of EXLX1-CelD were lower than that of CelD. The increase in the digestibility by EXLX1-GS3-CelD and EXLX1-GS6-CelD is because of the adequate affinity of the protein towards the cellulose substrate owing to the presence of EXLX1. Although EXLX1-CelD also acquired cellulose-binding ability, its activity was diminished when compared with CelD. It is well known that the linker peptide between two proteins has a strong influence on the activity of the original proteins.[26–28] The activity of CelD in EXLX1-CelD might have declined to some extent because of the direct fusion of the two proteins. The use of the linker peptide between the two proteins was important in maintaining their original activity and for acquiring the advantage of a fusion protein system. The activity of the mixed system (EXLX1 + CelD) was slightly lower than that of CelD. This result is inconsistent with a previous report by Kim *et al.*,[16] in which no negative effect of the addition of EXLX1 was observed in cellulose degradation by commercial cellulases. The slight decrease observed in the present

study may be ascribed to free EXLX1 hampering the ability of CelD to bind to cellulose. In the longer period of the reaction time (~48 h), digestibilities in each enzyme were comparable, with the exception of the low digestibility by EXLX1-CelD. Based on the results, the linker peptide connecting EXLX1 and CelD was shown to have a significant impact on cellulose digestibility of the fusion enzymes.

Further investigation was conducted to study the degradation activity of the fusion enzymes toward the cellulose substrate with different crystallinities. Figure 6(a) shows the degradation profiles of three different celluloses by the three kinds of enzymes, CelD, EXLX1-GS3-CelD and EXLX1-GS6-CelD. For all of the enzymes examined, cellulose with higher crystallinity was shown to be less degradable. As for the activity of each enzyme, EXLX1-GS3-CelD exhibited higher activity compared with CelD for all cellulose substrates, whereas the activity of EXLX1-GS6-CelD was variable and dependent on the crystallinity of the cellulose; EXLX1-GS6-CelD showed higher activity than CelD for the cellulose with a CrI 13.7, but comparable activity for the cellulose with a CrI 57.6 and lower activity for the cellulose with a CrI 74.0.

To evaluate the effectiveness of the fusion of EXLX1, the digestibility of the fusion enzymes (EXLX1-GS3-CelD and EXLX1-GS6-CelD) were divided by the digestibility of CelD to calculate the ratio of digestibility in each reaction system (Fig. 6b). The ratio of EXLX1-GS3-CelD/CelD increased as the CrI value increased, where an inverse trend was observed for the EXLX1-GS6-CelD/CelD ratio. This result indicates that excessively long and flexible linkers, such as the GS6 linker, are not preferable for linking EXLX1 and CelD. An appropriate length, such as the GS3 linker, appears to be crucial for ensuring the activity of EXLX1. It was also revealed that the EXLX1 domain in EXLX1-GS3-CelD enhanced CelD digestibility more effectively in higher crystalline cellulose. This may represent a

verification of the crystal-disrupting function of EXLX1. However, the observed improvement in the activity by fusing EXLX1 was still moderate, and further investigation, including protein conformation and degradation kinetics, is necessary to construct a powerful fusion enzyme that can disrupt and degrade highly crystalline cellulose.

Conclusion

We have prepared a series of fusion enzymes consisting of bacterial expansin EXLX1 and endoglucanase CelD connected with different peptide linkers that vary in length. Fused with EXLX1, CelD acquired high binding ability to various kinds of cellulose with different crystallinity, regardless of whether a linker was present. Among the fusion enzymes examined, EXLX1-GS3-CelD exhibited the highest activity, indicating that the linker peptide between the two proteins influences the activity of the fusion enzyme. Moreover, it was demonstrated that the EXLX1 domain worked more effectively toward higher crystalline cellulose.

Acknowledgements

This work was supported by the JSPS KAKENHI (Grant Number: 25820396) and JST A-STEP program (241FT0189). The authors would like to thank Profs. Tomohito Kameda and Toshiaki Yoshioka (Tohoku University) for the use of the XRD system.

