Tryptophan residue at active-site tunnel entrance of *Trichoderma reesei* cellobiohydrolase Cel7A is important to initiate degradation of crystalline cellulose*

Akihiko Nakamura†, Takeshi Tsukada, Sanna Auer, Tadaomi Furuta, Masahisa Wada, Anu Koivula, Kiyohiko Igarashi, and Masahiro Samejima

1Department of Biomaterials Sciences, Graduate School of Agricultural and Life Sciences, University of Tokyo, Tokyo 113-8657, Japan

2VTT Technical Research Centre of Finland, Espoo, FI-02044 VTT, Finland

3Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama 226-8501, Japan

4Department of Plant and Environmental New Resources, College of Life Sciences, Kyung Hee University, 1, Seocheon-dong, Giheung-ku, Yongin-si, Gyeonggi-do 446-701, Republic of Korea.

*Running title: TrCel7A W40A mutagenesis

†These authors contributed equally to this work.

‡Correspondence to either of these authors:
Masahiro Samejima, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan; Fax: 81-3-5841-5273; E-mail: amsam@mail.ecc.u-tokyo.ac.jp.
Anu Koivula, Espoo, 02044 VTT, Finland; E-mail: anu.koivula@vtt.fi

**Keywords:** cellobiohydrolase; crystalline cellulose degradation; *Trichoderma reesei*

Background: Mutation of W40 residue in a cellobiohydrolase TrCel7A causes a loss of the crystalline cellulose-degrading ability.

Results: W40A mutant showed reduced specific activity towards crystalline cellulose and diffused cellulose chain from entrance of active-site tunnel.

Conclusions: Trp 40 is essential for chain-end loading to initiate processive hydrolysis of TrCel7A.

Significance: The mechanisms of crystalline polysaccharides degradation are clarified.

SUMMARY

Glycoside hydrolase family-7 cellobiohydrolase from *Trichoderma reesei* (TrCel7A) is one of the best-studied cellulases with the ability to degrade highly crystalline cellulose (HCC). The catalytic domain (cat) and cellulose-binding domain (CBD) are both necessary for full activity on crystalline substrates. Our previous high-speed AFM studies showed that mutation of tryptophan residue (Trp40; W40) at the entrance of the catalytic tunnel drastically decreases the ability to degrade crystalline cellulose. Here, we examined the activities of the wild-type (WT) enzyme and tryptophan to alanine mutant (W40A) enzyme, with and without the CBD, towards various substrates. Evaluation and comparison of the specific activities of the enzymes (WT, WT<sub>cat</sub>, W40A, and W40A<sub>cat</sub>) adsorbed on crystalline cellulose indicated that Trp40 is involved in recruiting individual substrate chains into the active site tunnel to initiate processive hydrolysis. This idea was supported by molecular dynamics simulation study, i.e., the reducing-end glucose unit was effectively loaded into the active site of WT<sub>cat</sub>, but not into that of W40A<sub>cat</sub> when the simulation was started from subsite -7. However, when the similar simulations were carried out starting from subsite -5, both enzymes hold the substrate in 50 ns, indicating that the major difference between WT<sub>cat</sub> and W40A<sub>cat</sub> is the length of free chain-end of the substrate required to allow initiation of processive movements; this also reflects the difference between crystalline and
amorphous celluloses. The CBD is important for enhancing the enzyme population on crystalline substrate, but it also decreases the specific activity of the adsorbed enzyme, possibly by attaching the enzyme to non-optimal places on cellulose surface and/or hindering processive hydrolysis.

Cellulose is a linear polymer of β-1,4 linked D-glucose and is the main component of plant cell wall. The major cellulase produced by the cellulolytic ascomycete Trichoderma reesei is the glycoside hydrolase (GH) family-7 celllobiohydrolase (CBH), TrCel7A. TrCel7A has a two-domain structure composed of a catalytic domain (cat) and CBM1 family cellulose-binding domain (CBD) connected by a glycosylated linker peptide (1). The three-dimensional X-ray structure of the TrCel7A cat with cello-oligosaccharides has been determined, revealing a long active-site tunnel formed by loops that extend from the central β-sandwich fold and are stabilized by 10 disulfide bridges (2). The tunnel contains nine glucosyl binding sites, -7 to +2, and the catalytic site (-1/+1) is located at the far end, leading to production of cellobiose from the reducing end of the cellulose chain. The tunnel is aligned by hydrogen bond forming residues and in addition, four tryptophan residues provide hydrophobic stacking interactions for the glucosyl units at the tunnel entrance (Trp40), center of the tunnel (Trp38), and around the catalytic sites (Trp367 and Trp376), as shown in Figure 1. It is believed that the tight holding of the substrate and the asymmetry of the active site tunnel favor effective and continuous hydrolysis of crystalline cellulose in a processive mode (2,3).

