Doctoral Dissertation (Shinshu University)

Structure-function relations of family 1 carbohydrate-binding module from white-rot fungi

March 2015

Hiroto Nishijima

CONTENTS

Abbreviations

1. Terms

GH family	glycoside hydrolase family
Sup	supernatant
Amp^r	resistance gene to ampicillin
M.w.	molecular weight

2. Proteins

CBH	cellobiohydrolase
EG	endoglucanase
BGL	β-glucosidase
Xyn	xylanase
CBM	carbohydrate-binding module
CBM1	carbohydrate-binding module family 1
GFP	green fluorescent protein
CBM-GFP	chimeric protein of CBM1 and GFP
Xyn10B	xylanase belonging to GH10 from *Irpex lacteus*
*Tr*XYN III	xylanase belonging to GH10 from *Trichoderma reesei*
CBM_{Xyn10B}	CBM1 region of Xyn10B
CBM_{Xyn10B}-GFP	chimeric protein of CBM_{Xyn10B} and GFP
CBM_{ThEG1}	CBM1 region of EG1 from *Trametes hirsuta*
CBM_{ThEG1}-GFP	chimeric protein of CBM_{ThEG1} and GFP
CBM_{F42S}-GFP	phenylalanine 42 of CBM_{Xyn10B}-GFP mutated to serine
CBM_{Y52S}-GFP	tyrosine 52 of CBM_{Xyn10B}-GFP mutated to serine

3. Carbohydrates

CMC	carboxymethyl cellulose
G1	glucose
G2	cellobiose
MC	microcrystalline cellulose
*p*NP	*p*-nitrophenol
*p*NPG1	*p*NP-β-D-glycopyranoside
*p*NPG2	*p*NP-β-D-cellobioside
*p*NPL	*p*NP-β-D-lactopyranoside
*p*NPX1	*p*NP-β-D-xylopyranoside
X1	xylose
X2	xylobiose
X3	xylotriose
X4	xylotetraose
X5	xylopentaose
X6	xylohexose

4. Other reagents

AAB	ammonium acetate buffer
CBB-R250	Coomassie brilliant blue R250
IPTG	isopropyl β-D-thiogalactopyranoside
LB medium	Luria-Bertani medium
SAB	sodium acetate buffer
SDS	sodium dodecyl sulfate

5. Instrument and analytical methods

SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TLC	thin-layer chromatography

Chapter 1

General introduction

Chapter 1: General introduction

1.1. Carbohydrate-binding module

CBMs, which are present in enzymes that undertake plant cell wall degradation, are categorized into three types, type A, type B, and type C CBMs, based on their structure and function [1]. Type A CBMs have a flat hydrophobic surface composed of three aromatic amino acid residues binding to crystalline polysaccharides, such as cellulose and chitin (**Figure 1-1 (A)**) [2]. Type B CBMs can bind to amorphous cellulose or xylan, using open grooves to accommodate a single polysaccharide chain. Type C CBMs can bind only to the terminus of glycans or oligosaccharides as their binding sites are short pockets that cannot accommodate long glycan chains.

To date, CBMs have been categorized into 69 families based on the amino acid sequence, according to the CAZy (carbohydrate-active enzymes) database (http://www.cazy.org/) [3]. Carbohydrate-binding module family 1 (CBM1) belongs to the type A CBMs and is known to be a cellulose-binding domain (CBD) that is found in fungal cellulolytic and related proteins. CBM1 is a small stable domain that consists of around 30-40 amino acid residues, and it is connected to a catalytic domain of cellulolytic enzymes by a linker region. CBM1 facilitates adsorption on cellulose and enhances its degradation (**Fig. 1-1 (B)**) [4].

1.2. Distribution of carbohydrate-binding module

Cellulose is a native polysaccharide composed of more than 1000 β-D-glucose units linked by β-1,4-glycosidic bonds. Plant cell walls are mainly composed of cellulose, hemicellulose (xylan, mannan, xyloglucan, etc.), and lignin. Enzyme complexes of cellulases, hemicellulases, and/or lignin-degrading enzymes from fungi have the potential to hydrolyze these polymers to monomeric units. Several commercial enzyme preparations have been applied for quality improvement of detergent, in the paper-making industry, and for bioethanol production from cellulosic biomass. It is known that they contain many kinds of components of cellulases and related enzymes. Cellulases are categorized into four types, namely, cellobiohydrolase I (CBH I), cellobiohydrolase II (CBH II), endoglucanase (EG), and β-glucosidase (BGL), based on their roles in cellulose degradation. It is thought that the presence of these enzymes at an appropriate ratio can hydrolyze crystalline cellulose to

glucose effectively. Many of these cellulases have CBM1 in the N- or C-terminus region of proteins for effective cellulose degradation. CBM1s are also distributed in several hemicellulases, such as xylanase and mannanase, and non-hydrolytic proteins, such as swollenin, polysaccharide monooxygenase, cellobiose dehydrogenase, and carbohydrate esterase, although they have important roles in cellulose degradation [5-10]. CBM1s not only have important roles in cellulose degradation, but also in plant cell wall degradation. From this point of view, the role of CBM1s might vary from enzyme to enzyme, and the structure of each CBM1 might differ.

1.3. Variety of amino acid sequence of CBM1

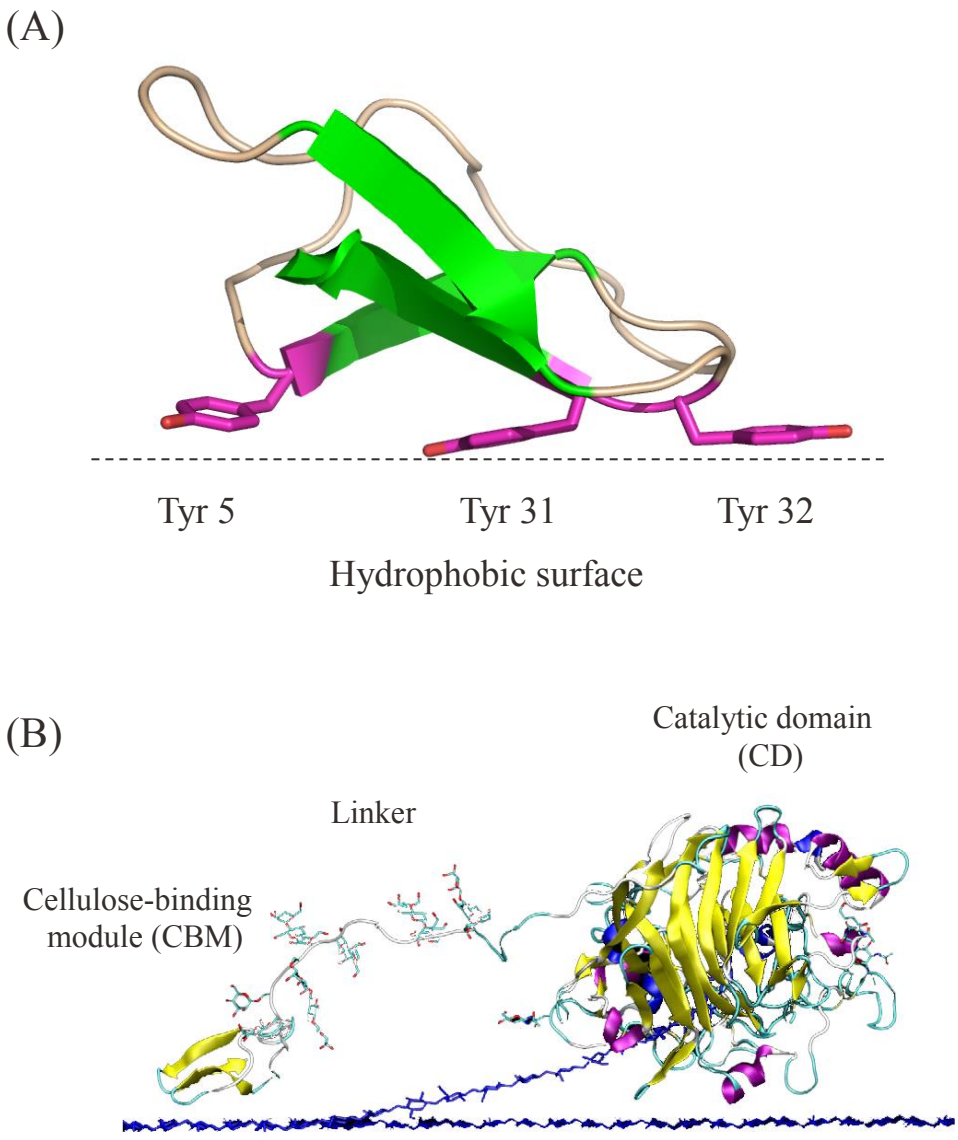
It has been proposed that three aromatic amino acid residues located on the flat surface of CBM1 are essential for the adsorption on cellulose. The motifs of these three amino acid residues were found to be Y-YY, W-WY, W-YY, Y-WY, and Y-FY, and the ability of adsorption of each CBM1 on cellulose is closely related to its hydrophobicity. These varieties of structure should be adapted to the variety of enzyme, as the mode of action of each enzyme for biomass degradation differs from one enzyme to another. On the other hand, the absence of CBM1 remarkably reduces crystalline cellulose degradation due to a decrease in adsorption [11-12]. As reported in *Trichoderma reesei* Cel7A (CBH I) and Cel6A (CBH II), the adsorption behaviors of the CBM1s on cellulose are not the same. The retention time of Cel7A on the cellulose surface during one action is relatively short, and the protein can move to the next region of the same cellulose chain. As such, Cel7A is called a processive enzyme and the typical motif of CBM1 in Cel7A on cellulose adsorption is Y-Y-Y. On the other hand, Cel6A can adsorb on cellulose for a longer period than Cel7A because of its higher binding parameters; its motif is W-Y-Y. In addition, CBM1 from Cel6A has an additional S-S bond, which produces a stable structure and conveys a binding capacity that differs from that of conventional CBM1 [13]. Furthermore, new types of CBM1 with unique motifs (Y-W-NY and Y-W-SF) have been reported recently in *Oomycetes* sp. [14]. Several hundred CBM1s have been identified to date on the basis of amino acid sequence similarities. The differences in the amino acid sequences of CBM1 should affect the binding ability; however, the relationship between structure and function has not been identified.

1.4. Purpose of this research

In nature, basidiomycetes and fungi play important roles in plant cell wall degradation. These microorganisms produce various types of cell wall-degrading enzyme, such as cellulases, hemicellulases, accessory enzymes, and some non-enzymatic proteins like swollenin. In many cases, these enzymes have CBM1 at the N- or C-terminal region. Presumably, the adsorption ability of CBM1 to insoluble substrates like the cell wall is a basic strategy and essential for efficient degradation among fungal cellulolytic enzymes. On the other hand, other CBM families, such as CBM2, CBM3, CBM4, CBM6, CBM8, and CBM9, that show the ability to absorb to cellulose have been reported, as found in bacterial enzymes. Actually, numerous CBM families were distributed in bacterial cellulolytic enzymes according to the type of polysaccharides, crystallinity of cellulose, and other factors. Against this background, it is an interesting question why basidiomycetes and fungi use only one type of CBM given the complicated structure of plant cell walls. In this context, it seems particularly important to clarify the structure-function relationships of CBM1 from white-rot fungi. Therefore, the mechanism of adsorption and desorption of CBM1 has been studied and its behavior is discussed in this thesis.

CBM1s from two basidiomycetes, *Trametes hirusta* and *Irpex lacteus*, have been focused on in this study. As the binding of an enzyme on its substrate is the first step of an enzymatic action, CBM1 could work effectively to concentrate it on the surface of cell walls, especially at low substrate concentration. In addition, desorption of enzyme is also important for enzyme recovery. Therefore, the control of adsorption/desorption of CBM1 with buffers having different ion strengths (using CBM1 of EG1 from *T. hirsuta*) was first investigated in this study. Recently, it was reported that there are some varieties of CBM1 based on their adsorption ability. It is well known that fungi produce many kinds of cellulolytic enzyme for plant cell wall degradation. The function of CBM1 should differ among these enzymes. White-rot fungus, *I. lacteus*, can also secrete different types of cellulase, most of which have CBM1, but some components in crude enzyme preparation were shown to lose CBM1 by proteolysis. This means that there are many types of enzyme in crude enzyme preparations, based on the adsorption ability. Their adsorption properties might be related to the function of enzymes. Here, I have focused on the fraction with strong adsorption ability on cellulose. The structure of CBM1 in this fraction has been analyzed and a new type of CBM1 has been found. I have studied the reason why this type of CBM1 can adsorb on cellulose strongly, as

described in the following chapter. In addition, similar types of CBM1 have been found in many enzymes from basidiomycetes and fungi. Therefore, the relationships between the structure and function of these enzymes are discussed in this thesis.



L. Zhong *et al*. Carbohydr Res, 344, 1984 (2009)

Fig. 1-1. (A) CBM1 (typical type A CBM) of Cel7A (CBH I) from *T. reesei* (PDB ID: 1CBH) [2] and (B) CBH I from *T. reesei* docked onto the cellulose surface by CBM1 [4]. The dotted line shows the flat hydrophobic surface consisting of three aromatic amino acids (Tyr 5-Tyr 31-Tyr 32) for adsorption to cellulose.

References

1. **Boraston AB, Bolam DN, Gilbert HJ, Davies GJ.** 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. **382:**769-781.
2. **Mattinen ML, Kontteli M, Kerovuo J, Linder M, Annila A, Lindeberg G, Reinikainen T, Drakenberg T.** 1997. Three-dimensional structures of three engineered cellulose-binding domains of cellobiohydrolase I from *Trichoderma reesei*. Protein Sci. **2:** 294-303.
3. **Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B.** 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic acids Res. **37:**233-238.
4. **Zhong L, Matthews JF, Hansen PI, Crowley MF, Cleary JM, Walker RC, Nimlos MR, Brooks CL 3rd, Adney WS, Himmel ME, Brady JW.** 2009. Computational simulations of the *Trichoderma reesei* cellobiohydrolase I acting on microcrystalline cellulose Iβ: the enzyme–substrate complex. Carbohydr. Res. **344:**1984-1992.
5. **Khan SN, Akter MZ, Sims PFG.** 2002. Cloning and expression analysis of two endo-1,4-β-xylanase genes from *Phanerochaete chrysosporium*. Plant Tissue Cult. **12**:57-68.
6. **Tenkanen M, Buchert J, Viikari L.** 1995. Binding of hemicellulases on isolated polysaccharide substrates. Enzyme Microb. Technol. **17**:499-505.
7. **Saloheimo M, Paloheimo M, Hakola S, Pere J, Swanson B, Nyyssönen E, Bhatia A, Ward M, Penttilä M.** 2002. Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. Eur. J. Biochem **269**:4202–4211.
8. **Svein JH, Gustav VK, Bjørge W, Vincent GHE.** 2012. Novel enzymes for the degradation of cellulose. Biotechnol. Biofuels, doi:10.1186/1754-6834-5-45.
9. **Igarashi K, Marc FJMV, Samejima M, Martin SL, Karl-Erik LE, Nishino T.** 1999. Cellobiose dehydrogenase from the fungi *Phanerochaete chrysosporium* and *Humicola insolens*. A flavohemoprotein from *Humicola insolens* contains 6-hydroxy-FAD as the dominant active cofactor. J. Biol. Chem. **274**:3338-44.
10. **Dominic WSW, Victor JC, Amanda A, McCormack JH, Peter B.** 2012. Functional cloning and expression of the *Schizophyllum commune* glucuronoyl esterase gene and characterization of the recombinant enzyme. Biotechnol. Res. Int. **2012**: ID 951267.
11. **Hamada N, Kodaira R, Nogawa M, Sinji K, Ito R, Amano Y, Shimosaka M, Kanda T, Okazaki M.** 2001. Role of cellulose-binding domain of exocellulase I from white rot basidiomycete *lrpex lacteus*. J. Biosci. Bioeng. **91:**359-362.
12. **Thongekkaew J, Ikeda H, Masaki K, Iefuji H.** 2013. Fusion of cellulose binding domain from *Trichoderma reesei* CBHI to *Cryptococcus* sp. S-2 cellulase enhances its binding affinity and its cellulolytic activity to insoluble cellulosic substrates. Enzyme Microb. Technol. **52:**241-246.
13. **Geraldine C, Markus L.** 1999. Widely different off rates of two closely related

cellulose-binding domains from *Trichoderma reesei*. Eur. J. Biochem. **262:**637-643.
14. **Larroque M, Barriot R, Bottin A, Barre A, Rouge P, Dumas B, Gaulin E.** 2012. The unique architecture and function of cellulose-interacting proteins in oomycetes revealed by genomic and structural analyses. BMC Genomics. **13:**605.

Chapter 2

Adsorption behavior of CBM1 of endoglucanase from *Trametes hirsuta* on cellulose

Chapter 2: Adsorption behavior of CBM1 of endoglucanase from *Trametes hirsuta* on cellulose

2. Introduction

Carbohydrate-binding module family 1 (CBM1) is widely distributed in fungal cellulases and proteins related to cellulose degradation, and CBM1 may be found at either the N- or the C-terminus of the protein. All CBM1 members show high primary structure similarity regardless of the type of catalytic domain. During the degradation of crystalline cellulose, CBM1 gives the catalytic domain a chance to approach the substrate, thereby enhancing degradation at the liquid-solid interface. The adsorption of CBM1 is thus necessary for efficient attack on insoluble substrates. The interaction between CBM1 and cellulose is diminished by the addition of ethylene glycol [1], cellodextrins, and an increase in temperature [2]. However, the additives have to be removed before the following round of saccharification, while taking care not to denature the enzymes. On the other hand, temperature is easy to regulate, but does not result in aggressive desorption. A simple and effective process, with mild conditions for enzymes and the environment, was required for the desorption of the enzyme. Some chaotropic agents, such as ammonium sulfate, acetate, urea, and iodine, are known to affect hydrophobic interactions [3-4]. These have been used in hydrophobic chromatography [5] and the purification of DNA fragments [6], among other uses. DNA is separated easily from cell lysates by increasing the hydrophobic interaction between nucleic acids and silica particles. In the same manner, the interaction between CBM1 and cellulose would also be controllable by changing the concentration of chaotropic agent. As described in this chapter, the adsorption behavior of CBM1 was studied using the CBM1 of EG1 from *Trametes hirsuta* (CBM_{ThEG1}) fused to GFP protein (CBM_{ThEG1}-GFP), and the possibility of recycling the enzyme was investigated.

2.1. Materials and methods

2.1.1. Chemicals and strains

Microcrystalline cellulose (MC) (column chromatography grade, Merck, Germany) was used as a substrate for cellulose-binding assay. *Escherichia coli* DH5α purchased from Takara (Japan) was used as the host strain for plasmid extraction, grown at 37˚C in Luria-Bertani medium containing 50 μg/mL ampicillin. Plasmid pBluescript II SK(+) was used as the

subcloning vector. *E. coli* BL21 (DE3) Origami strain (Novagen, NJ) and plasmid pRSET/EmGFP (Invitrogen, CA) were used for CBM1 of *T. hirsuta* from EG1 fused to green fluorescence protein (CBM_{ThEG1}-GFP) expression. *Aspergillus oryzae* niaD^{-} and expression vector pNAN-8142 were used as a host for the expression of recombinant cellulase (kindly provided by Ozeki Corporation, Japan). Carboxymethyl cellulose (CMC) (Tokyo Kasei, Japan) and MC were used for enzymatic assays. Schematic diagrams of the recombinant proteins are shown in **Fig. 2-1**.

2.1.2. Analytical methods

SDS-PAGE was performed by the method of Laemmli [7]. As the molecular weight standard, low range (BIO-RAD, USA) marker was used. The fluorescence intensities of CBM_{ThEG1}-GFP and GFP were measured using an FP-6200 spectrofluorometer (JASCO, Japan).

2.1.3. Plasmid construction and expression of CBM1 of *Th*EG1 and GFP fusion protein

The amino acid sequence of the CBM1 from *T. hirsuta* EG1 is VWGQCGGIGFSGDTTCTASTCVKVNDYYSQCQ (three aromatic amino acid residues that are essential for adsorption are underlined). The cDNA encoding the region of CBM1 of *T. hirsuta* EG1 (CBM_{ThEG1}) was amplified by PCR using PrimeSTAR HS DNA polymerase (Takara, Shiga, Japan) with sense primer (5'-GCCCGGATCCTCGACCGCCGTCTGGGG-3', *Bam*HI site is underlined) and reverse primer (5'-GTCGCCATGGGTGCGGACGCGCCAGC-3', *Nco*I site is underlined). The amplified CBM_{ThEG1} fragment was digested with *Bam*HI and *Nco*I, and ligated to the same site of pRSET/EmGFP for fusion to GFP with a histidine tag. The resulting plasmid named pRSET-CBM_{ThEG1}-GFP was used as an *E. coli* expression vector.

E. coli BL21 (DE3) Origami strain was transformed by pRSET-CBM_{ThEG1}-GFP and grown in 200 mL of Luria-Bertani medium containing 100 μg/mL ampicillin at 30˚C with rotary shaking at 180 rpm. When optical density (OD) 600 reached 0.6, isopropyl β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. The culture was further incubated overnight, and then cells were harvested by centrifugation and suspended in 20 mM sodium phosphate buffer (pH 7.4) containing 0.5 M NaCl. Resuspended cells were disrupted by ultrasonic treatment and each CBM_{ThEG1}-GFP was purified from

supernatant using a His-trap FF column (5 mL, GE Healthcare, UK) according to the manufacturer's instructions. Purified CBM$_{Th\text{EG1}}$-GFP was extensively dialyzed against 2 mM sodium acetate buffer (SAB) (pH 5.5). As a negative control, GFP was also expressed in *E. coli* using pRSET/EmGFP vector and was purified by the same procedure.

2.1.4. Plasmid construction and expression of recombinant enzyme in *Aspergillus oryzae*

The CBM1 deletion mutant of r*Th*EG1 (r*Th*EG1*Δ*CBM) was expressed in *A. oryzae*. The cDNA region of the *Th*EG1 catalytic domain was amplified by PCR using sense primer (5'-GCGTCAGCTGTCACGTCGACCGCCTC-3', *Pvu*II site is underlined) and antisense primer (5'-GGGAATTCGAGAAGAGATGAAGCGCTGTATTCC-3'). The amplified fragment was digested by *Pvu*II and *Hin*dIII (internal site of *Th*EG1). To fuse the signal sequence, the fragment was inserted into pBS/eg1, which is a plasmid vector containing the full-length *Th*EG1 cDNA, after the *Psh*AI- and *Hin*dIII-digested fragments had been removed. After the sequence was confirmed, the entire insertion was transferred into pNAN-8142 vector digested by *Xho*I and *Xba*I. Construction of the *A. oryzae* expression system of r*Th*EG1 and rEx-1 from *I. lacteus* MC-2 was performed as previously described by Nozaki [8] and Toda [9].

The *A. oryzae* transformants were grown in SPY medium containg 2% starch, 1% polypeptone, 0.5% yeast extract, 0.5% KH_2PO_4, and 0.05% $MgSO_4$ for 3 days at 30˚C with rotary shaking at 120 rpm. Purification from the culture medium supernatants of each recombinant cellulase was as described in previous studies [8] [9].

2.1.5. Measurement of protein amount

The reaction mixture (0.1 mL) consisted of 50 μg/mL CBM$_{Th\text{EG1}}$-GFP or recombinant enzymes, 10 mg/mL MC, 50 μg/mL bovine serum albumin (BSA), and 0.2 M Britton-Robinson's wide range buffer (pH 4-11) or 2-100 mM SAB (pH 5.5), in a 0.2 mL siliconized tube. It was incubated for 5 hours at 4˚C, being turned upside down once per min. After the suspension had been centrifuged, supernatant and precipitate were isolated, and precipitate was further washed twice with 100 mM SAB (pH 5.5).

The amounts of GFP and CBM$_{Th\text{EG1}}$-GFP adsorbed on MC were determined by measuring the decrease of fluorescent intensity (excitation/emission = 487/509 nm) in the supernatant of the reaction mixture. The amounts of protein adsorbed to MC, which

corresponds to the precipitate described above, were also estimated from the band intensity by using the image analysis software ImageJ (http://imagej.nih.gov/ij/) after electrophoresis.

2.1.6. Time course of adsorption and desorption

CBM_{ThEG1}-GFP, r*Th*EG1, and rEx-1 (50 μg/mL) were individually adsorbed to 1 mg of MC in the presence of 100 mM SAB (pH 5.5), in a total volume of 0.2 mL, for 5 hours at 30˚C. After centrifugation at 13,000 rpm for 5 min, the obtained precipitate was suspended in distilled water for 3 hours. Desorption was carried out at 45˚C with incubation for 3 hours. During this experiment, the reaction mixture was turned upside down once per min.

2.1.7. Enzymatic assay

r*Th*EG activity was determined in 0.2 M Britton-Robinson's buffer (pH 4-13) as previously described by Nozaki [8]. Either r*Th*EG1 or r*Th*EG1*Δ*CBM (50 μg/mL for MC or 250 ng/mL for CMC) was incubated with 10 mg/mL substrate and 50 mg/mL BSA in 2-100 mM SAB (pH 5.5) in a total volume of 0.3 mL. Incubation was carried out for 60 hours (for MC) or 30 min (for CMC) at 30˚C, with turning upside down once per min. After incubation, the enzymes were heat-denatured by boiling. All buffer concentrations and their volumes were then adjusted to 100 mM and 0.35 mL so as not to influence color development in the following analysis. The amount of reducing sugar produced in the following analysis was measured by the method of Somogyi-Nelson using D-glucose as a standard [10-11].

2.2 Results

2.2.1. Expression and purification methods of each protein

The GFP and CBM_{ThEG1}-GFP were expressed as 6 × His-tag fusion protein in *E. coli* and purified nickel affinity chromatography. Two proteins, r*Th*EG1 and r*Th*EG1*Δ*CBM, were expressed in *A. oryzae* and purified by DEAE Sepharose CL-6B and Toyopearl HW50S [8]. The rEx-1 was also expressed in *A. oryzae* and purified with a cellulose affinity column [9]. These proteins were purified to a single band on SDS-PAGE (**Fig. 2-2**).

2.2.2. Adsorption conditions of CBM_{ThEG1}-GFP on cellulose

Appropriate conditions for adsorption were determined first. CBM_{ThEG1}-GFP was incubated with MC under various pH conditions for 5 hours at 4˚C. As shown in **Fig. 2-3**, the

highest adsorption (about 65%) was observed around pH 4.5-5.5. At this pH, about 65% of CBM$_{Th\text{EG1}}$-GFP adsorbed to MC. This pH also corresponded to the optimum pH of r*Th*EG for MC hydrolysis. At higher pH, adsorption decreased gradually. About 50% of CBM$_{Th\text{EG1}}$-GFP adsorbed at pH 9.0 and r*Th*EG1 showed only 10% activity at this pH. This pH-dependent adsorption was very similar to that of CBM1 from *T. reesei* Cel7A (CBH I) [12]. Subsequent adsorption experiments were performed at pH 5.5 because the stability of fluorescence emission was remarkably low below pH 5.0. As for the incubation period, adsorption almost reached saturation by 3 hours at 4˚C, and increased only slightly after this time. The incubation time was therefore fixed at 5 hours at 4˚C in the following experiments.

2.2.3. Regulation of CBM$_{Th\text{EG1}}$-GFP adsorption

GFP and CBM$_{Th\text{EG1}}$-GFP were incubated with MC in various concentrations of SAB (pH 5.5). As a result, the amount of adsorbed protein was increased by increasing the concentration of SAB, as indicated by SDS-PAGE (**Fig. 2-4 (A)**). The CBM$_{Th\text{EG1}}$-GFP apparently adsorbed to MC in a manner dependent on the SAB concentration. The amount of protein reached ~65% at 100 mM SAB (**Fig. 2-4 (B)**), and remained constant at buffer concentrations of 100-300 mM (data not shown). On the other hand, ~8% GFP always adsorbed at any buffer concentration. Besides SAB, ammonium sulfate and NaCl concentrations in the presence of 2 mM SAB (pH 5.5) were also varied. Adsorption increased gradually until 50 mM, but adsorption peaked at only 35-40% (data not shown). It was reported that the adsorption of Cel7A increased 10% in the presence of 0.8 M ammonium sulfate [1]. However, we observed no remarkable increase, possibly because the buffer concentration used for the reaction might have itself been sufficient for adsorption.

