Functional analysis and application of fungal cellulose-binding domains

（糸状菌由来セルロース結合性ドメインの機能解析と応用に関する研究）

Naohisa Sugimoto

杉本　　直久

TABLE OF CONTENTS

Chapter 1. Introduction 1

1.1 Carbohydrate-protein interaction 1

1.1.1 Carbohydrate-binding proteins 1

1.1.2 Classification of non-catalytic CBMs 2

1.2 Degradation of cellulose in fungal cellulases 4

1.2.1 Cellulose structure 4

1.2.2 Degradation of crystalline cellulose by fungal cellulases 6

1.2.2.1 General degradation system 6

1.2.2.2 Cellobiohydrolase I from *Trichoderma reesei* 8

1.2.2.3 Adsorption studies on cellulases 10

1.3 Cellulose-binding domains 13

1.3.1 Structural features 13

1.3.1.1 Comparison of bacterial and fungal CBD 13

1.3.1.2 Phylogenetic analysis of fungal CBDs 15

1.3.2 Physicochemical properties 18

1.3.2.2 Reversibility 18

1.3.2.3 Localization 19

1.3.2.4 Behavior of CBDs on cellulose 20

1.3.2.5 Disruption activity of cellulose 20

1.4 Aim of thesis 21

1.5 References 22

Chapter 2. Production of fusion proteins 33

2.1 Introduction 33

2.2 Materials and Methods 33

2.3 Results 37

2.4 Discussion 39

2.5 References 40

Chapter 3. Cellulose Affinity Purification 41

3.1 Introduction 41

3.2 Materials and Methods 42
3.3 Results 47
3.4 Discussion 54
3.5 References 57
Chapter 4. Adsorption characteristics 61
4.1 Introduction 61
4.2 Materials and Methods 62
4.3 Results 68
4.4 Discussion 76
4.5 References 83
Chapter 5. Conclusion 87
Acknowledgements 91
Appendix 93

LIST OF ABBREVIATIONS

CBM	carbohydrate-binding module
CBH	cellobiohydrolase
CD	catalytic core domain
CBD	cellulose-binding domain
CAZy	carbohydrate-active enzymes
RFP	red-fluorescent protein
PASC	phosphoric acid swollen cellulose
BMCC	bacterial microcrystalline cellulose
DP	degree of polymerization
FPLC	fast protein liquid chromatography
HIC	hydrophobic interaction chromatography
AFM	atomic force microscopy
TEM	transmission electron microscopy

Chapter 1. Introduction

1.1 Carbohydrate-protein interaction

1.1.1 Carbohydrate-binding proteins

Interactions of protein with carbohydrate play pivotal roles in living cells. This includes degradation of many carbohydrates by enzymes, flocculation and protein folding. Carbohydrates are stored as energy sources (glycogen, starch, etc.) or composed cell wall where form the strong foundation in many living cells (β-1,3-glucans, chitin, cellulose, etc.) [1]. Moreover, some specific proteins such as secreted protein and/or cell surface protein attach carbohydrates (glycoproteins) to maintain the protein folds [2], or cell-cell interactions in eukaryotic cells [3].

Carbohydrates are often recognized by the specific proteins with catalytic function, so-called enzymes, or by non-catalytic carbohydrate-binding proteins. Investigation on structural features of binding interaction between proteins and carbohydrates is important not only from the view of biological chemistry, but also from that of engineering. The carbohydrate-binding proteins with catalytic function represent various types of carbohydrate-active enzymes which are classified into glycoside hydrolase (GH), glycosyltransferase (GT), carbohydrate oxidoreductase, glycosylsynthase, carbohydrate esterase (CE), polysaccharide lyase and membrane transporter. Non-catalytic carbohydrate-binding proteins often exist as carbohydrate-binding modules (CBM), which direct the carbohydrate-active enzymes to the corresponded substrates. Details of carbohydrate-active enzymes and carbohydrate-

binding modules are collected in Carbohydrate-Active EnZymes (CAZy) database (http://www.cazy.org/).

1.1.2 Classification of non-catalytic CBMs

The non-catalytic carbohydrate-binding modules (CBM) of glycoside hydrolases were originally defined as cellulose-binding domains (CBDs) [4, 5]. Subsequently, the more appropriate term of CBM has been introduced to reflect the diversity of specificity for targeted carbohydrates based on amino acid sequence similarity [6, 7]. The recognitions of CBMs are expanded to crystalline cellulose, amorphous cellulose, chitin, β-1,3-glucans, β-1,3-1,4-glucans, xylan, mannan, galactan, starch and other cell-surface glycans. At present, several thousands proteins are classified into over 64 CBM families in the CAZy database [8, 9]. Many three-dimensional structures of various CBMs have been solved, which provided foundation to elucidate the binding mechanisms of these on monosaccharides and oligo/poly-saccharides. Although CBM family is based on similarity of folding pattern of peptides, this classification does not necessarily represent the specificity of their ligands. Thereby, a new approach to predict the specificity of ligands related to possession of the particular fold of CBMs have been proposed by Boraston et al. [9] and CBMs have been categorized into three types: 'surface-binding' CBMs (Type A), 'glycan-chain-binding' CBMs (Type B) and 'small-sugar-binding' CBMs (Type C) [9]. These schematic architectures are shown in Fig. 1.

Type A: The architecture of this type CBM is arguably the most distinct relative to other carbohydrate-binding proteins. While the presence of aromatic side

chains in the binding sites of CBMs are consistent on the majority of carbohydrate-binding proteins, the planar binding face is the common architecture of 'surface-binding' CBM (Type A). CBMs of family 1, 2a, 3, 5, and 10 CBMs belonging to Type A bind to the flat surfaces of the crystalline form of cellulose or chitin. Type A CBMs show lower affinity for soluble carbohydrates [10]. The binding energy between Type A CBMs and crystalline cellulose might be associated with positive entropy of the adsorption by release of water molecules from the interfaces [11].

Type B: 'Glycan-chain-binding' CBMs (Type B) are often described as groove or cleft-like shapes comprised several subsite pockets able to accommodate the individual sugar units of glycan chains. The binding affinity is dependent on the degree of polymerization (DP) of the ligand. The affinity increases up to hexamers, while the interaction is negligible on the oligomer with DP3 or less. Family 2b, 4, 6, 15, 17, 20, 22, 27, 28, 29, 34 and 36 CBMs are classified in type B, which interact with individual glycan chains. Direct hydrogen bonds between protein and carbohydrate play a critical role in the binding mechanism of Type B CBMs as well as in the stacking effects of aromatic side chain [12-14].

Type C: 'Small-sugar-binding' CBMs (Type C) has a binding ability toward mono-, di-, or tri-saccharides. The sites of Type C CBMs for binding do not accommodate the extended glycan chains as Type B CBMs. Family 9, 13, 14, 18 and 32 CBMs are categorized in Type C. These CBMs are originally called lectins.

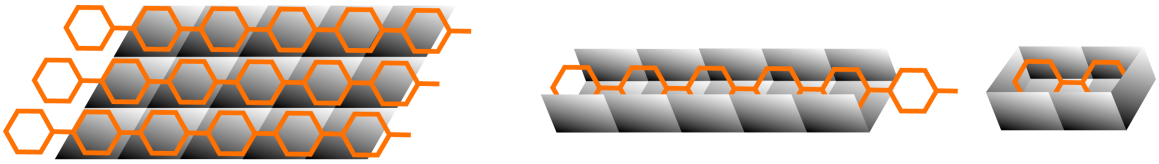


Fig.1. Schematic representation of the architecture of CBMs. Type A, suface-binding CBMs; Type B, glycan-chain-binding CBMs; and Type C, small-sugar-binding CBMs are built up the peculiar binding sites respectively.

1.2 Degradation of cellulose in fungal cellulases

1.2.1 Cellulose structure

Cellulose, the most abundant organic material in biosphere, exists in plant cell wall with coiled structure of hemicellulose and lignin. The content of cellulose in plant is 40-55 % on a dry-weight basis and the net cellulose production by plants on the earth is estimated to be more than 100 billion ton/year [15].

Cellulose is the homo-polymer of β-1,4-linked D-glucose and its degree of polymerization (DP) is estimated to be over 10,000 in the plant cell wall [16]. Though, loss of DP is observed during isolation of cellulose, leads DP of cellulose varying from a few hundred to a few thousand DP even in the best case. The pyranose ring of glucosyl residues in cellulose chain is rotated 180° relative to adjacent β-1,4-linked glucose ring, thus the minimum repeating unit of cellulose is cellobiose (Fig. 2, A). The crystalline form of cellulose in nature is called cellulose I. Each molecular chain of cellulose is stabilized with intermolecular hydrogen bonds among the adjacent molecular chains and the flat sheet structure is formed. Vertical packing of the sheets

with van der Waals interactions results in formation of highly-dense crystal structure of cellulose I (horizonal distance between the sheets is approximately 3.9 Å), in which the entire pyranose rings (hydrophobic face) are exposed on the two (110) corners (Fig. 2, B).

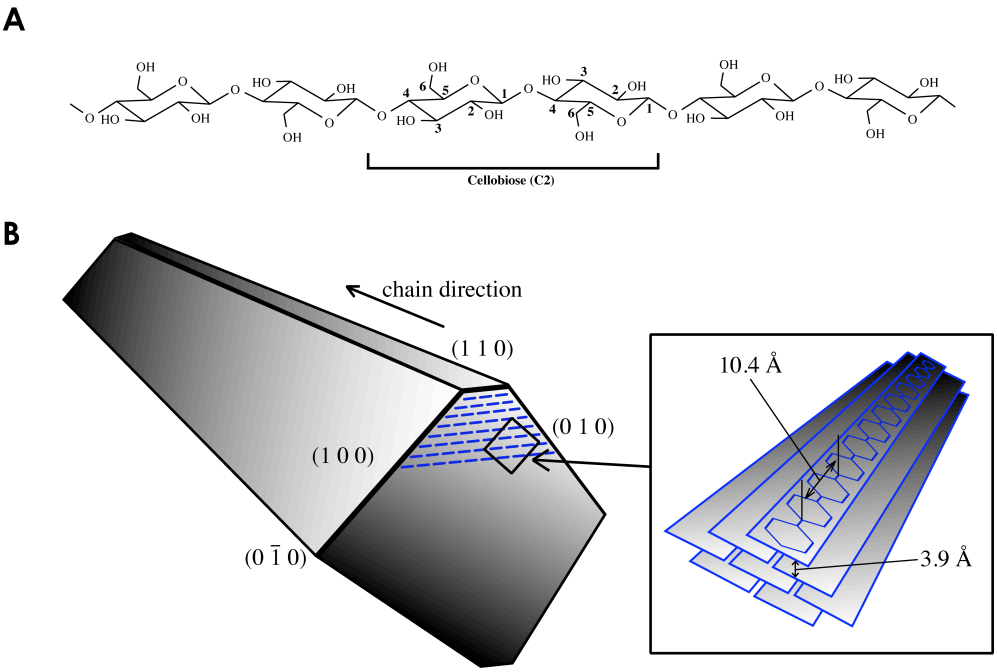


Fig. 2. Structural formula of cellulose and crystalline form of I_α. A: Structural formula of cellulose chain. The carbon numbers of two glucosyl residues are shown in the formula. Minimal repeating unit in cellulose is cellobiose (C2). B: Image of crystalline cellulose I_α. Hydrophobic face is (110) face where pyranose rings are fully exposed to solvent.

Moreover, cellulose I has been divided in the forms I_α or I_β [17], and their composite crystal are acceptable in natural cellulose. Cellulose I_α could be transformed (Fig. 3) to cellulose I_β by hydrothermal treatment over 220˚C in water [18]. Furthermore, cellulose I (I_α and I_β) is transformed to cellulose III_I by ammonia or ethylenediamine treatment, and the detailed structure of cellulose III_I has been determined [19].

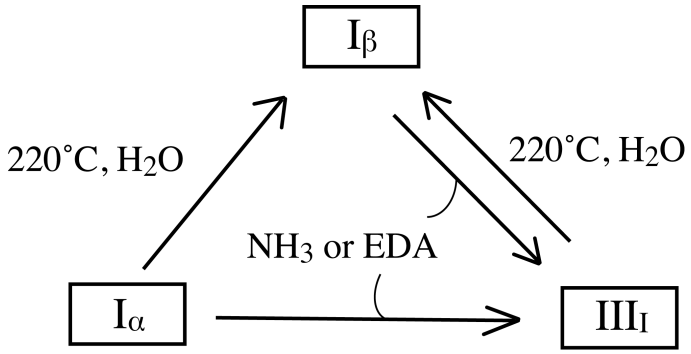


Fig. 3. Scheme of the morphological change of crystalline celluloses. NH_3 or EDA indicate the treatments by ammonia or ethylenediamine, respectively.

1.2.2 Degradation of crystalline cellulose by fungal cellulases

1.2.2.1 General degradation system

Fungal cellulases have been recognized to be effective degrading crystalline cellulose. Systematic studies on the degradation mechanism of cellulose by fungal cellulases started in 1950, by Reese's group [20]. In this report, it was postulated that native insoluble cellulose is first disrupted to shorter or linear cellulose chains without hydrolytic activity by the as called C_1 element (first step), and the shorter chains are then hydrolyzed into soluble shorter cellulose chains by the C_x element (second step). Then, soluble cellulose fragments are decomposed to glucose by *β*-glucosidase (BGL). This postulation has been called a theory of C_1-C_x elements (Fig. 4, A). However, C_1 enzyme destroys the hydrogen bonds network in crystalline cellulose (hydrogenbondase) in the C_1-C_x theory, later, instead of C_1 enzyme, cellobiohydrolases (CBHs), which have hydrolytic activities on crystalline cellulose from the chain ends,

were found [21, 22]. Thus, the endo-exo theory has been generally accepted in the mechanism of degradation of crystalline cellulose. According to this theory, endo type enzymes randomly break the non-crystalline or amorphous region, and exo type enzymes (CBHs) hydrolyze cellulose from the end of chains (Fig. 4, B). Moreover, it have been considered that endo and exo type enzymes act on crystalline cellulose together, cooperating in the effective degradation of cellulose called enzyme synergism.

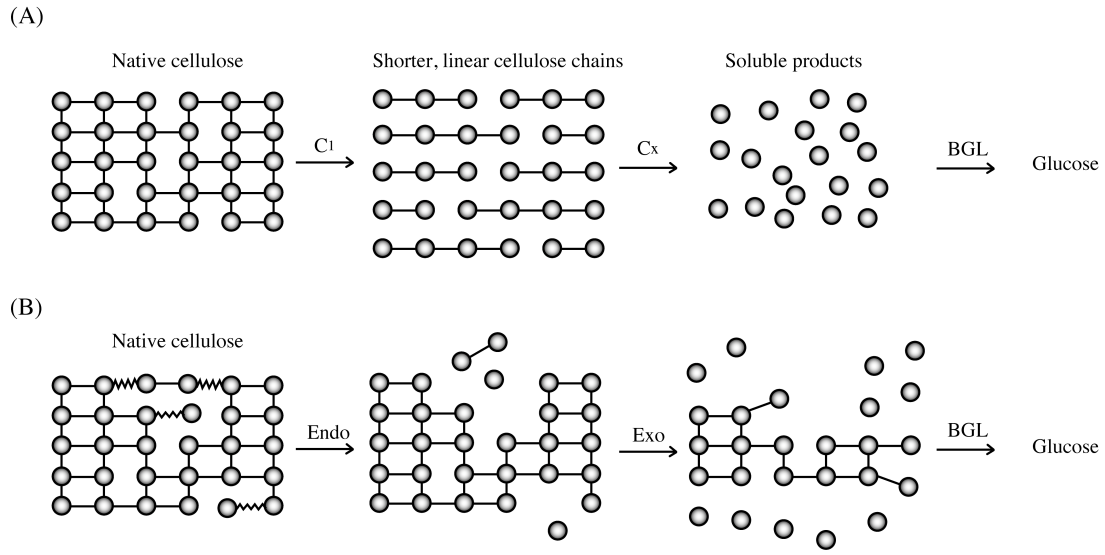


Fig. 4. The theories of C_1-C_x and endo-exo in hydrolysis of crystalline cellulose.
(A) A theory of C_1-C_x. C_1 component destroys the rigid packed cellulose, producing shorter and linear cellulose chains. C_x component hydrolyzes the glycosidic bonds in cellulose chains. (B) A theory of endo-exo. Endo-type enzymes hydrolyze amorphous regions of native cellulose and brake cellulose randomly. Exo-type cellulases hydrolyze the native cellulose from its chain ends. Co-operation with endo and exo type cellulases leads to an effective degradation of crystalline cellulose.

1.2.2.2 Cellobiohydrolase I from Trichoderma reesei

Cellobiohydrolase I (CBHI) is the major secreted protein in cellulolytic fungus *Trichoderma reesei* culture supernatants, it may constitute up to 60% the total secreted proteins. The CBHI from *T. reesei* (*Tr*CBHI), one of the most studied cellulase, have two distinct domains, which are the cellulose-binding domain (CBD in CBM family 1) and the catalytic core domain (CD in GH family 7), connected with *O*-linked glycosylated linker region. The CBD is important for the binding to crystalline cellulose. The lack of CBD from CBHI causes a decrease in degradability of crystalline cellulose [4, 23, 24]. The distance and/or relative orientation of CD and CBD are also precious factors for degradation of crystalline cellulose, which depends on length and arrangement of flexible linker region [25]. The long deletion removing practically whole linker of CBHI reduces degradability of crystalline cellulose even though binding properties of the enzyme are almost not altered [25]. Thus, it has been suggested that the two domains act in concert on the cellulose surface like caterpillar-like displacement of two-domains [26]. However, recent simulation based on molecular dynamics shows a different model in which stiff region of the linker was actually quite flexible [27]. Thus, *Tr*CBHI is considered an intrinsically disordered protein. Although the participation of CBD during catalytic activity of CBHI (directly movement) has not been clarified so far, detection of a single molecular CBHI using high-speed atomic force microscopy (HA-AFM) has revealed that the CD region obtained from CBHI by deletion of CBD has a similar velocity to move on crystalline cellulose as the intact CBHI [28]. Therefore, the processive motion of CBHI is essentially provided by the function of CD.

Furthermore, kinetic analysis on *Tr*CBHI corresponding to the surface density of the enzymes bound on cellulose has led to a concept of productive-binding and nonproductive-binding of cellulase [29-31]. Productive-binding and nonproductive-binding are distinguished by differences of the adsorption status of CD and CBD on cellulose, respectively (Fig. 5).

Ideally, increasing the adsorbed amount of CBHI on the substrate results in a linear increase of the production of cellobiose. However, further increase of the adsorbed amount causes a decrease of specific activity (catalytic turnover; min^{-1}) of CBHI and at a certain surface density, it reaches the maximum velocity of cellobiose production [31]. This results is considered to be a consequence of nonproductive-binding of CBD inhibiting the proper binding of CD owing to a steric interference on the surface of cellulose. On the other hand, linear relationship between the adsorption amount of the enzyme and activity (production of cellobiose/min) was observed in the result on hydrolysis on filter paper treated with both intact *Tr*CBHI or CD at 50°C [29]. Since the inhibition by nonproductive-binding might be depended on reaction temperature, further analytical experiments are needed to understand the relationship between activity and adsorbed amounts of CBHI.

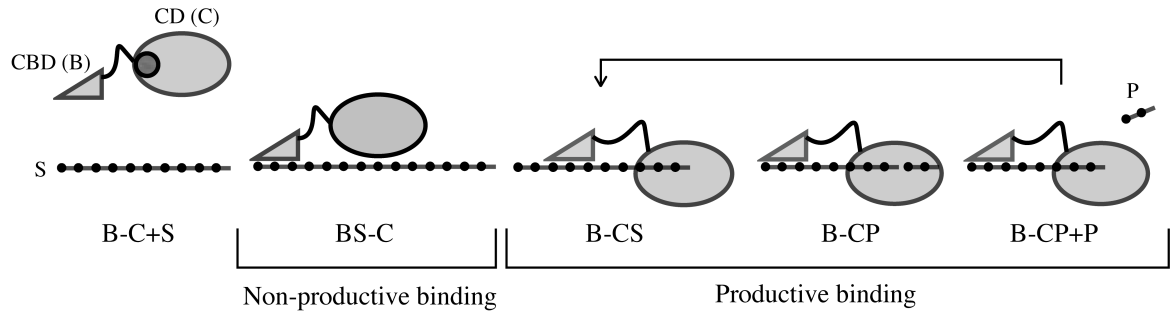


Fig. 5. Productive binding and non-productive binding of CBHI from *T. reesei*. S and P indicate cellulosic substrate and cellobiose, respectively.

