



Domain architecture divergence leads to functional divergence in binding and catalytic domains of bacterial and fungal cellobiohydrolases

Received for publication, June 15, 2020, and in revised form, August 15, 2020. Published, Papers in Press, August 18, 2020, DOI 10.1074/jbc.RA120.014792

Akihiko Nakamura^{1,*}, Daiki Ishiwata^{2,3}, Akasit Visootsat^{2,3}, Taku Uchiyama⁴, Kenji Mizutani⁵, Satoshi Kaneko⁶, Takeshi Murata⁷, Kiyohiko Igarashi⁴, and Ryota Iino^{2,3,*}

From the ¹Department of Applied Life Sciences, Faculty of Agriculture, Shizuoka University, Shizuoka, Shizuoka, Japan, the ²Department of Functional Molecular Science, School of Physical Sciences, SOKENDAI (The Graduate University for Advanced Studies), Hayama, Kanagawa, Japan, the ³Institute for Molecular Science, National Institutes of Natural Sciences, Okazaki, Aichi, Japan, the ⁴Department of Biomaterials Sciences, Graduate School of Agricultural and Life Sciences, University of Tokyo, Tokyo, Japan, the ⁵Graduate School of Medical Life Science, Yokohama City University, Yokohama, Japan, the ⁶Department of Subtropical Biochemistry and Biotechnology, Faculty of Agriculture, University of the Ryukyus, Nishihara, Okinawa, Japan, and the ⁷Department of Chemistry, Graduate School of Science, Chiba University, Inage, Chiba, Japan

Edited by Gerald W. Hart

Cellobiohydrolases directly convert crystalline cellulose into cellobiose and are of biotechnological interest to achieve efficient biomass utilization. As a result, much research in the field has focused on identifying cellobiohydrolases that are very fast. Cellobiohydrolase A from the bacterium *Cellulomonas fimi* (CfCel6B) and cellobiohydrolase II from the fungus *Trichoderma reesei* (TrCel6A) have similar catalytic domains (CDs) and show similar hydrolytic activity. However, TrCel6A and CfCel6B have different cellulose-binding domains (CBDs) and linkers: TrCel6A has a glycosylated peptide linker, whereas CfCel6B's linker consists of three fibronectin type 3 domains. We previously found that TrCel6A's linker plays an important role in increasing the binding rate constant to crystalline cellulose. However, it was not clear whether CfCel6B's linker has similar function. Here we analyze kinetic parameters of CfCel6B using single-molecule fluorescence imaging to compare CfCel6B and TrCel6A. We find that CBD is important for initial binding of CfCel6B, but the contribution of the linker to the binding rate constant or to the dissociation rate constant is minor. The crystal structure of the CfCel6B CD showed longer loops at the entrance and exit of the substrate-binding tunnel compared with TrCel6A CD, which results in higher processivity. Furthermore, CfCel6B CD showed not only fast surface diffusion but also slow processive movement, which is not observed in TrCel6A CD. Combined with the results of a phylogenetic tree analysis, we propose that bacterial cellobiohydrolases are designed to degrade crystalline cellulose using high-affinity CBD and high-processivity CD.

Cellobiohydrolases (CBHs) play key roles in degradation of crystalline cellulose, which is the homopolymer of β -1,4-linked glucose and fundamental component of plant cell wall (1). High hydrolytic activity of CBHs against crystalline cellulose is achieved by the unique structure of the catalytic domain (CD), which consists of tunnel-shaped substrate-binding sites cov-

ered by loops (2). In addition, many CBHs also have the cellulose-binding domain (CBD), and the CD and CBD are connected by the linker region (or domain). The CD and CBD are classified into glycoside hydrolase (GH) and carbohydrate binding module (CBM) families, respectively, according to the amino acid sequences (3). Although cellulases are classified into GH families 5–12, 44, 45, 48, 51, 74, 124, and 148, CBHs are only included in the members of GH6, 7, 9, and 48. The CBMs, which have flat surfaces for cellulose binding (called as type A CBM), are divided into CBM 1, 2, 3, 5, and 10 (4).

Cellulose is the most abundant biomass on earth and an important carbon source for fungi and bacteria. Cellulose degradation system of fungi has been well-known, and they produce many kinds of multidomain cellulases. Synergistic hydrolytic reactions between GH7 and GH6 CBHs (5) or CBHs and endoglucanases (EGs) have been studied in detail. An important cellulose degradation system of bacteria is cellulosome, which is the large complex of carbohydrate active enzymes anchored to the cell surface. The cellulosome system is employed by anaerobic bacteria, and only GH5, 8, 9, and 48 cellulases are reported as components (6). Another degradation system of bacteria is similar to the fungal one. For example, an actinomycetes *Cellulomonas fimi* produces free GH6 and GH48 CBHs (7).

In the process of crystalline cellulose hydrolysis, cellulases first bind on the cellulose surface. However, after the binding, not all cellulases can initiate hydrolysis, because accessible position is limited because of the tight packing of the cellulose chains in the crystal. When a cellulase molecule successfully catches a cellulose chain into the catalytic site, it can form productive complex. In other cases, cellulase binds nonproductively and then dissociates from the cellulose surface without hydrolysis. The unique function of CBH is a unidirectional movement on cellulose surface coupled with processive hydrolysis of the cellulose chain into cellobiose, the minimum repeating unit. The unidirectional movement of CBH has been directly observed by single-molecule imaging techniques recently. For the GH7 CBHs, which is a unique cellulase for fungi, the movement was first proved by high-speed atomic

This article contains supporting information.

✂ Author's Choice—Final version open access under the terms of the Creative Commons CC-BY license.

* For correspondence: Akihiko Nakamura, aki-naka@shizuoka.ac.jp; Ryota Iino, iino@ims.ac.jp.

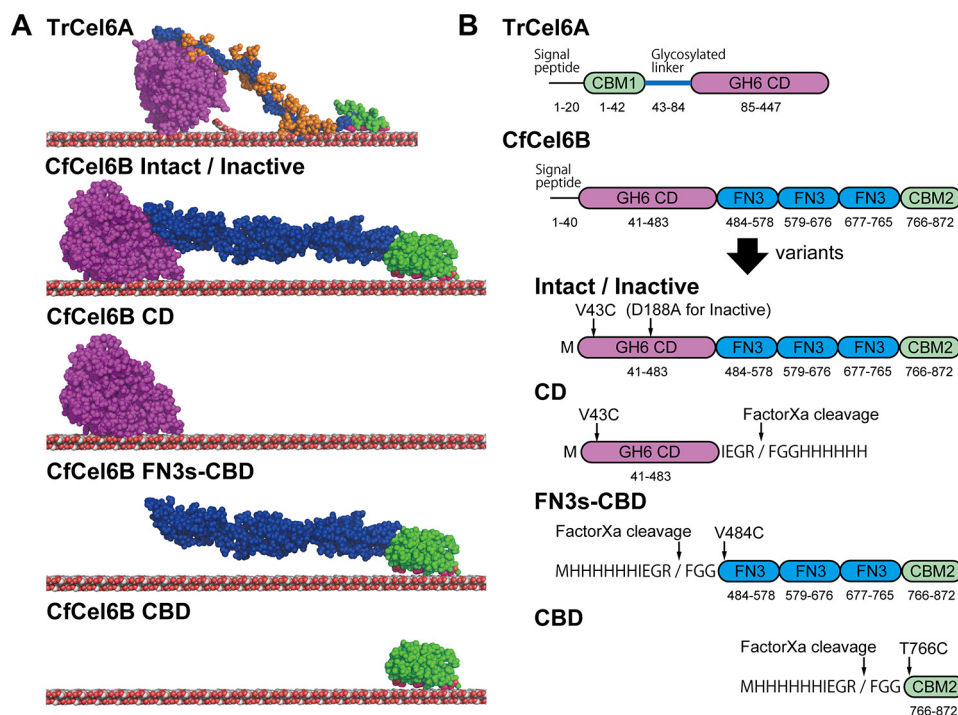


Figure 1. Structures of TrCel6 and CfCel6B. A, model structures of Intact TrCel6A, CfCel6B, and CfCel6B domain constructs used in this study. TrCel6A structure is the same as in the previous report (12). For CfCel6B, structure of CD is X-ray crystal structure (PDB code 7CBD), and FN3s and CBD are modeled by SWISS-MODEL server (38). Figures were prepared by PyMOL. B, detailed descriptions of domain compositions for each construct. Positions of mutation sites, histidine tags, and FaXa cleavage sites and estimated amino acid numbers for each domain are shown.

force microscopy (8, 9). Furthermore, the relationship between lengths of the tunnel-like structure of the CD and processivity has been analyzed experimentally and theoretically (10, 11). Although GH6 CBHs are common enzymes in fungi and bacteria, the movement of GH6 from only an ascomycete *Trichoderma reesei* (TrCel6A) has been observed by single-molecule fluorescence imaging (12). Because GH7 CBH hydrolyzes cellulose from reducing end and GH6 CBH hydrolyzes from nonreducing end of the cellulose chain, they show opposite directionality in the processive movements.

GH6 CBHs from fungi and bacteria are classified into the same family but are different in many points. First, the tunnel-like structure of the substrate-binding site of bacterial CD is longer than that of fungi CD (13). Therefore, bacterial GH6 CBH was expected to be more processive and less endolytic than fungal GH6 CBH. Second, the CBD and linker region are different. TrCel6A has CBM1-CBD connected to CD by a glycosylated linker region (Fig. 1A). The linker region is expected to be intrinsically disordered, and the interaction of sugars on the linker with the cellulose surface is investigated by the molecular dynamics simulation (14). In contrast, GH6 CBH from the bacterium *C. fimi* (CfCel6B) has CBM2-CBD and three fibronectin type 3 domains (FN3s) as a linker between CD and CBD. Both CBDs have flat surface with hydrophobic residues and are expected to bind on the hydrophobic surface of the crystalline cellulose (15, 16). CfCel6B is previously called CbhA and found as the first enzyme that is similar to the CBH II from *T. reesei* (another name of TrCel6A) (17). However, the effects of the different domain composition on the elementary steps of cellulose hydrolysis reaction are still elusive.