References

1. G. W. Huber, S. Iborra, and A. Corma, *Chem. Rev.*, 2006, **106**, 4044-4098.
2. Y.-H. Percival Zhang, *J. Ind. Microbiol. Biotechnol*., 2008, **35**, 367-375.
3. J. J. Bozell. and G. R. Petersen, *Green Chem*., 2010, **12**, 539-554.
4. M. Pauly and K. Keegstra, *Curr. Opin. Plant Biol*., 2010, **13**, 305-312.
5. L. R. Lynd, P. J. Weimer, W. H. van Zyl, I. S. Pretorius, *Microbiol. Mol. Biol. Rev.*, 2002, **66**, 506-577.
6. M. E. Himmel, S.-Y. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, J. W. Brady, T. D. Foust, *Science*, 2007, **315**, 804-807.
7. L. da Costa Sousa, S. PS Chundawat, V. Balan, B. E Dale, *Curr. Opin. Biotechnol*., 2009, **20**, 339-347.
8. N. Mosier, C. Wyman, B. Dale, R. Elander, Y. Y. Lee, M. Holtzapple, M. Ladisch, *Bioresour. Technol*., 2005, **96**, 673-686
9. D. J. Cosgrove, *Nature*, 2000, **407**, 321-326.
10. D. J. Cosgrove, *Nat. Rev. Mol. Cell Biol.*, 2005, **6**, 850-861.
11. M. Saloheimo, M. Paloheimo, S. Hakola, J. Pere, B. Swanson, E. Nyyssönen, A. Bhatia, M. Ward, M. Penttilä, *Eur. J. Biochem*., 2002, **269**, 4202-4211.
12. A. Levasseur, M. Saloheimo, D. Navarro, M. Andberg, F. Monot, T. Nakari-Setälä, M. Asther, E. Record, *Appl. Microbiol. Biotechnol*., 2006, **73**, 872-880.
13. X. Chen, N. Ishida, N. Todaka, R. Nakamura, J. Maruyama, H. Takahashi, K. Kitamoto, *Appl. Environ. Microbiol*., 2010, **76**, 2556-2561.
14. G. Jäger, M. Girfoglio, F. Dollo, R. Rinaldi, H. Bongard, U. Commandeur, R. Fischer, A. C. Spiess, Jochen Büchs, *Biotechnol. Biofuels*, 2011, **4**:33.
15. F. Kerff, A. Amoroso, R. Herman, E. Sauvage, S. Petrella, P. Filée, P. Charlier, B. Joris, A.

Tabuchi, N. Nikolaidis, D. J. Cosgrove, *Proc. Natl. Acad. Sci. U. S. A*., 2008, **105**, 16876-16881.

16. E. S. Kim, H. J. Lee, W.-G. Bang, I.-G. Choi, K. H. Kim, *Biotechnol. Bioeng*., 2009, **102**, 1342-1353.
17. N. Georgelis, A. Tabuchi, N. Nikolaidis, D. J. Cosgrove, *J. Biol. Chem*., 2011, **286**, 16814-16823.
18. K. Tokatlidis, S. Salamitou, P. Béguin, P. Dhurjati, J.-P. Aubert, *FEBS Lett*. 1991, **291**, 185-188.
19. Y.-H. P. Zhang, J. Cui, L. R. Lynd, L. R. Kuang, *Biomacromol*., 2006, **7**, 644-648.
20. M. Hall, P. Bansal, J. H. Lee, M. J. Realff, A. S. Bommarius, *FEBS J.*, 2010, **277**, 1571-1582.
21. A. Thygesen, J. Oddershede, H. Lilholt, A. B. Thomsen, Kenny Ståhl, *Cellulose*, 2005, **12**, 563-576.
22. A. Mittal, R. Katahira, M. E Himmel, D. K Johnson, *Biotechnol. Biofuels*, 2011, **4**:41.
23. N. Georgelis, N. H. Yennawar, D J. Cosgrove, *Proc. Natl. Acad. Sci. U. S. A*., 2012, **109**, 14830-14835.
24. A. B. Boraston, D. N. Bolam, H. J. Gilbert, G. J. Davies, *Biochem. J.* 2004, **382**, 769-781.
25. J. Hong, X. Ye, Y.-H. P. Zhang, *Langmuir*, 2007, **23**, 12535-12540.
26. X. Chen, J. L. Zaro, W. Shen, *Adv. Drug Deliv. Rev.*, 2013, **65**, 1357-1369.
27. G. K. Sonan, V. Receveur-Brechot, , C. Duez, N. Aghajari, M. Czjzek, R. Haser, C. Gerday, *Biochem. J*., (2007) **407**, 293-302.
28. N. Adlakha, S. Sawant, A. Anil, A. Lali, S. S. Yazdani, *Appl. Environ. Microbiol*., 2012, **78**, 7447-7454.