1.2.2.3 Adsorption studies on cellulases

Cellulases can be adsorbed to solid crystalline cellulose by the faculty of intra-molecular CBD. Adsorption is a prerequisite for enzymatic action and therefore a significant parameter to understand the degradation of cellulose. A number of studies have been published on the adsorption of cellulolytic enzymes on different substrates. The measurement is usually conducted by depletion methods in which the amount of bound protein is estimated by subtraction from initial protein concentration to free protein concentration in the reaction mixture after incubation. A binding isotherm is constructed by plotting the amount of bound protein (mol/g substrate) on free protein concentration (mol/L).

Adsorption of cellulases to cellulose is usually observed as a heterogeneous event. Some cellulosic substrates such as Avicel are heterogeneous on the macroscopic level and contain interparticular voids or pores. Moreover, cellulose heterogeneity is determined by crystal defects, and thus the amount of chain ends and amorphous areas vary depending on cellulose preparation. Some heterogeneity is also brought about by two-domain structure of the cellulases, which allows many different binding modes

involving one domain or both domains simultaneously (productive binding and non-productive binding; see section 1.2.2.2).

The heterogeneity of cellulose-cellulase interaction is tested by the concaveness of the Scatchard plot derived from the binding isotherm. Concaved Scatchard plot indicates that the adsorption does not obey the simple Langmuir-type adsorption involving only a single-binding site with fixed affinity. Stålberg and co-workers reported that binding data between *Tr*CBHI and Avicel show the concaveness [23], and applied the binding data to two-binding site model of the Langmuir-type adsorption (Table 1). The authors interpreted that fitting to this model is derived from two-domain structure of CBHI and/or also substrate heterogeneity. Thus, the quantitative interpretation of two-binding modes of intact CBHI is discussed only when the heterogeneity is vanishingly small. [31, 32].

However the above-mentioned binding analysis (two-binding site model), many researchers have used ‘‘relative affinity’’ to estimate the affinities of cellulases to celluloses due to the difficulty of interpretation. When the concentration of free protein is very small on the Langmuir’s isotherm, the denominator in the equation becomes near 1, and the adsorption isotherm should be a straight line. The initial slope of the binding isotherm can be used to estimate the relative affinity even in more complex binding interactions. Furthermore, a new approach to extend the relative affinity was presented by Gilkes and co-workers [33]. According to the model, the cellulose surface is considered as two-dimensional lattice of cellobiose units (C2). The size of a cellulase molecule exceeds the size of one cellobiose unit and consequently the potential binding site overlap. Thus the binding depends on the number of protein

molecules bound and on their distribution on the surface. At very low surface coverage, one enzyme molecule does not prevent the binding of a second molecule and thus no precise information is needed on the size, configuration or protein-protein interactions of the cellulase molecules on the cellulose surface. This model shown in Table 1 enables calculation of relative affinity (K_r) and estimation of the number of C2 lattices occupied by one adsorbate (α). This general idea about two-dimensional overlapping binding-site is confirmed by computer simulation involving size and shape of adsorbate [34].

Table 1. Adsorption models used in cellulase-cellulose interaction

Adsorption model	Expression	Discription
Langmuir (one-site model)	$[B] = \frac{A_{max} \times [F]}{K_{ad}^{-1} + [F]}$	A_{max}, maximum adsorption capacity (moles/g-substrate); $[F]$, free protein concentration (M); K_{ad}, affinity constant (M^{-1})
Langmuir (two-site model)	$[B] = \frac{A_1 \times [F]}{K_{ad1}^{-1} + [F]} + \frac{A_2 \times [F]}{K_{ad2}^{-1} + [F]}$ $(A_{max} = A_1 + A_2)$	A_1, adsorption capacity for site 1 (moles/g-substrate); A_2, adsorption capacity for site 2 (moles/g-substrate); A_{max}, maximum adsorption capacity (moles/g-substrate); $[F]$, free protein concentration (M); K_{ad1}, affinity constant for site 1 (M^{-1}); K_{ad2}, affinity constant for site 2 (M^{-1})
Gilkes' s overlapping site model	$[B] = \frac{[N_0] \times [F]}{K_{ad}^{-1} + \alpha[F]}$ $(K_r = [N_0] \times K_{ad})$	$[N_0]$, capacity of available binding sites (moles/g-substrate); $[F]$, free protein concentration (M); K_{ad}, affinity constant for site 1 (M^{-1}); α, number of lattice units occupied by a single adsorbate; K_r, relative affinity (L/g-substrate)

1.3 Cellulose-binding domains

1.3.1 Structural features

1.3.1.1 Comparison of bacterial and fungal CBD

A functional domain binding to cellulose was discovered by limited proteolysis of *T. reesei* CBH I (*Tr*CBHI, Cel7A) in middle of the 1980s [35]. Following that, the another several domains have been also confirmed not only in fungal cellulases but also in bacterial cellulases [4, 5, 36], and the domains were called cellulose-binding domains (CBDs).

Fungal CBDs belong to carbohydrate-binding module (CBM) family 1 (CBM1). The structure of the CBD from *Tr*CBHI has been determined two decade ago based on NMR analysis [37]. The fold of this CBD composed of 36 amino acid consists of three-stranded antiparallel β-sheet, and has three tyrosine residues forming a planar face with about 10.4 Å distance each other (Fig. 6). The distance is identical to an interval of cellobiose unit in a molecular chain of cellulose. Furthermore, the structure of fungal CBDs are commonly stabilized by two or three disulfide bridges within the compact folds [38]. Deletion of a disulfide bond in the $CBD_{TrCBHII}$ from *T. reesei* CBHII causes decrease affinity to crystalline cellulose, suggesting that the rigidity of the binding planar face of CBD is a beneficial property to decrease the binding entropy loss [39].

In addition to fungal CBD, bacterial CBDs belonging to CBM2a and CBM3 are both categorized to type A CBMs [9], which are composed of approximately 100 and 150 amino acids residues, respectively. Although the protein size and arrangements of

the aromatic residues in plane face are different among the three CBDs [40], the importance of these aromatic side chains for the binding to cellulose has been demonstrated by numerous researchers [10, 40-46]. On the other hand, calorimetric analysis of binding for bacterial CBD from *Cellulomonas fimi* Cex (CBM2a, CBD_{Cex}) towards to bacterial microcrystalline cellulose (BMCC) has revealed the thermodynamic binding character. In the case of binding of CBD_{Cex} to insoluble crystalline cellulose, entropic force by dehydration between the crystalline surface of cellulose and the binding face of CBD drives the adsorption [11]. The driving force of entropic adsorption is considered to derive from aromatic residues consist of binding planar face in both bacterial and fungal CBDs. In addition to the aromatic residues, hydrogen bonds between polar amino acid residues near the hydrophobic planar face of CBDs and cellulose also contribute to the interaction in some extent [40]. The mutations of the polar residues to alanine have been shown to decline the affinity in part [43, 47].

The affinities of bacterial CBDs to any cellulosic substrates mostly show higher values than fungal CBDs [48-50]. The reason can be explained in part by the different molecular surface area. The area of binding face of CBD_{Cex} showed approximately 5.4 nm^2, which is much more than deduced from the 3D structure of fungal CBD (2.3 nm^2). Moreover, the differences between bacterial and fungal CBDs are also observed in adsorption capacities. Bacterial CBDs have much better adsorption capacities than fungal CBDs toward any celluloses, especially crystalline cellulose, whereas the binding area of one molecule may thus have much higher than fungal CBDs [49]. The difference in activity of cellulase brought in different origin of CBD was demonstrated

using exoglucanase from *C. fimi* (Cex) [48]. The fusion protein replaced bacterial CBD in intact Cex to fungal CBD (CBM1) from *Tr*CBHI had two to three times less activity than intact Cex on BMCC, and had less binding capacity. Furthermore, the fusion protein did not release small particle from cotton fibers opposite to Cex (disruption activity; see section 1.3.2.4). The paradox between protein size and binding capacities of bacterial and fungal CBDs may be explained by the disruption activity of bacterial CBDs, thus creating additional surface area for CBD binding.

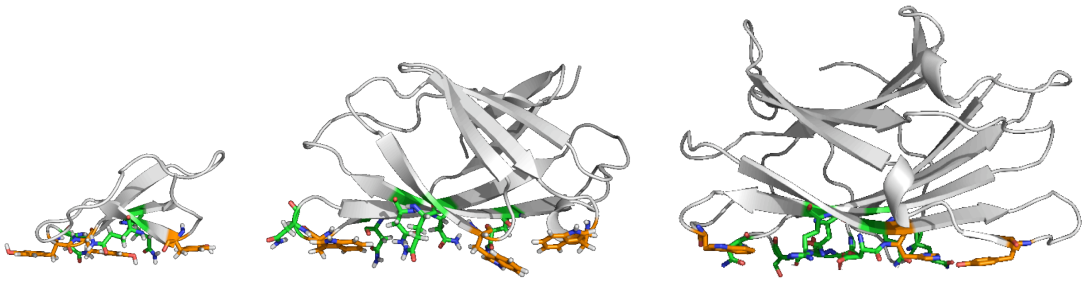


Fig. 6. Type A CBMs which bind to crystalline cellulose. Aromatic residues and polar residues in planar binding face are shown in orenge and green with stick-shape. The representatives of CBM1, CBM2, and CBM3 structures were built using PyMOL software [37, 40, 51].

1.3.1.2 Phylogenetic analysis of fungal CBDs

A lot of sequences of the CBM1s were observed in not only cellulases but also hemicellulases such as xylanase, mannanase, acetylxylan esterase as well as cellobiose dehydrogenase, chitinase etc [52]. Some of such hemicellulases were confirmed to adsorb to crystalline cellulose [53, 54]. Moreover, the presence of CBDs in plant polysaccharide hydrolysing enzymes which do not hydrolyze cellulose seems to be

common in fungi and also reported for bacterial enzymes. It has been considered that hemicelluloses are closely associated with cellulose in plant tissues [55], and the presence of CBDs in hemicellulases could help in bringing the enzyme closer to its substrate [56]. However, it has been also reported that the presence of fungal CBD did not give any significant role in the catalytic reaction [53].

The origin of this domain has been considered to predate the origin of the Dikarya, which estimated by using CAFÉ only because these domains are only about 40 amino acids in length (They contain insufficient phylogenetic information for topology-based gene tree/species tree reconciliation analyses) [57]. Therefore, I show consensus sequence logos of CBM1s and preliminary phylogenetic tree. The consensus sequence logos were drawn by use of Weblogo 2.8.2 server with respect to each preferential substrate for the catalytic domains (Fig.7, A). The phylogenetic tree of most probable sequences estimated by Weblogo was prepared by using MEGA5, and detail was shown in the figure's legend. There is no clear distinction among CBDs attached to cellulases and hemicelluloses containing GH5, 6, 7, 10, 11, 45, and 61 (Fig. 7, B), but not chitinases (GH18) and acetylxylanesterases (CE5).

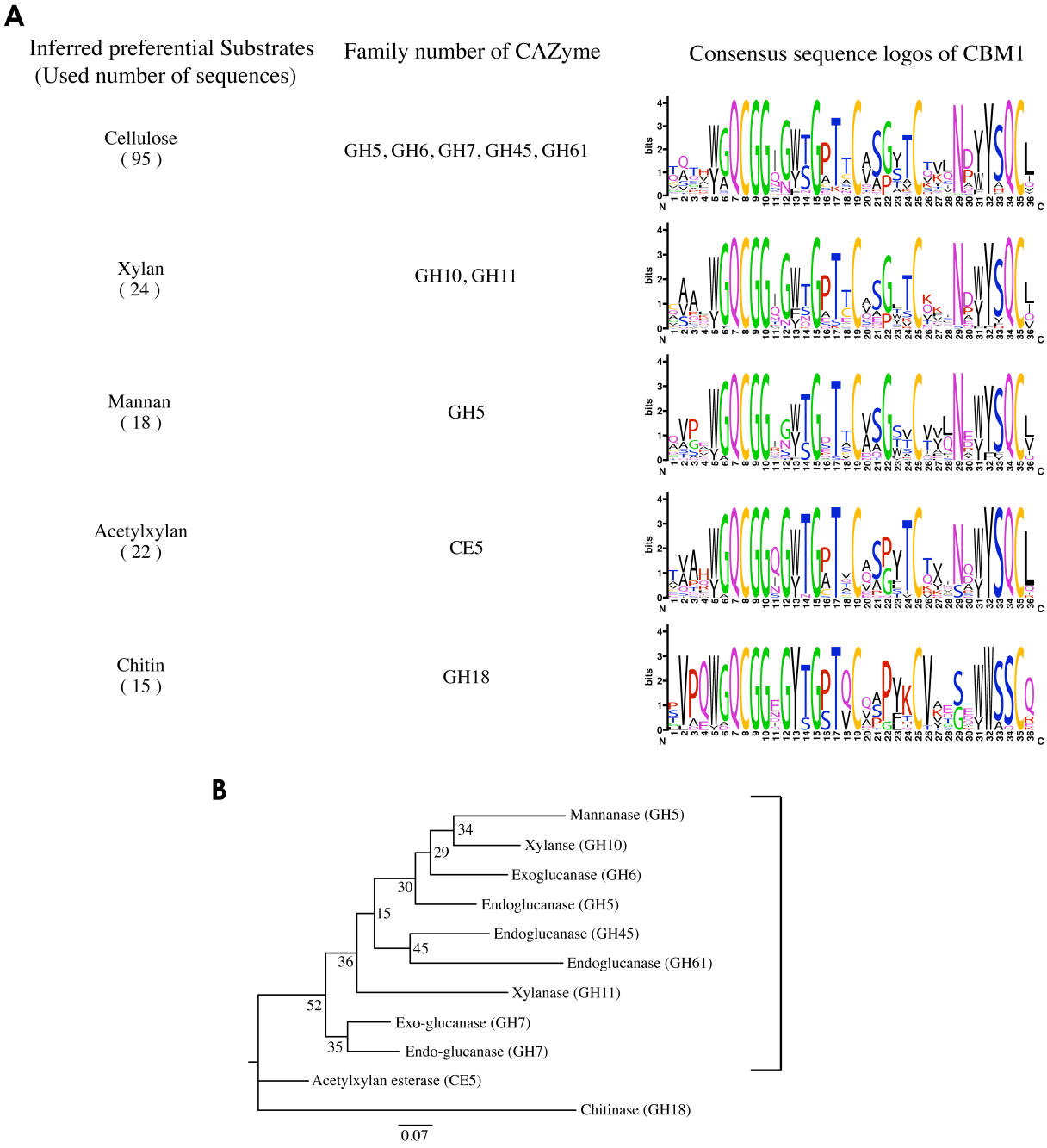


Fig. 7. Consensus sequences and phylogenetic tree of CBM1s. A: Consensus sequence logos were createed with respet to each inferred preferencial substrates for catalytic domains by using Weblogo 2.8.2. Used number of each sequence are also shown next to the substrate name. B: The evolutionary history was inferred using the Neighbor-Joining method [58]. The optimal tree with the sum of branch length = 2.98096938 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [59]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method [60] and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated. There were a total of 36 positions in the final dataset.

Evolutionary analyses were conducted in MEGA5 [61]. A clear divide of CBM1s as basis of the catalytic partners was not obtained inside of enclosure shown.

1.3.2 Physicochemical properties

1.3.2.1 Reversibility

The reversibility of binding of these CBDs has been investigated not only from the view point of carbohydrate-protein interactions but also the applications for immobilization or simplification of CBD fusion proteins on cellulosic materials. CBDs belonging to CBM family 1, 2a, 3 and 9, which have binding activities to insoluble celluloses (amorphous and crystalline celluloses), are important for the above application [62-66]. As an exploited example, immobilization of the CBD_{TrCBHI} fusion protein comprising a single-chain antibody fragment against 2-phenyloxazolone on bacterial cellulose could not be eluted by any solvent except for denaturing agents (urea, guanidine chloride, SDS or extreme pH) [38]. Moreover, the irreversible binding has been more improved by using 'double CBD', which means the connection of CBDs from *T. reesei* CBHI and CBHII with a linker peptide [67]. Besides these examples of fungal family 1 CBDs, it have been reported that CBD_{Cex} fused *Caldocellum saccharolyticum* β-glucosidase or its isolated CBD were partially eluted only with 6 M-guanidine hydrochloride [49], or CenAΔPT in which linker region is deleted from intact CenA could not be eluted from Avicel by water [68]. Thus, the binding behavior of these CBDs to crystalline cellulose seems to be 'quasi' irreversible. However, it has been confirmed that β-glucosidase harboring fungal CBD from *Phanerochaete chrysosporium* or intact *Tr*CBHI could be partially desorbed by

distilled water or elevated pH, and something by chaotropic agents (glycerol) in nature protein fold [69, 70]. As well as CBM1 CBD, it has also been reported that CBM2a and CBM3 are eluted in part by distilled water [8, 48]. Furthermore, the fact that the adsorption of the CBD_{TrCBHI} to crystalline cellulose is truly reversible in batch condition has been confirmed by Linder et al. [71]. Although the reversibility of CBDs from cellulose have been under going, these occasionally irreversible binding is considered that the intrinsic reversibly binding of CBDs tagged fusion proteins are influenced by its fusion partner in which interact to cellulose or shakiness of the whole protein.

1.3.2.2 Localization

Great interest in interaction of crystalline cellulose and CBDs is ‘‘Where do these CBDs adsorb on the crystalline face of cellulose?’’. It has been generally accepted that the planar ring-ring stacking interaction between the aromatic side chains of CBDs and pyranose rings of hydrophobic face manages the interaction. Thus, CBDs are thought to be bound to hydrophobic 110 face of crystalline cellulose, where the pyranose rings are fully exposed to solvent [38, 40, 72]. Though, it has been considered that these CBDs can also interact to hydrophilic faces of the crystal, since the binding amounts of CBD are too small when compared with the estimated surface area of the hydrophobic face of perfect crystal [43]. The experimental facts that these CBDs are almost completely bound to hydrophobic face of crystalline cellulose have been confirmed by single molecules detections using transmission electron microscopy (TEM) or total internal reflection fluorescence microscopy [73-75].

1.3.2.3 Behavior of CBDs on cellulose

The vigorous doubt whether or not these CBDs can move on cellulose surface has been investigated. The movements of bacterial family 2 CBDs, namely CBD_{Cex} and CBD_{CenA}, have been verified by means of fluorescence recovery after photobleaching (FRAP) technique. The fluorescent labelled CBDs bound on cellulose were moved to photobleaching area without clear dissociation from cellulose, indicating that these CBDs are diffused or sliding over a surface of crystalline cellulose [76]. Another demonstration comprising the diffusion of the CBDs by means of single molecule technique using quantum dots (QDs), CBD from *Acidothermus cellulolyticus* GuxA (CBM2) and CBD from *C. thermocelum* CipA (CBM3) exhibited a linear and directional motion along the cellulose crystal [77]. These findings indicate that CBDs may have the ability of sliding over a hydrophobic binding surface of cellulose.

1.3.2.4 Disruption activity of cellulose

In addition to the significant function of CBDs which leads the catalytic core domain to insoluble substrates, there has also been indicated that CBD_{CenA} acts as disruption activity of insoluble cellulose fiber (cotton) without hydrolytic activity, but not to BMCC [78]. The disruption ability of the CBD has been thought synergistic action with its catalytic domain involving hydrolysis, which called intramolecular synergism [78]. Such a disruption phenomena or dispersion of insoluble cellulose have been also observed over CBM1 CBD_{TrEGI} from *T. reesei* endoglucanase I (Cel7B), CBM2 CBD_{CenA} and CBM3 from *Thermomonospora fusca* E4 [25, 79, 80].

1.4 Aim of thesis

Fungal family 1 cellulose-binding domain (CBD) has a function to increase the local enzyme concentration on cellulose surface. The function is considered to be essential for efficient degradation of crystalline cellulose by CBHI, whereas it also provides a nonproductive-binding of CBHI on the surface of cellulose. Thus, reduction of nonproductive-binding by CBD would improve degradability of CBHI for crystalline cellulose. For this purpose, it is important to understand the binding mechanism and characteristics of CBD on crystalline cellulose.