In this study, by using single-molecule fluorescence imaging, we observed binding and dissociation of full-length CfCel6B and its domains (CD, CBD, and FN3s-CBD) on the crystalline cellulose to clarify the functions of CBD and FN3s. Furthermore, their movements on the crystalline cellulose were analyzed. Combined with crystal structures of the CD from bacterial and fungal GH6 CBHs (CfCel6B and TrCel6A, respectively), we successfully verified correlation between lengths of the tunnel-like structure of the CD and processivity. Furthermore, in our phylogenetic tree analysis, the CDs from bacteria and fungi were clearly separated. The fungal CBHs have glycosylated linker and CBM1, and bacterial CBHs have CBM2 except for the single domain enzymes. Given the difference of domain compositions between bacterial and fungal CBHs, CBM2 of bacterial CBHs compensates for facilitated initial binding on cellulose by glycosylated linkers of fungal CBHs. Our results indicate the difference of a design principle between bacterial CBHs and fungal CBHs.

Results

Preparation of fluorescently labeled samples

In this study, to conduct single-molecule fluorescence imaging, single free cysteines were introduced on the surface of full-length CfCel6B and its domains. We prepared full-length CfCel6B/V43C (termed as Intact), CD/V43C (CD), FN3s-CBD/V484C (FN3s-CBD), and CBD/T766C (CBD) (Fig. 1A). In addition, a catalytically inactive D188A mutant of full-length CfCel6B/V43C (Inactive), of which catalytic acid aspartate was mutated to alanine, was prepared as a negative control of processive movement coupled with catalysis (Fig. 1B). All of

Single-molecule fluorescence imaging of bacterial cellulase

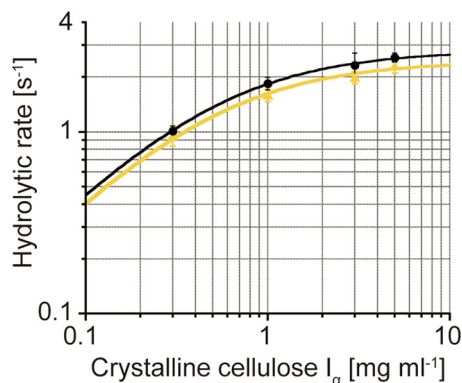


Figure 2. Michaelis–Menten plots of CfCel6B WT and Intact. Hydrolytic rates of WT (black circles) and Intact (yellow triangles) at various crystalline cellulose concentrations (0.3, 1.0, 3.0, and 5.0 mg ml⁻¹) were fitted by Michaelis–Menten equations. The values of K_m for WT and Intact were 0.52 and 0.51 mg ml⁻¹, respectively. The values of k_{cat} for WT and Intact were 2.8 and 2.4 s⁻¹, respectively. R^2 values of the fitting for WT and Intact were 0.98 and 0.99, respectively. Enzyme concentration was 100 nM.

proteins were successfully expressed in *Escherichia coli* and purified using cellulose affinity column or nickel–nitrilotriacetic acid affinity column. After the labeling of the free cysteine with Cy3-maleimide, hydrolytic activities of Cy3-labeled Intact and Inactive were compared with WT CfCel6B. 500 nM WT hydrolyzed 1 mg ml⁻¹ crystalline cellulose I_α at the rate of 0.068 ± 0.001 s⁻¹, and Intact showed comparative hydrolytic rate (0.059 ± 0.001 s⁻¹). On the other hand, hydrolytic rate of Inactive was very low (0.0080 ± 0.0010 s⁻¹), indicating that Inactive do not have cellulose hydrolytic activity. Then we further compared hydrolytic rates for WT and Intact at various concentrations of crystalline cellulose I_α to determine the K_m and k_{cat} values by the fitting with Michaelis–Menten equations (Fig. 2). The plots were fitted well (R^2 values for WT and Intact were 0.98 and 0.99, respectively); k_{cat} and K_m for WT were 2.8 s⁻¹ and 0.52 mg ml⁻¹, and those for Intact were 2.4 s⁻¹ and 0.51 mg ml⁻¹, respectively. Note that hydrolytic activity was also dependent on enzyme concentration, and 100 nM enzymes were used for these measurements. These results indicated that V43C mutation and labeling with Cy3 do not significantly affect the hydrolytic activity and affinity of CfCel6B against crystalline cellulose.

Binding rate constant

To determine the binding rate constant (k_{on}) for Intact, CD, FN3s-CBD, and CBD, enzymes of picomolar concentrations were dropped on the cover glass sparsely coated with crystalline cellulose microfibrils to clearly observe fluorescence signals from individual molecules. The bindings of enzymes were highly specific to the cellulose microfibrils, and almost no non-specific bindings to the glass surface were observed. These results indicate that all of the enzymes have correctly folded structures that recognize the surface of crystalline cellulose. The values of k_{on} were calculated as numbers of bound molecules divided by enzyme concentration, length of cellulose microfibril, and observation time (M⁻¹ μm⁻¹ s⁻¹). We could not directly estimate the number of bundles in the cellulose microfibrils from the fluorescence image stained with nanomo-

lar concentrations of enzymes, because of the limit of spatial resolution of optical microscopy. Therefore, we analyzed the distributions of k_{on} (Fig. 3). Distribution of k_{on} for Intact showed single peak at 4.3×10^8 M⁻¹ μm⁻¹ s⁻¹. Distributions of k_{on} for CD, FN3s-CBD, and CBD showed multiple peaks, which would correspond to number of bundles in cellulose microfibrils. The distributions were fitted well, and their R^2 values were better than 0.93. The smallest peak values for FN3s-CBD and CBD were 2.0×10^8 and 1.5×10^8 M⁻¹ μm⁻¹ s⁻¹, respectively. These values were almost half and one-third of that for Intact. On the other hand, distribution of k_{on} for CD showed a smallest peak at 1.7×10^7 M⁻¹ μm⁻¹ s⁻¹, which was less than one-twentieth of that for Intact. These results indicate that k_{on} of CfCel6B is highly dependent on the binding of the CBD, and FN3s and CD do not contribute significantly. However, if CD and CBD were connected by FN3s, these two domains seem to bind synergistically, because the k_{on} value for Intact is larger than the simple sum of those for CD and FN3s-CBD. Synergistic binding between CBD with linker region and CD has been also observed in TrCel6A (12).

Dissociation rate constant

Next, we analyzed distribution of binding time on cellulose surface. Distributions of binding times were better fitted by sum of two exponential decay functions than single exponential decay in Intact, FN3s-CBD, and CBD (Fig. 4 and Fig. S1), as reported previously for TrCel6A (12). These results indicate that at least two different binding modes of the enzyme exist. For CD, although the R^2 values were same for both fittings, the sum of the two exponential decay functions was used to estimate the fractions of two modes. Fast and slow components of dissociation rate constant (k_{off}^{fast} and k_{off}^{slow} , respectively) for Intact were 0.85 and 0.086 s⁻¹, respectively. The ratios of fast and slow components were 33 and 67% respectively. Those for CD were 1.7 s⁻¹ (81%) and 0.13 s⁻¹ (19%) and increased 1.5–2 times compared with those for Intact. In contrast, k_{off} values and ratios of fast components for FN3s-CBD and CBD were 2.7 s⁻¹ (26%) and 3.1 s⁻¹ (30%), respectively. In addition, those of slow components for FN3s-CBD and CBD were 0.29 s⁻¹ (74%) and 0.47 s⁻¹ (70%), respectively. These k_{off} values were comparable between FN3s-CBD and CBD but increased more than three times compared with those for Intact. These results indicated that cellulose-bound state of CfCel6B is stabilized by CD. We also found that CD showed much higher ratio of fast dissociation (81%) than that of slow dissociation (19%). This result was unique for CD among the four samples (*i.e.* Intact, CD, FN3s-CBD, and CBD), because the ratios of slow dissociation were almost 70% for the other three samples.

Affinity and dissociation constant

The values of the binding rate constant corresponding to the fast and slow components (k_{on}^{fast} and k_{on}^{slow} , respectively) were estimated from the k_{on} shown in Fig. 3 and the ratio of fast and slow components determined by the k_{off} analysis shown in Fig. 4 (Table 1). Then the values of the dissociation constant (K_d) for the fast and slow components (K_d^{fast} and K_d^{slow} , respectively)