The processivity of CBHs has been estimated by means of measurement of the ratio of released cellobiose to the sum of released glucose and cellotriose (4-6), product analysis from a mixture of cellulose and p-nitrophenyl lactoside (7) or the use of labeled cellulose substrates (8,9). Recently, we have demonstrated sliding movement of single TrCel7A molecules on a highly crystalline cellulose surface by means of real-time monitoring with a high-speed atomic force microscope (HS-AFM) (10-12). We showed that an inactive mutant did not slide on a cellulose surface, and also that the CBD was unnecessary for the sliding movement. These studies showed that the sliding movement reflected processive hydrolysis by the TrCel7A cat. Moreover, the AFM results demonstrated that TrCel7A W40A mutant could not slide on highly crystalline cellulose, even though it was active towards non-crystalline substrates. We interpreted these results as showing that not only catalysis, but also chain-end recognition is important for effective sliding on and degradation of crystalline cellulose.

In the present study, we prepared four samples of TrCel7A, i.e., wild-type (WT) and tryptophan 40 to alanine mutant (W40A), and the corresponding catalytic subunits, WTcat and W40Acat, without the CBD. We then characterized their hydrolytic activity and binding properties towards defined crystalline and amorphous cellulose substrates in order to examine the contributions of Trp40 and CBD to cellulose degradation. We also carried out molecular dynamic simulations of WTcat and W40Acat to elucidate the mechanism through which Trp40 promotes processive hydrolysis of cellulose chains.

**EXPERIMENTAL PROCEDURES**

Construction of mutant-encoding plasmid and transformation of T. reesei - E. coli strain DH5α (Promega) was used as the cloning host for DNA construction. The pEM-F5 plasmid containing the cel7A gene (cDNA) under its own promoter was used as an expression vector (13) and T. reesei ALKO 3413 (Roal Ltd.) strain lacking the genes coding for endogenous Cel7A and Cel7B was used as the host for production of the mutant enzyme. For selection of T. reesei transformants, hygromycin selection plasmid pRLMex30 was used (14). The mutant expression plasmid was constructed by overlap extension PCR and the DNA sequence of the whole mutated area was confirmed by DNA sequencing. Transformation and choice of the most productive transformant were performed basically as described earlier (15).

Isolation of TrCel7A and mutant enzymes - Production and purification of intact TrCel7A and the catalytic module of TrCel7A were done as described earlier (16).

TrCel7A W40A mutant was produced in a Chemap LF 20 fermenter (working volume 15 l). The cultivation medium used contained, in g* l⁻¹; whey powder 60, spent grain 30, KH₂PO₄ 5.0, (NH₄)₂SO₄ 5. Inocula (spore suspensions in 50% glycerol at – 80°C) were grown on the same medium with only 20 g l⁻¹ carbon source and 10 g l⁻¹ spent grain, buffered with 15 g l⁻¹ KH₂PO₄, in two stages of 4×50 ml (3 days, 200 rpm, 28°C) and 5×200 ml (10% v/v transfer, 2 days, 200 rpm, 28°C). The fermentation lasted for 192 hours and the other cultivation conditions were: T = 28°C, pH 4–5 (lower limit controlled by addition of NH₂OH, upper limit by...
addition of $\text{H}_3\text{PO}_4$), dissolved oxygen (DO) $>30\%$ (agitation 400–800 rpm), aeration 5 l min$^{-1}$. Foaming was controlled by automatic addition of Struktol J633 polyoleate antifoaming agent (Schill & Seilacher, Germany). After fermentation, the biomass was separated by centrifugation for 30 min at 4000 g, 4°C. The culture filtrate contained at least 1 g•L$^{-1}$ of W40A protein as judged from the Western blot analysis (see below).