I tried first to prepare the chimeric proteins by genetic recombination of red-fluorescent protein (RFP) and CBDs without CD of CBHs as shown in Fig. 8. In the second step, the fusion proteins were purified by affinity column chromatography of cellulose utilizing this family of CBM as affinity tags. Finally, analysis of the binding behavior of the fusion protein was conducted in order to demonstrate the adsorption kinetics on the nonproductive-binding of CBHI.

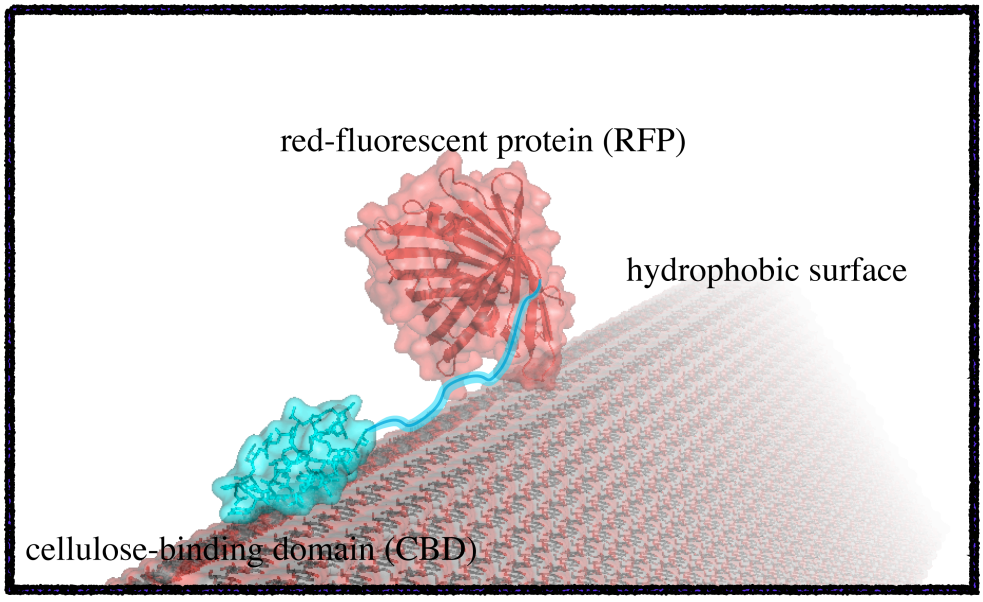


Fig. 8. Image of CBD fusion protein used in this study.

1.5 References

[1]. **Roberson MM, Armstrong PB**: Carbohydrate-binding component of amphibian embryo cell surfaces: restriction to surface regions capable of cell adhesion. *P Natl Acad Sci Usa* **(1980)**, 77:3460-3463.

[2]. **Shental-Bechor D, Levy Y**: Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *P Natl Acad Sci Usa* **(2008)**, 105:8256-8261.

[3]. **Van Mulders SE, Christianen E, Saerens SM, Daenen L, Verbelen PJ, Willaert R, Verstrepen KJ, Delvaux FR**: Phenotypic diversity of Flo protein family-mediated adhesion in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **(2009)**, 9:178-190.

[4]. **Tomme P, Van Tilbeurgh H, Pettersson G, Van Damme J, Vandekerckhove J, Knowles J, Teeri T, Claeyssens M**: Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. *Eur J Biochem* **(1988)**, 170:575-581.

[5]. **Gilkes NR, Warren RA, Miller RC, Kilburn DG**: Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. *J Biol Chem* **(1988)**, 263:10401-10407.

[6]. **Henrissat B, Popineau Y, Kader JC**: Hydrophobic cluster analysis of plant protein sequences. A domain homology between storage and lipid-transfer proteins. *Biochem J* **(1988)**, 255:901-905.

[7]. **Henrissat B, Claeyssens M, Tomme P, Lemesle L, Mornon JP**: Cellulase families revealed by hydrophobic cluster analysis. *Gene* **(1989)**, 81:83-95.

[8]. **Tomme P, Warren RAJ, Miller RC, Kilburn DG, Gilkes NR**: Cellulose-binding domains: Classification and properties. *Acs Sym Ser* (**1995**), 618:142-163.

[9]. **Boraston AB, Bolam DN, Gilbert HJ, Davies GJ**: Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* (**2004**), 382:769-781.

[10]. **Nagy T, Simpson P, Williamson MP, Hazlewood GP, Gilbert HJ, Orosz L**: All three surface tryptophans in Type IIa cellulose binding domains play a pivotal role in binding both soluble and insoluble ligands. *Febs Lett* **(1998)**, 429:312-316.

[11]. **Creagh AL, Ong E, Jervis E, Kilburn DG, Haynes CA**: Binding of the cellulose-binding domain of exoglucanase Cex from *Cellulomonas fimi* to insoluble microcrystalline cellulose is entropically driven. *P Natl Acad Sci Usa* **(1996)**, 93:12229-12234.

[12]. **Simpson PJ, Xie HF, Bolam DN, Gilbert HJ, Williamson MP**: The structural basis for the ligand specificity of family 2 carbohydrate-binding modules. *J Biol Chem* (**2000**), 275:41137-41142.

[13]. **Notenboom V, Boraston AB, Chiu P, Freelove ACJ, Kilburn DG, Rose DR**: Recognition of cello-oligosaccharides by a family 17 carbohydrate-binding module: An X-ray crystallographic, thermodynamic and mutagenic study. *J Mol Biol* (**2001**), 314:797-806.

[14]. **Pell G, Williamson MP, Walters C, Du H, Gilbert HJ, Bolam DN**: Importance of hydrophobic and polar residues in ligand binding in the family 15 carbohydrate-binding module from *Cellvibrio japonicus* Xyn10C. *Biochemistry-Us* **(2003)**, 42:9316-9323.

[15]. **Research Council (U.S.). Committee on Mineral Resources and the Environment N**: Mineral resources and the environment, supplementary report: Resource **(1975)**:416.

[16]. **Sjöström E**: Wood chemistry: fundamentals and applications. **(1993)**:293.

[17]. **Atalla RH, Vanderhart DL**: Native cellulose: a composite of two distinct crystalline forms. *Science* **(1984)**, 223:283-285.

[18]. **Yamamoto H, Horii F, Odani H**: Structural changes of native cellulose crystals induced by annealing in aqueous alkaline and acidic solutions at high temperatures. *Macromolecules* (**1989**), 22:4130-4132.

[19]. **Wada M, Chanzy H, Nishiyama Y, Langan P**: Cellulose III_I crystal structure and hydrogen bonding by synchrotron X-ray and neutron fiber diffraction. *Macromolecules* **(2004)**, 37:8548-8555.

[20]. **Reese ET, Siu RGH, Levinson HS**: The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. *J Bacteriol* (**1950**), 59:485-497.

[21]. **Halliwell G, Griffin M, Vincent R**: The role of component C1 in cellulolytic systems. *Biochem J* **(1972)**, 127:43P.

[22]. **Wood TM, McCrae SI**: Purification and properties of C1 component of *Trichoderma koningii* cellulase. *Biochem J* (**1972**), 128:1183-&.

[23]. **Ståhlberg J, Johansson G, Pettersson G**: A new model for enzymatic hydrolysis of cellulose based on the two-domain structure of cellobiohydrolase I. *Bio/Technol* (**1991**), 9:286-290.

[24]. **Nidetzky B, Steiner W, Hayn M, Claeyssens M**: Cellulose hydrolysis by the cellulases from *Trichoderma reesei*: A new model for synergistic interaction. *Biochem J* (**1994**), 298:705-710.

[25]. **Srisodsuk M, Reinikainen T, Penttilä M, Teeri TT**: Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase I in its interaction with crystalline cellulose. *J Biol Chem* (**1993**), 268:20756-20761.

[26]. **Receveur V, Czjzek M, Schülein M, Panine P, Henrissat B**: Dimension, shape, and conformational flexibility of a two domain fungal cellulase in solution probed by small angle X-ray scattering. *J Biol Chem* (**2002**), 277:40887-40892.

[27]. **Beckham GT, Bomble YJ, Matthews JF, Taylor CB, Resch MG, Yarbrough JM, Decker SR, Bu L, Zhao X, McCabe C *et al***: The *O*-glycosylated linker from the *Trichoderma reesei* Family 7 cellulase is a flexible, disordered protein. *Biophys J* (**2010**), 99:3773-3781.

[28]. **Igarashi K, Koivula A, Wada M, Kimura S, Penttila M, Samejima M**: High-speed atomic force microscopy visualizes processive movement of *Trichoderma reesei* cellobiohydrolase I on crystalline cellulose. *J Biol Chem* (**2009**), 284:36186-36190.

[29]. **Nidetzky B, Steiner W, Claeyssens M**: Cellulose hydrolysis by the cellulases from *Trichoderma reesei*: Adsorptions of 2 cellobiohydrolases, 2 endocellulases and their core proteins on filter paper and their relation to hydrolysis. *Biochem J* (**1994**), 303:817-823.

[30]. **Väljamäe P, Sild V, Pettersson G, Johansson G**: The initial kinetics of hydrolysis by cellobiohydrolases I and II is consistent with a cellulose surface-erosion model. *Eur J Biochem* (**1998**), 253:469-475.

[31]. **Igarashi K, Wada M, Hori R, Samejima M**: Surface density of cellobiohydrolase on crystalline celluloses. A critical parameter to evaluate enzymatic kinetics at a solid-liquid interface. *Febs J* (**2006**), 273:2869-2878.

[32]. **Igarashi K, Wada M, Samejima M**: Activation of crystalline cellulose to cellulose III_I results in efficient hydrolysis by cellobiohydrolase. *Febs J* **(2007)**, 274:1785-1792.

[33]. **Gilkes NR, Jervis E, Henrissat B, Tekant B, Miller RC, Warren RA, Kilburn DG**: The adsorption of a bacterial cellulase and its two isolated domains to crystalline cellulose. *J Biol Chem* **(1992)**, 267:6743-6749.

[34]. **Sild V, Ståhlberg J, Pettersson G, Johansson G**: Effect of potential binding site overlap to binding of cellulase to cellulose: a two-dimensional simulation. *Febs Lett* **(1996)**, 378:51-56.

[35]. **Vantilbeurgh H, Tomme P, Claeyssens M, Bhikhabhai R, Pettersson G**: Limited proteilysis of the cellobiohydrolase I from *Trichoderma ressei*. Separation of functional domains. *Febs Lett* (**1986**), 204:223-227.

[36]. **Langsford ML, Gilkes NR, Singh B, Moser B, Miller RC, Warren RA, Kilburn DG**: Glycosylation of bacterial cellulases prevents proteolytic cleavage between functional domains. *Febs Lett* **(1987)**, 225:163-167.

[37]. **Kraulis J, Clore GM, Nilges M, Jones TA, Pettersson G, Knowles J, Gronenborn AM**: Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry-Us* **(1989)**, 28:7241-7257.

[38]. **Mattinen ML, Kontteli M, Kerovuo J, Linder M, Annila A, Lindeberg G, Reinikainen T, Drakenberg T**: Three-dimensional structures of three engineered cellulose-binding domains of cellobiohydrolase I from *Trichoderma reesei*. *Protein Sci* **(1997)**, 6:294-303.

[39]. **Carrard G, Linder M**: Widely different off rates of two closely related cellulose-binding domains from *Trichoderma reesei*. *Eur J Biochem* (**1999**), 262:637-643.

[40]. **Tormo J, Lamed R, Chirino AJ, Morag E, Bayer EA, Shoham Y, Steitz TA**: Crystal structure of a bacterial family-III cellulose-binding domain: A general mechanism for attachment to cellulose. *Embo J* (**1996**), 15:5739-5751.

[41]. **Din N, Forsythe IJ, Burtnick LD, Gilkes NR, Miller RC, Warren RAJ, Kilburn DG**: The cellulose-binding domain of endoglucanase A (CenA) from *Cellulomonas fimi*: Evidence for the involvement of tryptophan residues in binding. *Mol Microbiol* (**1994**), 11:747-755.

[42]. **Linder M, Lindeberg G, Reinikainen T, Teeri TT, Pettersson G**: The difference in affinity between two fungal cellulose-binding domains is dominated by a single amino acid substitution. *Febs Lett* (**1995**), 372:96-98.

[43]. **McLean BW, Bray MR, Boraston AB, Gilkes NR, Haynes CA, Kilburn DG**: 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* (**2000**), 13:801-809.

[44]. **Bray MR, Johnson PE, Gilkes NR, McIntosh LP, Kilburn DG, Warren RAJ**: Probing the role of tryptophan residues in a cellulose-binding domain by chemical modification. *Protein Sci* (**1996**), 5:2311-2318.

[45]. **Ponyi T, Szabo L, Nagy T, Orosz L, Simpson PJ, Williamson MP, Gilbert HJ**: Trp22, Trp24, and Tyr8 play a pivotal role in the binding of the family 10 cellulose-binding module from *Pseudomonas* xylanase A to insoluble ligands. *Biochemistry-Us* (**2000**), 39:985-991.

[46]. **Takashima S, Ohno M, Hidaka M, Nakamura A, Masaki H**: Correlation between cellulose binding and activity of cellulose-binding domain mutants of *Humicola grisea* cellobiohydrolase 1. *Febs Lett* (**2007**), 581:5891-5896.

[47]. **Linder M, Mattinen ML, Kontteli M, Lindeberg G, Stahlberg J, Drakenberg T, Reinikainen T, Pettersson G, Annila A**: Identification of

functionally important amino acids in the cellulose-binding domain of *Trichoderma reesei* cellobiohydrolase I. *Protein Sci* (**1995**), 4:1056-1064.

[48]. **Tomme P, Driver DP, Amandoron EA, Miller RC, Antony R, Warren J, Kilburn DG**: Comparison of a fungal (family I) and bacterial (family II) cellulose-binding domain. *J Bacteriol* **(1995**), 177:4356-4363.

[49]. **Reinikainen T, Takkinen K, Teeri TT**: Comparison of the adsorption properties of a single-chain antibody fragment fused to a fungal or bacterial cellulose-binding domain. *Enzyme Microb Tech* (**1997**), 20:143-149.

[50]. **Tomme P, Boraston A, McLean B, Kormos J, Creagh AL, Sturch K, Gilkes NR, Haynes CA, Warren RAJ, Kilburn DG**: Characterization and affinity applications of cellulose-binding domains. *J Chromatogr B* (**1998**), 715:283-296.

[51]. **Xu GY, Ong E, Gilkes NR, Kilburn DG, Muhandiram DR, Harrisbrandts M, Carver JP, Kay LE, Harvey TS**: Solution structure of a cellulose-binding domain from *Cellulomonas fimi* by nuclear magnetic resonance spectroscopy *Biochemistry-Us* (**1995**), 34:6993-7009.

[52]. **Yoshida M, Igarashi K, Wada M, Kaneko S, Suzuki N, Matsumura H, Nakamura N, Ohno H, Samejima M**: Characterization of carbohydrate-binding cytochrome b(562) from the white-rot fungus *Phanerochaete chrysosporium*. *Appl Environ Microb* (**2005**), 71:4548-4555.

[53]. **Margolles-Clark E, Tenkanen M, Söderlund H, Penttilä M**: Acetyl xylan esterase from *Trichoderma reesei* contains an active-site serine residue and a cellulose-binding domain. *Eur J Biochem* **(1996**), 237:553-560.

[54]. **Ademark P, Varga A, Medve J, Harjunpaa V, Drakenberg T, Tjerneld F, Stålbrand H**: Softwood hemicellulose-degrading enzymes from *Aspergillus niger*: Purification and properties of a β-mannanase. *J Biotechnol* (**1998**), 63:199-210.

[55]. **Timell TE**: Recent progress in the chemistry of wood hemicelluloses. *Wood Sci and Technol* **(1967)**, 1:45-70.

[56]. **Pham TA, Berrin JG, Record E, To KA, Sigoillot J**: Hydrolysis of softwood by *Aspergillus* mannanase: role of a carbohydrate-binding module. *J Biotechnol* **(2010)**, 148:163-170.

[57]. **Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS *et al***: The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* **(2012)**, 336:1715-1719.

[58]. **Saitou N, Nei M**: The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **(1987)**, 4:406-425.

[59]. **Felsenstein J**: Confidence limits on phylogenies-An approach using bootstrap. *Evolution* (**1985**), 39:783-791.

[60]. **Jones DT, Taylor WR, Thornton JM**: The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* (**1992**), 8:275-282.

[61]. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S**: MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol* (**2011**), 28:2731-2739.

[62]. **Ong E, Greenwood J, Gilkes NR, Kilburn DG, Miller RC, Warren RA**: The cellulose-binding domains of cellulases: tools for biotechnology. *Trends Biotechnol* (**1989**), 7:239-243.

[63]. **Ong E, Gilkes NR, Miller RC, Warren RAJ, Kilburn DG**: Enzyme immobilization using a cellulose-binding domain: Properties of a β-glucosidase fusion protein. *Enzyme Microb Tech* (**1991**), 13:59-65.

[64]. **Levy I, Shoseyov O**: Cellulose-binding domains biotechnological applications. *Biotechnol Adv* (**2002**), 20:191-213.

[65]. **Kavoosi M, Meijer J, Kwan E, Creagh AL, Kilburn DG, Haynes CA**: Inexpensive one-step purification of polypeptides expressed in *Escherichia coli* as fusions with the family 9 carbohydrate-binding module of xylanase 10A from *T. maritima*. *J Chromatogr B Analyt Technol Biomed Life Sci* (**2004**), 807:87-94.

[66]. **Hildén L, Johansson G**: Recent developments on cellulases and carbohydrate-binding modules with cellulose affinity. *Biotechnol Lett* (**2004**), 26:1683-1693.

[67]. **Linder M, Salovuori I, Ruohonen L, Teeri TT**: Characterization of a double cellulose-binding domain. Synergistic high affinity binding to crystalline cellulose. *J Biol Chem* (**1996**), 271:21268-21272.

[68]. **Shen H, Schmuck M, Pilz I, Gilkes NR, Kilburn DG, Miller RC, Warren RA**: Deletion of the linker connecting the catalytic and cellulose-binding domains of endoglucanase A (CenA) of *Cellulomonas fimi* alters its conformation and catalytic activity. *J Biol Chem* (**1991**), 266:11335-11340.

[69]. **Otter DE, Munro PA, Scott GK, Geddes R**: Desorption of *Trichoderma reesei* cellulase from cellulose by a range of desorbents. *Biotechnol Bioeng* (**1989**), 34:291-298.

[70]. **Lymar ES, Li B, Renganathan V**: Purification and characterization of a cellulose-binding β-glucosidase from cellulose-degrading cultures of *Phanerochaete chrysosporium*. *Appl Environ Microb* (**1995**), 61:2976-2980.

[71]. **Linder M, Teeri TT**: The cellulose-binding domain of the major cellobiohydrolase of *Trichoderma reesei* exhibits true reversibility and a high exchange rate on crystalline cellulose. *P Natl Acad Sci Usa* (**1996**), 93:12251-12255.

[72]. **Reinikainen T, Teleman O, Teeri TT**: Effects of pH and high ionic strength on the adsorption and activity of native and mutated cellobiohydrolase I from *Trichoderma reesei*. *Proteins* (**1995**), 22:392-403.

[73]. **Lehtiö J, Sugiyama J, Gustavsson M, Fransson L, Linder M, Teeri TT**: The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *P Natl Acad Sci Usa* (**2003**), 100:484-489.

[74]. **Xu Q, Tucker MP, Arenkiel P, Ai X, Rumbles G, Sugiyama J, Himmel ME, Ding S**: Labeling the planar face of crystalline cellulose using quantum dots directed by type-I carbohydrate-binding modules. *Cellulose* (**2009**), 16:19-26.

[75]. **Dagel DJ, Liu Y, Zhong L, Luo Y, Himmel ME, Xu Q, Zeng Y, Ding S, Smith S**: In situ imaging of single carbohydrate-binding modules on cellulose microfibrils. *The journal of physical chemistry B* (**2011**), 115:635-641.

[76]. **Jervis EJ, Haynes CA, Kilburn DG**: Surface diffusion of cellulases and their isolated binding domains on cellulose. *J Biol Chem* (**1997**), 272:24016-24023.

[77]. **Liu Y, Zeng Y, Luo Y, Xu Q, Himmel ME, Smith SJ, Ding S**: Does the cellulose-binding module move on the cellulose surface? *Cellulose* (**2009**), 16:587-597.

[78]. **Din N, Damude HG, Gilkes NR, Miller RC, Warren RA, Kilburn DG**: C1-Cx revisited: intramolecular synergism in a cellulase. *P Natl Acad Sci Usa* (**1994**), 91:11383-11387.