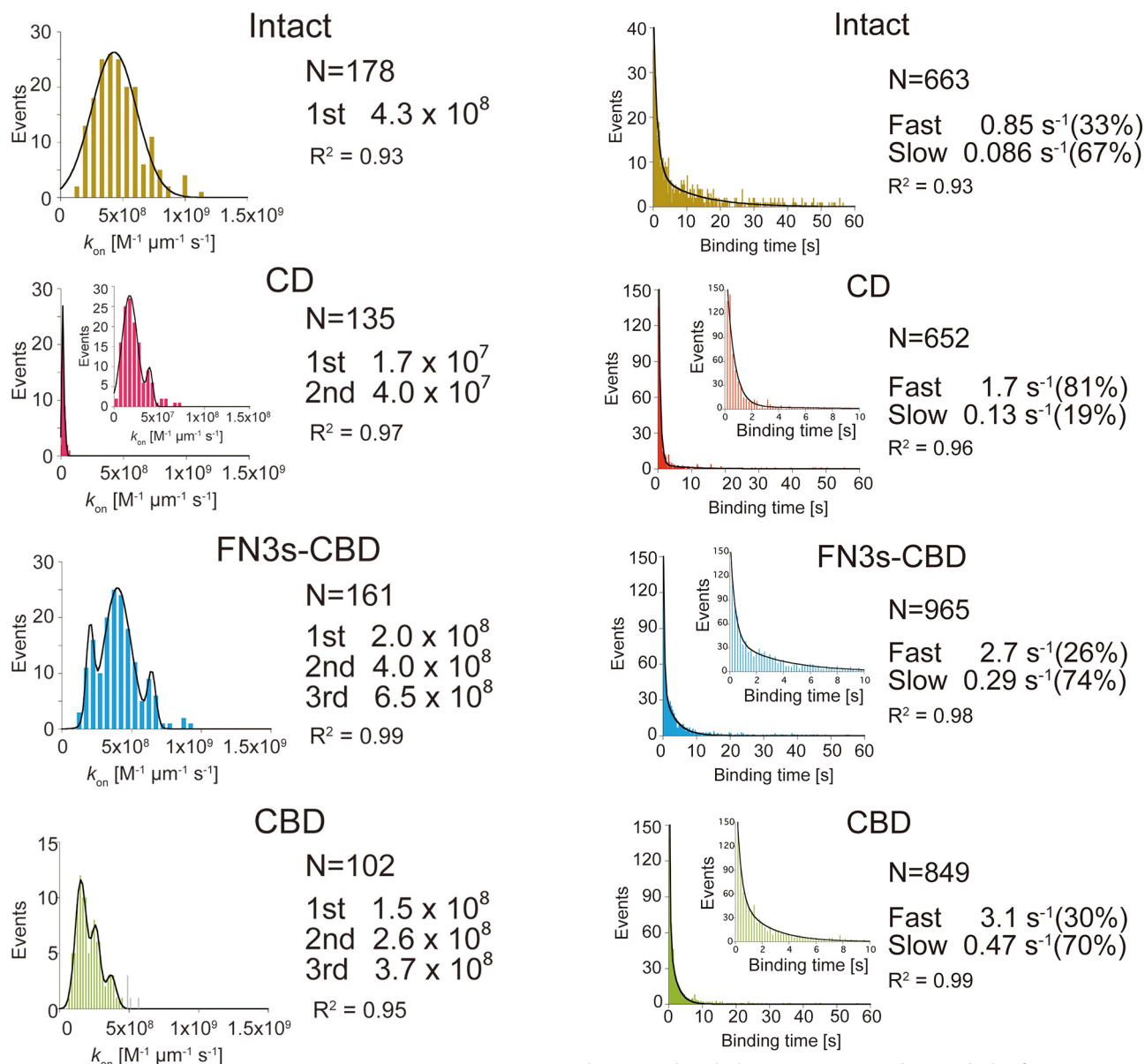


Figure 3. Binding rate constants (k_{on}) analysis of Intact, CD, FN3s-CBD, and CBD of CfCel6B. Distributions of k_{on} were fitted by Gaussian functions. Peak values of Gaussian fitting are shown in right. The N values represent numbers of cellulose microfibrils analyzed.

Figure 4. Dissociation rate constants (k_{off}) analysis of Intact, CD, FN3s-CBD, and CBD of CfCel6B. Distributions of binding times were fitted by sum of two exponential decay functions. The values and ratios of the fast and slow components of the dissociation rate constant (k_{off}^{fast} and k_{off}^{slow}) are shown on the right. The N values represent numbers of analyzed molecules.

were calculated from the ratio of k_{off} to k_{on} (k_{off}/k_{on}). Among them, K_d^{slow} for Intact showed the lowest value (3.0×10^{-10} M μ m). This value was 20 times lower than that of K_d^{fast} for Intact. The values of K_d^{fast} or K_d^{slow} for FN3s-CBD and CBD were comparable and less than 15 times higher than those of Intact. On the other hand, K_d^{slow} for CD was 133 times higher than that of Intact because of the low value of k_{on}^{slow} . The difference of K_d^{fast} values for CD and Intact was 20 times. These results indicate that CBD mainly contributes to the affinity of both fast and slow components of Intact.

Translational rate and processivity

Next, translational rate (k_{tr}) was measured from the distance between first and last positions of movement and moving time

(Fig. 5). We used higher laser power density (0.28μ W μ m $^{-2}$) and lower frame rate (1 fps) than those for binding and dissociation analyses, to achieve higher localization precision (4.5 ± 1.5 and 4.6 ± 1.4 nm for the x and y axes, respectively). Distributions of k_{tr} for Intact and CD could be fitted with sum of two Gaussians (R^2 values were 0.94 and 0.83, respectively). Peak values for Intact were 11.6 and 25.3 nm s^{-1} , and those for CD were 16.8 and 40.2 nm s^{-1} , respectively. On the other hand, distributions of k_{tr} for Inactive and FN3s-CBD could be fitted with a single Gaussian, and the peak values were 37.9 and 39.8 nm s^{-1} , respectively (R^2 values were both 0.82). Some Intact molecules moved more than 10 s, and all of these molecules showed lower k_{tr} near the first peak (Fig. 5, right and top panel). On the other hand, in other samples, no molecule moved more than 10 s (Fig. 5, right panel). These results strongly suggest that Intact

Single-molecule fluorescence imaging of bacterial cellulase

Table 1
Summary of binding-rate, dissociation-rate, and dissociation constants

Sample	k_{on}	Component	$k_{\text{off}}^{\text{fast}}$ or $k_{\text{off}}^{\text{slow}^a}$		$k_{\text{on}}^{\text{fast}}$ or $k_{\text{on}}^{\text{slow}^c}$	K_d^{fast} or $K_d^{\text{slow}^d}$
			Value	Ratio ^b		
	$M^{-1} \mu M^{-1} s^{-1}$		s^{-1}	%	$M^{-1} \mu M^{-1} s^{-1}$	$M \mu M$
Intact	4.3×10^8	Fast	0.85	33	1.4×10^8	6.0×10^{-9}
		Slow	0.086	67	2.9×10^8	3.0×10^{-10}
CD	1.7×10^7	Fast	1.7	81	1.4×10^7	1.2×10^{-7}
		Slow	0.13	19	3.2×10^6	4.0×10^{-8}
FN3s-CBD	2.0×10^8	Fast	2.7	26	5.2×10^7	5.2×10^{-8}
		Slow	0.29	74	1.5×10^8	2.0×10^{-9}
CBD	1.5×10^8	Fast	3.1	30	4.5×10^7	6.9×10^{-8}
		Slow	0.47	70	1.1×10^8	4.5×10^{-9}

^aThe $k_{\text{off}}^{\text{fast}}$ and $k_{\text{off}}^{\text{slow}}$ are the fast and slow components of the dissociation rate constant, obtained by the fitting of the distribution of binding time distribution (Fig. 4) with a double exponential decay function.

^bThe ratios of fast and slow components were calculated from the ratio of the area of each fitted exponential decay function.

^cThe k_{on} values determined in Fig. 3 were further divided into $k_{\text{on}}^{\text{fast}}$ and $k_{\text{on}}^{\text{slow}}$ by using the ratio of fast and slow components of k_{off} .

^dThe K_d^{fast} and K_d^{slow} values were calculated as $k_{\text{off}}^{\text{fast}}/k_{\text{on}}^{\text{fast}}$ and $k_{\text{off}}^{\text{slow}}/k_{\text{on}}^{\text{slow}}$, respectively.

molecules that showed a long moving time (>10 s) correspond to those moving processively, and Intact molecules that showed a short moving time (<10 s) are a mixture of those moving processively and diffusing on the cellulose surface. Therefore, the distribution of moving times for Intact was separately fitted to all ranges except for the first bin or the ranges longer than 10 s (Fig. 6). The time constant of moving time for the former was 6.6 s ($R^2 = 0.94$) and that for the latter was 4.6 s ($R^2 = 0.85$). These values were both shorter than the values for TrCel6A (7.7 s in both fittings) previously reported (12).

Crystal structure of CfCel6B CD

We solved a crystal structure of CfCel6B CD with 1.3 Å resolution (Table S1), to clarify the structural difference between CfCel6B and TrCel6A CDs (Fig. 7). The structure of CfCel6B CD was modeled by SWISS-MODEL server based on a GH6 CBH from a bacterium *Thermobifida fusca* (TfCel6B; PDB code 4AVO) and was used as a template of molecular displacement (13). In the determined crystal structure of CfCel6B CD, we confirmed that the Val⁴³, which is close to the N terminus of CfCel6B (without signal peptide) and mutated to Cys in the single-molecule fluorescence imaging, was located in the opposite side of the catalytic site.

In the crystal structure, CfCel6B has additional substrate-binding site (subsite) constructed by Trp³⁰³ at plus side, the same as TfCel6B (Fig. 7, right panel). This additional subsite is stabilized by two loops (Fig. 7, left panel, shown in blue) which are not found in TrCel6A (Fig. 7, middle panel). Near the product-binding site, a pair of exit loops has been also found in CfCel6B and TfCel6B. The exit loop 1 (shown in green) was capping the end of the substrate-binding tunnel. This exit loop 1 of CfCel6B is 4 amino acids shorter than that of TfCel6B, indicating that the tunnel of CfCel6B is more open than that of TfCel6B. In addition, the exit loop 2 (shown in cyan) of CfCel6B showed more open conformations compared with that of TfCel6B.

The conformation of the N-terminal loop of CfCel6B was more open compared with those of TrCel6A and TfCel6B. The serine residue (Ser¹⁰⁵, Ser¹⁸¹, and Ser²³² in CfCel6B, TrCel6A, and TfCel6B, respectively), which is thought to be important to form a hydrogen bond network among the water molecules

and catalytic residues, exists on the N-terminal loop (18). The Ser¹⁰⁵ and the N-terminal loop of CfCel6B, crystallized without ligand, stayed outside of the cleft. In contrast, the serine residues interact with the ligand, and the N-terminal loops face to inside of the cleft in the structures of TrCel6A and TfCel6B. The conformational change of the N-terminal loop caused by the interaction with ligand has been reported previously for the GH6 CBH from a basidiomycete (19).