2.2.4. Regulation of recombinant enzyme adsorption

As shown in **Fig. 2-5 (A)** and **(B)**, r*Th*EG1, which contains CBM1, adsorbed to MC in the same manner as did CBM$_{Th\text{EG1}}$-GFP. It was perfectly controllable in the range of 10-40% by changing the buffer concentration from 2 to 100 mM. On the other hand, r*Th*EG1*Δ*CBM hardly adsorbed to MC at any concentration. On the other hand, most of the rEx-1 adsorbed tightly to MC regardless of the buffer concentration (**Fig. 2-6**).

2.2.5. Time course of adsorption and desorption

As shown in **Fig. 2-7**, all of the studied proteins adsorbed to MC in 100 mM acetate buffer at 30˚C and rapidly desorbed with water at 45˚C. All adsorptions were saturated within 3 hours, while the time for desorption remained constant at under 30 min. After desorption, most of the MC remained without being degraded. When the temperature of incubation was 30˚C, adsorption was lower than seen at 4˚C in all cases. The decrease in adsorption was 30% for r*Th*EG1 and rEx-1, but only 5% for CBM_{ThEG1}-GFP. During the desorption step, adsorbed proteins were released and the adsorption levels of r*Th*EG1, rEx-1, and CBM_{ThEG1}-GFP decreased to 7, 20, and 35%, respectively. The low adsorption at 45˚C might have been due to the fact that adsorption affinity to cellulose decreased at higher temperatures and because cellulose was partially hydrolyzed during incubation.

2.2.6. Relationship between adsorption and hydrolysis

The hydrolytic ability of each enzyme for MC and CMC was determined at various adsorption levels by changing the buffer concentration. When degrading MC at an SAB concentration of 2 mM, r*Th*EG1 showed almost the same hydrolysis as r*Th*EG1*Δ*CBM (**Fig. 2-8 (A)**). This indicates that hydrolysis proceeded without adsorption at this concentration. At an SAB concentration of 100 mM, r*Th*EG1 showed 1.6-fold-higher hydrolysis than r*Th*EG1*Δ*CBM. This value was reflected in the specific activity of each enzyme. When CMC was used as the substrate, the hydrolysis ratio between r*Th*EG1 and r*Th*EG1*Δ*CBM was nearly constant at any concentration (**Fig. 2-8 (B)**). The hydrolysis of soluble cellulose did not seem to depend on adsorption. Interestingly, each hydrolysis behavior commonly fitted 2 lines, which crossed at around 10 mM. The first line was in the range of 0-10 mM, in which hydrolysis increased remarkably. In the second line (over 10 mM), the hydrolysis of MC alone by r*Th*EG increased gradually, whereas hydrolysis of CMC, which did not require adsorption, remained constant.

2.3. Discussion

The result from the effects of pH on CBM_{ThEG1}-GFP and r*Th*EG1 activity for MC (**Fig. 2-3**) showed a correlation between adsorption and enzymatic hydrolysis. In addition, CBM_{ThEG1} adsorbed to cellulose in a manner dependent on SAB concentration (**Fig. 2-4**). On the basis of the results from the adsorption of CBM_{ThEG1}-GFP, the adsorption behavior of the enzymes was also investigated. The amount of adsorbed CBM_{ThEG1}-GFP and r*Th*EG1 was

increased in proportion to the SAB concentration (**Fig. 2-4** and **Fig. 2-5**). These results indicate that ion-dependent adsorption of CBM1 was observed even when the catalytic domain was included. However, no influence on the adsorption to MC was observed in the case of rEx-1. rEx-1 contains CBM1 at the C-terminus, and the amino acids that are essential for adsorption are identical to those of r*Th*EG1. This type of enzyme has a tunnel-shaped active site at the catalytic domain, and hydrolyzes crystalline cellulose in a processive manner [13-14]. I suggest that the binding of the cellulose chain to the catalytic domain might be much stronger than that of CBM1, or alternatively that the binding is not easily influenced by buffer concentration.

CBM$_{Th\text{EG1}}$-GFP, r*Th*EG1, and rEx-1 were adsorbed to MC at 30˚C and rapidly desorbed at 45˚C (**Fig. 2-7**). The differences in adsorption or desorption among these proteins might also be explained by differences in the affinity of the catalytic domain, hydrolytic activity against cellulose, and variations in the CBM1 amino acid sequence between r*Th*EG1 and rEx-1. In particular, r*Th*EG1 belonging to the GH family 5 endoglucanase preferentially hydrolyzes the amorphous region of cellulose, to which cellulases adsorb with high affinity. The loss of the amorphous region might have resulted in the low adsorption of r*Th*EG1 (~50% adsorption in **Fig. 2-5**) compared with that of CBM$_{Th\text{EG1}}$-GFP, even though they contain the same CBM1.

The enzymatic hydrolysis of insoluble cellulose is strongly influenced by the adsorption capacity of the enzyme [15]. The adsorption of CBM1 is easily controllable by changing the SAB concentration. This allows the relationship between adsorption and hydrolysis activity to be revealed. Buffer concentration might affect the ionization of functional groups of the catalytic domain and the formation of an electrostatic bond, as well as hydrophobic interactions. At extremely low buffer concentrations, the active site of the enzyme might not be arranged in the concentration needed for catalysis. At buffer concentrations over 10 mM, the catalytic domain was maintained in the optimal conformation. This was supported by the result that the hydrolysis of CMC was not influenced at all (**Fig. 2-8 (B)**). In the range of 10-100 mM, hydrolysis of insoluble substrate was highly dependent on adsorption (**Fig. 2-8 (A)**).

In chapter 2, we have demonstrated that the adsorption and desorption of CBM1 can be easily regulated by changing SAB concentration. This regulation can be easily combined with existing saccharification and enzyme recycling processes. The advantage of this is that it

avoids the carry-over of hydrolysates such as cello-oligosaccharides and salts contained in cellulosic materials into the next saccharification step. However, some improvements are required before this process can be put to use. The desorption conditions have to be considerably improved to recover larger quantities of enzymes. Additionally, for β-glucosidase, which converts cellobiose into glucose, the presence of CBM1 might be available for the recovery process.

Finally, hydrolytic activity is strongly influenced by SAB concentrations when fungal cellulase is assayed using insoluble cellulose as the substrate. For the determination of reducing sugar, the buffer concentration tends to be kept as low as possible, so as not to disturb color development. However, buffer concentration should be over 100 mM when acetate buffer is used, as otherwise CBM1 might not adsorb maximally.

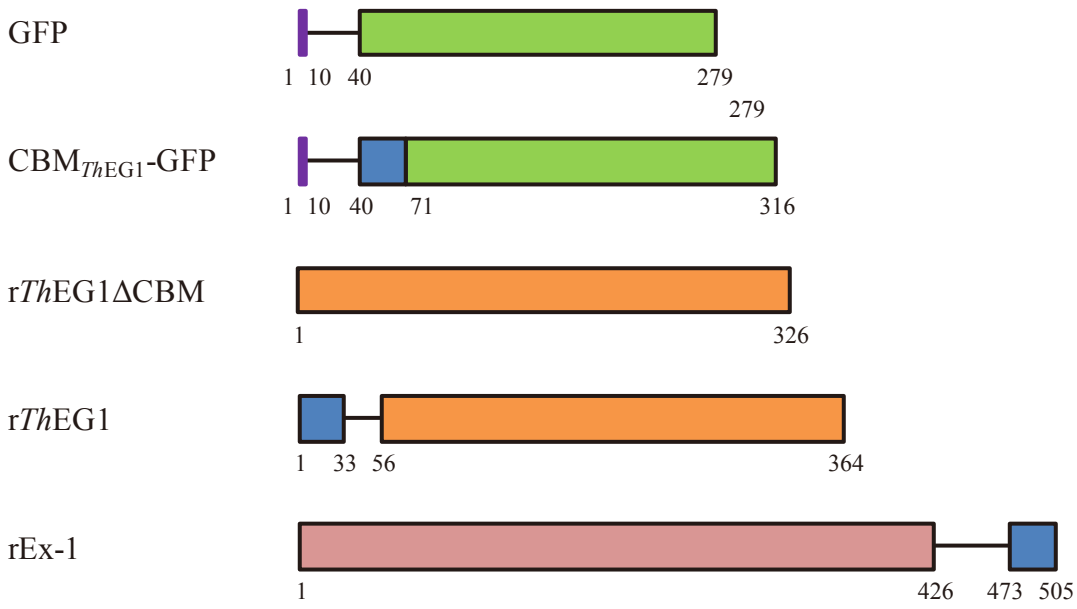


Fig. 2-1. Schematic diagrams of proteins used in the work described in this chapter. GFP and CBM$_{ThEG1}$-GFP were expressed in recombinant *E. coli*. r*Th*EG1*Δ*CBM, r*Th*EG1 (from *T. hirsuta*), and rEx-1 (from *I. lacteus*) were expressed in recombinant *A. oryzae.* The colors of each unit are as follows: purple, histidine tag; green, GFP; blue, CBM1; orange, GH family 5 catalytic domain; magenta, GH family 7 catalytic domain.

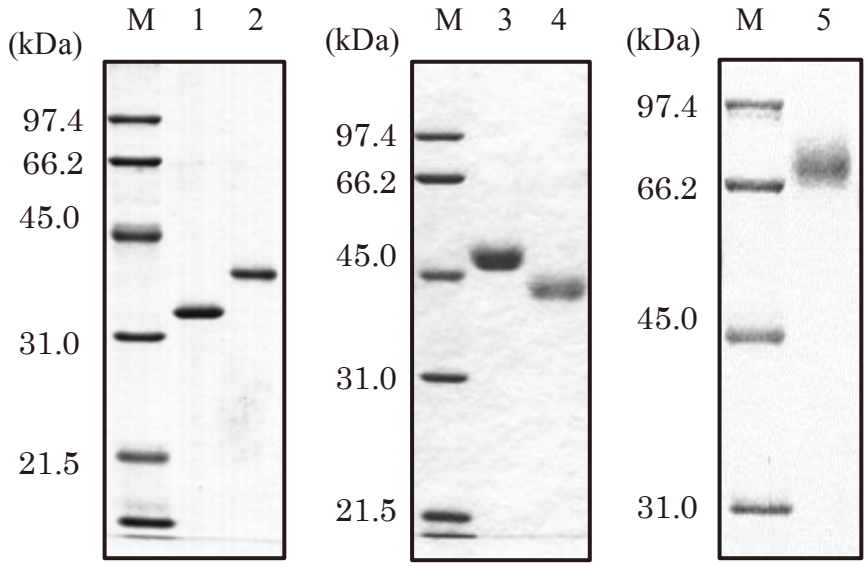


Fig. 2-2. SDS-PAGE of purified proteins used in the work described in this chapter.
M, molecular weight markers; lane 1, GFP; lane 2, CBM_{ThEG1}-GFP; lane 3, r*Th*EG1; lane 4, r*Th*EG1*Δ*CBM; lane 5, rEx-1.

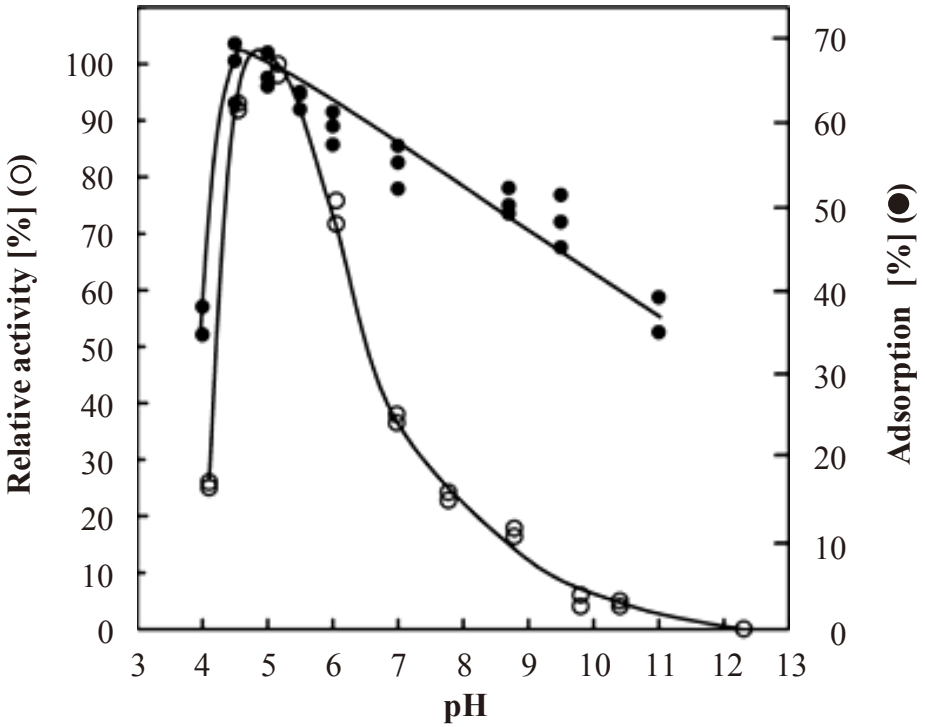


Fig. 2-3. Effects of pH on CBM$_{Th\text{EG1}}$-GFP adsorption and r*Th*EG1 activity.
●, CBM$_{Th\text{EG1}}$-GFP as a percentage of initial protein; ○, relative activity of r*Th*EG1 for MC.

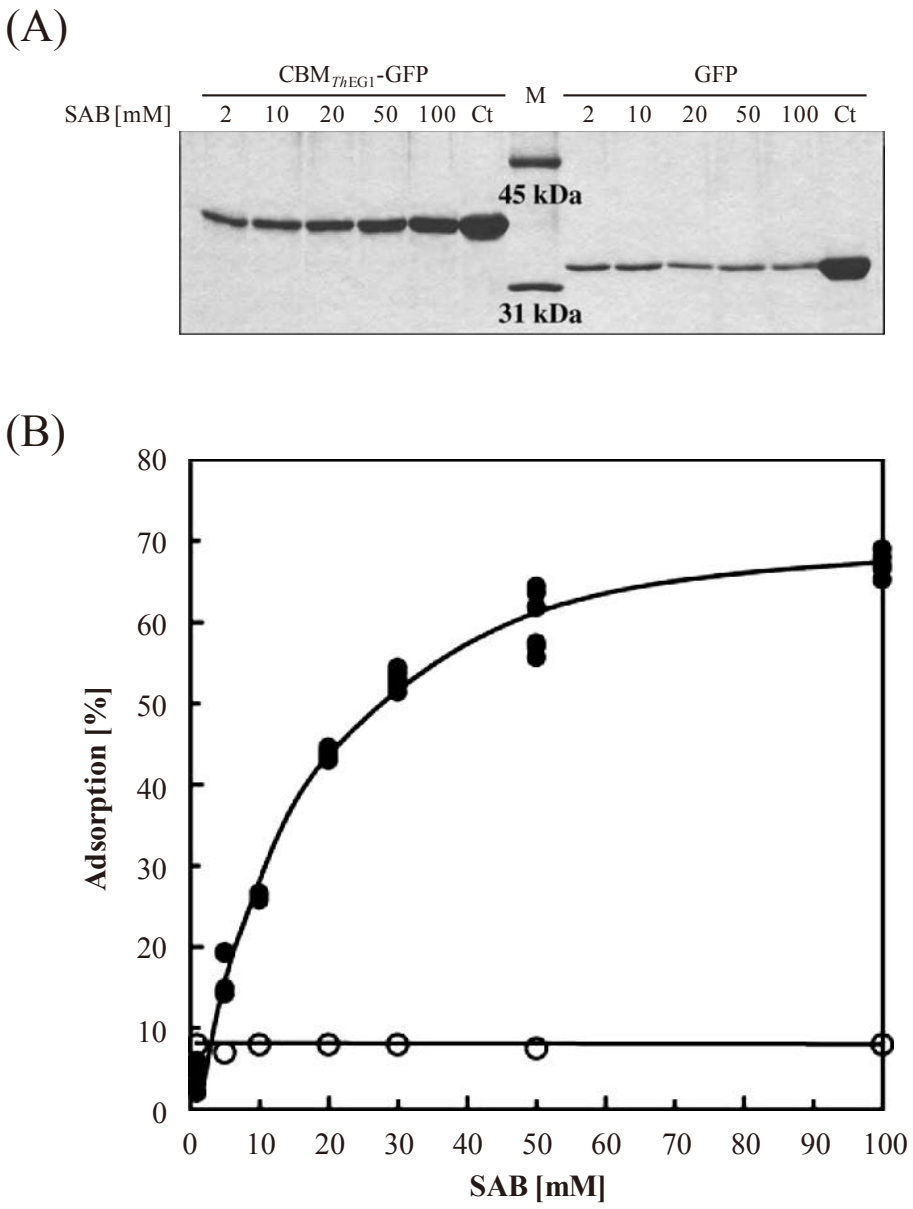


Fig. 2-4. Adsorption of CBM$_{ThEG1}$-GFP to MC at various buffer concentrations.
(A) Adsorbed proteins were subjected to SDS-PAGE. Ct, total amount of protein used in the experiment; M, molecular weight markers. (B) Percentage of adsorbed protein was quantified from the fluorescence data. ●, CBM$_{ThEG1}$-GFP; ○, GFP.

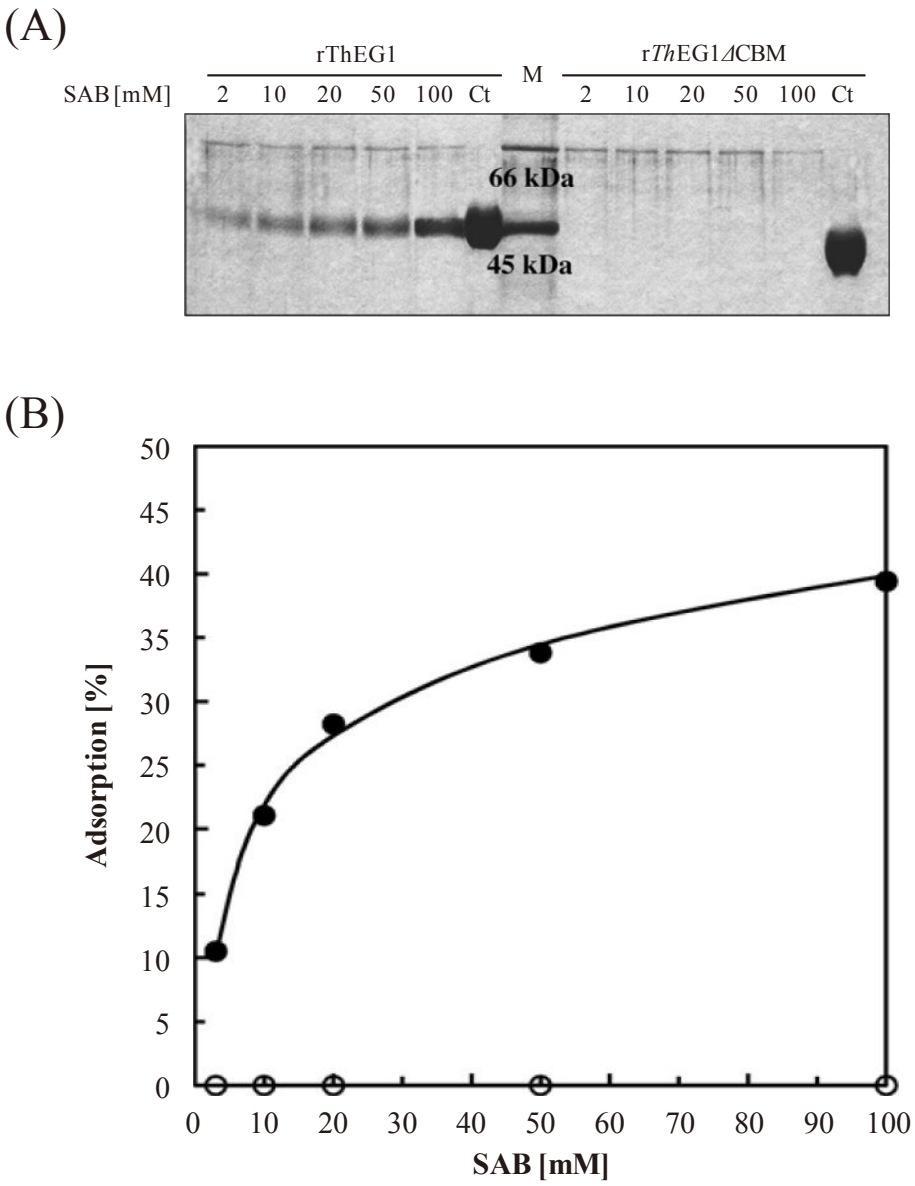


Fig. 2-5. Adsorption of r*Th*EG1 and r*Th*EG1*Δ*CBM.
(A) SDS-PAGE of the adsorbed proteins. Ct, total amount of protein used in the experiment; M, molecular weight markers. (B) Adsorption was estimated by measuring the intensity of the bands using ImageJ software. ●, r*Th*EG1; ○, r*Th*EG1*Δ*CBM.

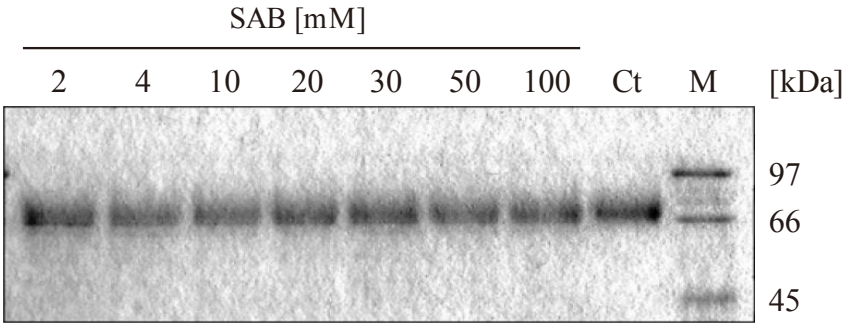


Fig. 2-6. Adsorption of rEx-1 to MC at various buffer concentrations.
Ct, total amount of protein used in the experiment; M, molecular weight markers.

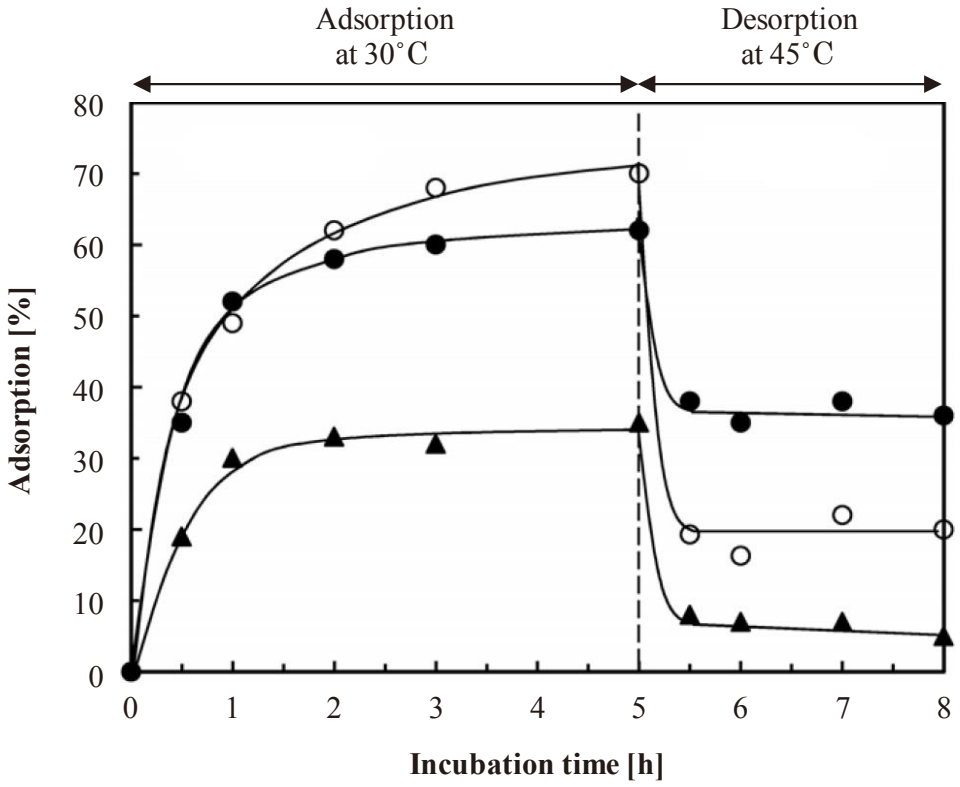


Fig. 2-7. Time course of adsorption and desorption of CBM$_{Th\text{EG1}}$-GFP, rEx-1, and r*Th*EG1 to MC.

CBM$_{Th\text{EG1}}$-GFP, rEx-1, and r*Th*EG1 were adsorbed to MC in the presence of 100 mM SAB (pH 5.5) at 30˚C. The protein was then desorbed with water at 45˚C. The vertical axis indicates the percentage of initial protein adsorbed at each time. ●, CBM$_{Th\text{EG1}}$-GFP; ○, rEx-1; ▲, r*Th*EG1.

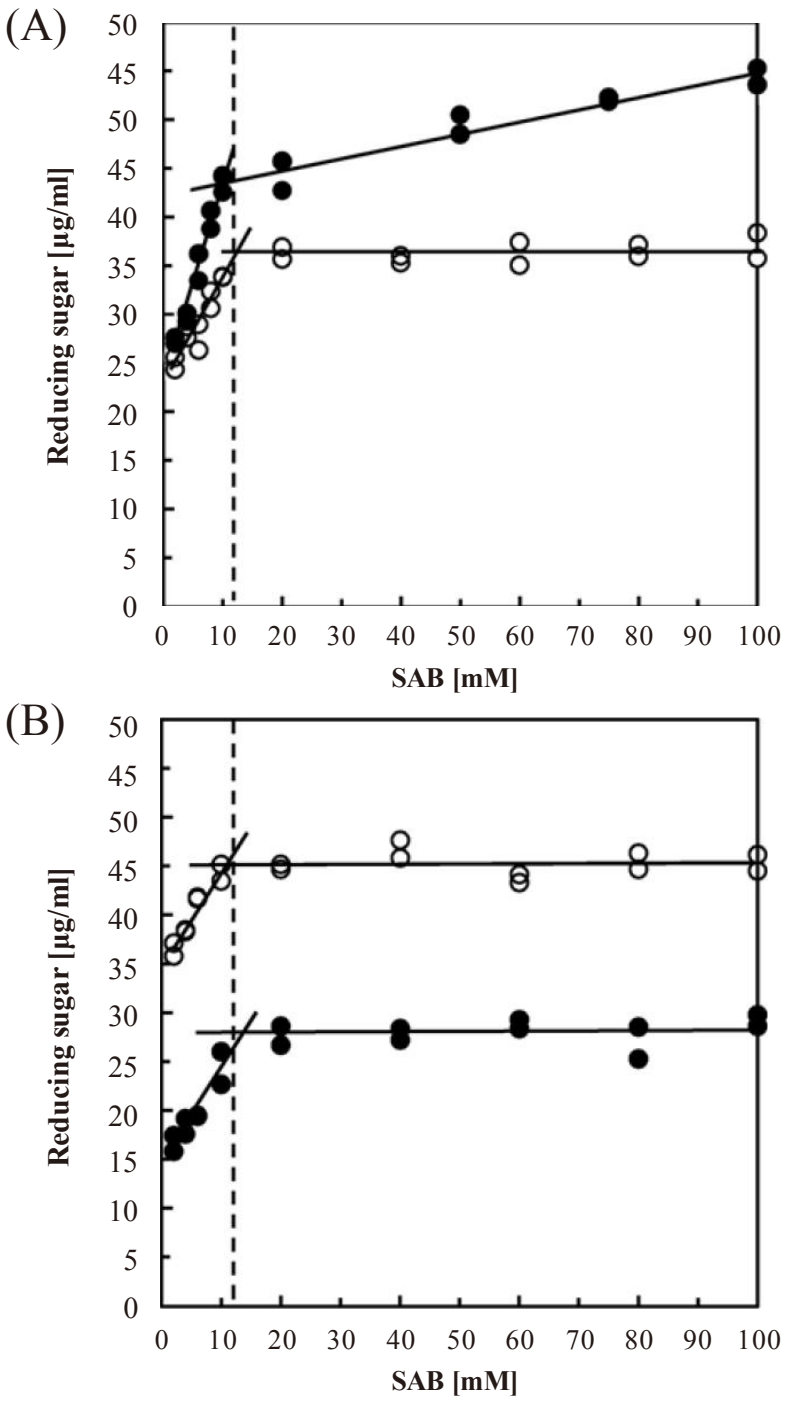


Fig. 2-8. Relationship between adsorption and hydrolysis.
Hydrolysis of MC (A) and CMC (B) was determined at various buffer concentrations. The dashed lines indicate the reflection points of the two lines. ●, r*Th*EG1; ○, r*Th*EG1⊿CBM.