[79]. **Gilkes NR, Kilburn DG, Miller RC, Warren RAJ, Sugiyama J, Chanzy H, Henrissat B**: Visualization of the adsorption of a bacterial endo-β-1,4-glucanase and its isolated cellulose-binding domain to crystalline cellulose. *Int J Biol Macromol* (**1993**), 15:347-351.

[80]. **Irwin D, Shin DH, Zhang S, Barr BK, Sakon J, Karplus PA, Wilson DB**: Roles of the catalytic domain and two cellulose binding domains of

Thermomonospora fusca E4 in cellulose hydrolysis. *J Bacteriol* (**1998**), 180:1709-1714.

Chapter 2. Production of fusion proteins

2.1 Introduction

Fungal cellulases often carry cellulose-binding domain (CBD) belonging to carbohydrate binding module (CBM) family 1. However, while the CBD binds to crystalline cellulose without hydrolytic activity, the catalytic domain (CD) attached to the CBD also binds to the cellulose via its substrate-binding site. Thus, in order to analyze the function of CBD precisely, fusion protein (chimeric protein) was used in which the CD was replaced by a fluorescent protein (FP) to avoid the interference of the CD binding. CBD fusion protein has an advantage because its size is similar to cellulase molecule. Moreover, it is expected that precise measurement of the protein concentration can be achieved by measuring the intensity or absorption of fluorescence originated in the fluorescent protein. Since the linker region of fungal cellulases are often having *O*-glycosylated, eukaryote expression system is needed to mimic the fungal cellulase. Therefore, I examined the production of fusion protein of the monomeric red FP (RFP) with four types of fungal CBDs at N- or C-terminus in the *Pichia pastoris* expression system.

2.2 Materials and Methods

2.2.1 Chemicals and strains

All the chemicals used in this study were laboratory grade from WaKo Pure Chemical Industries, Ltd. (Osaka, Japan) or Sigma-Aldrich (St. Louis, USA). *Escherichia coli*

strain JM109 was used as bacterial host for DNA cloning, pCR4Blunt-TOPO vector and the oligonucleotides were purchased from Invitrogen (Carlsbad, CA). For the protein expression, methylotoropic yeast *Pichia pastoris* KM71H and transforming vector and pPICZαA were purchased from also Invitrogen. Restriction enzymes and DNA ligase were obtained from TaKaRa Bio Inc. (Japan), and KOD plus DNA polymerase from TOYOBO CO., Ltd. (Japan) was used in polymerase chain reaction. Also pDsRed-Monomer vector coding monomeric red-fluorescent protein (RFP) was purchased from TaKaRa Bio Inc. (Clontech).

2.2.2 Plasmids construction

All of the fusion proteins used in this study were constructed by replacement of the catalytic core domain of cellobiohydrolases with RFP. Four types of CBD containing native linker regions were PCR-amplified with plasmid vector harboring four types of fungal cellobiohydrolases which had previously been cloned in our laboratory. The genes coding C-terminal CBDs from *T. reesei* CBHI (*Tr*Cel7A) and CBH58 from *Phanerochaete chrysosporium* (*Pc*Cel7D), and the genes coding N-terminal CBDs from *T. reesei* CBHII (*Tr*Cel6A) and CBH50 from *P. chrysosporium* (*Pc*Cel6A) were amplified by PCR with primers of 5-6 for *Tr*Cel7A and 9-10 for *Pc*Cel7D or with primers of 7-8 for *Tr*Cel6A and 11-12 for *Pc*Cel6A, as listed in Table 1. The PCR products were ligated into Blunt-TOPO vector by manufacturer's instructions. Following the same operation, RFP genes designed to connect at N- and C-terminus of the fusion proteins were amplified from pDsRed-Monomer with primer set 1-2 and 3-4 and also sub-cloned into Blunt-TOPO vector. These inserts ligated into Blunt-TOPO

vector were applied to DNA sequencing after mini-preparation of the plasmids. Designed Blunt-TOPO/CBDs and Blunt-TOPO/RFPs were double-digested as well as *E. coli-P. pastoris* shuttle vector pPICZαA using restriction enzymes sets *EcoR*I-*Kpn*I or *Kpn*I-*Not*I as listed in Table 1. These genes were ligated simultaneously into the digested pPICZαA after separation by agarose-gel electrophoresis, as down the stream of α-factor signal sequence that to be expressed in culture medium. *E. coli* JM109 were transformed by use of the ligated products, and zeocin resistant transformants were selected on low-salt LB agar plates supplemented with 200 μg/ml zeocin.

Table 1. Primers used in this study to amplify genes coding CBD-tagged RFP fusion proteins.

Primer name	Primer (5' → 3')	Restriction enzyme
DsRed-C terminal type-forward (1)	TTTGAATTC***AAAAGA***ATGGACAACACCGAGGACGTCATC	*EcoR*I
DsRed-C terminal type-reverse (2)	TTTGGTACCCTGGGAGCCGGAGTGGCGG	*Kpn*I
DsRed-N terminal type-forward (3)	TTTGGTACCATGGACAACACCGAGGACGTCATC	*Kpn*I
DsRed-N terminal type-reverse (4)	TTTGCGGCCGC***CTA***CTGGGAGCCGGAGTGG	*Not*I
*Tr*CBHI-C terminal type-forward (5)	TTTGGTACCGGCAACCCTAGCGGCGGC	*Kpn*I
*Tr*CBHI-C terminal type-reverse (6)	TTTGCGGCCGC***TTA***CAGGCACTGAGAGTAGTAAGGG	*Not*I
*Tr*CBHII-N terminal type-forward (7)	TTTGAATTC***AAAAGA***CAAGCTTGCTCAAGCGTCTGGG	*EcoR*I
*Tr*CBHII-N terminal type-reverse (8)	TTTGGTACCCGATCCGACTGGAGGTACTCTG	*Kpn*I
*Pc*CBHI-C terminal type-forward (9)	TTTGGTACCTTCAGCGGCACCTCCTCCC	*Kpn*I
*Pc*CBHI-C terminal type-reverse (10)	TTTGCGGCCGC***TTA***GTAGCACTGCGAGTAGTAAGG	*Not*I
*Pc*CBHII-N terminal type-forward (11)	TTTGAATTC***AAAAGA***CAGGCGTCGGAGTGGGGAC	*EcoR*I
*Pc*CBHII-N terminal type-reverse (12)	TTTGGTACCCGACGGAGGAGGAGGAGGG	*Kpn*I

The restriction enzymes sites are underlined. Kex2 protease cleavage site (*AAAAGA*) or the terminal codon site are shown in bold italic.

2.2.3 Expression of fusion proteins

The constructed plasmids named pPICZαA/RFP-CBD_{TrCBHI}, pPICZαA/RFP-CBD_{PcCBHI}, pPICZαA/$CBD_{TrCBHII}$-RFP, pPICZαA/$CBD_{PcCBHII}$-RFP were used for transformation of *P. pastoris* cells with the standard electroporation procedure described earlier [1]. Zeocin-resistant transformants were cultured in 200 ml of fresh YPG medium (YPG, 1% yeast extract, 2% polypeptone, and 1% glycerol, w/v) in scale of 500 ml-erlenmeyer flask for one day at 30°C with shaking at 150 rpm. The microbial cells in culture mediums harvested by centrifugation for 7 minutes at 1,750 g were washed by 40 ml of distilled water at twice, and then suspended in 40 ml of induction medium (YP, 1% yeast extract, 2% polypeptone, w/v). In order to induce the fusion proteins, 500 μl of 100%-methanol was added in the induction mediums for one time a day. The fluorescent intensity and protein concentration of the supernatants of each culture were measured over six days after the first methanol addition. The fluorescent intensity of medium was measured by Fluoroskan Ascent Microplate Fluorometer (Thermo Fisher Scientific Inc., Japan) with filter pair (excitation at 544 nm and emission at 590 nm). 180 μl of 100 mM sodium phosphate buffer (pH 8.0) were added to 20 μl of the supernatant of each culture media in 96-well black micro-plate (Costar, Product Code 3925), and the plate was subjected to fluorescent measurement after incubation for 5 min at room temperature. Protein concentrations were assayed by Bio-Rad Protein Assay (Bio-Rad Laboratories, Inc.). SDS-PAGE analyses of the each culture medium were performed as the standard procedure in our laboratory [2] to evaluate the expression.

2.3 Results

Fusion proteins connected with four CBDs were secreted in culture medium by all transformants after methanol addition. The production levels of proteins were followed by protein assay and fluorescent intensity measurement. The concentrations of all CBD-RFPs in the medium had been increasing over the 6 days period (Fig. 1, A and B). The relative fluorescent intensities toward to protein concentration of all expressed proteins increased untill day four and then remained unchanged (Fig. 1, C). The concentrations of RFP-CBD$_{Tr\text{CBHI}}$, CBD$_{Tr\text{CBHII}}$-RFP, RFP-CBD$_{Pc\text{CBHI}}$ and CBD$_{Pc\text{CBHII}}$-RFP fusion proteins reached on day 6 were 1.2 g/L, 0.18 g/L, 1.0 g/L and 0.79 g/L, respectively. The production level of CBD$_{Tr\text{CBHII}}$-RFP was lesser relative to other fusion proteins. The molecular weights of the proteins as estimated by SDS-PAGE analysis were approximately 40 kDa, 48 kDa, 43 kDa and 52 kDa for the CBD$_{Tr\text{CBHI}}$, CBD$_{Tr\text{CBHII}}$, CBD$_{Pc\text{CBHI}}$ and CBD$_{Pc\text{CBHII}}$ fused proteins, respectively (Fig. 2, C). Predicted molecular weights for CBD$_{Tr\text{CBHI}}$, CBD$_{Tr\text{CBHII}}$, CBD$_{Pc\text{CBHI}}$ and CBD$_{Pc\text{CBHII}}$ fused proteins without any glycosylation were 32.2 kDa, 33.8 kDa, 32.2 kDa and 33.7 kDa, respectively, and the number of predicted *O*-linked glycosylation sites in linker regions were 12, 23, 17 and 28 in the same order, which the glycosylation sites agreed well with the observed molecular weights.

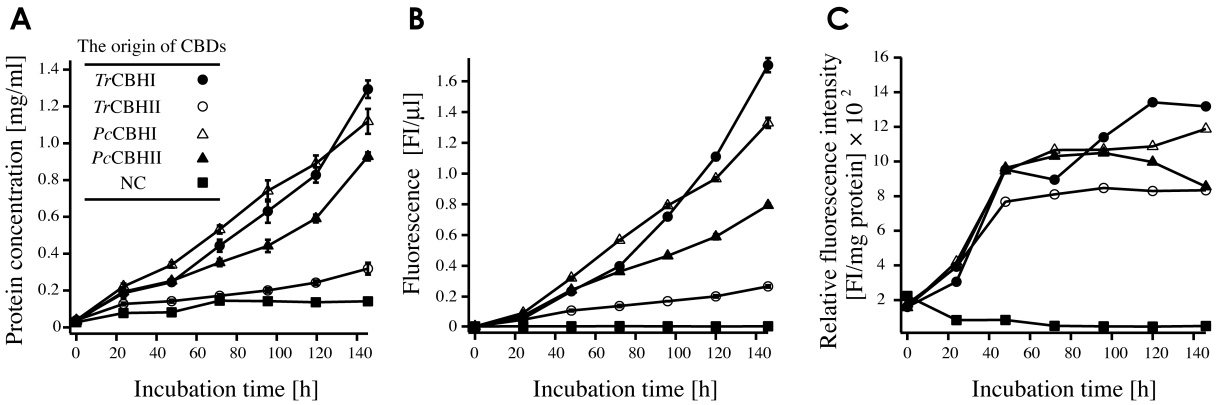


Fig. 1. Production of various CBD fusion proteins. Protein concentrations (A) and fluorescence intensity (B) in culture media. Relative fluorescence intensities of CBD fusion proteins are also shown (C). RFP tagged with CBD$_{Tr\text{CBHI}}$ and CBD$_{Tr\text{CBHII}}$ are shown with filled and open circles, respectively, and RFP tagged with CBD$_{Pc\text{CBHI}}$ and CBD$_{Pc\text{CBHII}}$ are shown with open and filled triangles, respectively. Negative control without transformation is shown with closed squares. The symbols are also shown in inset of panel A.

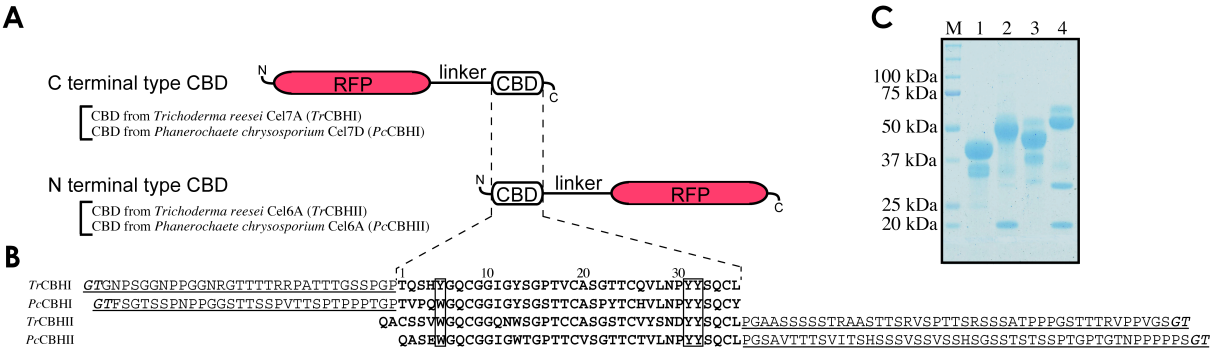


Fig. 2. Construction of expressed fusion proteins and SDS-PAGE results. (A) Domain constructions of RFP C-terminally fused with CBDs from *T. reesei* CBHI and *P. chrysosporium* CBH58 and N-terminally fused with CBDs from *T. reesei* CBHII and *P. chrysosporium* CBH50. Native linker regions of these cellobiohydrolases were used. (B) The amino acid sequences of CBD and linker regions. Boldface type indicates CBD sequences. The linker regions and the sites of restriction enzyme *Kpn*I (Gly-Thr) are shown in underlined and italic letters, respectively. (C) SDS-PAGE of culture media containing expressed fusion proteins. Samples of approximately 10 μg were separated on 12.5% (w/v) polyacrylamide gel. Lane M, molecular weight standards: sizes are shown to the left; lane 1, RFP-CBD$_{Tr\text{CBHI}}$; lane 2, CBD$_{Tr\text{CBHII}}$-RFP; lane 3, RFP-CBD$_{Pc\text{CBHI}}$; lane 4, CBD$_{Pc\text{CBHII}}$-RFP.

2.4 Discussion

High-level secretions of the CBD fusion proteins into the culture medium were achieved in a *P. pastoris* expression system. Production levels of fusion proteins in the culture medium reached 1.2 g/L (RFP-CBD_{TrCBHI}). Reinikainen and co-workers have reported expression of a similar fusion protein in *E. coli* [3], OxscFv-CBD_{CBHI}, with the CBD from *T. reesei* CBHI, but the production level obtained in a fermenter cultivation was only 0.09 g/L. The genome of *P. pastoris* has been disclosed recently [4], and several genes are predicted to carry a family 1 CBM, such as *mlp* gene coding hypothetical yeast flocculation protein and hypothetical glycoside hydrolase family 45 endo-glucanase. These results indicate that *P. pastoris* is a suitable host for the production of fusion proteins with a fungal CBD. The observation that production of $CBD_{TrCBHII}$ was lower than that of other CBDs might reflect difficulty in correct protein folding, since $CBD_{TrCBHII}$ has three putative disulfide bonds [5, 6], whereas other family 1 CBDs have only two putative disulfide bonds. At any rate, the expression of fungal CBD-tag with fusion proteins of interest using *P. pastoris* might be provided a success procedure for industrial manufacturing.

2.5 References

[1]. **Igarashi K, Yoshida M, Matsumura H, Nakamura N, Ohno H, Samejima M, Nishino T**: Electron transfer chain reaction of the extracellular flavocytochrome cellobiose dehydrogenase from the basidiomycete *Phanerochaete chrysosporium*. *Febs J* (**2005**), 272:2869-2877.

[2]. **Kawai R, Yoshida M, Tani T, Igarashi K, Ohira T, Nagasawa H, Samejima M**: Production and characterization of recombinant *Phanerochaete chrysosporium* β-glucosidase in the methylotrophic yeast *Pichia pastoris*. *Biosci Biotechnol Biochem* (**2003**), 67:1-7.

[3]. **Reinikainen T, Takkinen K, Teeri TT**: Comparison of the adsorption properties of a single-chain antibody fragment fused to a fungal or bacterial cellulose-binding domain. *Enzyme Microb Tech* (**1997**), 20:143-149.

[4]. **De Schutter K, Lin YC, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, Rouzé P, Van de Peer Y, Callewaert N**: Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat Biotechnol* (**2009**), 27:561-566.

[5]. **Hoffrén AM, Teeri TT, Teleman O**: Molecular dynamics simulation of fungal cellulose-binding domains: differences in molecular rigidity but a preserved cellulose binding surface. *Protein Eng* (**1995**), 8:443-450.

[6]. **Carrard G, Linder M**: Widely different off rates of two closely related cellulose-binding domains from *Trichoderma reesei*. *Eur J Biochem* (**1999**), 262:637-643.

Chapter 3. Cellulose Affinity Purification

3.1 Introduction

The production of CBD fusion proteins was successful and the details were described in chapter 2. In this chapter, I focused on the potential application of CBD, and investigated the feasibility of purification of fusion proteins. The most realistic application is their use of affinity chromatography. Although oligo-histidine, glutathione S-transferase (GST), and maltose-binding protein (MBP) have been used as affinity tags for the purification of fusion protein [1-3], they all can be considered for high cost affinity purification systems which is a common disadvantage. In fact CBD can be genetically fused to a protein of interest for specific binding to cheap and versatile cellulose matrices, suggesting that CBD-tags have a potential for affinity purification and/or immobilization using 'inexpensive' cellulose matrices [4-8]. Bacterial CBDs have often been used as the affinity tags with the combination of *Escherichia coli* protein expression system, and the CBD_{Cex} (CBM2a) or CBD_{CenA} from *Cellulomonas fimi* and CipA CBD (CBM3) from *Clostridium* spp. became commercially available [9-13]. In contrast, fungal CBDs belonging to CBM family 1 are not common as a tag for fusion protein maybe because of their inability of heterologous expression in *E. coli* system. Although CBD from cellobiohydrolase I from *Trichoderma reesei* (CBD_{TrCBHI}) has been used for cellulose affinity chromatography [14-16], the application of the system is quite limited.

CBDs belonging to CBM family 1 share common feature of rather small molecular weight with less than 40-amino acids residues (approximately 3 kDa), with

two dithiol bonds in the domain, and they contain three hydrophobic amino acids, which form a hydrophobic flat face to bind a hydrophobic surface of crystalline cellulose [17, 18]. Various reports describe apparent reversible adsorption of fungal CBD on cellulose [16, 19, 20]. Therefore, fungal CBDs could be suitable for affinity purification of fusion proteins, if the proteins are correctly folded during protein production.

In this study, I examined the feasibility of protein purification by means of expression of a target protein as an N- or C-terminal fusion protein with CBM family 1 CBDs from fungal cellobiohydrolases (CBHs) in the methylotropic yeast *Pichia pastoris*, followed by affinity separation of the fusion protein on a cellulose column. The affinity of the fusion protein for the cellulose column was enhanced by the addition of kosmotropic salt, ammonium sulfate, and the fusion protein could be easily eluted with water. Expression level, extent of purification (fold) and purification yield were compared among CBD-tagged fusion proteins.

3.2 Materials and Methods

3.2.1 Chemicals

All chemicals used in this study were laboratory-grade products from WAKO Pure Chemical Industries, Ltd. (Osaka, Japan) or Sigma-Aldrich (St. Louis, USA). Cellulose powders, CF11 (fibrous), CC31 and CC41 (microgranular) were from

Whatman Ltd. (England) and Avicel (Funacel II, average grain size 80 μm) was from Funakoshi Ltd. (Tokyo, Japan).

3.2.2 Production of CBD fusion proteins

The production of all fusion proteins was described in chapter 2.