Comparison of GH6 CDs from fungi and bacteria

To discuss the relationship between the structure and function of CD, we compared all CDs of GH6 enzymes listed in CAZy database (RRID:SCR_012909), except the enzymes with only patent information or including unknown residues in the sequence. By the structural alignment of X-ray crystal structures and homology modeled structures of CD, five groups were identified (Fig. 8). In the phylogenetic tree diagram, bacterial and fungal cellulases were clearly separated. Interestingly, EGs and CBHs from aerobic bacteria and fungi formed different groups, but those of anaerobic fungi were mixed.

Each group showed characteristic domain compositions, although the phylogenetic tree was prepared based on the sequences of only CD (Fig. 8, right panel). For example, fungal EGs did not have CBD, and many of CBHs were constructed by CD and CBM1-CBD with serine- or threonine/proline-rich, glycosylated linker region. Many anaerobic fungal cellulases had a glycosylated linker and two CBM10-CBDs that show ~6 times weaker affinity than that of CBM2-CBD (20). On the other hand, bacterial EGs had three types of domain compositions. One had only CD, and the other two had CD with CBM2-CBD in the N or C terminus. Furthermore, bacterial CBHs showed completely different compositions. CBM2-CBD was the major component, although CBM3-CBD and CBM10-CBD were also found. In addition, two of them had an additional CD domain classified into GH5 or GH12 EGs.

Given the domain composition lists of CBHs from fungi and bacteria, the CBM1 with the glycosylated linker is the common domains for fungal CBHs. The CBM2 is the common domain for bacterial CBH, but the FN3s is not. Another interesting point is the order of domains. The fungal CBHs had CBM1 on the N terminus of CD. Although bacterial CBHs did not show

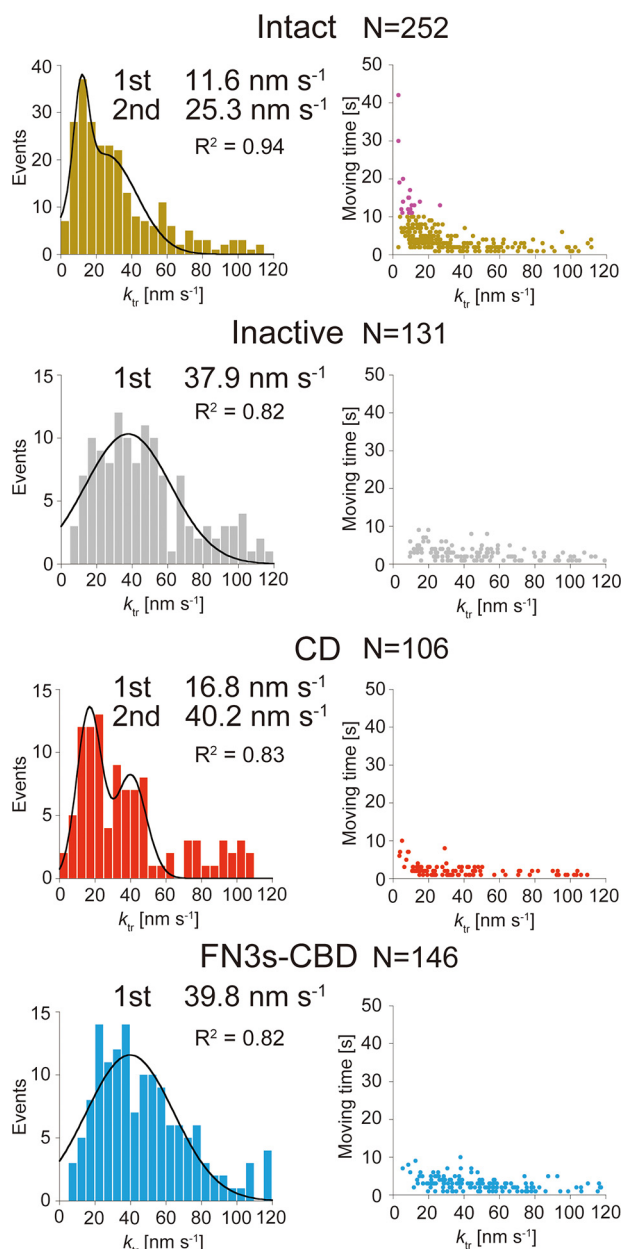


Figure 5. Translational rate (k_{tr}) and moving time analyses of Intact, Inactive, CD, and FN3s-CBD. Left panels, distributions of k_{tr} fitted by Gaussian functions. Peak values of distributions are shown. Right panels, plots of k_{tr} versus moving time. For Intact, molecules moved more than 10 s are shown in purple. The N values represent numbers of analyzed molecules.

clear order, further characterizations of GH6 CBHs from other bacteria is required to draw a conclusion.

Discussion

In this study, we found that Intact CfCel6B exhibits similar k_{cat} to TrCel6A (2.4 and 2.8 s^{-1} for CfCel6B and TrCel6A, respectively) (Fig. 2), although CfCel6B has largely different domain composition from that of TrCel6A (12). On the other hand, the K_m value for Intact CfCel6B (0.51 $mg\ ml^{-1}$) was much lower than that for TrCel6A (2.7 $mg\ ml^{-1}$), indicating that the affinity of CfCel6B to the crystalline cellulose is higher

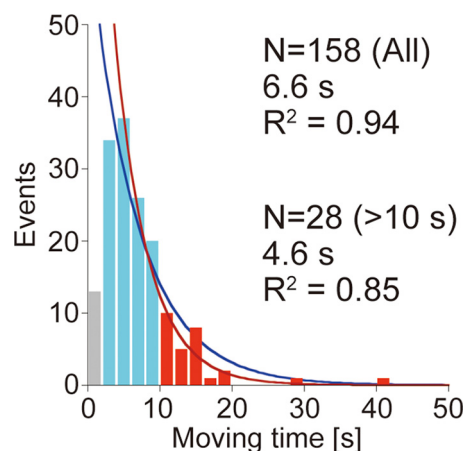


Figure 6. Distribution of moving time. Distribution of moving time for Intact was fitted by single exponential decay functions. Blue and red lines are fittings with the whole range and using only the range more than 10 s, respectively. The first bin was excluded from fittings. The N values represent numbers of analyzed molecules.

than that of TrCel6A. To understand the differences in the mechanisms of crystalline cellulose hydrolysis by these enzymes, here we quantitatively compare the kinetic parameters of elementary reaction steps such as binding (k_{on}), translational movement (k_{tr} and processivity), and dissociation (k_{off}) determined in the present and previous studies (12).

Our single-molecule fluorescence imaging enables direct estimations of the k_{on} and k_{off} of processive cellulases separately to understand which parameter mainly affects the affinity to the crystalline cellulose (Figs. 3 and 4). For both CfCel6B and TrCel6A, we found fast and slow components that correspond to the bindings on hydrophilic and hydrophobic crystal surfaces of the cellulose, respectively (Table 1) (12). Considering that cellulose hydrolysis will occur on hydrophobic, high-affinity crystal surface, the slow component is more relevant to the productive binding, although only a few binding events will lead to the hydrolysis. The values of k_{on}^{slow} for Intact CfCel6B and TrCel6A were comparable and $2.9 \times 10^8\ M^{-1}\ \mu M^{-1}\ s^{-1}$ (Fig. 3 and Table 1) and $2.3 \times 10^8\ M^{-1}\ \mu M^{-1}\ s^{-1}$ (12), respectively. The values of k_{off}^{slow} for Intact CfCel6B and TrCel6A were 0.086 and 0.10 s^{-1} , respectively (Fig. 4 and Table 1) (12). Thus, the values of k_{off}^{fast} or k_{off}^{slow} between Intact CfCel6B and TrCel6A are similar. However, the ratio of slow dissociation component for CfCel6B was 67% and much higher than that for TrCel6A (30%). The higher ratio of the slow dissociation component indicates that CfCel6B more specifically binds to the hydrophobic surface than TrCel6A. This difference is one of the reasons that CfCel6B showed a lower K_m value than TrCel6A. However, quantitative comparison is not easy, because direct measurement of the k_{on} for productive binding was difficult because of the low frame rate and localization precisions, similar to the case for our recent single-molecule fluorescence imaging of chitinase A from *Serratia marcescens* (21).

To understand role of the CD in the binding and dissociation of CfCel6B and TrCel6A on the cellulose surface, we also measured k_{on} and k_{off} for CfCel6B CD (Figs. 3 and 4) and solved its crystal structure (Fig. 7). Overall structure of CfCel6B CD was similar to that of the TrCel6A CD except for the additional

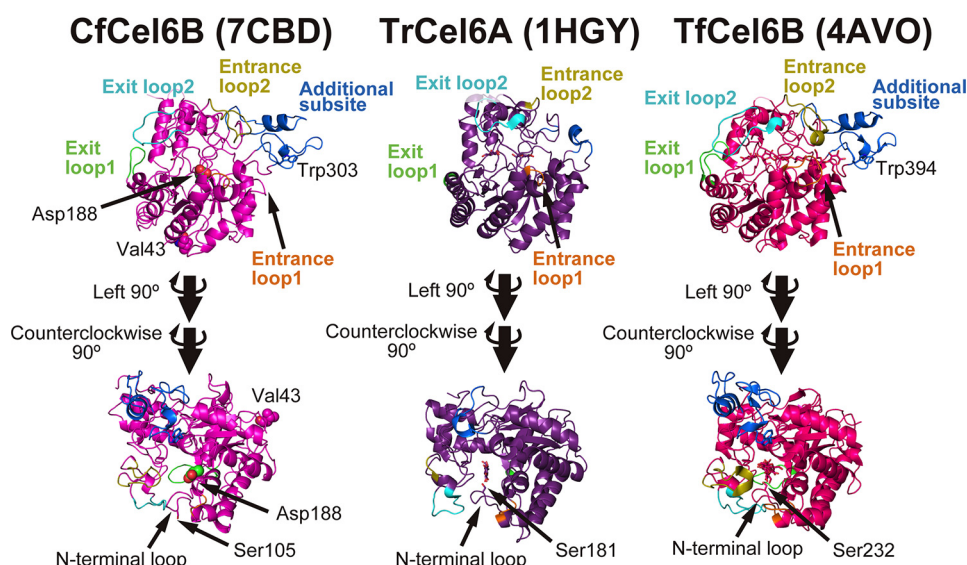


Figure 7. Structural comparison of bacterial and fungal GH6 CDs. *Left panel*, crystal structure of apo CfCel6B CD (PDB code 7CBD). Entrance loops 1 and 2 are shown in orange and yellow, and exit loops 1 and 2 are green and cyan, respectively. Loops constructing additional subsite are shown in blue. Val⁴³ and Asp¹⁸⁸ of CfCel6B, which are mutated to cysteine for fluorescent labeling and to alanine for inactivation, respectively, are shown as spheres. *Middle and right panels*, crystal structures of TrCel6A CD (PDB code 1HGY) and TrCel6B CD (PDB code 4AVO) shown from same directions and viewpoints. The active serine residues of three enzymes and the ligands in 1HGY and 4AVO are shown by stick model.