Figure captions

Fig. 1 Schematic representation of enzymes in this study.

Fig. 2 SDS-PAGE analysis of recombinant proteins after purification. Lane M, molecular-weight marker; lane 1, EXLX1; lane 2, CelD; lane 3, EXLX1-CelD; lane 4, EXLX1-GS3-CelD; lane 5, EXLX1-GS6-CelD.

Fig. 3 Photographs (**a**) and XRD patterns (**b**) of cellulose pretreated with different concentration of phosphoric acid. The concentration of phosphoric acid used for the pretreatment was 76.9%, 77.7%, and 82.1%.

Fig. 4 Time courses of the ratio of bound protein to cellulose with different CrI. Cellulose concentration (5 g/L) and protein concentration (1.0 μmol/L) were same in all experiments.

Fig. 5 Time courses of cellulose degradation by cellulases. Cellulose with the lowest crystallinity (CrI = 13.7) was used as a substrate in this experiment.

Fig. 6 Digestibility of cellulose with different crystallinity (**a**), and the ratio of digestibility of CelD and the fusion enzymes (**b**).

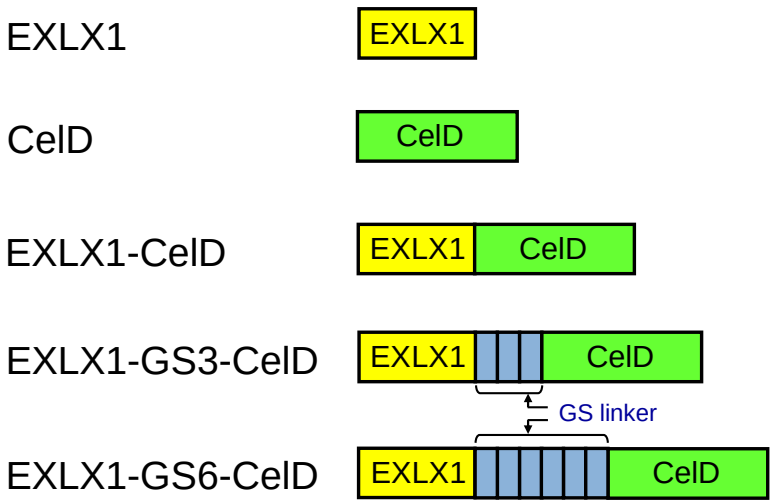

Fig. 1 Schematic representation of enzymes in this study.