Purification of the W40A mutants was started by adjusting the clarified *Trichoderma* culture filtrate (300 ml) in a Biogel P6 gel filtration column to 50 mM Na acetate buffer, pH 5.6. The sample, containing both intact W40A and its catalytic domain (W40A$^{\text{cat}}$), was then applied to a DEAE Sepharose FF column equilibrated with 50 mM sodium acetate, pH 5.6. The enzyme was eluted with a linear gradient of 0 to 0.5 M NaCl in equilibration buffer. The fractions containing the highest activity towards 4-methylumbelliferyl-$\beta$-D-lactoside (MULac) were pooled and adjusted to pH 6.0. Ammonium sulfate was added to 0.65 M and the sample was loaded on a Phenyl-Sepharose FF column equilibrated with 50 mM sodium acetate, pH 5.6. The enzyme was eluted with a linear gradient to 50 mM potassium phosphate buffer, pH 6.0. Ammonium sulfate was added to 0.65 M and the sample was loaded on a Phenyl-Sepharose FF column equilibrated with 50 mM potassium phosphate pH 6.0 containing 0.65 M ammonium sulfate. The unadsorbed proteins, containing both intact W40A and its catalytic domain (W40A$^{\text{cat}}$), was then applied to a DEAE Sepharose FF column equilibrated with 50 mM sodium acetate, pH 5.6. The enzyme was eluted with a linear gradient of 0 to 0.5 M NaCl in equilibration buffer. The fractions containing the highest activity towards 4-methylumbelliferyl-$\beta$-D-lactoside (MULac) were pooled and adjusted to pH 6.0. Ammonium sulfate was added to 0.65 M and the sample was loaded on a Phenyl-Sepharose FF column equilibrated with 50 mM potassium phosphate pH 6.0 containing 0.65 M ammonium sulfate. The unadsorbed proteins, including W40A catalytic module (W40A$^{\text{cat}}$) were eluted with equilibrating buffer and the adsorbed W40A was eluted with a linear decreasing gradient to 50 mM potassium phosphate buffer, pH 6.0. Analysis of the fractions was performed using MULac as a substrate. SDS-PAGE and Western blotting were also employed, as described earlier (15). The purest fractions containing intact, two-domain W40A and W40A$^{\text{cat}}$ were separately combined into two pools. The buffer of both pools was exchanged to 50 mM sodium acetate buffer pH 5.0 in a Biogel P6 column equilibrated with the same buffer. The fractions containing the highest protein concentration were pooled, concentrated, and used for the biochemical characterization described below. The result of SDS-PAGE of the purified enzymes is shown in Supplemental Figure S1. The contaminating activities towards 4-methylumbelliferyl-$\beta$-D-glucoside and hydroxyethyl cellulose (HEC) were also checked as described earlier, and the protein preparations were estimated to contain less than 0.1% $\beta$-glucosidase or endoglucanase contamination (17).

Protein concentration of *TrCel7A* preparations was measured in terms of UV absorption at 280 nm. For WT, a value of $\varepsilon$=83,000 M$^{-1}$cm$^{-1}$ was used, and for WT$^{\text{cat}}$ a value of $\varepsilon$=80,000 M$^{-1}$cm$^{-1}$, based on quantitative amino acid analysis, was used (A. Koivula, unpublished work). For W40A and W40A$^{\text{cat}}$ theoretical epsilon values of 77,500 M$^{-1}$cm$^{-1}$ and 74,500 M$^{-1}$cm$^{-1}$ were used, respectively; these were calculated from the measured values for the wild-type proteins by subtracting $\varepsilon$=5,500 M$^{-1}$cm$^{-1}$ for a tryptophan residue (18).

### Degradation of 4-methylumbelliferyl $\beta$-D-lactoside

Enzyme activities toward 4-methylumbelliferyl $\beta$-D-lactoside (MULac) were determined by monitoring the amount of released methylumbiferone at various substrate concentrations (0-3.0 mM) in 50 mM NaAc buffer, pH 5.0, at 27°C. Enzyme concentrations used were 0.3 µM for WT, WT$^{\text{cat}}$, W40A intact, and W40A$^{\text{cat}}$. The enzyme reactions were stopped by adding 0.5 M Na$_2$CO$_3$, and liberation of MU was measured as described by Voutilainen et al (19). The obtained data were fitted to the Michaelis-Menten equation. To calculate the kinetic parameters for substrate inhibition, the following equation was used for curve fitting:

$$v=k_{\text{cat}}E_0S/(K_m+S+S^2/K_{m2})$$

where $E_0$ is enzyme concentration, $S$ is substrate concentration, $K_m$ is the Michaelis constant, $K_{m2}$ is the substrate inhibition constant, and $k_{\text{cat}}$ is the catalytic constant. All kinetic parameters were calculated using Delta Graph 5.5.5 (Red Rock, USA) or KaleidaGraph 3.6.4 (Synergy, USA).

### Cellulose preparation

Highly crystalline cellulose (HCC) was prepared from green algae *Cladophora* spp. as described previously (20). Phosphoric acid-swollen cellulose (PASC) was prepared from Avicel (Funakoshi Corporation, Japan) as follows: Avicel was mixed with 85% (w/w) phosphoric acid and a completely clear solution was obtained by smashing with a glass stick. After overnight incubation at 4°C, cellulose was regenerated in water and a suspension was prepared with a high-speed blender. The cellulose suspension was washed with water and stored at 4°C. Crystallinity of cellulose substrates was estimated by X-ray powder diffraction (21), and showed that HCC is about 99% crystalline while PASC is non-crystalline.