3.2.3 Purification of RFP-CBD$_{TrCBHI}$ with a CC31 column

The crude extracellular proteins obtained by centrifugation (6,000 *g*, 30 min) were loaded onto a manually packed CC31 cellulose column (ϕ10 x 20 mm) and eluted with distilled water at 25˚C. The protein sample containing 1 M ammonium sulfate was also loaded on the column, which had been pre-equilibrated with the same concentration of ammonium sulfate, and the column was washed with 1 M ammonium sulfate solution then eluted with distilled water.

The crude fusion protein sample (1.0 ml, approximately 0.5 mg/ml) was applied to a CC31 column (ϕ10 x 50 mm) and eluted with water at the flow rate of 0.5, 1.0 or 1.5 ml/min after a 5 min wash with 1 M ammonium sulfate at 4˚C, in order to examine the effect of flow rate. The effect of column volume (CV) was examined with the same protein sample, using ϕ10 x 20, ϕ10 x 55 and ϕ30 x 90 mm columns (CV=1.6, 4.3 and 60 ml, respectively), which were washed with 3 CV of 1 M ammonium sulfate and then eluted with distilled water at 1.5 ml/min. Purity of the product and purification yield were determined by measurement of fluorescence intensity and

protein concentration of the collected fractions and loading samples, as described above.

3.2.4 Characterization of celluloses as affinity column materials

The performance of crystalline celluloses (CF11, CC31, CC41 and Avicel) as cellulose affinity column chromatography materials was tested with a manually packed open column (ϕ10 x 5 mm) using RFP-CBD$_{TrCBHI}$. The cellulose columns were equilibrated with 1 M ammonium sulfate after being washed with sufficient water to remove fine cellulose particles, followed by the loading of protein samples (0.16 mg/ml, 1 ml) in the presence of 1 M ammonium sulfate and washing with 12 CV of the same solution. RFP-CBD$_{TrCBHI}$ was eluted with 12 CV of distilled water. To estimate the binding capacity of celluloses, protein loading was increased until red fluorescence was detectable in the collected fractions. The fluorescence intensity and protein concentration of each fraction were measured with a Fluoroskan Ascent Microplate Fluorometer and a Bio-Rad Protein Assay kit, respectively.

3.2.5 Purification of RFP-CBDs on an Avicel column

Avicel was suspended in at least ten times its weight of water, and the supernatant, containing fine cellulose particles, was removed. This was repeated 5 times until the supernatant was clear, and the rinsed Avicel was packed into open columns (ϕ10 x 10 mm). RFP-CBD$_{TrCBHI}$, -CBD$_{TrCBHII}$, -CBD$_{PcCBHI}$, -CBD$_{PcCBHII}$ were purified on columns

prepared in this way at 4˚C or 20˚C. The fluorescence intensity and protein concentration of the elute fractions and loading samples were measured as described above.

A sample of RFP-CBD$_{TrCBHI}$ was subjected to Avicel column chromatography with a stepwise gradient elution from 1 M ammonium sulfate to distilled water (as in hydrophobic chromatography). Purified RFP-CBD$_{TrCBHI}$ in the presence of 1 M ammonium sulfate was loaded onto the column, which was washed with 12 CV of 0.5 M ammonium sulfate, followed by 12 CV of 0.25 M ammonium sulfate and 12 CV of 0.125 M ammonium sulfate, and finally eluted with distilled water.

3.2.6 Adsorption and desorption analysis of RFP-CBD$_{TrCBHI}$

RFP-CBD$_{TrCBHI}$ purified on a cellulose column was cleaved by treatment with papain and the RFP region without the CBD-linker region was purified on a DEAE column. Its purity was confirmed by SDS-PAGE analysis. The prepared RFP region and bovine serum albumin (BSA), in addition to intact RFP-CBD$_{TrCBHI}$, were subjected to Avicel column chromatography (ϕ10 x 50 mm for BSA, ϕ10 x 5 mm for other samples) under the same conditions described above. These proteins samples were also subjected to chromatography on a phenyl-Toyopearl 650S column (ϕ10 x 5 mm) equilibrated with 1 M ammonium sulfate; the column was washed with the same solution and immediately eluted with distilled water. The fluorescence intensity and protein

concentration of fractions were measured to evaluate the loss during washing with 1 M ammonium sulfate and the yield on elution with water.

3.3 Results

3.3.1 Cellulose affinity purification of RFP-CBD$_{TrCBHI}$ fusion proteins

All fusion proteins constructed in the present study were secreted into the culture medium of *P. pastoris*, as shown in chapter 2. The crude RFP-CBD$_{TrCBHI}$ fusion protein (culture filtrate) was directly loaded onto a cotton cellulose (CC31) column, and eluted with water from the column. When the protein sample was loaded in the presence of 1 M ammonium sulfate after equilibration of the column with the same solution, the fusion protein was apparently bound in the upper part of the column, and was not desorbed during further washing with 1 M ammonium sulfate. The protein was eluted immediately when distilled water was employed, as shown in Fig. 1. The yield of RFP-CBD$_{TrCBHI}$ varied depending on the flow rate, and 36, 52, and 49% of the applied protein was recovered at the flow rates of 0.5, 1.0 and 1.5 ml/min, respectively. The volume of the column also affected the yield, and 61, 54 and 25% of the loaded protein was recovered from 1.6, 4.3 and 60 ml CC31 columns, respectively.

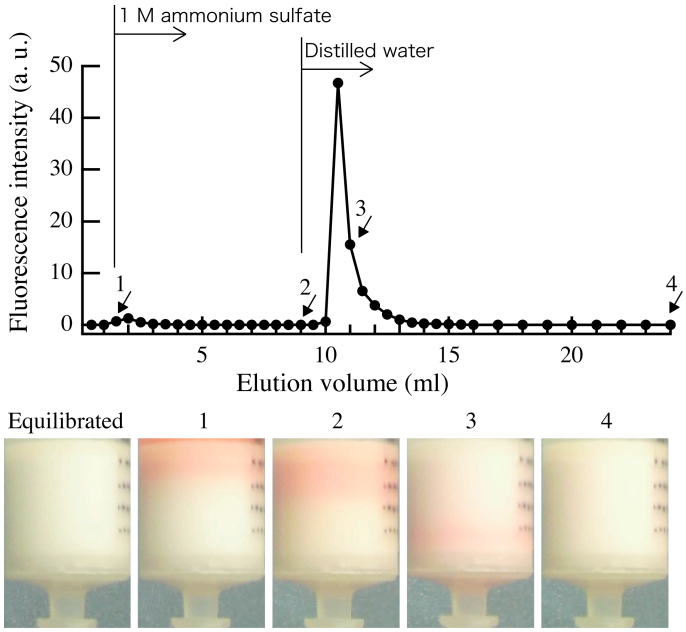


Fig. 1. Separation of RFP-CBD$_{TrCBHI}$ on a CC31 column equilibrated with 1 M ammonium sulfate and eluted with distilled water. Upper: Fluorescence intensity of fractions during separation; bottom: changes of color upon adsorption and desorption of the fusion protein. The numbers of pictures correspond to those in upper chromatogram.

3.3.2 Characterization of celluloses as affinity column materials

The culture filtrate containing RFP-CBD$_{TrCBHI}$ was diluted to 0.16 mg/ml with distilled water containing 1 M ammonium sulfate and loaded onto columns of four types of cellulose (CF11, CC31 and CC41 from cotton, and Avicel from wood), followed by elution by water. As shown in Table 1, the yield varied from 43% (CF11) to 79% (Avicel) although the extent of purification (fold) was similar in all cases. Purity, as assessed by SDS-PAGE analysis, was almost the same in all cases, as shown in Fig. 2. As shown in Table 1, the maximum capacity of the column for the fusion protein was

greatest on Avicel, followed by cotton celluloses with small particle size (CC41>CC31>CF11). The yields and purities of the fusion protein using CC31 and Avicel were similar, whereas the maximum capacity of Avicel was about 5 times higher than that of CC31.

Table 1. Comparison of cellulose matrices as affinity materials for CBD-tagged fusion protein.

Cellulose	Protein concentration (mg/ml)	Total protein (mg)	Specific fluorescence (FI/mg)	Total fluorescence (FI)	Yield (%)	Purity (Fold)	Maximum desorption capacity	
							Volume basis (mg/ml-cellulose)	Weight basis (mg/g-cellulose)
	0.16	0.16	1.8×10^3	2.8×10^2	100	1.0		
CF11	0.017	0.058	2.1×10^3	1.2×10^2	43	1.2	0.54	2.4
CC31	0.026	0.090	2.1×10^3	1.9×10^2	68	1.2	0.89	2.2
CC41	0.022	0.077	2.4×10^3	1.9×10^2	68	1.5	1.4	3.4
Avicel	0.031	0.094	2.3×10^3	2.2×10^2	79	1.3	4.8	17

RFP-CBD$_{TrCBHI}$ was used for the evaluation. Crude solution of RFP-CBD$_{TrCBHI}$ in the presence of 1 M ammonium sulfate was loaded onto four kinds of cellulose column equilibrated with the same solution at room temperature (22˚C).

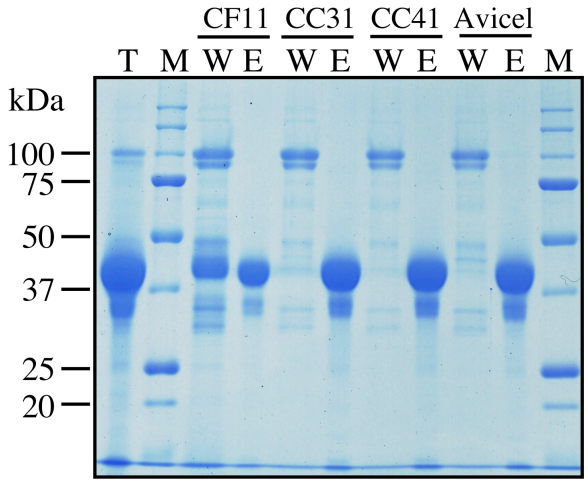


Fig. 2. SDS-PAGE of unbound and bound fractions after separation on various cellulose matrices. Total protein obtained from culture medium of *P. pastoris* harboring RFP-CBD$_{TrCBHI}$ fusion protein (T), wash fraction with 1 M ammonium sulfate (W) and eluate with water (E) were analyzed by 12.5% (w/v) polyacrylamide gel electrophoresis. The target protein of approximately 40 kDa was recovered in the eluate in all cases.

3.3.3 Affinity purification of RFP-CBD$_{TrCBHI}$ on an Avicel column

The purification yield and the extent of purification (fold) of RFP-CBD$_{TrCBHI}$ on an Avicel column at 4˚C were 61 % and 1.5, respectively, whereas the corresponding values at 20˚C were 76% and 1.4, respectively, as shown in Table 2. Next, the elution conditions were evaluated by decreasing the concentration of ammonium sulfate from 1 M to 0 M (water). The column was washed with 12 CV of 1.0 M, 0.5 M 0.25 M, and 0.125 M ammonium sulfate after application of the fusion protein, and the amounts of eluted protein were estimated to be 0%, 0%, 0.06% and 2.7% of the applied amount, respectively. Finally, elution with water desorbed 62% of the applied protein. Direct elution with water afforded approximately 80% recovery at room temperature, suggesting that the purification yield is higher when the ammonium sulfate solution is immediately changed to water, i.e., without gradient elution.

Purified intact RFP-CBD$_{TrCBHI}$ and the truncated RFP region without the CBD and linker regions were loaded on Avicel columns equilibrated with 1 M ammonium sulfate. The fusion protein with CBD completely bound to the column, as shown in Fig. 3A, whereas most of the truncated RFP region was not adsorbed on the column under these conditions. Similarly, about 95% of BSA loaded on the Avicel column was not bound in the presence of 1 M ammonium sulfate, and only 4% of loaded BSA was subsequently eluted with water. Only 0.009 mg BSA/ml-Avicel was bound in 1 M ammonium sulfate, although 4.8 mg RFP-CBD$_{TrCBHI}$ was trapped, as shown in Table 1. For comparison, similar experiments were carried out using a hydrophobic interaction (phenyl) column, as shown in Fig. 3B. Intact RFP-CBD$_{TrCBHI}$ and RFP region showed

similar behavior to that on the cellulose column, i.e., the fusion protein was bound to the matrix, but RFP without CBD was not. In contrast, BSA was bound to the phenyl column in the presence of 1 M ammonium sulfate and eluted with water (Fig. 3B), clearly indicating that the cellulose column is superior to purify CBD-tagged fusion protein.

Table 2. Temperature dependency of affinity purification of RFP-CBD$_{Tr\text{CBHI}}$ on an Avicel column.

Temperature (°C)		Protein concentration (mg/ml)	Total protein (mg)	Specific fluorescence (FI/mg)	Total fluorescence (FI)	Purification (Fold)	Yield (%)	Desorption yield (%)
	crude	0.58	0.29	1.8×10^3	5.2×10^2	1.0	100	— —
4°C		0.049	0.12	2.6×10^3	3.2×10^2	1.46	61.2	70.8
20°C		0.064	0.16	2.5×10^3	4.0×10^2	1.36	75.7	81.6

Fluorescence measurement was carried out with excitation at 544 nm and emission at 590 nm. Desorption yield was evaluated by dividing the total fluorescence after water elution by the fluorescence of totally adsorbed protein after washing with 1 M ammonium sulfate.

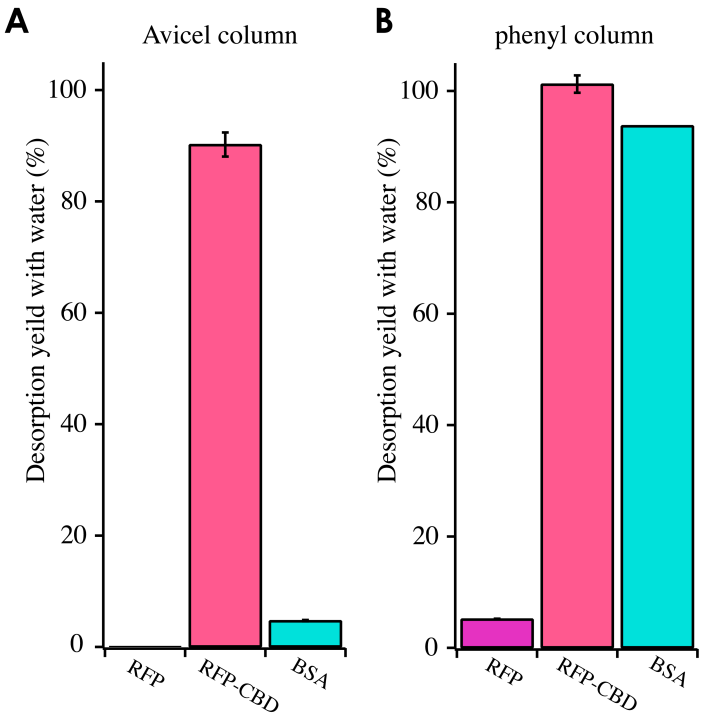


Fig. 3. Adsorption and desorption of RFP-CBD$_{Tr\text{CBHI}}$ from cellulose (A) and phenyl column (B). RFP region without the CBD-linker region (cleaved with papain), purified intact RFP-CBD$_{Tr\text{CBHI}}$, or BSA was loaded onto an Avicel column or a phenyl column in the presence of 1 M ammonium sulfate and then eluted with water.

3.3.4 Affinity purification of various CBD fusion proteins

Four kinds of CBD-fused RFPs were applied to an Avicel column to investigate the effect of various sequences of CBD on cellulose affinity chromatography. All the fused proteins were eluted with water after injection with 1 M ammonium sulfate, as shown in Fig. 4A. The yields of CBD$_{Tr\text{CBHI}}$, CBD$_{Tr\text{CBHII}}$, CBD$_{Pc\text{CBHI}}$ and CBD$_{Pc\text{CBHII}}$ were 76%, 62%, 65% and 51%, respectively, as listed in Table 3. As shown in Table 3, the desorption ratio (amount of eluted protein/amount of bound protein) was highest for CBD$_{Tr\text{CBHI}}$ (82%). All fusion proteins exhibited a single band on SDS-PAGE, as shown Fig. 4B.

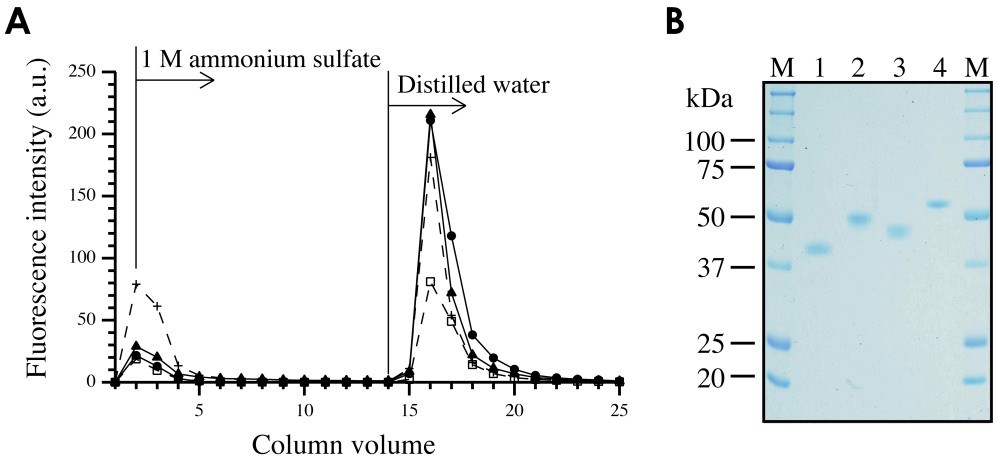


Fig. 4. Cellulose affinity chromatography of various CBD fusion proteins (A) and SDS-PAGE analysis of eluate fractions (B). Four types of CBD fusion proteins were bound to an Avicel column in the presence of 1 M ammonium sulfate and eluted with distilled water. RFP tagged with CBD_{TrCBHI} and $CBD_{TrCBHII}$ is shown with filled circles and open squares, respectively, and RFP tagged with CBD_{PcCBHI} and $CBD_{PcCBHII}$ is shown with filled triangles and crosses, respectively. Approximately 1 μg of sample was injected into a 12.5% (w/v) polyacrylamide gel. Lane M, molecular weight standards: sizes are shown to the left; lane 1, RFP-CBD_{TrCBHI}; lane 2, $CBD_{TrCBHII}$-RFP; lane 3, RFP-CBD_{PcCBHI}; lane 4, $CBD_{PcCBHII}$-RFP.

Table 3. Comparison of various CBDs from fungal cellobiohydrolases.

Type of CBDs	Production in culture filtrate (g/L)	Purity (Fold)	Yield (%)	Desorption ratio (%)
CBD_{TrCBHI}	1.2	1.36	75.7	81.6
$CBD_{TrCBHII}$	0.18	1.40	62.4	71.9
CBD_{PcCBHI}	1.0	1.26	65.3	76.9
$CBD_{PcCBHII}$	0.79	1.24	50.8	75.8

Cellulose affinity chromatography was performed at 20°C as described in Materials and Methods.

3.4 Discussion

Bacterial CBDs, such as CBM2 and CBM3, have been used as affinity tags for purification and/or immobilization of proteins of interest, often in combination with the *E. coli* expression system. Although fungal CBM family 1 CBDs have similar adsorption characteristics to bacterial CBDs of much smaller molecular size, they have not generally been considered as tools for affinity purification, possibly because they are not necessarily compatible with the *E. coli* expression system. In this study, I investigated fungal CBDs as fusion tags for affinity chromatography, using the yeast *P. pastoris* expression system.

The reversibility of binding of fungal CBD to cellulose has been controversial, with some authors indicating that the binding is reversible [14, 15, 19, 20], and others finding that it is irreversible [21-24]. In the present study, I had no difficulty in adsorbing and desorbing the fusion proteins from the cellulose column, and the recovered protein could be re-adsorbed on cellulose, indicating that the adsorption of fungal CBD on cellulose is reversible, and the elution is not simply due to denaturation of the protein. However, an increase of column volume appeared to decrease the recovery of fusion proteins, suggesting that irreversible entrapment may also occur. To minimize this, I examined various conditions of adsorption and desorption using 4 commercially available cellulose matrices. The best performance was obtained with an Avicel column pre-equilibrated with ammonium sulfate. This effect of ammonium sulfate may be due to increased hydrophobicity in the presence of the kosmotropic salt, and blocking of negative charges of the cellulose matrices by the ammonium salt. Cellulose carriers derived from cotton (CF11, CC31, CC41) were less effective,

possibly because their surface area is smaller than that of Avicel. It has been reported that the binding of CBD from *T. reesei* CBHI and CBHII to crystalline cellulose was increased at low temperature, i.e., that the affinity was higher at low temperature [19, 20]. For the Avicel column tested in this study, the yield obtained at 4˚C was lower than that at 20˚C, in agreement with the former studies. This suggests that conducting affinity purification in a cold room will decrease the product yield.