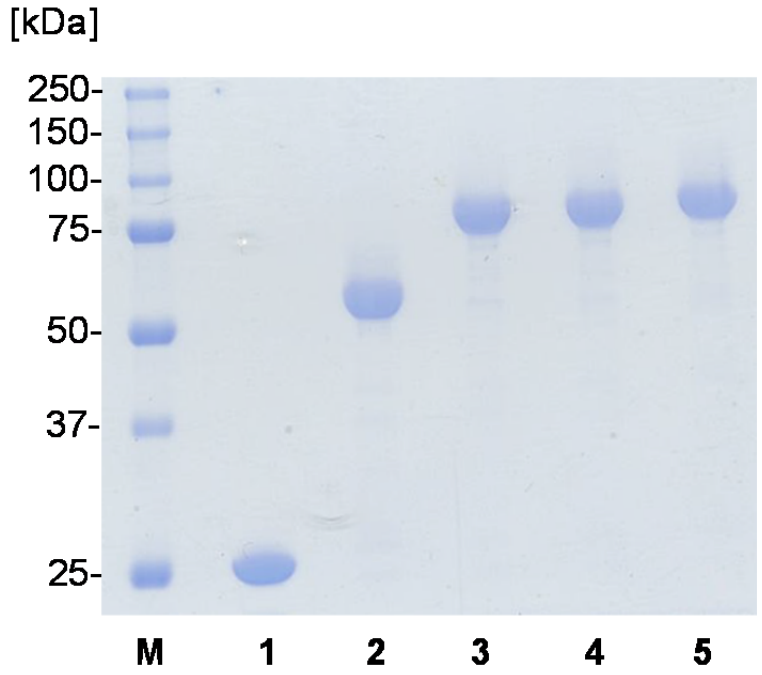

Fig. 2 SDS-PAGE analysis of recombinant proteins after purification. Lane M, molecular-weight marker; lane 1, EXLX1; lane 2, CelD; lane 3, EXLX1-CelD; lane 4, EXLX1-GS3-CelD; lane 5, EXLX1-GS6-CelD.

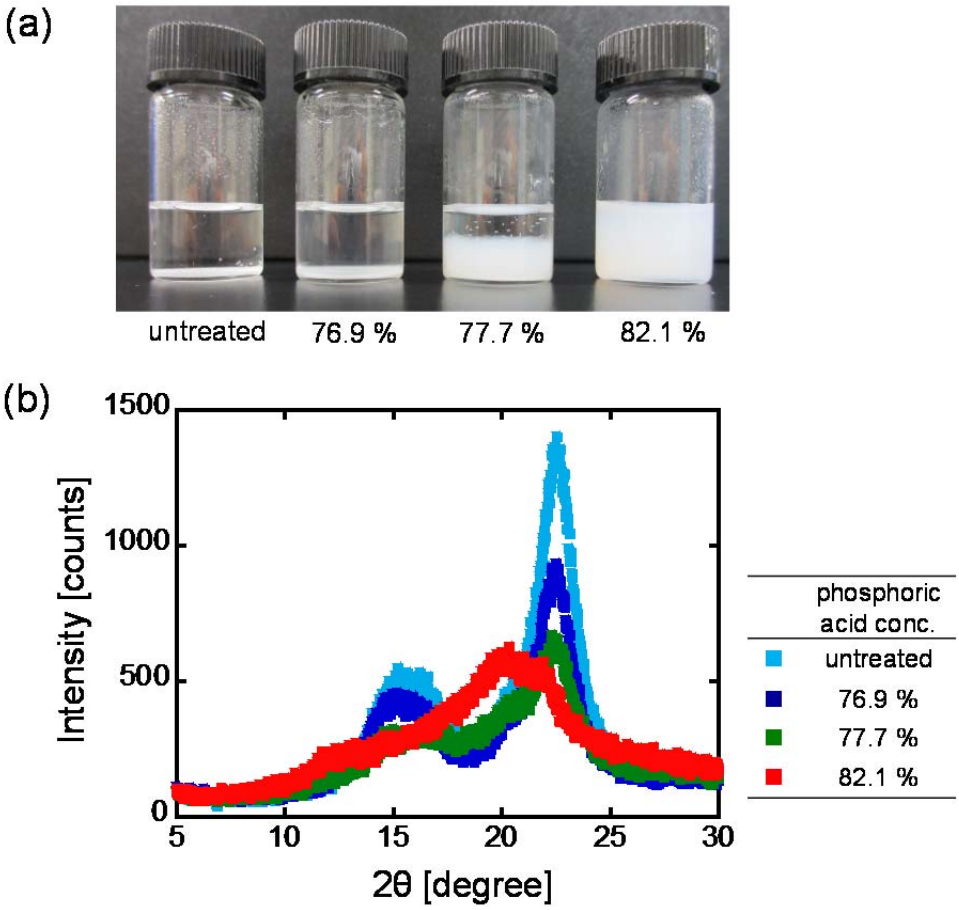

Fig. 3 Photographs (**a**) and XRD patterns (**b**) of cellulose pretreated with different concentration of phosphoric acid. The concentration of phosphoric acid used for the pretreatment was 76.9%, 77.7%, and 82.1%.