### Cellulose hydrolysis

The hydrolytic activities towards HCC and PASC were determined by shaking each enzyme (final concentration of 1.4 µM) and 0.1% substrate at 27°C in 50 mM NaAc, pH 5.0. The hydrolysis was evaluated at nine different time points up to 24 h and the formation of soluble reducing sugars was determined with the *p*-hydroxybenzoic
acid hydrazide (PAHBAH) method using cellobiose as a standard (22). The amount of released cello-oligosaccharides was also determined by high-performance anion-exchange chromatography at three or four time points, basically as described earlier (23). The standards were glucose from Fluka (Buchs, Switzerland), cellobiose, cellotetraose and cellopentaose from Serva (Heidelberg, Germany), and cellotriose and cellohexaose from Seikagaku (Seikagaku America, Associates Cape Cod, Inc., Massachusetts, USA). In all cases, cellobiose was the major soluble product formed. Plots of produced cellobiose were fitted to an exponential and linear equation:

\[ q(t) = a(1 - \exp(-b*\mu t)) + c\mu t \]

and a differential equation:

\[ v(t) = a\mu b\exp(-b*\mu t) + c \]

where \( a, b \) and \( c \) are constant, \( t \) is incubation time, \( q(t) \) is product concentration, and \( v(t) \) is production velocity.

Adsorption analysis — Adsorption on HCC and PASC (0.1% (w/v)) was measured under similar conditions to those used for the hydrolysis with 1.4 \( \mu M \) enzyme at pH 5.0, 27°C. After 1 h and 2 h incubation, the mixtures were filtered through a Millex GV 0.22 \( \mu m \) membrane (Millipore, USA) to terminate the reaction. Then 700 \( \mu l \) of 50 mM NaAc (pH 5.0) was added to 300 \( \mu l \) filtrate and the protein concentration was measured using spectrofluorometer (Varian Gary Eclipse). A separate standard curve was prepared for each protein. The binding experiments were done in triplicate. The amount of adsorbed enzyme was calculated from the initial protein concentration, and the remaining cellulose contents at time points 1 h and 2 h were calculated from the amount of produced cellobiose.

Molecular dynamics simulation — The initial coordinates of the TrCel7A cat were taken from Protein Data Bank (PDB) structure 8CEL (2), and the structure of the W40A mutant was created by in silico mutation. The protonation states at pH 7.0 were determined using the PDB2PQR server (24), where the ten disulfide bonds described in the PDB were assigned to the relevant cysteine residues. To compare the behaviors of cellulose chain at the tunnel entrances of WT cat and W40A cat, the reducing end glucose (head) of cellononaose was positioned at subsite -7 (other parts were created automatically), which is the non-reducing end (tail) position in the original PDB structure. Then the systems were fully solvated with explicit solvent (including crystal water) and 19 Na\(^+\) counter ions were added to obtain electrostatic neutrality. All the simulations were performed using the Amber 11 package (25). The Amber ff03 force field was used for proteins. We employed the general Amber force field (GAFF) with the AM1-BCC partial charges for the N-terminal pyroglutamic acid (PCA), the GLYCAM_06 force field for the cellononaose, and the TIP3P model for water molecules. The systems were energetically minimized for 300 steepest-descent steps and equilibrated for 1 ns with gradually reducing restraints. Finally two 50 ns production runs were performed with different initial velocities for each system. The temperature and the pressure were controlled by using the Berendsen rescaling method (26) and the long-range electrostatic forces were calculated using the particle-mesh Ewald method (27). The trajectory analysis was conducted by using the Amber module ptraj, and the snapshot structures were visualized with VMD (28). In addition, the chain-end of cellononaose was initially positioned at subsite -5 and molecular dynamics simulations were performed using same methods as described above for both the WT cat and W40A cat.

RESULTS

Hydrolysis of soluble substrate - The kinetic constants (\( k_{\text{cat}} \), \( K_{\text{m1}} \), and \( K_{\text{m2}} \)) of the four enzyme variants (WT, WT cat, W40A, and W40A cat) with a soluble substrate, MULac, were measured and the results are shown in Table 1. There is no significant difference detected in the values. The W40A mutant would seem to have slightly increased substrate inhibition constant \( K_{\text{m2}} \), which may hint that the MULac substrate has also non-productive binding modes involving the subsite -7.