References

1. **Henriksson G, Salumets A, Divne C, Pettersson G.** 1997. Studies of cellulose binding by cellobiose dehydrogenase and a comparison with cellobiohydrolase 1. Biochem. J. **324:**833-838.
2. **Linder M, Teeri TT.** 1996. The cellulose-binding domain of the major cellobiohydrolase of *Trichoderma reesei* exhibits true reversibility and a high exchange rate on crystalline cellulose. Proc. Natl. Acad. Sci. U S A. **93:**12251-12255.
3. **Salvi G, Rios PDL, Vendruscolo M.** 2005. Effective interactions between chaotropic agents and proteins. Proteins **61:**492-499.
4. **Déjugnat C1, Dufrêche JF, Zemb T.** 2011. Ion-specific weak adsorption of salts and water/octanol transfer free energy of a model amphiphilic hexapeptide. Phys. Chem. Chem. Phys. **13:**6914-6924.
5. **Cummins P, O'Connor B,.** 2011. Hydrophobic interaction chromatography. Methods in Molecular Biology **681:**431-437.
6. **Boom R, Sol CJA, Salimans MMM,Jansen CL,PMEWV Dillen,Noordaa JVD.** 1990. Rapid and simple method for purification of nucleic acids. J. Clin. Microbiol. **28:**495-503.
7. **Laemmli UK.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680-685.
8. **Nozaki K, Seki T, Matsui K, Mizuno M, Kanda T, Amano Y.** 2007. Structure and characteristics of an endo-β-1,4-glucanase, isolated from *Trametes hirsuta*, with high degradation to crystalline cellulose. Biosci. Biotechnol. Biochem. **71:**2375-2382.
9. **Toda H, Takada S, Amano Y, Kanda T, Okazaki M, Shimosaka M.** 2005. Expression of a cellobiohydrolase cDNA from the white rot fungus *Irpex lacteus* in *Aspergillus oryzae*. J. Appl. Glycosci. **52:**23-26.
10. **Somogyi M.** 1952. Note on sugar determination. J. Biol. Chem. **195:**19-23.
11. **Nelson N.** 1952. A photometric adaptation of somogyi method for the determination of glucose. J. Biol. Chem. **15:**375-380.
12. **Lindera M, Nevanena T, Teeri TT.** 1999. Design of a pH-dependent c ellulose-binding domain. FEBS. **447:**13-16.
13. **Divne C, Stahiberg J, Reinikainen T, Ruohonen L, Pettersson, Knowles JKC, Teeri TT, Jones TA.** 1994. The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. Science. **265:**524-528.
14. **Divne C, Stahiberg J , Teeri TT, Jones TA.** 1998. High-resolution crystal structures reveal how a cellulose chain is bound in the 50 AÊ long tunnel of cellobiohydrolase I from *Trichoderma reesei*. J. Mol. Biol. **275:**309-325.
15. **Takashima S, Ohno M, Hidaka M, Nakamura A, Masaki H, Uozumi T.** 2007. Correlation between cellulose binding and activity of cellulose-binding domain mutants of *Humicola grisea* cellobiohydrolase 1. FEBS letters. **581:**5891-5896.

Chapter 3

Screening of proteins strongly adsorbed on cellulose from Driselase and elucidation of their properties

Chapter 3: Screening of proteins strongly adsorbed on cellulose from Driselase and elucidation of their properties

3. Introduction

White-rot basidiomycetes are a group of fungi characterized by the ability to depolymerize and mineralize lignin in wood using a set of extracellular ligninolytic enzymes [1]. *Irpex lacteus*, a kind of white-rot fungus, is known as a strong cellulase and hemicellulase producer and used for the production of a commercial enzymatic preparation, Driselase. *I. lacteus* has also been found to be very efficient in lignocellulose degradation compared with other fungi, and it is expected that fungal biochemical pretreatment by *I. lacteus* would enhance the conversion of cellulosic biomass [2-3]. In our research, several cellulases and hemicellulases were purified from Driselase to study a biomass-degrading system of *I. lacteus*.

In our previous research, several cellulases and hemicellulases were purified from Driselase, and their enzymatic properties were studied for both basic research and application processes in a biomass decomposition system. To date, two CBH I-type cellulases (named Ex-1 and Ex-2) [4-5], one CBH II-type cellulase (Ex-4) [6], and EG-type cellulase (En-1) [7-8] have been purified from Driselase. These cellulases from *I. lacteus* MC-2 have also been heterologously expressed as recombinant enzymes and their enzymatic properties have been revealed. In addition to the cellulolytic enzymes, some xylanolytic enzymes like xylanase have been reported. In 1985, Amano and Kanda reported two endo-1,4-β-xylanases, designated xylanase I and III, purified from Driselase and their enzymatic properties [9-12]. Xylanase is one of the representative hemicellulases and plays important roles in biomass degradation in nature.

Upon the degradation of insoluble substrates like crystalline cellulose and also the plant cell wall, the adsorption of enzymes to the surface of the substrate is one of the key steps for efficient hydrolysis. As described in chapter 2, CBM1 plays very important roles in the adsorption and desorption behaviors of enzymes on insoluble substrates. Recently, draft genome sequence analysis of *I. lacteus* NK-1 was carried out (unpublished data) and showed that at least 12 cellulolytic and hemicellulolytic enzymes have CBM1 with different amino acid sequences (**Table 3-1**). Interestingly, non-cellulolytic enzymes, one β-glucosidase belonging to GH family 3 and three xylanases belonging to GH family 10, also have different

motifs of CBM1 in the N- or C-terminal region. Actually, many cellulose-binding proteins were also found in the culture broth of *I. lacteus* NK-1 (**Fig. 3-1**). As described in this chapter, proteins strongly adsorbed on microcrystalline cellulose were screened from Driselase, produced by *I. lacteus* MC-2, by using the cellulose affinity method developed as described in chapter 2.

3.1. Materials and methods

3.1.1. Chemicals and strains

Driselase was kindly provided by Kyowa Hakko Co. (Japan). Microcrystalline cellulose (MC) (column chromatography grade, Merck, Germany) was used as a carrier for fractionation and a substrate for enzymatic assay. Birchwood xylan (Sigma-Aldrich, USA), oatspelt xylan (Sigma-Aldrich, USA), and carboxymethyl cellulose (CMC) (Tokyo Kasei, Japan) were also used for enzymatic assays. Ion liquid-pretreated erianthus (hardy pampas grass), consisting of 49% cellulose, 28% hemicellulose, and 23% lignin, was provided by Hayakawa, and used as a biomass substrate. Pretreatment of erianthus was performed in 16.2 mL of 1-ethyl-3-methylimidazolium acetate (120˚C) with 0.5 g of erianthus at 90 min. Dissolved erianthus was precipitated in addition to 40 mL of deionized water. The regenerated erianthus was centrifuged at 3,000 rpm for 10 min and washed with deionized water, twice. Chromogenic substrates for enzymatic assay, such as *p*NP-β-D-glycopyranoside (*p*NPG1), *p*NP-β-D-cellobioside (*p*NPG2), *p*NP-β-D-lactopyranoside (*p*NPL), and *p*NP-β-D-xylopyranoside (*p*NPX1), were purchased from Sigma-Aldrich (USA). CBB-R250 (Wako, Japan) and AE-1360 EzStain Silver (Atto, Japan) were used for SDS-PAGE gel staining. PVDF membrane (GE Healthcare, UK) was used for the transfer of digested proteins after SDS-PAGE. DEAE Sephadex A-50 (Pharmacia Biotech AB, Sweden) was used for column chromatography. Silica Gel 60 A (Merck, German) was used for thin-layer chromatography (TLC). Xylose (X1) (Wako, Japan) and xylooligosaccharides (X2~X6) (Biocon, Japan) were used as substrates for TLC.

pBluescript II SK(+) (Stratagene, USA) was used for subcloning and sequencing of the Xyn10B cDNA fragment. *Escherichia coli* DH5α was used as plasmid host and grown in LB medium at 37˚C on a rotary shaker. Resistance to ampicillin (Amp^r) was selected by using the antibiotic at a concentration of 50 μg/mL. Expression vector pNAN-8142 and expression host *Aspergillus oryzae* $niaD^-$ were kindly provided by Ozeki Corporation (Japan). Expression

vector pET-23d and pCold I (Takara, Japan) and expression host *E. coli* Origami B strain (Novagen, NJ) were purchased. GH family 10 xylanase from *Trichoderma reesei* (named *Tr*XYN III, GenBank accession number is BAA89465) was kindly provided by the Ogasawara laboratory, Nagaoka University of Technology [13], and used for comparison.

3.1.2. Analytical methods

SDS-PAGE was performed by the method of Laemmli using 12.5% gel [14]. After electrophoresis, the gels were stained with 0.25% CBB-R250 (Wako, Japan). Low range protein marker (BIO-RAD, USA) was used as a molecular weight standard. The amount of protein was determined by the method of Lowry [15] using bovine serum albumin (BSA) as a standard.

3.1.3. Driselase fractionation

Ten grams of MC was added to 1,000 mL of 1% Driselase solution, and this solution was agitated overnight at 4˚C. The reaction mixture was filtrated and the residual MC was washed with 600 mL of cold 0.1 M SAB (pH 5.0). Solutions filtrated by the above step were collected as fraction 1. Proteins adsorbed on MC were eluted with 500 mL of cold water at 4˚C (fraction 2), and then eluted with the 1 mM SAB (pH 5.0) at 37˚C (fraction 3). Finally, the residual MC was used as fraction 4. Proteins remaining in fraction 4 were forcibly eluted with 2% SDS containing 0.06% 2-mercaptoethanol by boiling for 3 min for SDS-PGE. A flow chart of the cellulose affinity fractionation of Driselase is shown in **Fig. 3-2**.

3.1.4. Enzymatic assay

In the cellulose affinity fractionation described in 3.1.3, fractions 1 to 3 were directly used for enzyme activity measurement. Fraction 4 was suspended in 40 mL of 1 mM SAB (pH5.0) and the suspension was used as the enzyme solution. Purified Xyn10B from Driselase and recombinant *Tr*XYN III were used in solutions at a concentration of 10 μg/mL. These enzymatic solutions were incubated at 30˚C with various substrates, which were 0.125% birchwood xylan, 0.125% oatspelt xylan, 0.125% CMC, and 0.125% ion liquid-pretreated erianthus contained in 25 mM SAB (pH 5.0). The amount of reducing sugar produced was measured by the method of Somogyi-Nelson using D-glucose or D-xylose as a standard [16-17]. The enzyme activity on a chromogenic substrate was measured as follows.

The enzyme solutions were incubated at 30˚C with 1.25 mM *p*NPL, *p*NPG1, *p*NPG2, and *p*NPX1 contained in 25 mM SAB (pH 5.0). The reaction was stopped by the addition of 2.0 mL of 1% sodium carbonate solution and then the amount of liberated *p*-nitrophenol was estimated calorimetrically at 420 nm. All activity units are defined as the amount of enzyme that released 1 μmol product per min.

3.1.5. Amino acid sequencing of 42 kDa protein

Fraction 4 in cellulose affinity fractionation was subjected to SDS-PAGE and stained by AE-1310 EzStaine Reverse (ATTO, Japan). Then, proteins were transferred to a PVDF membrane (GE Healthcare, UK). The piece of PVDF membrane containing the electroblotted 42 kDa protein was subjected to a PPSQ-31 protein sequencer (Shimadzu, Japan) to determine the N-terminal amino acid sequence of the native 42 kDa protein. In order to determine the internal amino acid sequence of the 42 kDa protein, this electroblotted protein was extracted from the PVDF membrane and incubated with 0.25 μg of V8 protease (Wako, Japan) contained in 50 mM sodium phosphate buffer (pH7.8) at 37˚C for 16 hours. After that, the digested protein was applied to SDS-PAGE and stained by AE-1310 EzStaine Reverse. The major fragments corresponding to sizes of 10 and 22 kDa were also subjected to the protein sequencer by the same method as described above. The resulting sequences were collated to the database of the *I. lacteus* NK-1 draft genome sequence.

3.1.6. Cloning of cDNA encoding 42 kDa protein (Xyn10B) from *I. lacteus* MC-2

The strategy for the cDNA cloning of Xyn10B is shown in **Fig. 3-3 (A)**. Full-length cDNA of Xyn10B was amplified by PCR from first-strand cDNA of *I. lacteus* using Prime STAR HS DNA polymerase (Takara, Japan) with the sense primer (5'-AAGCTTAAGCTCAACCTCACAAGA-3', *Hin*dIII site is underlined) and reverse primer (5'-TCTAGATTACGAGAGAGAGCCTGCATGA-3', *Xba*I site is underlined) as per the following PCR conditions.

(PCR conditions: 95˚C for 5 min, followed by 30 cycles at 98˚C for 10 s, 49˚C for 15 s, and 72˚C for 90 s, followed by a final extension step at 72˚C for 5 min.) The amplified fragment (1,181 bp, **Fig. 3-3 (B)**) was ligated to the *Sma*I site of pBluescript II SK (+) (named pBS-Xyn10B). The DNA sequence of the amplified fragment was determined by Hokkaido System Science Co., Ltd. (Japan).

3.1.7. Construction and expression of the recombinant Xyn10B in *Aspergillus oryzae*

Construction of a recombinant Xyn10B expression system in *A. oryzae* was attempted. For secreted protein expression, two signal peptide sequences were used. One is the native signal peptide sequence of Xyn10B from *I. lacteus* MC-2, and the other is the Taka-amylase (TAA) signal peptide sequence (Genbank accession number AAA32708) from *A. oryzae*. The construction maps of the recombinant Xyn10B expression vectors for *A. oryzae* are shown in **Fig. 3-4 and Fig. 3-5.** Briefly, a plasmid pBS-Xyn10B was digested by *Hin*dIII and *Xba*I, and Xyn10B fragment was inserted into the same sites of pNAN-8142. The resulting plasmid was named pNAN-Xyn10B (**Fig. 3-4**).

To construct the expression vector using TAA signal peptide, signal-lacking cDNA of Xyn10B (Xyn10B*Δ*Sig) was amplified by PCR from first-strand cDNA of *I. lacteus* MC-2 using PrimeSTAR HS with the sense primer (5'-CCCCGGGGGCAGTGGCCTCGATG-3', *Sma*I site is underlined) and reverse primer (5'-TCTAGATTACGAGAGAGAGCCTGCATGA-3', the same primer as described above, *Xba*I site is underlined) as per the following PCR conditions.

(PCR conditions: 95˚C for 5 min, followed by 30 cycles at 98˚C for 10 sec, 49˚C for 15 sec, and 72˚C for 80 s, followed by a final extension step at 72˚C for 5 min.) PCR product was bluntly ligated into the *Sma*I site of pBluescript II SK+ and the resulting plasmid was named pBS-Xyn10B*Δ*Sig. TAKA-amylase A signal peptide gene (TAASig) (AAGCTTTTATGATGGTCGCGTGGTGGTCTCTATTTCTGTACGGCCTTCAGGTCGCG GCACCTGCTTTGGCTGCAACGCCCGGG, start codon is boxed, *Hin*dIII and *Sma*I sites are underlined) cloned pBluescript II SK+ (named pBS-TAASig) vector was provided by Dr. Nozaki. pBS-Xyn10B*Δ*Sig was then digested by *Sma*I and *Xba*I, and Xyn10B*Δ*Sig fragment was inserted into the same sites of pBS-TAASig. The resulting plasmid named pBS-TAASig-Xyn10B was then digested at *Hin*dIII and *Xba*I sites and the TAASig fused Xyn10B fragment was inserted into the same site of pNAN-8142. The resulting plasmid was named pNAN-TAASig-Xyn10B (**Fig. 3-5**).

pNAN-Xyn10B and pNAN-TAASig-Xyn10B were used as *A. oryzae* expression vectors. Methods of transforming *A. oryzae* were performed according to Gomi *et al.* [18]. The *A. oryzae* transformants were grown in SPY medium containg 2% starch, 1% polypeptone, 0.5% yeast extract, 0.5% KH_2PO_4, and 0.05% $MgSO_4$ for 3 days at 30˚C with rotary shaking at 120 rpm. The supernatant was used for enzymatic assay and SDS-PAGE for

the examination of recombinant Xyn10B expression.

3.1.8. Construction and expression of the recombinant Xyn10B in *Escherichia coli*

The construction of a recombinant Xyn10B expression system in *E. coli* was attempted using two expression vectors. One was pET-23d vector (Takara, Japan) and the other was pCold I (Takara) vector. For construction of the pET-23d expression system, pET-bgl-base vector based on pET-23d was kindly provided by Dr. Nozaki. pBS-Xyn10B*Δ*Sig was digested at *Sma*I and *Xba*I sites, and Xyn10B*Δ*Sig fragment was inserted into *Eco*RV and *Spe*I sites of the pET-bgl-base vector (named pET-Xyn10B) **(Fig. 3-6)**.

For construction of the pCold I vector expression system, signal-lacking and *Hin*dIII-site-added cDNA of Xyn10B (Xyn10B*Δ*Sig$_{HindIII}$) was amplified by PCR from pBS-Xyn10B using PrimeSTAR HS DNA polymerase with the sense primer (5'-AAGCTTCAAGTGGAAGCTGTCGCTG-3', *Hin*dIII site is underlined) and reverse primer (5'-TCTAGATTACGAGAGAGAGCCTGCATGA-3', the same primer as described above, *Xba*I site is underlined) as per the following PCR conditions.

(PCR conditions: 95˚C for 3 min, followed by 30 cycles at 98˚C for 10 sec, 58˚C for 10 sec, and 72˚C for 88 s, followed by a final extension step at 72˚C for 2 min.) PCR product (Xyn10B*Δ*Sig$_{HindIII}$) was bluntly ligated into the *Eco*RV site of pBluescript II SK+ and the resulting plasmid was named pBS-Xyn10B*Δ*Sig$_{HindIII}$. The pBS-Xyn10B*Δ*Sig$_{HindIII}$ was digested by *Hin*dIII and *Xba*I, and Xyn10B*Δ*Sig$_{HindIII}$ fragment was inserted into the same sites of pCold I. The resulting plasmid was named pCold-Xyn10B **(Fig. 3-7)**.

pET-Xyn10B and pCold-Xyn10B were used as *E. coli* expression vectors. The cultivation method of *E. coli* Origami B strain transformed by pET-Xyn10B was the same as for CBM$_{ThEG1}$-GFP described in chapter 2. *E. coli* transformed by pCold-Xyn10B was grown in 100 mL of Luria-Bertani medium containing 100 μg/mL ampicillin at 37˚C with rotary shaking at 180 rpm. When optical density (OD) 600 reached 0.4~0.6, the culture was cooled in a water bath at 15˚C for 30 min and isopropyl β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. The culture was further incubated overnight at 15˚C, and then cells were harvested by centrifugation and suspended in 20 mM sodium phosphate buffer (pH 7.4) containing 0.5 M NaCl. Cells were disrupted by ultrasonic treatment and supernatants were recovered by centrifugation. The expression of recombinant Xyn10B in each *E. coli* was judged by enzymatic assay and SDS-PAGE analysis using the supernatants.

3.1.9. Purification of Xyn10B from Driselase

Driselase (10 g) was dissolved in 200 mL of 20 mM ammonium acetate buffer (AAB) (pH 5.0). Insoluble materials contained in dissolved Driselase solution were removed by centrifugation at 8,000 rpm for 15 min. The supernatant was fractionated using 40% saturated ammonium sulfate, and the resulting precipitate was dissolved in 500 mL of 20 mM AAB (pH 5.0). The solution was applied to a DEAE-Sephadex A-50 column (5.0 × 47 cm) equilibrated with the same buffer at 4˚C. Elution of the proteins was also performed with 20 mM AAB (pH 5.0) and recovered in a test tube as each fraction (10 mL/tube). Each fraction was applied to SDS-PAGE for confirmation of purity.

3.1.10. Enzymatic hydrolysis and thin-layer chromatography (TLC) analysis of reaction products

Purified Xyn10B was incubated at 30˚C with 0.125% birchwood xylan or 0.2 mg/mL xylooligosaccharides (X1 to X6) containing 25 mM SAB (pH 5.0). The reaction mixture was spotted on Silica-gel 60 (Merck, Germany) and well dried, and then developed two times with a solvent system of isopropanol/butanol/water (60/15/12, vol/vol/vol) at 30˚C. Sugar spots of the TLC plate were detected by heating the plate at 120˚C for 5 min after the spraying of 30% sulfuric acid.

3.2 Results

3.2.1. Component analysis of four fractions from Driselase

To screen the proteins that absorb to MC, the developed method described in chapter 2 was adapted to the commercial enzymatic preparation Driselase. Four fractions obtained by cellulose affinity fractionation were analyzed by SDS-PAGE (**Fig. 3-8**). From the previous studies about each enzyme contained in Driselase, it is possible to some extent to assign the band on the SDS-PAGE based on the molecular size. Most of the endoglucanase I and CBH II (designated as En-1 and Ex-4, 35 and 39 kDa, respectively) were detected in fraction 1, without adsorption to MC. Although both enzymes have CBM1 at the N-terminal, both enzymes were mostly removed by proteolysis at the linker region connecting CBM1 and the catalytic domain [6, 8]. As a result, it was thought that En-1 and Ex-4 flowed through without adsorption to MC. On the other hand, CBH I a (designated as Ex-1, 53 kDa) and CBH I b

(designated as Ex-2, 55 kDa), which seem to be isoforms with different levels of glycosylation, bound to the MC. Both types of CBH I have CBM1 in their C-terminal regions [5]. Fraction 4 was the residual MC containing some strongly adsorbed proteins. Interestingly, it was not eluted with 1 M NaCl, 1% (v/v) Triton X-100, or 0.06% 2-mercaptoethanol (data not shown). After boiling fraction 4 with 2% SDS and 0.06% 2-mercaptoethanol, some proteins eluted from MC were detected by SDS-PAGE. The majority of the proteins with a mass of 42 kDa remained bound to the cellulose without elution.

3.2.2. Enzymatic activity of four fractions on various substrates

Specific activities of each fraction on various cellulosic and xylanolytic substrates are shown in **Table 3-2**. Compared with the specific activity of crude Driselase on MC, fraction 1 showed lower activity, but fraction 2 and fraction 3 showed high activities. Despite strong adsorption on cellulose, fraction 4 containing the 42 kDa protein showed significant xylan degradation activity (0.217 units/mL), which was 2.4 times higher than the CMC degradation activity (0.0894 units/mL). It corresponded to about 0.4% and 0.04% xylanase and CMC-degrading activity of total activity including the start enzyme solution. The *p*NPX1 degradation activities of fraction 2 and fraction 3 were three times higher than that in crude Driselase. The *p*NPG2 and *p*NPX1 degradation activities were maintained (about 15%) in fraction 4 more than the *p*NPG1 and *p*NPL degradation activities (about 5%).

3.2.3. Amino acid sequencing of 42 kDa protein fragments digested by protease

To identify the 42 kDa protein, the N-terminal amino acids of the 42 kDa protein without peptidase digestion were firstly sequenced. As a result, the N-terminal sequence was determined to be Val-Ala-Glu-Trp-Gly-Gln-X-Gly-Gly-Ile-Gly-Phe-Thr (X is unidentified). Subsequently, to determine the internal amino acid sequence of the 42 kDa protein, protease digestions using trypsin and V8 protease were carried out. Although no fragments appeared after trypsin treatment, five fragments (V8-I to V8-V) were obtained by digestion using V8 protease (**Fig. 3-9**). As shown by N-terminal sequencing of all five fragments, three fragments (V8-I, V8-II, and V8-V) were coincident with the N-terminal sequence of the native 42 kDa protein already determined as described above. Two fragments of 10 kDa (V8-V) and 22 kDa (V8-III) were determined to be Asn-Ser-Met-Lys-Trp-Asp-Ala-Thr-Glu-Ans-Thr-Arg-Gly-Gln-Phe-Thr and

Ile-Phe-Asn-Glu-Asp-Gly-Thr-Phe-Arg-Ser-Ser-Val-Phe-Tyr-Asn-Val-Leu, respectively. These three sequences of the N-terminal and internal region of the 42 kDa protein were collated to the database of the *I. lacteus* NK-1 draft genome sequence and shown to be completely identical to the partial sequences of endo-1,4-β-xylanase belonging to GH family 10. This result agrees with the high xylan degradation activity in fraction 4 of Driselase described above (**Table 3-2**). Thus, the 42 kDa protein was designated as Xyn10B.

3.2.4. Cloning of cDNA encoding Xyn10B from *I. lacteus* MC-2 and heterologous expression

It is necessary to elute Xyn10B from MC for the further elucidation of protein function. However, it is too difficult to obtain Xyn10B from residual MC in the non-denatured state. Thus, an attempt was made to express Xyn10B from *I. lacteus* MC-2 heterologously. First, the cDNA encoding Xyn10B was amplified by PCR using primers designed based on the gene encoding Xyn10B of *I. lacteus* NK-1. The successfully amplified PCR product (1,181 bp) was cloned into plasmid pBluescript SK and then sequenced. The cDNA sequence of Xyn10B was composed of 1,152 bp and the amino acid sequence deduced from it was composed of 383 amino acid residues (**Fig. 3-10**). The three peptide fragments of the 42 kDa protein sequenced by peptide sequencing were contained in the deduced amino acid sequence and completely coincident. Thus, the 42 kDa protein was genetically identified as Xyn10B from *I. lacteus* MC-2.

The primary structure of Xyn10B is composed of three units. From the result of N-terminal sequencing of the native 42 kDa protein, it was suggested that the first 20 amino acid residues (Met 1 to Ala 20) form a signal sequence. The regions of CBM1 and the catalytic domain are from Val 21 to Leu 55 and from Leu 88 to Ser 383, respectively. The actual molecular mass (42 kDa) determined by SDS-PAGE was about 4 kDa higher than that calculated from the deduced amino acid sequence (38 kDa). Because there are no possible sites modified by *N*-glycosylation, it is thought that this difference was caused by *O*-glycosylation in the linker region. A Blast search revealed that the sequence of the Xyn10B core domain showed high identity with those of GH10 xylanases, such as *Phanerochaete carnosa* HHB-10118-sp xylanase (83%, Genbank accession no. EKM58934), *Thermoascus aurantiacus* xylanase (63%, PDB no. 2BNJ_A), and *T. reesei* xylanase III (56%, BAA89465.2) (**Fig. 3-11**). Among these, Xyn10B and *P. carnosa* xylanase have CBM1. It

was found that *I. lacteus* has two other GH10 xylanases (Xyn10A and Xyn10C), as revealed by the draft genome sequence (**Fig. 3-12**). Their catalytic domains showed low homology (42% and 41% compared with Xyn10B, respectively) and also contained CBM1 at the N-terminal.

3.2.5. Expression of recombinant Xyn10B

To elucidate the properties of Xyn10B from *I. lacteus* MC-2, an attempt was made to express the whole length of Xyn10B heterologously using *E. coli* and *A. oryzae* as the host organism. In the *Aspergillus* expression system, two signal peptide sequences were used for secretion. However, recombinant Xyn10B was not detected in the culture supernatant of both *A. oryzae* transformants (**Fig. 3-13 (A), (B)**). Xyn10B was also not detected in *E. coli* with the pET vector expression system (**Fig. 3-13 (B)**). On the other hand, Xyn10B was expressed as an insoluble body using the pCold vector expression system in *E. coli* (**Fig. 3-13 (B)**). Various conditions, such as variations of IPTG concentration, co-expression by chaperone plasmid set, and host strain, were tested for the expression into the soluble fraction; however, no expression of soluble recombinant Xyn10B in the soluble fraction was found. Although several refolding methods were also attempted for the insoluble Xyn10B, Xyn10B showing enzymatic activity was not obtained as a result. Since Xyn10B is a multi-domain structured protein composed of CBM1 and catalytic domain and the flexibility of the whole structure of Xyn10B would be large, it seems to be difficult to express Xyn10B as a properly folding structure. Therefore, Xyn10B would be purified from Driselase for elucidation of its enzymatic properties.

3.2.6. Purification of Xyn10B from Driselase

A summary of the purification of Xyn10B from crude Driselase is shown in **Table 3-3**. Most of the Xyn10B in Driselase was precipitated using a 40% concentration of ammonium sulfate (data not shown). A DEAE Sephadex A-50 column chromatogram of Driselase is shown in Fig. 3-14 (A). In terms of the hydrolysis activity for birchwood xylan, fractions were separated into two peaks (Peak 1: fractions 40-80, Peak 2: fractions 81-105, **Fig. 3-14 (A)**). These peaks were subjected to SDS-PAGE analysis (**Fig. 3-14 (B)**). As a result, almost all Xyn10B (42 kDa protein) appeared in Peak 1. Fraction 50 was used as purified Xyn10B for the following experiments, since Xyn10B was purified as almost a single band in fraction 50.