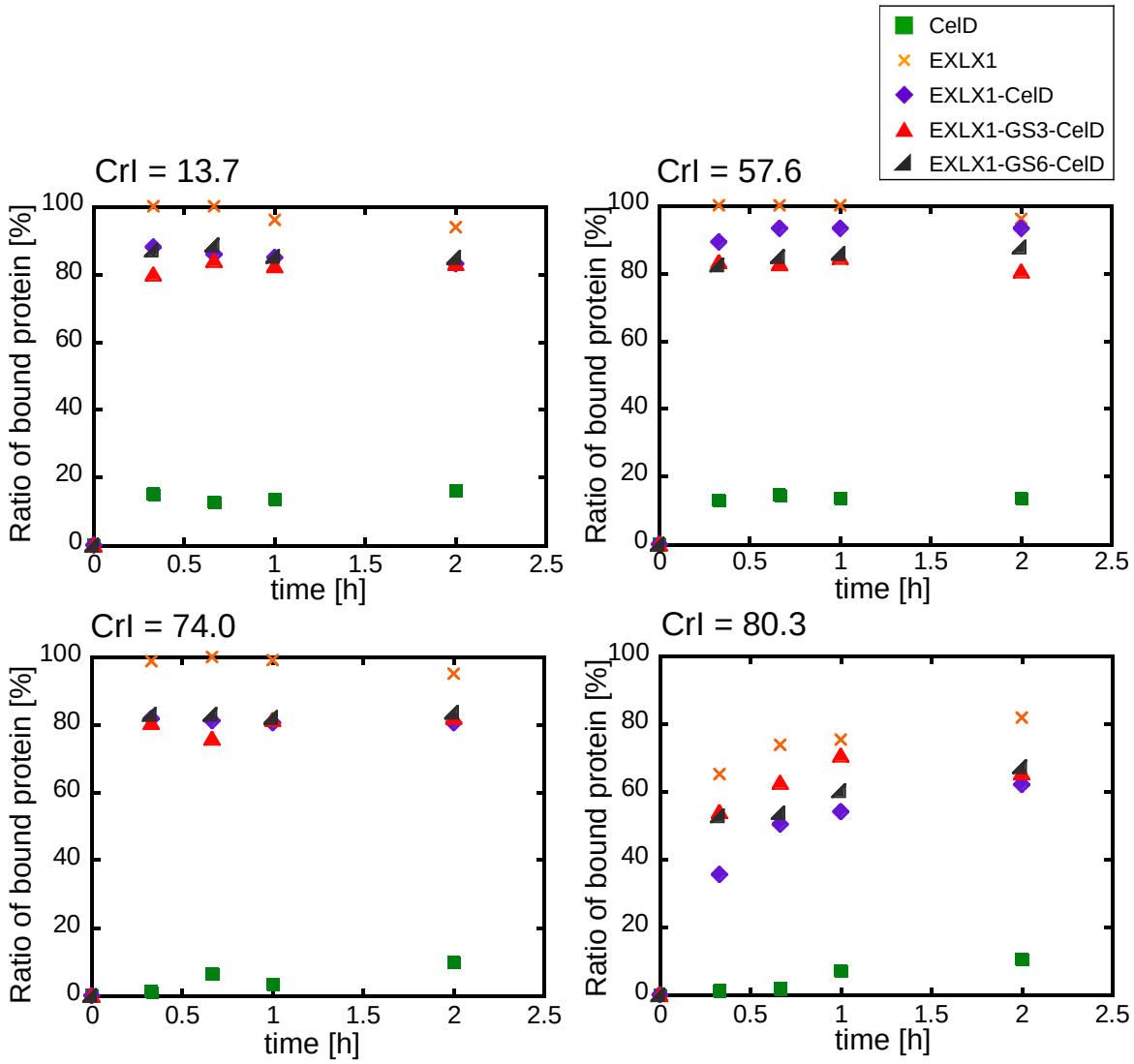

Fig. 4 Time courses of the ratio of bound protein to cellulose with different CrI. Cellulose concentration (5 g/L) and protein concentration (1.0 μmol/L) were same in all experiments.