Hydrolysis of amorphous cellulose — In order to study the hydrolysis of insoluble cellulose in detail, we first used amorphous cellulose (PASC) as a substrate and determined the amount of bound protein as well as the velocity of cellobiose production, which was calculated from the progress curves. As shown in Figure 2A, all variants showed hydrolytic activity on PASC and furthermore, W40A mutant enzyme was over 2-fold more efficient than the WT in terms of cellobiose production (Table 2). The specific activities of the bound enzymes were also determined by dividing the cellobiose...
production rates by the amount of adsorbed enzyme at time points 1h and 2h (Table 2). Since the WT and W40A bound equally well to PASC, the specific activity of adsorbed W40A was 2.7 (1 h) and 2.2 (2 h) times greater than that of WT, respectively, indicating that the Trp40 residue slows down the hydrolysis of amorphous cellulose. Although the single-domain version W40Acat had quite low activity towards PASC, its specific activity was similarly 2.7 (1 h) and 2.6 (2 h) times higher than that of WTcat, indicating that the negative effect of Trp40 was independent of the presence or absence of CBD. However, as can be seen in Table 2, W40Acat was significantly less well adsorbed on PASC when compared to WT, W40A or WTcat.

Hydrolysis of highly crystalline cellulose – In contrast to PASC, cellobiose production by WTcat, W40A, and W40Acat was significantly decreased compared with that by WT, when HCC was used as a substrate (Fig. 2B). The results indicate important roles for both Trp40 and CBD in the hydrolysis of crystalline cellulose. As shown in Table 2, the velocities of cellobiose production by W40A and WTcat on HCC were reduced to approximately 20% and 30% of that of WT, respectively. The specific activity of bound W40A was 45% (1 h) or 21% (2 h) of that of WT, whereas the WTcat showed a specific activity 4.4 (1 h) or 2.6 (2 h) times higher than that of WT. The adsorption of W40A was 48% (1 h) or 73% (2 h) of that of WT. These results indicate that Trp40 may play a role in initiating degradation of HCC as well as in processive hydrolysis of HCC, whereas CBD hinders the hydrolysis.

Molecular dynamics simulation – In order to examine the the function of Trp40 at the atomic level, molecular dynamics (MD) simulations were carried out for WTcat and W40Acat with cellononaose, with the reducing-end glucose initially positioned at Trp40 or Ala40. Two simulation runs for 50 ns were performed for each system (Fig. 3A), and the distances between the reducing-end glucose, and the catalytic carboxylic acid Glu212 are plotted in Fig. 3B. In the case of WTcat, cellononaose entered the catalytic tunnel and finally penetrated about 11 Å within the tunnel, approximately two glucose units, towards the catalytic residue Glu212 during the first simulation (Supplemental Video S1). The cellononaose had finally entered from subsite -7 to subsite -5. The second trial gave similar results, although the approach towards the Glu212 was much faster than in the first trial (Supplemental Video S2).

On the other hand, in the case of W40Acat, the cellononaose did not move into the active site tunnel but diffused away from the alanine residue at subsite -7. The substrate chain bound on the back side of the catalytic domain in the first trial, and on one of the loops forming the active site tunnel in the second trial (Supplemental Video S3 and S4). In contrast, W40Acat as well as WTcat kept the cellononaose when reducing-end of the chain was initially positioned at subsite -5 as shown in Fig. 3A. In the simulations of WTcat, glucose rings were stuck to subsite -5, -6, and -7 (Supplemental Video S5 and S6), which is similar to the simulations started form subsite -7. However, the simulation of W40Acat started from -5 showed the strong enough interactions at subsites -5 and -6 to keep the cellononaose chain in the tunnel (Supplemental Video S7 and S8). In the second trial of W40Acat, the cellononaose processed into the subsite tunnel much earlier than WTcat (Fig. 3C). Consequently, these results indicate that Trp40 at subsite -7 may assist entry of the cellulose into the subsite tunnel with much shorter length of the free chain-end and in initiating the processive reaction of TrCel7A.

DISCUSSION
The protein-carbohydrate interactions of carbohydrate-active enzymes (CAZymes) are characterized by a network of hydrogen bonds and hydrophobic stacking interactions involving aromatic amino acid residues (29,30). Stacking interactions seem to be especially important in the cases of enzymes and individual protein domains that act at the surface of insoluble substrates, e.g. carbohydrate-binding modules (CBM) (31), chitinases (32) and cellulases (33). TrCel7A has a two-domain structure composed of a catalytic domain (cat) and a cellulose-binding domain (CBD), both of which have aromatic amino acid residues that serve to align the substrate-binding area and form hydrophobic stacking interactions with the glucosyl units of cellulose (2,31). In processive cellulose hydrolysis the TrCel7A catalytic domain slides along the cellulose chain and releases the product (cellobiose) from the subsites +1 and +2, then proceeds to the next cellobiose unit from the reducing end. To investigate this process, we focused here on the roles of TrCel7A Trp40 and CBD in cellulose degradation. The Trp40 residue is located in an overhang at the entrance in the active site tunnel, and in the productive binding mode it makes contact with the sugar ring (2).