subsite and loops. Although Trp³⁰³ in additional subsite is exposed to the solvent, the values of k_{on} and k_{off} for CfCel6B CD were similar to those for TrCel6A CD measured in our previous study (12). In addition, the ratio of k_{off}^{slow} for CfCel6B CD (19%, Table 1) was not largely different from that for TrCel6A (28%). Therefore, the binding and dissociation are not largely affected by the additional subsite and loops of CfCel6B CD.

The most crucial domain for binding and dissociation of CfCel6B is CBD (Figs. 3 and 4). In the present study, the CBD showed 10 times higher k_{on} than CD, and the value was more than one-third of that for Intact (Table 1). These results clearly indicate that CBD has a critical role in the initial interaction of CfCel6B with crystalline cellulose. On the other hand, comparison of k_{off}^{slow} or k_{off}^{fast} between CD and CBD showed that binding of CD to hydrophobic or hydrophilic surface of crystalline cellulose is more stable than those of CBD (Fig. 4 and Table 1). However, the ratio of slow dissociation component for CD was 19% and much lower than that for CBD (70%), which was close to the value for Intact (67%). Therefore, stability of binding for Intact is a result of the cooperation between CD and CBD, but the specificity of binding to the hydrophobic surface arises from the binding by CBD. The CBD of CfCel6B is a member of type A CBM2, which is specific to cellulose and has a flat surface with aromatic residues (22). Binding of CBD in this group is expected to be driven by an increase of entropy (23), and this would be a reason for the high specificity of CfCel6B CBD to the hydrophobic surface of crystalline cellulose (Table 1). In case of TrCel6A CBD, the ratio of slow dissociation component was only 30% in our previous report (12). Because TrCel6A CBD belongs to CBM1, this result indicates that CBM2 is more specific to the hydrophobic surface than CBM1. Higher affinity of CBM2 was reported by Tomme *et al.* (24) using the CBM1 of the GH7 CBH from *T. reesei* and CBM2 of a GH10 xylanase from *C. fimi*. The preference of CBD binding to the hydrophobic surface has

been reported previously by Nimlos *et al.* (25). In their simulation, within few hundred nanoseconds, CBM1 moved to the hydrophobic surface from the hydrophilic surface. Similar events might also occur for CfCel6B CBD, although we could not resolve such a short time event in our single-molecule fluorescence imaging. As a conclusion, both CBDs of CfCel6B and TrCel6A have a role to lead the CD to the hydrophobic surface of crystalline cellulose, on which an accessible chain end exists.

Another large difference between TrCel6A and CfCel6B is the linker region. The contribution of the glycosylated linker of fungal cellulase to the binding on the cellulose surface has been reported previously (12, 14). On the other hand, in the present study, FN3s of CfCel6B do not contribute to the interaction with the cellulose surface, because FN3s-CBD showed similar k_{on} and k_{off} values with CBD (Table 1). Although FN3s are one of a common domains in bacterial glycoside hydrolases, such as polygalacturonosidase, chitinase, pullulanase, amylase, and cellulase (26), FN3s were not conserved in bacterial GH6 CBHs (Fig. 8). Recently, Valk *et al.* (27) proposed that the FN3s in bacterial GHs work as stable linkers connecting functional domains (CD and CBD) to keep the relative orientation and distance. Although our results support this idea, FN3-like domain of chitinase A from *S. marcescens* has a function as a binding domain on chitin (28). Therefore, we need to carefully compare the sequence, structure, and function of FN3s in other cases.

Other important parameters we need to compare are k_{tr} and processivity. The k_{tr} for Intact CfCel6B showed two slow and fast components with peak values of 11.6 and 25.3 nm s⁻¹, respectively (Fig. 5). We attribute the slow component of Intact to processive movement coupled with the hydrolysis of cellulose chain, because the value of k_{tr} is similar to the moving velocity (12.7 nm s⁻¹) observed by high-speed atomic force microscopy (29), and Inactive showed only fast component (37.9 nm s⁻¹). From this result, the hydrolytic activity of productively

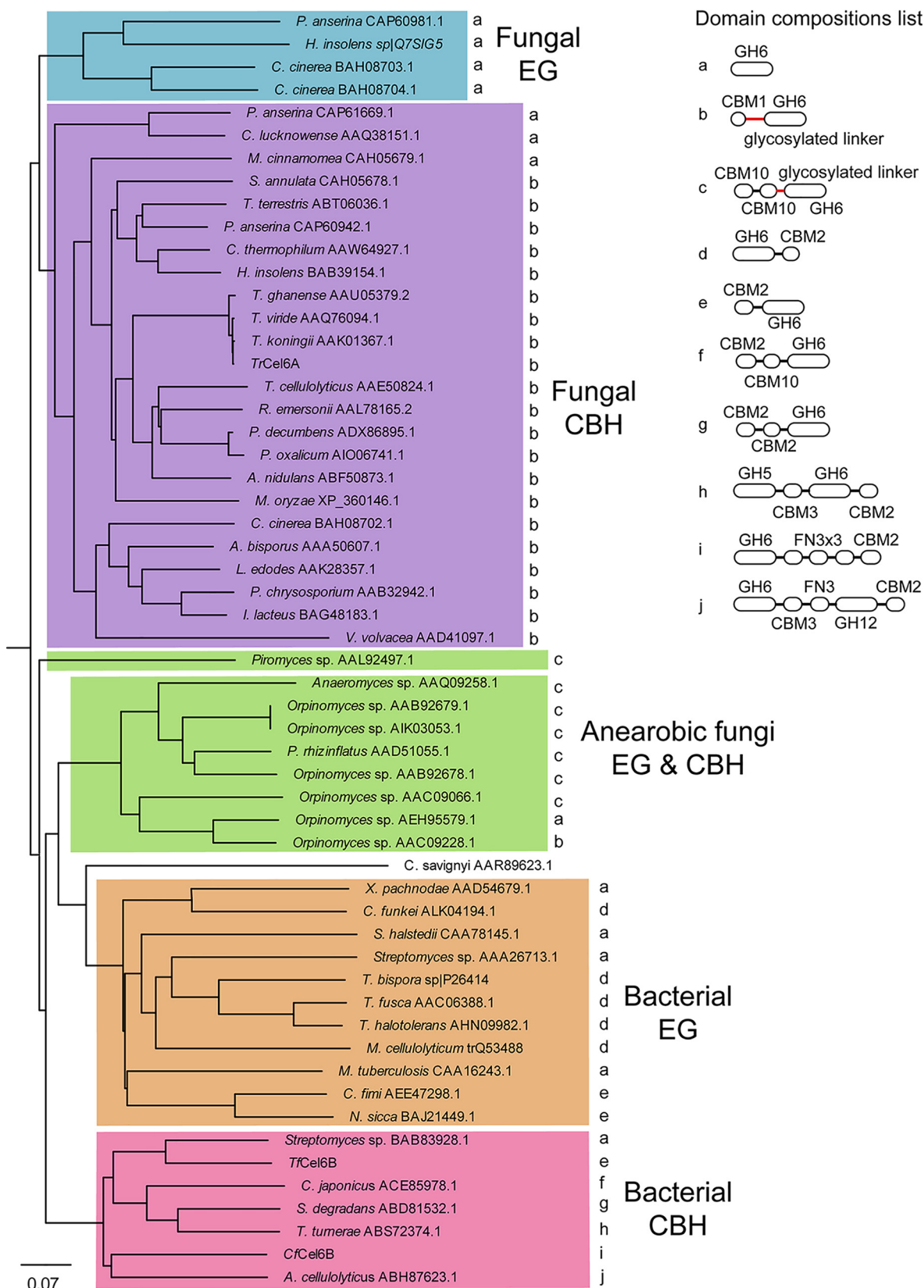


Figure 8. Phylogenetic tree of GH6 CD from bacteria and fungi and their domain compositions. GH6 enzymes in CAZy database were analyzed. GH6 regions were determined by homology modeling by SWISS-MODEL server and aligned depending on their structures using MODELLER. CBH and EG were classified according to the description in CAZy or original papers. Domain compositions were extracted from NCBI database or results of homology modeling. Glycosylated linkers are shown in red.

Single-molecule fluorescence imaging of bacterial cellulase

bound Intact CfCel6B molecule can be estimated to be 11.6 s^{-1} , because the length of product cellobiose is $\sim 1 \text{ nm}$. The large gap between the activities estimated from biochemical analysis (2.4 s^{-1} ; Fig. 2), and single-molecule analysis (11.6 s^{-1}) is presumably due to the low fraction of productive binding, as previously demonstrated for chitinase A from *S. marcescens* (30). From the values described above, only 20% of the Intact molecules is estimated to be productively bound even at high substrate concentration. This low ratio is caused by the limited numbers of the accessible chain ends on the surface of crystalline cellulose. Slow and fast movements in the translational movement on crystalline cellulose have been also observed in TrCel6A, and k_{tr} for slow and fast components were 8.8 and 34.9 nm s^{-1} , respectively (12). The fast movements correspond to surface diffusion without hydrolysis of cellulose chain as reported previously on the CBM2 using fluorescence recovery after the photobleaching method (31).