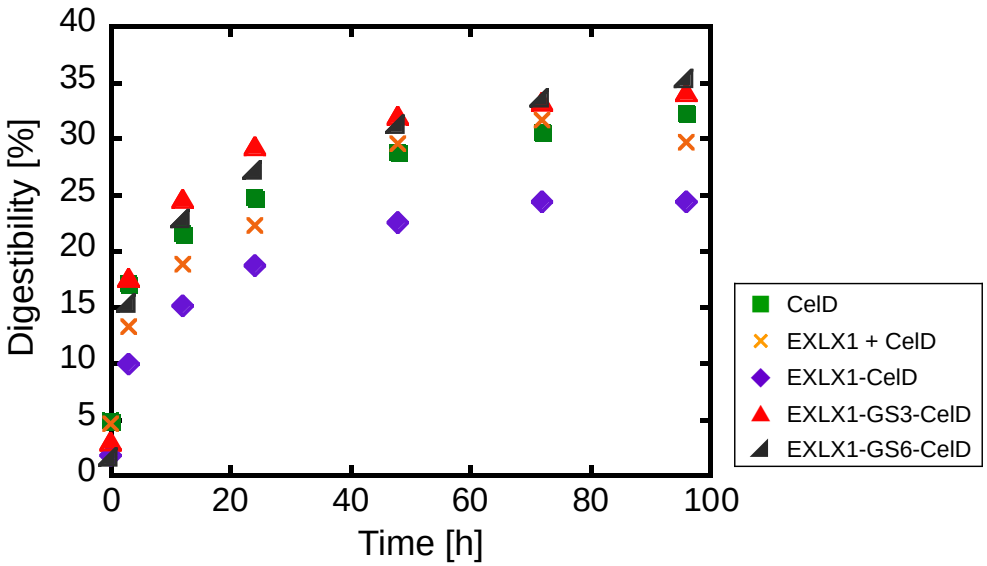

Fig. 5 Time courses of cellulose degradation by cellulases. Cellulose with the lowest crystallinity (CrI = 13.7) was used as a substrate in this experiment.