With soluble MULac as a substrate, no significant difference in activity was detected, as
expected (Table 1). Elimination of the CBD did not significantly affect the PASC hydrolysis by the WT, whereas mutation of Trp40 to Ala caused a clear increase in cellobiose production from amorphous substrates (Table 2). Since PASC is an insoluble substrate, adsorption behavior was also compared and is summarized in Scheme 1. The results indicate that CBD had a very small, if any, effect to the adsorption of the WT on PASC, which was also an expected result. However, the W40A cat was binding clearly worse to PASC than the WT cat, or CBD remaining mutant (W40A). This suggests that adsorption of the TrCel7A catalytic domain on amorphous cellulose is dependent on the Trp40. The CBD can, however, compensate the binding defect caused by the W40A mutation. Concerning the specific activity of the adsorbed enzymes on PASC, the Trp40 residue had a negative effect as shown in Scheme 1. We assume that this is due to the stacking -7, which interferes with the cellulose chain slide within the tunnel and with the dissociation of the catalytic domain from a cellulose chain.

The hydrolysis of HCC, in contrast, showed completely different behavior from those of MULac and PASC. Enzyme adsorption was clearly promoted by the CBD whereas Trp40 had less effect to the overall binding. However, the specific activity of adsorbed enzyme on HCC was greatly enhanced by the presence of Trp40 but hindered by the CBD. This is an important finding, because there is likely to be a very sensitive balance between the Trp40 and CBD affecting to the processes of processivity and adsorption. Considering the MD simulation results, Trp40 should have a particularly important role when a limited number of reducing-glucose units are available to the active site. The essential difference between amorphous and crystalline celluloses is an availability of the free chain-ends. Therefore, HCC has only small number of glucose units free at the chain-end, whereas PASC has longer chain-ends available. The free reducing-end of the chain might not reach to deep position of subsite without Trp40 residue, which is critical for HCC degradation. On the other hand, the chain-ends of PASC easily reach to subsite -5, and Trp40 residue is unnecessary for chain loading. These results are also supported by previous molecular dynamic study calculating binding energy of subsite -5 to -7 (3,34). Concerning the role of CBD, the population of the adsorbed enzyme is clearly higher in the presence of the CBD, as can be seen when comparing the adsorption of WT cat and WT. On the other hand, the CBD domain lowers the specific activity of the adsorbed enzyme. There may be two reasons for this: 1) the CBD may restrict the intramolecular processive movement of the enzyme on cellulose surface; 2) the CBD may lead to the non-productive binding of enzyme on places where the catalytic module does not find a free chain end. In our high-speed AFM study, we observed the apparent movement of WT cat on HCC, and suggested that the major contribution of the CBD is simply to increase the concentration of adsorbed enzyme on the crystalline substrate (10). The present experiments with WT and WT cat further suggest that CBD-mediated adsorption involves non-productive binding and/or disturbance of the hydrolytic activity, which is seen as reduced specific activity of the adsorbed enzyme.

In our previous high-speed AFM work, we observed a “wriggling” movement of the W40A mutant enzyme on the HCC surface and speculated that this is due to inability to recruit the substrate chain into the active site tunnel (10). Our MD simulation results here also support this hypothesis, and the biochemical data suggests a function of Trp40 in initiating the processive sliding movement on HCC. The importance of this Trp residue in crystalline cellulose degradation is further supported by the fact that this residue is conserved among the sequences of GH7 family cellobiohydrolases (http://www.cazy.org/) as shown in Supplemental Figure S2 (35). Furthermore, there is a conserved Trp residue at the subsite entrance among GH6 family CBHs but not in GH6 endoglucanases. It has been shown previously that the T. reesei GH6 family cellobiohydrolase (TrCel6A) variants having this corresponding Trp residue mutated to alanine or aspartic acid have clearly decreased hydrolysis activity of bacterial microcrystalline cellulose but not of PASC (36). This suggests that a tryptophan residue located at the entrance of the subsite is essential for substrate loading in all CBHs that degrade crystalline celluloses.

In conclusion, our present study indicates that the CBD of TrCel7A serves to increase the enzyme concentration on the HCC surface, but at the expense of reducing the specific activity of the adsorbed enzyme. The main function of the conserved Trp40 residue at the edge of the subsite tunnel appears to be recruitment of cellulose chain reducing-ends into the active site tunnel and initiation of processive hydrolysis of cellulose chain on HCC. This is important because long free cellulose chain ends are not frequently available on HCC, but again there is a trade-off due to the stacking interaction of the Trp40 residue, which impedes progression of the
cellulose chain within the active-site tunnel. Consequently, removal of Trp40 leads to more efficient degradation of less crystalline substrates.