3.2.7. Xyn10B hydrolysis product from xylan

The TLC analysis of hydrolysis products from birchwood xylan by Xyn10B is shown in **Fig. 3-15**. Xyn10B produced X2 and X3 at the early stage of hydrolysis. On the other hand, the amount of X3 was decreased and small amounts of X1 were produced after the stage of action, and finally X2 was observed as the main product at the end of hydrolysis. In the case of xylooligosaccharide degradation, X2 was produced mainly from X3, X4, X5, and X6 for a 24 hour reaction, although a small amount of xylose was produced (**Fig. 3-16 (A)**). The time course of hydrolysis products from X5 by Xyn10B is shown in **Fig. 3-16 (B)**. X2~X4 were produced by Xyn10B at the early stage of hydrolysis. In particular, a small amount of X4 was produced but X1 was not produced at the first stage. The final products from birchwood xylan were the same as those from xylooligosaccharides, although oligosaccharide that had higher DP was detected. This is thought to be heterooligosaccharide, as birchwood xylan contained glucuronic residues as a side chain. The modes of action mentioned above coincided with those of typical family 10 xylanases. From these results, Xyn10B shows endo-type activity for xylan, which is similar to *Tr*XYN III and Xyn3 from *A. oryzae* [19].

3.2.8. Comparison of hydrolysis behavior between Xyn10B and *Tr*XYNIII

GH10 xylanase (*Tr*XYN III) from *T. reesei* is only composed of a catalytic domain, and the homology of the catalytic domain between Xyn10B and *Tr*XYN III is very high. Here, to confirm the effect of CBM1 in xylanase, Xyn10B and TrXYNIII were reacted to pure xylans such as birchwood and oatspelt xylans and biomass treated with ionic liquid.

First, the specific activities of Xyn10B on birchwood xylan and oatspelt xylan were 24.3 U/mg and 8.14 U/mg, respectively. Those of *Tr*XYN III were about 1.5-fold higher than those of Xyn10B (32.9 U/mg and 12.1 U/mg, respectively), but the differences were not so large. The time courses of reducing sugar production from both xylans are shown in **Fig. 3-17 (A)** and **(B)**. The velocity of *Tr*XYNIII was slightly higher than that of Xyn10B at the initial stage of the reaction (~2 hours). It was also revealed that CBM1 of Xyn10B showed no ability to adsorb to insoluble xylan. Thus, it is thought that the effect of CBM1 for hydrolysis of both soluble and insoluble xylans was extremely low.

On the other hand, the production of reducing sugar from biomass treated with ionic liquid after a 24 hour reaction by Xyn10B was 3.3-fold higher than that of *Tr*XYN III (**Fig.**

3-17 (C)). The levels of reducing sugars produced at the end of hydrolysis of Xyn10B and *Tr*XYN III were 2.44 μmol/mL and 0.74 μmol/mL, respectively. The erianthus treated with ionic liquid contained 28% hemicelluloses and 49% cellulose. In this case, CBM1 might enable the catalytic domain of Xyn10B to attack the xylan chain efficiently by adsorption to cellulose around the xylan.

3.3. Discussion

As described in this chapter, an interesting protein that strongly adsorbed to MC without desorption was detected using a cellulose affinity column. Furthermore, this protein was identified as an endo-1,4-β-xylanase by N-terminal amino acid sequencing and cDNA cloning, and designated as Xyn10B.

Xyn10B was shown to be composed of two domains, namely, CBM1 at the N-terminal and a catalytic domain at the C-terminal, which supports the adsorption ability of Xyn10B to MC. Draft genome sequence analysis has already been carried out using *I. lacteus* NK-1 (unpublished data). The homologous sequences to CBM1 from ## were found from a total of 12 enzymes, including not only cellulose, but also xylanase, β-glucosidase, and xyloglucanase (**Table 3-1**). Although xylanases from *T. reesei* and *Thermoascus aurantiacus* have no CBM1, it is interesting that *Phanerochaete chrysosporium* xylanase (XynA, GenBank ID: AAG44993.1) has CBM1 at the N-terminal. As just described, CBM1 is comparatively distributed in hemicellulases and non-hydrolytic proteins, as well as most cellulases among various fungi.

In general, CBM1 does not adsorb to hemicelluloses such as xylan, mannan, and other amorphous polysaccharides. Actually, CBM1 of Xyn10B shows no adsorption ability to insoluble xylans such as oatspelt xylan. The residual MC called fraction 4 in cellulose affinity fractionation showed hydrolytic activities for birchwood and oatspelt xylans (**Table 3-2**). This means that the adsorption ability to cellulose and the hydrolytic activity of Xyn10B function independently. In the degradation of erianthus treated with ionic liquid, Xyn10B showed higher activity than *Tr*XYNIII, while the hydrolytic activities for pure xylan were not so different between Xyn10B and *Tr*XYNIII (**Fig. 3-17**). It was reported that CBM 2 of xylanases from *Pseudomonas fluorescens* (XYLA and XYLC) could increase the degradation rate of plant cell wall [20] because XYLA and XYLC attached to cellulose via CBM2, and the linker region enabled the catalytic domain to access multiple hemicellulose chains. It was

also reported that CflXyn11A from *Cellulomonas flavigena*, which also contains CBM2, showed a high synergistic effect with *T. reesei* EG I (Cel7B) in the initial stage of sugar cane bagasse degradation [21]. On the other hand, only the catalytic domain of CflXyn11A showed lower levels of synergy and hydrolytic activity. It was discussed that the cellulose-binding domain of xylanase may be an advantage for the removal of xylan found in the close vicinity of cellulose, and make cellulase more accessible on cellulose. Moreover, it was reported that not only xylan degradation but also the degradation of other hemicelluloses in biomass was increased by a cellulose-binding domain. CBM1 of Man5A could increase mannan/cellulose complex degradation. As just described, the existence of CBM1 in hemicellulases and non-hydrolytic proteins except cellulase might contribute to the efficient degradation of the plant cell wall in a native environment [22].

Sugimoto *et al.* reported that more than 80% of cellulases with CBM1s were eluted from a cellulose affinity column at room temperature [23]. However, Xyn10B was not eluted from MC by changing the temperature and concentration of salt after adsorption. It has also been reported that several cellulose-binding modules irreversibly adsorb onto cellulose without desorption. For example, CBM1 of CBH II from *T. reesei* has an additional S-S bond, which may stabilize the steric structure, resulting in the high capacity for adsorption on cellulose [24]. In a second case, CBM2a of xylanase 10A from *Cellulomonas fimi* has a more planar binding face and the thermodynamic driving force is dominated by entropic effects, notably dehydration of polar surface residues [25-27]. However, it seemed that the mode of action of those irreversible CBMs from *T. reesei* or *C. fimi* differs from Xyn10B adsorption to cellulose because it was completely immobilized on the cellulose surface.

Table 3-1. Alignment of the primary structure of CBM1 from *Irpex lacteus*

GH Family	Gene	Enzyme	Strain [a]	Location	Primary structure of CBM1
3	-	-	NK-1	N	SGQ**Y**GQCGGIGWTGATTCISGWTCTLLN----------D**YY**SQCL
5	-	-	NK-1	N	AQE**W**QQCGGIGWSGATSCVSGSVCTELN----------S**YY**SQCI
5	*cen1*	En-1	MC-2	N	VSV**W**GQCGGIGFTGSTTCDAGTSCVHLN----------D**YY**FQCQ
6	*cex4*	Ex-4	MC-2	N	AQT**W**AQCGGIGFTGPTTCVAGSVCTKQN----------D**YY**SQCI
7	*cel1*	-	MC-2	C	VAQ**W**GQCGGTGFTGPTVCASPFTCHVVN----------P**YY**SQCY
7	*cel2*	Ex-1, Ex-2	MC-2	C	VAQ**W**AQCGGIGYSGATTCVSPYTCHVVN----------A**YY**SQCY
7	*cel3*	-	MC-2	C	AAQ**W**AQCGGMGFTGPTVCASPFTCHVLN----------P**YY**SQCY
10	*xyn10a*	Xyn10A	NK-1	N	SQV**W**GQCGGIGWTGATTCVSASSCVKSN----------D**YY**SQCI
10	*xyn10b*	Xyn10B	MC-2	N	VAE**W**GQCGGIGFTGSTTCDSPFVCTVIN----------S**YYY**QCL
10	*xyn10c*	Xyn10C	NK-1	N	SQV**W**GQCGGEGWTGATTCVSGSTCVAQN----------Q**WY**SQCL
74	-	-	NK-1	C	QSY**Y**AQCGGTNYIGPTICADGSTCVAQNACKPFISSSEFRQ**YY**----
LPMO [b]	-	-	NK-1	C	VQK**Y**GQCGGQGYSGASICASGTTCTALN----------D**YY**SQCL

Gray boxes indicate conserved amino acid residues. Three aromatic residues located on the flat surface of CBM1 are shown in bold.
a: NK-1, *I. lacteus* strain NK-1 draft genome sequence; MC-2, *I. lacteus* strain MC-2 cDNA sequence
b: Copper-dependent lytic polysaccharide monooxygenase

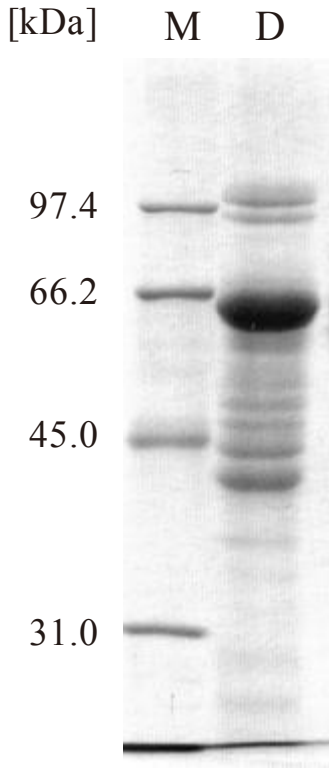


Fig. 3-1. SDS-PAGE of proteins adsorbed to microcrystalline cellulose from culture broth of *Irpex lacteus* NK-1.

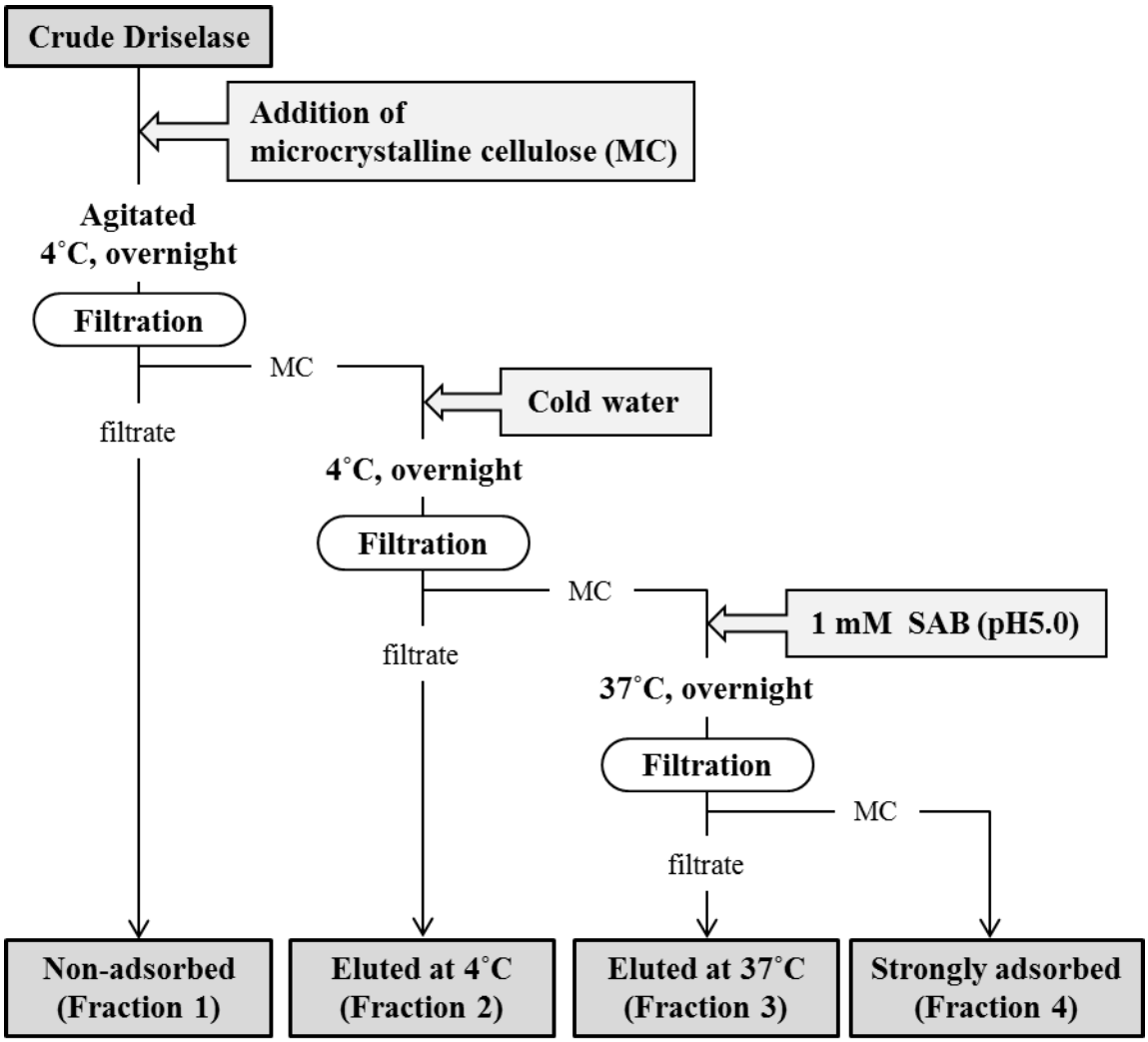


Fig. 3-2. Flow chart of the cellulose affinity fractionation of Driselase.

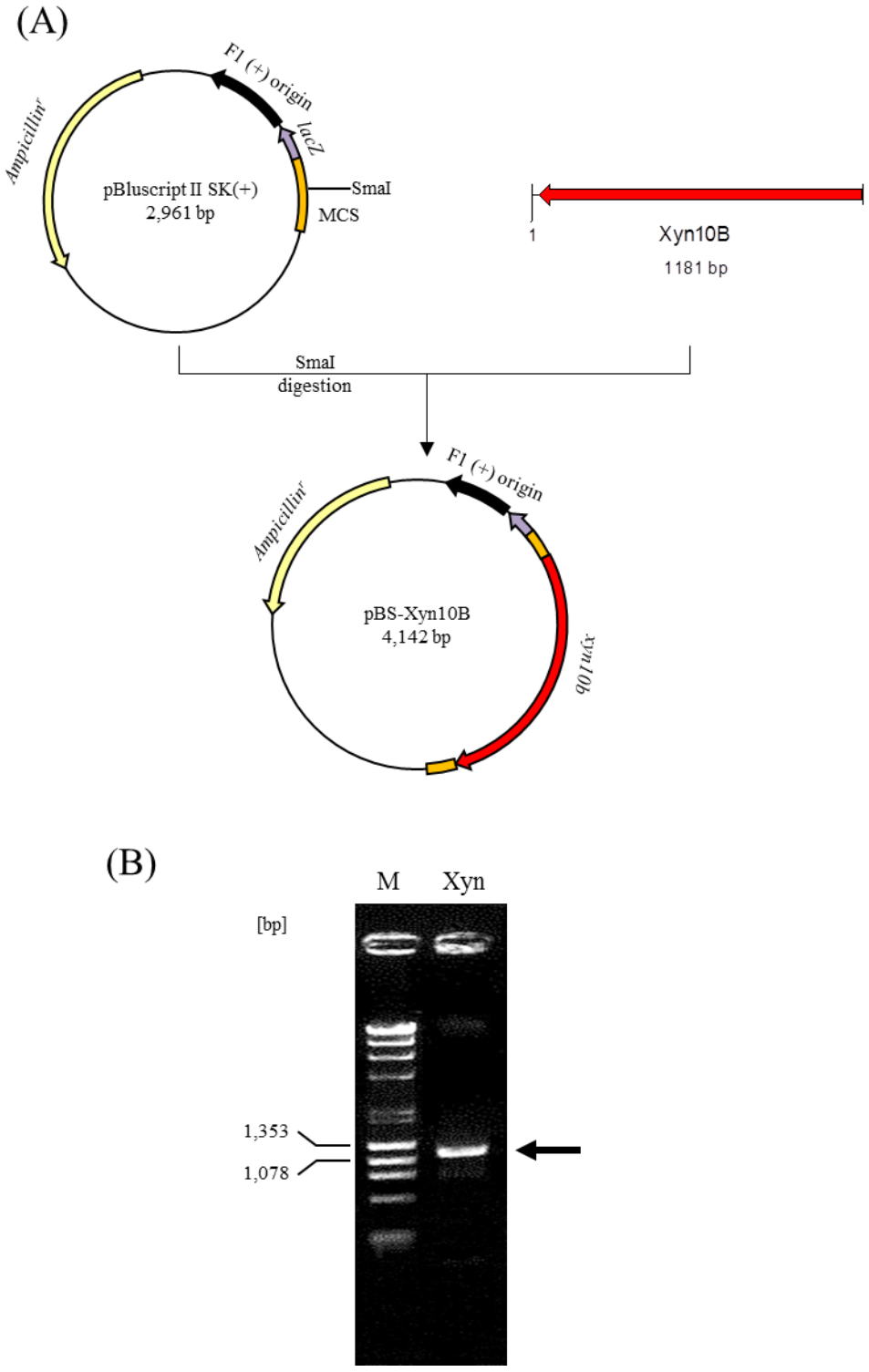


Fig. 3-3. Construction map of pBS-Xyn10B (A) and agarose gel electrophoresis of PCR fragment of Xyn10B cDNA gene amplification.

M, DNA digested markers; Xyn, PCR product of Xyn10B cDNA gene amplification. A mixture of λ/*Hin*dIII and φ174/*Hae*III was used as size markers. The arrow shows amplified cDNA of Xyn10B.

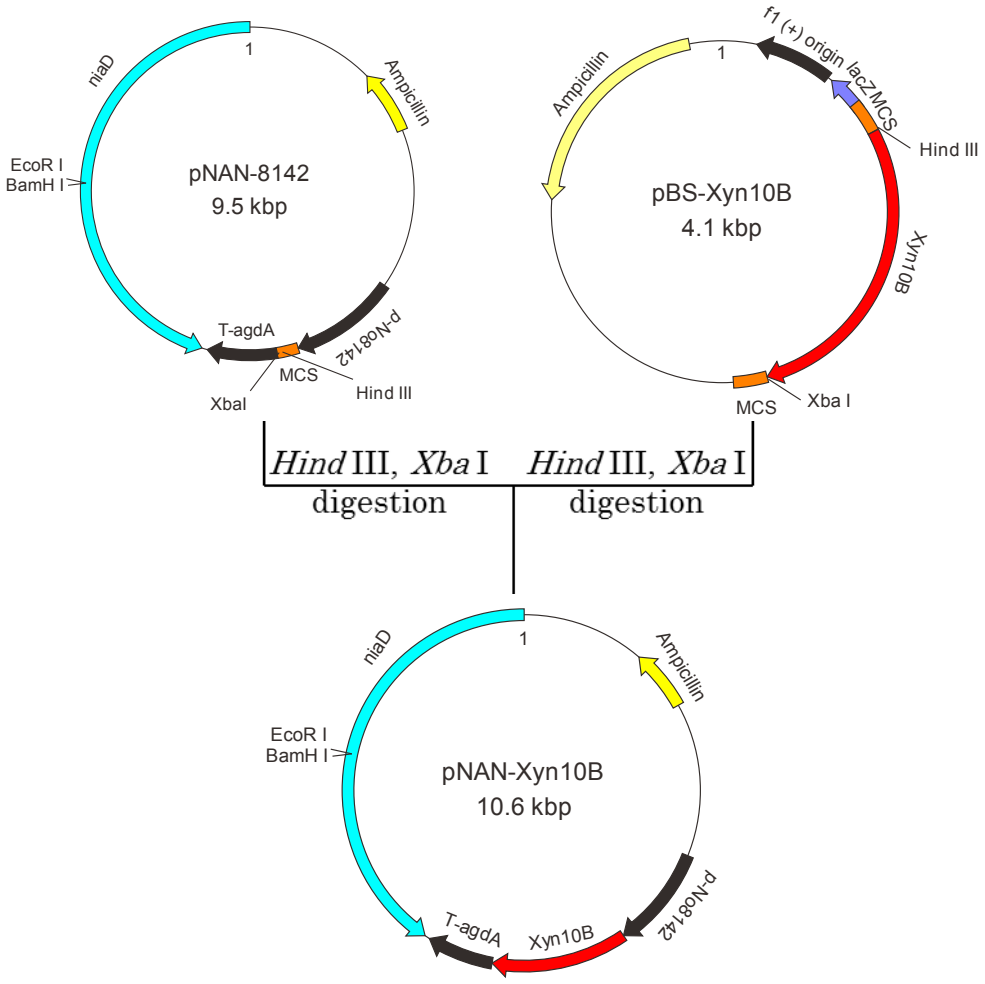


Fig. 3-4. Construction map of pNAN-Xyn10B

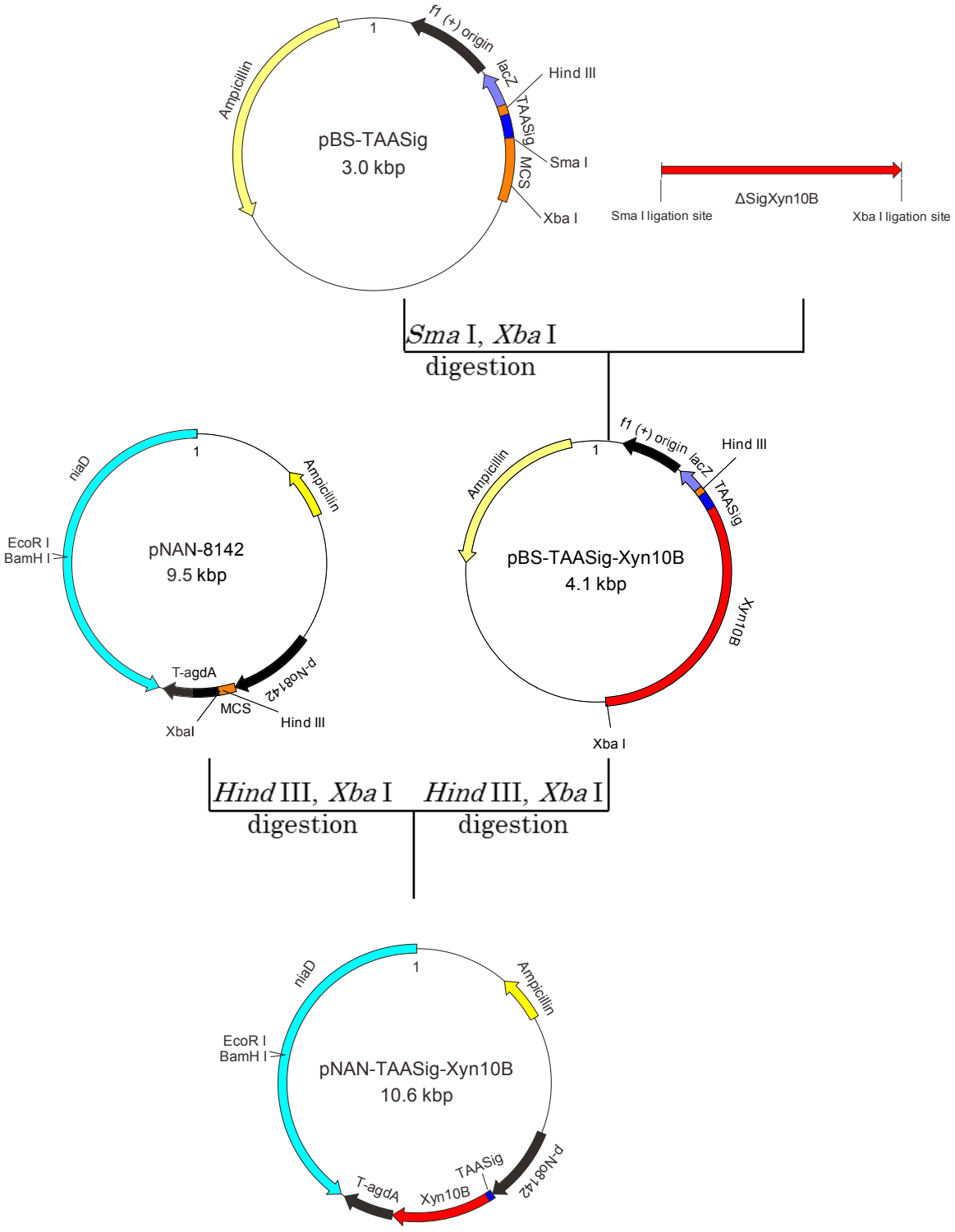


Fig. 3-5. Construction map of pNAN-TAASig-Xyn10B
*Δ*SigXyn10B fragment was obtained from pBS-*Δ*SigXyn10B by *Sma* I and *Xba* I digestion.

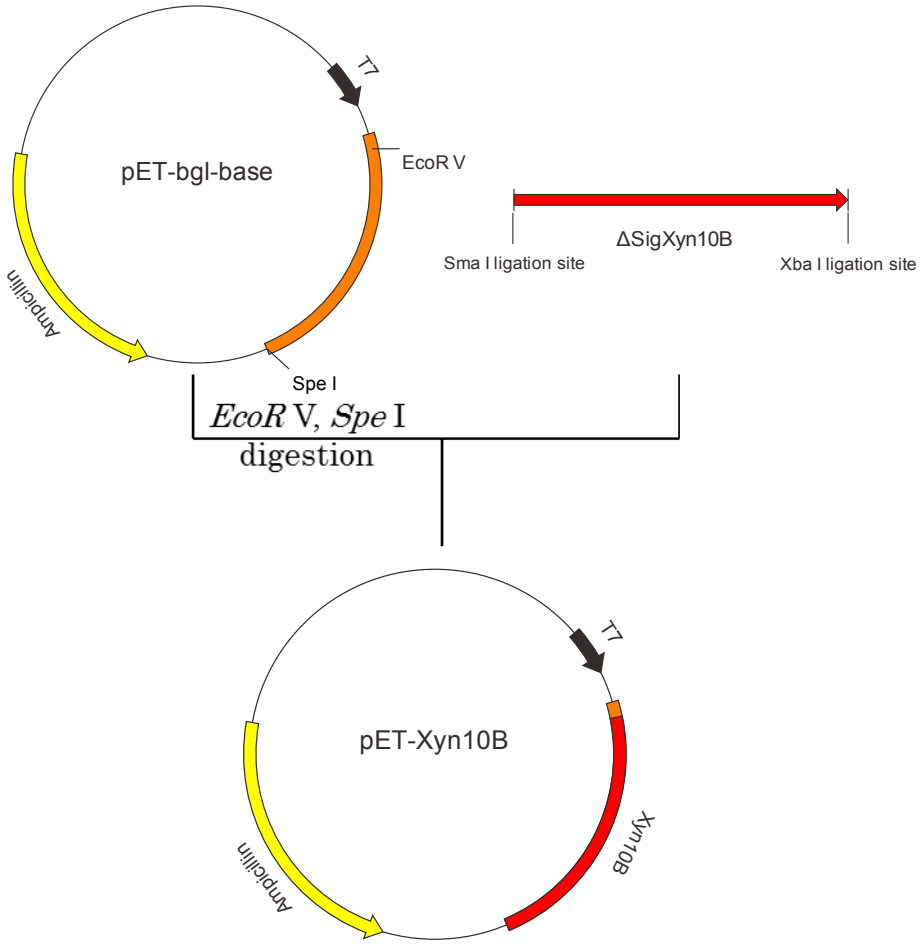


Fig. 3-6. Construction map of pET-Xyn10B
*Δ*SigXyn10B fragment was obtained from pBS-*Δ*SigXyn10B by *Sma* I and *Xba* I digestion. The *Xba* I ligation site and the *Sma* I ligation site of *Δ*SigXyn10B were ligated to *Spe* I site and *EcoR* V of pET-bgl-base respectively.