The best performance in cellulose affinity purification of fungal CBD-tagged protein was obtained by using 1 M ammonium sulfate, which is known to increase hydrophobic interaction, to maximize binding of the fusion protein, followed by elution with water. However, the adsorption and desorption behaviors of CBD_{TrCBHI}-tagged RFP, RFP without CBD, and BSA were somewhat different between a cellulose column and a phenyl column. RFP alone was eluted in the void volume, whereas the fusion protein with CBD was significantly adsorbed on the phenyl column. This may imply that CBD simply enhances the hydrophobicity of the target protein, thereby increasing absorption on the phenyl column. However, the behavior of BSA, a well-known hydrophobic protein, was completely different from that of CBD-tagged protein, i.e., BSA was tightly bound on the phenyl column in the presence of 1 M ammonium sulfate and was eluted with water, whereas it was present in the unadsorbed flow-through fractions on an Avicel column. These results indicate that the binding of CBD-tagged protein to crystalline cellulose is not simply a hydrophobic interaction, and rather suggest that fungal family 1 CBDs are adsorbed on the cellulose surface not only via hydrophobic interaction, but also via other interactions, such as hydrogen-bonding and/or ionic interactions.

Fungal CBDs contain three aromatic residues arranged to form a plane [17, 18]. In CBD from *T. reesei* Cel7A, the residues involved are Tyr5, Tyr31, and Tyr32, although in the other three CBDs used in this study, one Trp residue is involved in place of the Tyr near the N-terminus. Linder and coworkers [25, 26] reported that replacement of one tyrosine residue with tryptophan resulted in higher affinity for cellulose, so that CBD_{TrCBHI} may be less favorable for adsorption than the other CBDs tested in this study. As in the case of the temperature dependence of cellulose affinity column, higher affinity may cause lower recovery of fusion protein, because the present results show that the highest yield was obtained when CBD_{TrCBHI} used as the fusion tag. Other CBDs with a tryptophan residue commonly gave low recovery on elution with water. Therefore, mutation of these residues to other amino acids, such as phenylalanine and histidine, might provide a better yield in affinity purification using cellulose columns.

In conclusion, fusion proteins of RFP with fungal family 1 CBDs were highly expressed in a *P. pastoris* expression system, and could be purified by single-step cellulose affinity purification in high yield. The recovery was increased when the culture filtrate was applied to the column in the presence of 1 M ammonium sulfate, which enhances hydrophobicity as well as blocking negative charges on the cellulose. The bound fusion proteins could be easily and quickly eluted with water. This system appears to be suitable for large-scale purification of target proteins.

3.5 References

[1]. **Porath J, Carlsson J, Olsson I, Belfrage G**: Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* (**1975**), 258:598-599.

[2]. **Bedouelle H, Duplay P**: Production in *Escherichia coli* and one-step purification of bifunctional hybrid proteins which bind maltose. Export of the Klenow polymerase into the periplasmic space. *Eur J Biochem* **(1988**), 171:541-549.

[3]. **Smith DB, Johnson KS**: Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **(1988**), 67:31-40.

[4]. **Ong E, Greenwood J, Gilkes NR, Kilburn DG, Miller RC, Warren RA**: The cellulose-binding domains of cellulases: tools for biotechnology. *Trends Biotechnol* (**1989**), 7:239-243.

[5]. **Ong E, Gilkes NR, Miller RC, Warren RAJ, Kilburn DG**: Enzyme immobilization using a cellulose-binding domain: Properties of a β-glucosidase fusion protein. *Enzyme Microb Tech* (**1991**), 13:59-65.

[6]. **Levy I, Shoseyov O**: Cellulose-binding domains: biotechnological applications. *Biotechnol Adv* **(2002**), 20:191-213.

[7]. **Kavoosi M, Meijer J, Kwan E, Creagh AL, Kilburn DG, Haynes CA**: Inexpensive one-step purification of polypeptides expressed in *Escherichia coli* as fusions with the family 9 carbohydrate-binding module of xylanase 10A from *T. maritima*. *J Chromatogr B Analyt Technol Biomed Life Sci* **(2004**), 807:87-94.

[8]. **Hildén L, Johansson G**: Recent developments on cellulases and carbohydrate-binding modules with cellulose affinity. *Biotechnol Lett* **(2004**), 26:1683-1693.

[9]. **Owolabi JB, Beguin P, Kilburn DG, Miller RC, Warren RA**: Expression in *Escherichia coli* of the *Cellulomonas fimi* structural gene for endoglucanase B. *Appl Environ Microb* **(1988)**, 54:518-523.

[10]. **Boraston AB, McLean BW, Guarna MM, Amandaron-Akow E, Kilburn DG**: A family 2a carbohydrate-binding module suitable as an affinity tag for proteins produced in *Pichia pastoris*. *Protein Expres Purif* **(2001)**, 21:417-423.

[11]. **Kwan EM, Boraston AB, McLean BW, Kilburn DG, Warren RA**: *N*-Glycosidase-carbohydrate-binding module fusion proteins as immobilized enzymes for protein deglycosylation. *Protein Eng Des Sel* **(2005)**, 18:497-501.

[12]. **Rodriguez B, Kavoosi M, Koska J, Creagh AL, Kilburn DG, Haynes CA**: Inexpensive and generic affinity purification of recombinant proteins using a family 2a CBM fusion tag. *Biotechnol Progr* **(2004)**, 20:1479-1489.

[13]. **Hong J, Ye X, Wang Y, Zhang YP**: Bioseparation of recombinant cellulose-bindning module-proteins by affinity adsorption on an ultra-high-capacity cellulosic adsorbent. *Anal Chim Acta* (**2008**), 621:193-199.

[14]. **Otter DE, Munro PA, Scott GK, Geddes R**: Desorption of *Trichoderma reesei* cellulase from cellulose by a range of desorbents. *Biotechnol Bioeng* (**1989**), 34:291-298.

[15]. **Tomme P, Driver DP, Amandoron EA, Miller RC, Antony R, Warren J, Kilburn DG**: Comparison of a fungal (family I) and bacterial (family II) cellulose-binding domain. *J Bacteriol* **(1995)**, 177:4356-4363.

[16]. **Lymar ES, Li B, Renganathan V**: Purification and characterization of a cellulose-binding β-glucosidase from cellulose-degrading cultures of *Phanerochaete chrysosporium*. *Appl Environ Microb* **(1995)**, 61:2976-2980.

[17]. **Kraulis J, Clore GM, Nilges M, Jones TA, Pettersson G, Knowles J, Gronenborn AM**: Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. A study

using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry-Us* (**1989**), 28:7241-7257.

[18]. **Mattinen ML, Kontteli M, Kerovuo J, Linder M, Annila A, Lindeberg G, Reinikainen T, Drakenberg T**: Three-dimensional structures of three engineered cellulose-binding domains of cellobiohydrolase I from *Trichoderma reesei*. *Protein Sci* (**1997**), 6:294-303.

[19]. **Linder M, Teeri TT**: The cellulose-binding domain of the major cellobiohydrolase of *Trichoderma reesei* exhibits true reversibility and a high exchange rate on crystalline cellulose. *P Natl Acad Sci Usa* (**1996**), 93:12251-12255.

[20]. **Carrard G, Linder M**: Widely different off rates of two closely related cellulose-binding domains from *Trichoderma reesei*. *Eur J Biochem* (**1999**), 262:637-643.

[21]. **Beldman G, Voragen AG, Rombouts FM, Searle-van Leeuwen MF, Pilnik W**: Adsorption and kinetic behavior of purified endoglucanases and exoglucanases from *Trichoderma viride*. *Biotechnol Bioeng* (**1987**), 30:251-257.

[22]. **Reinikainen T, Takkinen K, Teeri TT**: Comparison of the adsorption properties of a single-chain antibody fragment fused to a fungal or bacterial cellulose-binding domain. *Enzyme Microb Tech* (**1997**), 20:143-149.

[23]. **Linder M, Nevanen T, Soderholm L, Bengs O, Teeri TT**: Improved immobilization of fusion proteins via cellulose-binding domains. *Biotechnol Bioeng* (**1998**), 60:642-647.

[24]. **Lehtiö J, Wernerus H, Samuelson P, Teeri TT, Stahl S**: Directed immobilization of recombinant staphylococci on cotton fibers by functional display of a fungal cellulose-binding domain. *Fems Microbiol Lett* (**2001**), 195:197-204.

[25]. **Linder M, Lindeberg G, Reinikainen T, Teeri TT, Pettersson G**: The difference in affinity between two fungal cellulose-binding domains is dominated by a single amino-acid substitution. *Febs Lett* (**1995**), 372:96-98.

[26]. **Linder M, Mattinen ML, Kontteli M, Lindeberg G, Ståhlberg J, Drakenberg T, Reinikainen T, Pettersson G, Annila A**: Identification of functionally important amino acids in the cellulose-binding domain of *Trichoderma reesei* cellobiohydrolase I. *Protein Sci* (**1995**), 4:1056-1064.

Chapter 4. Adsorption characteristics

4.1 Introduction

Fungal cellulases that hydrolyze crystalline cellulose share a common two-domain structure consisting of a catalytic domain (CD) and a CBD, which promotes degradation by mediating adsorption of the cellulases on the cellulose surface [1, 2]. Cellobiohydrolase (CBH) I from *Trichoderma reesei* (*Tr*Cel7A or *Tr*CBHI), one of the best-studied cellulases, is known to have two binding modes on crystalline cellulose, i.e., productive binding in which both CD and CBD participate in adsorption and non-productive binding involving only CBD [3, 4]. These two modes of adsorption have been considered to correspond to high- and low-affinity adsorption in Langmuir's two-binding-site model, and the relationship between cellulose-degrading activity and adsorption has been discussed on this basis [4, 5]. However, it is still unclear whether the two modes of adsorption of *Tr*CBHI can be appropriately described in terms of the parameters of Langmuir's two-site model.

In addition to the binding mode of cellulase (productive or non-productive binding), the crystalline form of cellulose also strongly influences the decomposition of cellulose. When cellulose I (naturally occurring crystalline cellulose) is treated with supercritical ammonia, it is converted to cellulose III_I, which is more susceptible to degradation by cellulases [5]. In previous report, Igarashi and co-workers showed that the adsorption parameters for high-affinity adsorption of *Tr*CBHI were significantly larger for ammonia-treated cellulose III_I than for cellulose I, and they proposed that the phenomenon can be interpreted in terms of the combined effects of increased surface

area and decreased interaction between cellulose chains in the crystal, although the precise mechanism remains unclear.

In the present study, I examined the adsorption kinetics of fungal family 1 CBD on crystalline celluloses I_α and III_I, in addition to amorphous cellulose. In order to exclude the contribution of the CD to the adsorption, I used a fusion protein of the CBD of *Tr*CBHI with red-fluorescent protein (RFP), which is not adsorbed on celluloses.

4.2 Materials and Methods

4.2.1 Regents and strains

Restriction enzymes, DNA ligase, and *Escherichia coli* JM109 were obtained from TaKaRa Bio Inc. (Japan). pCR4Blunt-TOPO vector for sequencing, oligonucleotides, *Pichia pastoris* KM71H for protein production and pPICZαA transforming vector were purchased from Invitrogen (Carlsbad, CA). KOD plus DNA polymerase from Toyobo Co., Ltd. (Japan) was used for PCR. PreScission protease was purchased from GE Healthcare Life Sciences. Avicel cellulose powder (Funacel II, average grain size 80 μm) was bought from Funakoshi Ltd. (Tokyo, Japan). Crystalline cellulose I_α and III_I were prepared from green algae (*Cladophora* spp.) as described previously [6]. The amorphous cellulose (PASC) was prepared from Avicel by phosphoric acid treatment. Amorphous cellulose (PASC) was prepared from Avicel as follows. Avicel (2 g) was mixed with 85% (w/w) phosphoric acid (20 ml) and smashed with grass stick until the solution was to be clear. The cellulose solution was regenerated in water

and equalized with a high-speed blender after incubation through overnight. Cellulose suspension was washed with sufficient amount of water [7].

4.2.2 Construction and production of fusion proteins

Fungal CBD gene from *T. reesei* CBHI (*Tr*Cel7A) and its construct were already explained in Chapter 2, i. e., RFP-CBD_{TrCBHI}. To obtain RFP coupled to the *O*-linked glycosylated linker region (RFP-linker), the gene of the RFP-containing linker region was amplified by PCR with primers 5'-TTT*GAATTC*AAAAGAATGGACAACACCGAGGACGTCATC -3' (primer 1) and 5'- AGGTCCTTGAAACAAAACTTCCAAGGGAGAGCTTCCAGTGGTAGTGGC -3' (primer 2) using plasmid vector pPICZαA/RFP-CBD as a template. The CBD region at the C terminus was also amplified with primers 5'-TTGGAAGTTTTGTTTCAAGGACCTACCCAGTCTCACTACGGCCAGTG -3' (primer 3) and 5'-TTT*GCGGCCGC*TTAATGATGATGATGATGATGAGCACCTGGCAGGCACTGA GAGTAGTAAGGGTTCAGGAC -3' (primer 4). The restriction enzyme site and cleavage site of PreScission protease are indicated in italic and by underlining, respectively. The PCR products were subjected to second PCR with primers 1 and 4 to obtain the entire gene, which was inserted into Blunt-TOPO vector. The structure was confirmed by DNA sequencing. The gene was transferred to the *Eco*RI-*Not*I site of *E. coli-P. pastoris* shuttle vector pPICZαA. The constructed plasmid (pPICZαA/RFP-PreScission-CBD-his6) was used to transform *P. pastoris* cells and the fusion protein was expressed as described previously (see chapter 2).

4.2.3 Preparation of fusion proteins

The fusion proteins (RFP-PreScission-CBD-his6 and RFP-CBD) were purified by Avicel column chromatography as described in Chapter 3. In addition, the RFP region was obtained from RFP-CBD by papain cleavage as also described in Chapter 3. To obtain the RFP-linker region without CBD, purified RFP-PreScission-CBD-his6 was cleaved with an appropriate amount of PreScission protease in 20 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl and 1 mM DTT for 24 hour at 5°C. Further purification on an anion-exchange column (DEAE-Toyopearl, Tosoh Corporation, Japan) was performed, after replacement of the solvent with 20 mM Tris-HCl buffer (pH 8.0). All protein samples were concentrated by ultrafiltration and reconstituted in 20 mM sodium acetate buffer (pH 5.0). The purity of the three proteins was confirmed by SDS-PAGE.

4.2.4 Adsorption experiments

Adsorption study of the fusion proteins was performed as follows: various concentrations of RFP-CBD, RFP-linker region, or RFP (40 μL) were mixed with 80 μL of a suspension of cellulose (I_α, III_I or PASC) in 20 mM acetate buffer (pH 5.0) at a final concentration of 0.1 %. After incubation for 2 hours at 5°C, the suspension was centrifuged (14,500 rpm for 3 minutes at the same temperature) and the concentration of protein in the supernatant was calculated from the UV absorption at 559 nm using a molar extinction coefficient of 26,100 $M^{-1}cm^{-1}$ for RFP in 20 mM acetate buffer (pH 5.0). The adsorption amounts were evaluated by subtraction to amounts in supernatant

of the samples with cellulose from its without cellulose. Other series of adsorption measurements were also performed at 15 and 25°C for RFP-CBD. The measurements of pH dependency were performed using approximately 3.4 μM of RFP-CBD_{TrCBHI} in 100 μL volume containing 0.1% substrates at room temperature. 20 mM of each sodium acetate buffer in range of pH 3.89-5.5 and sodium phosphate buffer in range of pH 5.54-7.98 were tested.

4.2.5 Analysis of adsorption parameters

All fittings were performed on IGOR Pro software (version 6.22J) and the correlation coefficients for following models were calculated by using Delta Graph software (version 5). For the monomial Langmuir's model, the following equation was used:

$$[B] = \frac{Amax[F]}{Kad^{-1}+[F]} \quad \text{(Equation 1.)}$$

where [B] and [*F*] are the bound amount and free concentration of RFP-CBD, respectively. A_{max} and K_{ad} are the maximal concentration of bound protein (nmol/mg-cellulose) and the association constant for the binding site (μM^{-1}), respectively. Single reciprocal plots in the Scatchard form ([B]/[*F*] versus [B]) were also fitted to a model of one-dimensional overlapping binding sites with non-cooperative binding [8, 9] as follows:

$$\frac{[\theta]}{[F]} = Kad(1-n[\theta])\left(\frac{1-n[\theta]}{1-(n-1)[\theta]}\right)^{n-1} \quad \text{(Equation 2.)}$$

where [θ] and [*F*] are coverage of bound CBD per maximal binding sites and free concentration, respectively. K_{ad} and *n* are the association constant and the number of lattices occupied by a single CBD. If there is cooperativity of binding of CBD

molecules on the binding surface, the binding would change with surface coverage of CBD fusion protein. Thus, the data points were also fitted to the empirical Hill's equation [10].

$$[B] = \frac{Amax[F]^n}{K_{ad}^{app^{-1}} + [F]^n} \quad \text{(Equation 3.)}$$

$$Kd = (K_{ad}^{app^{-1}})^{1/n} \quad \text{(Equation 4.)}$$

where K_{ad}^{app} and n are apparent association constant and parameter for cooperativity, respectively, and A_{max} is the maximal adsorption capacity. The dissociation constant K_d was calculated from the values of K_{ad}^{app} and n using Eq. 4. Furthermore, the empirical Freundlich equation, which is often used for heterogeneous surfaces, was also tested:

$$[B] = a[F]^{1/m} \quad \text{(Eq. 5)}$$

where a and m are the Freundlich equilibrium constant and the power term of the Freundlich isotherm [11].

4.2.6 Estimation of isosteric enthalpy of adsorption

The enthalpy change of adsorption was estimated indirectly by means of the Clausius-Clapeyron equation (the natural logarithm of free concentration is plotted against the reciprocal of absolute temperature) as follows:

$$\frac{d \ln [F]}{dT} = \frac{\Delta H^0}{RT^2} \quad \text{(Eq. 6)}$$

where $[F]$ and ΔH^0 are free concentration and enthalpy change of adsorption. R and T are the gas constant (8.31451 $JK^{-1}mol^{-1}$) and absolute temperature, respectively. In

order to estimate the dependence of ln [F] on surface coverage (or surface density), Hill's equation was transformed as follows:

$$\ln[\mathrm{F}] = \frac{1}{\mathrm{n}} \ln\left(\left(\frac{[\rho]}{1\text{-}[\rho]} \right) \left(\frac{1}{K_{ad}^{app}} \right) \right) \quad \text{(Eq. 7)}$$

where [ρ] is the surface density of bound protein on cellulose substrate [4]. The isosteric enthalpies were plotted versus surface density.

4.3 Results

4.3.1 pH dependency on binding to celluloses

RFP-CBD was bound on all cellulose matrices I tested in range of pH 4 to 8. The amount of the fusion protein bound to PASC was higher than to I_α or III_I at any pH range (Fig. 1). Although the relative maximum adsorbed amounts to I_α and PASC were achieved at pH 4.68, and their optimal pH values of binding estimated by fitting to single gaussian equation were shown similar values (4.90 and 4.97 for cellulose I_α and PASC, respectively), but its for III_I was slightly different (achieved at pH 5.5.).

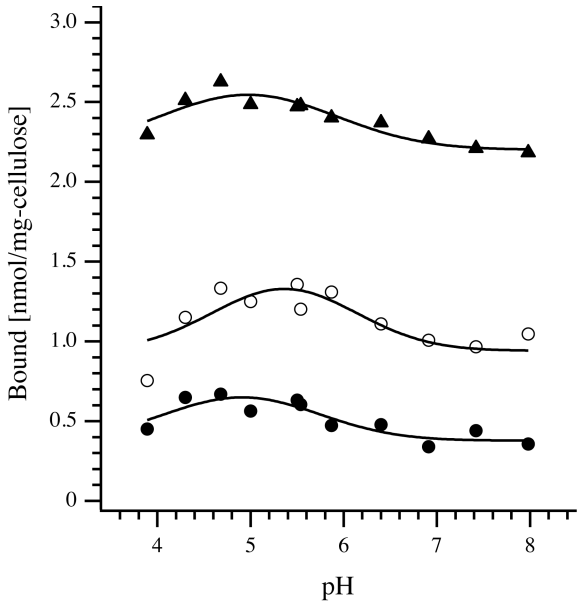


Fig. 1. pH dependency of the adsorption of RFP-CBD to celluloses at room temperature. The amounts of RFP-CBD adsorbed to cellulose I_α, III_I and PASC are shown in closed circles, open circles and closed triangles, respectively. The solid lines were drawn by single gaussian.