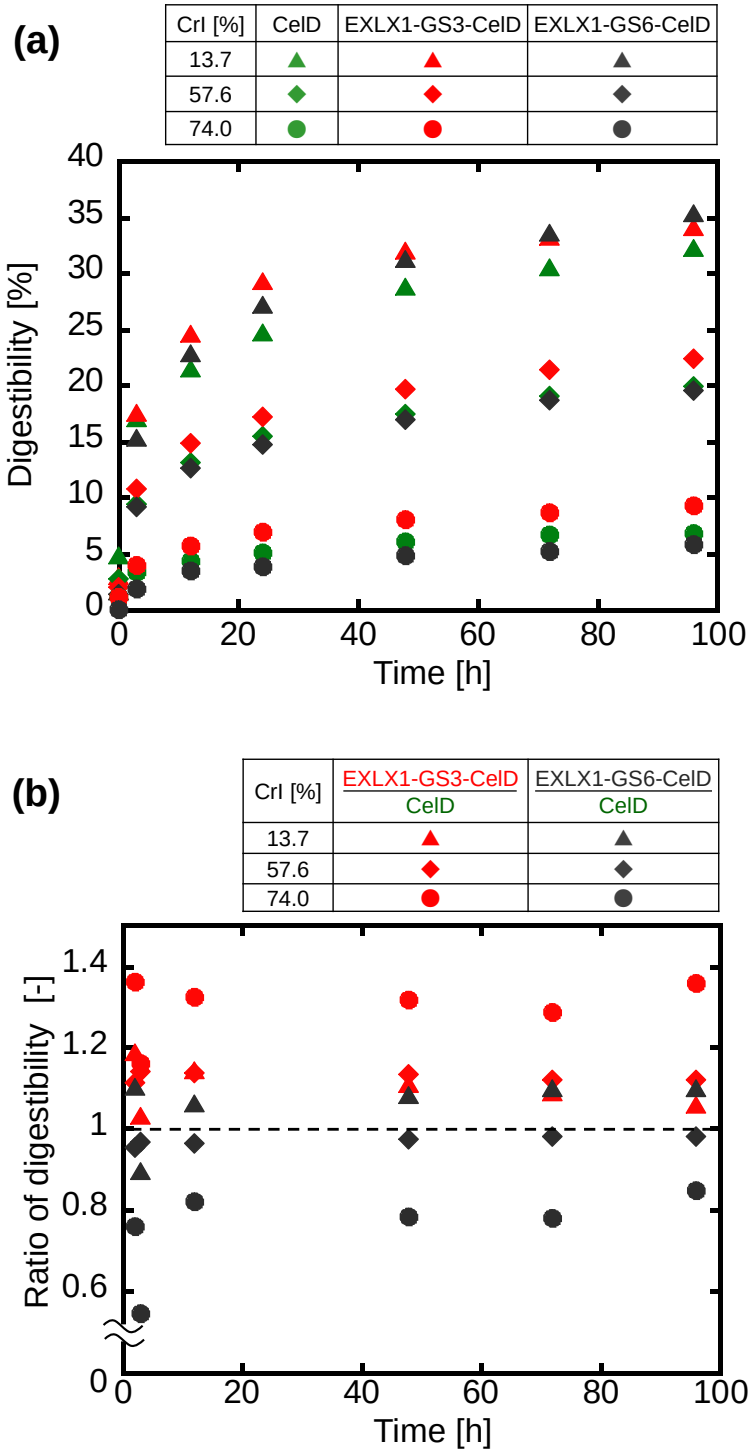

Fig. 6 Digestibility of cellulose with different crystallinity (**a**), and the ratio of digestibility of CelD and fusion enzymes (**b**).

Table 1 Nucleotide sequences of primers used in this study

Primer	Nucleotide sequence
p1-F	5’-AAGGAGATATACATATGGCATATGACGACCTGCATGA-3’
p2-R	5’-GGTGGTGGTGCTCGAGTTCAGGAAACTGAACATGGC-3’
p3-F	5’-AAGGAGATATACATATGGCAAAAATAACGGAGAATTA-3’
p4-R	5’-GGTGGTGGTGCTCGAGGTACAGTACTTCATTTTGAG-3’
p5-F	5’-GTTTCCTGAACTCGAGGCAAAAATAACGGAGAATTA-3’
p6-F	5’-GTTTCCTGAACTCGAGGGTGGCGGTGGCTCGGGCGGTGGTGGGTCGGGTGGCGGCGGATCCGCAAAAATAACGGAGAATTA-3’
p7-R	5’-GGATCCGCCGCCACCCGACCCACCACCGCCCGAGCCACCGCCACCGAATTCTTCAGGAAACTGAACATGG-3’
p8-F	5’-GGTGGCGGCGGATCCGGTGGCGGTGGCTCGGGCGGTGGTGGGTCGGGTGGCGGCGGATCCGCA-3’
p9-R	5’-CACCGCCACCGAATTCGGATCCGCCGCCACCCGACCCACCACCGCCCGAGCCACCGCCACCTTCAGGAAACTGAACATGGC-3’

Table 2 Crystallinity index (CrI) of cellulose pretreated with phosphoric acid with different concentration.

Concentration of phosphoric acid [wt%]	CrI
untreated	80.3
76.9	74.0
77.7	57.6
82.1	13.7

Table of contents entry

Fusion enzymes composed of bacterial expansin (EXLX1) and cellulase (CelD) exhibited high binding ability and increased digestibility to cellulose with various crystallinity.