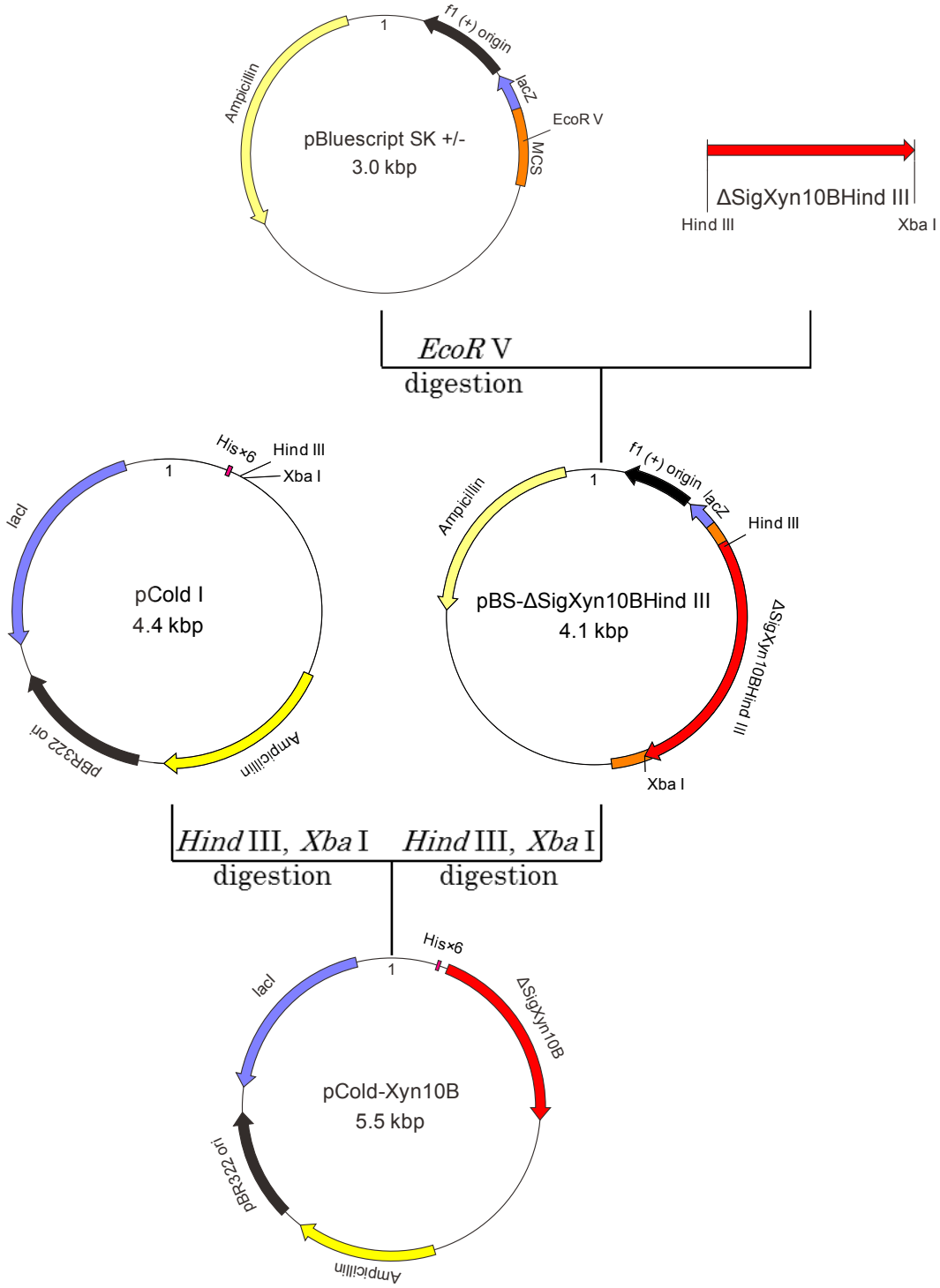


Fig. 3-7. Construction map of pCold-Xyn10B
⊿SigXyn10B fragment was amplified by PCR.

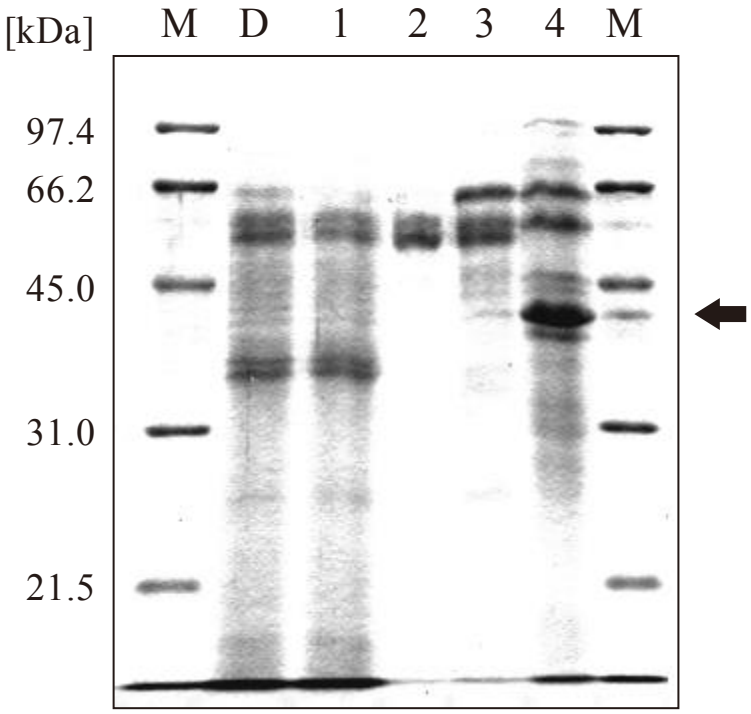


Fig. 3-8. SDS-PAGE of fractionated components adsorbed to MC.
M, molecular weight markers; D, Driselase, which is a commercial enzyme preparation; lane 1, non-adsorbed fraction; lane 2, eluted fraction at 4˚C; lane 3, eluted fraction at 37˚C; lane 4, eluted fraction with 2% SDS containing 0.06% 2-mercaptoethanol by boiling for 3 min. Known enzyme appearance position and 42 kDa protein (Xyn10B) are indicated by arrows. The acrylamide concentration of the gel was 12.5%.

Table 3-2. Specific activities of each Driselase fraction on various substrates

Substrate	Specific activity (mU/mg)				
	Driselase	Fraction 1	Fraction 2	Fraction 3	Fraction 4 [c]
MC [a]	19.7	16.3	34.7	53.4	3.80
CMC [a]	3920	4700	950	1455	64.5
Birchwood xylan [a]	954	N.D.	N.D.	N.D.	157
Oatspelt xylan [a]	585	N.D.	N.D.	N.D.	43.9
*p*NPG1 [b]	83.6	85.2	17.5	39.4	4.18
*p*NPG2 [b]	29.1	25.0	24.6	17.4	4.50
*p*NPL [b]	8.97	8.41	14.3	7.57	0.442
*p*NPX1 [b]	2.11	1.87	6.15	5.87	0.338

N.D.: not determined

a: One unit is defined as the amount of enzyme that produces reducing power equivalent to 1.0 μmol D-glucose per min.

b: One unit is defined as the amount of each *p*NP-substrate-digesting activity that releases 1.0 μmol *p*NP.

c: The MC residue of Driselase fractionation was suspended with 40 mL of 1 mM SAB (pH5.0), which was used as an enzymatic solution of fraction 4.

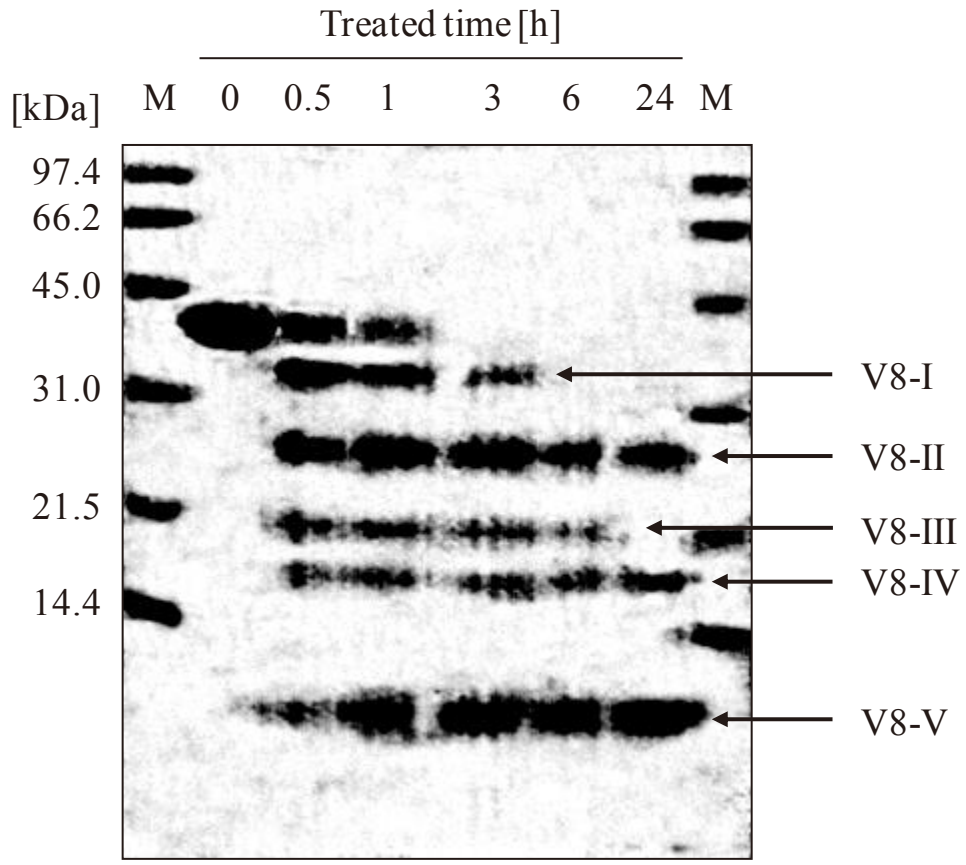


Fig. 3-9. SDS-PAGE of proteolysis pattern of 42 kDa protein digested by V8 protease.
M, molecular weight markers. The arrow indicates V8-protease-digested fragments from Xyn10B. The acrylamide concentration of the gel was 15.0%.

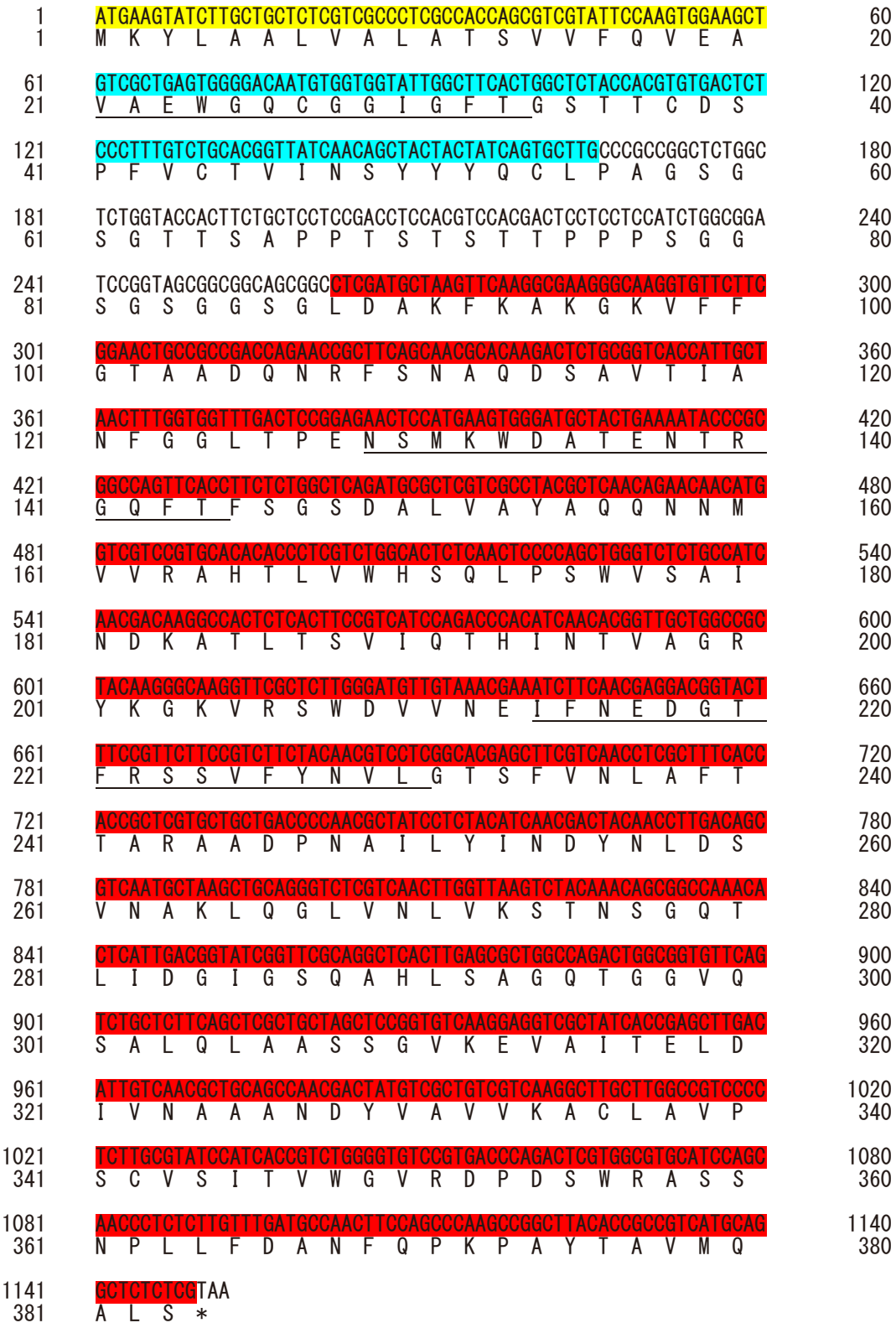


Fig. 3-10. cDNA sequence and deduced amino acid sequence of Xyn10B gene from *Irpex lacteus* MC-2.

The cDNA sequences encoding Xyn10B are colored yellow (signal sequence), cyan (CBM1), and red (catalytic domain) according to the domain construction of the deduced primary structure. The amino acid sequences corresponding to the results of N-terminal amino acid sequence analyses are underlined.

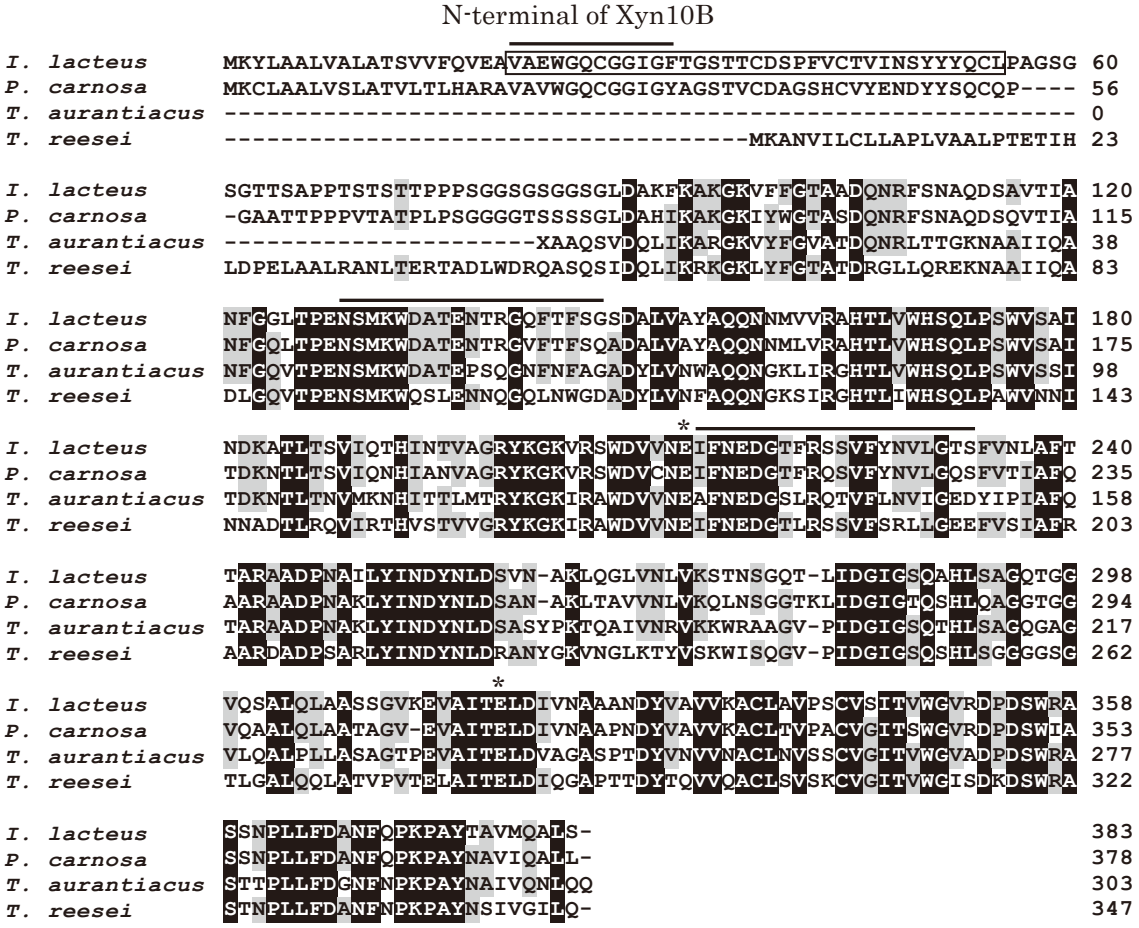


Fig. 3-11. Alignment of the amino acid sequences of Xyn10B and the homologous xylanases belonging to GH10.

**I. lacteus*, Xyn10B; *P. carnosa*, *Phanerochaete carnosa* HHB-10118-sp xylanase; *T. aurantiacus*, *Thermoascus aurantiacus* xylanase; *T. reesei*, *Trichoderma reesei* XynIII.

**Genbank accession numbers of each xylanase are *P. carnosa*, EKM58934; *T. aurantiacus*, 2BNJ_A; and *T. reesei*, BAA89465.2.

Black and gray boxes indicate identical and similar amino acids to those of *Ipx*Xyn10B, respectively. The underlined sequences of *Ipx*Xyn10B were determined by protein sequencing. The boxed region indicates CBM1. *, two deduced catalytic amino acids.

Fig. 3-12. Alignment of the GH family 10 xylanases from *I. lacteus*.
Three aromatic amino acids on a flat surface for adsorption to cellulose are shown by dots. Two deduced catalytic amino acids are shown by asterisks.

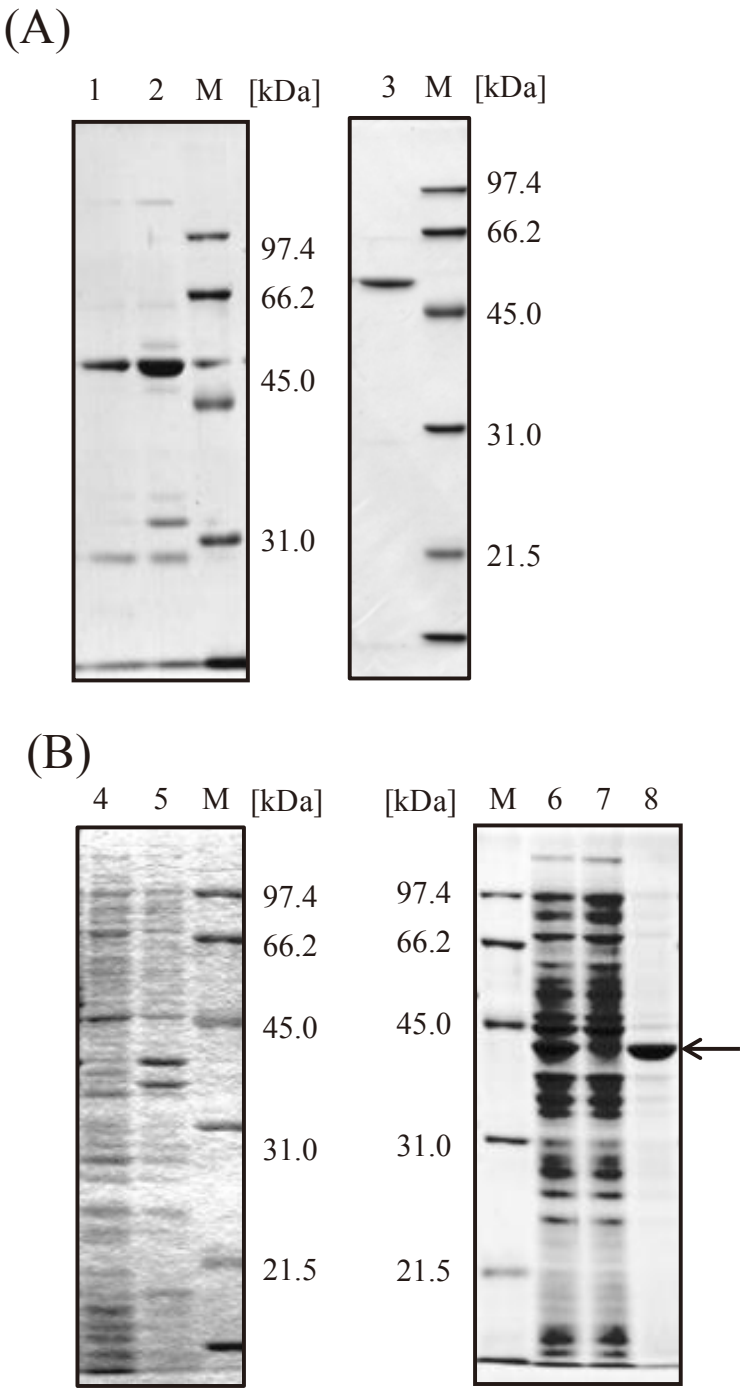


Fig. 3-13. SDS-PAGE of recombinant Xyn10B expressed by *Aspergillus oryzae* (A) and *Escherichia coli* (B).

M, molecular weight markers; lane 1, culture supernatant of *A. oryzae* transformant with pNAN-Xyn10B; lane 2, culture supernatant of *A. oryzae* transformant with pNAN-8142; lane 3, culture supernatant of *A. oryzae* transformant with pNAN-TAASigXyn10B; lane 4, soluble fraction after sonication of cells cultured with *E. coli* transformant with pETd-Xyn10B; lane 5, insoluble fraction; lane 6, whole cell lysate of cultured *E. coli* transformant with pCold-Xyn10B; lane 7, soluble fraction after sonication of cells; lane 8, insoluble fraction after sonication of cells. The arrow shows insoluble recombinant Xyn10B.

Table 3-3. Summary of purification of Xyn10B from crude Driselase

Step	Total activity [$\times 10^3$ U]	Total protein [$\times 10^4$ mg]	Specific activity [a] [U/mg]	Purification [fold]	Recovery [%]
Crude Driselase	98.4	9.2	1.07	1.0	100
40% saturated $(NH_4)_2SO_4$ precipitation	39.6	2.3	1.72	1.6	41.8
DEAE Sephadex A-50	13.4	0.055	24.3	23	14.1

a: Birchwood xylan was used as a substrate.

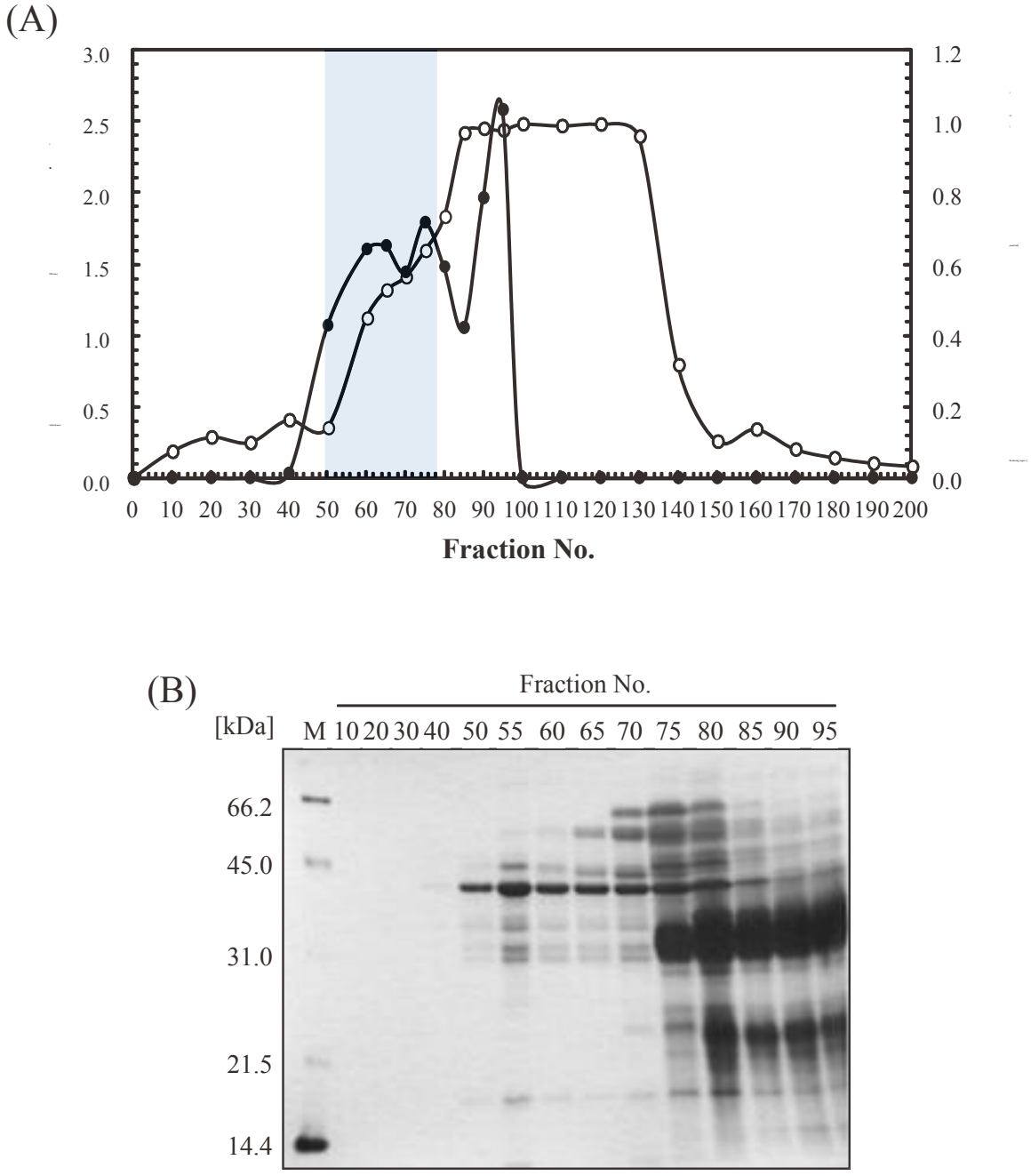


Fig. 3-14. Elution pattern of Driselase (after 40% ammonium sulfate fractionation) on DEAE Sephadex A-50 column (A) and SDS-PAGE analysis of eluted fractions (B).
○, Absorbance (280 nm); ●, enzymatic activity for birchwood xylan of each fraction with 200-fold dilution with 20 mM AAB (pH5.0). Elution of the proteins was performed by 20 mM AAB (pH 5.0) and recovered in a test tube as each fraction (10 ml/tube). The arrow shows Xyn10B in fraction 50.

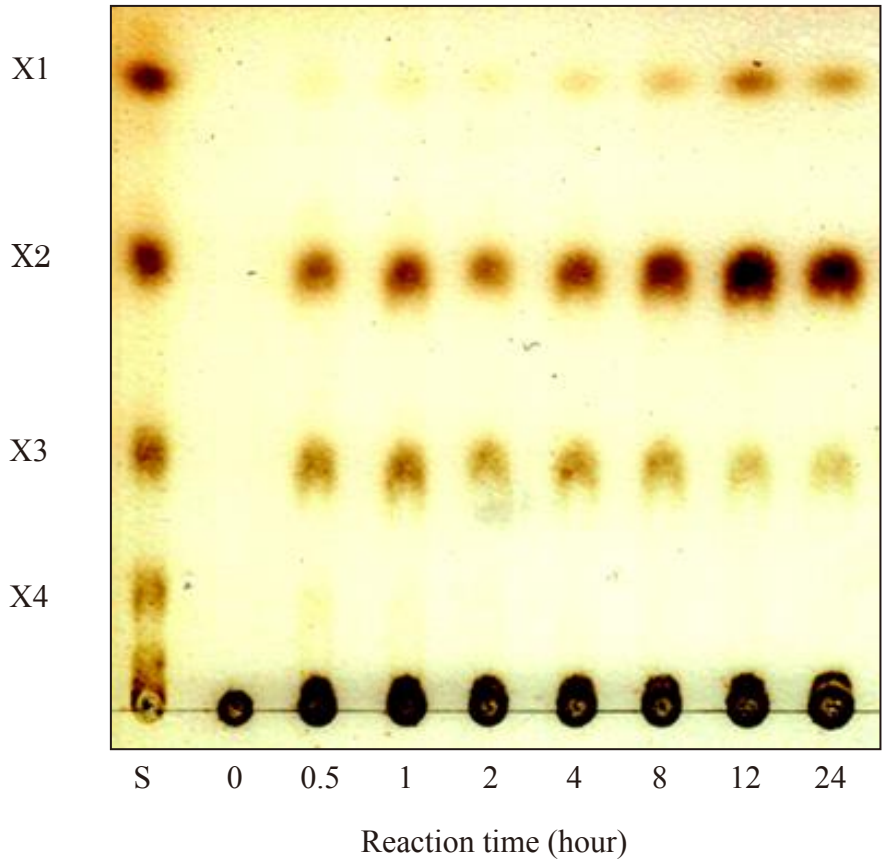


Fig. 3-15. Thin-layer chromatograph of time course pattern of hydrolysates from birchwood xylan by Xyn10B.