4.3.2 Adsorption isotherms

All adsorption measurements were performed in 20 mM sodium acetate buffer (pH 5.0) containing 0.1% matrices. The dependence of the amount of RFP-CBD adsorbed on crystalline cellulose I_α on a free protein concentration is plotted in Fig. 2A and 2B. The corresponding Scatchard plot is shown in Fig. 2C. Although RFP-CBD does not have a catalytic domain, the Scatchard plot of this fusion protein was not linear, but was concave-downward, as has been reported for the intact enzyme containing both CD and CBD. Since RFP and RFP-linker could not bind to any of the celluloses used here (data not shown), the non-linear relationships in the Scatchard plot do not arise from interaction of RFP or the linker region with cellulose, but may be a general feature of binding of proteins to a limited surface area. Therefore, I applied several adsorption models to analyze the data shown in Fig. 2A-C. The correlation coefficients are listed in Table 1. Among the models tested, the Hill's model with negative cooperativity gave a better correlation with the observed data than did the Langmuir model, overlapping-site model, or the Freundlich model.

Curve fitting of the experimental data to the Hill's model was also employed to examine the temperature and substrate dependences, as shown in Fig. 3. The adsorption parameters obtained are summarized in Table 2. The values of A_{max} differed among the substrates tested, but were independent of temperature. The A_{max} values for cellulose III_I and PASC were 1.4-1.5 and 7.0-7.3 times higher than that of cellulose I_α, respectively, suggesting that the surface areas available for adsorption of the RFP-CBD fusion protein on these substrates are larger than in the case of native crystalline cellulose. In contrast, the apparent association constant (K_{ad}^{app}) tended to decline with

increasing temperature. This phenomenon agrees well with the previous report confirming the temperature dependency of the adsorption to crystalline cellulose I [12], indicating that the interaction of fungal CBD with celluloses has a negative enthalpy change of adsorption. The K_{ad}^{app} values for cellulose I_α and PASC were quite similar, while that for cellulose III_I was 2.3-3.1 times higher than the value for cellulose I_α at the same temperature. The value of the dissociation constant (K_d; free protein concentration affording half-maximal adsorption) for cellulose III_I was apparently lower than those for the other cellulose samples, indicating that CBD is adsorbed with higher affinity on the surface of cellulose III_I than on that of cellulose I_α. The cooperativity (n) was less than 1.0 for all cellulose samples tested, suggesting that the negative cooperative adsorption is not substrate-dependent, but is a general feature of RFP-CBD adsorption on celluloses. Notably, the cooperative parameters (n) were increased relative to cellulose I_α, while the values for III_I and amorphous cellulose were remained unchanged.

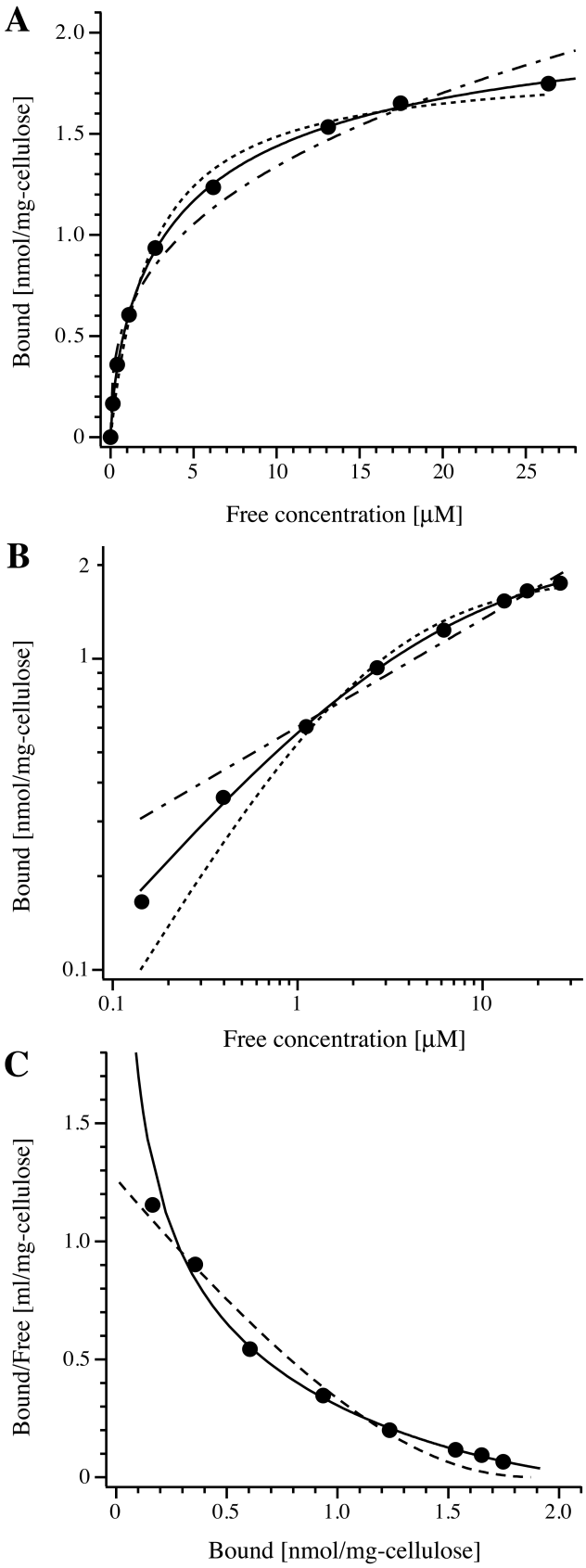


Fig. 2. Analysis of the adsorption of RFP-CBD on crystalline cellulose I_a at 5°C. A, concentration dependence of adsorption ([*F*]-B plot); B, double-logarithmic plot of data in A; C, Scatchard plot. Fittings to Hill's model, the Langmuir single-binding

model, Stankowski's overlapping-site model, and the Freundlich model are shown by solid, dotted, dashed, and dotted-dashed lines, respectively.

Table 1. Correlation coefficients obtained from non-linear regression analysis of binding data points for RFP-CBD to crystalline cellulose at 5°C.

Binding model	Correlation coefficient (R^2)
Hill	0.9995
Langmuir	0.9915
Stankowski overlapping site	0.9770
Freundlich	0.9763

Correlation coefficients were calculated with Delta Graph version 5.

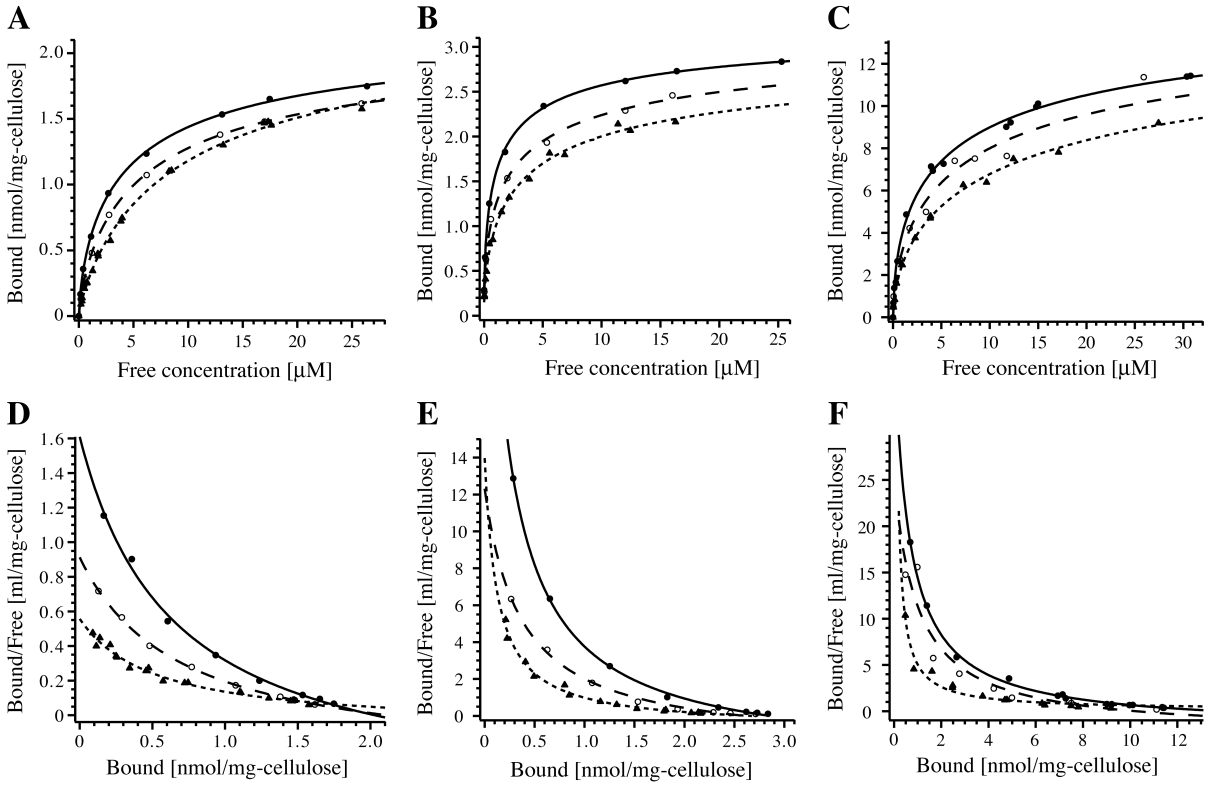


Fig. 3. [*F*]-B plots (A-C) and Scatchard plots (D-F) of RFP-CBD fusion protein at various temperatures on various cellulose substrates. The binding experiments were performed using cellulose I_a (A, D), cellulose III_I (B, E), and PASC (C, F), respectively. Binding data were obtained at 5˚C (closed circular), 15˚C (open circular), and 25˚C (closed triangles), respectively.

Table 2. Adsorption parameters of RFP-CBD on celluloses estimated according to Hill's model.

Cellulose	Temperature	K_{ad}^{app}	n	A_{max} (nmol/mg-cellulose)	K_d (μM)
I_α	5˚C	0.34 ± 0.02	0.70 ± 0.03	2.3 ± 0.1	4.7
	15˚C	0.25 ± 0.01	0.76 ± 0.02	2.2 ± 0.04	6.2
	25˚C	0.15 ± 0.01	0.83 ± 0.03	2.3 ± 0.1	9.8
III_I	5˚C	0.91 ± 0.05	0.58 ± 0.02	3.3 ± 0.1	1.2
	15˚C	0.58 ± 0.10	0.54 ± 0.05	3.3 ± 0.3	2.7
	25˚C	0.47 ± 0.06	0.58 ± 0.04	3.1 ± 0.3	3.7
PASC	5˚C	0.34 ± 0.03	0.58 ± 0.04	16 ± 1	6.4
	15˚C	0.26 ± 0.06	0.61 ± 0.11	16 ± 3	9.1
	25˚C	0.20 ± 0.03	0.60 ± 0.04	16 ± 2	15

4.3.3 Thermodynamic analysis of the adsorption

The value of *n* for cellulose I_α was temperature-dependent, i.e., an increase of the temperature resulted in a decrease of negative cooperativity, whereas the *n* values for cellulose III_I and PASC were almost independent of temperature (Table 2). In order to clarify the difference of adsorption kinetics among the cellulose samples tested, the surface density dependence of the enthalpy change of adsorption (isosteric enthalpy) was evaluated. The slopes of Clausius-Clapeyron plots were calculated at various levels of surface coverage (Fig. 4A-C), and the estimated isosteric enthalpy was plotted versus arbitrary surface density (Fig. 4D-F). The isosteric enthalpy for cellulose I_α decreased markedly with increasing surface density (Fig. 4D). In contrast, III_I and PASC showed little change of the isosteric enthalpy in response to change of surface density; the value remained in the range between -40 kJ/mol and -30 kJ/mol (Fig. 4E and F).

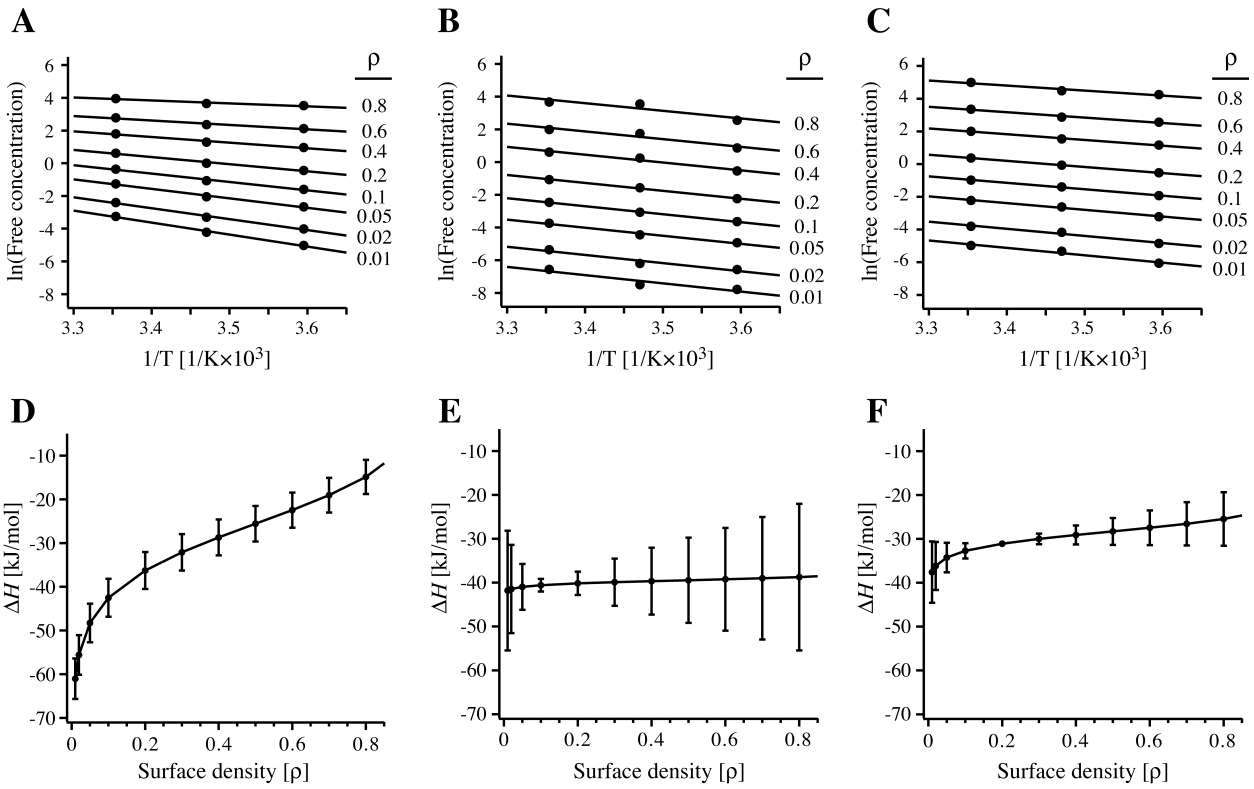


Fig. 4. Dependence of isosteric enthalpy of adsorption on surface density. Clausius-Clapeyron plots (A-C) and plots of the slopes of the Clausius-Clapeyron equation versus arbitrary surface density (D-F) are shown for cellulose I_a (A, D), cellulose III_I (B, E), and PASC (C, F). The values of ln[*F*] were calculated according to Hill's model with arbitrary surface density values of 0.01 to 0.8.

4.4 Discussion

Efficient enzymatic hydrolysis of crystalline cellulose requires effective adsorption of cellobiohydrolases, which is mediated by the cellulose-binding domain (CBD). However, the mechanism of adsorption and the factors determining the efficiency of degradation of crystalline cellulose are not yet fully understood. In the present study, I examined the kinetics of adsorption on several types of cellulose by using a CBD fusion protein, in which the catalytic domain (CD) of *Tr*CBHI is replaced with RFP, a protein of a similar size that does not interact with cellulose.

I initially expected that RFP-CBD would show monomial Langmuir-type adsorption on crystalline cellulose, with a linear Scatchard plot, because the fusion protein has only a single domain with affinity for cellulose. However, the Scatchard plot of this fusion protein was not linear, but was concave-downward, like that reported for the intact enzyme with two binding domains (i.e., CD and CBD). Among several models tested, I found that Hill's equation showed the best fit to the obtained data. This model, with a negative cooperativity parameter (n<1), is often applied to interpret the affinity change of substrates (equilibrium constant) to multimeric proteins that show an induced-fit mechanism, whereby the first bound substrate induces a structural change of a neighboring binding pocket for a second substrate having lower affinity [13]. In the present study, the negatively cooperative adsorption of CBD on cellulose could be interpreted in terms of steric exclusion and/or repulsion by bound CBD with respect to unbound CBD approaching nearby regions of the surface, as already suggested for lectin-carbohydrate interactions [14]. This steric exclusion effect was supported by geometrical feature of the interaction between the CBD and

cellulose. Fungal CBD principally belongs to carbohydrate-binding module (CBM) family 1 (CBM1), and the structures from *Tr*Cel7A had been determined by nuclear magnetic resonance analysis [15]. The CBD is composed of 36 amino acids, and each distance of solvent-exposed three aromatic residues formed planar binding face is consisted with an interval of cellobiose unit in a cellulose chain. Thereby, it has been assumed that the CBD is bound to the hydrophobic 110 face of crystalline cellulose, which agrees with microscopic observations [16, 17]. Namely, the binding of CBD is theoretically expected to occupy an area of 2 nm x 3 nm at the hydrophobic surface of cellulose (at least 2.5 cellobiose units of cellulose chain) [15, 18], then CBD molecule could be closely approached another CBD molecules on the surface. Moreover, intermolecular RFP has a diameter of 4 nm and a length of 5 nm [19], similar size to CD of *Tr*Cel7A [20] as shown in Scheme 1A. Therefore, the RFP domain of the fusion protein as well as CD could interact with nearby molecules before CBD occupies the whole surface of cellulose. In the RFP-CBD fusion protein, moreover, the two domains are connected by a flexible linker region [21], as in the intact enzyme, so the surface area effectively occupied by a single molecule should be larger than the actual size of the fusion protein, as represented in Scheme 1B. Indeed, a recent single-molecule study of binding of CBD fused with green-fluorescent protein (GFP) on crystalline cellulose suggested that a larger space than the theoretical size of the molecule is occupied by the fusion protein because of the fluctuation of the GFP region [17]. Therefore, I conclude that the reason for the non-linear Scatchard plot of RFP-CBD is a steric exclusion effect.

The negative cooperativity was relieved with increase of temperature when cellulose I_α was used as a substrate, but not cellulose III_I or PASC. Namely, the binding site of the CBD to cellulose I_α was more independent at high temperature. However, since the A_{max} was not decreased at higher temperature, the observed more independent adsorption was suggesting that relieve the steric exclusion effect on the same surface area.