Furthermore, an interesting difference was found between CDs of CfCel6B and TrCel6A. The slow component for CD, which was not found in TrCel6A, has been observed in CfCel6B (Fig. 5). This result strongly suggests that CfCel6B CD has higher processivity than TrCel6A CD by the additional subsite and loops observed in the crystal structure (Fig. 7). Another possibility is that additional Trp increases efficiency to catch a cellulose chain end to form a productive complex. On the other hand, the value of processivity for CfCel6B Intact, estimated from k_{tr} (11.6 nm s^{-1} ; Fig. 5), moving time (4.6 s ; Fig. 6), and size of the product (cellobiose, 1.0 nm), was 53, which is smaller than 68, the value for TrCel6A Intact estimated from the same analysis (12). These results strongly suggest that contribution of glycosylated linker and CBD of TrCel6A to the processivity is larger than that of FN3s-CBD of CfCel6B. Our results also indicate that not only the structure of CD but also the linker and CBD are important for the processivity.

Domain compositions of GH6 cellulases clearly showed different tendencies between fungi and bacteria. When the phylogenetic tree of GH6 CD was prepared, groups of fungal EG and CBH, bacterial EG and CBH, and anaerobic fungal cellulase were separated as expected (Fig. 8). However, this phylogenetic tree also indicates an interesting relationship between the domain compositions and the function of CD. For instance, fungal CBH basically has glycosylated linker and CBM1-CBD to achieve high-affinity binding and high processivity for efficient degradation of crystalline cellulose. In contrast, bacterial CBHs have large variety in combinations of domains and length of linker regions. The common properties among them are CBM2-CBD with high affinity and CD with a long substrate-binding tunnel (and high processivity, presumably). These two domains, important for crystalline cellulose degradation, were highly conserved except for the CBH from *Streptomyces* sp. Therefore, bacterial CBHs seem to compensate for weak interaction of nonglycosylated linker region with crystalline cellulose surface by strong binding of CBM2-CBD and high processivity of CD. As a result, bacterial and fungal GH6 CBHs have a similar function: processive hydrolysis of the crystalline cellulose from the nonreducing end. The phylogenetic tree analysis suggests that the bacterial and fungal CBHs have been generated by a divergent evolution from an ancestor, but converged

functionally. The glycosylated linker and CBM1-CBD of fungal GH6 CBHs are not common among the bacterial GH6 CBHs, and the CBM2-CBD of bacterial CBHs is rarely observed in eukaryote either, although the catalytic domain has the same fold. Therefore, the two groups may have evolved independently, and both obtain the binding function using the unique ways for fungi or bacteria to degrade the crystalline cellulose.

Recently, a cellulase containing a GH6 CBH from *Reconciliobacillus cellulovorans* (RcCelC) has been reported (32). Domain composition of this enzyme is GH6-GH5-CBM3, and GH6 CD is similar to the bacterial one. Although RcCelC does not have CBM2, it contains Ser-Pro-Thr-rich linkers between GH6 and GH5, and GH5 and CBM3, respectively. In their study, glycosylation of RcCelC has been confirmed by Periodic acid-Schiff staining. Recently, glycosylation of bacterial proteins also has been considered to be common, especially in pathogenic bacteria (33). The cellulases with glycosylated linker and bacterial GH6 CBH are one of the next interesting targets to analyze the elementary steps of the reaction and processivity by using the single-molecule imaging analysis.

Experimental procedures

Mutant design and gene preparation

The CfCel6B gene without signal peptide (Ala⁴¹ to Gly⁸⁷² of CfCel6B WT) was amplified with primers including NcoI (CCATGG) and HindIII (AAGCTT) recognition sites for forward and reverse primers, respectively, and ligated to pET27b vector. As a result, one methionine residue was added in the N terminus of enzyme. The gene of CfCel6B CD (Met plus Ala⁴¹–Thr⁴⁸³ of CfCel6B WT) for crystallization was amplified with primers containing NcoI and HindIII recognition sites. Six histidine residues and the stop codon were connected to the codon corresponding to Thr⁴⁸³ and ligated with pET27b after digestion by NcoI and HindIII. The genes of CfCel6B V43C was amplified by same forward primer additionally containing V43C mutation (GTC to TGC). The reverse primer was prepared at RsrII recognition site (1514 bp downstream from Ala⁴¹). Part of CfCel6B gene was swapped with amplified fragment. For CD-V43C, the same forward primer and reverse primers including the codons for FaXa protease recognition site (IEGRFVG: ATCGAAGGCCGCTTTGGCGGC) between the codons of Thr⁴⁸³ and the His₆ tag were used. CfCel6B V43C/D188A gene was prepared by additional a pair of primers, which has an 18-bp overlap region, for the D188A mutation. The two fragments of DNA from Ala⁴¹ to Ala¹⁸⁸ and Ala¹⁸⁸ to the RsrII site (near Val⁵⁴⁶) were mixed and amplified with the same primers for CfCel6B V43C. The fragment was ligated with the plasmid of CfCel6B V43C treated by NcoI and RsrII. The gene of FN3-CBD (Val⁴⁸⁴–Gly⁸⁷²) V484C was amplified by primers with NcoI and HindIII sites. The start codon, His₆ tag, and FaXa recognition site were added to the N terminus of the protein. The gene was swapped with CfCel6B. CBD (Thr⁷⁶⁶–Gly⁸⁷²) T766C was also amplified with similar sets of primers. The gene was also ligated to pET27B. Primester GXL polymerase (Takara) was used for all of PCRs. Restriction enzymes were purchased from NEB. PCR products were purified by agarose gel electrophoresis and Wizard® SV Gel and PCR clean-up system

(Promega) according to the manufacturer's instruction. Ligations of DNA fragments were achieved by Mighty Mix kit (Takara). All of ligated plasmids were transformed into Tuner (DE3) (Merck Millipore) by electroporation using MicroPulser (Bio-Rad) according to the setting for *E. coli* transformation. 50 μl of transformed competent cell was mixed with 200 μl of SOC medium and incubated at 37 °C for 1 h. All of suspension were spread on agarose LB plate containing 25 $\mu\text{g ml}^{-1}$ kanamycin and incubated for a night at 37 °C. Three or four colonies were cultivated in 10 ml of LB medium containing 25 $\mu\text{g ml}^{-1}$ kanamycin at 37 °C and 300 rpm for 16 h. The plasmids were purified from the cells by FastGene plasmid mini kit (NIPPON Genetics). Whole sequences of genes were verified, and the plasmids were stored in $-30\text{ }^{\circ}\text{C}$.

Expression, purification, and Cy3 labeling of protein samples

Plasmids were transformed into *E. coli* Tuner (DE3) by electroporation in the same way above. Single colonies of one-fourth of the plate were inoculated in 10 ml of LB medium with 25 $\mu\text{g ml}^{-1}$ kanamycin and incubated at 37 °C and 200 rpm for 1 h. 3 ml of preculture medium was added in 50 ml of Overnight Express instant LB medium (Novagen) containing 25 $\mu\text{g ml}^{-1}$ of kanamycin and incubated at 25 °C and 130 rpm for a night. The cell was harvested by centrifuge at $3000 \times g$ for 10 min at 4 °C. Harvested cell was stored in $-80\text{ }^{\circ}\text{C}$ until purification.

For the purification of CfCel6B WT, Intact, and inactive, ~ 7 g of cell was suspended in 70 ml of 100 mM Tris-HCl, pH 8.0, containing 100 mM sodium chloride. The cells were disrupted by sonication for 15 min on ice. Suspension was mixed with 30 μl of Benzonase (Merck Millipore) and 710 μl of 2 M magnesium chloride, and precipitant was removed by centrifuge at $8,000 \times g$ for 10 min and $30,000 \times g$ for 10 min sequentially at 37 °C. Two times the amount of 3 M ammonium sulfate was added to supernatant and centrifuged at $8,000 \times g$ 10 min. Supernatant was loaded on cellulose column equilibrated with 1 M ammonium sulfate (34). Unbound protein was washed out by 1 M ammonium sulfate, and bound protein was eluted by milliQ water. Purity of proteins were analyzed by SDS-PAGE, and the fractions containing the ~ 80 -kDa protein were collected and concentrated by 30-kDa cut Vivaspin 20 column at $6000 \times g$. Buffer was changed to 20 mM Tris-HCl, pH 7.5, by Econo-Pack 10DC column (Bio-Rad). The enzyme was loaded on the Toyoperl DEAE 650-S (Tosho) and eluted by the linear gradient of sodium chloride from 0 to 300 mM. Target proteins were concentrated by Vivaspin. CfCel6B WT was further loaded into YMC-Pack Diol-200G (YMC) and eluted by 20 mM sodium phosphate, pH 7.0, with 100 mM sodium chloride. Other free cysteine mutants were reduced by 10 mM DTT for 2 h at 25 °C before being put into a size-exclusion column. Reduced protein was loaded to YMC-Pack Diol-200G and eluted by 20 mM sodium phosphate, pH 7.0, with 100 mM sodium chloride. The protein in the fraction showed peak absorption at 280 nm chromatogram was reacted with five times higher moles of Cy3-maleimide for a night at room temperature. Unreacted Cy3 was removed by Econo-Pack 10DC column with 20 mM sodium phosphate, pH 7.0, with 100 mM sodium chloride. Labeled enzyme was concentrated, and labeling ratio was calculated. Molecular extinction coefficient of $131,650\text{ M}^{-1}\text{ cm}^{-1}$ at 280 nm was

used for CfCel6B, and those of $12,000\text{ M}^{-1}\text{ cm}^{-1}$ at 280 nm and $150,000\text{ M}^{-1}\text{ cm}^{-1}$ at 550 nm were used for Cy3. Purified protein was kept at $-80\text{ }^{\circ}\text{C}$ after flash freezing with liquid nitrogen.