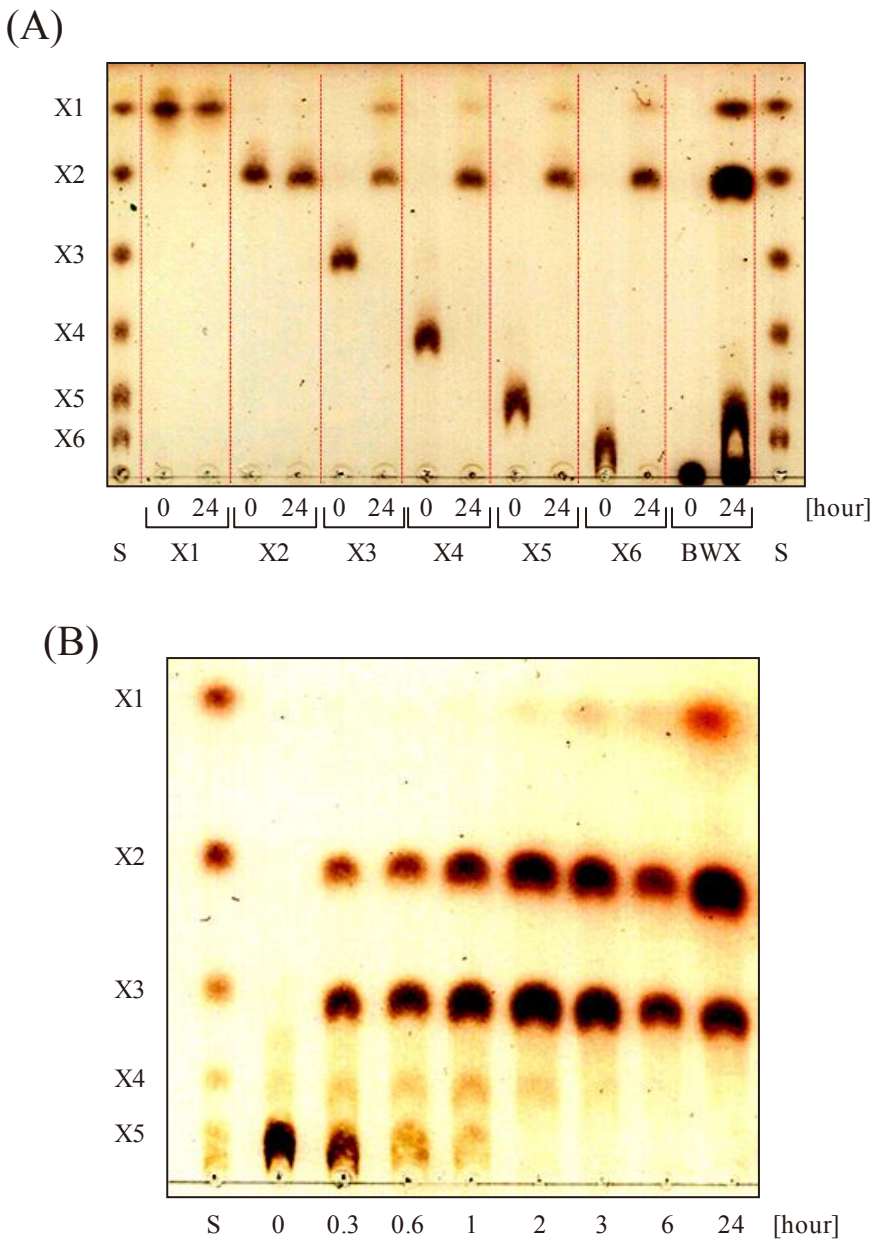


Fig. 3-16. Thin-layer chromatograph of hydrolysis pattern of xylooligosaccharides and birchwood xylane by Xyn10B (A) and time course pattern of hydrolysis products from xylopentaose by Xyn10B (B).

In (A), xylose, each xylooligosaccharide and birchwood xylan (BWX) were reacted with Xyn10B at 37˚C for 24 hours, and then 60 µl of reaction mixtures were spotted. In (B), xylopentaose was reacted with Xyn10B at 37˚C for 24 hours, and the reaction mixtures were sequentially sampled and 50 µl of reaction mixtures were spotted.

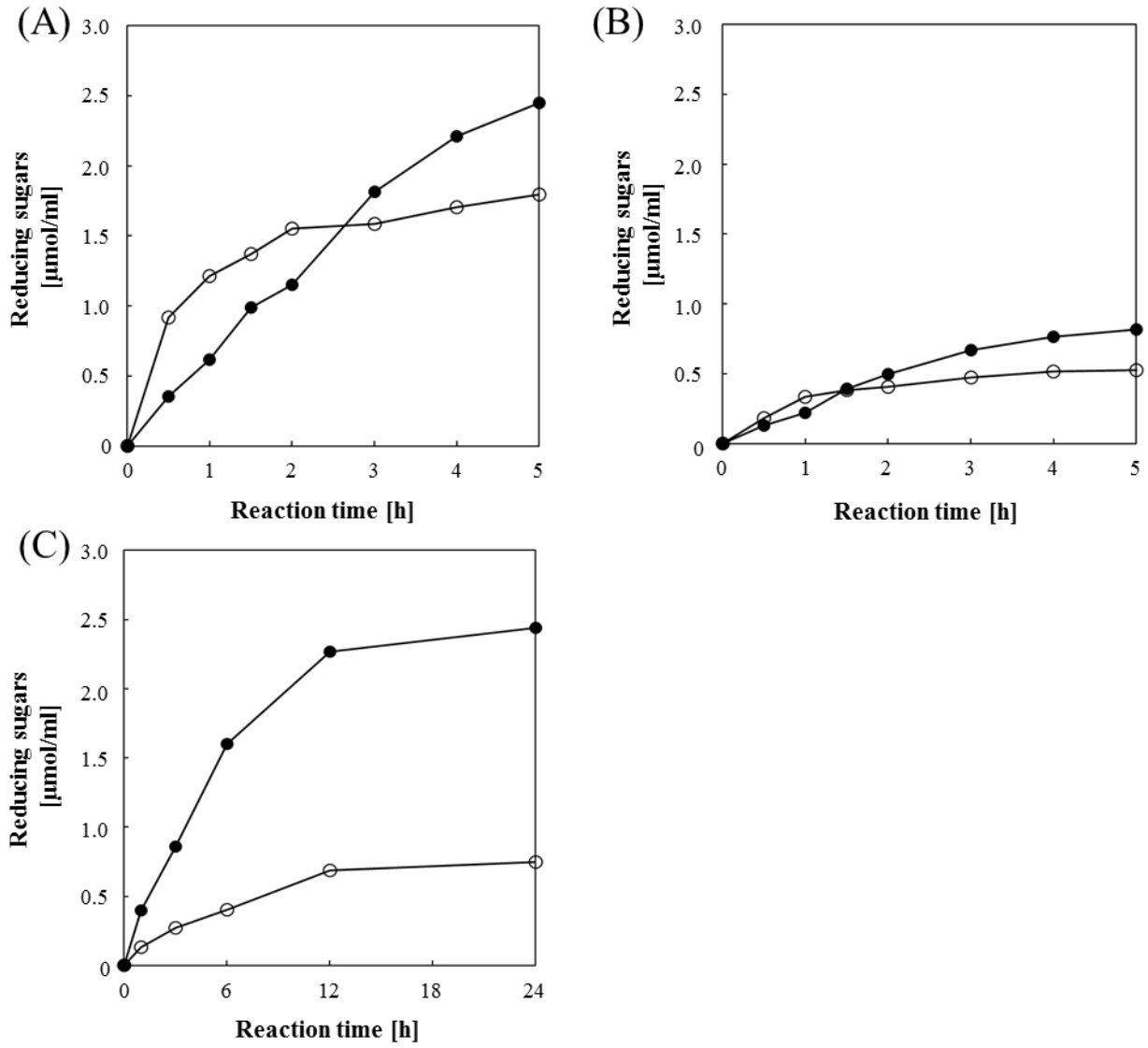


Fig. 3-17. Hydrolysis behavior of Xyn10B (●) and *Tr*XYN III (○) on (A) birchwood xylan, (B) oatspelt xylan, and (C) biomass (ionic liquid-pretreated erianthus).

References

1. **Novotný Č, Cajthaml T, Svobodová K, Šušla M, Šašek V.** 2009. *Irpex lacteus*, a white-rot fungus with biotechnological potential. Folia Microbiol. **54**:375-390.
2. **Chunyan X, Fuying M, Xiaoyu Z, Shulin C.** 2010. Biological pretreatment of corn stover by *Irpex lacteus* for enzymatic hydrolysis. J. Agric. Food Chem. **58**:10893-10898.
3. **Yanqing Y, Yelin Z, Jiane Z, Fuying M, Xuewei Y, Xiaoyu Z, Yujue W.** 2013. Improving the conversion of biomass in catalytic fast pyrolysis via white-rot fungal pretreatment. Bioresour. Technol. **134**:198-203.
4. **Kanda T, Nakakubo S, Wakabayashi K, Nisizawa K.** 1978. Purification and properties of exo-cellulase of avicelase type from a wood-rotting fungus, *Irpex lacteus* (Polyporus tulipiferae). J. Biochem. **84:**1217-1226.
5. **Hamada N, Ishikawa K, Okumura R, Fuse N, Kodaira R, Shimosaka M, Amano Y, Kanda T, Okazaki M.** 1999. Purification, characterization and gene analysis of exo-cellulase II (Ex-2) from the white rot basidiomycete *Irpex lacteus*. J. Biosci. Bioeng. **87:**442-451.
6. **Toda H, Nagahata N, Amano Y, Nozaki K, Kanda T, Okazaki M, Shimosaka M.** 2008. Gene cloning of cellobiohydrolase II from white rot fungus *Irpex lacteus* MC-2 and its expression in *Pichia pastoris*. Biosci. Biotechnol. Biochem. **72:**3142-3147.
7. **Kanda T, Wakabayashi K, Nisizawa K.** 1980. Purification and properties of a lower-molecular-weight endo-cellulase from *Irpex lacteus* (Polyporus tulipiferae). J. Biochem. **87:**1625-1634.
8. **Toda H, Takada S, Oda M, Amano Y, Kanda T, Okazaki M, Shimosaka M.** 2005. Gene cloning of an endoglucanase from the basidiomycete *Irpex lacteus* and its cDNA expression in *Saccharomyces cerevisiae*. Biosci. Biotechnol. Biochem. **69:**1262-1269.
9. **Amano Y, Kanda T, Okazaki M.** 1994. Mode of action of two endo-1,4-β-xylanases from *Irpex lacteus* toward rice-straw xylan . Mokuzai gakkaishi. **40:**57-63.
10. **Amano Y, Okazaki M, Mitani M, Kanda T.** 1994. Transglycosyl reaction of endo-1,4-β-xylanases from *Irpex lacteus* on aryl β-xyloside. Mokuzai gakkaishi. **40:**308-314.
11. **Amano Y, Okazaki M, Kanda T.** 1994. Effect of substituted groups of aryl β-D-xyloside on transglycosyl reaction of two different xylanases from *Irpex lacteus*. Mokuzai gakkaishi. **40:**315-320.
12. **Kanda T, Amano Y, Nishizawa K.** 1985. Purification and properties of two endo-1,4-β-xylanases from *Irpex lacteus* (*Polyporus tulipiferae*). J. Biochem. **98:**1545-1554.
13. **Ogasawara W, Shida Y, Furukawa T, Shimada R, Nakagawa S, Kawamura M, Yagyu T, Kosuge A, Xu J, Nogawa M, Okada H, Morikawa Y.** 2006. Cloning, functional expression and promoter analysis of xylanase III gene from *Trichoderma reesei*. Appl. Microbiol. Biotechnol. **72:**995-1003.
14. **Laemmli UK.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680-685.

15. **Lowry OH, Rosebrough NJ, Farr AL, Randall RJ.** 1951. Protein measurement with the folin phenol reagent. J. Biol. Chem. **193:**265-275.
16. **Somogyi M.** 1952. Note on sugar determination. J. Biol. Chem. **195:**19-23.
17. **Nelson N.** 1952. A photometric adaptation of somogyi method for the determination of glucose. J. Biol. Chem. **15:**375-380.
18. **Gomi K, Iimura Y, Hara S.** 1987. Integrative transformation of *Aspergillus oryzae* with a plasmid containing the *Aspergillus nidula*ns *argB* gene. Agric. Biol. Chem. **51:**2549-2555.
19. **Kimura I, Oshima H, Sasahara H, Kimura Y, Tajima S.** 2004. Cloning, sequencing and expression analysis of an endo-1,4-β-D-xylanase gene from *Aspergillus sojae*. J. Biol. Chem. **51:**237-240.
20. **Black GW, Rixon JE, Clarke JH, Hazlewood GP, Theodorou MK, Morris P, Gilbert H. J.** 1996. Evidence that linker sequences and cellulose-binding domains enhance the activity of hemicellulases against complex substrates. Biochem. J. **319:**515-520.
21. **Pavón-Orozco P, Santiago-Hernández A, Rosengren A, Hidalgo-Lara ME, Stålbrand H.** 2012. The family II carbohydrate-binding module of xylanase CflXyn11A from *Cellulomonas flavigena* increases the synergy with cellulase TrCel7B from *Trichoderma reesei* during the hydrolysis of sugar cane bagasse. Bioresour. Technol. **104:**622-630.
22. **Hagglund P, Eriksson T, Collen A, Nerinckx W, Claeyssens M, Stalbrand H.** 2003. A cellulose-binding module of the *Trichoderma reesei* β-mannanase Man5A increases the mannan-hydrolysis of complex substrates. J. Biotechnol. **101:**37-48.
23. **Sugimoto N, Igarashi K, Samejima M.** 2012. Cellulose affinity purification of fusion proteins tagged with fungal family 1 cellulose-binding domain. Protein Expr. Purif. **82**: 290-296.
24. **Geraldine C, Markus L.** 1999. Widely different off rates of two closely related cellulose-binding domains from *Trichoderma reesei*. Eur. J. Biochem. **262:**637-643.
25. **Creagh AL, Ong E, Jervise 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. Biochemistry **93:**12229-12234.
26. **Jervis EJ, Haynes CA, Kilburn DG.** 1997. Surface diffusion of cellulases and their isolated binding domains on cellulose. J. Biol. Chem. **272:**24016-24023.
27. **Mclean BW, Bray MR, Boraston AB, Gilkes NR, Haynes CA, Kilburn DG.** 2000. Analysis of binding of the family 2a carbohydrate-binding module from *Cellulomonas fimi* xylanase 10A to cellulose: specificity and identification of functionally important amino acid residues. Protein Eng. **13:**801-809.

Chapter 4

Analysis of strong adsorption of CBM1 from Xyn10B on cellulose

Chapter 4: Analysis of strong adsorption of CBM1 from Xyn10B on cellulose

4. Introduction

In chapter 3, it was revealed that the 42 kDa protein showing high affinity to microcrystalline cellulose (MC) was an endo-1,4-β-xylanase belonging to GH10. This protein designated as Xyn10B is composed of 383 amino acid residues and constructed of two domains, which are CBM1 (residues 21 to 55) and a catalytic domain (residues 87 to 383) located at the N- and C-termini, respectively. This result supports the finding that Xyn10B strongly adsorbs to cellulose. However, it is still unclear why Xyn10B is not desorbed from cellulose once it has been adsorbed. From the sequence analysis of CBM1 of Xyn10B (CBM_{Xyn10B}), it was found that CBM_{Xyn10B} has additional aromatic amino acid residues (Phe 42 and Tyr 52) in addition to the highly conserved three-aromatic-amino-acid motif among CBM1s (**Table 4-1**). In several studies, it has been reported that the binding ability of CBM1 to crystalline cellulose is affected by the type and configuration of aromatic acid residues. Against this background, it seems to be important to examine the effects of Phe 42 and Tyr 52 on the adsorption to cellulose.

As described in this chapter, CBM_{Xyn10B} and CBM_{ThEG1}, which is CBM1 of endo-1,4-β-glucanase from *Trametes hirsuta* as a control, were expressed as fusion proteins with GFP. In addition, two mutated CBM_{Xyn10B} fusion proteins of which Phe 42 and Tyr 52 were replaced by serine were constructed to elucidate strong adsorption of Xyn10B.

4.1. Materials and methods

4.1.1. Chemicals and strains

Microcrystalline cellulose (MC) (column chromatography grade, Merck, Germany) and cotton (Hakujuji Co., Ltd., Japan) were used as substrates for cellulose-binding assays. *Escherichia coli* DH5α purchased from Takara (Shiga, Japan) was used as the host strain for plasmid extraction, grown at 37˚C in LB medium containing 50 μg/mL ampicillin. Plasmid pBluescript II SK(+) was used as the subcloning vector. *E. coli* Origami B strain (Novagen, NJ) and plasmid pRSET/EmGFP (Invitrogen, Carlsbad, CA) were used for analysis of the expression of CBM1 fused to green fluorescence protein (CBM-GFP).

4.1.2. Analytical methods

SDS-PAGE was performed by the method of Laemmli using 12.5% gel [1]. After electrophoresis, the gel was stained with 0.25% CBB-R250. As the molecular weight standard, low range (BIO-RAD, USA) markers were used. Protein content was determined by the method of Lowry [2] with BSA as a standard. The fluorescence intensity of CBM-GFPs was measured using an FP-6200 spectrofluorometer (JASCO Japan).

4.1.3. Plasmid construction of CBM1 of Xyn10B and GFP fusion protein

The cDNA region corresponding to CBM1 of Xyn10B (CBM_{Xyn10B}) was amplified from pBS-Xyn10B (described in chapter 4) by PCR reaction using PrimeSTAR HS DNA polymerase (PCR conditions: 95˚C for 5 min, followed by 30 cycles at 98˚C for 10 s, 60˚C for 15 s, and 72˚C for 30 s, followed by a final extension step at 72˚C for 5 min) with the sense primer (5'-AGGATCCGCTGTCGCTGAGTGG-3', *Bam*HI site is underlined) and the reverse primer (5'-GCCATGGAGCCAGAGCCGGCGG-3', *Nco*I site is underlined). The amplified CBM_{Xyn10B} fragment (138 bp) was bluntly ligated into the *Sma*I site of pBluescript II SK(+). The resulting plasmid was then digested by *Bam*HI and *Nco*I, and the CBM_{Xyn10B} fragment was inserted into the same sites of pRSET/EmGFP for fusion to GFP with a histidine tag. The resulting plasmid (named pRSET-CBM_{Xyn10B}-GFP) was used as an *E. coli* expression vector. The fusion protein expression vector of the CBM1 derived from *T. hirsuta* endoglucanase 1 and GFP (pRSET-CBM_{ThEG1}-GFP) was prepared as described by Nozaki *et al.* [3] and in chapter 2.

4.1.4. CBM-GFP mutant construction

The expression vectors for CBM1 mutated fusion proteins (CBM_{F42S}-GFP and CBM_{Y52S}-GFP) were constructed according to the manual of the PrimeSTAR Mutagenesis basal kit (Takara, Japan) using mutation primers (5'-TCTCCCTCTGTCTGCACGGTTATCAAC-3' and 5'-GCAGACAGAGGGAGAGTCACACGTGGT-3' for CBM_{F42S}-GFP; 5'-TACTACTCTCAGTGCTTGCCCGCCGGC-3' and 5'-GCACTGAGAGTAGTAGCTGTTGATAAC-3' for CBM_{Y52S}-GFP, serine mutation site is underlined) and pRSET-CBM_{Xyn10B}-GFP as a template (PCR conditions: 95˚C for 5 min,

followed by 30 cycles at 98˚C for 10 s, 52˚C for 15 s, and 72˚C for 30 s, followed by a final extension step at 72˚C for 5 min).

The nucleotide sequences of the obtained plasmids were confirmed by sequencing and then transformed into *E. coli* Origami B strain (Novagen, NJ) for expression.

4.1.5. Expression and purification of CBM-GFPs

E. coli Origami B strains were grown in 100 ml of LB medium containing 100 μg/mL ampicillin with shaking at 180 rpm and 30˚C, induced with 1 mM IPTG after 12 hours, and then grown after an additional 24 hours. Cells were harvested by centrifugation and suspended in 20 mM sodium phosphate buffer (pH 7.4) containing 0.5 M NaCl. Cells were disrupted by ultrasonic treatment and each CBM-GFP was purified from supernatant using a His-trap FF column (GE Healthcare, UK) according to the manual. Purified CBM-GFPs were confirmed by SDS-PAGE as a single protein band.

4.1.6. Binding study of CBM-GFPs

The reaction mixture (1 mL) consisted of 0.02% MC and various amounts of protein in 10 mM SAB (pH 5.0). The reaction mixture was incubated at 4˚C for 2 hours with rotation at 12 rpm. The amount of adsorbed protein was measured by determining the fluorescence intensity (excitation/emission = 487/509 nm) in the supernatant of the reaction mixture. Binding parameters were determined by fitting the data to the Langmuir equation.

$$[B]=A_{max}\,[F]\,K_{ad}/(1+K_{ad}\,[F])$$

In this equation, [B] and [F] are the concentrations of bound and free CBM-GFPs (mol per gram of cellulose), respectively. A_{max} and K_{ad} are the maximal concentration of bound protein (μmol/g of cellulose) and the association constant for the binding site (μM^{-1}), respectively. The A_{max} and K_{ad} values in the equation can be calculated by single reciprocal plots ([B]/[F] versus [F]) [4].

CBM-GFPs adsorbing to the MC were washed with 0.1 mL of deionized water two times and subjected to SDS-PAGE after elution (described above).

4.1.7. Fluorescent images of cotton fiber with CBM-GFPs

The CBM-GFP fusion protein (1 μM) was incubated with 1 mg of cotton in 10 mM SAB (pH 5.0), in a total volume of 1 mL. After 2 hours at 4˚C with rotation at 12 rpm, the cotton was washed twice with 1 mL of the same buffer. Adsorbed CBM-GFP fusion protein was observed using a fluorescence microscope, BX-51 (Olympus, Japan).

4.2 Results

4.2.1. Binding behavior of each CBM-GFP on cellulose

A construction map of pRSET-CBM_{Xyn10B}-GFP is shown in **Fig. 4-1**. Amplified CBM_{Xyn10B} is shown in **Fig. 4-2**. To clarify the observed strong adsorption of Xyn10B on crystalline cellulose, only the CBM1 region of Xyn10B was expressed as a fusion protein with GFP (CBM_{Xyn10B}-GFP), as shown in **Fig. 4-3 (A)**, and the adsorption behavior was investigated. An additional fusion protein between GFP and CBM1 from *Th*EG1 (designated as CBM_{ThEG1}-GFP) was also constructed to compare the adsorption ability on cellulose. The adsorption behavior of CBM_{Xyn10B}-GFP was similar to that of Xyn10B, which had the same CBM, showing apparently higher adsorption than CBM_{ThEG1}-GFP. This result suggests that the strong adsorption of Xyn10B arises from the CBM1, not from the catalytic domain (**Fig. 4-3 (B)**)

Adsorption isotherms were also determined according to the one-site Langmuir model (**Fig. 4-4**). The A_{max} and K_{ad} values of each CBM-GFP fusion protein are shown in **Table 4-2**. The A_{max} (7.8 μmol/g) and K_{ad} (2.0 L/μmol) values of CBM_{Xyn10B}-GFP were about 2-fold higher than those of CBM_{ThEG1}-GFP (3.4 μmol/g and 1.2 L/μmol, respectively), even though the two proteins had the same binding motif (W-YY). The isotherms of CBM_{ThEG1}-GFP were almost the same as those reported in the CBM1s from *T. reesei* Cel7A and Cel5A [5]. To investigate the role of the additional Phe 42 and Tyr 52 residues, mutant proteins, CBM_{F42S}-GFP and CBM_{Y52S}-GFP, were constructed. The adsorption isotherms of CBM_{Y52S}-GFP decreased significantly when compared with the native protein, and were very similar to those observed for CBM_{ThEG1}-GFP. On the other hand, the adsorption isotherm of CBM_{F42S}-GFP was not changed from that of native CBM_{Xyn10B}-GFP. From these results, it is suggested that the aromatic residue Tyr 52 plays an important role in defining the binding affinity. Thus, CBM_{Xyn10B} contains a unique motif (W-YYY) that facilitates strong cellulose adsorption.

4.2.2. Microscopic images of the adsorption of CBM-GFPs on cotton

CBM-GFPs adsorbed on cotton were observed under a fluorescence microscope and the distribution of the fusion proteins was also observed (**Fig. 4-5**). GFP (A, B) without a CBM was not adsorbed onto the cotton, resulting in the observation of no fluorescence on cellulose. CBM_{ThEG1}-GFP (E, F) and the mutant protein, CBM_{Y52S}-GFP (G, H), showed uniform adsorption on cotton surfaces. On the other hand, the fluorescence of CBM_{Xyn10B}-GFP (C, D) was detected as a spot on the cotton due to bunched adsorption. This set of observations suggests that the adsorption behaviors of the native and mutant CBMs are completely different from each other. This may be due to the high binding capacity, A_{max}, on cellulose. The aggregation of CBM_{Xyn10B}-GFP was observed in solution also in the absence of cellulose, and could be enhanced by increasing the protein and salt concentrations (**Fig. 4-6**). On the other hand, CBM_{Y52S}-GFP did not show signs of aggregation in solution under similar conditions.

4.3. Discussion

On the basis of the results from the draft genome sequencing, all of the GH family 10 xylanases (Xyn10A, Xyn10B, and Xyn10C) from *I. lacteus* NK-1 have a CBM1 at their N-terminal region. Given the difference among these CBMs, two aromatic residues, Phe 42 and Tyr 52, which are replaced with a serine in Xyn10B, could have some influence on the structure of CBM1 (**Table 3-1**). As shown in **Table 4-1**, three contiguous aromatic residues in the CBM1 region are not found in other origins. Consequently, we mutated Tyr 52 to another amino acid. To investigate the role of Tyr 52 in cellulose binding, a mutant protein, CBM_{Y52S}-GFP, was constructed. The adsorption isotherms of CBM_{Y52S}-GFP decreased significantly when compared with those of the native protein (**Fig. 4-4**), with values that closely matched the values of CBM_{ThEG1}-GFP (**Table 4-2**).

The adsorption ability of the CBM_{Xyn10B}-GFP on cellulose was found to be 2-fold greater than that of CBM_{ThEG1}-GFP, which is used as a general CBM1 (**Table 4-2**). Compared with other CBMs from several microorganisms and the mutant proteins, the CBM_{Xyn10B}-GFP showed high adsorption capacity on cellulose. Although the mutant protein CBM_{F42S}-GFP showed similar adsorption parameter values to the native CBM_{Xyn10B}-GFP, CBM_{Y52S}-GFP showed almost the same adsorption capacity as the other proteins presented in **Table 4-2**.

From the CAZy database ((http://www.cazy.org/), we found many W-YYY motif

CBM1s in hemicellulolytic enzymes, such as xylanase and mannanase (19/55) (**Table 4-3**). On the other hand, only 2 W-YYY motif CBM1s were found in CBH I-type cellulases, despite most types of CBH I containing at least one CBM1. As the W-YYY motif CBM1s have strong ability to adsorb on cellulose, these properties can be suitable for hemicellulases such as xylanase and mannanase because of the location of hemicellulose, which is a neighbor to cellulose [6-8]. On the other hand, CBH I can actively degrade crystalline cellulose with processive attack, so movement on the surface of cellulose is required. The property for weak adsorption on cellulose should be suitable for this type of mode of action. It was reported that the progress of processive cellulase and reducing sugar production from filter paper was reduced by the addition of CBM1 [9]. Actually, CBM_{Xyn10B} is not a suitable CBM1 for CBH I-type cellulase. We constructed a chimera CBH I of *T. reesei* (CBM1 of CBH I replaced by CBM_{Xyn10B}) expression system in *A. oryzae*. This chimera CBH I showed strong adsorption to cellulose, but cellulose degradation activities were significantly decreased compared with those of native CBH I, despite soluble substrate degradation activity being almost unchanged (data not shown). This finding coincides with our hypothesis presented above.