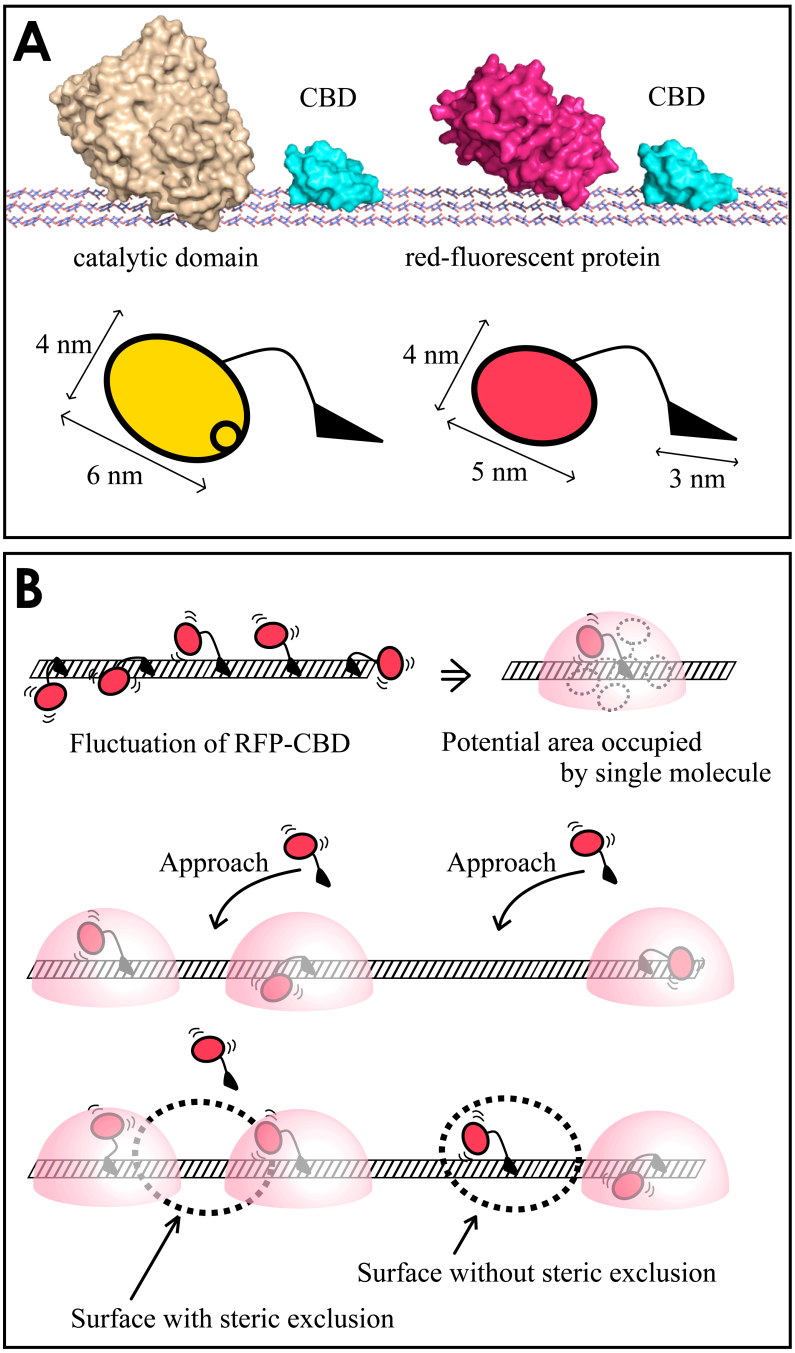
The values of isosteric enthalpy for these substrates were compared in order to clarify the reason for this difference of adsorption kinetics. The isosteric enthalpy for cellulose I_α was highly dependent on surface density, whereas that of cellulose III_I or PASC was more or less coverage-independent. It has been suggested that CBD molecules can diffuse or slide on the substrate surface under the influence of neighboring CBD molecules [22, 23]. Such surface diffusion of CBD molecules would increase the entropy, and may account for the reduction of enthalpy of adsorption [14, 24, 25]. In other words, steric exclusion could be relieved, and the *n*-value would become close to 1.0, on cellulose I_α as a result of surface diffusion of CBD. These results could be interpreted as indicating that fungal family 1 CBD shows uni-order adsorption on cellulose I_α, but shows less-ordered and/or multi-patterned adsorption on cellulose III_I and amorphous cellulose.

In the case of intact *Tr*Cel7A, high- and low-affinity adsorption according to Langmuir's two-binding-site model has so far been assumed to correspond to productive- and non-productive binding, respectively. However, the results obtained in the present study suggest that Langmuir's two-site model cannot represent the actual populations of productive and non-productive enzyme molecules. Therefore, I

reanalyze the previously reported adsorption data for *Tr*Cel7A [5] in terms of Hill's model with negative cooperativity. These data were well fitted by Hill's equation with negative cooperativity, and the adsorption parameters obtained from the reanalysis are listed in Table 3. The adsorption maximum (A_{max} value) for cellulose I_α is the same as that of RFP-CBD fusion protein, though A_{max} for cellulose III_I was slightly increased by the presence of the CD. The K_{ad}^{app} and n values were slightly affected by the exchange of RFP for CD, but the general tendency is similar for the two adsorbates, i.e., the n value for cellulose I_α is higher than that for cellulose III_I, whereas the K_{ad}^{app} value for cellulose III_I is higher than that for cellulose I_α, indicating that *Tr*Cel7A has higher affinity for cellulose III_I, although the negative cooperativity of its adsorption is higher on cellulose III_I than on cellulose I_α.

The negative cooperativity for cellulose III_I was shown to be higher than cellulose I_α as like to RFP-CBD. This high steric exclusion effect of *Tr*Cel7A on the surface of III_I might be leading the efficient decomposition of cellulose because of the ease of traffic congestion on the surface. Anyhow, all the adsorption parameters thus obtained for intact *Tr*Cel7A are similar to those obtained here for RFP-CBD fusion protein, which is able to bind on the cellulose surface only in a non-productive manner, suggesting that the population of "truly" productive molecules is quite small. The above conclusion is also supported by high-speed atomic force microscopic (HS-AFM) observations of *Tr*Cel7A, which indicated that the apparent velocity of movement of enzyme on the substrate surface does not reflect the specific activity of adsorbed enzymes [4, 7, 26]. It seems that further work is needed in order to

understand the relationship between adsorption and catalysis in sufficient detail to optimize cellulose hydrolysis by cellulases.



Scheme 1. Schematically representation of two-domain CBD molecues on cellulose surface. A, comparison of molecular size of intact *Tr*Cel7A (left) and RFP-CBD (right) on cellulose; B, occupancy of two-domain RFP-CBD on a cellulose surface. The affinity for the adsorption of fusion protein decreases with increase of the surface density, even though the surface is not saturated.

Table 3. Adsorption parameters of intact *Tr*Cel7A estimated according to Hill's model.

		K_{ad}^{app}	n	A_{max} (nmol/mg-cellulose)	K_d (μM)
*Tr*Cel7A[a]	I_α	0.49 ± 0.07	0.78 ± 0.08	2.3 ± 0.2	2.5
	III_I	1.2 ± 0.6	0.51 ± 0.15	4.0 ± 0.8	0.73

[a] These values were obtained by reanalysis of previous data [5].

4.5 References

[1]. **Tomme P, Van Tilbeurgh H, Pettersson G, Van Damme J, Vandekerckhove J, Knowles J, Teeri T, Claeyssens M**: Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. *Eur J Biochem* (**1988**), 170:575-581.

[2]. **Reinikainen T, Teleman O, Teeri TT**: Effects of pH and high ionic strength on the adsorption and activity of native and mutated cellobiohydrolase I from *Trichoderma reesei*. *Proteins* (**1995**), 22:392-403.

[3]. **Ståhlberg J, Johansson G, Pettersson G**: A new model for enzymatic hydrolysis of cellulose based on the two-domain structure of cellobiohydrolase I. *Bio/Technol* (**1991**), 9:286-290.

[4]. **Igarashi K, Wada M, Hori R, Samejima M**: Surface density of cellobiohydrolase on crystalline celluloses. A critical parameter to evaluate enzymatic kinetics at a solid-liquid interface. *Febs J* (**2006**), 273:2869-2878.

[5]. **Igarashi K, Wada M, Samejima M**: Activation of crystalline cellulose to cellulose III_I results in efficient hydrolysis by cellobiohydrolase. *Febs J* (**2007**), 274:1785-1792.

[6]. **Wada M, Chanzy H, Nishiyama Y, Langan P**: Cellulose III_I crystal structure and hydrogen bonding by synchrotron X-ray and neutron fiber diffraction. *Macromolecules* (**2004**), 37:8548-8555.

[7]. **Igarashi K, Koivula A, Wada M, Kimura S, Penttilä M, Samejima M**: High-speed atomic force microscopy visualizes processive movement of *Trichoderma reesei* cellobiohydrolase I on crystalline cellulose. *J Biol Chem* (**2009**), 284:36186-36190.

[8]. **McGhee JD, von Hippel PH**: Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J Mol Biol* (**1974**), 86:469-489.

[9]. **Stankowski S**: Large-ligand adsorption to membranes. 1. Linear ligands as a limiting case. *Biochim Biophys Acta* (**1983**), 735:341-351.

[10]. **Hill AV**: The possible effects of the aggregation of the molecules of hemoglobin on its dissociation curves. *Proc physiol soc* (**1910**):4-7.

[11]. **Medve J, Ståhlberg J, Tjerneld F**: Isotherms for adsorption of cellobiohydrolase I and II from *Trichoderma reesei* on microcrystalline cellulose. *Appl Biochem Biotech* (**1997**), 66:39-56.

[12]. **Linder M, Teeri TT**: The cellulose-binding domain of the major cellobiohydrolase of *Trichoderma reesei* exhibits true reversibility and a high exchange rate on crystalline cellulose. *P Natl Acad Sci Usa* (**1996**), 93:12251-12255.

[13]. **Levitzki A, Koshland DE**: Negative cooperativity in regulatory enzymes. *P Natl Acad Sci Usa* (**1969**), 62:1121-1128.

[14]. **Dam TK, Brewer CF**: Multivalent lectin-carbohydrate interactions energetics and mechanisms of binding. *Adv Carbohydr Chem Biochem* (**2010**), 63:139-164.

[15]. **Kraulis J, Clore GM, Nilges M, Jones TA, Pettersson G, Knowles J, Gronenborn AM**: Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry-Us* (**1989**), 28:7241-7257.

[16]. **Lehtiö J, Sugiyama J, Gustavsson M, Fransson L, Linder M, Teeri TT**: The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *P Natl Acad Sci Usa* (**2003**), 100:484-489.

[17]. **Dagel DJ, Liu Y, Zhong L, Luo Y, Himmel ME, Xu Q, Zeng Y, Ding S, Smith S**: In situ imaging of single carbohydrate-binding modules on cellulose microfibrils. *J phys chem B* (**2011**), 115:635-641.

[18]. **Tormo J, Lamed R, Chirino AJ, Morag E, Bayer EA, Shoham Y, Steitz TA**: Crystal structure of a bacterial family-III cellulose-binding domain: A general mechanism for attachment to cellulose. *Embo J* (**1996**), 15:5739-5751.

[19]. **Yarbrough D, Wachter RM, Kallio K, Matz MV, Remington SJ**: Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-Å resolution. *P Natl Acad Sci Usa* (**2001**), 98:462-467.

[20]. **Divne C, Ståhlberg J, Reinikainen T, Ruohonen L, Pettersson G, Knowles JK, Teeri TT, Jones TA**: The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science* (**1994**), 265:524-528.

[21]. **Beckham GT, Bomble YJ, Matthews JF, Taylor CB, Resch MG, Yarbrough JM, Decker SR, Bu L, Zhao X, McCabe C *et al***: The *O*-glycosylated linker from the *Trichoderma reesei* Family 7 cellulase is a flexible, disordered protein. *Biophys J* (**2010**), 99:3773-3781.

[22]. **Jervis EJ, Haynes CA, Kilburn DG**: Surface diffusion of cellulases and their isolated binding domains on cellulose. *J Biol Chem* (**1997**), 272:24016-24023.

[23]. **Liu Y, Zeng Y, Luo Y, Xu Q, Himmel ME, Smith SJ, Ding S**: Does the cellulose-binding module move on the cellulose surface? *Cellulose* (**2009**), 16:587-597.

[24]. **Tomme P, Creagh AL, Kilburn DG, Haynes CA**: Interaction of polysaccharides with the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC. 1. Binding specificity and calorimetric analysis. *Biochemistry-Us* (**1996**), 35:13885-13894.

[25]. **Boraston AB, Creagh AL, Alam MM, Kormos JM, Tomme P, Haynes CA, Warren RAJ, Kilburn DG**: Binding specificity and thermodynamics of a family 9 carbohydrate-binding module from *Thermotoga maritima* xylanase 10A. *Biochemistry-Us* (**2001**), 40:6240-6247.

[26]. **Igarashi K, Uchihashi T, Koivula A, Wada M, Kimura S, Okamoto T, Penttilä M, Ando T, Samejima M**: Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface. *Science* **(2011)**, 333:1279-1282.

Chapter 5. Conclusion

Population of people in the world has been currently increasing to over 7 billions, and securing of resources such as water, foodstuff and energy is essential to save our future. In this view, plant biomass should be utilized in a more efficient way because of its bounty on the earth. Especially, cellulose is the most abundant material in nature, we should consider effective use of this attractive resource. In these days, efficient decomposition of biomass by cellulolytic enzymes has attracted much attention due to policy for utilizing the renewable resource, which crucially depends on efficient degradation of crystalline cellulose.

CBD plays an essential role in cellulase hydrolysis of crystalline cellulose. Non-catalytic binding proteins containing CBDs have a great advantage for fundamental protein applications such as immobilization/purification of enzymes or identification of the corresponding substrates. Thus, understanding of their function is of fundamental importance for rational utilization of cellulases and their applications. In this thesis, I investigated the adsorption behavior of fungal CBDs and feasibility of their affinity purification using cellulose-CBD interactions.

In chapter 2 and 3, I have described the production and affinity purification of CBD fusion proteins in which the catalytic core domain of cellulases was replaced by red-fluorescent protein. These constructs can be used to investigate the binding characteristics of cellulases depending on the function of CBD. *Pichia pastoris* was used for the expression of the recombinant proteins. The production levels of fusion proteins in the culture filtlate reached over 1 g/L both in the case of RFP-CBD$_{Tr\text{CBHI}}$

and -CBD$_{Pc\text{CBHI}}$. The genome of *P. pastoris* has several genes predicted to contain CBM1 (fungal type CBD). These results indicated that *P. pastoris* is likely to be a very suitable host for the production of fungal type CBD fusion proteins on industrial scale.

CBDs can be used for various applications such as purification and immobilization of enzymes as affinity-tags. Since purification of proteins has been little conducted by use of fungal CBDs, I investigated whether fungal CBDs can be used as fusion tags for affinity chromatography (chapter 3). All the cellulose matrices I tested behaved as affinity carriers, that is, all of the CBD fusion proteins were tightly bound on cellulose in the presence of 1 M ammonium sulfate, and they could be easily eluted from cellulose by distilled water. Hydrophobic proteins such as bovin serum albumin almost did not bind to cellulose under this condition. Thus blocking of negatively charged group on cellulose and increasing the hydrophobic interaction by ammonium sulfate may be a good condition to repel the most proteins except for CBDs. Of the tested celluloses, Avicel was found to be the best affinity carrier to purify the CBD fusion proteins. The yield was approximately 80% for RFP-CBD$_{Tr\text{CBHI}}$ at room temperature without denaturing of the CBD. This simple and efficient purification system appears to be suitable for large-scale purification of target proteins.

Degradability of cellulose by cellulolytic enzymes from the fungus *Trichoderma reesei* is most powerful and among them, cellobiohydrolase I (CBHI) is the most studied cellulase so far. Fungal CBHs hydrolyzing crystalline cellulose share a two-domain structure of catalytic domain (CD) and non-catalytic cellulose-binding domain (CBD). The binding of CD to cellulose is considered to be a ‘‘productive-

binding'', while binding of CBD as a ''nonproductive-binding''. Unfortunately, since the nonproductive-binding interferes with the activity of CD at high surface coverage, understanding and controling of this nonproductive-binding is one of the options for improvement of this generally slow decomposition process. Thereby, in chapter 4, the adsorption characteristics of CBD fusion protein from *Tr*CBHI was studied mainly in order to understand the binding mechanism of CBD and thus two-domain structure. Binding data were better fitted by the Hill's model with negative cooperativity than by other adsorption models, suggesting the occurrence of a steric exclusion effect among the fusion molecules on the cellulose surface. All cellulose samples I tested were well fitted to Hill's model, and the degree of negative cooperativity depended on the nature of cellulose. Notably, the negative cooperativity decreased with increasing temperature only to crystalline cellulose I_α in spite of the affinities were not a little influenced toward all substrates tested by change of temperature. This phenomenon indicates occurrence of a relieve mechanism of steric exclusion on cellulose I_α by the surface diffusion or sliding of CBD fused molecules, which was supported by estimation of isosteric enthalpy. Interestingly, the binding data of intact *Tr*CBHI was also fitted to Hill's model regardless the type of crystalline cellulose. The adsorption parameters of *Tr*CBHI, especially negative cooperativity and maximum adsorption capacity, were similar to CBD fusion protein without CD. The strength of affinity does not participate in the cooperativity at equilibrium condition, i.e., cooperativity may be independent parameter corresponded to nature of cellulose and/or shape, flexibility, and mobility of adsorbed molecules on cellulose surface. Thus, these results indicate that the binding based on CBD predominates in the case of adsorption on cellulose by intact *Tr*CBHI.

Further investigation is needed in order to understand the cooperative relationship between adsorption behavior and catalytic function of cellulase for the efficient degradation of cellulose.

Acknowledgements

本研究を進めるにあたり、実験や論文作成のご指導に加え、学外での研究の機会を与えてくださいました東京大学農学生命科学研究科　森林化学研究室の鮫島正浩教授に心より感謝いたします。本学博士課程入学の際、当分野に関する知識に乏しい私を受け入れてくださり、３年間辛抱強く見守り、折に触れて励まして頂いたことも重ねて感謝申し上げます。また、同研究室の五十嵐圭日子准教授には、日頃から研究に関する様々な助言を頂きました。心より感謝いたします。私の拙い試行錯誤を暖かく見守り、時に厳しい助言も頂き、本当に多くのことを学ぶ機会を与えてくださったことも重ねて御礼申し上げます。また、研究室において様々な形でご助力くださり、よく励ましの言葉をかけてくださった寺田珠美助教、同研究室に入りたての私に、やさしく実験を教えてくださった丸山道子さん、様々なご助言をくださりました黒澤美幸さんに心より感謝申し上げます。

東京大学農学生命科学研究科　生物素材科学研究室の和田昌久准教授には、セルロースの調製や、その構造について多くのご助言を頂きました。心より感謝申し上げます。和田先生には論文の審査において副査も務めて頂きました。重ねて御礼申し上げます。また、同研究室の木村聡助教には、SEM 観察や蛍光顕微鏡観察などの際、快くご指導頂きました。心より感謝いたしております。

産業技術総合研究所の分子スケール光デバイスグループの吉川佳宏博士には、原子間力顕微鏡を用いたセルロースの観察ならびに CBD の一分子脱着力測定の際に大変お世話になりました。実験について熱く指導、ご助言を頂いただけでなく、未熟な私に研究のあり方についてもご指導頂きました。心より感謝申し上げます。また、金里雅敏グループ長をはじめ同グループのみなさまにも、楽しいお茶会や、やさしく接して頂き御礼申し上げます。

また木材化学研究室出身の岡本哲明博士には、蛍光顕微鏡についてのご指導と研究に関する様々なアドバイスを頂きました。心より感謝申し上げます。また、同研究室の山ギシ崇之くんには、いきなり押し掛ける私に嫌な顔一つせず、顕微鏡実験を手伝ってもらいました。心より感謝申し上げます。

森林化学研究室の同期である堀千明さん、石野貴久くんには、研究の話だけでなく、様々な場面で相談に乗って頂き心より感謝申し上げます。また、博士課程１年の中村彰彦くん、修士２年の櫻木潔くんには、様々な実験でお世話になりました。心より感謝申し上げます。また、同研究室の先輩である石田卓也博士、鈴木一史博士、和田朋子博士、石黒真希博士、そして平石正男さんには、研究の様々な点でサポートをして頂くだけでなく、研究者としての姿勢や考え方を学ばせて頂きました。心より感謝いたします。

また名城大学農学部応用微生物学研究室の志水元亨助教には本研究室に来るきっかけを作ってくださっただけでなく、研究についてや日頃から様々な面で支えて頂きました。この場を借りて心より感謝申し上げます。

JX ホールディングス株式会社様より多くの研究活動費を支援して頂きました。御礼申し上げます。

I would be deeply grateful to Dr. Peter Biely in Institute of Chemistry Slovak Academy of Sciences to correct my thesis and to discuss about biomass degradation. I would like to thank Dr. Markus Linder, Dr. Anu Koivula, and Dr. Merja Penttilä in VTT, for grateful discussion about function of cellulose-binding domain and cellulase, and also to help when I trip to Finland. I would also like to thank all members in their laboratories in VTT, especially Jenni Rahikainen for all of your help.

最後に、これまで私の学生生活を支え、常に暖かく見守り励ましてくれた家族 、いつも心配をかけてしまった母に心より感謝し、謝辞とさせて頂きます。

2012 年 10 月

Appendix