E. coli cells expressing CfCel6B CD, FN3-CBD, and CBD were disrupted and centrifuged by the same method without Benzonase and magnesium chloride. Supernatant was loaded on nickel-nitrilotriacetic acid-agarose column (Qiagen) and washed by 50 mM sodium phosphate, pH 7.0, with 100 mM sodium chloride. The column was washed by the buffer containing 20 mM imidazole, and target proteins were eluted by 50 and 100 mM imidazole-containing buffer. Collected proteins were concentrated to 200 μl by ultracentrifuge, and 4 μl of 100 mM calcium chloride and 20 μl of 1 mg ml^{-1} FaXa protease (NEB) were added to sample. After incubation at 23 °C for a night, and target protein was reduced by 10 mM DTT at 25 °C for 1 h. Reduced protein was injected to HPLC equipped with YMC-Pack Diol-200G. Target protein was eluted by 50 mM sodium phosphate, pH 7.0, with 100 mM sodium chloride. For labeling, the procedures were same to those of CfCel6B Intact and inactive. Molecular coefficient of CD was $\epsilon_{280} = 71,930\text{ M}^{-1}\text{ cm}^{-1}$, those of FN3-CBD and CBD were $\epsilon_{280} = 59,720\text{ M}^{-1}\text{ cm}^{-1}$ and $\epsilon_{280} = 29,850\text{ M}^{-1}\text{ cm}^{-1}$. CfCel6B CD for crystallization was purified without FaXa treatment and reduction. Molecular coefficient of CD for crystallization was $\epsilon_{280} = 1.49\text{ mg}^{-1}\text{ ml cm}^{-1}$. TrCel6A Intact was the same sample used in the previous report (12).

Activity measurement of enzymes

Purified 0.5 μM of CfCel6B WT, Intact, and inactive were reacted with 1 mg ml^{-1} of crystalline cellulose I_{α} in 100 μl of 100 mM sodium acetate buffer, pH 5.0, at 30 °C. After 1 h of incubation, 100 μl of sodium hydroxide was added to stop the reaction. Suspension was centrifuged at $15,000 \times g$ for 10 min, and 120 μl of supernatant was mixed with same volume of 4-hydroxybenzoic acid hydrazide (PAHBAH) solution (35). The mixture was heated at 95 °C for 10 min, and absorbance at 405 nm of 220 μl of solution was taken by 96-well plate reader. Standard curve was prepared using glucose, and activity was calculated as a normalized value by time and enzyme concentration.

To determine the turnover and affinity of CfCel6B WT and Intact to crystalline cellulose I_{α} , 0.1 μM of CfCel6B WT and Intact were incubated with 0.3, 1.0, 3.0, and 5.0 mg ml^{-1} of crystalline cellulose in 100 mM sodium acetate buffer, pH 5.0, at 25 °C for 2 min. Supernatant as collected after centrifuge and concentration of products were analyzed by HPLC (11).

Observation of binding, dissociation, and translational movement

Cover glass (24 mm \times 32 mm, Matsunami Glass) was incubated in 10 M potassium hydroxide for a night and washed by milliQ water. For the single-molecule observation, we used annular illumination single-molecule fluorescence microscopy with EM-CCD camera (Andor) (12). 20 μl of 0.2 mg ml^{-1} crystalline cellulose I suspension was coated on the glass at 3000 rpm by spin coater (Mikasa) as described previously (12). For analyses of binding rate and dissociation rate constants, 20 μl of 50 pM Intact, FN3-CBD, 100 pM CBD, or 600 pM CD was dropped on the glass. The power density of a 532-nm laser was

Single-molecule fluorescence imaging of bacterial cellulase

set at $0.14 \mu\text{W} \mu\text{m}^{-2}$, and the frame rate was 5 fps. The rate constant of photobleaching for Cy3 conjugated with CfCel6B was $0.053 \pm 0.002 \text{ s}^{-1}$ (or time constant of $18.9 \pm 0.7 \text{ s}$) under this observation condition. Crystalline cellulose microfibrils were stained by $10 \mu\text{l}$ of 10 nM TrCel6A-S386C-Cy3, and the movie was overlaid with the image to analyze the molecules bound on the cellulose. Lengths of the crystalline cellulose microfibrils were measured by using ImageJ as previously described (36). Binding rates were corrected by the labeling ratio. For the analysis of translational movement and k_{tr} , the laser power density of $0.28 \mu\text{W} \mu\text{m}^{-2}$ and frame rate of 1 fps were used to improve the localization precision. Under this observation condition, localization precisions in the x and y axes were 4.5 ± 1.5 and $4.6 \pm 1.4 \text{ nm}$, respectively. Time constant of photobleaching for Cy3 conjugated with CfCel6B was $14.9 \pm 0.6 \text{ s}$.

X-ray crystal structure analysis of CfCel6B CD

$1 \mu\text{l}$ of 10 mg ml^{-1} purified CfCel6B CD was mixed with $1 \mu\text{l}$ of 21% PEG3350 and 10 mM sodium chloride in 100 mM sodium acetate buffer, pH 5.0, on the sitting-drop plate (griner). The drop was equilibrated with $100 \mu\text{l}$ of 21% PEG3350 and 10 mM sodium chloride in 100 mM sodium acetate buffer, pH 5.0, for a week at 20°C . Formed rod-like crystal was soaked in 40% PEG3350, 10 mM sodium chloride, and 100 mM sodium acetate, pH 5.0. Diffraction of 1.0 \AA X-ray was measured from 0 to 360° with 0.5° oscillation. Diffraction spots were observed up to 1.3 \AA resolution, the diffraction images were processed by HKL2000, and the phase was determined by Phaser in Phenix suite (37). The template for molecular replacement was prepared by the SWISS-MODEL server using TfCel6B (PDB code 4AVO) as a template for modeling (38). Structural refinement and model editing was done by Phenix refine and Coot (39). CDs of CfCel6B, TfCel6B, and TrCel6A were compared and visualized by PyMOL.

Comparison of GH6 enzymes from bacteria and fungi

Amino acid sequences of GH6 cellulase classified as characterized enzymes in CAZy database were downloaded. Some enzymes containing X residues or reported only as patent were rejected. Homology model structures of GH6 CD were prepared by SWISS-MODEL server with default settings (38). GH6 CDs were aligned by multiple structural alignment function in MODELLER (40). Phylogenetic tree was calculated by Clustal Omega server using the neighbor-joining method (41). The tree file was visualized by Figtree. Domain configurations were determined according to NCBI, Uniprot database, and results of modeling.

Data availability

The structure presented in this article has been deposited into the Protein Data Bank (PDB) with the following ID: 7CBD. All remaining data are contained within the article.

Acknowledgments—We thank all beamline staff members at the Photon Factory for their kind help with data collection.

Author contributions—A. N. and R. I. conceptualization; A. N., D. I., A. V., T. U., K. M., and R. I. formal analysis; A. N., T. M., K. I.,

and R. I. funding acquisition; A. N. writing-original draft; S. K. resources; T. M. methodology; K. I. and R. I. supervision.

Funding and additional information—This study was supported by the Leading Initiative for Excellent Young Researchers program (to A. N.) and Grant-in-Aid for Scientific Research on Innovative Areas “Molecular Engine” grants JP18H05424 (to R. I.) and JP18H05425 (to T. M.), Grants-in-Aid for Scientific Research Grants 15H06898 and 19H03094 (to A. N.), 18H05494 (to K. I.), and 18H02418 and 18H04755 (to R. I.) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. K.I. was supported by Business Finland (formerly the Finnish Funding Agency for Innovation (TEKES)) via the Finland Distinguished Professor Program Advanced Approaches for Enzymatic Bio-mass Utilization and Modification.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: CBH, cellobiohydrolase; CBD, cellulose-binding domain; CD, catalytic domain; CBM, carbohydrate-binding module family; GH, glycoside hydrolase family; FaXa, factor Xa protease; CfCel6B, glycoside hydrolase family 6 cellobiohydrolase from *C. fimi*; TrCel6A, glycoside hydrolase family 6 cellobiohydrolase from *T. reesei*; TfCel6B, glycoside hydrolase family 6 cellobiohydrolase from *T. fusca*; PDB, Protein Data Bank; EG, endoglucanase; FN3, fibronectin type 3 domain; Intact, full-length CfCel6B/V43C; Inactive, catalytically inactive D188A mutant of full-length CfCel6B/V43C; fps, frames/s.