The adsorption behavior of the unique CBM presented above was observed under a fluorescence microscope. The CBM_{Xyn10B}-GFP could adsorb on cotton in an aggregated form, and be visualized as green fluorescence spots. The aggregates of the CBM_{Xyn10B}-GFP were found under high salt concentration conditions, even in the absence of cellulose (data not shown). In addition, CBM_{Xyn10B}-GFP showed no specific adsorption to insoluble xylan, but showed stronger adsorption to regenerated amorphous cellulose than CBM_{ThEG1}-GFP (data not shown). This property is likely to be closely related to the hydrophobicity of the CBM.

We undertook homology modeling of the CBM_{Xyn10B} using Swiss Model (http://swissmodel.expasy.org/) and Pymol (http://www.pymol.org/) with the CBM1 of *T. reesei* CBH I (PDB ID is 1CBH) as a template (**Fig. 4-7**). As a result, despite Tyr 52 having an important role in the strong adsorption of CBM_{Xyn10B}, it was suggested that the aromatic side chain of additional Tyr 52 is not located on the flat surface for cellulose binding. The CBM_{Xyn10B}-GFP was aggregated in the absence of cellulose (**Fig. 4-6**), which we expected to be one reason for the strong adsorption of CBM_{Xyn10B}. The additional tyrosine Tyr 52 plays an important role not only in the adsorption process but also in the aggregation of CBM_{Xyn10B}-GFP molecules. One possibility that cannot be ruled out is that Tyr 52 is not involved directly in hydrophobic interaction with the cellulose surface, but is responsible for

the formation of CBM1 aggregates. It is noteworthy that CBM_{Xyn10B}-GFP shows remarkable aggregation and still maintains the binding ability. Markus *et al.* reported that double CBM1, which was made by fusing the N-terminal CBM1 of *T. reesei* CBHII to the C-terminal CBM1 of CBHI, showed remarkably high affinity for cellulose because of the double-binding model [10]. Although there might be various reasons why Xyn10B has strong adsorption ability on cellulose, it is suggested that the aggregation of CBM_{Xyn10B} domains can form a multiple binding face; this seemed to be one of the possibilities explaining the strong adsorption.

The reason for the strong cellulose adsorption of the CBM1 of Xyn10B remains unclear and needs to be resolved. To elucidate the precise structure of the CBM1 of Xyn10B, X-ray crystallography is needed; however, no one has succeeded in obtaining the suitable crystal of the CBM1 domain for X-ray diffraction.

In conclusion, Xyn10B showed strong adsorption ability on crystalline cellulose with a novel W-YYY CBM1 motif. Interestingly, the CBM1 of Xyn10B has an additional aromatic amino acid, Tyr 52, which can participate in cellulose binding. However, the role of Tyr 52 of the CBM_{Xyn10B} domain in binding to cellulose is unclear. Furthermore, the reason why hemicellulases such as xylanases and mannanases have strong adsorption abilities remains to be resolved. The strong cellulose adsorption properties of the CBM1 of xylanase from *I. lacteus* may provide new insights into biomass degradation by fungi.

Table 4-1. Alignment of the amino acid sequences of the general CBM1s from several sources

Organism [a]	Enzyme [b]	CBM1 sequence [c]
I. lacteus	Xyn10B	VAE**W**GQCGGIGFTGSTTCDSP**F**VCTVINS**YY**Y QCL
T. hirsuta	EG1	-AV**W**GQCGGIGFSGDTTCTA-STCVKVND**YY**SQCQ
T. reesei	CBH I	--H**Y**GQCGGIGYSGPTVCASGTTCQVLNP**YY**SQCL
T. reesei	EG I	--H**W**GQCGGIGYSGCKTCTSGTTCQYSND**YY**SQCL
P. chrysosporium	XYNA	-PV**W**GQCGGIGWTGPTTCTAGNVCQEYSA**YY**SQCI
P. chrysosporium	XYNB	-AL**Y**GQCGGQGWTGPTCCSSGT-CKFSNN**WY**SQCL

a: *I. lacteus*, *Irpex lacteus*; *T. hirsuta*, *Trametes hirsuta*; *T. reesei*, *Trichoderma reesei*; *P. chrysosporium*, *Phanerochaete chrysosporium*.

b: Genbank accession numbers of each enzyme are: Xyn10B, AB933639; Ex-1, BAA76364; EG1, BAD01163.1; CBH I, CAH10320; EG I, AAA34212; XYNA, AAG44992; and XYNB, AAG44995.

c: Gray boxes indicate identified amino acid residues. Three aromatic residues located on the flat surface of CBM1 are shown in bold. Phe 42 and Tyr 52 of Xyn10B CBM1 are shown in boxes.

Table 4-2. Adsorption parameters of the wild-type and mutated CBM1s to microcrystalline cellulose and other CBM adsorption parameters from several sources

CBMs	Organism [a]	CBM family	A_{max} (μmol/g)	K_{ad} (l/μmol)	Ref.
CBM_{Xyn10B}-GFP	*I. lacteus*	1	7.8	2.0	This study
CBM_{F42S}-GFP	*I. lacteus*	1	7.9	1.8	This study
CBM_{Y52S}-GFP	*I. lacteus*	1	4.4	1.5	This study
CBM_{ThEG1}-GFP	*T. hirsuta*	1	3.4	1.2	This study
CBH I [b]	*T. reesei*	1	1.4	3.3	11
CBH II [b]	*T. reesei*	1	0.73	1.2	11
EG II [b]	*T. reesei*	1	3.2	9.0	12
CipA	*C. thermocellum*	3	0.6	0.7	13
Xylanase A	*P. fluorescens*	10	2.7	6.7	13

[a] *C. thermocellum*, *Clostridium thermocellum*; *P. fluorescens*, *Pseudomonas fluorescens*.
[b] CBM1s fused with catalytic domain.

Table 4-3. Distribution of CBM1s that have the W-YYY motif in GH family enzymes and other proteins related to cellulose degradation found in the CAZy database (http://www.cazy.org/)

Enzyme [a]	GH family	Number of W-YYY motif CBM1 genes [b]
cellulases		
endoglucanase	5, 7, 45	9
CBH I	7	2
CBH II	6	3
hemicellulases		
xylanase	10	12
mannanase	5, 26	7
xyloglucanase	74	4
Related enzymes		
CDH	-	5
esterase	-	2
others	-	11
Total		55

[a] CDH, cellobiose dehydrogenase; CBH, cellobiohydrolase

[b] Genbank accession numbers of each enzyme are: endoglucanase: BAB62317, BAD67544, CAP99211, AFV30215, CCA67360, AEO67421, CCA40496, CAY71902, AFG25592; xylanase: BAN09067, AEN99941, CAC18990, ADX07328, CCT68268, CCT67762, AAT84257, AAA65588, BAA19220, AEO60457, CCA71122, AEO67457; mannase: AFP95336, ABG79370, AEO67907, AGH62580, ADN93457, AAA34208, CCA73405; CDH, CCD47495, EAA27355, BAD95668, ADT70777, ADT70778; xyloglucanase: CAP66717, AEO64184, AAP57752, ABH71452; CBH II, BAH59082, CCA68892, AAT64008; esterase: CCD56636, AEO68387; CBH I, CAD79778, AEO67172; others: BAI83433, CCA74446, EAU86087, CCA37660, CAY68599, CCA73931, CCA73934, AFJ54163, BAB62318, CAK46515, BAE55582.

Table 4-4. Alignment of the amino acid sequences of the W-YYY motif CBM1s of glycoside hydrolase family 10 xylanases from the CAZy database (http://www.cazy.org/)

Organism	Accession number	CBM1 sequences [a]
Acremonium cellulolyticus	BAN09067	**W**GQCGGIGWNGATT--CVSPYTCQQVNP**YY**YQCL
Chrysosporium lucknowense	AEN99941	**W**GQCGGIGWTGPTQ--CESPWTCQKLND**WY**WQCL
Coniothyrium minitans	CAC18990	**W**SQCGGNGWSGPTT--CVSGSVCSKVND**WY**FQCI
Flammulina velutipes	ADX07328	**W**AQCGGINWTGATT--CVSGYTCTFQND**WF**YQCL
Fusarium fujikuroi	CCT68268	**Y**YQCGGKNYKGPTE--CEKPFKCVEHNE**YY**FQCV
Fusarium fujikuroi	CCT67762	**W**GQCGGNGWTGPTT--CQSGLTCQKIND**WY**YQCV
Fusarium graminearum	AAT84257	**W**GQCGGTGWSGSTT--CQSGLKCEKIND**FY**YQCI
Fusarium oxysporum	AAA65588	**W**GQCGGNGWTGATT--CASGLKCEKIND**WY**YQCV
Humicola grisea	BAA19220	**W**GQCGGIGWNGPTK--CQSPWTCTRLND**WY**FQCL
Myceliophthora thermophila	AEO60457	**W**GQCGGIGWTGPTQ--CESPWTCQKLND**WY**WQCL
Piriformospora indica	CCA71122	**W**QQCGGITYSGPSHPHCVDGAICQEWNP**YY**FQCI
Thielavia terrestris	AEO67457	**W**GQCGGIGWTGATQ--CQSPYTCQKLND**WY**YQCL

[a] Gray boxes indicate identified amino acid residues. Three aromatic residues located on the flat surface of the CBM1 are shown in bold. Additional aromatic amino acids that participate to strengthen the adsorption are shown in boxes.

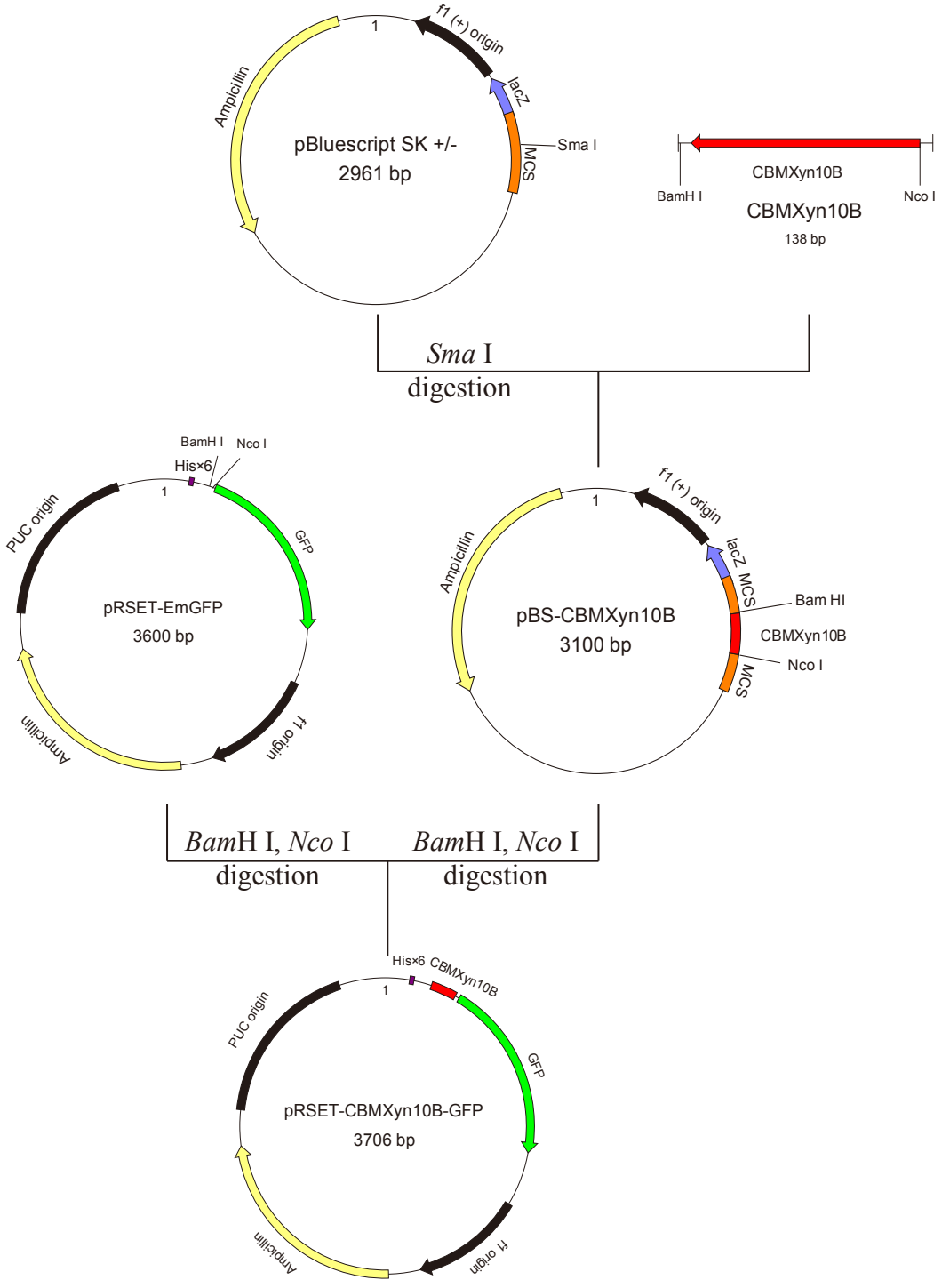


Fig. 4-1. Construction map of pBS-CBM$_{Xyn10B}$ and pRSET-CBM$_{Xyn10B}$-GFP. CBM$_{Xyn10B}$ gene was amplified from pBS-Xyn10B by PCR.

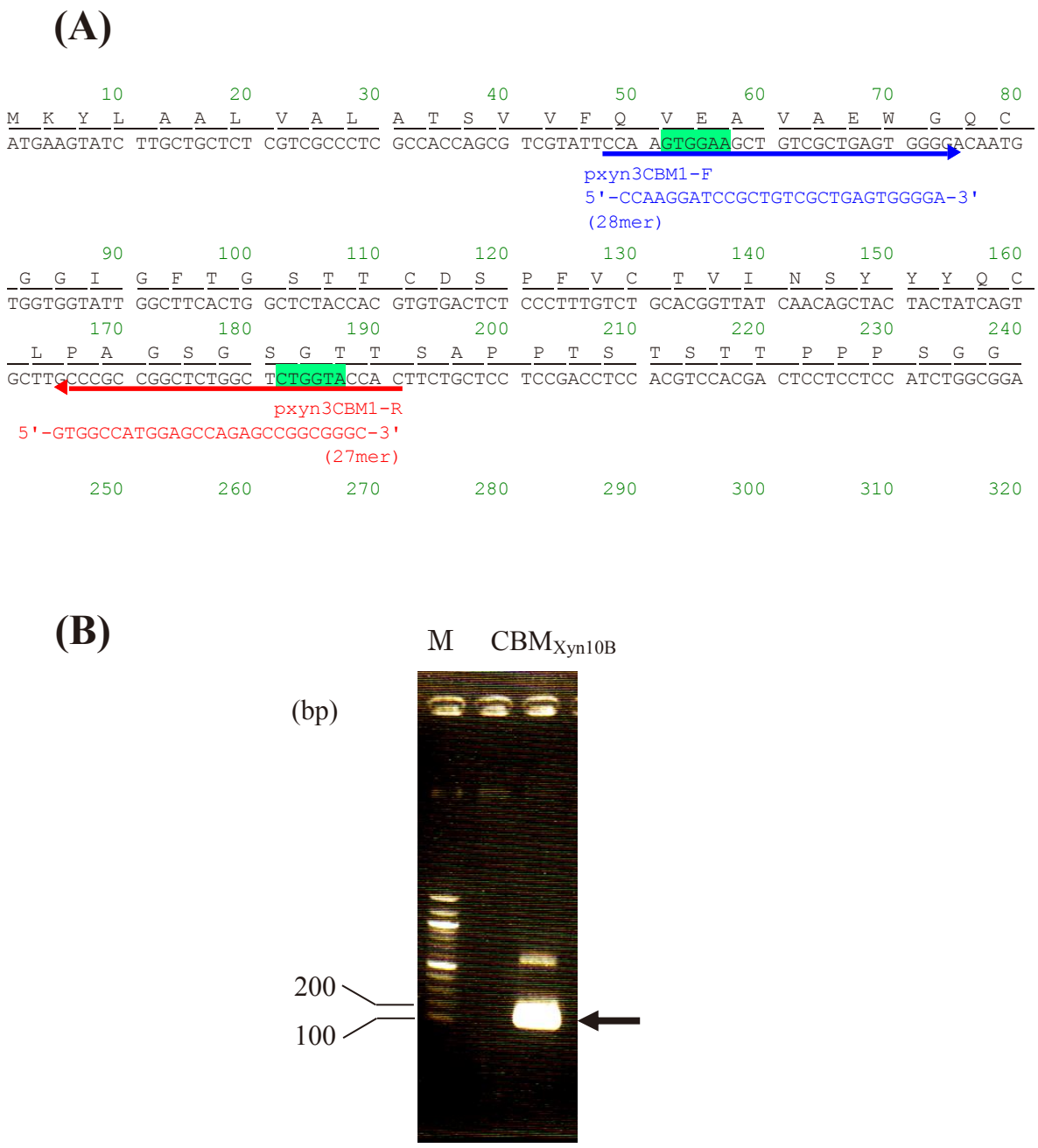


Fig. 4-2. Amplification of CBM_{Xyn10B} gene.
(A) Amplified region of CBM_{Xyn10B} gene. (B) Agarose gel electrophoresis of PCR fragment of CBM_{Xyn10B} gene amplification. M, 100 bp DNA ladder; CBM_{Xyn10B}, PCR product of CBM_{Xyn10B} gene amplification. The arrow shows the amplified CBM_{Xyn10B} gene.

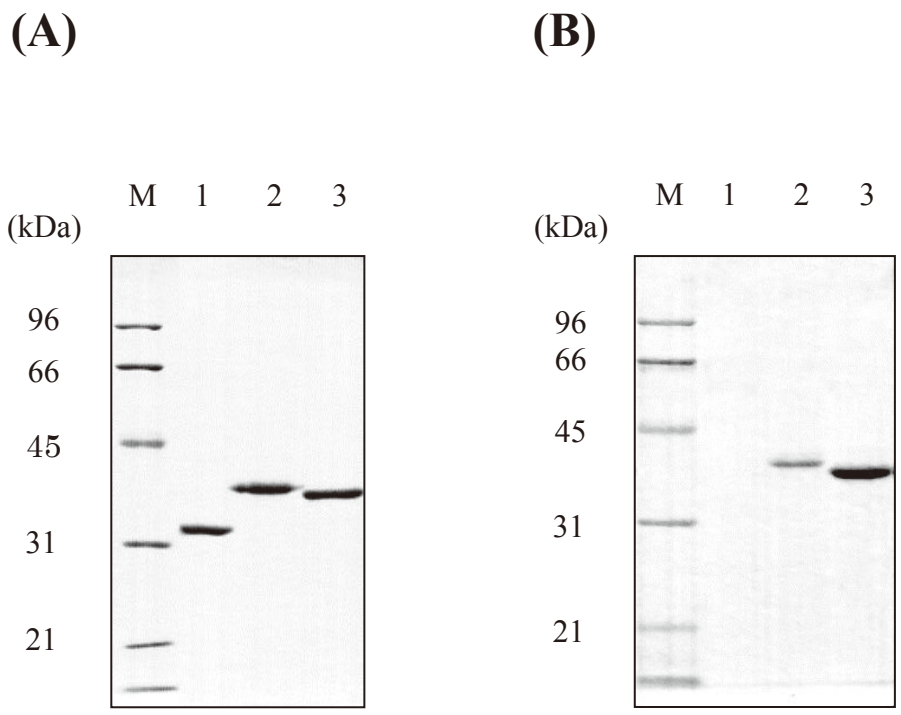


Fig. 4-3. SDS-PAGE analysis of 50 pmol purified CBM1-GFP fusion proteins (A) and their adsorption to microcrystalline cellulose (B).
M, molecular weight markers; Lane 1: GFP, Lane 2: $CBM_{\mathit{Th}EG1}$-GFP, Lane 3: CBM_{Xyn10B}-GFP. The acrylamide concentration of the gel was 12.5%.

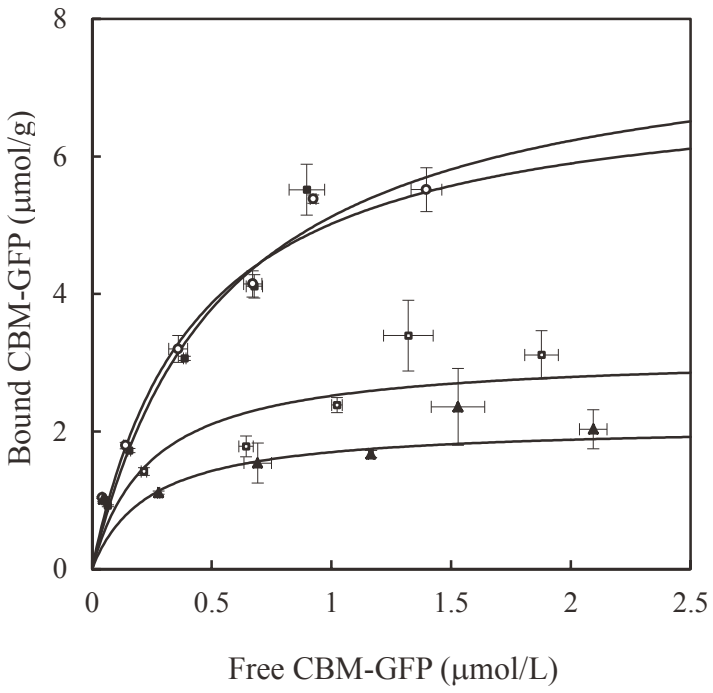


Fig. 4-4. Adsorption isotherms of the wild-type and mutated CBM1-GFPs to microcrystalline cellulose.

○, CBM_{Xyn10B}-GFP; ▲, CBM_{ThEG1}-GFP; ■, CBM_{F42S}-GFP; □, CBM_{Y52S}-GFP

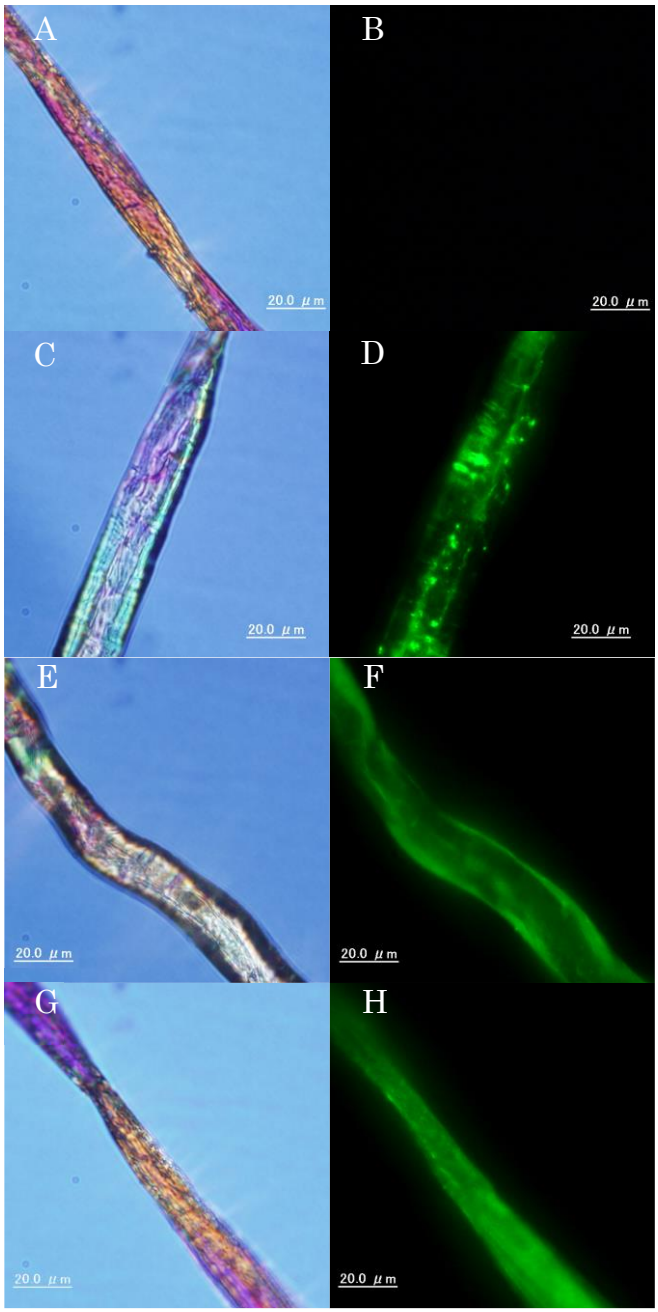


Fig. 4-5. Bright field (left) and fluorescent detection (right) image of each CBM-GFP adsorbed on cotton.
(A, B) GFP without CBM1, (C, D) CBM_{Xyn10B}-GFP, (E, F) CBM_{ThEG1}-GFP, (G, H) CBM_{Y52S}-GFP

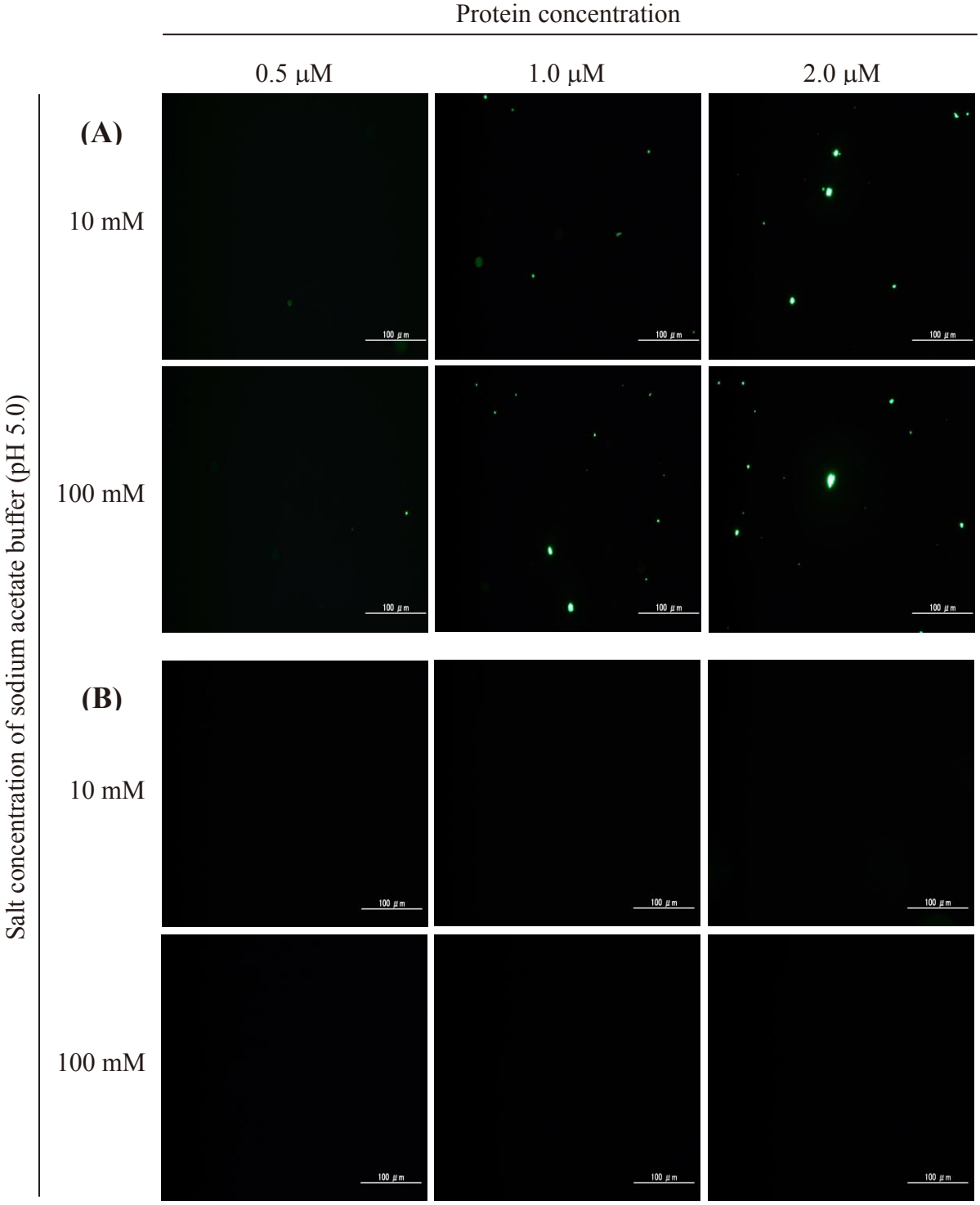


Fig. 4-6. Fluorescent detection image of (A) CBM_{Xyn10B}-GFP (aggregated) and (B) CBM_{Y52S}-GFP (non-aggregated) in the absence of cellulose.