REFERENCES

TrCel7A W40A mutagenesis


Acknowledgements
We thank Michael Bailey, Harry Boer, Tiina Hartikainen, Tarmo Pellikka, Tiina Liljankoski, and Riitta Suikkonen at VTT for help in production, purification and initial characterisation of the TrCel7A W40A mutant enzymes. This research was supported by a Grant-in-Aid for Innovative Areas (No. 24114008) to K.I. from the Japanese Ministry of Education, Culture, Sports, and Technology (MEXT), by a Grant-in-Aid for Scientific Research (B: No. 24380089), by Grants-in-Aid for Scientific Research for Young Scientists (A: No. 19688016 and 21688023) and (B: No. 17780245) to K.I., Grant-in-Aid for JSPS Fellows to T.T. (No. 08J07454) from Japan Society for the Promotion of Science (JSPS), and by a grant from the Advanced Low Carbon Technology Research and Development Program (ALCA) of the Japan Science and Technology Agency (JST) to K.I.

FOOTNOTES
Abbreviations
CBD: cellulose-binding domain; CBH: cellobiohydrolase; cat: catalytic domain; GH: glycoside hydrolase; CBM: carbohydrate-binding module; MULac: 4-methylumbelliferyl β-D-lactoside; PASC: phosphoric acid-swollen cellulose; HCC: highly crystalline cellulose; WT: wild-type
**Table 1.** Kinetic constants measured on MULac (27 °C, pH 5.0).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$k_{cat}$ (min⁻¹)</th>
<th>$K_m$ (mM)</th>
<th>$k_{cat}/K_m$ (min⁻¹ mM⁻¹)</th>
<th>$K_m^2$ (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>29.5 ± 1.7</td>
<td>0.358 ± 0.03</td>
<td>82 ± 1.5</td>
<td>4.70 ± 1.6</td>
</tr>
<tr>
<td>WT$_{cat}$</td>
<td>24.9 ± 1.6</td>
<td>0.335 ± 0.03</td>
<td>74 ± 1.6</td>
<td>6.22 ± 2.7</td>
</tr>
<tr>
<td>W40A</td>
<td>19.8 ± 1.5</td>
<td>0.318 ± 0.03</td>
<td>62 ± 1.6</td>
<td>7.07 ± 4.1</td>
</tr>
<tr>
<td>W40A$_{cat}$</td>
<td>21.7 ± 1.3</td>
<td>0.293 ± 0.02</td>
<td>74 ± 1.7</td>
<td>7.77 ± 4.2</td>
</tr>
</tbody>
</table>

$K_m^2$: substrate inhibition constant

**Table 2** Hydrolysis velocity and adsorption of the TrCel7A enzyme variants on HCC and PASC measured at 27°C in 50 mM NaAc, pH 5.0. The enzyme concentration was 1.4 µM and substrate concentration was 0.1% (w/v).

<table>
<thead>
<tr>
<th>Cellulose substrate</th>
<th>Enzyme</th>
<th>Time point (h)</th>
<th>Hydrolysis velocity$^a$ (µM•min⁻¹)</th>
<th>Bound enzyme (nmol/mg-cellulose)$^b$</th>
<th>Specific activity of adsorbed enzyme$^c$ (min⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PASC</td>
<td>WT</td>
<td>1</td>
<td>1.8 ± 1.1</td>
<td>1.4</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>1.5 ± 0.6</td>
<td>1.5</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>W40A</td>
<td>1</td>
<td>4.4 ± 1.1</td>
<td>1.3</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>3.0 ± 0.5</td>
<td>1.6</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>WT$_{cat}$</td>
<td>1</td>
<td>1.4 ± 0.0</td>
<td>1.1</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.72 ± 0.01</td>
<td>1.2</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>W40A$_{cat}$</td>
<td>1</td>
<td>0.69 ± 0.10</td>
<td>0.20</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.34 ± 0.02</td>
<td>0.24</td>
<td>1.7</td>
</tr>
<tr>
<td>HCC</td>
<td>WT</td>
<td>1</td>
<td>0.76 ± 0.02</td>
<td>0.62</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.47 ± 0.01</td>
<td>0.56</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>W40A</td>
<td>1</td>
<td>0.16 ± 0.00</td>
<td>0.30</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.08 ± 0.01</td>
<td>0.41</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>WT$_{cat}$</td>
<td>1</td>
<td>0.21 ± 0.06</td>
<td>−0.05$^d$</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.15 ± 0.02</td>
<td>−0.07$^d$</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>W40A$_{cat}$</td>
<td>1</td>
<td>0.06 ± 0.01</td>
<td>0.10</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.03 ± 0.01</td>
<td>0.15</td>
<td>0.3</td>
</tr>
</tbody>
</table>

$^a$Hydrolysis velocity, shown as produced cellobiose/min, was calculated at time points 1 h and 2 h from the progress curves shown in Fig. 2.