References

1. Teeri, T. T. (1997) Crystalline cellulose degradation: new insight into the function of cellobiohydrolases. *Trends Biotechnol.* **15**, 160–167 [CrossRef](#)
2. Davies, G., and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853–859 [CrossRef Medline](#)
3. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490–D495 [CrossRef Medline](#)
4. Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* **382**, 769–781 [CrossRef Medline](#)
5. Henrissat, B., Driguez, H., Viet, C., and Schülein, M. (1985) Synergism of cellulases from *Trichoderma reesei* in the degradation of cellulose. *Bio/Technology* **3**, 722–726 [CrossRef](#)
6. Demain, A. L., Newcomb, M., and Wu, J. H. (2005) Cellulase, clostridia, and ethanol. *Microbiol. Mol. Biol. Rev.* **69**, 124–154 [CrossRef Medline](#)
7. Spertino, S., Boatti, L., Icardi, S., Manfredi, M., Cattaneo, C., Marengo, E., and Cavaletto, M. (2018) *Cellulomonas fimi* secretomes: *in vivo* and *in silico* approaches for the lignocellulose bioconversion. *J. Biotechnol.* **270**, 21–29 [CrossRef Medline](#)
8. Igarashi, K., Uchihashi, T., Koivula, A., Wada, M., Kimura, S., Okamoto, T., Penttilä, M., Ando, T., and Samejima, M. (2011) Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface. *Science* **333**, 1279–1282 [CrossRef Medline](#)
9. Igarashi, K., Koivula, A., Wada, M., Kimura, S., Penttilä, M., and Samejima, M. (2009) High speed atomic force microscopy visualizes processive movement of *Trichoderma reesei* cellobiohydrolase I on crystalline cellulose. *J. Biol. Chem.* **284**, 36186–36190 [CrossRef Medline](#)
10. Payne, C. M., Jiang, W., Shirts, M. R., Himmel, M. E., Crowley, M. F., and Beckham, G. T. (2013) Glycoside hydrolase processivity is directly related

- to oligosaccharide binding free energy. *J. Am. Chem. Soc.* **135**, 18831–18839 [CrossRef Medline](#)
11. Nakamura, A., Watanabe, H., Ishida, T., Uchihashi, T., Wada, M., Ando, T., Igarashi, K., and Samejima, M. (2014) Trade-off between processivity and hydrolytic velocity of cellobiohydrolases at the surface of crystalline cellulose. *J. Am. Chem. Soc.* **136**, 4584–4592 [CrossRef Medline](#)
 12. Nakamura, A., Tasaki, T., Ishiwata, D., Yamamoto, M., Okuni, Y., Visootsat, A., Maximilien, M., Noji, H., Uchiyama, T., Samejima, M., Igarashi, K., and Iino, R. (2016) Single-molecule imaging analysis of binding, processive movement, and dissociation of cellobiohydrolase *Trichoderma reesei* Cel6A and its domains on crystalline cellulose. *J. Biol. Chem.* **291**, 22404–22413 [CrossRef Medline](#)
 13. Sandgren, M., Wu, M., Karkehabadi, S., Mitchinson, C., Kelemen, B. R., Larenas, E. A., Ståhlberg, J., and Hansson, H. (2013) The structure of a bacterial cellobiohydrolase: the catalytic core of the *Thermobifida fusca* family GH6 cellobiohydrolase Cel6B. *J. Mol. Biol.* **425**, 622–635 [CrossRef Medline](#)
 14. Payne, C. M., Resch, M. G., Chen, L., Crowley, M. F., Himmel, M. E., Taylor, L. E., 2nd, Sandgren, M., Ståhlberg, J., Stals, I., Tan, Z., and Beckham, G. T. (2013) Glycosylated linkers in multimodular lignocellulose-degrading enzymes dynamically bind to cellulose. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 14646–14651 [CrossRef Medline](#)
 15. Gilkes, N. R., Kilburn, D. G., Miller, R. C., Warren, R. A. J., Sugiyama, J., Chanzy, H., and Henrissat, B. (1993) 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.* **15**, 347–351 [CrossRef Medline](#)
 16. Lehtiö, J., Sugiyama, J., Gustavsson, M., Fransson, L., Linder, M., and Teeri, T. T. (2003) The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 484–489 [CrossRef Medline](#)
 17. Meinke, A., Gilkes, N. R., Kwan, E., Kilburn, D. G., Warern, R. A., and Miller, R. C., Jr. (1994) Cellobiohydrolase A (CbhA) from the cellulolytic bacterium *Cellulomonas fimi* is a β -1,4-exocellobiohydrolase analogous to *Trichoderma reesei* CBH II. *Mol. Microbiol.* **12**, 413–422 [CrossRef Medline](#)
 18. Burgin, T., Ståhlberg, J., and Mayes, H. B. (2018) Advantages of a distant cellulase catalytic base. *J. Biol. Chem.* **293**, 4680–4687 [CrossRef Medline](#)
 19. Tachioka, M., Nakamura, A., Ishida, T., Igarashi, K., and Samejima, M. (2017) Crystal structure of a family 6 cellobiohydrolase from the basidiomycete *Phanerochaete chrysosporium*. *Acta Crystallogr. F Struct. Biol. Commun.* **73**, 398–403 [CrossRef Medline](#)
 20. Gill, J., Rixon, J. E., Bolam, D. N., McQueen-Mason, S., Simpson, P. J., Williamson, M. P., Hazlewood, G. P., and Gilbert, H. J. (1999) The type II and X cellulose-binding domains of *Pseudomonas xylanase* A potentiate catalytic activity against complex substrates by a common mechanism. *Biochem. J.* **342**, 473–480 [CrossRef Medline](#)
 21. Visootsat, A., Nakamura, A., Vignon, P., Watanabe, H., Uchihashi, T., and Iino, R. (2020) Single-molecule imaging analysis reveals the mechanism of a high-catalytic-activity mutant of chitinase A from *Serratia marcescens*. *J. Biol. Chem.* **295**, 1915–1925 [CrossRef Medline](#)
 22. Xu, G.-Y., Ong, E., Gilkes, N. R., Kilburn, D. G., Muhandiram, D. R., Harris-Brandts, M., Carver, J. P., Kay, L. E., and Harvey, T. S. (1995) Solution structure of a cellulose-binding domain from *Cellulomonas fimi* by nuclear magnetic resonance spectroscopy. *Biochemistry* **34**, 6993–7009 [CrossRef Medline](#)
 23. Creagh, A. L., Ong, E., Jarvis, E., Kilburn, D. G., and Haynes, C. A. (1996) Binding of the cellulose-binding domain of exoglucanase Cex from *Cellulomonas fimi* to insoluble microcrystalline cellulose is entropically driven. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12229–12234 [CrossRef Medline](#)
 24. Tomme, P., Driver, D. P., Amandoron, E. A., Miller, R. C., Jr., Antony, R., Warren, J., and Kilburn, D. G. (1995) Comparison of a fungal (family I) and bacterial (family II) cellulose-binding domain. *J. Bacteriol.* **177**, 4356–4363 [CrossRef Medline](#)
 25. Nimlos, M. R., Beckham, G. T., Matthews, J. F., Bu, L., Himmel, M. E., and Crowley, M. F. (2012) Binding preferences, surface attachment, diffusivity, and orientation of a family 1 carbohydrate-binding module on cellulose. *J. Biol. Chem.* **287**, 20603–20612 [CrossRef Medline](#)
 26. Little, E., Bork, P., and Doolittle, R. F. (1994) Tracing the spread of fibronectin type III domains in bacterial glycohydrolases. *J. Mol. Evol.* **39**, 631–643 [CrossRef Medline](#)
 27. Valk, V., Kaaij, R. M. V. D., and Dijkhuizen, L. (2017) The evolutionary origin and possible functional roles of FNIII domains in two *Microbacterium aurum* B8A granular starch degrading enzymes, and in other carbohydrate acting enzymes. *Amylase I*, 1–11 [CrossRef](#)
 28. Oyeleye, A., and Normi, Y. M. (2018) Chitinase: diversity, limitations, and trends in engineering for suitable applications. *Biosci. Rep.* **38**, BSR2018032300 [CrossRef](#)
 29. Uchiyama, T., Uchihashi, T., Nakamura, A., Watanabe, H., Kaneko, S., Samejima, M., and Igarashi, K. (2020) Convergent evolution of processivity in bacterial and fungal cellulases. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 19896–19903 [CrossRef Medline](#)
 30. Nakamura, A., Tasaki, T., Okuni, Y., Song, C., Murata, K., Kozai, T., Hara, M., Sugimoto, H., Suzuki, K., Watanabe, T., Uchihashi, T., Noji, H., and Iino, R. (2018) Rate constants, processivity, and productive binding ratio of chitinase A revealed by single-molecule analysis. *Phys. Chem. Chem. Phys.* **20**, 3010–3018 [CrossRef Medline](#)
 31. Jarvis, E. J., Haynes, C. A., and Kilburn, D. G. (1997) Surface diffusion of cellulases and their isolated binding domains on cellulose. *J. Biol. Chem.* **272**, 24016–24023 [CrossRef Medline](#)
 32. Kolinko, S., Wu, Y. W., Tachea, F., Denzel, E., Hiras, J., Gabriel, R., Bäcker, N., Chan, L. J. G., Eichorst, S. A., Frey, D., Chen, Q., Azadi, P., Adams, P. D., Pray, T. R., Tanjore, D., et al. (2018) A bacterial pioneer produces cellulase complexes that persist through community succession. *Nat. Microbiol.* **3**, 99–107 [CrossRef Medline](#)
 33. Lu, Q., Li, S., and Shao, F. (2015) Sweet talk: protein glycosylation in bacterial interaction with the host. *Trends Microbiol.* **23**, 630–641 [CrossRef Medline](#)
 34. Sugimoto, N., Igarashi, K., and Samejima, M. (2012) Cellulose affinity purification of fusion proteins tagged with fungal family 1 cellulose-binding domain. *Protein Exp. Purif.* **82**, 290–296 [CrossRef Medline](#)
 35. Lever, M. (1977) Carbohydrate determination with 4-hydroxybenzoic acid hydrazide (PAHBAH): effect of bismuth on the reaction. *Anal. Biochem.* **81**, 21–27 [CrossRef Medline](#)
 36. Shibafuji, Y., Nakamura, A., Uchihashi, T., Sugimoto, N., Fukuda, S., Watanabe, H., Samejima, M., Ando, T., Noji, H., Koivula, A., Igarashi, K., and Iino, R. (2014) Single-molecule imaging analysis of elementary reaction steps of *Trichoderma reesei* cellobiohydrolase I (Cel7A) hydrolyzing crystalline cellulose I_a and III₁. *J. Biol. Chem.* **289**, 14056–14065 [CrossRef Medline](#)
 37. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., et al. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 [CrossRef Medline](#)
 38. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 [CrossRef Medline](#)
 39. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 [CrossRef Medline](#)
 40. Webb, B., and Sali, A. (2016) Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics* **54**, 5.6.1–5.6.37 [CrossRef Medline](#)
 41. Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D., and Lopez, R. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 [CrossRef Medline](#)