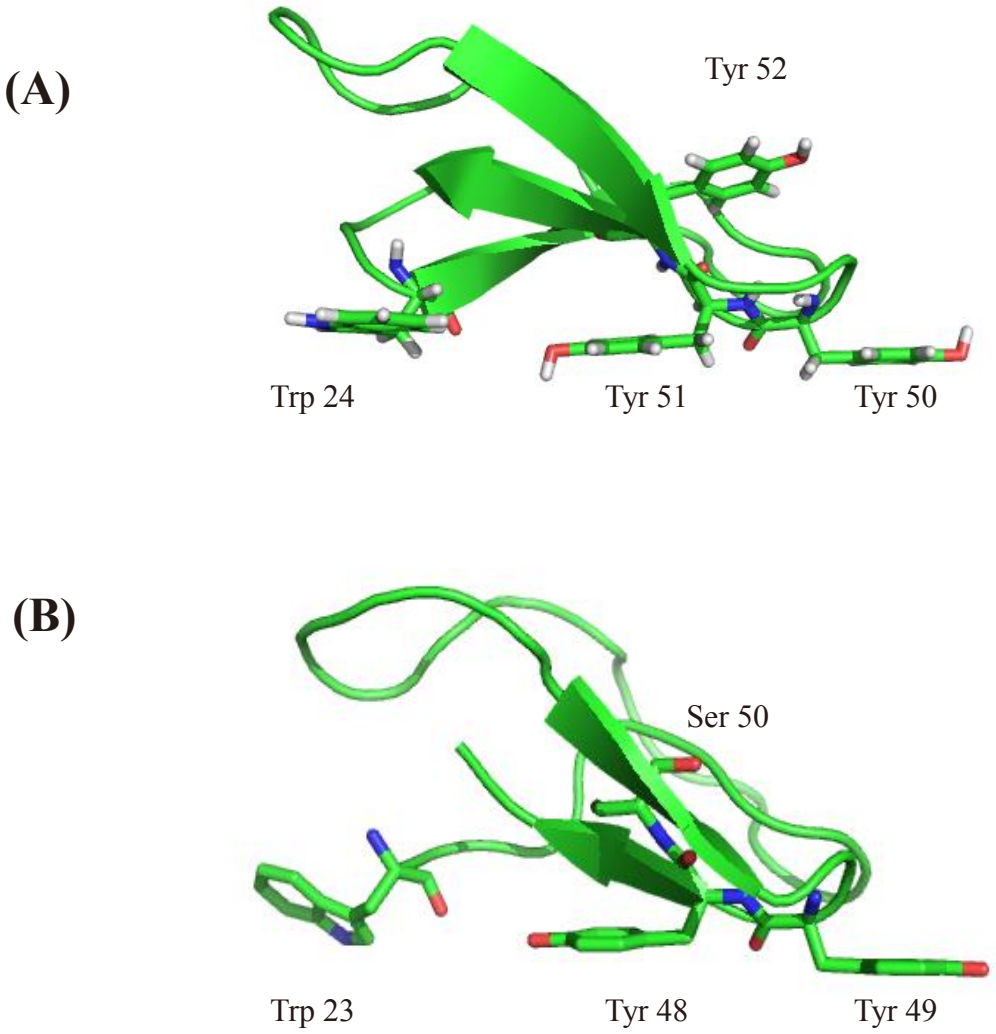


Fig. 4-7. Homology modeling of the CBM1 of Xyn10B (A) and *Th*EG1 (B).

References

1. **Laemmli UK.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680-685.
2. **Lowry OH, Rosebrough NJ, Farr AL, Randall RJ.** 1951. Protein measurement with the folin phenol reagent. J. Biol. Chem. **193:**265-275.
3. **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-β-1,4-glucanase onto crystalline cellulose. J. Appl. Glycosci. **58:**133-138.
4. **Tomme P, Boraston A, McLean B, Kormos J, Creagh AL, Sturch K, Gilkes NR, Haynes CA, Warrena RAJ,Kilburn DG.** 1998. Characterization and affinity applications of cellulose-binding domains. J. Chromatogr. B. Biomed. Sci. Appl. **715:**283-296.
5. **Medve J, Karlsson J, Lee D, Tjerneld F.** 1998. Hydrolysis of microcrystalline cellulose by cellobiohydrolase I and endoglucanase II from *Trichoderma reesei*: adsorption, sugar production pattern, and synergism of the enzymes. Biotechnol. Bioeng. **59:**621-634.
6. **Gill J, Rixon JE, Bolam DN, Mcqueen-Mason S, Simpson PJ, Williamson MP, Hazlewood GP, Gillbert HJ. Creagh AL, Ong E, Jervise E, Kilburn DG, Haynes CA.** 1999. The type II and X cellulose-binding domains of *Pseudomonas* xylanase A potentiate catalytic activity against complex substrates by a common mechanism. Biochem. J. **342:**473-480.
7. **Black GW, Rixon JE, Clarke JH, Hazlewood GP, Theodorou MK, Morris P, Gilbert H. J.** 1996. Evidence that linker sequences and cellulose-binding domains enhance the activity of hemicellulases against complex substrates. Biochem. J. **319:**515-520.
8. **Hagglund P, Eriksson T, Collen A, Nerinckx W, Claeyssens M, Stalbrand H.** 2003. A cellulose-binding module of the *Trichoderma reesei* β-mannanase Man5A increases the mannan-hydrolysis of complex substrates. J. Biotechnol. **101:**37-48.
9. **Zheng F, Ding S.** 2013. Processivity and enzymatic mode of a glycoside hydrolase family 5 endoglucanase from *Volvariella volvacea*. Appl. Environ. Microbiol. **79:**989-996.
10. **Linder M, Salovuori I, Ruohonen L, Teeri TT.** 1996. Characterization of a double cellulose-binding domain. synergistic high affinity binding to crystalline cellulose. J. Biol. Chem. **271:**21628-21272.
11. **Medve J, Ståhlberg J, Tjerneld F.** 1997. Isotherms for adsorption of cellobiohydrolase I and II from *Trichoderma reesei* on microcrystalline cellulose. Appl. Biochem. Biotechnol. **66:**39-56.
12. **Medve J, Karlsson J, Lee D, Tjerneld F.** 1998. Hydrolysis of microcrystalline cellulose by cellobiohydrolase I and endoglucanase II from *Trichoderma reesei*: adsorption, sugar production pattern, and synergism of the enzymes. Biotechnol. Bioeng. **59:**621-634.
13. **Georgelisa N, Yennawar NH, Cosgrove DJ.** 2012. Structural basis for entropy-driven

cellulose binding by a type-A cellulose-binding module (CBM) and bacterial expansin. Proc. Natl. Acad. Sci. USA. **109:**14830-14835.

Chapter 5

General conclusions

Chapter 5: General conclusions

Carbohydrate-binding module (CBM) has been recognized as a protein unit having carbohydrate-binding activity. Several CBM1s with different amino acid sequences and different functions have been reported [1]. Genetic information revealed that at least 12 cellulolytic and hemicellulolytic enzymes from *I. lacteus* have CBM1s with different amino acid sequences. The role in biomass degradation has been thought to differ among these enzymes. Analysis of the adsorption behaviors of CBM1s from various types of enzyme has been carried out in this study.

The adsorption behavior of CBM1 of endo-glucanase from *T. hirsuta* has been analyzed under different conditions, as shown in Chapter 2. The adsorption of $CBM_{Th\text{EG1-GFP}}$ and *Th*EG1 could be controlled by regulating the concentration of sodium acetate buffer and temperature. However, this system cannot be applied for CBH-type cellulases, which have a tunnel-like structure in the active site, because their binding on cellulose is stronger than the endo-type one. The adsorption/desorption regulation of CBM1 can be easily utilized for enzyme recycling processes in bioethanol production. These results and the finding of adsorption/desorption behavior of CBM1 were applied to fractionation of proteins in Driselase from *I. lacteus*.

The cellulose-binding proteins in Driselase from *I. lacteus* have also been discussed in Chapter 3. At least 12 enzymes containing CBM1 were found from *I. lacteus* draft genome sequences (**Table 3-1**). Although both Xyn10B and *Tr*XYN III belong to the GH family 10 xylanases, the specific activity of Xyn10B on soluble xylan was shown to be lower than that of *Tr*XYN III. In the case of endo-type cellulases, it was reported that the presence of CBM1 in enzymes decreases the hydrolysis activity for soluble substrates [2-3]. As Xyn10B from *I. lacteus* contains CBM1 in the N-terminus, it seems that the hydrolysis activity of Xyn10B on soluble xylan decreases. On the other hand, Xyn10B shows approximately 3-fold higher activity on insoluble substrates such as pretreated biomass than *Tr*XYN III, which does not contain CBM1. It was reported that the presence of CBM1 in hemicellulases could promote the hydrolysis activity on biomass degradation because CBM can accelerate susceptibility to biomass [4-5]. In addition, it was also reported that biomass degradation activity by xylanase from *Pseudomonas fluorescens* (XYLA, contains CBM2 and CBM10) promoted the increase in the levels of cellulose adsorption parameters because of CBM. It is suggested that strong

adsorption of CBM_{Xyn10B} on cellulose helps the effective biomass degradation of Xyn10B.

As described in chapter 4, as a difference in the amino acid sequence of CBM1 of Xyn10B to those of other CBM1s, an expression system of fused protein of CBM_{Xyn10B} and GFP protein (CBM_{Xyn10B}-GFP) was constructed in *E. coli* in order to determine the adsorption behavior. CBM_{ThEG1}-GFP (adsorption and desorption behavior is described in chapter 2) was also produced for use as a general control of CBM1. It is noteworthy that CBM_{Xyn10B} has an additional aromatic amino acid, Tyr 52, involved in strong adsorption as a W-YYY motif. In the case of CBM1 of CBH I from *T. reesei*, it was reported that the mutation of amino acids located on the molecular surface significantly affected cellulose binding [6]. Although the role of Y52 of CBM_{Xyn10B} is still unclear and the structure of CBM_{Xyn10B} is also not well understood, another hydrophobic face produced by an aromatic residue in CBM1 may induce the interaction of CBM1 by itself and/or interaction between CBM1 and cellulose. In fact, the CBM_{Xyn10B}-GFP could adsorb on cotton in an aggregated form, and be visualized as green fluorescence spots, as shown in this study. It was also reported that cellulose-binding domain enhanced the degradation of hemicellulose in plant biomass by xylanase and mannanase. In plant biomass, those CBM1s have difficulty binding on cellulose because the cellulose surface is poorly exposed. The aggregation of CBM_{Xyn10B} may have a role in concentrating enzymes on the poorly exposed cellulose. In fact, Xyn10 B showed high hydrolysis activity on pretreated biomass rather than *Tr*XYN III from *T. reesei* (**Fig. 3-14**).

Many families of CBM have been found in bacteria; for example, CBM2 (type A), CBM27 (type B), and CBM36 (type C) have been reported as bacterial CBMs. On the other hand, only CBM1 has been reported as a fungal CBM that can bind to only cellulose (or chitin). Bacteria use CBMs belonging to several families, which can bind to corresponding substrates of each catalytic domain. On the other hand, fungi may use several types of CBM1 that have different structures and functions suitable for corresponding substrates. CBM2b of *Cellulomonas fimi* Xylanase11A belongs to the CBM2 sub-family and it can adsorb on xylan rather than cellulose, which is different from general CBM2 [7-8]. It is possible that CBM1 can be classified into several types, which belong to sub-families.

Distribution of CBM1 motifs from several enzymes, which may relate to cellulose binding are shown in Fig. 5-1. 81 CBMs from CAZy database (http://www.cazy.org) are classified by amino acid sequence which related to adsorption ability to cellulose. W-YY and W-WY motif CBM1s are evenly found from various types of enzymes. Y-YY and Y-FY motif

CBM1s were mainly found from CBH I type cellulases. On the other hand, CBM1s which contains additional S-S bond are mainly found from CBH II type cellulases and swollenin. In addition, CBM1s which contains additional aromatic amino acid (such as CBM_{Xyn10B}) are found from GH family 5 endoglucanases and hemicellulases, especially xylanases. The function of CBM1 for processive enzymes is not the same as that for non-processive enzymes, as the adsorption/desorption behaviors of these types of enzyme are completely different. Processive enzyme, CBH I from *T. reesei*, contains Y-YY motif CBM1, the adsorption ability of which is lower than that of W-YY motif CBM1. This may favor the short-term adsorption required to move smoothly along the cellulose chain. On the other hand, CBH II from *T. reesei* contains an additional S-S bond in CBM1, which is associated with irreversible adsorption upon a change of salt concentration but desorption upon a temperature change. It may also be beneficial for limiting the processive action of CBH II. The third type of CBM1, which shows strong adsorption on cellulose, has been found in this study, and is mainly distributed to hemicellulases such as xylanases, mannanases, and acetylxylan esterases. Interestingly, Sugimoto reported that CBM1 of GH family 18 chitinase has a unique W-WW-SS motif [9]. This motif of CBM1 has the potential to cause high chitin adsorption compared with other CBM1s. From the above discussion, the author suggests that CBM1 should be classified into at least three types, although there are not major differences between them in terms of the amino acid sequences. The relations between mode of action of enzymes and adsorption ability of CBM1s are graphically represented in Fig. 5-2 as considered from variety of CBM1s sequences from fungi. Conceivably, lower adsorption ability of Y-YY and Y-FY motif CBM1s are adapted for CBH I type cellulase (processive enzymes), On the other hands, CBM1s which contains an additional S-S bond are used for CBH II type cellulases (endoprocessive enzymes) and swollenin as limited and irreversible adsorption on cellulose is suitable for these proteins. Furthermore CBM1s which have W-YYY motif are found from hemicellulases and endoglucanases (nonprocessive enzymes) as they need to stay on the surface of cellulosic biomass.

In conclusion, the author has found unique fungal CBM_{Xyn10B} from *I. lacteus* and also found that an extra tyrosine (Tyr 52) can enhance the binding ability of CBM_{Xyn10B} on cellulose. Compared with typical CBM1 from *T. reesei*, CBM_{Xyn10B} showed strong adsorption on cellulose with aggregation. From the comparison of structure and function relations of CBM1s, it is proposed that enzyme can adopt the suitable CBM1 among the several types of

CBM1s. Finally, it is anticipated that the strong cellulose binding ability of CBM_{Xyn10B} can be applied to a unique strategy of a biomass degradation system by fungi and also for protein immobilization [8] [10]. However, to achieve these applications, the relationship between the precise structure of CBM_{Xyn10B} and the mechanism of adsorption on cellulose and other substrates should be resolved.

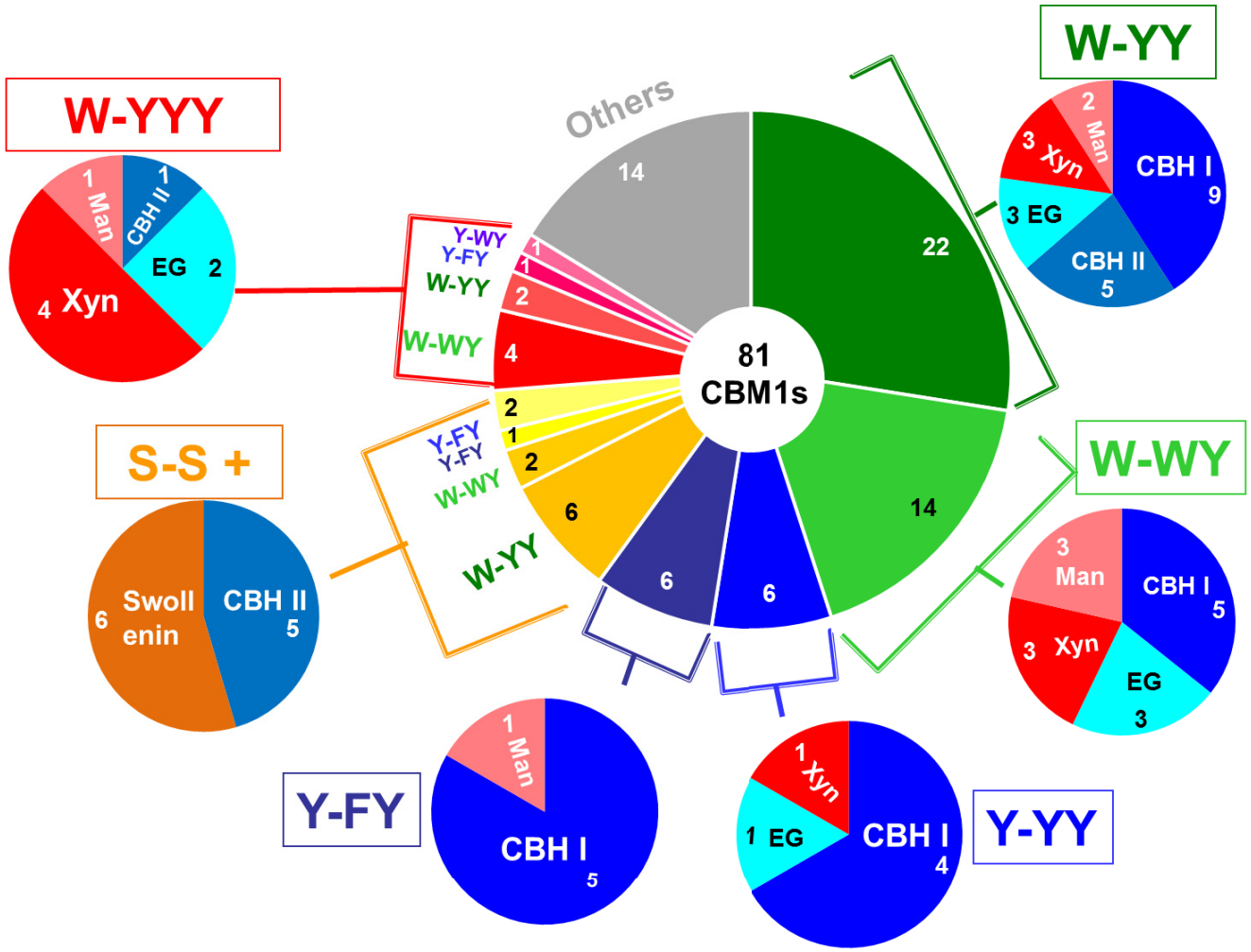


Fig. 5-1. Distribution of CBM1 motifs from several enzymes.
CBM1s from CBH I, CBH II, EG, xylanase, mannanase and swollenin (total 81 CBM1) were found in the CAZy database (http://www.cazy.org). CBM1 motifs were classified by construction of three aromatic amino acids on flat surface (W-YY, W-WY, Y-YY, and Y-FY motifs), contains an additional S-S bond (S-S +), and contains an additional aromatic amino acid (W-YYY).

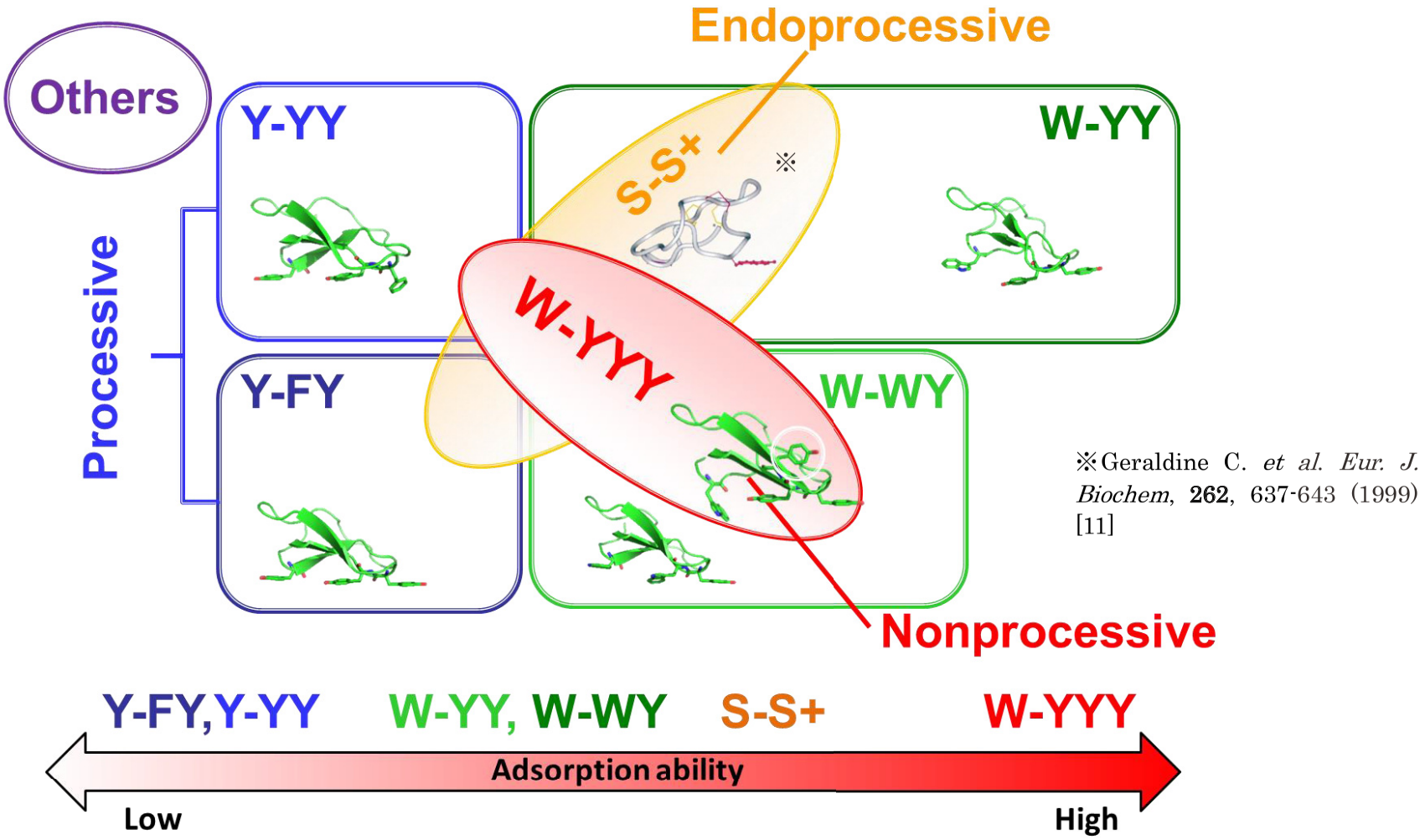


Fig. 5-2. Schematic diagram of grouping of CBM1 motifs based on enzymatic action manner and adsorption ability of CBM1s.

References

1. **Takashima S, Ohno M, Hidaka M, Nakamura A, Masaki H, Uozumi T.** 2007. Correlation between cellulose binding and activity of cellulose-binding domain mutants of *Humicola grisea* cellobiohydrolase 1. FEBS letters. **581:**5891-5896.
2. **Nozaki K, Seki T, Matsui K, Mizuno M, Kanda T, Amano Y.** 2007. Structure and Characteristics of an Endo-β-1,4-glucanase, Isolated from *Trametes hirsuta*, with High Degradation to Crystalline Cellulose. Biosci. Biotechnol. Biochem. **71:**2375-2382.
3. **Toda H, Takada S, Oda M, Amano Y, Kanda T, Okazaki M, Shimosaka M.** 2005. Gene Cloning of an Endoglucanase from the Basidiomycete *Irpex lacteus* and Its cDNA Expression in *Saccharomyces cerevisiae*. Biosci. Biotechnol. Biochem. **69:**1262-1269.
4. **Hagglund P, Eriksson T, Collen A, Nerinckx W, Claeyssens M, Stalbrand H.** 2003. A cellulose-binding module of the *Trichoderma reesei* β-mannanase Man5A increases the mannan-hydrolysis of complex substrates. Journal of Biotechnology. **101:**37-48.
5. **Black GW, Rixon JE, Clarke JH, Hazlewood GP, Theodorou MK, Morris P, Gilbert H. J.** 1996. Evidence that linker sequences and cellulose-binding domains enhance the activity of hemicellulases against complex substrates. Biochem. J. **319:**515-520.
6. **Linder M, Mattinen ML, Kontteli M, Lindeberg G, Ståhlberg J, Drakenberg T, Reinikainen T, Pettersson G, Annila A.** 1995. Identification of functionally important amino acids in the cellulose-binding domain of *Trichoderma reesei* cellobiohydrolase I. J. Biol. Chem. **4:**1056-1064.
7. **Bolam DN, Xie H, White P, Simpson PJ, Hancock SM, Williamson MP, Gilbert HJ.** 2001. Evidence for synergy between family 2b carbohydrate binding modules in *Cellulomonas fimi* xylanase 11A. Biochemistry. **40:**2468-2477.
8. **Pavón-Orozco P, Santiago-Hernández A, Rosengren A, Hidalgo-Lara ME, Stålbrand H.** 2012. The family II carbohydrate-binding module of xylanase CflXyn11A from Cellulomonas flavigena increases the synergy with cellulase TrCel7B from *Trichoderma reesei* during the hydrolysis of sugar cane bagasse. Bioresource Technology. **104:**622-630.
9. **Sugimoto N.** 2012. Functional analysis and application of fungal cellulose-bindingdomains. Doctoral dissertation.
10. **Richins RD, Mulchandani A, Chen W.** 2000. Expression, immobilization, and enzymatic characterization of cellulose-binding domain-organophosphorus hydrolase fusion enzymes. Biotechnol Bioeng. **69:**591-596.
11. **Geraldine C, Markus L.** 1999. Widely different off rates of two closely related cellulose-binding domains from *Trichoderma reesei*. Eur. J. Biochem. **262:**637-643.

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Prof. Yoshihiko Amano, Department of Material Engineering and Chemistry, Faculty of Engineering, Shinshu University, who gave me the opportunity to pursue this study, for his kindest and timely guidance, valuable suggestions and discussions throughout this study.

I am indebt to Dr. Masayuki Nozue, Dr. Masahiro Nogawa, professor and associate professor in the Department of Applied Biology, Faculty of Textile Science and Technology, Shinshu University, and Dr. Satoshi Kaneko, professor in the Department of Subtropical Biochemistry and Biotechnology, Faculty of Agriculture, University of Ryukyus, for their reviewing this thesis and providing useful suggestion and advice to improve this thesis.

I deeply grateful to Dr. Kouichi Nozaki, associate professor in the Department of Material Engineering and Chemistry, Faculty of engineering, Shinshu University, whose enormous support and insightful comments were invaluable during the course of my study. I learned from him many techniques to perform the research. I was taught by him a lot about research. I also very thanks to about using his paper for my graduation requirement.

I would like to thank Dr. Masahiro Mizuno, assistant professor in the Department of Material Engineering and Chemistry, Faculty of engineering, Shinshu University, who gave me invaluable comments for research, his experiments. His corrections of this thesis were very helpful for me.

I would like to express my gratitude to Dr. Peter Biely, Institute of Chemistry, Slovak Academy of Science, for his valuable and helpful support and discussions. He visited our laboratory during his work and gave me a lot of advice, suggestions and investigative techniques. His corrections of my paper were very helpful for me.

I also would like to thank Ogasawara laboratory, Nagaoka University of Technology and Shimosaka laboratory, Shinshu University about kindly provided recombinant *Tr*XYHN III and rEx-1 expression system in *A. oryzae*, respectively.

I appreciate to Dr. Tsutomu Arai, Dr. Naoto Fuzino, and Ms. Aki Nakano for many supports about my research. Their assists were helpful for my research.

It is a great pleasure to acknowledge Dr. Takahisa Kanda, previous professor in the Department of Material Engineering and Chemistry, Faculty of Engineering, Shinshu University, who gave me invaluable comments and warm encouragements.

This study was performed under the Development of Technology for High Efficiency Bioenergy Conversion Project supported by New Energy and Industrial Technology Development Organization (NEDO), and also supported by Grant-in-Aid for Global COE Program by Ministry of Education, Culture, Sports, Science, and Technology in JPN.

I also express my gratitude to all past and present members of Prof. Yoshihiko Amano's Laboratory, the Department of Material Engineering and Chemistry, Faculty of Engineering, Shinshu University. My study was achieved by grace of their helpful supports and encouragements.

Lastly, my special thanks to my family for their moral support and warm encouragements.

March, 2015
Hiroto Nishijima