$^b$The bound enzyme concentration was determined from the supernatant after 1 h and 2 h incubation with HCC or PASC as explained in the materials and methods section, and then dividing this concentration value by the remaining cellulose amount.

$^c$Specific activity of the adsorbed enzyme was calculated by dividing the hydrolysis rate at time points 1 h and 2 h by the bound enzyme concentration. The uncertainty in the specific activity values is estimated to be 20% based on three repeated experiments.

$^d$The error is estimated to be 50%, since the measured concentration was lower than 0.1 µM. Therefore the amount of adsorbed WT$_{cat}$ on HCC are given as rough estimates.
FIGURE LEGENDS

FIGURE 1. Three-dimensional structure of the catalytic domain (cat) of *Trichoderma reesei* Cel7A celllobiohydrolase. The α-carbon chain structure of *Tr* Cel7A catalytic domain (cat) is shown, together with a modelled continuous cellononaose (Glc9) chain in the active site tunnel (Protein Data Bank code: 8CEL). The tryptophan 40 (W40) residue at subsite -7 is shown in red and glutamic acid residue E212, which acts as the catalytic nucleophile between subsites -1/+1, used as a reference position in the molecular dynamics simulations, is shown in magenta.

FIGURE 2. Progress curves and product (cellobiose) velocity curves of cellulose hydrolysis by WT, W40A, WTcat, and W40Acat. Phosphoric acid-swollen cellulose (PASC, A) and *Cladophora* spp. (HCC, B) were incubated at 27°C in 50 mM NaAc, pH 5.0, with 1.4 µM of each enzyme at 0.1 % (w/v) substrate concentration. Duplicate assays were performed at each time point. Soluble products were determined by the PAHBAH method using cellobiose as a standard, as described in the Materials and Methods section. WT, filled square; W40A, filled circle; WTcat, filled triangle; W40Acat, filled diamond. Cellobiose production velocities were calculated from progress curves (PASC; C and HCC; D).

FIGURE 3. Initial and final snapshots of WTcat and W40Acat with cellononaose, and the distances between cellononaose and the catalytic site in the molecular dynamics (MD) simulations. The reducing-end glucose of cellononaose was initially placed at subsite -7 or -5 of WTcat and W40Acat. Simulations were performed twice (trials a and b) for each enzyme. Initial and final structures were depicted with the close-up views (A). The distances between the catalytic residue Glu212 and the reducing-end glucose of cellononaose during the simulations for subsites -7 and -5 are plotted in (B) and (C), respectively.

Scheme 1. Effect of CBD and Trp40 residue (W40) on adsorption and specific activity of the adsorbed enzyme (velocity/adsorption) during HCC and PASC hydrolysis. The number written besides the arrows means the ratio calculated from the 1h (top) and 2h (bottom) data. Amount of adsorbed enzyme for HCC (A) and PASC (C) and specific activity for HCC (B) and PASC (D) were calculated from Table 1. Values of front edge of the arrow were divided by values of tail edge of the arrow.
Figure 1

TrCel7A W40A mutagenesis
Figure 2

TrCel7A W40A mutagenesis

A

B

C

D

Cellulose (µM)

Cellulose (µM)

Velocity (µM/min)

Velocity (µM/min)

Reaction time (h)

Reaction time (h)

Reaction time (h)

Reaction time (h)
Figure 3

A

Subsite -7 start

Initial (0 ns)  Final (50 ns)

WTcat

W40Acat

Subsite -5 start

Initial (0 ns)  Final (50 ns)

WTcat

W40Acat

B

distance [Å] vs time [ns]

WTcat a  WTcat b  W40Acat a  W40Acat b

C

distance [Å] vs time [ns]

WTcat -5a  WTcat -5b  W40Acat -5a  W40Acat -5b
Scheme 1

TrCel7A W40A mutagenesis

A  HCC

Adsorption

WTcat

WT

W40Acat

2.1

1.4

16

7.8

3.0

2.7

0.4

0.5

C  PASC

1.3

1.3

1.1

1.3

6.5

6.7

5.5

5.0

Specific Activity

B

0.2

0.4

9.5

7.9

2.2

4.8

1.0

0.6

D

1.0

1.7

0.4

0.4

0.4

0.4

1.0

1.4
Tryptophan residue at active-site tunnel entrance of *Trichoderma reesei* cellobiohydrolase Cel7A is important to initiate degradation of crystalline cellulose

Akihiko Nakamura, Takeshi Tsukada, Sanna Auer, Tadaomi Furuta, Masahisa Wada, Anu Koivula, Kiyohiko Igarashi and Masahiro Samejima

*J. Biol. Chem.* published online March 26, 2013

Access the most updated version of this article at doi: 10.1074/jbc.M113.452623

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2013/03/26/M113.452623.DC1

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2013/03/26/jbc.M113.452623.full.html#ref-list-1